Advanced receivers for distributed cooperation in mobile ad hoc networks

Serdar Sahin

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Institut National Polytechnique de Toulouse (Toulouse INP)

Discipline ou spécialité :

Informatique et Télécommunication

Présentée et soutenue par :

M. SERDAR SAHIN

le jeudi 10 octobre 2019

Titre :

Advanced receivers for distributed cooperation in mobile ad hoc networks

---

Ecole doctorale :

Mathématiques, Informatique, Télécommunications de Toulouse (MITT)

Unité de recherche :

Institut de Recherche en Informatique de Toulouse (IRIT)

Directeur(s) de Thèse :

M. CHARLY POULLIAT

MME MARIE LAURE BOUCHERET

Rapporteurs :

Mme FLORENCE ALBERGE, UNIVERSITE PARIS 11

Mme INBAR FIJALKOW, ENSEA

Membre(s) du jury :

M. LUC VANDENDORPE, ECOLE POLYTECHNIQUE DE LOUVAIN, Président

M. ANTONIO MARIA CIPRIANO, THALES, Membre

M. BERNARD HENRI FLEURY, AALBORG UNIVERSITY, Membre

M. CHARLY POULLIAT, TOULOUSE INP, Membre

Mme MARIE LAURE BOUCHERET, TOULOUSE INP, Membre

M. RAPHAEL VISOZ, ORANGE ISSY LES MOULINEAUX, Invité

M. RAYMOND KNOPP, INSTITUT EURECOM SOPHIA ANTIPOLIS, Membre

Remerciements

Il ne me reste plus qu'une dernière responsabilité pour achever cette aventure de plus de 3 ans : de remercier les gens qui m'ont tendu la main à des moments où je n'en pouvais plus.

J'aimerais d'abord remercier mon encadrant industriel Antonio M. Cipriano et mes directeurs de thèse Charly Poulliat et Marie-Laure Boucheret, pour m'avoir formé et pour m'avoir guidé et soutenu à travers mes travaux de recherche, toute en me laissant une grande liberté. Je remercie Charly pour sa grande disponibilité, pour m'avoir toujours poussé à essayer d'aller plus loin et pour l'enthousiasme scientifique qu'il m'a fait part lors de nos nombreuex échanges. Je remercie également Antonio pour sa grande disponibilité, pour sa grande rigueur et pertinence scientifique. Je vous aussi exprimer ma gratitude envers Marie-Laure pour m'avoir formé, pour ses conseils techniques et ses encouragements pendant ces trois ans de recherche. Je vous remercie tous pour la confiance que vous avez placée en moi.

Je souhaite remercier Florence Alberge et Inbar Fijalkow pour avoir rapporté mon manuscrit, et j'aimerais également remercier les autres membres de mon jury Bernard-Henri Fleury, Raymond Knopp, Luc Vanderdorpe et Raphaël Visoz pour avoir accepté d'examiner ma thèse et d'assister à ma soutenance.

Je remercie Charly et Marie Chabert pour m'avoir accueilli au sein de l'équipe Signal et Communications de l'IRIT. Je tiens aussi à remercier Marie-Laure pour avoir partagé son bureau et sa bonne humeur avec moi, et je remercie également à mes collègues et ex-collègues de l'équipe: Adrien, Asma, Baha, Bilel, Camille, Cédric, Charles-Ugo, Claire, Dana, Etienne, Lorenzo, Louis, Maxime, Mouna, Nathalie, Nicolas, Olivier, Pierre-Hugo, Selma, Tarik, Thomas, Vinicius et les autres... je vous souhaite à tous une très bonne continuation.

Je remercie Jean-Luc Peron, le responsable de l'équipe Waveform Design à Thales, pour m'avoir accueilli au sein de son service et pour son intérêt à mes travaux. J'aimerais remercier mes collègues, ex-collègues et amis de Thales pour leur humour, soutien et leur bonne humeur. Merci Adrien, Alice, Anaël, Arnaud, Arthur, Benoit, Christophe, Dorin, Elie, Elisa, Eric, Florian, Hélène, Jérôme, Kévin, Laurent, Lélio, Lorenzo, Ludovic, Luxmiram, Marie, Olivier, Philippe, Ranga, Raphaël, Romain, Simon, Sylvain, Titouan et Xavier. Je remercie de nouveau à Eric pour avoir complété mon code pendant son stage et d'avoir survécu!

J'aimerais vivement remercier Mathieu Raimondi, Pascal Acco et Benjamin Gadat pour toutes leurs aides qui m'ont permis de trouver cette opportunité de thèse.

Pour leur soutien et disponibilité en dehors du travail, je tiens à remercier mes amis, et en particulier à Baptiste, Carole, Guillaume, Marion, Olivier, Thomas et Yohan.

Ve son olarak, bana olan eksiksiz destekleri ve sevgileri için anneme, babama ve kardeşime minnettarım. Benim bu birikimleri kazanmamı ve gelebildiğim bu noktaya ulaşmamı siz sağladınız. İyi ki varsınız ve her şey için çok sağ olun!

Dünyada her şey için, maddiyat için, maneviyat için, muvaffakiyet için, en hakikî mürşit ilimdir, fendir. İlim ve fennin haricinde mürşit aramak gaflettir, cehalettir, dalâlettir. Yalnız, ilim ve fennin yaşadığımız her dakikadaki safhalarının tekamülünü idrak etmek ve terakkiyatını zamanında takip eylemek şarttır. Bin, iki bin, binlerce sene evvelki ilim ve fen lisanının çizdiği düsturları şu kadar bin sene sonra bugün aynen tatbikata çalışmak elbette ilim ve fennin içinde bulunmak değildir.[0]

Mustafa Kemal Atatürk, 22 Eylül 1924, Samsun

[0]For everything in the world, for wealth, for the soul, for success, the truest guide is science. To seek guidance in other things is foolishness, ignorance and heresy. But, it is necessary to perceive the developments of science at each moment of our lives, and to follow its progress on time. Attempting to keep relying on the same scientific postulates formulated thousands of years ago would certainly not constitute a scientific contribution.

Contents

List of Figures

List of Tables

Abbreviations

3G	3rd Generation
3GPP	3rd Generation Partnership Project
4G	4th Generation
5G	5th Generation
A2A	Air to Air
AF	Amplify & Forward
AMP	Approximate Message Passing
APP	A Posteriori Probability
ARQ	Automatic Repeat Request
ASIC	Application Specific Integrated Circuit
AWGN	Additive White Gaussian Noise
BCJR	Balk Cocke Jelinek Raviv
BER	Bit Error Rate
BICM	Bit Interleaved Coded Modulation
BLAST	Bell-Labs Layered Space-Time
BP	Belief Propagation
BPSK	Binary Phase Shift Keying
CB	Cooperative Broadcasting
CDMA	Code Division Multiple Access
CF	Compress & Forward
CFO	Carrier Frequency Offset
CFR	Channel Frequency Response
CIFRE	Convention Industrielle de Formation par la Recherche
CIR	Channel Impulse Response
CNRS	Centre National de Recherche Scientifique

CP	Cyclic Prefix
CRC	Cyclic Redundancy Check
CSI	Channel State Information
CSMA	Carrier Sense Multiple Access
D2D	Device-to-Device
DF	Decode & Forward
DFE	Decision Feedback Equalizer
DFT	Discrete Fourier Transform
DL-SEP	Double Loop - Scalar EP
DSP	Digital Signal Processing
ECA	Expectation Consistent Approximation
EESM	Exponential ESM
EP	Expectation Propagation
ESM	Effective Signal-to-noise Metric
EXIT	Extrinsic Information Transfer
EXT	Extrinsic
EXVIT	Extrinsic Variance to Information Transfer
FD	Frequency Domain
FDE	Frequency Domain Equalizer
FDMA	Frequency Division Multiple Access
FDO	Frequency Domain Oversampling
FEC	Forward Error Correction
FFT	Fast Fourier Transform
FH	Frequency Hopping
FIR	Finite Impulse Response
FLOP	Floating Point Operation
FPGA	Field Programmable Gate Array
FS	Fractionally-Spaced

GaBP	Gaussian-approximated Belief Propagation
GAMP	Generalized Approximate Message Passing
G2G	Ground to Ground
HARQ	Hybrid Automatic Repeat Request
IBI	Inter-Block Interference
IBDFE	Iterative Block DFE
IC	Interference Cancellation
ICI	Inter-Carrier Interference
IDMA	Interleave Division Multiple Access
IEEE	Institute of Electrical and Electronics Engineers
IFFT	Inverse Fast Fourier Transform
IID	Identically and Independently Distributed
INPT	Institut National Polytechnique de Toulouse
IRIT	Institut de Recherche en Informatique de Toulouse
ISI	Inter-Symbol Interference
IST	Iterative Soft Thresholding
IV	Iteration Variant
LDPC	Low-Density Parity Check
LE	Linear Equalizer
LLR	Log-Likelihood Ratio
LTE	Long Term Evolution
LTE-A	Long Term Evolution Advanced
LoS	Line of Sight
LUT	Lookup Table
MAC	Medium Access Control
MAI	Multiple-Antenna Interference
MANET	Mobile Ad Hoc Network
MAP	Maximum A Posteriori

MCS	Modulation and Coding Scheme
MF	Mean Field
MFB	Matched Filter Bound
MI	Mutual Information
MIESM	Mutual Information ESM
MIMO	Multiple Input Multiple Output
MISO	Multiple Input Single Output
ML	Maximum Likelihood
MLSE	Maximum Likelihood Sequence Estimation
MMSE	Minimum Mean Square Error
MSE	Mean Square Error
MTC	Machine Type Communications
MUD	Multiple-User Detector
MUI	Multiple-User Interference
MU-MIMO	Multi-User Multiple Input Multiple Output
NLoS	Non Line of Sight
NOMA	Non-orthogonal Multiple Access
NRNSC	Non-Recursive Non-Systematic Convolutional
OAMP	Orthogonal Approximate Message Passing
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
OLA	Overlap and Add
OMA	Orthogonal Multiple Access
OSI	Open Systems Interconnection
PAPR	Peak to Average Power Ratio
PDA	Probabilistic Data Association
PDF	Probability Density Function
PER	Packet Error Rate

PHY	Physical
PIC	Parallel Interference Cancellation
PMF	Probability Mass Function
PMR	Private Mobile Radio
PSK	Phase Shift Keying
RSC	Recursive Systematic Convolutional
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying
QoS	Quality of Service
SC	Single Carrier
SC-FDE	Single-Carrier Frequency Domain Equalization
SC-FDMA	Single-Carrier Frequency Division Multiple Access
SI	Self-Iterated
SIC	Successive Interference Cancellation
SIMO	Single Input Multiple Output
SINR	Signal to Interference and Noise Ratio
SIR	Symmetric Information Rate
SISO	Soft Input Soft Output
SNR	Signal to Noise Ratio
STBICM	Space-Time Bit Interleaved Coded Modulation
SVD	Singular Value Decomposition
TD	Time Domain
TDMA	Time Division Multiple Access
TI	Time Invariant
TV	Time Varying
UHF	Ultra High Frequency
UW	Unique Word
VAMP	Vector Approximate Message Passing

VHF	Very High Frequency
WLAN	Wireless Local Area Network
WMAN	Wireless Metropolitan Area Network
WSN	Wireless Sensor Network
ZF	Zero Forcing
ZP	Zero Padding

Notations

Generalities

Bold lowercase letters are used to denote vectors: let $\mathbf{u}$ be a $N \times 1$ vector, then scalars $u_n, n = 1, \ldots, N$ are its entries. Capital bold letters denote matrices: for a $N \times M$ matrix $\mathbf{A}$, $[\mathbf{A}]_{n,:}$ and $[\mathbf{A}]_{:,m}$ respectively denote its n^{th} row and m^{th} column, and $a_{n,m} = [\mathbf{A}]_{n,m}$ is the entry (n, m). When dealing with indexing vectors and matrices, the first element has the index 1, by following a MATLAB®-like convention. Similarly, when using brackets to define the contents of a matrix, ; for vertical concatenation (over the columns) and , is used for horizontal concatenation (over the rows). Given the vector $\mathbf{x} \in \mathbb{C}^K$, we use $\mathbf{x}^{\backslash k}$ or $\mathbf{x}^{\backslash x_k}$ to denote the size $K - 1$ vector, corresponding to $\mathbf{x}$ without its k^{th} element, i.e. x_k.

$\mathbf{I}_N$ is the $N \times N$ identity matrix, $\mathbf{0}_{N,M}$ and $\mathbf{1}_{N,M}$ are respectively all zeros and all ones $N \times M$ matrices. The operator $\mathbf{Diag}(\mathbf{u})$ denotes the diagonal matrix whose diagonal is defined by $\mathbf{u}$, the operator $\mathbf{diag}(\mathbf{U})$ denotes the column vector defined by the diagonal of the matrix $\mathbf{U}$, and $\mathbf{Diag}(\mathbf{U})$ denotes the matrix whose diagonal is defined by the $\mathbf{U}$, but its other entries are zeroes. $\mathbf{X}^T$ denotes the transpose operation, $\mathbf{X}^H$ denotes the Hermitian transposition and $\mathbf{X}^*$ denotes the complex-conjugate operation.

$\mathbb{R}, \mathbb{C}$, and $\mathbb{F}_k$ are respectively real, complex and k^{th} order Galois fields. $(n)_N$ denotes the classical modulo operation, that yields the number between 0 and $N - 1$ congruent to n, and $\langle n \rangle_N \triangleq 1 + (n - 1)_N$ denotes modulo for congruence to the numbers between 1 and N. Standard fonts are used for denoting most parameters, or sampled processes, but time-continuous processes are denoted with roman, i.e. serifed fonts. For instance, samples of $\mathrm{u}(t)$ at $t = kT_s$, are denoted u_k, for integer k, and T_s being a sampling period.

$\mathcal{F}_K$ is the normalized K-point Discrete Fourier Transform (DFT) matrix with its elements given by $[\mathcal{F}_K]_{k,l} = \exp(-2j\pi(k-1)(l-1)/K)/\sqrt{K}$, for $k, l = 1, \ldots, K$, and such that $\mathcal{F}_K \mathcal{F}_K^H = \mathbf{I}_K$. Moreover, if $\mathbf{x} \in \mathbb{C}^K$ is a time-domain process, then its frequency-domain response is denoted with an underbar, such that $\underline{\mathbf{x}} = \mathcal{F}_K \mathbf{x}$ (however there are some exceptions to the use of underbar during message passing derivations). In particular, we have $\mathcal{F}_K^H \mathbf{Diag}(\sqrt{K}\underline{\mathbf{x}}) \mathcal{F}_K = \mathbf{Circ}(\mathbf{x})$, where $\mathbf{Circ}(\cdot)$ is the $K \times K$ circulant matrix whose first row is given by the argument.

Probability and Statistics

Let x and y be two random variables, $\mu_x = \mathbb{E}[x]$ is the expected value, $\sigma_x^2 = \mathrm{Var}[x]$ is the variance and $\sigma_{x,y} = \mathrm{Cov}[x, y]$ is the covariance. The probability of the discrete random variable x taking the value α is $\mathbb{P}[x = \alpha]$. For random vectors $\mathbf{x}$ and $\mathbf{y}$, $\boldsymbol{\mu}_{\mathbf{x}} = \mathbb{E}[\mathbf{x}]$ is the expected value, and $\boldsymbol{\Sigma}_{\mathbf{x},\mathbf{y}} = \mathbf{Cov}[\mathbf{x}, \mathbf{y}]$ is the covariance matrix and $\boldsymbol{\Sigma}_{\mathbf{x}} = \mathbf{Cov}[\mathbf{x}, \mathbf{x}]$. The squared Mahalanobis distance is denoted $\|\mathbf{x} - \boldsymbol{\mu}\|_{\boldsymbol{\Sigma}}^2 = (\mathbf{x} - \boldsymbol{\mu})^H \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})$. For the sake of notational simplicity, we do not distinguish a random variable and its values.

Moreover, within the context of variational Bayes methods, following notations are used.

$\propto$	Relation operator for two functions scaled by a constant factor
$\delta(x)$	Dirac delta distribution
$m_{\mathrm{F}_j \to x_k}(x_k)$	Message from the factor node F_j to the variable node x_k
$m_{x_k \to \mathrm{F}_j}(x_k)$	Message from the variable node x_k to the factor node F_j
$p(x)$	PDF of x
$q(x)$	Approximated PDF of x
$\tilde{q}(x)$	Pre-projection posterior PDF of x (in the context of EP)
$\mathcal{V}(\cdot)$	Set of neighbours of a node
$\mathbf{x}_{\mathrm{F}_i}$	Vector of variables related to the factor F_i

If $\mathbf{x} \in \mathbb{R}^k$ is Gaussian distributed of mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$, then it is denoted $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and its PDF is denoted $\mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$, and

$$\mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{\sqrt{(2\pi)^k |\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right). \tag{1}$$

If $\mathbf{x} \in \mathbb{C}^k$ is symmetrically circular complex Gaussian distributed of mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$, then it is denoted $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and its PDF is denoted $\mathcal{CN}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$, and

$$\mathcal{CN}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(\pi)^k |\boldsymbol{\Sigma}|} \exp\left(-(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right), \tag{2}$$

and $\mathcal{R}(\mathbf{x}), \mathcal{I}(\mathbf{x}) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}/2)$. If v is inverse-gamma distributed with shape parameter α and scale parameter β, then it is denoted $v \sim \mathcal{IG}(\alpha, \beta)$ and its PDF is denoted

$$\mathcal{IG}(v; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \left(\frac{1}{v}\right)^{(\alpha+1)} \exp\left(-\frac{\beta}{v}\right), \tag{3}$$

with a mean $\beta/(\alpha-1)$ (if $\alpha > 1$) and a variance $\beta^2/[(\alpha-1)^2(\alpha-2)]$ (if $\alpha > 2$). When $\mathbf{x}$ follows the complex multivariate Student distribution with location parameter $\boldsymbol{\mu}$, the shape matrix $\boldsymbol{\Sigma}$ and with ν degrees of freedom, it is denoted $\mathbf{x} \sim \mathcal{CT}_\nu(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and its PDF is denoted

$$\mathcal{CT}_\nu(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{\Gamma(\nu+k)}{(\pi\nu)^k \Gamma(\nu) |\boldsymbol{\Sigma}|} \left(1 + \frac{1}{\nu}(\mathbf{x}-\boldsymbol{\mu})^H \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)^{-(\nu+k)}. \tag{4}$$

The mean of this distribution is $\boldsymbol{\mu}$ and its covariance matrix is $\frac{\nu}{\nu-1}\boldsymbol{\Sigma}$. Bernoulli distribution of success probability p is denoted by $\mathcal{B}(b; p)$.

Network Parameters

U	Number of users
$\mathcal{Z}_{\mathrm{rx}}$	Set of listening nodes
$\mathcal{Z}_{\mathrm{tx}}$	Set of transmitting nodes
$\mathcal{Z}_{u,\mathrm{dec}}$	Set of nodes to be decoded by the u^{th} node

Radio Parameters

α_0	Roll-off factor
$\mathbf{b}_u$	Transport block of node u
B	Number of data blocks
$\mathrm{BER}_{u,v}$	Bit error rate at node u for decoding node v's message
$\mathcal{C}_u$	Channel code of node u
$\mathbf{d}_u^{(b)}$	b^{th} coded block of node u
$\mathbf{d}_{u,k}^{(b)}$	Q-bit vector corresponding to the k^{th} symbol of the b^{th} data block
$\mathcal{E}_b$	Per-bit energy
$\mathcal{E}_s$	Per-symbol energy
$\mathcal{E}_{\text{tx},u}$	Effective isotropic radiated power of node u
f_c	Carrier frequency
$\mathbf{F}_{\text{WF},u}$	Single-carrier waveform modulator
$\mathbf{G}_{\text{WF},u}$	Single-carrier waveform demodulator
φ	Memoryless mapper
$\varphi_q^{-1}(x)$	q^{th} bit of the symbol x, according to φ
$\mathrm{h}_{u,v}(\tau, t)$	Time-varying channel impulse response between nodes u and v
K	Number of symbols per data block
K_b	Number of bits per transport block
K_c	Number of bits per codeword
K_d	Number of coded bits per data block
L	Delay spread
M	Order of the constellation
m_{up}	Receiver upsampling factor
N'	Number of samples used to transmit a data block
N_0	Thermal noise power spectral density
$\mathrm{PER}_{u,v}$	Packet error rate at node u for decoding node v's message
P_u	u^{th} user's transmit power
Q	Number of bits for encoding a symbol, i.e $Q = \log_2 M$
R	Overall coding and modulation rate
R_c	Channel coding rate
$\mathrm{r}_u^{(b)}(t)$	Signal received by the u^{th} node
$\mathrm{s}_u^{(b)}(t)$	Signal transmitted by the u^{th} node
$\mathrm{SNR}_{u,v}^{(b)}$	Signal-to-noise ratio of node v's b^{th} data block at node u
U	Number of users
W	Effective bandwidth
W_0	Occupied bandwidth
$\mathbf{x}_u^{(b)}$	b^{th} data block of user u
$\mathcal{X}$	Constellation

Small-case scalars corresponding to the definitions above are their iterator indexes, unless specified otherwise. For instance b indexes data blocks from 1 to B.

Introduction

This Ph.D. thesis has been carried out under the collaboration between the "Signal & Communications" group of IRIT-INPT, a CNRS laboratory at the University of Toulouse, France, and the "Waveform Design" team of Thales SIX GTS (former Thales Communications & Security), Gennevilliers, France, through a CIFRE funding, from June 2016 to May 2019.

Context of the Thesis

Wireless communications systems have significantly increasing influence in the modern world, with both the widespread use of innovative consumer goods, and with growing industrial applications and services. The wireless medium is a particularly challenging environment for high-throughput systems: transmitted signals are subject to a variety of non-trivial propagation phenomena such as multi-path fading, shadowing or interference caused by other users. These difficulties require carefully designed protocols at Physical (PHY) and Medium Access Control (MAC) layers, with the use of advanced signal processing algorithms at the emitter and the receiver.

There has been an emerging design paradigm of communications protocols with cooperation at the PHY layer (or cross-layer), in order to cope with the need for relaying techniques in civilian broadband systems such as the 3rd Generation Partnership Project (3GPP) cellular networks or the Institute of Electrical and Electronics Engineers (IEEE) 802.16m WiMAX Wireless Metropolitan Area Network (WMAN). Such transmission strategies challenge the conventional interference avoidance design principle, and instead they seek to create interference intentionally, to either harvest additional information, or to economize spectral and temporal resources, through advanced signal processing and coding schemes.

These techniques naturally gather a significant amount of attention for their applicability in Mobile Ad Hoc Networks (MANETs), which are networks that consist of peer-to-peer transmissions. The mobility of user nodes makes the network topology subject to variations in time, and such networks play a crucial role for many wireless systems that require fast deployment and scalability, such as Wireless Sensor Networks (WSNs) and Private Mobile Radios (PMRs) used in civilian protection, disaster-relief, search-and-rescue operations, and also for tactical networks [Tob87; Haa00; Lu+07]. The use of multi-hop relaying is crucial to cope with mobility and unstable topologies in such networks, hence using cooperative strategies could further increase robustness, or enable sharing radio resources for multiplexing new data. This thesis investigates the design of innovative PHY layer receiver algorithms for future tactical MANETs.

Distributed Cooperation in MANETs

Maintaining a low end-to-end latency in wireless MANETs while ensuring a certain level of quality of service is a challenge due to the vulnerability of multi-hop routes. Increasing the link robustness requires heavy control signalling and coordination, which ends up reducing the spectral efficiency of the system [GW02; Bur+06].

In order to alleviate these issues, distributed cooperation strategies with minimal overhead have been explored in order to either improve robustness or the throughput [And+08]. Indeed, the broadcast and superposition nature of the wireless medium has to be exploited through non-orthogonal transmissions, which has to be met with advanced PHY techniques capable of mitigating intra-system interference.

A fundamental example in this objective is the cooperative broadcast strategy, where every node that has correctly received a message, becomes a relay and actively helps to the successful packet delivery, by re-transmitting simultaneously with the source and other relays [MY04; SSM06; HC10; Cha+19]. A signal received through this process is naturally subject to a significant amount of interference, that is prone to rapid changes, and PHY receivers need to be able to cope with such considerations.

More advanced distributed cooperation schemes use partial Channel State Information (CSI) at the receiver, and control channel feedback in order to coordinate non-orthogonal transmission strategies for increasing the system load [DPP15; ZLL06; HMH16] . Such signalling methods require state-of-the-art PHY signal processing capabilities and moreover, the benefit in their real-world implementation is non-trivial to conclude upon, as they often involve a non-negligible amount of overhead.

In this thesis, the aforementioned challenges for enabling the design of cooperative MANETs will be addressed from two fronts :

- identify from the state-of-the-art of PHY receiver design techniques, the strategies that are likely to be realistic for near-future implementation in real-world platforms,
- ensure that selected receivers' behavior can be predicted with link abstraction techniques for enabling system-level simulation strategies for assessing the impact of the receiver on a cooperative protocol.

Approximate Bayesian Inference for PHY Design

All these challenges for future cooperative MANETs underline that the received signal is subject to a significant amount of perturbations, such as self-interference (e.g. Inter-Symbol Interference (ISI)), intra-system interference or jamming. To mitigate or tolerate such issues, advanced PHY layer signal processing is required. In particular, iterative detection and/or decoding techniques play a significant role at the PHY layer [HEA11; Ngu+15]. These techniques, that originated from turbo codes [BGT93], are well suited to Bit Interleaved Coded

Modulation (BICM) systems [CTB98] and bring significant improvement over conventional signal processing thanks to probabilistic Soft Input Soft Output (SISO) algorithms. Although exact implementation of such algorithms can incorporate prohibitive complexity, there are some low-complexity implementations for realistic PHY design for MANETs [CPC02].

In recent years, there are emerging variational Bayesian methods that enable deriving iterative detection and estimation algorithms for a large variety of signal processing problems. In particular the Belief Propagation (BP) algorithm, which had enabled the design early iterative turbo detection structures [WP99; TSK02], has been extended with several novel iterative techniques [Sen+11; QM07; Han+18]. Hence it is important to assess the impact of these developments for PHY layer receiver design with innovative SISO modules, and identify techniques with attractive complexity-performance trade-off. To better understand these methods, we review the fundamental properties of variational Bayesian inference techniques [Naj11; Rie+13; Per+15]. Within this family of message passing techniques, Expectation Propagation (EP)-based doubly iterative receivers are at the heart of this thesis' major technical contributions.

Alternatively there is a category of emerging low complexity parsimonious signal recovery techniques, called Approximate Message Passing (AMP), which increasingly gathers attention for iterative receiver design [Guo+13; Wu+14; Zha+15a]. There exists many strong connections between these methods, and algorithms based on variational Bayesian methods, hence they should be investigated for clearing the dust on their strengths and weaknesses.

Furthermore, the use of deep learning for PHY layer design is also another hot topic [OH17]. Designer must understand whether these learning algorithms could provide a sensible advantage to the conventional model-based techniques [SDW19; HLW14].

Advanced Turbo Receivers

It has been just over twenty years [GLL97] since filter-based turbo equalization, and more generally, turbo detection with filter-banks have been introduced. Focusing on their use for equalization, Finite Impulse Response (FIR) structures have been widely accepted as a reference, following the works of *Tüchler et al.* [TSK02; TK+02; TS11], based on the methodology of *Wang & Poor* [WP99] or with the works of *Roumy et al.* [Rou00; Rou+01]. While these structures appear to be near optimal for low or medium-low data rates (i.e. for low order constellations with a code rate up to 1/2), their performance is strongly degraded when high spectral efficiency applications are considered. There has been various attempts to improve upon these structures by using turbo Decision Feedback Equalizer (DFE) [LB06; LX11; JM11; Tao16], however they often incorporate a considerable amount of robustness, and the limits and usefulness of DFE-like SISO structures are not well-known.

On the other hand, from an implementation point of view, FIR structures poses various difficulties. In the exact turbo equalization scheme, a covariance matrix has to be inverted for each equalized symbols, and even with complexity-efficient implementation methods [TSK02], they can be prohibitive for hardware. Another issue with the FIR DFE structures, is limitation

caused by the decision feedback latency, which can severely limit the operating frequency of the algorithm.

To this end, Frequency Domain (FD) equalization structures are attractive alternatives, and they are widely used in low-cost mobile terminals for the uplink channel, through Single-Carrier Frequency Domain Equalization (SC-FDE) transmissions [SKJ94]. These structures are attractive due to the parallel processing of each symbol with single-tap filters in the FD, and by using Fast Fourier Transforms (FFTs). The corresponding linear FD turbo equalization structures [TH00; TH01] also achieves near-optimal performance at low spectral efficiency applications. To improve the detection capabilities for high rate systems, several FD structures use equalization and Interference Cancellation (IC) steps successively, by using a type of a soft *block* decision feedback [BT05; Tao15], to which we refer as self-iterations, to avoid confusions with Time Domain (TD) block DFE structures [Kal95].

These two categories of receivers suffer from the use of hard or A Posteriori Probability (APP) based feedback, which violate the extrinsic message principles of turbo structures, and results in unpredictable and irregular behavior due to decisions being strongly correlated to equalized estimates. EP-based receiver design of these structures could be investigated to improve the properties of the soft feedback.

Semi-Analytical Link Abstraction

As announced at the earlier paragraphs of this introduction, one of the main design challenges of cooperative protocols for MANETs, is being able to evaluate the impact of a realistic PHY receiver and channel model. This cannot be carried out with the actual PHY receivers, as multiple links could be active at the same time, and simulating transmissions and the receiver on each link has prohibitive complexity.

PHY layer abstraction methods aim to describe the link behavior with semi-analytical models [Bru+05], this is a further more complex channels for cooperative links where signals may combine coherently. Moreover, with the emergence of novel categories of receivers, the predictability of their asymptotic and finite-length behavior has to be investigated.

Dissertation Outline and Main Contributions

The main objective of this thesis is to explore advanced PHY layer receivers for mitigating crippling amounts of interference that can occur in MANETs with loosely-coordinated cooperative protocols. A secondary objective is to provide a tool to accurately assess the system-level impact of iterative PHY algorithms and cooperative MAC protocols within suitable MANET channel models. The dissertation outline is given in Figure 1 and main contributions are summarized below.

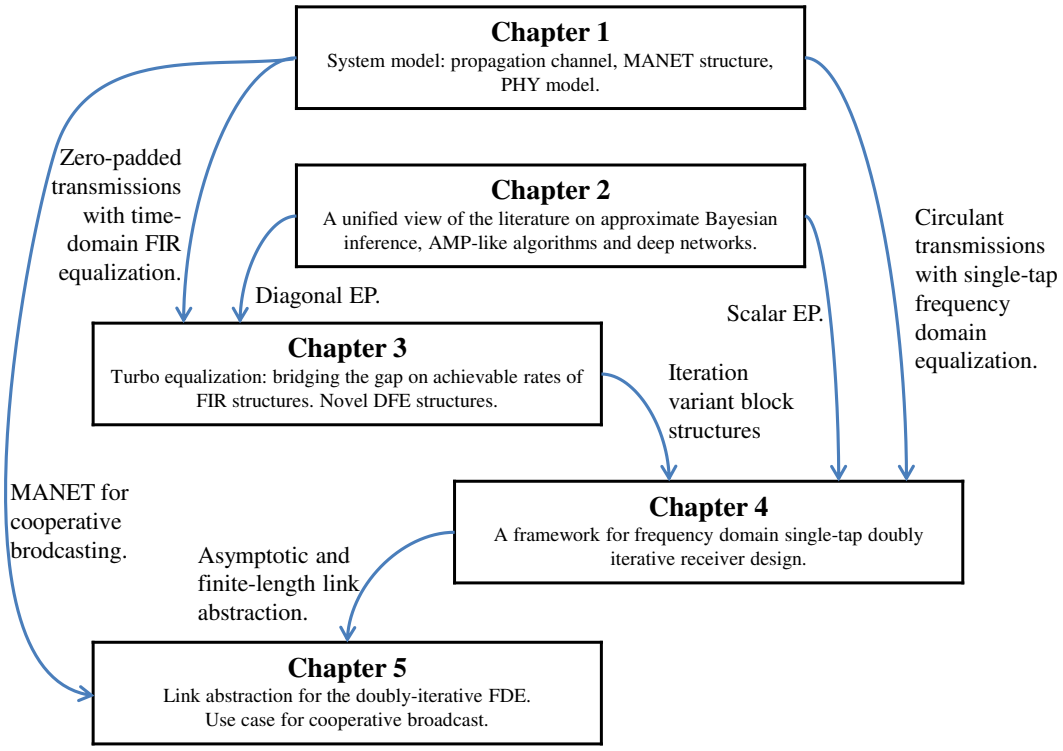


Figure 1: Thesis outline and relationships between different chapters.

Chapter 1: In this context chapter, an introduction to the technical challenges of MANETs is given in order to provide the motivations behind our system model. Notations for a generic single-carrier BICM transmitter and receiver is given, along with definitions of performance metrics, along with some notions on theoretical limitations of communications systems. The gap of progress between asymptotic cooperative schemes and practical relaying protocols is pointed out, in order to justify the need for the design tools for fairly comparing different protocols.

Chapter 2: This chapter is a work of synthesis that summarizes major advances and emerging iterative PHY receiver design techniques, while analytically outlining the similarities and differences between various approximate Bayesian inference techniques (Mean Field (MF), BP and EP) and message passing algorithms (Generalized Approximate Message Passing (GAMP), Vector Approximate Message Passing (VAMP) and Orthogonal Approximate Message Passing (OAMP)). The similarities between such iterative methods, and the state-of-the-art neural network structures, are also underlined, to nuance the potential benefits of using deep learning techniques for receiver design.

The contributions of this chapter have not been yet published as a stand-alone work, but parts of it has been used to provide necessary background for the works in [Şa+19b; Şa+19a].

Chapter 3: This chapter points out several inconsistencies observed in the literature of FIR turbo equalizers, regarding the contradictory behaviors of Linear Equalizer (LE) and DFE. In this regard, a generalized turbo FIR model is exposed, in order to clearly lay out differences and similarities between different algorithms, and to understand the impact of using a hard,or soft decision feedback, and the impact of perfect decision assumption.

A novel APP based and EP-based exact FIR turbo DFE are proposed, along with a novel efficient algorithmic implementation for such structures, based on Cholesky decomposition and Givens rotations. Finally, self-iterated EP-based FIR turbo DFE is explored to further improve the achievable rates of these structures. These contributions have been published in the journal [Şa+18c].

Another original contribution in this chapter aims to reduce the computational complexity of FIR turbo DFE, based on EP or APP, by predicting the reliability of the decision feedback, before using the equalizer, in order to optimize its static filters accordingly. This contribution has been published in the journal [Şa+20a].

Chapter 4: This chapter aims to drastically reduce self-iterated equalization complexity, compared to FIR DFE, through the consideration of low complexity FD receiver design, by following a "scalar EP" framework with double-loop scheduling. To illustrate the capabilities of this framework, first a basic turbo Frequency Domain Equalizer (FDE) is derived, and its finite-length and asymptotic performance is analyzed. This contribution has been published in the conference paper [Şa+18b].

Next, through comparison to AMP-like low-complexity algorithms, our proposal appears to be able to reach lower error rates that structures that would have been based on this literature, if its parameters are properly tuned. Hence, the concept of deep unfolding is used to generate deep neural network based on this receiver, and we have proposed a turbo-receiver-oriented loss function in order optimize this receivers' parameters through learning. This contribution has been published in the conference papers [Şa+19a; Şa+19c].

There has also been various extensions of this framework for different circularized waveforms, for time-varying channel equalization, for multiple-user or multiple-antenna systems has been derived or discussed. This has led to various publications such as [Şa+18a; Şa+18e; Şa+18d].

Finally, the impact of channel and Signal to Noise Ratio (SNR) estimation through mismatched and robust FDE has been analyzed, and a groundwork has been laid out for deriving a joint channel estimator and equalizer based on a three-loop receiver. These contributions are being completed for future publications such as [Şa+20b].

The unifying technical aspect of all these structures is a novel SISO demapper, based on EP, which enabled computing robust extrinsic symbol feedback, for interference regeneration and cancellation. This element has been subject of the patent [Şa+b].

Chapter 5: In this chapter, a simulation framework is proposed for enabling system-level evaluation of cooperative MANET protocols. The core of such methods consist in semi-analytical abstraction of physical links, and to this end, the prediction of the proposed FD method is investigated through asymptotic evolution analysis [YGP07], then by extending effective prediction methods from the literature [VB+10]. Contributions in these topics have partially been presented in [Şa+19b].

An original finite-length prediction method is also presented in this chapter, by exploiting the dispersion of the soft feedback reliability. This method is the subject of the patent [Şa+a].

Finally, in order to illustrate the use of these techniques in an explicit application, the notion of cooperative broadcasting is presented, and the impact of the proposed receiver is investigated at the link level and then system level. Part of this work that focused on modeling the behavior of the cooperative link has been published in [Şa+18d].

Introduction (Français)

Cette thèse de doctorat a été menée grâce à la collaboration entre le groupe "Signal & Communications" du laboratoire CNRS, IRIT-INPT, à l'Université de Toulouse, France, et l'équipe "Waveform Design" de Thales SIX GTS (précédemment Thales Communications & Security), Gennevilliers, France, à travers un financement CIFRE de Juin 2016 à Mai 2019.

Le Contexte de la Thèse

Les systèmes de communications sans-fils ont un impact de plus en plus marquants dans le monde moderne, avec à la fois avec l'usage répandu de produits de consommation innovantes, et avec l'expansion des cas d'usages industrielles et de services. Le milieu sans-fil est un environnement particulièrement éprouvant pour les systèmes haut-débits: les signaux transmises sont susceptibles à une variété de phénomènes de propagation non-triviaux tel que l'évanouissement multi-trajets, le masquage ou l'interférence généré par d'autres usagers. Ces défis nécessitent des protocoles conçus avec soin aux couches PHY et MAC avec l'utilisation des algorithmes de traitement de signal avancés à l'émetteur et au récepteur.

Il y a une paradigme de conception émergeant des protocoles de communication avec coopération au niveau de la couche PHY (ou inter-couches), afin de répondre aux besoins d'utilisation de techniques de relayage dans les systèmes civiles large-bandes tels que le réseau cellulaire du 3GPP ou le WMAN 802.16m WiMAX du IEEE. Tels stratégies de transmission s'oppose aux principes conception conventionnels fondés sur l'évitement d'interférence, en plutôt créant de l'interférence intentionnellement, afin de soit pouvoir récolter de l'information supplémentaire, ou afin d'économiser des ressources spectrales et temporelles, grâce à des techniques avancées de traitement du signal et de codage.

Naturellement, ces techniques attirent beaucoup d'attention par rapport à leur applicabilité dans les réseaux ad-hoc mobiles (MANETs), puisque ceux-ci sont principalement constitués des transmissions pair-à-pair. La mobilité des utilisateurs rendent la topologie du réseau susceptible à des variations dans le temps, et de tels réseaux jouent un rôle clé dans tout système sans-fil nécessitant un déploiement rapide et de la capacité de passage à l'échelle, tels que les réseaux de capteur (WSNs) et les radios mobiles professionnelles (PMRs) utilisés pour la protection des civiles, les opérations d'aide humanitaire ou de sauvetage et aussi pour les réseaux tactiques [Tob87; Haa00; Lu+07]. L'utilisation du relayage par sauts multiples est un point clé pour faire face à la mobilité et des topologies instables dans de telles réseaux, donc l'utilisation des stratégies coopérations pourrait augmenter la robustesse davantage, ou permettre de partager les ressources radios pour multiplexer plusieurs flux de données. Cette thèse étudie la conception d'algorithmes innovants en couche PHY pour les récepteurs des MANETs tactiques futurs.

La Coopération Distribuée dans les MANETs

Maintenir un faible délai de transmission bout-à-bout dans les MANETs sans-fil, tout en assurant un certain niveau de qualité de service, est un défi majeur à cause de la vulnérabilité des trajets à saut-multiples. Augmenter la robustesse du lien nécessite de la coordination et une signalisation excessive sur un lien de contrôle, ce qui réduit l'efficacité spectrale du système [GW02; Bur+06].

Pour remédier ces problèmes, les stratégies coopératives distribués ayant les charges de signalisation les plus légers ont été explorées, afin de soit améliorer la robustesse ou le débit du système [And+08]. En effet, l'aspect favorable du milieu sans-fil à la diffusion et à la superposition des signaux peut être exploité par des transmissions non-orthogonaux avec l'utilisation des techniques couche PHY avancés, capable de mitiger l'interférence intra-système.

Un exemple fondamental à cet égard, est la technique de la diffusion coopérative, où chaque noeud ayant correctement reçu une message, devient un relai et aide activement à la transmission du paquet en le retransmettant simultanément avec la source et les autres relais [MY04; SSM06; HC10; Cha+19]. Un signal reçu par ce moyen subit naturellement une quantité significative d'interférence, qui peut rapidement évoluer, et les récepteurs couche PHY doivent donc affronter ces considérations.

Des stratégies coopératives distribuées plus avancées utilise une connaissance partielle de l'état du canal (CSI) au niveau du récepteur, et utilise un canal de retour pour coordonner les transmissions non-orthogonaux afin d'optimiser la charge du système [DPP15; ZLL06; HMH16]. De telles méthodes de signalisation nécessitent l'utilisation des techniques de traitement du signal de l'état de l'art en couche PHY, et par ailleurs, leur intérêt pour une implémentation réelle est non-trivial à déterminer, puisque les coûts de signalisation sont souvent négligés dans les études.

Dans cette thèse, les défis précédemment cités pour la conception des MANETs coopératives seront adressés sur deux fronts:

- identifier, à partir de l'état de l'art sur les techniques de conception de récepteurs PHY, les stratégies qui sont le plus vraisemblable pour une implémentation dans le futur proche sur des plateformes réelles,
- s'assurer que le comportement des récepteurs choisis peut être prédit avec des techniques d'abstraction de lien afin de pouvoir mener des simulations système permettant d'évaluer l'impact des récepteurs à l'échelle d'un protocole coopératif.

Inférence Bayésienne Approximée pour la Conception PHY

Tous ces défis cités pour les MANETs coopératives futurs soulignent que le signal reçu est sujet à de nombreuses perturbations, tel que l'auto-interférence (e.g. l'interférence entre symbols (ISI)), l'interférence intra-système ou le brouillage. Afin de mitiger, ou tolérer ces

nuisances, on utilise du traitement de signal avancé au niveau de la couche PHY. En particulier, les techniques de détection et/ou de décodage itératives jouent un rôle significatif à cet égard [HEA11; Ngu+15]. Ces techniques, qui proviennent des turbo-codes [BGT93], son bien adaptés aux systèmes de modulation codé à bit entrelacés (BICM) [CTB98] et apportent des gains significatifs comparés aux techniques de traitement classiques, grâce à des modules probabilistes à entrée et à sortie souples (SISO). Tandis que l'implémentation exacte de tels algorithmes peut relever une complexité prohibitive, il existe des implémentations faible-complexités pour permettre leur utilisation dans la couche PHY des MANETs réalistes [CPC02].

Récemment, il y a des méthodes bayésiennes variationnelles émergentes qui permettent de dériver des algorithmes de détection et d'estimation itératives pour une grande variété de problèmes de traitement du signal. En particulier, l'algorithme de propagation de croyance (BP), qui avait permis la conception des premières structures de détection turbo [WP99; TSK02], a été étendu grâce à des nouvelles techniques itératives [Sen+11; QM07; Han+18]. Il est donc important d'évaluer l'impact de ces développements pour la conception de récepteurs PHY avec des modules SISO innovantes, et identifier ceux qui permettent d'avoir un compromis performance-complexité attractive. Afin de mieux comprendre ces méthodes, nous effectuerons la revue des propriétés fondamentales des techniques d'inférence bayésienne variationnelles [Naj11; Rie+13; Per+15]. Parmi cette famille de techniques de passage de message, les récepteurs doublement itératifs à base de la propagation d'espérance (EP) sont au coeur de contributions techniques majeures de cette thèse.

Alternativement, il y a une catégorie émergents de techniques de récupération du signal parcimonieux, à faible complexité, appelées passage de message approximée (AMP), qui attirent de plus en plus l'attention pour la conception de récepteurs itératives [Guo+13; Wu+14; Zha+15a]. Il existe de nombreuses connections entre ces méthodes et les algorithmes basés sur les méthodes bayésiennes variationnelles, et donc ils méritent d'être examinés afin de clarifier leur points forts et leurs faiblesses.

Par ailleurs, l'utilisation de l'apprentissage profond pour la conception de la couche PHY est un autre sujet d'actualité [OH17]. On doit alors comprendre si ces algorithmes peuvent fournir un avantage considérable face aux méthodes conventionnelles basés sur les modèles [SDW19; HLW14].

Turbo Récepteurs Avancés

Cela a fait un peu plus de vingt années depuis l'émergence de turbo égalisation à base de filtres [GLL97], ou de manière plus générale, l'utilisation de banc de filtres pour la détection turbo. Si on se limite à leur utilisation en égalisation, les structures à base de filtres à réponse impulsionnelles finis (FIR) sont largement acceptés comme une référence fondamentale, suivant les travaux de *Tüchler et al.* [TSK02; TK+02; TS11], basés sur la méthodologie de *Wang & Poor* [WP99] ou avec les travaux de *Roumy et al.* [Rou00; Rou+01]. Alors que ces structures semblent être quasi-optimales pour des systèmes à base ou moyen débits (i.e. pour des

constellations d'ordre faibles avec un rendement de code inférieure à 1/2), leur performances sont fortement dégradés lorsqu'on considéré des applications à haut débits. Il y a eu plusieurs tentatives d'amélioration de ces structures en utilisant des turbo égaliseurs avec un retour de décisions (DFE) [LB06; LX11; JM11; Tao16], par contre ces méthodes se focalisent souvent sur la robustesse avec des hypothèses simplificatrices, et les limites et l'utilité des structure SISO semblables à un DFE no sont pas bien connus.

D'autre part, de point de vue de l'implémentation matérielle, les structures FIR posent des divers problèmes. Pour le turbo égaliseur exact, une matrice de covariance doit être inversée pour chaque symbole égalisé, et même avec des approches d'implémentation efficaces en complexité [TSK02], ceci peut s'avérer trop complexe pour implémentation matérielle. Un autre soucis se présente pour les DFE FIR où la latence de retour de décision constitue le chemin critique et peut sévèrement limiter la fréquence de fonctionnement de l'algorithme.

Dans cet objectif, les structures d'égalisation en domaine fréquentielle (FD) sont des alternatives très attractives, et ils sont utilisés de manière répandus dans des bornes mobiles à bas coût pour le canal de la liaison montante, à travers des transmissions mono-porteuses avec égalisation fréquentielle (SC-FDE) [SKJ94]. L'attractivité de ces structures est notamment dû à l'utilisation de filtres à un coefficient pour le traitement en parallèle de toutes les symboles, avec l'utilisation de la transformée de Fourier rapide (FFT). Les structures de turbo-égalisation associées [TH00; TH01] atteignent aussi une performance quasi-optimale pour des applications à bas débit. Afin d'améliorer leurs performances pour des systèmes hauts débits, plusieurs structures FD utilisent les étapes d'égalisation et d'annulation d'interférence (IC) successivement, avec une variante de retour de décision souple sur *block* [BT05; Tao15], lequel on décrira comme une auto-itération, pour éviter la confusion avec les structures DFE block de [Kal95].

Ces deux catégories de récepteurs subissent le même sort, en violant le principe de messages intrinsèque échangés équitables, et les résultats sont souvent imprédictibles, et avec le retour dur du démodulateur est fortement corrélée aux estimés égalisés. La conception de récepteurs à base d'EP peut être investigué pour améliorer les propriétés du retour souple.

Abstraction de Lien Semi-Analytique

Comme mentionné au début de cette introduction, un des défis de conception des protocoles coopératifs pour les MANETs est d'être capable d'évaluer l'impact d'une récepteur PHY et une modèle de canal réalistes. Cela ne peut être effectué avec les vrais algorithmes des récepteurs PHY, puisqu'à cette échelle plusieurs liens peuvent être simultanément actives et simuler ces transmissions et chaque récepteur a une complexité excessive.

Les méthodes d'abstraction de la couche PHY ont pour but de décrire le comportement du lien en utilisant des modèles semi-analytiques [Bru+05], et cela est encore plus compliqué lorsqu'on considère les canaux compliquées mise en jeu par des liens coopératives où les signaux peuvent s'interférer et/ou combiner de façon cohérent. D'autre part, avec l'émergence de nouvelles catégories de récepteurs, leur prédictibilité asymptotique et à taille-fini devront

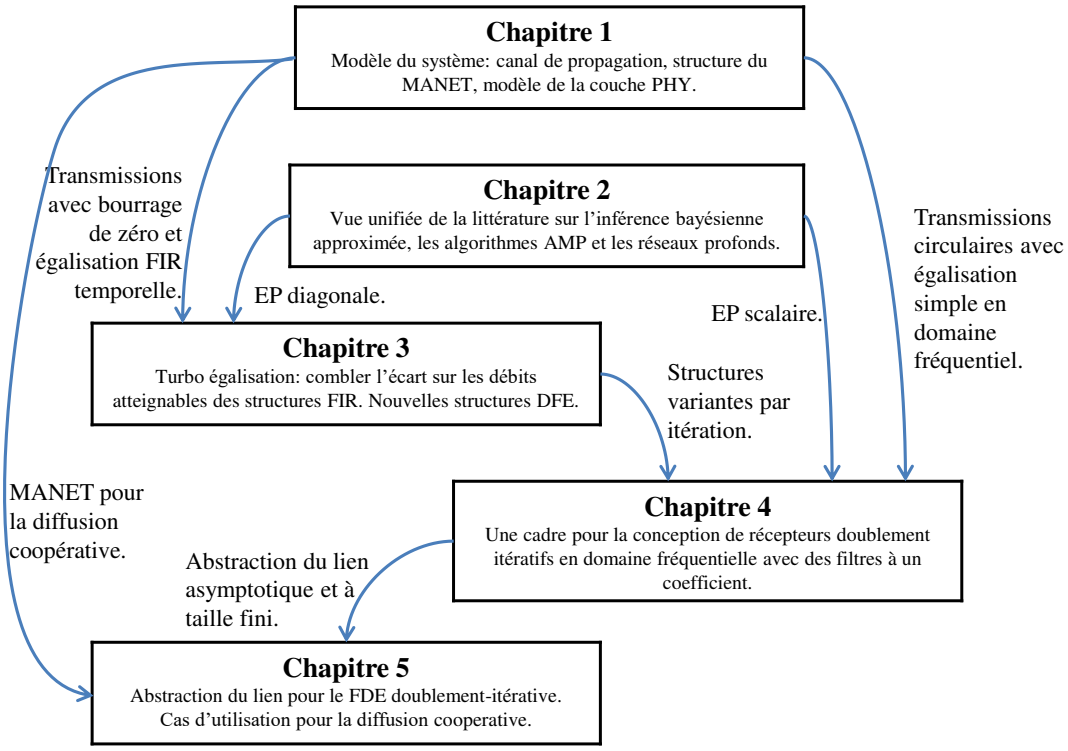


Figure 2: Le plan de la thèse et les liens entre les différents chapitres.

être investigué.

Le Plan de la Thèse et les Contributions Principales

L'objectif principal de cette thèse est d'explorer de récepteur couche PHY avancés pour mitiger la quantité écrasante d'interférence qui peut survenir dans des MANETs utilisant des protocoles coopératives peu coordonnées. Un objectif secondaire est de fournir un outil permettant d'évaluer fidèlement l'impact à l'échelle du système des algorithmes PHY itératives et des protocoles MAC coopératives avec des modèles de canal convenables pour MANET. Le plan de la thèse est illustré dans la Figure 2 et les contributions principales sont résumées ci-dessous.

Chapitre 1: Dans ce chapitre de contexte, on fournit une introduction sur les défis techniques pour les MANETs afin d'expliquer notre motivation sur le modèle de système considéré. Les notations pour un émetteur et récepteur mono-porteuse générique utilisant de la modulation codée à bit entrelacés, i.e. BICM, est présenté, avec la définition des métriques de performance et avec quelques notions théoriques sur les limites systèmes de télécommunica-

tions. On souligne l'écart de progrès entre les schémas coopératifs théoriques, asymptotiques et des protocoles de relayage coopératif, afin de pouvoir comparer les différentes méthodes disponibles pour effectuer un choix de conception.

Chapitre 2: Ce chapitre est une thèse qui résume les avancés majeurs sur les techniques itératives en couche PHY, ainsi que les méthodes émergents, toute en effectuant une comparaison analytique entre divers techniques d'inférence bayésienne approximées (MF, BP et EP) et les algorithmes de passage de massage approximés (GAMP, VAMP et OAMP). Les similarités entre ces structures itératives et les réseaux neuronales de l'état de l'art sont aussi évoqués afin de nuancer le potentiel apport de l'utilisation l'apprentissage profond pour la conception de récepteurs.

Les contributions de ce chapitre n'ont pas été encore publiées comme un article séparé, mais certains passages ont été utilisés pour fournir le fond nécessaire pour les travaux dans [Şa+19b; Şa+19a].

Chapitre 3: Ce chapitre soulève certaines inconsistances, constatées dans la littérature sur les turbo égaliseurs à réponse impulsionnelle fini (FIR), à propos de comportement contradictoire des structures des égaliseurs linéaires (LE) et des égaliseurs à retour de décision (DFE). À cet égard, une modèle de turbo égaliseur FIR générique est donnée afin de plus facilement identifier les similarités et les différences entre les différents algorithmes de l'état de l'art, et de comprendre l'impact de l'utilisation du retour dur ou souple, ainsi que celui de l'hypothèse de retour de décision parfaite.

Nouveaux turbo égaliseurs FIR exactes à base de retours APP et basés sur EP sont proposés, avec une nouvelle proposition d'implémentation efficace pour de tels structures, en exploitant la décomposition de Cholesky et les rotations de Givens. Pour finir, un turbo FIR DFE, auto-itérée, à base d'EP, est explorée afin de davantage pousser les limites de débits atteignables avec ce type de structures. Ces contributions ont été publiées dans l'article de journal [Şa+18c].

Une autre contribution originale de ce chapitre cherche à réduire la complexité calculatoire des turbo DFE FIR, basés sur EP ou APP, en prédisant la fiabilité du retour de décision, avant d'utiliser l'égaliseur, afin d'optimiser ses filtres statiques de manière appropriée. Cette contribution a été publiée dans un article de journal [Şa+20a].

Chapitre 4: Ce chapitre vise à radicalement réduire la complexité de l'égalisation auto-itérée, compare à celle des DFE FIR, à travers l'étude de méthodes de conception pour des récepteurs domaine fréquentielle (FD) à faible complexité, en suivant une approche EP "scalaire" avec un ordonnancement à deux boucles. Afin d'illustrer les capacités de cette approche, dans un premier temps un turbo égaliseur domaine fréquentielle (FDE) simple est dérivée, et ses performances à taille-fini et asymptotique sont analysés. Cette contribution a été publiée dans l'article de conférence [Şa+18b].

Puis, à travers une comparaison avec les algorithmes faible complexité basés sur l'AMP, on montre que notre proposition est capable d'atteindre des taux d'erreur plus faibles que ce que ces structures auraient donnée pour une application en turbo FDE, lorsqu'on calibre correctement leurs paramètres. Ainsi, le concept du dépliement profond est utilisé pour générer un réseau de neurones basé sur ce récepteur, et nous avons proposé un fonction de perte dédiée pour les récepteurs turbo afin d'optimiser les paramètres du turbo FDE à base de EP à travers l'apprentissage. Ces contributions ont été publiées dans les papiers de conférence [Şa+19a; Şa+19c].

Il y a eu aussi divers extension de cette approche fréquentielle pour d'autres formes d'onde circulaires, comme pour l'égalisation de canaux temps-variant, ou pour des systèmes à plusieurs utilisateurs ou antennes. Cela a mène d'autre publications tel que [Şa+18a; Şa+18e; Şa+18d].

Pour finir, l'impact de l'estimation du canal et du SNR à travers des FDE en disparité ou robuste set analysé, et un travail de fond est mise en place pour dériver un égaliseur et estimateur de canal conjointe avec un récepteurs à trois boucles. Ces contributions sont en train d'être complétés pour des publications futures telle que [Şa+20b].

L'aspect unifiant de toutes ces structures fréquentielles est le nouveau démodulateur SISO basé sur EP, permettant de calculer un retour extrinsèque sur les symboles à détecter, ce qui peut être utilisé pour l'annulation ou la régénération de l'interférence. Cet élément est le sujet du brevet [Şa+b].

Chapitre 5: Dans ce chapitre, une méthodologie de simulation est proposé pour permettre l'évaluation au niveau système des protocoles coopératives pour MANETs. Le coeur de cette méthode réside dans l'abstraction semi-analytique des liens physiques et dans cet objectif, on a étudié la prédiction de la méthode FD proposé à travers l'analyse de son évolution asymptotique [YGP07], et puis en étendant des méthodes de prédiction efficaces de la littérature [VB+10]. Les contributions sur ces sujets ont été partiellement présentés dans [Şa+19b].

Une méthode de prédiction originale à taille-fini est aussi présentée dans ce chapitre, en exploitant la dispersion de la fiabilité du retour souple. Cette méthode est le sujet du brevet [Şa+a].

Pour finir, afin d'illustrer l'utilisation de ces techniques dans une application explicite, la notion de la diffusion coopérative est présentée, et l'impact du récepteur proposé est étudié au niveau du lien, puis à l'échelle du système. Une partie de ce travail sur la modélisation des liens coopératifs a été publiée dans [Şa+18d].

List of publications

Peer-Reviewed Journal Articles

J1- S. Şahin, A. M. Cipriano, C. Poulliat, M.-L. Boucheret, "Iterative Equalization with Decision Feedback based on Expectation Propagation", *IEEE Transactions on Communications*, vol. 66, no. 10, pp. 4473-4487, Oct. 2018. [Şa+18c]

J2- S. Şahin, A. M. Cipriano, C. Poulliat, M.-L. Boucheret, "A Framework for Iterative Frequency Domain EP-based Receiver Design", *IEEE Transactions on Communications*, vol. 66, no. 12, pp. 6478-6493, Dec. 2018. [Şa+18a]

J3- S. Şahin, C. Poulliat, A. M. Cipriano, M.-L. Boucheret, "Iterative Decision Feedback Equalization Using Online Prediction", *IEEE Access*, vol. 8, pp. 23638-23649, Jan. 2020. [Şa+20a]

International Conference Articles

C1- S. Şahin, A. M. Cipriano, C. Poulliat, M.-L. Boucheret, "Iterative Equalization Based on Expectation Propagation: a Frequency Domain Approach", *26th European Signal Processing Conference (EUSIPCO)*, Rome, Italy, Sep. 2018, pp. 932-936. [Şa+18b]

C2- S. Şahin, C. Poulliat, A. M. Cipriano, M.-L. Boucheret, "Spectrally Efficient Iterative MU-MIMO Receiver for SC-FDMA based on EP", *IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Bologna, Italy, Sep. 2018. [Şa+18e]

C3- S. Şahin, A. M. Cipriano, C. Poulliat, M.-L. Boucheret, "On Cooperative Broadcast in MANETs with Imperfect Clock Synchronization", *IEEE Military Communications Conference (MILCOM)*, Los Angeles, CA, USA, Oct. 2018. [Şa+18d]

C4- S. Şahin, A. M. Cipriano, C. Poulliat, M.-L. Boucheret, "Evolution Analysis of Iterative BICM Receivers with Expectation Propagation over ISI Channels", *IEEE International Symposium on Information Theory (ISIT)*, Paris, France, Jul. 2019. [Şa+19b]

C5- S. Şahin, C. Poulliat, A. M. Cipriano, M.-L. Boucheret, "Doubly Iterative Turbo Equalization: Optimization Through Deep Unfolding", *IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, İstanbul, Turkey, Sep. 2019. [Şa+19a]

C5- S. Şahin, A. M. Cipriano, C. Poulliat, M.-L. Boucheret, "Doubly Iterative Turbo Equalization: Optimization Through Deep Unfolding", *IEEE 45th International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Barcelona, Spain, May 2020. [Şa+20b]

National Conference Articles

N1- S. Şahin, A. M. Cipriano, C. Poulliat, M.-L. Boucheret, "Récepteurs Itératifs par Propagation de l'Espérance : Optimisation par Dépliage Profond", *Actes du XXVIIème Colloque du Groupement de Recherche en Traitement du Signal et des Images (GRETSI)*, Lille, France, Aug. 2019. [Şa+19c]

N2- Eric Soubigou, Serdar Şahin, Antonio Maria Cipriano, Charly Poulliat and Romain Chayot, "Analyse Multicritères des Performances et de la Complexité des Turbo-égaliseurs à Complexité Réduite à base de Treillis et de Filtres", *Actes du XXVIIème Colloque du Groupement de Recherche en Traitement du Signal et des Images (GRETSI)*, Lille, France, Aug. 2019. [Sou+19]

Patents

P1- S. Şahin, A. M. Cipriano, C. Poulliat, M.-L. Boucheret, "Procédé pour calculer une estimation d'un signal numérique modulé et de sa fiabilité", awaiting publication., time-stamped: Feb. 2018. [Şa+b]

P2- S. Şahin, A. M. Cipriano, C. Poulliat, M.-L. Boucheret, "Méthode de prédiction des performances d'un récepteur itératif basé sur la propagation d'espérance", awaiting publication., time-stamped: Oct. 2019. [Şa+a]

Articles in Preparation or Under Review

R1- S. Şahin, A. M. Cipriano, C. Poulliat, M.-L. Boucheret, "Finite-length Pseudo-random Semi-analytic Prediction for Doubly Iterative Turbo Receivers", in preparation.

Chapter 1

On Distributed Cooperation in Mobile ad hoc Networks

Contents

1.1 Introduction

Common wireless systems, such as the 3GPP cellular networks or the IEEE 802.11 Wireless Local Area Network (WLAN) Wi-Fi, rely on fixed infrastructures, where base stations or access points play a key role to establish connection between nodes. This is the classical "star topology" where a fixed terminal has the computational power to carry out heavy signal processing, or to handle complex network access management for mobile, low-cost nodes. However 3GPP's latest updates in Long Term Evolution Advanced (LTE-A), for 4th Generation (4G) cellular networks, include Device-to-Device (D2D) links where mobile terminals are

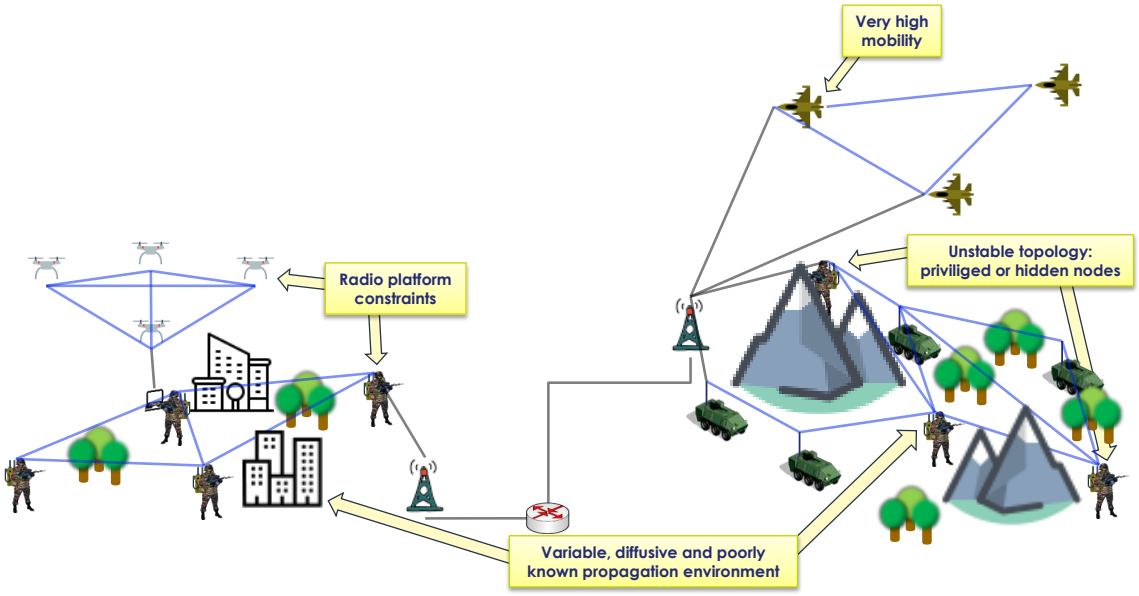


Figure 1.1: Illustration of a tactical network.

used as relays for range extension, on top of base stations [TUY14; KK14; Baz+17]. Similarly, 3GPP's 5th Generation (5G) New Radio's the Machine Type Communications (MTC) use case for future WSN with heterogeneous nodes, considers the use of peer-to-peer relaying, as emerging cannot always rely only on a fixed infrastructure.

In ad hoc networks, as opposed to the fixed topology systems, users communicate directly with each other, i.e. with peer-to-peer transmissions, leading to a mesh network infrastructure [AWW05]. Such systems are self-organizing and they use multi-hop connectivity across the network through user terminals which can act both as a source or as a relaying node for other sources. MANETs are of significant importance for military applications, as they can adapt to very different deployment requirements, as illustrated in Figure 1.1. A widespread application is voice broadcasting among soldiers or vehicles with push-to-talk radios on the tactical edge, but more advanced command-and-control networks also connect MANETs to a command center through satellite or cabled networks [HHK05; LN09; Lar+10; Elm12].

Applications involving low-latency, wide coverage and robust connectivity requirements, such as push-to-talk, pose many design challenges for MANETs [GW02; Bur+06; Fis16]:

- handle issues caused by an unstable dynamic topology,
- ensure robustness in harsh propagation channels with uncertainties,
- mitigate intra-system interference,
- adapt to the challenges of radio hardware constraints.

Multi-hop relaying provides a certain robustness to channel impairments, thanks to possibly available alternative routes, but the resulting end-to-end latency and jitter is non-negligible

[BFY04; GT01]. Clustering is used to address these issues, to some extent, by establishing a sort of infrastructure on a large scale, but within each cluster, distributed protocols and signal processing algorithms are preferred for minimizing overhead [OLT07; MLC17; Gav18]. Finally, the network scalability is another major issue, as the number of nodes increases protocols should be capable of supplying required resources, without causing a significant loss of spectral efficiency due to coordination overhead of the control channel.

Considering the broadcast and superposition nature of radio transmissions in the wireless medium, cooperation and relaying appear to be necessary for achieving the network capacity [OLT07; ADT11; Lim+11; LKK17; LC15a]. Cooperative communications strategies construct virtual Multiple Input Multiple Output (MIMO) transmissions within a network, and bring considerable flexibility on protocol design for improving the throughput or the diversity [Lan02; TV05; KMY+07; DL10]. In various wireless systems with infrastructure, cooperation and relaying concepts have been proven to be of primordial importance, both as a fundamental means to achieve network capacity, or as a means to satisfy highly demanding Quality of Service (QoS) requirements, such as the coverage [Li+12; Cip+12; Zha+15b; HMH16; Tho+17]. But the above-mentioned issues of MANETs increase the toll of control signalling on the spectral efficiency, and hence, conventional strategies which require a high amount of coordination are not applicable for such networks [Tse+02; Ram05; And+08].

Theoretical works on the per-user capacity in wireless ad hoc networks also show that broadcast traffic makes a more efficient use of available resources, compared to unicast transmissions [GK00; KHRR06; WSG08]. This observation has led to the emergence of protocols that extend the notion of "physical link" from a conventional peer-to-peer routes to a multi-nodal routes [LTW04; SGL06; Jak+06; RT07]. Such concepts rely on cross-layer design of MAC and PHY layers to establish connections that are robust to the difficulties posed by the wireless medium [DPB13], but low-cost oscillators and amplifiers of MANETs also induce a significant amount of interference and limitations for the cooperation. While these issues can be avoided for low throughput services [SH03; Cha+19], they are problematic for high-spectral efficiency applications and require advanced receivers for ISI, Inter-Block Interference (IBI) or Inter-Carrier Interference (ICI) mitigation [WGV06; Bla+07; WXY09; Yen+11].

These challenges for future cooperative MANETs show that the received signal is subject to various perturbations, whether it be self-interference, intra-system interference or jamming. To mitigate or tolerate such issues, advanced PHY layer signal processing is required. In particular, iterative detection and/or decoding techniques play a significant role at the PHY layer [HEA11; Ngu+15]. These techniques, that originated from turbo codes [BGT93], are well-suited for BICM [CTB98] and bring significant improvements over conventional methods thanks to probabilistic SISO algorithms. Although exact usage of such algorithms can have prohibitive costs, there are some low-complexity implementations for MANETs [CPC02].

In this chapter, MANET and radio system model assumptions used in this thesis are provided, along with some fundamental notations and definitions. Next, our hypotheses on the wireless propagation channel are exposed. Finally, an overview of challenges for cooperative communications in MANETs is given, in order to emphasize the motivations behind design choices and the PHY layer models used throughout this thesis.

1.2 Network and Radio Model

In this section, network model assumptions and high-level radio properties are exposed for the considered MANET.

1.2.1 Network Structure and Services

In this thesis, the considered MANET is comprised of low-cost terminals with limited capabilities for a vast range of applications including, but not limited to mobile, vehicular or tactical networks. Moreover, the network consists of homogeneous nodes and thus, peer-to-peer physical links are subject to the same set of limitations. These users can be deployed on an area ranging from a few to a hundred square kilometers, with possible clustering into smaller groups.

MANETs for military or civilian protection use a high amount of services that are based on group communications where multiple users are interested in decoding a single source's information [HHK05; Lar+10]. Hence these applications require by nature broadcast or multicast traffic where destination nodes are also well-suited to serve as relay nodes for multi-hop transmissions. However these services can have different latency and robustness requirements:

- push-to-talk (PTT) voice transmissions are often narrowband and involve stringent delay constraints and require a significant amount of robustness versus channel outage,
- situational awareness messages (information sharing) have looser latency constraints,
- video streaming requires wideband transmissions, also with looser latency constraints,
- augmented-reality could require both wideband signals and strict latency constraints.

From the considerations above, robustness, latency and throughput constitute the key performance metrics for MANETs.

Although multi-hop relaying appear as a beneficial factor for increasing coverage and peer-to-peer link robustness [BFY04], as mentioned in the introduction, unicast traffic does not scale well with such networks [GK00], and the network capacity decreases with the number of nodes. This issue is less severe for broadcast traffic [WSG08], and the capacity scaling further improves with cooperative signalling [OLT07]. Consequently, many MAC protocols for unicast traffic in MANETs exploit an underlying broadcast mechanism, along with cooperation, in order to increase efficiency of the overall traffic.

1.2.2 Protocol Stack and Access

As communications system design involves a high variety of disciplines, they are conventionally modeled as a vertically stacked layered architecture to address this complex engi-

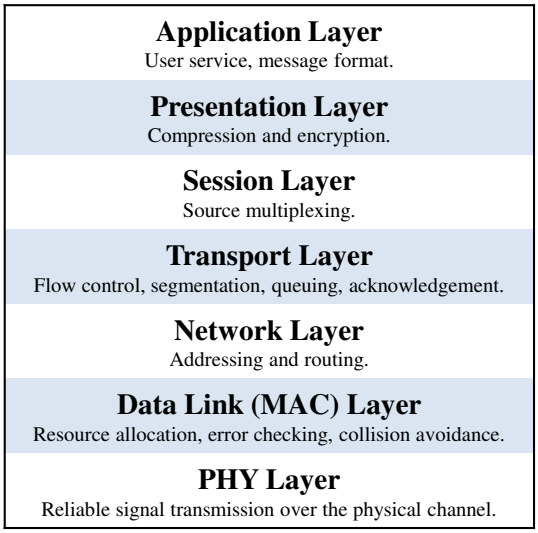


Figure 1.2: Open Systems Interconnection layers.

neering problem. A widespread example is the Open Systems Interconnection (OSI) model which enables interoperability of protocol at various abstraction levels [Elm12], as shown in Figure 1.2. In this thesis, only MAC and PHY layers are considered, with a strong focus on PHY, and a few references to the network layer. To this end, some vocabulary is introduced to be specific about the notions that will be discussed.

The relevance of the traditional OSI model and its derivatives have become questionable for MANETs, due to multi-hop transmissions. For instance, frequent use of relaying causes increased latency, as packets have to go at least up to the network layer for multi-hop routing [Cla+03]. Moreover, the use of cooperative techniques blurs the borders between PHY and MAC layers [Ram05; Elm12]. Hence, MANETs require innovative adaptive protocol stack, depending on whether a node is relaying or not, and they need to use advanced cross-layer algorithms that exploit the notion of cooperative link [RT07; KK14]

The packet that upper layers provide to the MAC layer is called the *payload*, and the MAC layer possibly carries out pre-processing at MAC with segmentation, control signalling for resource allocation or Automatic Repeat Request (ARQ) [Let18], one-or-two-hop topology discovery, or network coding, and produces one or more *transport blocks* as its output. MAC layer adds to the segmented and pre-processed data an header for protocol control and also appends a Cyclic Redundancy Check (CRC) code, for error detection. PHY layer processes a transport block to produce a set of *data blocks*, which are the elementary blocks of symbols that is transmitted on the physical medium. Data blocks are assigned to physical channel resources, according to a cross PHY-MAC resource allocation and transmission protocol, and Frequency Hopping (FH) with a pseudo-random pattern is used at the PHY output, for increasing jamming resilience [AWH07].

Another important functionality of MAC layer is the access regulation for collision avoidance, and intuitive multiple access techniques for MANETs are based on random access,

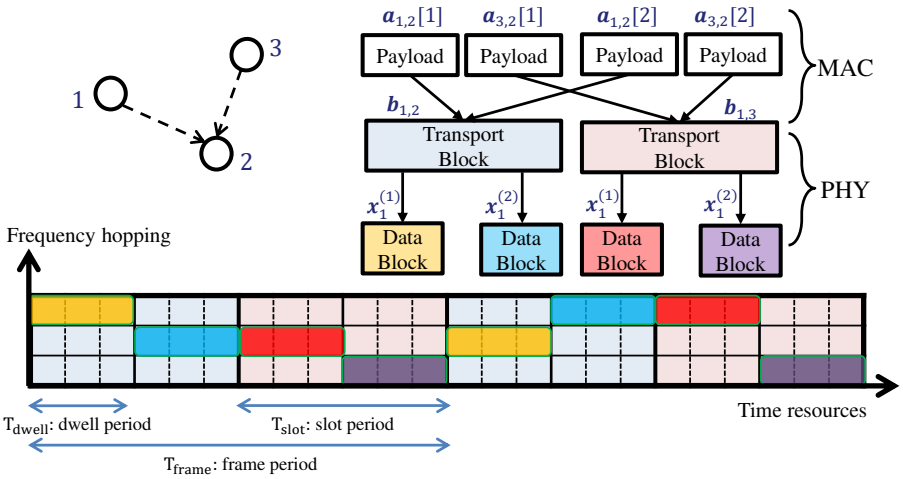


Figure 1.3: Illustration of the resource grid and notations for simple TDMA.

where the medium is attempted to be allocated when needed, by ensuring no other nodes are transmitting. Although this type of protocols are particularly interesting for delay-constrained applications, they require coordination procedures, such as Carrier Sense Multiple Access (CSMA), which can cause large contention and collision resolution periods, and which is vulnerable to the hidden terminal problem. Alternatively, reservation-based MAC protocols with strict schedules, such as Time Division Multiple Access (TDMA) or Frequency Division Multiple Access (FDMA), increase predictability and avoid collision and hidden node issues. In this thesis, the network is considered to be synchronous, with a resolution of a few tens of microseconds, and for simplicity we limit ourselves to MAC protocols with predetermined static schedules. Such access schedules can be combined with spatial reuse or Non-orthogonal Multiple Access (NOMA) for further increased spectral efficiency [BS08].

1.2.3 High-Level Network and Radio Model

Following the assumptions above, here a high-level model of a baseline MANET and radios is described, along with some definitions and notations, and an illustration on Figure 1.3.

The considered MANET has U homogeneous user terminals with varying degrees of mobility, depending on whether radios are on foot (0 - 15 km/h), on terrestrial vehicles (50 - 130 km/h) or on flying vehicles (0 - 400 km/h). Transmitters operate around a carrier frequency f_c, which is typically in the Ultra High Frequency (UHF) band. Radios are equipped with a Global Navigation Satellite System (GNSS) receiver for sharing a large-scale time reference, but the local oscillators of radios can have small clock offsets and oscillator drifts. A network-wide reference time and frequency synchronization procedure is periodically carried out by the upper layers, which allows to assume that residual timing errors are in the order of a microsecond outdoors, and up to a few tens of microseconds otherwise.

Next, we discuss the PHY and MAC resource grid on time, frequency and power domains.

The time duration during which a round of MAC protocol is carried out is called a frame; and with TDMA, each frame is sliced into MAC slots, such that each elementary access operation is scheduled within a slot. Each slot is further divided into dwell periods T_{dwell}, which corresponds to the lowest granularity of time resource, such that the PHY layer carries out its elementary operations at each dwell. From the frequency point of view, the available bandwidth is sliced into elementary frequency bands (similar to the subcarriers or resource blocks in Long Term Evolution (LTE)) and different set of users are assigned to specific sub-bands through FDMA. Finally, a transmit power of P_u is assigned to each radio.

At a given MAC slot, depending on the considered protocol, transmitting radio set $\mathcal{Z}_{\text{tx}}$ and receiving radio set $\mathcal{Z}_{\text{rx}}$ are established. Moreover, for each $u \in \mathcal{Z}_{\text{rx}}$, the set of radios whose message are to be decoded is denoted $\mathcal{Z}_{u,\text{dec}}$, with $\mathcal{Z}_{u,\text{dec}} \subset \mathcal{Z}_{\text{tx}}$. Then, at a node $u \in \mathcal{Z}_{\text{tx}}$, the MAC layer receives payloads $\mathbf{a}_{u,v}$, destined to nodes $v \in \mathcal{Z}_{\text{rx}}$, and produces a set of transport blocks $\mathbf{b}_u$, scheduled to be transmitted on the ongoing time slot. At the physical layer, a transport block $\mathbf{b}_u$ is encoded through a BICM scheme into a set of B data blocks $\mathbf{x}_u^{(b)}$, with $b = 1, \ldots, B$, consisting of complex symbols which are transmitted on the wireless medium with the baseband signal $\mathrm{s}_u^{(b)}(t)$, within a dwell-period, on the assigned band. The signal received at a receiver in the baseband is denoted $\mathrm{r}_u^{(b)}(t)$.

In the following, wireless channel modeling within MANETs is discussed, followed by a description of the PHY layer in Section 1.4. As the thesis is focused on PHY receivers, specific parameters of network and radio models will be revealed gradually as needed throughout the remainder of this chapter, and when will discuss a cross-layer model in the final chapter.

1.3 The Wireless Propagation Channel

In this section, we discuss the modeling of the wireless propagation channel between two single-antenna radios in a MANET. Extension to channels for multiple-antenna radios is briefly covered in the next section.

In particular, we overview some modeling issues that differ from the traditional assumptions used in cellular networks. For instance, D2D nature of transmissions makes the antenna height of user terminals more relevant for path-loss considerations, and the slope of the path-loss changes with distance, according to the amount of non-line-of-sight reflections in the environment. Another limiting issue is caused by the shadowing, which has a very important role in describing the spatial correlation of peer-to-peer links in MANETs, which severely impacts the realistic assessment of communications schemes [Fis+13b]. Finally, the mobility characteristics of such networks also differ from traditional models, as both the transmitter and the receiver are mobile [DL10], and the multi-path propagation characteristics are often harsher than those encountered in cellular networks, due to the nature of the terrain [LN+14].

The focus of this thesis is not on the channel modeling, therefore most of these issues are covered in a simpler way. But the major modeling aspects that will deeply impact the system performance have to be addressed. In order to keep channel parameters simple, we rely on

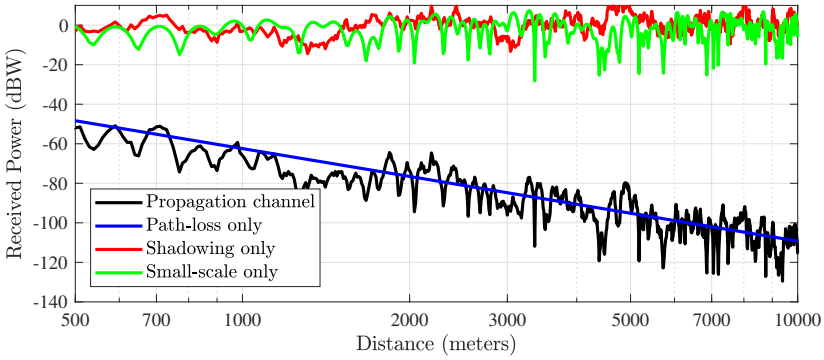


Figure 1.4: Large-scale and small-scale components of a channel.

stochastic linear time variant channel modeling, close to conventional approaches, but with some modifications and parameter-selection based on existing works on MANET channels [DL10; Eri+14; Fis16].

Let the time-varying impulse response of a baseband wireless channel from a node v to a node u, be $\mathrm{h}_{u,v}(\tau, t)$, t being the time parameter and τ being the delay variable over which the channel impulse response is given. In the absence of noise or interference, the baseband signal at the radio u is given by the time-varying convolution product

$$\mathrm{r}_u(t) \triangleq (\mathrm{s}_v * \mathrm{h}_{u,v})(t) = \int_{-\infty}^{\infty} \mathrm{s}_v(t-\tau)\mathrm{h}_{u,v}(\tau, t)\,\mathrm{d}\tau. \tag{1.1}$$

We consider that the propagation medium is symmetric, hence $\mathrm{h}_{v,u}(\tau, t) = \mathrm{h}_{u,v}(\tau, t)$. Moreover, we consider that the channel model can be decomposed as contributions from large-scale and small-scale propagation effects, such that

$$\mathrm{h}_{u,v}(\tau, t) \triangleq (\mathrm{h}^{\mathrm{LS}}_{u,v} * \mathrm{h}^{\mathrm{SS}}_{u,v})(\tau, t) \triangleq \int_{-\infty}^{\infty} \mathrm{h}^{\mathrm{LS}}_{u,v}(\tau-\tau', t-\tau')\mathrm{h}^{\mathrm{SS}}_{u,v}(\tau', t)\,\mathrm{d}\tau' \tag{1.2}$$

In the following, we will discuss the characteristics of these two components.

1.3.1 Large-Scale Propagation Channel

The large-scale wireless propagation effects include a distance-dependent attenuation component, the path-loss, and a position dependent shadow fading component. The path-loss attenuation captures the average attenuating behavior of a channel, and it is impacted by many factors, such as the transmission frequency, antenna heights and the nature of the propagation environment. On the other hand, shadow fading captures the slow variations of channel attenuation around the value of path-loss, which is caused by multiple reflections of transmitted signal from randomly located obstacles in the environment.

The resulting large-scale channel component is considered as the product of these two

attenuating components, i.e.

$$h^{\mathrm{LS}}_{u,v}(\tau,t) \triangleq h^{\mathrm{sh}}_{u,v}(t)h^{\mathrm{PL}}_{u,v}(\tau), \tag{1.3}$$

where $h^{\mathrm{PL}}_{u,v}(\tau)$ and $h^{\mathrm{sh}}_{u,v}(t)$ are respectively the path-loss and the shadowing components.

1.3.1.1 Path-Loss

The path-loss is a specular channel with the attenuating power gain $g^{\mathrm{PL}}_{u,v}$, and the propagation delay $\tau^{\mathrm{prop}}_{u,v}$ such that

$$h^{\mathrm{PL}}_{u,v}(\tau) \triangleq \sqrt{g^{\mathrm{PL}}_{u,v}}\delta(\tau - \tau^{\mathrm{prop}}_{u,v}), \tag{1.4}$$

where $\delta(t)$ is the Dirac delta function. Denoting the distance between u and v as $d_{u,v}$, the propagation delay is given by $\tau^{\mathrm{prop}}_{u,v} \triangleq d_{u,v}/c$, with the velocity of light, $c = 3 \times 10^8$ m/s.

Most widespread models of the path-loss gain are given in the logarithmic domain, hence the log-domain path-loss gain is denoted $\mathrm{PL}_{u,v}/10$, and we have $g^{\mathrm{PL}}_{u,v} \triangleq 10^{-\mathrm{PL}_{u,v}/10}$. Friis' transmission equation describes the free-space electromagnetic wave propagation law between the transmission and the reception antenna, and it follows

$$\mathrm{PL}^{\mathrm{FS}}_{u,v} \triangleq 20\log_{10}\frac{4\pi 10^3}{c} + 20\log_{10}\frac{f_c}{1\mathrm{MHz}} + 20\log_{10}\frac{d_{u,v}}{1\mathrm{km}}. \tag{1.5}$$

However, in real-world propagation environments, the refractive index of different propagation mediums change the slope of the log-domain path loss, especially in terrestrial communications, there is a very sharp change of slope between the Line of Sight (LoS) and Non Line of Sight (NLoS) propagation regions [DL10] [Itub].

Network configurations considered in this thesis are either exclusively located in LoS region (e.g. Air to Air (A2A) communications), or in NLoS region (e.g. terrestrial Ground to Ground (G2G) communications with dense obstacles), hence we keep a simpler, single-slope, frequency-dependent log-distance path-loss model given by

$$\begin{aligned}\mathrm{PL}_{u,v} &\triangleq \mathrm{PL}^{\mathrm{FS}}_{u,v} + \mathrm{PL}^{\mathrm{corr}} + 10(\alpha-2)\log_{10}\frac{d_{u,v}}{1\mathrm{km}} \\ &= 32.44 + \mathrm{PL}^{\mathrm{corr}} + 20\log_{10}\frac{f_c}{1\mathrm{MHz}} + 10\alpha\log_{10}\frac{d_{u,v}}{1\mathrm{km}},\end{aligned} \tag{1.6}$$

where the values of $(\mathrm{PL}^{\mathrm{corr}}, \alpha)$ are dependent on the propagation medium.

For the A2A channels in airborne MANETs, the free-space path-loss model is sufficiently accurate, according to experiments on the field [GRW12], hence $(\mathrm{PL}^{\mathrm{corr}}, \alpha) = (0, 2)$.

On the other hand, for terrestrial communications, many parameters, such as the antenna height or the propagation environment [Fis16] [Itua], impact the path-loss model parameters. Here we are interested in radios with low antenna height, typically at 2 meters high above ground. ITU-R-1546 recommendation provides reference measurements and prediction meth-

ods for point-to-area terrestrial transmissions, however results therein has to be extrapolated to low antenna heights, which involves a lot of environmental assumptions. To simplify the modeling, we rather rely on D2D channel measurements carried out in [Fis16], where 2 meters high antennas were used, and four different environments were tested within the Very High Frequency (VHF) and the UHF bands. Indeed, we will use

- $(\mathrm{PL}^{\mathrm{corr}}, \alpha) = (30.0, 4.68)$ for urban environments,
- $(\mathrm{PL}^{\mathrm{corr}}, \alpha) = (26.0, 3.36)$ for rural environments,
- $(\mathrm{PL}^{\mathrm{corr}}, \alpha) = (19.5, 3.34)$ for hilly environments,
- $(\mathrm{PL}^{\mathrm{corr}}, \alpha) = (17.3, 3.23)$ for mountainous environments.

The urban case is seemingly in accordance with the ITU-R-1546 recommendations, which indicates $(\mathrm{PL}^{\mathrm{corr}}, \alpha) = (27.5, 4.6)$, for 3 meters high antennas, when measured at 400 MHz.

1.3.1.2 Shadowing

Shadowing is of crucial importance for modeling the behavior of wireless mobile networks, as it captures spatial correlations of different links. If neglected, a considerable difference between simulated and real-word performance would be observed. The disruptive effect of this category of fading has to be mitigated with smart protocol design, and it has been the subject of studies on cooperative communications in order to increase link availability [DPB13]. This component is treated a real-valued random variable as the result of a multitude of obstacle layout configurations on the channel, strongly correlated in time and in space.

When the presence of a high number of random obstacles is considered, the shadow fading can be modeled as a zero-mean Gaussian distributed variable in the log-domain path-loss, due to the central limit theorem, and thus $g^{\mathrm{sh}}_{u,v}(t) \triangleq (h^{\mathrm{sh}}_{u,v}(t))^2$ is a log-normal distributed random process, i.e. $10\log_{10}(g^{\mathrm{sh}}_{u,v}(t)) \sim \mathcal{N}(0, \sigma_S^2)$. It is possible in some simulation scenarios to consider that $10\log_{10}(g^{\mathrm{sh}}_{u,v}(t))$ has non-zero mean μ_S, in order to account for fixed obstacles.

The standard deviation of the log-domain variations is both dependent on the environment, and on the frequency. For instance, based on mobile to mobile measurements in UHF, [Fis16] proposes the following model

$$\sigma_S^2 = 0.65 \log^2 \frac{f_c}{1\mathrm{MHz}} - 1.3 \log \frac{f_c}{1\mathrm{MHz}} + K_S \tag{1.7}$$

with $K_S = 6.1$ for urban, $K_S = 6.9$ for rural, $K_S = 7.9$ for hilly and $K_S = 6.2$ for mountainous environments. Alternatively, ITU-R-1546 recommendations indicate, $\sigma_S^2 = 1.6 \log \frac{f_c}{1\mathrm{MHz}} + K_S$, with $K_S = 2.1$ for urban and $K_S = 3.8$ for sub-urban or hilly environments [Itua] . In our simulations, the model in [Fis16] will be used, as its measurements are made in a more similar context to our topic of interest.

The shadowing environment is assumed to be temporally fully correlated, as environmental obstacles are mostly static, or moving slowly relative to the users. Hence, when evaluating the asymptotic behavior of a protocol, the environment will remain static for the duration of a whole MAC frame, and then change randomly, otherwise, when evaluating a certain mobility configuration, the environment can remain static over multiple MAC frames.

As previously stated, spatial correlation plays a significant role in modeling MANETs behavior A simple approach to account for these effects is the use of an exponential auto-correlation on each link (auto-regressive process), which follows adequately the evolution of shadow fading as nodes move [Eri+14]. Nevertheless this approach neglects inter-link correlations, which is detrimental for predicting the outage behaviors [Fis+13b]. Such effects can be captured through two-dimensional mapping of shadow fading by also accounting for the angular orientation of links for with spatial link correlation [FSA04]. This practical approach, originally devised for cellular networks, causes the link between two radios to be asymmetric in MANETs, which goes against common design assumptions used for protocol design.

An alternative approach uses stochastic two-dimensional mapping, through the use of a random obstacle generator, where obstacles have different shapes and sizes, and are modeled as random processes [Fis+13a]. In this case simulated shadow fading can emulate outage behavior that is very close to real-terrain measurements, by determining the obstacle shapes, sizes and density through D2Ds measurements [Fis16]. Such considerations are well-suited to the terrestrial scenarios that will be considered in this thesis, but for aeronautical scenarios there is no shadowing.

1.3.2 Small-Scale Propagation Channel

The small-scale propagation component $h^{\text{ss}}_{u,v}(\tau, t)$ is caused by scattering and reflections of the transmitted wave, which induces time and frequency selectivity. This component is modeled as a wide-sense stationary channel with uncorrelated scattering [TV05].

In this thesis, a channel model with time-invariant L_{ss} discrete components and with time-invariant delays is considered, such that

$$h^{\text{ss}}_{u,v}(\tau, t) \triangleq \sum_{l=1}^{L_{\text{ss}}} a^{\text{ss}}_{u,v,l}(t)\delta(\tau - \tau^{\text{ss}}_l), \tag{1.8}$$

where $a^{\text{ss}}_{u,v,l}(t) \sim \mathcal{CN}(m^{\text{ss}}_l, \sigma^{\text{ss}2}_l)$, and the power spectral density of $a^{\text{ss}}_{u,v,l}(t)$ is the Doppler spectrum $f \mapsto S^{\text{ss}}_l(f)$. $K^{\text{ss}}_l \triangleq |m^{\text{ss}}_l|^2/\sigma^{\text{ss}2}_l$ denotes the Rice factor of the l^{th} tap, with delay τ^{ss}_l and $\mathcal{E}^{\text{ss}}_l \triangleq |m^{\text{ss}}_l|^2 + \sigma^{\text{ss}2}_l$ denotes the l^{th} tap power. In this case, the channel is fully characterized by the channel power-delay profile given by $\{\tau^{\text{ss}}_l, \mathcal{E}^{\text{ss}}_l, K^{\text{ss}}_l, S^{\text{ss}}_l(f)\}_{l=1}^{L_{\text{ss}}}$.

With this model, $h^{\text{ss}}_{u,v}(\tau, t)$ is independent from radios u and v, and it can be considered as an independent and identically distributed random process, between any links, for a given environmental power-delay profile. For accounting for static nodes, it is made dependent on node positions, through a simplified spatial "correlation model", such that, for any node v,

Model	Small-Scale Parameters	
EPA	Delay (μs)	$\{0, 0.03, 0.07, 0.09, 0.11, 0.19, 0.41\}$
	Power (dB)	$\{0.0, -1.0, -2.0, -3.0, -8.0, -17.2, -20.8\}$
	Rice (dB)	$-\infty$
	Doppler	Jakes
EVA	Delay (μs)	$\{0, 0.03, 0.15, 0.31, 0.37, 0.71, 1.09, 1.73, 2.51\}$
	Power (dB)	$\{0.0, -1.5, -1.4, -3.6, -0.6, -9.1, -7.0, -12.0, -16.9\}$
	Rice (dB)	$-\infty$
	Doppler	Jakes / Bi-Gaussian 1
ETU	Delay (μs)	$\{0, 0.05, 0.12, 0.20, 0.23, 0.50, 1.60, 2.30, 5.00\}$
	Power (dB)	$\{-1.0, -1.0, -1.0, 0.0, 0.0, 0.0, -3.0, -5.0, -7.0\}$
	Rice (dB)	$-\infty$
	Doppler	Jakes / Bi-Gaussian 1

Table 1.1: 3GPP Standardized Terrestrial Radio Access Channels.

Model	Small-Scale Parameters	
Hilly	Delay (μs)	$\{0, 0.9, 1.5, 4.8, 5.4, 6.4, 7.2, 8.8, 10.3, 12.2, 13.8, 16.5\}$
	Power (dB)	$\{0.0, -1.0, -7.6, -5.4, -1.6, -2.5, -4.0,$ $-2.7, -6.0, -17.1, -14.3, -14.9\}$
	Rice (dB)	$-\infty$
	Doppler	Jakes / Bi-Gaussian 1 / Bi-Gaussian 2
Mountain	Delay (μs)	$\{0, 1.1, 3.2, 4.7, 13.0, 15.1, 16.1, 21.4, 25.4, 26.9, 31.7, 35.6\}$
	Power (dB)	$\{-5.1, -0.9, -0.5, -8.2, -5.1,$ $0.0, -0.5, -6.3, -7.3, -8.5, -11.4, -13.6\}$
	Rice (dB)	$-\infty$
	Doppler	Jakes / Bi-Gaussian 1 / Bi-Gaussian 2

Table 1.2: Measured Terrestrial Radio Access Channels for D2D.

$h^{\text{ss}}_{u,v}(\tau, t) = h^{\text{ss}}_{u',v}(\tau, t)$, if $d_{u,u'} \leqslant d_\epsilon$, where d_ϵ is a small constant [Eri+14].

In the following, we expose some channel profiles that will be considered in this thesis. For G2G and other terrestrial environments, the standardized models of the 3GPP for LTE are considered with the pedestrian, vehicular and typical urban channels, as seen in Table 1.1 [3gpb]. These models can be complemented with the hilly and mountainous terrain models, shown in Table 1.2 and suited to D2D transmissions in the UHF band, based on measurements in [Fis16]. Note that all taps in these models are assumed to be NLoS components.

For A2A transmissions modeling, a two-tap power-delay profile is considered [Haa02], such that the delay spread is given by the geometric ray tracing, the first tap is a fully deterministic LoS component and the second path is a Rayleigh fading NLoS component. The ratio of power between the two paths corresponds to a more generalized definition of the Rice factor, and it increases with the height of vehicles [GW15a].

The category of Doppler spectrum for NLoS paths, referenced in the terrestrial channel tables above, are selected with the following rule: for taps with delays below 1 μs, Jakes spectrum is user, for those with delay between 1 μs and 10 μs Bi-Gaussian 1, and for those with delay above 10 μs, Bi-Gaussian 2 spectrum are used. For LoS paths, the Doppler spectrum is equivalent to a frequency shift, and the expression of these spectra are given in [Fis16]. These models are characterized by the maximum Doppler shift frequency, which depends on the relative speed between the two nodes.

1.3.3 Overall Propagation Channel Expression

To conclude on this section, the complete wireless propagation channel model is summarized. Recalling that the propagation channel is related to the large and small scale components with $\mathrm{h}_{u,v}(\tau,t) \triangleq (\mathrm{h}^{\mathrm{LS}}_{u,v} * \mathrm{h}^{\mathrm{SS}}_{u,v})(\tau,t)$, the propagation channel is as follows

$$\mathrm{h}_{u,v}(\tau,t) = \sum_{l=1}^{L_{\mathrm{ss}}} a_{u,v,l}(t)\delta(\tau - \tau_{u,v,l}), \tag{1.9}$$

where $\tau_{u,v,l} \triangleq \tau^{\mathrm{ss}}_l + \tau^{\mathrm{prop}}_{u,v}$, and $a_{u,v,l}(t) \triangleq \sqrt{g^{\mathrm{PL}}_{u,v} g^{\mathrm{sh}}_{u,v}(t-\tau^{\mathrm{ss}}_l)}\, a^{\mathrm{ss}}_{u,v,l}(t)$, where the path-loss, shadowing and small-scale parameters are computed according to configurations above.

1.4 Physical Layer Model

In this section we discuss the considered physical layer assumptions for addressing a multi-user Single Carrier (SC) TDMA system over channels selective in time and frequency. With the objective of having a sufficiently flexible layout, a structure slightly inspired from the uplink architecture of LTE [Zha+15b] is used.

For the core PHY operation, we consider the use of BICM [CTB98] with memoryless Quadrature Amplitude Modulation (QAM). A generic SC waveform is considered, in order to make a more efficient use of the high-power amplifier, thanks to lower Peak to Average Power Ratio (PAPR) [SKJ94]. This formalism will allow later in this dissertation to address different instantiates of block SC transmissions. The considered system typically operates around the carrier frequency $f_c = 400$ MHz, where we focus on transmissions over a frequency band of width W, which is shared by all the users. The pseudo-random frequency-hopping is carried out, at the level of each transmitted data block, with shifts greater than W_0, with $W_0 > W$ being the occupied bandwidth. In order to have a reasonable ratio between the occupied and effective bandwidth, the use of pulse-shaping is considered.

A CRC-based error detection functionality is also used along with channel coding at PHY, in order to be able to carry out selective Decode & Forward (DF) relaying capability for cooperation (explained in detail in the upcoming section).

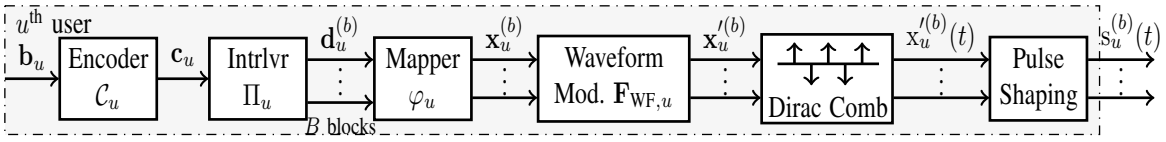


Figure 1.5: Block diagram of a single carrier transmitter with BICM encoding.

In the following, the mathematical description of baseband PHY transmitter and receiver models is provided for the transmission of a data block.

1.4.1 Transmitter

For nodes $u \in \mathcal{Z}_{\text{tx}}$, the transport block $\mathbf{b}_u$, of length K_b, is first processed by a BICM encoder to obtain after modulation B encoded data-blocks $\mathbf{x}_u^{(b)}$ of length K each. Next, a generic SC waveform is used to generate baseband a transmit signal $\mathrm{s}_u^{(b)}(t)$, for each data block. The illustration of the transmitter structure is given in Figure 1.5.

First $\mathbf{b}_u$ is encoded into a codeword $\mathbf{c}$, of length K_c, through forward error correction with a channel code $\mathcal{C}_u : \mathbb{F}_2^{K_b} \to \mathbb{F}_2^{K_c}$, of rate $R_c = K_b/K_c$. Code families that will be typically considered in this thesis are (punctured) convolutional codes, LTE turbo-code (parallel concatenation) with rate-matching or Low-Density Parity Check (LDPC) codes [RL09; RK18]. The resulting codeword is interleaved by an interleaver Π_u which is a bijective function from $\{1, \ldots, K_c\}$ to itself, and we abusively denote the interleaved codeword $\mathbf{d}_u = \Pi_u(\mathbf{c}_u)$ to intend $d_{u,j} = c_{u,\Pi_u(j)}$, for $j = 1, \ldots, K_c$. In numerical simulations, besides LTE rate-matcher, random interleavers are typically used, either using a uniformly random permutation generator or a uniformly s-random permutation generator [DD95].

This interleaved codeword is then segmented into B coded blocks, $\mathbf{d}_u^{(b)}$ of length K_d, $b = 1, \ldots, B$, such that $K_c = BK_d$ and $d_{u,j}^{(b)} \triangleq d_{u,(b-1)K_d+j}$. Next we consider a QAM constellation $\mathcal{X} \subset \mathbb{C}$, with $M = |\mathcal{X}|$, such that memoryless mapper $\varphi : \mathbb{F}_2^Q \to \mathcal{X}$, with $Q = \log_2 M$, maps the sub-vector $\mathbf{d}_{u,k}^{(b)} \triangleq [d_{u,Q(k-1)+1}^{(b)}; \ldots; d_{u,Q(k+1)}^{(b)}]$ of the b^{th} coded block to the symbol $x_{u,k}^{(b)} \in \mathcal{X}$, and we use $\varphi_q^{-1}\left(x_k^{(b)}\right)$ or $d_{u,k,q}^{(b)}$ to refer to $d_{u,k(Q-1)+q}^{(b)}$. The resulting data blocks $\mathbf{x}_u^{(b)}$, $b = 1, \ldots, B$, are of length $K = K_d/Q$, and assuming Identically and Independently Distributed (IID) symbols, the constellation $\mathcal{X}$ is such that, the symbol random process has zero mean (i.e. $\mathbb{E}[x_{u,k}^{(b)}] = \sum_{\alpha \in \mathcal{X}} \alpha/M = 0$) and unitary variance (i.e. $\sigma_x^2 = \text{Var}[x_{u,k}^{(b)}] = \sum_{\alpha \in \mathcal{X}} |\alpha|^2/M = 1$). Hence, the overall encoding rate of the transport block can be also rewritten as $R = R_c Q$.

At this point of the thesis a generic single-carrier waveform modulator $\mathbf{F}_{\text{WF},u} \in \mathbb{C}^{N' \times K}$ is considered, this will enable focusing on Zero Padding (ZP)-SC transmissions in Chapter 3 and on circulant transmissions in Chapter 4, by either using Cyclic Prefix (CP) or ZP-SC with Overlap and Add (OLA). Each data block $\mathbf{x}_u^{(p)} \in \mathcal{X}^K$ is mapped to $N' \geqslant K$ complex

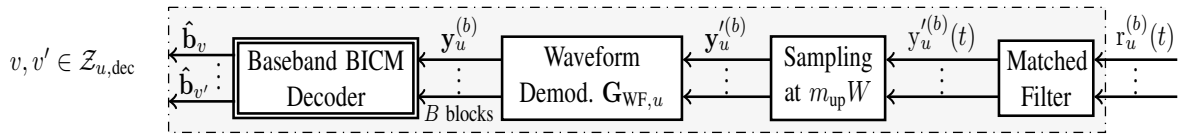


Figure 1.6: Block diagram of a single carrier receiver.

baseband samples by a block linear operator with

$$\mathbf{x}_u'^{(b)} \triangleq \mathbf{F}_{\text{WF},u}\mathbf{x}_u^{(b)} = [x_{u,1}'^{(b)}; \ldots; x_{u,N'}'^{(b)}]. \tag{1.10}$$

The signal constituted by these samples, at sample rate W, is represented by a Dirac comb

$$\mathrm{x}_u'^{(b)}(t) \triangleq \sum_{n=1}^{N'} x_{u,n}'^{(b)} \delta\left(t - \frac{n-1}{W}\right), \tag{1.11}$$

which is an infinite-bandwidth signal but with limited channel uses, such that $N'/W \leqslant T_{\text{dwell}}$. This waveform can be an up-sampled or a spread version of $\mathbf{x}_u^{(b)}$, or its transposition into a specific frequency band.

In order to transmit a band-limited signal, for respecting available spectral resources, and to minimize the ISI at the receiver following Nyquist criterion, a root raised-cosine filter is used for pulse shaping, and the baseband output of the emitter is

$$\mathrm{s}_u^{(b)}(t) \triangleq (\mathrm{h}_{\text{RRC}}^{(\alpha_0)} * \mathrm{x}_u'^{(b)})(t), \tag{1.12}$$

with $\mathrm{h}_{\text{RRC}}^{(\alpha_0)}(\tau)$ being the normalized root-raised-cosine filter of roll-off $0 \leqslant \alpha_0 \leqslant 1$, with a Nyquist bandwidth of W and the occupied bandwidth of the system is $W_0 \triangleq (1+\alpha_0)W$.

Considering the transmit power P_u (in Watts) of the u^{th} radio, assuming a power loss due to a back-off of OBO_{dB} is required at the output of the high power amplifier and a transmit antenna gain of $G_{\text{dB}}^{\text{tx}}$, the effective isotropic radiated power (EIRP) is

$$\mathcal{E}_{\text{tx},u} \triangleq P_u 10^{(G_{\text{dB}}^{\text{tx}} - \text{OBO}_{\text{dB}})/10}. \tag{1.13}$$

In conclusion, the baseband signal transmitted by the node u is $\sqrt{\mathcal{E}_{\text{tx},u}}\mathrm{s}_u^{(b)}(t)$.

1.4.2 Receiver

For nodes $u \in \mathcal{Z}_{\text{rx}}$, assuming an ideal carrier recovery, the observed equivalent baseband signal at the receiver is given by

$$\mathrm{r}_u^{(b)}(t) = \sum_{v \in \mathcal{Z}_{\text{tx}}} \sqrt{\mathcal{E}_{\text{rx},u}} \left(\mathrm{h}_{u,v}^{(b)} * \mathrm{s}_v^{(b)}\right)(t) + \mathrm{z}_u^{(b)}(t), \tag{1.14}$$

where, $\mathcal{E}_{\mathrm{rx},u} = 10^{(G^{\mathrm{rx}}_{\mathrm{dB}})/10}\mathcal{E}_{\mathrm{tx},v}$, $G^{\mathrm{rx}}_{\mathrm{dB}}$ is the receive antenna gain, and $\mathrm{z}_u(t)$ is the noise process at the receiver antenna output, which is a zero-mean Additive White Gaussian Noise (AWGN) with variance whose doubly-sided power spectral density is $N_0/2$. The channel impulse response $\mathrm{h}^{(b)}_{u,v}(\tau,t)$ is obtained with the propagation phenomena described in Section 1.3, and the superscript with the block index b is added in order to distinguish baseband channels used for each data block, in case they are separated through TDMA and the FH.

The considered receiver architecture, given in Figure 1.6, consists in matched filtering and sampling stages, followed by waveform processing for demodulating signal to the baseband and a BICM decoder. Although, for optimal detection, a full match filter on the channel and the pulse shaping is needed, but as the availability continuous channel state is unlikely for many practical real-world systems, we consider a partial matched-filtered receiver [Lat98], i.e. matched-filtering only to the transmit pulse shape filter, with

$$\mathrm{y}'^{(b)}_u(t) = \left(\mathrm{h}^{(\alpha_0)(-*)}_{\mathrm{RRC}} * \mathrm{r}^{(b)}_u\right)(t), \tag{1.15}$$

$$= \sum_{v\in\mathcal{Z}_{u,\mathrm{dec}}} \left(\mathrm{h}'^{(b)}_{u,v} * \mathrm{x}'^{(b)}_v\right)(t) + \mathrm{w}'^{(b)}_u(t), \tag{1.16}$$

where $\mathrm{h}'^{(b)}_{u,v}(\tau,t) \triangleq \sqrt{\mathcal{E}_{\mathrm{rx},u}}\left(\mathrm{h}^{(\alpha_0)(-*)}_{\mathrm{RRC}} * \mathrm{h}^{(b)}_{u,v} * \mathrm{h}^{(\alpha_0)}_{\mathrm{RRC}}\right)(\tau,t)$, with $\mathrm{h}^{(\alpha_0)(-*)}_{\mathrm{RRC}}(t) = \mathrm{h}^{(\alpha_0)}_{\mathrm{RRC}}(-t)^*$. User signals to be decoded, by the BICM decoder, are collected in the set $\mathcal{Z}_{u,\mathrm{dec}}$, and the remaining interferers are treated as noise and grouped into $\mathrm{w}'_u(t)$, along with noise

$$\mathrm{w}'^{(b)}_u(\tau,t) = \sum_{v\in\mathcal{Z}_{\mathrm{tx}}\setminus\mathcal{Z}_{u,\mathrm{dec}}} \left(\mathrm{h}'^{(b)}_{u,v} * \mathrm{x}'^{(b)}_v\right)(t) + \left(\mathrm{h}^{(\alpha_0)(-*)}_{\mathrm{RRC}} * \mathrm{z}^{(b)}_u\right)(t). \tag{1.17}$$

In the following, the delay spread of $\mathrm{h}'^{(b)}_{u,v}(\tau,t)$ is assumed to be less than a constant limit, known by the receiver.

The received continuous signal is sampled at a rate $m_{\mathrm{up}}W$, where $m_{\mathrm{up}} \geqslant 1$ is an up-sampling factor, which will is equal to 1 for the remainder of this chapter and along most of the thesis[1]. Hence, with $m_{\mathrm{up}} = 1$, when considering sampling at an instant t_0, the observed samples are given by

$$y'^{(b)}_{u,n} = \mathrm{y}'^{(b)}_u(t_0 + n/W), \tag{1.18}$$

$$= \sum_{v\in\mathcal{Z}_{u,\mathrm{dec}}} \left(\mathrm{h}'^{(b)}_{u,v} * \mathrm{x}'^{(b)}_v\right)(t_0 + n/W) + \mathrm{w}'^{(b)}_u(t_0 + n/W), \tag{1.19}$$

$$\triangleq \sum_{v\in\mathcal{Z}_{u,\mathrm{dec}}} \sum_{l=1}^{L} h'^{(b)}_{u,v,n,l} x'^{(b)}_{v,n-l+1} + w'^{(b)}_{u,n}, \tag{1.20}$$

for $n = 1,\dots N'$, and where $0 \leqslant t_0 < 1/W$ is given by an adequate fine synchronization algorithm[2], and the equivalent discrete baseband channel has a delay spread limited to L

[1] It will serve its purpose when a fractionally-spaced equalization case will be discussed in Chapter 4.

[2] Here we have assumed that a coarse synchronization algorithm already compensates any delays above a sample period $1/W$, through the use of synchronization signalling.

samples, i.e. $h^{\prime(b)}_{u,v,n,l} = 0$ for $l > L$. The detailed relationship between the discrete and the time-continuous channel models is given in the Appendix A, which is followed by some insights on synchronization and model uncertainties.

In this case, the Channel Impulse Response (CIR) is also described by the vector $\mathbf{h}^{\prime(b)}_{u,v,n} = [h^{\prime(b)}_{u,v,n,1}, \ldots, h^{\prime(b)}_{u,v,n,L}]^T$, and the convolution operations on discrete sequences can be rewritten with the help of a channel matrix as

$$\mathbf{y}^{\prime(b)}_u = \sum_{v \in \mathcal{Z}_{\text{dec}}} \mathbf{H}^{\prime(b)}_{u,v} \mathbf{x}^{\prime(b)}_v + \mathbf{w}^{\prime(b)}_u, \tag{1.21}$$

where baseband samples are vectorized $\mathbf{x}^{\prime(b)}_u = [x^{\prime(b)}_{u,1}; \ldots; x^{\prime(b)}_{u,N'}]$, $\mathbf{y}^{\prime(b)}_u = [y^{\prime(b)}_{u,1}; \ldots; y^{\prime(b)}_{u,N'+L-1}]$ and $\mathbf{w}^{\prime(b)}_u = [w^{\prime(b)}_{u,1}; \ldots; w^{\prime(b)}_{u,N'+L-1}]$ and the $(N' + L - 1) \times N'$ channel matrix is given by

$$\mathbf{H}^{\prime(b)}_{u,v} = \begin{bmatrix} h^{\prime(b)}_{u,v,1,1} & & & & \\ h^{\prime(b)}_{u,v,2,2} & h^{\prime(b)}_{u,v,2,1} & & & \\ \vdots & \ddots & \ddots & & \\ h^{\prime(b)}_{u,v,L,L} & \cdots & h^{\prime(b)}_{u,v,L,1} & & \\ & \ddots & \ddots & \ddots & \\ & h^{\prime(b)}_{u,v,N,L} & \cdots & h^{\prime(b)}_{u,v,N,2} & h^{\prime(b)}_{u,v,N,1} \\ & & \ddots & \ddots & \vdots \\ & & & h^{\prime(b)}_{u,v,N+L-1,L} & h^{\prime(b)}_{u,v,N+L-1,L-1} \\ & & & & h^{\prime(b)}_{u,v,N+L,L} \end{bmatrix}, \tag{1.22}$$

considering that the transmitted signals are not subject to any interference from a previous block. When the channel is quasi-static, i.e. $h^{\prime(b)}_{u,v,k,l} = h^{\prime(b)}_{u,v,k',l}$, for all k, k', l, this channel matrix is a tall Toeplitz matrix.

The noise component in the equation above, when $m_{\text{up}} = 1$, the impact of $\mathrm{z}_u(t)$ remains AWGN, after the receiver matched filter, thanks root-raised cosine filters' properties. Hence $\mathbf{w}^{\prime(b)}_u$ is assimilated to a zero-mean Gaussian noise process with the covariance matrix

$$\mathbf{\Sigma}^{(b)}_{\mathbf{w}',u} = \sigma_w^2 \mathbf{I}_{N'+L-1} + \sigma_x^2 \sum_{v \in \mathcal{Z}_{\text{tx}} \setminus \mathcal{Z}_{u,\text{dec}}} \mathbf{H}^{\prime(b)}_{u,v} \mathbf{H}^{\prime(b)H}_{u,v}, \tag{1.23}$$

with $\sigma_w^2 = N_0 W$ being the contribution of the AWGN noise power.

The baseband waveform demodulator applies the block linear operator $\mathbf{G}_{\text{WF},u} \in \mathbb{C}^{K \times N'}$ in order to yield the effective observation model

$$\mathbf{y}^{(b)}_u \triangleq \mathbf{G}_{\text{WF},u} \mathbf{y}^{\prime(b)}_u = \sum_{v \in \mathcal{Z}_{\text{dec}}} \mathbf{H}^{(b)}_{u,v} \mathbf{x}^{(b)}_u + \mathbf{w}^{(b)}_u, \tag{1.24}$$

where the equivalent baseband channel matrix is given by

$$\mathbf{H}_{u,v}^{(b)} = \mathbf{G}_{\mathrm{WF},u}\mathbf{H}_{u,v}^{\prime(b)}\mathbf{F}_{\mathrm{WF},u}, \tag{1.25}$$

and the noise and interference covariance matrix is $\mathbf{\Sigma}_{\mathbf{w},u}^{(b)} = \mathrm{Var}[\mathbf{w}_u^{(b)}] = \mathbf{G}_{\mathrm{WF},u}\mathbf{\Sigma}_{\mathbf{w}',u}^{(b)}\mathbf{G}_{\mathrm{WF},u}^{H}$.

Finally, the last step at the receiver of the node u consists in BICM detection and decoding, which attempts to recover estimates of $\mathbf{b}_v$ for $v \in \mathcal{Z}_{\mathrm{dec}}$, denoted $\hat{\mathbf{b}}_{u,v}$. The design of this detection and encoding stage is the main focus of this thesis, with a focus on frequency-selective channels.

For evaluating the performance of a such receiver, SNR, Bit Error Rate (BER) and Packet Error Rate (PER) metrics are introduced. By considering the transmission of K_{tr} transport blocks at the PHY layer, we have

$$\mathrm{PER}_{u,v} \triangleq \frac{1}{K_{\mathrm{tr}}}\sum_{k_{\mathrm{tr}}=1}^{K_{\mathrm{tr}}} \mathbb{1}_{\hat{\mathbf{b}}_{u,v}[k_{\mathrm{tr}}]\neq\mathbf{b}_v[k_{\mathrm{tr}}]} \approx \mathbb{P}\left[\hat{\mathbf{b}}_{u,v} \neq \mathbf{b}_v\right], \tag{1.26}$$

and

$$\mathrm{BER}_{u,v} \triangleq \frac{1}{K_b K_{\mathrm{tr}}}\sum_{k_{\mathrm{tr}}=1}^{K_{\mathrm{tr}}}\sum_{j=1}^{K_b} \mathbb{1}_{\hat{b}_{u,v,j}[k_{\mathrm{tr}}]\neq b_{v,j}[k_{\mathrm{tr}}]} \approx \mathbb{P}\left[\hat{b}_{u,v,j}[k_{\mathrm{tr}}] \neq b_{v,j}[k_{\mathrm{tr}}]\right], \tag{1.27}$$

where brackets are used to index different transmit blocks.

There can be multiple definitions of the SNR, depending on the considered receiver stage of node u for observing the signal from a source node v. At the output of the matched-filter, the continuous signal is observed within the band B, and we have

$$\mathrm{SNR}_{u,v}^{(b)} \triangleq \frac{\|\mathrm{h}_{u,v}^{\prime(b)}\|^2\sigma_x^2}{(1+\alpha_0)\sigma_w^2 + \sum_{v'\in\mathcal{Z}_{\mathrm{tx}}\setminus v}\|\mathrm{h}_{u,v'}^{\prime(b)}\|^2\sigma_x^2}, \tag{1.28}$$

where the norm is defined on the second order Lebesgue space, i.e. $\|\mathrm{h}_{u,v}^{\prime(b)}\|^2 = \int_{-\infty}^{\infty}|\mathrm{h}_{u,v}^{\prime(b)}(t)|^2\mathrm{d}t$ and following baseband sampling of the complex envelope (i.e. band W) we define

$$\left.\frac{\mathcal{E}_s}{N_0}\right|_{u,v}^{(b)} \triangleq \frac{\left\|\mathbf{H}_{u,v}^{(b)}\right\|^2\sigma_x^2}{\sum_{v'\in\mathcal{Z}_{\mathrm{dec}}\setminus v}\left\|\mathbf{H}_{u,v'}^{(b)}\right\|^2\sigma_x^2 + \sigma_w^2}, \tag{1.29}$$

where the matrix norm function is the Frobenius norm. Finally, the definition of the SNR at the bit level is given by defining the energy of a bit as $\mathcal{E}_b \triangleq \mathcal{E}_s/R$, and

$$\left.\frac{\mathcal{E}_b}{N_0}\right|_{u,v}^{(b)} \triangleq \frac{1}{R}\left.\frac{\mathcal{E}_s}{N_0}\right|_{u,v}^{(b)}. \tag{1.30}$$

1.5 Cooperation in MANETs

In this section we overview the fundamental limits and metrics for wireless communications, and then discuss major developments related to cooperative transmissions and their implications. A main observation is that successful cooperative protocols require a significant amount of coordination overhead, which is prohibitive for a variety of real-world applications. Finally, examples of practical distributed cooperative protocols suited to MANETs are given.

1.5.1 Fundamental Limits on Peer-to-Peer Communications

For the past 30 years, there has been a surge of innovative radio access techniques at the PHY and MAC layers, which allowed the spectral efficiency to increase by few orders of magnitude. Part of this success is due to state-of-the-art signal-processing algorithms for precoding, detection or channel estimation and Forward Error Correction (FEC) techniques. These innovative techniques are the fruit of years of theoretical research which has looked for solutions for maximizing the spectral efficiency of communications systems. This sub-section provides a summary of fundamental metrics for characterizing wireless communications.

1.5.1.1 Mutual Information and Channel Capacity

A key metric for evaluating communications efficiency is the Mutual Information (MI). Considering random variables x and y that model respectively a transmitted and a received signal, with a joint Probability Density Function (PDF) $p(x, y)$, then MI is defined as

$$I(x;y) \triangleq \mathbb{E}_{p(x,y)}\left[\log_2 \frac{p(x,y)}{p(x)p(y)}\right] = \mathbb{E}_{p(x,y)}\left[\log_2 \frac{p(y|x)}{p(y)}\right], \tag{1.31}$$

which is a non-negative quantity, measured in bits, and the term $p(y|x)$ is the channel transition probability density for a discrete memoryless channel [CT12]. $I(x;y) = 0$ if and only if x and y are independent, and otherwise this quantity can be interpreted as the amount of information y (or x) conveys about x (or y). More formally, the mutual information is a pseudo-distance between distributions $p(x, y)$ and the product of the marginals $p(x)p(y)$, with $I(x;y) = D_{\mathrm{KL}}(p(x,y)\|p(x)p(y))$, where D_{KL} is the Kullback-Leibler divergence, defined for two not-necessarily normalized PDFs $p(x)$ and $q(x)$ by

$$D_{\mathrm{KL}}(p(x)\|q(x)) \triangleq \mathbb{E}_{p(x)}\left[\log_2 \frac{p(x)}{q(x)}\right] + \mathbb{E}_{q(x)}[1] - \mathbb{E}_{p(x)}[1]. \tag{1.32}$$

Although this measure is a directed divergence, i.e. $D_{\mathrm{KL}}(p(x)\|q(x)) \neq D_{\mathrm{KL}}(q(x)\|p(x))$ in general, mutual information is a symmetric measure. Note that, the two latter terms in Equation (1.32) are correction factors for improper distributions (i.e. un-normalized) [Min+05].

The capacity of a discrete memoryless channel $p(y|x)$ is the maximum amount of mutual information possible, between the input x and the output y, for any possible input distribu-

tions $p(x)$, with a power constraint $\mathbb{E}[|x|^2] \leqslant P_x$, i.e.

$$C \triangleq \max_{p(x), \mathbb{E}[|x|^2] \leqslant P_x} I(x;y). \tag{1.33}$$

Given these definitions, the formalization of the communication theory and information theory began with the fundamental *noisy-channel coding theorem* [Sha48] , which states the existence of a channel code of rate $R \leqslant C$ which can achieve error-free communications over this channel. These concepts can be directly extended to distributions on random vectors.

In the case of a real AWGN channel, with input $x \in \mathbb{R}$ and output $y \in \mathbb{R}$, such that $y = hx + w$, with $w \sim \mathcal{N}(0, \sigma_w^2)$, and $h \in \mathbb{R}$, then the maximum achievable channel capacity is reached for $p(x) = \mathcal{N}(0, P_x)$ and it is given by

$$C = \frac{1}{2} \log_2 \left(1 + h^2 \frac{P_x}{\sigma_w^2} \right). \tag{1.34}$$

The extension to the complex AWGN channel, with input $x \in \mathbb{C}$ and output $y \in \mathbb{C}$, such that $y = hx + w$, with $w \sim \mathcal{CN}(0, \sigma_w^2)$, and $h \in \mathbb{C}$, results in the following capacity expression

$$C = \log_2 \left(1 + |h|^2 \frac{P_x}{\sigma_w^2} \right), \tag{1.35}$$

which is reached for $p(x) = \mathcal{CN}(0, P_x)$. For a Gaussian vector channel with input $\mathbf{x} \in \mathbb{C}^k$, and output $\mathbf{y} \in \mathbb{C}^n$, such that $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}$, $\mathbf{w} \sim \mathcal{CN}(\mathbf{0}_n, \mathbf{\Sigma_w})$, and $\mathbf{H} \in \mathbb{C}^{n \times k}$, the maximum achievable capacity is reached for $p(\mathbf{x}) = \mathcal{CN}(\mathbf{0}_n, \mathbf{\Sigma_x})$, and it is given by

$$C = \max_{\mathrm{tr}(\mathbf{\Sigma_x}) \leqslant n P_x} \log_2 \det \left(\mathbf{I} + \mathbf{H} \mathbf{\Sigma_x} \mathbf{H}^H \mathbf{\Sigma_w}^{-1} \right). \tag{1.36}$$

This maximization problem can be solved by applying the water-filling algorithm on the LDL decomposition of covariance matrices [CT12].

In many real-world systems, the input symbol distribution of a discrete memoryless channel cannot be optimized due to many hardware constraints, such as the PAPR limitations caused by the high power amplifiers. In this case, constellation-constrained achievable rates of the system yield a more meaningful measure of system capacity. For instance, the coded modulation capacity is computed by considering uniform IID symbols from the constellation $\mathcal{X}$, hence $C_{\mathrm{CM}} \triangleq I(x;y)$ with $p(x) = \sum_{\alpha \in \mathcal{X}} |\mathcal{X}|^{-1} \delta(x - \alpha)$. In this case, the capacity on a scalar memoryless channel $p(y|x)$ is rewritten as

$$C_{\mathrm{CM}} = Q - \frac{1}{M} \sum_{x \in \mathcal{X}} \mathbb{E}_{p(y|x)} \left[\log_2 \frac{\sum_{x' \in \mathcal{X}} p(y|x')}{p(y|x)} \right]. \tag{1.37}$$

Another practical capacity measure is the BICM capacity, C_{BICM} which is computed at the

bit-level, considering uniform IID bits [CTB98],

$$C_{\text{BICM}} = Q - \frac{1}{2}\sum_{q=1}^{Q}\sum_{\beta=0}^{1}\mathbb{E}_{p(y|x)}\left[\log_2\frac{\sum_{x'\in\mathcal{X}}p(y|x')}{\sum_{x\in\mathcal{X}_q^\beta}p(y|x)}\right], \tag{1.38}$$

with $\mathcal{X}_q^\beta = \{\alpha \in \mathcal{X}, \varphi_q^{-1}(\alpha) = \beta\}$, $\beta \in \mathbb{F}_2$. In general, constellation-constrained achievable rates do not have closed-form expressions, and although they can be computed for discrete memoryless channels with small dimensions, their computation in more real-world situations (e.g. frequency-selective channels) is quite involved. In Chapter 2, a practical approach for estimating C_{BICM} through Extrinsic Information Transfer (EXIT) analysis will be presented.

1.5.1.2 Characterization of Fading Channels

Previously described notions of capacity were given with the assumption that the channel remains static for the whole duration of communications. This is rarely the case in wireless networks, as explained in sections above, and in particular, the considered system uses P separate data blocks to transmit a codeword, along with FH, during which channel state is likely to change abruptly. Hence alternative capacity metrics are needed for evaluating the achievable performance in such channels.

In the following, in addition to the noise, other channel parameters are also random processes, and to discuss the characteristics of a system, the instantaneous information rate $I(\bar{\mathcal{E}}_s/N_0)$ is defined, as a random process parameterized by the average received baseband SNR $\bar{\mathcal{E}}_s/N_0$ [TV05; KMY+07].

When the coherence time of the channel is small enough with regards to the transmission length, the channel is fast fading, and typical values of the channel have likely occurred on the codeword. Hence the expected value of the mutual information is defined as the ergodic capacity of the channel, with $C_{\text{erg}} \triangleq \mathbb{E}[I(\bar{\mathcal{E}}_s/N_0)]$.

Alternatively, when a received codeword has only gone through a small number of random channel states, the notion of delay-limited capacity is used, along with the concept of outage probability [TV05; KH00]. The outage probability is defined as $P_o(\bar{\mathcal{E}}_s/N_0, R) = \mathbb{P}[I(\bar{\mathcal{E}}_s/N_0) \leqslant R]$, for $R \geqslant 0$, and the system is characterized by the triplet $(\bar{\mathcal{E}}_s/N_0, R, P_o)$. Then, the ϵ-outage capacity is defined as follows

$$C_\epsilon = \max_R \left[P_o(\bar{\mathcal{E}}_s/N_0, R) = \epsilon\right], \tag{1.39}$$

and note that this notion of capacity does not guarantee error-free communications.

Let us consider that a system is designed to communicate at a rate $R(\bar{\mathcal{E}}_s/N_0)$, a function of the $\bar{\mathcal{E}}_s/N_0$, then, the multiplexing gain ρ of the system is the SNR-normalized rate

$$\rho(\bar{\mathcal{E}}_s/N_0) \triangleq \frac{R(\bar{\mathcal{E}}_s/N_0)}{\log \bar{\mathcal{E}}_s/N_0}. \tag{1.40}$$

Moreover the triplet $(\bar{\mathcal{E}}_s/N_0, R, P_o)$ of slow fading channels, imposes the notion of the diversity gain Δ, through the diversity-multiplexing trade-off

$$\Delta(\rho) \triangleq \lim_{\bar{\mathcal{E}}_s/N_0 \to \infty} -\frac{\log P_o(\bar{\mathcal{E}}_s/N_0, \rho(\bar{\mathcal{E}}_s/N_0) \log \bar{\mathcal{E}}_s/N_0)}{\log \bar{\mathcal{E}}_s/N_0}. \tag{1.41}$$

The diversity gain is the asymptotic slope of the outage probability at high SNR, and the fundamental trade-off above describes the link between the spectral efficiency and the robustness of a wireless communication scheme [TV05]. This metric is used to evaluate the efficiency of multi-antenna or multi-user wireless transmission protocols, which is especially practical for cooperative communications [LTW04], however, some recent schemes have generalized this concept to finite-SNR diversity-multiplexing trade-off [FL15].

1.5.2 On Cooperative Communications and Network Information Theory

In this section, a overview of major information theoretical results is provided for communications networks with more than two nodes. This discussion outlines how conventional multiple-access and broadcast operations were affected by the concept of cooperation.

First, we summarize known results on the capacity of multiple access and broadcast channels. For evaluating the limits of multiple-user systems, the concept of capacity region is used, which consists in set of simultaneously achievable rates of involved users. For instance, when considering a two-user access channel with input messages $\mathbf{x_1}$ and $\mathbf{x_2}$, and output $\mathbf{y}$, such that $\mathbf{y} = \mathbf{H_1 x_1} + \mathbf{H_2 x_2} + \mathbf{w}$, with $\mathbf{w} \in \mathcal{C}(\mathbf{0}, \mathbf{\Sigma_w})$, the achievable rate regions with perfectly known channels at the receivers is the set of tuples (R_1, R_2) that satisfy

$$R_1 < I(\mathbf{x_1}; \mathbf{y} | \mathbf{H_1}, \mathbf{H_2}, \mathbf{x_1}), \tag{1.42}$$

$$R_2 < I(\mathbf{x_2}; \mathbf{y} | \mathbf{H_1}, \mathbf{H_2}, \mathbf{x_2}), \tag{1.43}$$

$$R_1 + R_2 < I(\mathbf{x_1}, \mathbf{x_2}; \mathbf{y} | \mathbf{H_1}, \mathbf{H_2}), \tag{1.44}$$

and such regions for multiple-access can be generalized to any number of users [TV05]. The capacity of such channels can be achieved through superposition coding and either joint decoding, or separate decoding with Successive Interference Cancellation (SIC), and practical mechanisms for doing doing so, is among the focus of ongoing research on NOMA [Wan+15]. A PHY design problem with multiple-access is discussed in the Section 4.6 of this thesis.

The dual problem of multiple-access is broadcasting from one node to multiple nodes. The capacity of such channels are not known to this day, but the best known rates are given by the dirty paper coding technique [Cos83], which uses superposition coding at the transmitter, with the full channel knowledge, and also requires either a joint decoder or a SIC at the receivers. Rate-splitting is a sub-optimal variant, which uses simpler linear precoding and partial CSI, and it is being theoretically evaluated for future networks [Cle+16; Liu+19a].

There are only few configurations where the explicit capacity of wireless networks are known, however, the proofs used for achievable rates still provide hints on operations required by spectrally-efficient protocols. Alternatively, cut-set bound provides an upper bound to the

capacity, which is an important reference for evaluating transmission strategies [KMY+07].

In multi-node transmissions mechanisms based on broadcasting or multiple-access, the notion of cooperation is weak, and it is often simply called facilitation. The explicit notion of cooperative transmissions arose in the information theory community with the three terminal models of Van der Meulen [Meu71], where a source communicates with a destination, with the help a of third node, which is dedicated to relaying. The capacity of this "relay channel" is still not known to this day, there are well-known achievable rate theorems which have gathered a lot of attention in the literature [CG79]. In particular,

- Decode & Forward (DF) is a two round scheme where the source transmits its message in two phases. The relay decodes the message during the first phase, and then helps to the source on the second phase with a well-suited codebook. This scheme operates well when the relay is close to the source, otherwise it can even degrade the direct link.
- Compress & Forward (CF) removes decoding at the relay, such that the relay reliably transmits its compressed observations in the second phase, alongside the source's message. This scheme offers interesting gains when the relay is close to the destination.
- Partial DF is a variant of DF, where the relay only decodes a portion of the source's message, to avoid degrading the performance when the source-relay channel is poor.
- Hybrid CF-DF scheme combines partial DF and CF, by enabling the relay to joint decode a portion of source's message, and compress the remainder. This scheme achieves the best known rates for a full-duplex relay.

The schemes described above are information theoretical theorem proofs, enabled through block-Markov and superposition encoding, and backward joint-decoding. Moreover, partial DF based schemes require perfect CSI at the transmitter and the relay, and CF based schemes require perfect CSI at the destination and the relay. These schemes were extended to full and half duplex relay networks [KGG05], but their closeness to the cut-set bound is unknown.

On the other hand, the concept of network coding [ALY00] proved that cut-set bound is achievable on multi-source, multi-relay networks with noiseless links, through block-Markov superposition coding across messages of different sources. A decade later, quantize-map & forward (QMF) has extended both network coding and CF, through the use of a novel approach to modeling Gaussian channels as deterministic linear models [ADT11], which also enabled proving that partial DF and CF achieve capacity within a few bits [OD13]. Alternatively, noisy network coding (NNC) has also emerged [Lim+11; KH11], and it appears to be one of the multiple-access schemes closest to the relay network capacity. The hybrid CF-DF scheme for single relay channel is extended to general networks by combining noisy network coding with partial DF [LC15b]. Distributed DF (DDF) has been proposed recently as an extension of partial DF for one-source multi-destination relay networks [LKK17], and it appears to be one of the techniques closest to the capacity, for carrying out broadcast on relay networks. The coding theorems above were later unified by a general coding scheme in [LC15a].

Nevertheless, the coding strategies above are not mature enough for practical code design in various situations. For instance, all the schemes above are for full-duplex radios, and many of them require perfect CSI on each link at the transmitter, or worse, at each active node of the network. To address these issues, some recent works focus on more specifically in finding schedules for optimally using half-duplex nodes [CTK16], whereas others attempt to find practical encoding and decoding schemes for block-Markov superposition codes [Ma+15]. But there needs still to be years of investigations before characterizing, from an information theoretical point of view, the robustness or vulnerability of these schemes for the wireless medium. To this end, a significant portion of the literature focused on sub-optimal, but simpler cooperative protocols,by using practical coding schemes and half-duplex relaying.

A practical cooperation scheme based on DF has been proposed for code division multiple access in [SEA03], with the use of block Markov coding and interference cancellation. Cooperative protocols for TDMA and FDMA is proposed in [LW03; LTW04], using selective DF and amplify & forward relaying with distributed space-time block codes. Selective DF is a simplified sub-optimal implementation of partial DF, where the relay uses a CRC check to decide whether to forward the decoded codeword or not, and amplify & forward behaves similarly to a naive CF relaying with a unitary compression rate and some power allocation. Asymptotic diversity-multiplexing trade-off of these schemes are evaluated in fading channels. A coded cooperation scheme was proposed in [HN02], where the concept of DF is incorporated into channel code design, which provides jointly diversity gain and coding gain. A soft DF strategy is proposed in [SV05], where probabilistic, "soft" symbol estimates are relayed, outperforming both selective DF and Amplify & Forward (AF). A practical coding scheme was also proposed for CF relaying in [HL06], through Wyner-Ziv coding. In [ZV05], the asymptotic performance of various selective DF relaying strategies are evaluated on a network scale, as a generalized Hybrid Automatic Repeat Request (HARQ), that can use diversity combining or code combining, along with different cross-layer relay selection strategies. On another note, [WGV06] showed that cooperative diversity is also available through delay dithering, as it can create frequency-selective diversity. A review of practical network coding techniques is given in [TB+16], and implementations of noisy network coding are given in [MVB16; HMH16]. DF-based cooperation has also been theoretically incorporated in a variant of 5G NOMA [DPP15], and partial DF has been theoretically evaluated for D2D transmissions [Tho+17].

1.5.3 Some Distributed Cooperative Cross-Layer Protocols for MANETs

The non-exhaustive examples of practical cooperative schemes above do indeed show promising results for the communications networks they have been investigated for, by abandoning the traditional concept of MAC-layer, based on interference avoidance. Cooperative approaches either enable constructive interference at the receiver, or they facilitate interference cancellation at the destinations.

However, even with practical code design and half-duplex radios, most of the schemes above require perfect CSI between many channels, and optimal code parameters are strongly dependent on topology, and coordination required for this is ill-suited to MANETs. On top of

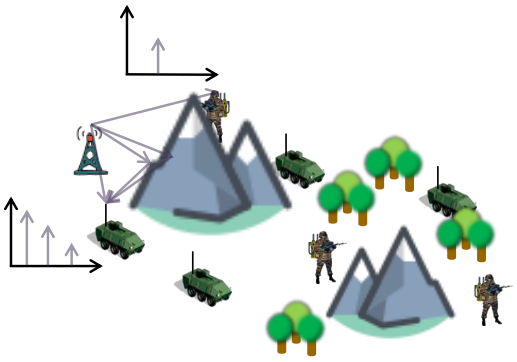
node scheduling overhead, one needs to account for the loss of spectral efficiency that would be caused by such coordination and information exchange for networks where the topology and physical channels are dynamic. Such issues have been considered as major limitations for system design [And+08], and the need for reconsidering the concept of "information theory" for MANETs was underlined. In particular, for fair comparisons of protocols, three items are needed for evaluating network capacity

- finite-length/delay-constrained outage evaluation,
- accounting for the concept of multi-nodal and multi-hop links, and exploit them to more efficiently (e.g. constructive interference, virtual MIMO or network coding),
- accounting for overhead and channel estimation, as they have significant impact on the protocol performance.

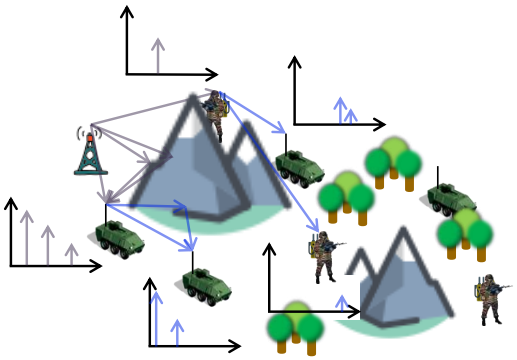
This is a very complex design problem, and such complete studies have only been carried out for some particular protocols to this day. Moreover, differences in modeling physical layer abstraction limit the ability to compare different works. It is important to define appropriate performance metrics for evaluating the capabilities and limits of cooperative protocols for MANETs [Rau13; NS12]. Due to these difficulties, the literature on cooperation in MANET is more limited, especially when high throughput or high fidelity applications are targeted, and most of the practical schemes seek distributed DF relay selection techniques either to perform unicast or multicast/broadcast transmissions.

For instance, for unicast transmissions, various centralized and decentralized protocols for DF relay selection are compared in [XC11; XC12; FXC13], by accounting for the overhead for relay selection and the channel estimation, whereas the literature on relay selection for 4G D2D has moved on the study higher layer aspects related to buffering, power allocation or network coding etc. While in the centralized scheme, the destination uses a control signal to elect the best relays, depending of the CSI it has at hand, in the decentralized scheme, relays use timers that are initialized according the CSI they have obtained from destination's pilot sequence. In both cases there are risks of collision or interference and for fair comparison, these strategies have to be compared with relevant metrics and accurate PHY model. Relay selection techniques for unicast can further be enhanced with HARQ strategies, by carrying out code or diversity combining at potential relays [ZV05].

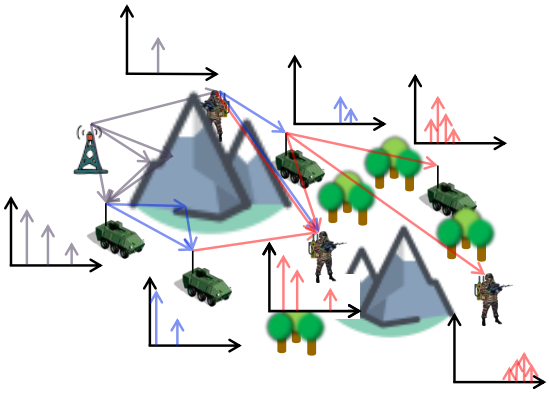
As evoked in Section 1.2.1, for MANETs considered in this thesis, broadcast traffic plays a more significant role, cooperative techniques for flooding is an important research axis. Indeed traditional non-cooperative flooding is inefficient due to excessive occurrence of contention periods [Tse+02] or due to unnecessary repetitions caused by resource-inefficient naive flooding [WC02]. A more popular method to achieve broadcast traffic in MANETs is to use optimized link state routing (OLSR) protocol's multi-point relay functionality [Cla+03]. In this category of algorithms, each radio elects its relay nodes in order to cover its entire two-hop neighborhood. Such protocols use control messages to update each radio's elected relay topology, and for optimizing these protocols' parameters, and for fairly assessments their performance, refined system-level simulators are needed [NS12; NS15].



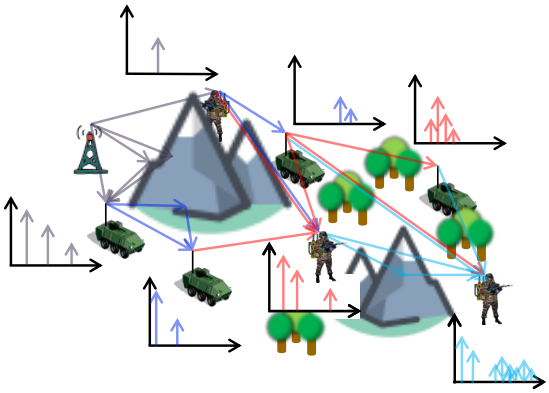
(a) Phase 1: a source transmits and two radios decode, and become relays.



(b) Phase 2: source and two relays transmit, two radios decode, another one fails.



(c) Phase 3: source and four relays transmit, two radios decode and the last node fails.



(d) Phase 4: source and six relays transmit, the last node decodes.

Figure 1.7: Cooperative broadcast example, with the illustration of CIR.

Alternatively, in [MY04], accumulative broadcast was considered for minimum-energy broadcasting in loosely-synchronized networks with limited local information, thanks to the use of selective DF [LTW04]. Other approaches [SH03; Jak+06] use simultaneous participation (i.e. non-orthogonal access) of multiple nodes for the re-transmission of a broadcast packet, which is referred to as *cooperative broadcast*. Such techniques are attractive for MANETs [Ram05]. However non-orthogonal cooperative broadcast generates at the receiver an artificial multi-path channel given by the combinations of the signals from all active relays, thus potentially increasing the frequency and time selectivity perceived by the receiver [WGV06; HKW04].

This concept is illustrated in Figure 1.7 and it has been formulated for WSNs with Opportunistic Large Array concept [SH03], for tactical networks with Barrage Relay Networks [HC10] and its variants [GHK16; Yen+11], and for WLANs with Constructive Interference for

Low Rate Wireless Personal Area Networks [Cha+19]. However while Cooperative Broadcasting (CB) achieves a significant amount of robustness, there is a loss of maximum throughput, due to the underlying diversity-multiplexing trade-off [LW03].

1.6 Conclusion

In this chapter, we have discussed the ongoing change of design paradigm based on cooperation in wireless networks, through the use of relaying functionalities at the PHY-layer.

Theoretical works on these techniques reveal that cooperative transmissions could significantly increase the effective system load in wireless networks, by making use of non-orthogonal access techniques, and advanced PHY layer coding and interference mitigation algorithms. As such techniques often use theoretical bounds to assess their gain in system performance, the problem of designing near-capacity PHY receiver algorithms needs to be addressed, in order to cope with the intra-system interference generated by such protocols.

Hence, the first objective of this thesis is to identify powerful techniques from the state-of-the-art PHY signal processing literature, and to propose practical low-complexity solutions.

Another major roadblock for cooperative wireless transmissions is brought by the amount of control signalling they require, in order to share parameters such as the CSI. When considering decentralized networks, with MANETs being an example, even the most fundamental aspects of cooperation, such as relay selection, becomes a highly complex problem to solve due to the lack of infrastructure. Indeed, it is non-trivial to determine whether overhead required by cooperative protocols remain reasonable when compared to the real-world potential improvements brought by cooperative relaying, with the use of realistic PHY behavior. Hence it is necessary for MANET protocol design, to be able to evaluate the impact of cooperative cross-layer transmissions within a system-level simulator.

Consequently, the second major objective of this thesis is to propose receivers with predictable behavior for enabling assessing distributed cooperative protocols' performance with system level simulators.

With these two objectives at hand, this chapter introduced a MANET system model with homogeneous radios and single-carrier waveforms with BICM-based PHY layer, and the related channel model is presented.

Chapter 2

Advanced Receiver Design: from Approximate Inference to Learning

Contents

2.1 Introduction

This chapter discusses the advances in PHY layer iterative receiver design techniques, with a promising category of algorithms for achieving near-optimal detection performance with reasonable complexity. Although our major technical contributions in this thesis revolve around the use of Expectation Propagation (EP) message passing algorithm for turbo detection, there exists a shared background with a variety of other signal processing techniques which are increasingly used in communications theory. This chapter provides a synthesis on fundamentals of iterative detection and decoding, and as we were often required to compare our approach to these techniques, this chapter goes beyond a simple state of the art: it analytically discusses the differences and similarities of some categories of approximate statistical inference and message-passing algorithms in the context of turbo detection. Moreover these algorithms are discussed along with other emerging techniques from the broad area of machine learning, with similar objectives and notations, to discuss the potential use of emerging deep learning techniques for receiver design.

Indeed, while receiver techniques that achieve optimal detection and decoding performance are well-founded within the classical estimation and detection theory [Kay93], their computational and memory complexity is a major roadblock for the majority of practical applications with real-time embedded receiver platforms. In some cases, even the complexity involved for offline computation of optimal performance bounds is prohibitive. Hence, overcoming this complexity limitation of modern receivers is a major objective for enabling their implementation on practical radio platforms with scarce resources.

With the emergence of turbo codes, and the associated probabilistic decoding algorithm in early nineties [BGT93], near-optimal and practical channel coding techniques have become possible, with an affordable complexity, after nearly half a century later from the formulation of the noisy channel coding theorem of Shannon [Sha48]. This has led to a surge of interest in the *turbo principle* and the related "soft" signal processing algorithms at PHY-layer receivers that aim to approach the theoretical link capacity by iteratively refining detection and/or decoding performance. More specifically, we have witnessed the re-discovery of LDPC codes, with effective iterative decoders and optimized code design techniques [Gal62; MN96; RSU01], and the emergence of turbo demodulators [BSY98], turbo equalizers [GLL97; Fij+00; LGL01; TSK02; LB03; TS11], multi-antenna or multi-user turbo detectors [WP99; BC02; Wit+02; CVB04] that achieve outstanding performance, even in extreme channels.

In this chapter, different theoretical concepts that has led to development of receivers based on the turbo principle are discussed, and recent advances that can further refine these techniques are investigated. Beyond a simple overview of state of the art techniques, one major objective of this chapter is to provide a unified view of different existing and emerging detection techniques, and these fundamentals simplify the derivations in next chapters and also provide a clearer view on our contributions. To this end, we first define the optimal receiver and then discuss sub-optimal but practical receiver architectures that are investigated in the literature. In particular, Bit Interleaved Coded Modulation (BICM) systems play an enabling role for the aforementioned turbo principle and iterative detection and decoding

architectures constitute a promising approach to carry out near-optimal decoding [CTB98].

We then take a step back and discuss the concept of *probabilistic graphical models*, in order to introduce the well-established Belief Propagation (BP) algorithm and show how such models are used for probabilistic receiver design [KFL01a]. The resulting BICM receivers consists of the well-known iterative structure with a Soft Input Soft Output (SISO) detector and a SISO decoder that exchange probabilistic information. The success and the limitations of BP for deriving these structures is discussed, along with their asymptotic analysis, through the concept of Extrinsic Information Transfer (EXIT) function analysis [TB00; Hag04].

Next, the broader area of approximate Bayesian inference methods is summarized in order to outline the design of alternative SISO detector algorithms, which in particular use low-complexity *self-iterations*. Unlike traditional turbo receivers, BICM receivers involving these detectors have a double-loop architecture and the inner-detection loop can bring several performance enhancements. Although Bayesian inference problems arise in many artificial intelligence topics, most of the theoretical results on algorithm analysis, such as BP, was enabled by the progress made for free energy approximation in statistical physics [YFW03]. Indeed, several techniques used for the optimization of the approximate free energy has contributed to the generalizations of BP and to alternative approximate inference algorithms. In order to clarify the specificity of these algorithms, we briefly summarize conclusions drawn in this field and overview developments that has led to the Expectation Propagation (EP) algorithm [Min01; Min+05]. Receivers based on EP has started seeing light even in 3GPP specification procedures [3gpd] , hence investigating its uses goes beyond academic ventures. Recent trends in parsimonious signal detection with Approximate Message Passing (AMP) algorithms are also discussed [DMM09; Ran11; MP17], along with their implications for self-iterated receiver design. In particular, using *state-evolution* techniques devised for these algorithms, the asymptotic behavior of EP-based self-iterated algorithms appear to be predictable.

Finally, we also discuss the implications of recent developments on deep learning for PHY layer receiver design [OH17], by drawing attention to some interesting parallels between successful deep convolutional neural networks and the self-iterated detectors based on approximate inference and AMP. Next, the concept of deep unfolding is discussed, which appears to be a promising approach for near-optimal practical receiver design [HLW14].

To illustrate these concepts without loss of generality and clearly, through this chapter we will consider a single-user BICM system which encodes a transport block $\mathbf{b}$ into a single complex data block $\mathbf{x} \in \mathbb{C}^K$, following the procedure depicted in Section 1.4.1, with $B = 1$. The receiver observes, at the baseband, $\mathbf{y} \in \mathbb{C}^N$ with a linear detection model

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}, \tag{2.1}$$

where $\mathbf{H} \in \mathbb{C}^{N \times K}$ and $\mathbf{w} \sim \mathcal{C}(\mathbf{0}_{N\times 1}, \mathbf{\Sigma}_{\mathbf{w}})$. This is a simplified presentation of the receiver model in Section 1.4.2, in Equation (1.21), but even more complex multi-user detection problem could be rewritten as a linear detection problem through vectorization, and the physical distinctions between systems are incorporated in $\mathbf{H}$ and on the encoding of $\mathbf{x}$[1].

[1] This is illustrated in an example in Chapter 4.

2.2 On Receiver Structures

In this section, some major categories of receiver architectures are discussed with regards to their closeness to the optimal detection performance and computational complexity.

2.2.1 Optimal Detection and Decoding

The receiver that operates with minimum PER performance will be called the optimum receiver and it consists in satisfying the Maximum A Posteriori (MAP) criterion. Indeed, assuming the knowledge of perfect CSI, we have

$$\begin{aligned}\hat{\mathbf{b}} = \arg\min_{\tilde{\mathbf{b}}} \mathbb{P}\left[\tilde{\mathbf{b}} \neq \mathbf{b}|\mathbf{y}, \mathbf{H}, \mathbf{\Sigma_w}\right] &= \arg\max_{\tilde{\mathbf{b}}} \mathbb{P}\left[\tilde{\mathbf{b}} = \mathbf{b}|\mathbf{y}, \mathbf{H}, \mathbf{\Sigma_w}\right], \\ &= \arg\max_{\mathbf{b}} p\left(\mathbf{b}|\mathbf{y}, \mathbf{H}, \mathbf{\Sigma_w}\right), \end{aligned} \tag{2.2}$$

where $p\left(\mathbf{b}|\mathbf{y}\right)$ is the a posteriori Probability Density Function (PDF) on $\mathbf{b}$ given the observation $\mathbf{y}$. Moreover, as transport block bits are uniformly and identically distributed, this criterion is also equivalent to Maximum Likelihood (ML) criterion, i.e.

$$\hat{\mathbf{b}} = \arg\max_{\mathbf{b}\in\mathbb{F}_2^{K_b}} p\left(\mathbf{y}|\mathbf{H}, \mathbf{\Sigma_w}, \mathbf{b}\right). \tag{2.3}$$

The implementation of this receiver for any generic transceiver would require evaluating values of $\mathbf{y}$ for 2^{K_b} possible transport block $\mathbf{b}$. Although this might be doable for very short-block communications systems, in general, its implementation complexity is prohibitive.

2.2.2 Separate Detection and Decoding

Earlier tricks to circumvent optimal receiver's complexity consisted in decoupling the problem of detection and decoding. Indeed, due to the transmitter structure, ML criterion can be rewritten as a marginal of the joint distribution on intermediary variables

$$p(\mathbf{y}|\mathbf{H}, \mathbf{\Sigma_w}, \mathbf{b}) = \int_{\mathbf{x},\mathbf{d},\mathbf{c}} p(\mathbf{y}, \mathbf{x}, \mathbf{d}, \mathbf{c}|\mathbf{H}, \mathbf{\Sigma_w}, \mathbf{b})\, \mathrm{d}\mathbf{c}\, \mathrm{d}\mathbf{d}\, \mathrm{d}\mathbf{x} \tag{2.4}$$

where the argument can be factored through conditional independence as

$$p(\mathbf{y}, \mathbf{x}, \mathbf{d}, \mathbf{c}|\mathbf{H}, \mathbf{\Sigma_w}, \mathbf{b}) = \underbrace{p(\mathbf{y}|\mathbf{x}, \mathbf{H}, \mathbf{\Sigma_w})}_{\text{channel}} \underbrace{p(\mathbf{x}|\mathbf{d})}_{\text{mapping}} \underbrace{p(\mathbf{d}|\mathbf{c})}_{\text{interleaver}} \underbrace{p(\mathbf{c}|\mathbf{b})}_{\text{FEC}}. \tag{2.5}$$

Thanks to BICM encoding, considering that

$$p(\mathbf{d}|\mathbf{c}) = \delta(\mathbf{d} = \Pi(\mathbf{c})) = \prod_{j=1}^{K_c} \delta(d_j = [\Pi(\mathbf{c})]_j) = \prod_{j=1}^{K_c} \delta(c_j = [\Pi^{-1}(\mathbf{d})]_j), \tag{2.6}$$

when K_c is sufficiently large, and the interleaver is random, coded bits at either side of the interleaver are considered to be uncorrelated. Separate detection and decoding takes advantage of this property and further assumes that the detection stage ignores the channel code by considering the coded bits to be IID. Thanks to this approximation, the detection is carried out only based on observations and the channel, and then the outcome is used for decoding the transport block. However, an important limitation of this approach comes from the loss of capacity due to the use of BICM [CTB98], compared to the coded modulation capacity. The use of Gray mapping significantly reduces the loss of capacity, but the loss can still remain important for energy-efficient systems operating at low spectral efficiency.

In earlier approaches, the separate detectors seek to satisfy MAP or ML criteria (which are equivalent as $p(\mathbf{d})$ is treated as a uniform distribution) by constraining estimated code bits to lie in $\mathbb{F}_2$, which results in hard decisions. For instance, in the context of ISI channels, ML sequence estimation has been carried out by the Viterbi algorithm [Vit67]. Alternatively, a sub-optimal approach is to perform ML solely on the channel likelihood to estimate a transmitted symbol x_k, constrained to belong to $\mathbb{C}$ rather than $\mathcal{X}$, and then perform ML bit demapping to get $\mathbf{d}_k$. This latter approach leads to filter-based structures.

Soft-Output Detection However, more advanced receivers with separate detection architectures avoid the use of hard decisions in order to prevent loss of information at the detection step. To circumvent this issue, "soft" information on detected coded block bits is carried by bit Log-Likelihood Ratios (LLRs)

$$L(d_j|\mathbf{y},\mathbf{H},\mathbf{\Sigma_w}) \triangleq \log \frac{\mathbb{P}[d_j=0|\mathbf{y},\mathbf{H},\mathbf{\Sigma_w}]}{\mathbb{P}[d_j=1|\mathbf{y},\mathbf{H},\mathbf{\Sigma_w}]}, \tag{2.7}$$

which convey probabilistic information on the quality of estimation of binary variables. This output is characteristic of a Bernoulli-distributed prior PDF, with

$$p(d_j) = \frac{1}{1+e^{-L(d_j)}}\delta(d_j) + \frac{e^{-L(d_j)}}{1+e^{-L(d_j)}}\delta(d_j-1), \tag{2.8}$$

which is to be used for channel decoding through the PDF $p(c_j) = p(d_{\Pi^{-1}(j)})$. This operation can either be carried out with soft-output estimators to compute the Equation (2.7), which, for the ISI channel case consists in soft-output Viterbi or Balk Cocke Jelinek Raviv (BCJR) algorithms [HH89; Bah+74]. For filter-based structures, the filtering stage also estimates uncertainty on equalized symbols $\hat{\mathbf{x}} \in \mathbb{C}^K$, i.e. $\mathbf{\Sigma}_{\hat{\mathbf{x}}} = \mathrm{Cov}(\hat{\mathbf{x}})$, in order to carry out soft-output ML demapping.

Soft-Input Channel Decoding Soft-input channel decoding is derived from the marginal b_j of the joint MAP estimation of having a codeword $\mathbf{c}$ and a transport block $\mathbf{b}$, i.e.

$$p(b_j) = \sum_{\mathbf{b}^{\setminus b_j}} \sum_{\mathbf{c}} p(\mathbf{b}|\mathbf{c}) \prod_{j=1}^{K_c} p(c_j), \tag{2.9}$$

which yield APP LLRs

$$L(b_j) = \log \frac{\mathbb{P}[b_j = 0]}{\mathbb{P}[b_j = 1]}. \tag{2.10}$$

Note that $p(\mathbf{b}|\mathbf{c}) = \delta(\mathbf{c} = \mathcal{C}(\mathbf{b}))$ is another hard constraint based on Dirac distribution, and the computation of this marginals changes drastically according to the structure and properties of the codebook [RL09]. Hence, final decisions are taken on the sign of LLRs with

$$\hat{b}_j = \begin{cases} 0 & \text{if} \quad L(b_j) \geqslant 0, \\ 1 & \text{otherwise.} \end{cases} \tag{2.11}$$

In this situation, the soft-input decoding require probabilistic algorithms that account for probabilistic information on coded bits in order to carry out channel decoding.

2.2.3 Joint Detection and Decoding

The concept of joint detection and decoding architectures consists in designing detectors and decoders that are aware of one another, and that carry out their operations accordingly.

For BICM transmissions, the coded bits at either side of the interleaver can no longer be considered to be IID, as the correlations brought by the code or the channel has to be accounted for. Hence a particular instance of this category of receivers are iterative detection and decoding architectures, which uses separate modules for detection and decoding but by ensuring that they exchange probabilistic information on coded bits [TB00; Hag04]. This process is implemented through Soft Input Soft Output (SISO) modules and through Belief Propagation (BP), a.k.a. the sum-product message passing algorithm, which leads to *turbo-like receivers* that will be the focus of remainder of this chapter.

In detail, a SISO detector and a SISO decoder approximately resolve the optimal estimation problem, through iterative processing where they exchange *extrinsic* information on coded bits. More specifically, as c_j and d_j are binary variables, the evolution of their estimated statistics are characterized by the bit LLRs, as illustrated in Figure 2.1. A SISO module treats extrinsic LLRs provided by the other SISO module as its *prior* information, and it computes extrinsic LLRs on coded bits, $L_e(\cdot)$, by ensuring that they are independent from the prior LLRs, denoted by $L_p(\cdot)$. The interleaver ties extrinsic and prior LLRs through $L_e(d_j) = L_p(c_{\Pi(j)})$ and $L_e(c_j) = L_p(d_{\Pi^{-1}(j)})$.

The reasoning behind the derivation of such architectures will be discussed in the next section, Example 1, with the introduction of graphical models and the belief propagation algorithm. Some examples of SISO decoders are provided in [RL09] and regarding SISO modules for detection, the optimal SISO receiver for ISI channels, and some examples of sub-optimal filter-based SISO equalizers are exposed in the next chapter. The remainder of this chapter will discuss the derivation of SISO detectors in general, through the use of approximate statistical inference or estimation techniques.

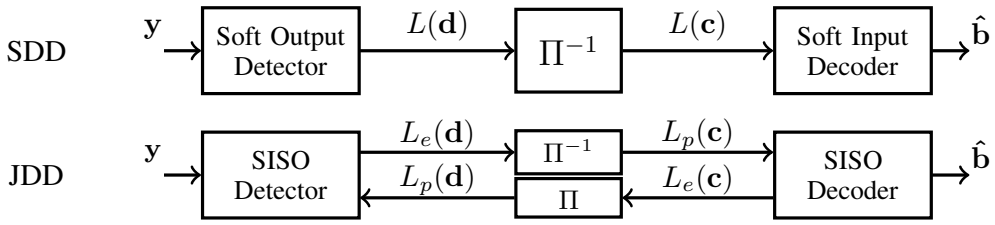


Figure 2.1: Comparison of separate and iterative detection and decoding architectures.

2.3 Conventional SISO Detection for BICM with Iterative Decoding

In the previous section the concept of BP, also known as the sum-product algorithm was mentioned as the justification for the use of iterative detection and decoding architectures for BICM. This section discusses graphical stochastic models, and in particular factor graphs, in order to introduce BP more formally, in the context of the design of a SISO detector.

2.3.1 Probabilistic Graphical Models

Probabilistic graphical models provide a means to describe complex statistical systems that incorporate observed and latent variables, and constraints that govern relationships between them. Through specific factorization of joint probability distributions that describe the stochastic model, such graphs can be constructed and then used for designing Bayesian inference algorithms. Our motivation being the computation of near-optimal SISO detectors, let us recall the optimum detector based on MAP inference

$$\hat{x}_k = \arg\max_{x_k} p(x_k|\mathbf{y}). \tag{2.12}$$

Moreover, the desired marginal PDF needs to be computed through

$$p(x_k|\mathbf{y}) = \sum_{\mathbf{x}^{\setminus x_k}} p(\mathbf{x}|\mathbf{y}), \tag{2.13}$$

where $p(\mathbf{x}|\mathbf{y})$ is the joint posterior PDF, which involves K discrete-valued latent random variables $\{x_k\}_{k=1}^{K}$, and N complex-valued observed variables $\{y_n\}_{n=1}^{N}$. Both of the equations above, and the joint posterior itself, are often computationally intensive, and approximate estimates are computed through message passing algorithms. These algorithms use the underlying structure within the joint PDF in order to estimate approximate posteriors with lower complexity. In order to discuss this structure concept, we will first briefly review different graphical models that are used on such problems [YFW03].

Bayesian networks and *Markov random fields* are two categories of graphical models which have played important roles in the development and analysis of message passing techniques

[CR14]. However the graphical models we will discuss are *factor graphs*, a category of bipartite undirected graphs consisting of squares and circles which respectively represent factor nodes and variable nodes. Moreover, observation variables are illustrated as filled dark circles and latent variables by empty circles, and when a factor depends on a variable, they are connected with an edge. In this case, the joint posterior is factorized into N_f factors as

$$p(\mathbf{x}) \propto \prod_{i=1}^{N_f} f_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i}), \tag{2.14}$$

where f_{F_i} is a non-negative function which depends on $\mathbf{x}_{\mathrm{F}_i}$, a sub-vector of $\mathbf{x}$, with F_i being the denomination of the factor node, and the operator $\mathcal{V}(\cdot)$ denotes the set of neighbors of a node. Factor graphs possess a more flexible capacity for modeling stochastic systems, and Bayesian networks or Markov random fields have equivalent factor graph models [CR14].

As the use of factor graphs are more widespread in the communication theory community [KFL01b; Loe+07; WP99; Wal06; Sen+11; HEA11], in the remainder of this thesis inference problem are modeled with this tool, in order to discuss approximate iterative algorithms. Moreover, in Section 2.4.1, an important property of factor graphs is discussed with regards to the BP and EP algorithms, whose fixed points correspond to the local minima of the Bethe free energy, which is directly depends on the selected factor graph.

2.3.2 Loopy Belief Propagation

Belief propagation, also known as the sum-product algorithm aims to provide an approximate posterior PDF $q(\mathbf{x})$ of the true posterior $p(\mathbf{x})$. This algorithm's development is directly related to the use of graphical models, and its particular cases have been re-discovered independently in different scientific communities. An example of its earliest uses is in coding theory with Tanner graphs [Tan81], derived for Gallager's LDPC codes [Gal62], or with Trellis diagrams [Bah+74] for representing convolutional codes. Pearl's approach to Bayesian networks is an alternate formulation where BP has been derived for solving various artificial intelligence problems [Pea88], and another formalism is given through Markov random fields for image processing [YFW03]. It has been also formulated as the "generalized distributive law" [AM00], which rather considers modeling of stochastic processes as junction trees. As briefly covered in the previous sub-section, different graphical models can be converted among each other, and we will primarily focus on factor graph formalism of BP [KFL01a], as it enables to explicitly expose how graphs differ from one-another, by exposing different factor PDFs constituting the true joint posterior. Further motivation for the use of this model will be presented in the next section which discusses approximate Bayesian inference from a slightly different point of view.

This algorithm computes *extrinsic* messages from a nodes to its adjacent nodes, which are perceived as approximate *prior* distributions at the receiving end. Then the destination node computes a "local" approximation of its associated marginal density, called belief, which is later used to carry out approximate inference. These messages are improper (i.e. non-

normalized) distributions but beliefs are normalized as regular probability distributions.

The extrinsic message $m_{\mathrm{F}_i \to x_k}(x_k)$ sent from a factor node F_i towards a variables node x_k consists in the product of the true factor f_{F_i} and all incoming messages from its neighboring variable nodes, except for x_k itself, marginalized down to the variable x_k.

More explicitly, BP prior and extrinsic messages, from a factor node's point of view, exchanged between v_n and F_i are respectively

$$m_{x_k \to \mathrm{F}_i}(x_k) \triangleq \prod_{\mathrm{F}_j \in \mathcal{V}(x_k) \setminus \{\mathrm{F}_i\}} m_{\mathrm{F}_j \to x_k}(x_k), \tag{2.15}$$

$$m_{\mathrm{F}_i \to x_k}(x_k) \triangleq \sum_{\mathbf{x}_{\mathrm{F}_i}^{\setminus x_k}} f_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i}) \prod_{x_l \in \mathcal{V}(\mathrm{F}_i) \setminus \{x_k\}} m_{x_l \to \mathrm{F}_i}(x_l). \tag{2.16}$$

The "sum-product algorithm" denomination comes from the expression of the message from a factor node to a variable node, and this algorithm is executed usually by initializing all messages with $m_{x_k \to \mathrm{F}_i}(x_k) = m_{\mathrm{F}_i \to x_k}(x_k) = 1$ for all k and i, and then iteratively updating messages over the factor graph. The *belief* of a variable node x_k is the *approximate posterior* $q(x_k)$ of the true posterior $p(x_k)$, and it is given by

$$q(x_k) = \frac{1}{Z_k} \prod_{\mathrm{F}_j \in \mathcal{V}(x_k)} m_{\mathrm{F}_j \to x_k}(x_k), \tag{2.17}$$

where Z_k is the normalization constant, called *evidence*, given by

$$Z_k = \sum_{x_k} \prod_{\mathrm{F}_j \in \mathcal{V}(x_k)} m_{\mathrm{F}_j \to x_k}(x_k). \tag{2.18}$$

BP message passing achieves exact inference on acyclic graphs, when the message update schedule is well-adapted to the graph structure. In such cases, the sum-product algorithm is merely an optimization algorithm for the marginalization of $p(\mathbf{x})$, by organizing the order of operations. In particular, for tree-structured graphs, exact inference with BP message-passing is completed by a single forward and then backward pass from one end to the other of the graph. This has been known as the forward-backward algorithm for hidden Markov models.

In the more general case, where the factor graph has cycles, BP message passing is carried out iteratively until convergence, or until a stop criterion. In this case, this algorithm is rather called "loopy BP", and its convergence is not guaranteed, and its performance is sub-optimal compared to exact inference. Indeed, loopy BP computes an approximate posteriors which is potentially a solution to a distributed optimization problem over each factor of the true posterior, but the local optima of factors might not coincide with the global optimum.

Example 1 (Iterative Detection and Decoding)
Here we provide the proof that the use of BP on the optimal detection (joint ML in Equation (2.3)) leads to an iterative structure with SISO modules. Thanks to BICM, the joint likelihood is equivalent to the factorized as in Equation (2.5), and an associated factor graph, illustrated

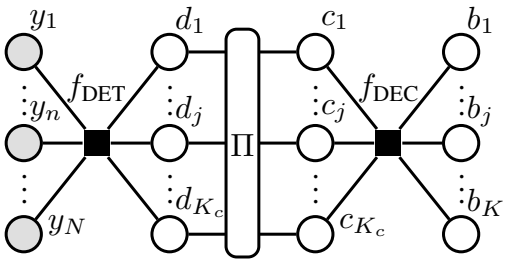


Figure 2.2: Factor graph of the BICM system, for iterative detection and decoding.

in Figure 2.2, is constructed with

$$p(\mathbf{y},\mathbf{x},\mathbf{d},\mathbf{c}|\mathbf{H},\boldsymbol{\Sigma}_{\mathbf{w}},\mathbf{b}) = \underbrace{p(\mathbf{y}|\mathbf{x},\mathbf{H},\boldsymbol{\Sigma}_{\mathbf{w}})p(\mathbf{x}|\mathbf{d})}_{\triangleq f_{DET}(\mathbf{y},\mathbf{d})}\underbrace{p(\mathbf{d}|\mathbf{c})}_{\Pi}\underbrace{p(\mathbf{c}|\mathbf{b})}_{\triangleq f_{DEC}(\mathbf{c},\mathbf{b})}, \tag{2.19}$$

where we introduce the factors $f_{DET}(\mathbf{y},\mathbf{d})$ and $f_{DEC}(\mathbf{c},\mathbf{b})$, respectively for detection and decoding, by omitting some parameters, to alleviate their notations and focus on estimation of $\mathbf{c}$ *and* $\mathbf{d}$.

When BP is applied to this factor graph, the update messages of the coded and interleaved bit $d_{k,q}$ is given by

$$m_{DET\to d_{k,q}}(d_{k,q}) \propto \sum_{\mathbf{d}^{\setminus d_{k,q}}} f_{DET}(\mathbf{y},\mathbf{d}) \prod_{k=1}^{K}\prod_{q'\neq q} m_{d_{k,q'}\to DET}(d_{k,q'}), \tag{2.20}$$

and similarly the message from DEC to $c_{k,q}$ is as follows

$$m_{DEC\to c_{k,q}}(c_{k,q}) \propto \sum_{\mathbf{c}^{\setminus d_{k,q}}} f_{DEC}(\mathbf{c},\mathbf{b}) \prod_{k=1}^{K}\prod_{q'\neq q} m_{c_{k,q'}\to DEC}(c_{k,q'}), \tag{2.21}$$

under the interleaving constraints

$$m_{d_j\to DET}(d_j) = m_{DEC\to c_{\Pi(j)}}(c_{\Pi(j)}), \tag{2.22}$$

$$m_{c_j\to DEC}(c_j) = m_{DET\to d_{\Pi^{-1}(j)}}(d_{\Pi^{-1}(j)}). \tag{2.23}$$

Hence the optimal detection problem can be approximated through this iterative process, by initializing $m_{d_j\to DET}(d_j) \propto 1$ and $m_{c_j\to DEC}(c_j) \propto 1$, i.e. with uniform IID distributions, and following a serial scheduling where DET and DEC are successively updated.

From the point of view of the factor node DET, $m_{d_j\to DET}(d_j)$ and $m_{DET\to d_{k,q}}(d_{k,q})$ are respectively considered as a priori *and* extrinsic *PDFs, and this constitutes the principle of operation of a Soft Input Soft Output (SISO) module. Moreover, as c_j and d_j are binary variables, the evolution of their extrinsic messages is characterized by the bit LLRs, as illustrated in Figure 2.1 where, $L_e(d_j) = L_p(c_{\Pi(j)})$ and $L_e(c_j) = L_p(d_{\Pi^{-1}(j)})$ are respectively LLRs of the extrinsic PDF $m_{d_j\to DET}(d_j)$ and a priori PDF $m_{c_j\to DEC}(c_j)$.*

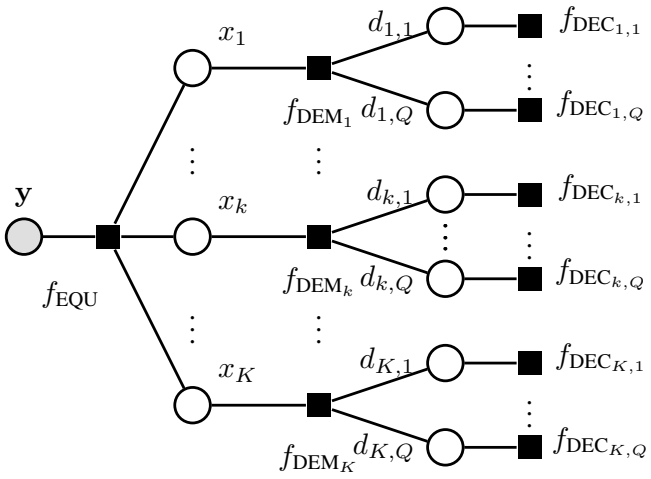


Figure 2.3: Factor graph of the BICM decoder with a SISO detector at symbol level.

Despite its limitations, loopy BP has been widely used for near-optimal resolution of a variety of complex probabilistic problem, from statistical physics to signal processing, information theory or artificial intelligence fields. In some particular graph structures, loopy BP can offer more interesting performance complexity trade-offs, than the general situation above. For instance, in the context of communication theory, it is used for practical SISO channel decoding of different family of codes, such as LDPC codes or turbo-codes [RL09]. In the context of detector design for ISI channels, the underlying graph can be represented as an acyclic factor graph, or more traditionally a Trellis, which leads to the derivation of optimal MAP detection algorithms such as BCJR.

2.3.3 Generic MAP SISO Detector

Let us consider a SISO detector based on BP by addressing the factor graph of the detection factor $f_{\mathrm{DET}}(\mathbf{y}, \mathbf{x}, \mathbf{d})$ of the generalized linear system in Equation (2.1). From the system model, we have the observation model

$$p(\mathbf{y}|\mathbf{x}, \mathbf{H}, \mathbf{\Sigma_w}) = \mathcal{CN}(\mathbf{y}; \mathbf{Hx}, \mathbf{\Sigma_w}) = \frac{1}{\pi \det(\mathbf{\Sigma_w})} \exp\left(-\|\mathbf{y} - \mathbf{Hx}\|^2_{\mathbf{\Sigma_w}}\right), \tag{2.24}$$

the mapping constraints

$$p(x_k|\mathbf{d}_k) = \delta(x_k - \varphi(\mathbf{d}_k)) = \prod_{q=1}^{Q} \delta(d_{k,q} - \varphi_q^{-1}(x_k)), \tag{2.25}$$

and the prior information from the decoder

$$p(d_{k,q}) = \frac{1}{1 + e^{-L_p(d_{k,q})}} \delta(d_{k,q}) + \frac{e^{-L_p(d_{k,q})}}{1 + e^{-L_p(d_{k,q})}} \delta(d_{k,q} - 1) = \frac{e^{-d_{k,q} L_p(d_{k,q})}}{1 + e^{-L_p(d_{k,q})}}. \tag{2.26}$$

Then the joint posterior at the SISO detector is factorized as

$$\begin{aligned} p(\mathbf{x},\mathbf{d}) &= p(\mathbf{y}|\mathbf{x},\mathbf{H},\boldsymbol{\Sigma}_\mathbf{w}) \prod_{k=1}^{K} p(x_k|\mathbf{d}_k) \prod_{q=1}^{Q} p(d_{k,q}), \\ &\propto \underbrace{\exp\left(-\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2_{\boldsymbol{\Sigma}_\mathbf{w}}\right)}_{=f_{\mathrm{EQU}}(\mathbf{x})} \prod_{k=1}^{K} \underbrace{\prod_{q=1}^{Q} \delta(d_{k,q}-\varphi_q^{-1}(x_k))}_{=f_{\mathrm{DEM}_k}(x_k,\mathbf{d}_k)} \underbrace{\exp(-d_{k,q}L_p(d_{k,q}))}_{=f_{\mathrm{DEC}_{k,q}}(d_{k,q})}, \end{aligned} \tag{2.27}$$

with $x_k \in \mathcal{X}$, $\forall k$ and $d_{k,q} \in \mathbb{F}_2$, $\forall k,q$, illustrated in Figure 2.3. In this case, messages from factor nodes DEM_k and EQU are

$$m_{\mathrm{DEC}\to d_{k,q}}(d_{k,q}) = f_{\mathrm{DEC}_{k,q}}(d_{k,q}), \tag{2.28}$$

$$\begin{aligned} m_{\mathrm{DEM}_k\to x_k}(x_k) &= \sum_{\mathbf{d}_k} f_{\mathrm{DEM}_k}(x_k,\mathbf{d}_k) \prod_{q=1}^{Q} m_{\mathrm{DEC}\to d_{k,q}}(d_{k,q}), \\ &= \sum_{\alpha\in\mathcal{X}} \prod_{q=1}^{Q} \exp\left(-\varphi_q^{-1}(\alpha)L_p(d_{k,q})\right) \delta(x_k-\alpha), \end{aligned} \tag{2.29}$$

$$\begin{aligned} m_{\mathrm{EQU}\to x_k}(x_k) &= \sum_{\mathbf{x}^{\setminus x_k}} f_{\mathrm{EQU}}(\mathbf{x}) \prod_{k'\neq k} m_{\mathrm{DEM}_{k'}\to x_{k'}}(x_{k'}), \\ &= \sum_{\alpha\in\mathcal{X}} \sum_{\mathbf{x}^{\setminus x_k}} \exp\left(-\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2_{\boldsymbol{\Sigma}_\mathbf{w}} - \sum_{k'\neq k}\sum_{q=1}^{Q} \varphi_q^{-1}(x_{k'})L_p(d_{k',q})\right) \delta(x_k-\alpha), \end{aligned} \tag{2.30}$$

$$\begin{aligned} m_{\mathrm{DEM}_k\to d_{k,q}}(d_{k,q}) &= \sum_{x_k\in\mathcal{X}} \sum_{\mathbf{d}_k^{\setminus d_{k,q}}} f_{\mathrm{DEM}_k}(x_k,\mathbf{d}_k) m_{\mathrm{EQU}\to x_k}(x_k) \prod_{q'\neq q} m_{\mathrm{DEC}\to d_{k,q'}}(d_{k,q'}), \\ &= \sum_{\beta=0}^{1} \sum_{x_k\in\mathcal{X}_q^\beta} \sum_{\mathbf{x}^{\setminus x_k}} \exp\left(-\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2_{\boldsymbol{\Sigma}_\mathbf{w}} - \sum_{(k',q')\neq(k,q)} \varphi_{q'}^{-1}(x_{k'})L_p(d_{k',q'})\right) \delta(d_{k,q}-\beta), \end{aligned} \tag{2.31}$$

with $\mathcal{X}_q^\beta = \{\alpha\in\mathcal{X}, \varphi_q^{-1}(\alpha)=\beta\}$, $\beta\in\mathbb{F}_2$. The latter message being proportional to a Bernoulli distribution, it is used to compute extrinsic LLRs $L_e(d_{k,q})$ of the SISO module.

Note that the marginalization over $\mathbf{x}^{\setminus x_k}$ constitutes the computational bottleneck of loopy BP for this generalized linear model, as it implies computation of KM^{K-1} metrics (i.e. arguments of the exponential). For some well-structured channel matrices $\mathbf{H}$, these marginals can be computed more elegantly, as we will discuss in the next chapter for ISI channels with the derivation of the SISO BCJR algorithm [Bah+74; DJB+95; CB05].

2.3.4 Gaussian Approximated Belief Propagation

The generalization of the sum-product algorithm, through factor graphs has allowed extending the message passing concept beyond discrete-valued distributions, with a formalism on the summary operation. Indeed, message-passing rules of BP in Equations (2.15-2.16) are extended to continuous PDFs with

$$m_{x_k \to \mathrm{F}_i}(x_k) \triangleq \prod_{\mathrm{F}_j \in \mathcal{V}(x_k) \setminus \{\mathrm{F}_i\}} m_{\mathrm{F}_j \to x_k}(x_k), \tag{2.32}$$

$$m_{\mathrm{F}_i \to x_k}(x_k) \triangleq \int_{\mathbf{x}_{\mathrm{F}_i}^{\setminus x_k}} f_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i}) \prod_{x_l \in \mathcal{V}(\mathrm{F}_i) \setminus \{x_k\}} m_{x_l \to \mathrm{F}_i}(x_l) \, \mathrm{d}\mathbf{x}_{\mathrm{F}_i}^{\setminus x_k}. \tag{2.33}$$

This generalization can be of significant interest, in situation where the product of a family of PDFs and their marginals remain tractable through computations. This is often the case with random variables belonging to the exponential family, and in particular, for Gaussian-distributed random variables for whom computed messages always describe a Gaussian PDF.

In this particular context, when considering a Gaussian distribution, MAP inference is equivalent to the Minimum Mean Square Error (MMSE) estimate of the posterior, i.e. if $\mathbf{x}$ is jointly Gaussian, then

$$\arg\max_{\mathbf{x}} p(\mathbf{x}) = \mathbb{E}\left[\mathbf{x}\right], \tag{2.34}$$

which leads to attractive computational properties when carrying out approximate inference with Gaussian-distributed variables. This has led to a category of message passing algorithms, called Gaussian-approximated Belief Propagation (GaBP), where prior messages are projected to Gaussian distributions, when computing extrinsic message on a factor which represents a Gaussian PDF.

More explicitly, let us consider GaBP on a factor graph where factors are either Gaussian or categorical PDFs, and denote the BP extrinsic messages $m^*_{\mathrm{F}_i \to x_k}(x_k)$ (from Eq. (2.33)) where x_k is assumed to be Gaussian distributed. The original BP message is then transformed into Gaussian distribution $m_{\mathrm{F}_i \to x_k}(x_k)$, with reverse-information projection, also known as moment projection, with

$$m_{\mathrm{F}_i \to x_k}(x_k) = \mathrm{proj}_{\mathcal{CN}}\left[m^*_{\mathrm{F}_i \to x_k}(x_k)\right] \triangleq \arg\min_{q(x_k) \in \mathcal{CN}} D_{\mathrm{KL}}(m^*_{\mathrm{F}_i \to x_k}(x_k) \| q(x_k)). \tag{2.35}$$

This criteria is equivalent to carrying out moment-matching, and by denoting $m_{\mathrm{F}_i \to x_k}(x_k) = \mathcal{CN}(x_k; \mu_{x,k}, \sigma^2_{x,k})$, we have

$$\mu_{x,k} = \mathbb{E}_{m^*_{\mathrm{F}_i \to x_k}}[x_k], \quad \sigma^2_{x,k} = \mathrm{Var}_{m^*_{\mathrm{F}_i \to x_k}}[x_k]. \tag{2.36}$$

If the factor node F_i is already a Gaussian PDF, this operation is identity, otherwise if F_i is a categorical PDF, then its first and second order statistics have to be computed.

The use of GaBP on linear detection structures yields receiver structures that share similarities with well known Wiener/Kalman filters from the estimation theory [Kay93; WP99;

TSK02], or with Kalman smoothing when considering a dynamic linear system [Loe+07; SSS02; OT02]. In communications theory, this approach has been used to derive sub-optimal SISO detectors for linear channel channels, which has led to very popular turbo-equalization structures based on MMSE filters with prior information.

2.3.5 Generic Linear MMSE SISO Detector

This section discusses a less complex alternative to the optimal MAP SISO detector described in subsection 2.3.3, for the problem of SISO detector design for the the BICM system with generic linear observation model of Equation (2.1). Here, through the use of GaBP, an MMSE-like filter-bank structure is obtained with the use of Gaussian-distributed assumption on the symbol variable x_k.

Let us recall the factorization of the model

$$p(\mathbf{x},\mathbf{d}) \propto \underbrace{\exp\left(-\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2_{\mathbf{\Sigma_w}}\right)}_{=f_{\text{EQU}}(\mathbf{x})} \prod_{k=1}^{K} \underbrace{\prod_{q=1}^{Q} \delta(d_{k,q}-\varphi_q^{-1}(x_k))}_{=f_{\text{DEM}_k}(x_k,\mathbf{d}_k)} \underbrace{\exp(-d_{k,q}L_p(d_{k,q}))}_{=f_{\text{DEC}_{k,q}}(d_{k,q})}, \tag{2.37}$$

where $d_{k,q} \in \mathbb{F}_2$, $\forall k,q$ but we assume unconstrained $x_k \in \mathbb{C}$. This factor graph involving also the data symbols x_k explicitly is illustrated in Figure 2.3. This assumption relaxes the constraints of the factor node EQU but it does not affect DEM, as f_{DEM_k} uses hard constraints to force $x_k \in \mathcal{X}$. The extrinsic BP message from DEM

$$m^*_{\text{DEM}_k\to x_k}(x_k) = \sum_{\alpha\in\mathcal{X}} \prod_{q=1}^{Q} \exp\left(-\varphi_q^{-1}(\alpha)L_p(d_{k,q})\right)\delta(x_k-\alpha), \tag{2.38}$$

which is an improper PDF, to be normalized for computing its sufficient statistics. Let us denote $\mathcal{P}_k(\alpha)$ the normalized prior Probability Mass Function (PMF) associated to $m^*_{\text{DEM}_k\to x_k}(x_k)$, for $\alpha \in \mathcal{X}$

$$\mathcal{P}_k(\alpha) \triangleq \frac{1}{\sum_{\alpha'\in\mathcal{X}}\prod_{q=1}^{Q}\exp\left(-\varphi_q^{-1}(\alpha')L_p(d_{k,q})\right)} \prod_{q=1}^{Q}\exp\left(-\varphi_q^{-1}(\alpha)L_p(d_{k,q})\right), \tag{2.39}$$

and its projection onto the Gaussian PDF $\mathcal{CN}(x_k; x_k^p, v_{x,k}^p)$, yields

$$x_k^p \triangleq \mathbb{E}_{\mathcal{P}_k}[x_k] = \sum_{\alpha\in\mathcal{X}} \alpha\mathcal{P}_k(\alpha), \quad v_{x,k}^p \triangleq \text{Var}_{\mathcal{P}_k}[x_k] = \sum_{\alpha\in\mathcal{X}} |\alpha - x_k^p|^2\mathcal{P}_k(\alpha). \tag{2.40}$$

Consequently, the projected message is as follows

$$m_{\text{DEM}_k\to x_k}(x_k) \propto \exp\left(-\frac{|x_k - x_k^p|^2}{v_{x,k}^p}\right), \tag{2.41}$$

and, the message from the node EQU becomes

$$m_{\text{EQU}\to x_k}(x_k) = \int_{\mathbf{x}\backslash x_k} f_{\text{EQU}}(\mathbf{x}) \prod_{k'\neq k} m_{\text{DEM}_{k'}\to x_{k'}}(x_{k'})\, \mathrm{d}\mathbf{x}\backslash x_k,$$
$$\propto \frac{1}{m_{\text{DEM}_k\to x_k}(x_k)} \underbrace{\int_{\mathbf{x}^{\backslash x_k}} \exp\left(-\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2_{\mathbf{\Sigma_w}} - \sum_{k'} \frac{|x_{k'}-x^p_{k'}|^2}{v^p_{x,k'}}\right) \mathrm{d}\mathbf{x}^{\backslash x_k}}_{\triangleq q_k(x_k)}. \tag{2.42}$$

Indeed, we have

$$q_k(x_k) \propto \int_{\mathbf{x}^{\backslash x_k}} \exp\left(-\mathbf{x}^H\left[\mathbf{H}^H\mathbf{\Sigma}_\mathbf{w}^{-1}\mathbf{H} + \mathbf{V}_\mathbf{x}^{\mathbf{p}-1}\right]\mathbf{x} + 2\Re\left(\mathbf{x}^H\left[\mathbf{H}^H\mathbf{\Sigma}_\mathbf{w}^{-1}\mathbf{y} + \mathbf{V}_\mathbf{x}^{\mathbf{p}-1}\mathbf{x}^\mathbf{p}\right]\right)\right) \mathrm{d}\mathbf{x}^{\backslash x_k},$$
$$\propto \int_{\mathbf{x}^{\backslash x_k}} \mathcal{CN}(\mathbf{x};\boldsymbol{\mu}_\mathbf{x},\mathbf{\Gamma}_\mathbf{x})\, \mathrm{d}\mathbf{x}^{\backslash x_k},\ \propto \mathcal{CN}(x_k;\mu_{x,k},\gamma_{x,k}), \tag{2.43}$$

where $\mathbf{V}_\mathbf{x}^\mathbf{p} \triangleq \mathbf{Diag}([v^p_{x,1},\ldots,v^p_{x,K}])$ and

$$\mathbf{\Gamma}_\mathbf{x} \triangleq (\mathbf{H}^H\mathbf{\Sigma}_\mathbf{w}^{-1}\mathbf{H} + \mathbf{V}_\mathbf{x}^{\mathbf{p}-1})^{-1} = \mathbf{V}_\mathbf{x}^\mathbf{p} - \mathbf{V}_\mathbf{x}^\mathbf{p}\mathbf{H}^H(\mathbf{\Sigma}_\mathbf{w} + \mathbf{H}\mathbf{V}_\mathbf{x}^\mathbf{p}\mathbf{H}^H)^{-1}\mathbf{H}\mathbf{V}_\mathbf{x}^\mathbf{p}, \tag{2.44}$$
$$\boldsymbol{\mu}_\mathbf{x} \triangleq \mathbf{\Gamma}_\mathbf{x}(\mathbf{H}^H\mathbf{\Sigma}_\mathbf{w}^{-1}\mathbf{y} + \mathbf{V}_\mathbf{x}^{\mathbf{p}-1}\mathbf{x}^\mathbf{p}) = (\mathbf{I}_K - \mathbf{V}_\mathbf{x}^\mathbf{p}\mathbf{H}^H(\mathbf{\Sigma}_\mathbf{w} + \mathbf{H}\mathbf{V}_\mathbf{x}^\mathbf{p}\mathbf{H}^H)^{-1}\mathbf{H})(\mathbf{V}_\mathbf{x}^\mathbf{p}\mathbf{H}^H\mathbf{\Sigma}_\mathbf{w}^{-1}\mathbf{y} + \mathbf{x}^\mathbf{p})$$
$$= \mathbf{x}^\mathbf{p} + \mathbf{V}_\mathbf{x}^\mathbf{p}\mathbf{H}^H(\mathbf{\Sigma}_\mathbf{w} + \mathbf{H}\mathbf{V}_\mathbf{x}^\mathbf{p}\mathbf{H}^H)^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}^\mathbf{p}). \tag{2.45}$$

Consequently, the marginal PDF on x_k can be obtained by

$$\gamma_{x,k} \triangleq \mathbf{e}_k^H\mathbf{\Gamma}_\mathbf{x}\mathbf{e}_k = v^p_{x,k}(1 - v^p_{x,k}\mathbf{e}_k^H\mathbf{H}^H(\mathbf{\Sigma}_\mathbf{w} + \mathbf{H}\mathbf{V}_\mathbf{x}^\mathbf{p}\mathbf{H}^H)^{-1}\mathbf{H}\mathbf{e}_k), \tag{2.46}$$
$$\mu_{x,k} \triangleq \mathbf{e}_k^H\boldsymbol{\mu}_\mathbf{x} = x^p_k + v^p_{x,k}\mathbf{e}_k^H\mathbf{H}^H(\mathbf{\Sigma}_\mathbf{w} + \mathbf{H}\mathbf{V}_\mathbf{x}^\mathbf{p}\mathbf{H}^H)^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}^\mathbf{p}). \tag{2.47}$$

Finally, the message from EQU becomes

$$m_{\text{EQU}\to x_k}(x_k) \propto \frac{q_k(x_k)}{m_{\text{DEM}_k\to x_k}(x_k)} = \frac{\mathcal{CN}(x_k;\mu_{x,k},\gamma_{x,k})}{\mathcal{CN}(x_k;x^p_k,v^p_{x,k})} \propto \mathcal{CN}(x_k;x^e_k,v^e_{x,k}), \tag{2.48}$$

with

$$x^e_k \triangleq \frac{v^p_{x,k}\mu_{x,k} - \gamma_{x,k}x^p_k}{v^p_{x,k} - \gamma_{x,k}},\quad v^e_{x,k} \triangleq \frac{v^p_{x,k}\gamma_{x,k}}{v^p_{x,k} - \gamma_{x,k}}. \tag{2.49}$$

By using Equations (2.46-2.47) with Equation (2.49), and $\xi_k \triangleq \mathbf{e}_k^H\mathbf{H}^H(\mathbf{\Sigma}_\mathbf{w} + \mathbf{H}\mathbf{V}_\mathbf{x}^\mathbf{p}\mathbf{H}^H)^{-1}\mathbf{H}\mathbf{e}_k$, we have

$$v^e_{x,k} = 1/s_k - v^p_{x,k}, \tag{2.50}$$
$$x^e_k = x^p_k + \xi_k^{-1}\mathbf{e}_k^H\mathbf{H}^H(\mathbf{\Sigma}_\mathbf{w} + \mathbf{H}\mathbf{V}_\mathbf{x}^\mathbf{p}\mathbf{H}^H)^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}^\mathbf{p}). \tag{2.51}$$

Note that GaBP parameters can be computed with a linear filter-bank, that involves a non-negative symmetric matrix inversion with quadratic complexity $\mathcal{O}(K^2)$, followed by a stage of parallelized arithmetic operations. Hence it is much more affordable than the exponential complexity of loopy BP.

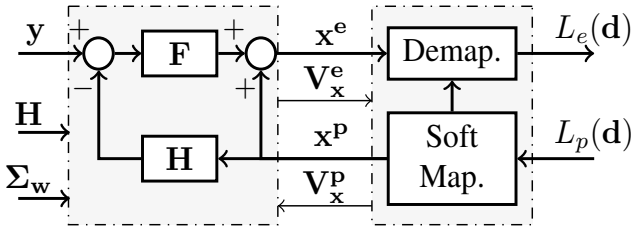


Figure 2.4: Turbo receiver based on the generic linear MMSE SISO detector.

Once the message from EQU is computed, the extrinsic output of this SISO module can be computed through factor node DEM, with

$$\begin{aligned} m_{\mathrm{DEM}_k \to d_{k,q}}(d_{k,q}) &= \int_{x_k} \sum_{\mathbf{d}_k \setminus d_{k,q}} f_{\mathrm{DEM}_k}(x_k, \mathbf{d}_k) m_{\mathrm{EQU} \to x_k}(x_k) \prod_{q' \neq q} m_{\mathrm{DEC} \to d_{k,q'}}(d_{k,q'}) \, \mathrm{d}x_k, \\ &= \sum_{\beta=0}^{1} \sum_{\alpha \in \mathcal{X}_q^\beta} \exp\left(-\frac{|x_k^e - \alpha|^2}{v_{x,k}^e} - \sum_{q' \neq q} \varphi_{q'}^{-1}(\alpha) L_p(d_{k,q'}) \right) \delta(d_{k,q} - \beta). \end{aligned} \tag{2.52}$$

By denoting $\mathcal{D}_k(\alpha)$ the normalized posterior PMF, such that

$$\mathcal{D}_k(\alpha) \propto \exp\left(-\frac{|x_k^e - \alpha|^2}{v_{x,k}^e} - \sum_{q=0}^{Q} \varphi_q^{-1}(\alpha) L_p(d_{k,q}) \right), \tag{2.53}$$

for $\alpha \in \mathcal{X}$, the extrinsic LLRs can be computed as

$$L_e(d_{k,q}) = \ln \frac{\sum_{\alpha \in \mathcal{X}_q^0} \mathcal{D}_k(\alpha)}{\sum_{\alpha \in \mathcal{X}_q^1} \mathcal{D}_k(\alpha)} - L_p(d_{k,q}). \tag{2.54}$$

The equations above correspond to the symbol-wise MAP SISO demapper [BSY98], and the turbo detection scheme with the linear MMSE SISO detector is illustrated in Figure 2.4.

2.3.6 Asymptotic Analysis for Iterative BICM Receiver: EXIT

For facilitating the design of SISO modules, and for the link abstraction with receivers based on iterative detection and decoding, prediction and analysis tools are needed. Hence various approaches were proposed to analyze the decoding dynamics of BP based BICM systems, and to characterize the behavior of SISO modules. For instance, in channel coding, density evolution was proposed for tracking the dynamics of PDFs of exchanged bit LLRs within advanced channel decoders for turbo codes or LDPC codes [RSU01] and by extension, of turbo receivers [BC02].

Density evolution is affordable on sparse and asymptotically large graphs, that involve random variables characterized by a small amount of parameters. However, the signal processing dynamics of such algorithms are often computationally prohibitive on more general

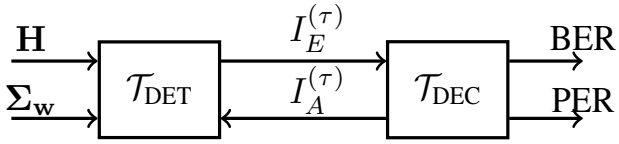


Figure 2.5: EXIT function equivalent model of a turbo receiver.

factor graphs and simpler alternatives are used. Extrinsic Information Transfer (EXIT) analysis simplified density evolution to an asymptotic single-parameter tracking problem, and it has lead to great achievements on the design and analysis of BICM systems with iterative detection [Bri99; TB00].

EXIT functions are MI transfer functions of a SISO module, that track the evolution of MI of extrinsic LLRs and the transmitted bits, as a function of the MI of prior LLRs and the transmitted bits [TB00].To illustrate how EXIT charts work, let us consider a turbo receiver with a SISO detector and a SISO decoder, that is iterated for $\tau = 0, \ldots, \mathcal{T}$ turbo-iterations. The iteration index is incremented after each round of decoding, until $\mathcal{T}$ is reached, and the extrinsic LLRs of the detector (also the prior LLRs of the decoder) are $L_e^{(\tau)}(d_j)$ (i.e. $L_p^{(\tau)}(c_{\Pi^{-1}(j)})$) and the extrinsic LLRs of the decoder (also the prior LLRs of the detector) are $L_e^{(\tau)}(c_j)$ (i.e. $L_p^{(\tau)}(d_{\Pi(j)})$), for all $j = 1, \ldots, K_c$.

The SISO detector is then described by the transfer function $\mathcal{T}_{\text{DET}}$, which depends on the channel parameters with

$$I_E^{(\tau)} = \mathcal{T}_{\text{DET}}(I_A^{(\tau)}; \mathbf{H}, \mathbf{\Sigma_w}), \tag{2.55}$$

where the prior information I_A and the extrinsic information I_E are the average MI between coded bits and respectively the a priori and extrinsic LLRs of the module, given as

$$I_A^{(\tau)} \triangleq \frac{1}{KQ} \sum_{k,q} I(d_{k,q}; L_p^{(\tau)}(d_{k,q})), \tag{2.56}$$

$$I_E^{(\tau)} \triangleq \frac{1}{KQ} \sum_{k,q} I(d_{k,q}; L_e^{(\tau)}(d_{k,q})). \tag{2.57}$$

The SISO decoder's EXIT function $\mathcal{T}_C$ is similarly given by

$$I_A^{(\tau+1)} = \mathcal{T}_{\text{DEC}}(I_E^{(\tau)}), \tag{2.58}$$

and the correspondence between the iterative turbo receiver and the EXIT model is illustrated on Figure 2.5.

EXIT functions are synthesized through a Monte-Carlo method, by isolating the concerned SISO module and feeding it with prior LLRs, matching the desired I_A, and then computing an histogram on extrinsic LLRs, to estimate I_E. Hence, to implement this approach, a prior LLRs sequence generator is needed, and the following assumption and property provide a solution to this end [TB00].

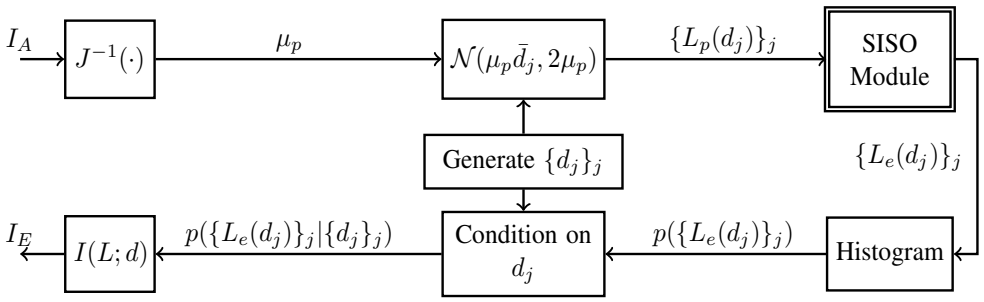


Figure 2.6: EXIT function synthesis method.

Assumption 1
A priori LLRs of a SISO module are IID with $L_p^{(\tau)}(d_{k,q}) \sim \mathcal{N}(\bar{d}_{k,q}\mu_p^{(\tau)}, \sigma_p^{2(\tau)})$, *such that* $\bar{d}_{k,q} \triangleq 1 - 2d_{k,q}$ *and* $\mu_p^{(\tau)}, \sigma_p^{2(\tau)} \geqslant 0$.

Property 1
If Assumption 1 holds, then, as LLRs are symmetrically distributed (more specifically, in the sense of exponential symmetry), i.e. $\mathbb{P}(L_p^{(\tau)}(d_{k,q})|d_{k,q} = 0) = \mathbb{P}(-L_p^{(\tau)}(d_{k,q})|d_{k,q} = 1)$, *then we have* $\sigma_p^{2(\tau)} = 2\mu_p^{(\tau)}$. *Hence, we have*

$$I(d_{k,q}; L_p^{(\tau)}(d_{k,q},)) = J(\mu_p^{(\tau)}), \tag{2.59}$$

with $J(\mu) \triangleq 1 - \int_L \log_2(1 + e^{-L})\mathcal{N}(L; \mu, 2\mu)\,\mathrm{d}L$.

Property 1 states that μ_a needed for input LLR generation at I_A can be obtained by the binary MI function $J^{-1}(\cdot)$. Hence an experimental setup would generate N_{try} blocks of KQ LLRs, $L_p^{(\tau)}(d_{k,q})[n]$, $n = 1, \ldots, N_{\text{try}}$ from a pseudo-random generator which follows the PDF $\mathcal{N}(\bar{d}_{k,q}\mu_p^{(\tau)}, 2\mu_p^{(\tau)})$, with $\mu_p^{(\tau)} = J^{-1}(I_A)$. The considered SISO module is fed with $L_p^{(\tau)}(d_{k,q})[n]$ and outputs $L_e^{(\tau)}(d_{k,q})[n]$, which does not follow a conditional Gaussian distribution in general. Hence, an histogram is computed for $p(L_e^{(\tau)}(d_{k,q})|d_{k,q})$, which is used to estimate I_E through the use of Equation (1.31). Indeed, the MI between an LLR L and a bit d is

$$I(L;d) = \mathbb{E}_{p(d,L)}\left[\log_2 \frac{p(L|d)}{p(l)}\right], \tag{2.60}$$

$$= \frac{1}{2}\sum_{d=0}^{1}\int_{L\in\mathbb{R}} \log_2 \frac{p(L|d)}{p(L|d=0) + p(L|d=1)} p(L|d)\,\mathrm{d}L, \tag{2.61}$$

$$= 1 - \int_{L\in\mathbb{R}} \log_2\left(1 + e^{-L}\right) p(L|d=0)\,\mathrm{d}L, \tag{2.62}$$

where respectively hypotheses of d being i.i.d. uniform distributed, and of L being distributed with exponential symmetry, i.e. $p(L|d=0) = p(-L|d=1)$, are used. Hence by inserting the measured histogram of $p(L_e^{(\tau)}(d_{k,q})|d_{k,q})$ into equation above, $I_E^{(\tau)} = I(L_e^{(\tau)}(d_{k,q}); d_{k,q})$ is estimated. This EXIT synthesis procedure is illustrated in Figure 2.6.

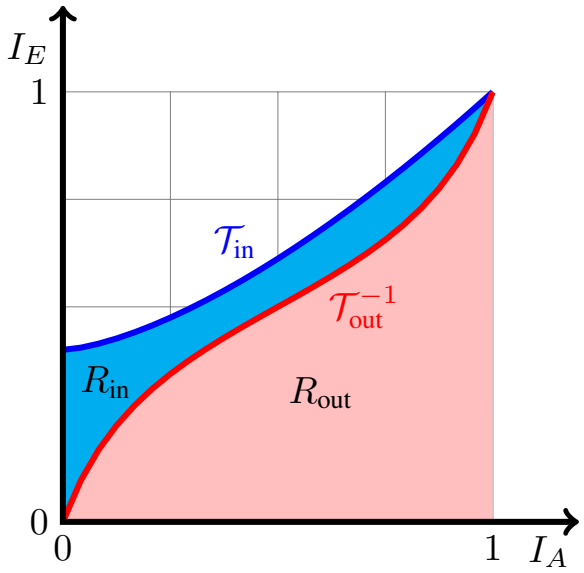


Figure 2.7: EXIT function area theorem.

The *area theorem* [AKt02] is a major property of EXIT analysis for assessing the asymptotic limits of serially-concatenated SISO modules, consisting of an inner and an outer decoder (in the general sense). The achievable rate R_{in} over the channel through the inner code satisfies

$$R_{\text{in}} \approx \int_0^1 \mathcal{T}_{\text{in}}(I_A)\,\mathrm{d}I_A, \tag{2.63}$$

such that if the code rate of the inner module is the unity, then C_{in} is also the *receiver-constrained achievable rate* over the channel. Moreover, the code rate of the outer decoder satisfy

$$R_{\text{out}} \approx \int_0^1 \mathcal{T}_{\text{out}}^{-1}(I_A)\,\mathrm{d}I_A, \tag{2.64}$$

and if $R_{\text{out}} < R_{\text{in}}$, such that the curves of $\mathcal{T}_{\text{in}}$ and $\mathcal{T}_{\text{out}}^{-1}$ do not intersect, then iterative detection and decoding can achieve asymptotically error-free communications at rate R_{out}, as illustrated in Figure 2.7. This important result has paved the way for code design techniques which adapt the shape of $\mathcal{T}_{\text{out}}^{-1}$ to the channel transfer function $\mathcal{T}_{\text{in}}$ in order to achieve the desired rate at a target operating point [TB00; RL09]. Moreover, if $\mathcal{T}_{\text{out}}^{-1}(1) = 1$, then the communications are asymptotically (infinite-length codeword) error-free, which is practical indicator for assessing the presence of an error-floor in BER or PER performance. When the EXIT of an optimal MAP SISO detector is computed, the receiver-constrained achievable rate coincides with the *symmetric information rate* of the channel, under BICM constraints, which generalizes the BICM capacity evoked in Section 1.5.1.1 of the previous chapter. Note that, if $\mathbf{H} = \mathbf{I}_K$, $\mathbf{\Sigma}_{\mathbf{w}} = \sigma_w^2 \mathbf{I}_K$ and MAP symbol demapper [BSY98] is used as the inner code, then BICM capacity of the constellation $\mathcal{X}$ can be estimated through the area theorem.

Another major use of EXIT analysis is for the asymptotic prediction of decoding performance through the use of a Lookup Table (LUT) that links the decoder prior information to

its BER or PER. In this case, for large K, the asymptotic error-rate performance of turbo receivers can be evaluated across turbo-iterations τ.

The consistent Gaussian model (Assumption 1) accurately characterizes EXIT functions for many cases, and in particular for the MAP detector, but some sub-optimal receivers' EXIT charts yield accurate predictions only when this technique is limited to small values of τ or for the asymptotic limit $\tau \to +\infty$. Indeed, as one SISO module's prior inputs are provided from another SISO module's extrinsic outputs, the prediction accuracy depends on whether the following assumption is true.

Assumption 2
There exists $\mu_e^{(\tau)} > 0$*, such that extrinsic LLR of a SISO module are approximately IID with* $L_e^{(\tau)}(d_{k,q}) \sim \mathcal{N}(\bar{d}_{k,q}\mu_e^{(\tau)}, 2\mu_e^{(\tau)})$*, and* $\mu_e^{(\tau)} \approx J^{-1}(I_E^{(\tau)})$.

The IID assumption is often true, in the large system limit thanks to interleaving, and Gaussian model holds for demapper output with Binary Phase Shift Keying (BPSK) and Gray-mapped Quadrature Phase Shift Keying (QPSK) constellations. However, the consistent Gaussian approximation of LLRs at the decoder output is lost across turbo iterations due to non-linear dynamics of channel decoders [Fu05]. Hence, EXIT can remain accurate for a few iterations, but its accuracy needs to be experimentally evaluated for sub-optimal detectors.

2.4 Self-Iterated SISO Detection with Approximate Inference

Within the last decade, there has been a variety alternative techniques to BP that are used for the design of emerging SISO detectors [QM07; Sen+11; SA11; SA12; Rie+13; SZW+15; Han+18]. These techniques belong to the general field of *approximate Bayesian inference*, which simplifies MAP inference in a complex stochastic system, through an approximated estimate of the posterior distribution. In particular, we are interested in variational Bayesian methods, which use deterministic approximations for carrying out inference, as opposed to sampling-based pseudo-random methods [Per+15]. Indeed, while the latter have a potential of yielding more precise approximations, they are a poor candidate for usage in embedded real-time platforms of practical receivers.

In this section, we will discuss how various approximate inference techniques could be used to derive *self-iterated* SISO receivers, which leads to a category double-loop turbo receiver architectures, as illustrated in Figure 2.8. Traditionally, algorithms with multiple-loops were avoided in turbo receiver design, due to drawbacks in computational complexity or due to absence of performance prediction mechanisms [TK+02; BC02; IB10; NVB12]. Such structures have recently gathered interest in the literature, due to their attractive performance, and thanks to the emergence of low-complexity implementations.

To this end, we first expose the concept of variational approximate inference on graphical models, which has its roots in statistical physics, and which seeks for the optimal approximate

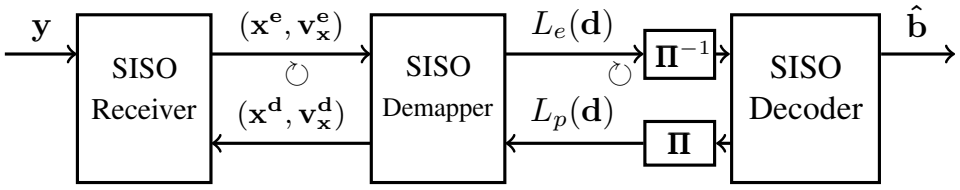


Figure 2.8: Doubly iterative turbo detection with self-iterated SISO detection.

but tractable posterior for carrying out inference on the true posterior. Through the use of *mean-field* and *Bethe* methods on graphical models, this optimization problem is transformed into a distributed local optimization problem whose optimum values are given by fixed-points of message passing algorithms. We show that BP belongs to this category of algorithms and discuss its potential, limitations and its variants that can overcome its issues [Min+05].

In particular, EP is a category of solutions that generalizes BP through a very flexible approach for computing local approximations, and we thus review some low-complexity instances of EP, mainly diagonal-EP and scalar-EP, for the considered linear detection model. The resulting SISO detector is a self-iterated module that generalizes the detector based on GaBP. We then discuss alternative self-iterated detector design techniques, based on Probabilistic Data Association (PDA) and an emerging category of parsimonious signal recovery algorithms called Approximate Message Passing (AMP). In particular, AMP-like algorithms have been thoroughly investigated for compressed sensing, and partially for communications. AMP has strong links to variational inference techniques and their asymptotic behavior can be predicted through *state-evolution*, which is a counterpart of density evolution or EXIT analysis.

2.4.1 On Approximate Inference with Variational Bayesian Methods

Variational methods are iterative techniques, used along with graphical models, for analytically approximating some intractable integrals that appear during Bayesian MAP inference. In this section, we have chosen to outline the differences of these algorithms from this free energy minimization point of view [YFW03; OW05], that has its roots in statistical physics.

Notion of free energy emerges in statistical mechanics, when considering a discrete physical system (e.g. modeling particles in a fluid) described by a discrete-valued state vector $\mathbf{x} = [x_1, x_2, \ldots, x_K]$, with an energy functional E, which maps each possible state $\mathbf{x}$ to an energy $E(\mathbf{x})$ [YFW05]. In thermal equilibrium, the PDF of the state vector follows a Boltzmann distribution

$$p(\mathbf{x}) = \frac{1}{Z} \exp\left(-\frac{E(\mathbf{x})}{k_B T}\right), \tag{2.65}$$

where T is the temperature, k_B is the Boltzmann's constant and Z is the normalization

constant of the PDF (i.e. evidence in statistics), called the partition function

$$Z = \sum_{\mathbf{x}\in\{\text{states}\}} \exp\left(-\frac{E(\mathbf{x})}{k_B T}\right). \tag{2.66}$$

Then the Helmholtz free energy $F_H \triangleq -k_B T \log Z$ is the useful work obtainable from a closed system, at constant temperature and volume, and it can be used to identify macroscopic characteristics of the system. However, the direct computation of Z is prohibitive for systems with a large number of states and to approximate $\log Z$, statistical physicists have developed variational techniques which replaces $p(\mathbf{x})$ by a simpler distribution $q(\mathbf{x})$. To this end, for a given proper PDF $q(\mathbf{x})$ (i.e. $\int_{\mathbf{x}} q(\mathbf{x})\,\mathrm{d}\mathbf{x} = 1$), the variational Gibbs energy is defined by

$$F_G(q) = \underbrace{\mathbb{E}_{q(\mathbf{x})}\left[E(\mathbf{x})\right]}_{\text{variational average energy}} - \underbrace{k_B T \mathbb{E}_{q(\mathbf{x})}\left[-\log q(\mathbf{x})\right]}_{\text{variational entropy}}, \tag{2.67}$$

and considering that $E(\mathbf{x}) = -k_B T(\log p(\mathbf{x}) + \log Z)$, Gibbs and Helholtz free energies follow

$$F_G(q) = F_H + k_B T D_{\mathrm{KL}}(q(\mathbf{x}) \| p(\mathbf{x})). \tag{2.68}$$

Variational techniques aim to estimate Helmholtz free energy by attempting to minimize the variational Gibbs free energy, its minimum being $F_G(p) = F_H$, as $D_{\mathrm{KL}}(q(\mathbf{x})\|p(\mathbf{x})) \geqslant 0$. To this end $q(\mathbf{x})$ is constrained to belong to a particular family of PDFs which simplifies computation of the evidence and marginals of $q(\mathbf{x})$.

Variational Free Energy Variational Bayesian inference methods have been inspired from this approach, when considering an inference problem with the joint posterior $p(\mathbf{x}, \mathbf{y}) = p(\mathbf{x}|\mathbf{y})p(\mathbf{y}) = p(\mathbf{y}|\mathbf{x})p_0(\mathbf{x})$, and assuming, as in many real-world problems, that the likelihood $p(\mathbf{y}|\mathbf{x})$ and the prior $p_0(\mathbf{x})$ are available. The computation of the posterior $p(\mathbf{x}|\mathbf{y})$ and its marginals is intractable due to integrals over high-dimensional latent variables, and to alleviate this, variational methods aim to approximate the true posterior $p(\mathbf{x}|\mathbf{y})$ by an approximate posterior $q(\mathbf{x})$, whose eventual dependence on $\mathbf{y}$ is omitted for notations, and which has a structure that ensures low-complexity computation of marginals $q(x_k)$, for $k = 1, \ldots, K$.

Hence, for any given true posterior $p(\mathbf{x}|\mathbf{y})$, the following optimization problem has to be addressed

$$q^\star(\mathbf{x}) = \arg\min_{q\in\mathcal{Q}} D_{\mathrm{KL}}(q(\mathbf{x})\|p(\mathbf{x}|\mathbf{y})), \tag{2.69}$$

where $\mathcal{Q}$ is the family of PDFs of the approximate posterior, and $F(q) \triangleq D_{\mathrm{KL}}(q(\mathbf{x})\|p(\mathbf{x}|\mathbf{y}))$ is known as the *variational free energy* [YFW01; Per+15]. In the following, we will discuss how the choice of $\mathcal{Q}$ impacts the free energy, to introduce popular message passing methods such as the MF approximation, loopy BP or EP, and then briefly discuss their generalizations.

When using a factor graph $\{\mathrm{F}_i\}_i$, $\{x_k\}_k$ to model the true posterior, it has been shown that common instances of the variational energy minimization problem is solved by message passing algorithms [Min+05; KFL01a], that carry out local energy minimization in a

distributed manner, on factor marginals $p_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i}) \triangleq \int_{\mathbf{x}^{\setminus \mathbf{x}_{\mathrm{F}_i}}} p(\mathbf{x})\, \mathrm{d}\mathbf{x}^{\setminus \mathbf{x}_{\mathrm{F}_i}}$, and/or on variable marginals $p_k(x_k) \triangleq \int_{\mathbf{x}^{\setminus x_k}} p(\mathbf{x})\, \mathrm{d}\mathbf{x}^{\setminus x_k}$. The general form of these algorithms exchange messages that satisfy

$$q_k(x_k) \propto m_{\mathrm{F}_i \to x_k}(x_k) m_{x_k \to \mathrm{F}_i}(x_k), \forall \mathrm{F}_i \in \mathcal{V}(x_k), \tag{2.70}$$

$$q_k(x_k) \propto \prod_{\mathrm{F}_i \in \mathcal{V}(x_k)} m_{\mathrm{F}_i \to x_k}(x_k), \tag{2.71}$$

where $q_k(x_k)$ is an approximation of $p_k(x_k)$. Then, we have [Min+05]

$$m_{x_k \to \mathrm{F}_i}(x_k) \triangleq \prod_{\mathrm{F}_j \in \mathcal{V}(x_k) \setminus \{\mathrm{F}_i\}} m_{\mathrm{F}_j \to x_k}(x_k), \tag{2.72}$$

and it is sufficient to define $m_{\mathrm{F}_i \to x_k}(x_k)$, in order to characterize a message passing algorithm. The expression of approximate posterior $q_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i})$ of factors marginals depends on the specific energy minimization method.

Mean-Field Energy The simplest category of approximate posteriors, for carrying out variational inference, is the set of fully-factorized posteriors without any constraints, i.e.

$$q(\mathbf{x}) = \prod_{k=1}^{K} q_k(x_k), \text{ such that } \int_{x_k} q_k(x_k)\, \mathrm{d}x_k = 1 \tag{2.73}$$

and resulting optimized distributions yield the *mean-field* approximation. In this case, the optimization problem becomes that of minimization of the *mean-field energy*, given by

$$F_{\mathrm{MF}}(q) \triangleq \int_{\mathbf{x}} \left[\prod_{k=1}^{K} q_k(x_k) \right] \log \frac{1}{p(\mathbf{x}, \mathbf{y})}\, \mathrm{d}\mathbf{x} - \sum_{k=1}^{K} \mathbb{E}\left[-\log q_k(x_k) \right]. \tag{2.74}$$

When the minimization of this energy is considered on the factor-graph of $p(\mathbf{x}|\mathbf{y})$, its optima are given by the fixed points of the following message passing algorithm

$$m_{\mathrm{F}_i \to x_k}(x_k) \triangleq \exp\left(\int_{\mathbf{x}_{\mathrm{F}_i}^{\setminus x_k}} \left[\log f_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i}) \right] \prod_{x_l \in \mathcal{V}(\mathrm{F}_i) \setminus \{x_k\}} q_l(x_l)\, \mathrm{d}\mathbf{x}_{\mathrm{F}_i}^{\setminus x_k} \right). \tag{2.75}$$

This is known as *variational message passing*, and it is a message passing algorithm that operates within the log-domain, which is particularly suited for many practical PDFs belonging to the exponential family. This algorithm is always convergent and it uses *posterior* information for updating factor nodes. Moreover, if the true factors are all of the same exponential family of distributions, then by restricting $q_k(x_k)$ to belong to the same family, the steady-state solution is given by moment matching with the true factor's marginals [Per+15; Min+05].

When there are hard constraints (e.g. delta distributions) underlying in the true posterior $p(\mathbf{x}, \mathbf{y})$, or more generally, when the true posterior is a multi-modal distribution, the MF

approximation becomes loose and its use is undesirable. Indeed, this technique locks on a mode of a PDF, and it approximates its non-zero behavior around the mode, ignoring any information that might lie elsewhere.

Bethe Free Energy with Strong Constraints (BP) Another intuitive category of approximate posteriors consists in mimicking the factorization structure of the true posterior, following a factor graph Equation (2.14), with local posterior approximations, such that $q_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i})$ approximates $f_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i})$, for $i = 1, \ldots, N_f$, and $q_k(x_k)$ approximates $p_k(x_k)$, for $k = 1, \ldots, K$, such that $\int_{x_k} q_k(x_k)\,\mathrm{d}x_k = 1$, and $\int_{\mathbf{x}_{\mathrm{F}_i}} q_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i})\,\mathrm{d}\mathbf{x}_{\mathrm{F}_i} = 1$ [YFW03]. Moreover, the following *strong consistency constraints* need to be satisfied

$$q_k(x_k) = \int_{\mathbf{x}_{\mathrm{F}_i}^{\setminus x_k}} q_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i})\,\mathrm{d}\mathbf{x}_{\mathrm{F}_i}^{\setminus x_k}, \quad \forall x_k \in \mathcal{V}(\mathrm{F}_i), \forall i = 1, \ldots, N_f. \tag{2.76}$$

If the factor graph has a tree structure, the joint approximate posterior is given by

$$q(\mathbf{x}) = \frac{\prod_{i=1}^{N_f} q_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i})}{\prod_{k=1}^{K} q_k(x_k)^{|\mathcal{V}(x_k)|-1}}, \tag{2.77}$$

and the variational Bayes problem reduces to the minimization of *Bethe free energy*

$$F_{\mathrm{B}}(q) \triangleq \sum_{i=1}^{N_f} D_{\mathrm{KL}}(q_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i}) \| f_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i})) + \sum_{k=1}^{K} (|\mathcal{V}(x_k)| - 1)\mathbb{E}\left[-\log q_k(x_k)\right]. \tag{2.78}$$

A major property is that, for tree-structured (acyclic) factor graphs, Bethe free energy is strictly equivalent to the Gibbs free energy $F_G(q)$, and approximating local posteriors of factors will yield the optimal solution, and $q(\mathbf{x}) = p(\mathbf{x}|\mathbf{y})$. In the general case, where the factor graph has cycles, the Bethe approach is only an approximation of the Gibbs free energy, and the original global optimization problem is replaced by sub-optimal local approximation problems on each factor, and the joint approximate posterior expression is unavailable.

It was shown that the fixed-points $\{q_k(x_k)\}$, $\{q_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i})\}$ of the Bethe free energy optimization coincides exactly with the fixed-points of the loopy Belief Propagation (BP) algorithm. As previously introduced, the message passing rules of BP are

$$m_{\mathrm{F}_i \to x_k}(x_k) \triangleq \int_{\mathbf{x}_{\mathrm{F}_i}^{\setminus x_k}} f_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i}) \prod_{x_l \in \mathcal{V}(\mathrm{F}_i)\setminus\{x_k\}} m_{x_l \to \mathrm{F}_i}(x_l)\,\mathrm{d}\mathbf{x}_{\mathrm{F}_i}^{\setminus x_k}, \tag{2.79}$$

and the approximate posteriors on factors are given by

$$q_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i}) \propto f_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i}) \prod_{x_l \in \mathcal{V}(\mathrm{F}_i)} m_{x_l \to \mathrm{F}_i}(x_l). \tag{2.80}$$

Unlike mean-field approximation, message updates are based on *extrinsic* information, and it is very-well suited for handling hard constraints.

However, computations in BP can become intractable in the general case, if underlying PDFs, are not Gaussian or categorical distributions, and its complexity significantly increases when continuous and discrete distributions are mixed. This greatly limits the scope of application of BP, and it ends up being commonly used either on full discrete graphs or full-Gaussian graphs. Gaussian-approximated Belief Propagation (GaBP) somewhat goes beyond this limitation by projecting extrinsic Gaussian messages and discrete messages from one to another.

Generalizations of BP address the sub-optimality of the Bethe approach, by grouping variable and factor nodes into clusters, such that each factor or variable only belong to a cluster [YFW05]. If the clusters are chosen such that there are no cycles between regions, then Gibbs free energy is replaced by the region-based Bethe free energy, which can be minimized. And if there is no way to construct cycle-free regions, then it is still possible to optimize BP by finding the region which has the lowest region-based Bethe free energy.

Bethe Free Energy with Weak Constraints (EP) Due to the attractive capability of Bethe free energy being equal or close to the Gibbs free energy, and the limited scope of BP, alternative techniques that have looser constraints were proposed. In particular, analytic tractability of these constraints enable working with clusters of variable nodes more easily, for modeling correlations among the variables. Following [YFW05], if there are different variable nodes x_k and $x_{k'}$ that appear together, i.e. $\mathcal{V}(x_k) = \mathcal{V}(x_{k'})$, it is possible to create a non-overlapping partition $\{\mathbf{x}_{\mathrm{V}_k}\}_{k=1}^{K_c}$ of $\mathbf{x}$ by grouping them together, with K_c being the number of clusters. Simple clustering options are the trivial partition with $\{\mathbf{x}\}$, or the full partition $\{x_k\}_{k=1}^{K}$.

The computationally intensive parts of BP reside in the strong consistency constraints, which require the marginals of the approximate posterior of a factor, i.e. $q_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i})$, to be equal to the approximate posteriors of connected variable clusters, i.e. $q_k(\mathbf{x}_{\mathrm{V}_k})$. An alternative approach, called *weak consistency constraint*, consists in simplifying this into a moment matching constraint between the marginals, for approximate factors belonging to the exponential family [RMO14; Per+15].

The exponential family of PDFs is defined by any distribution belonging to the set

$$\mathcal{Q} = \left\{ q(\mathbf{x}) = \prod_{k=1}^{K_c} h_k(\mathbf{x}_{\mathrm{V}_k}) \exp\left(\boldsymbol{\theta}_k^T \boldsymbol{\phi}_k(\mathbf{x}_{\mathrm{V}_k}) + g_k(\boldsymbol{\theta}_k)\right) \right\}, \tag{2.81}$$

for any complex vector $\mathbf{x} = \{\mathbf{x}_{\mathrm{V}_k}\}_{k=1}^{K_c}$, with each cluster having a real-valued natural parameter vector $\boldsymbol{\theta}_k$, a real-valued sufficient statistics function $\boldsymbol{\phi}_k(\mathbf{x}_{\mathrm{V}_k})$, of same dimension as $\boldsymbol{\theta}_k$, a real scalar log-partition function $g_k(\boldsymbol{\theta}_k)$ and a real positive base measure $h_k(\mathbf{x}_{\mathrm{V}_k})$. Real or complex valued Gaussian distributions and gamma distributions belong to this family, and as a counter-example the student's t-distribution does not belong to this family.

Then, in this case, the factor graph based approximation of local posteriors is such that $q_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i})$ approximates $f_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i})$, for $i = 1, \ldots, N_f$, and $q_k(\mathbf{x}_{\mathrm{V}_k})$ approximates $p_k(\mathbf{x}_{\mathrm{V}_k})$, for $k = 1, \ldots, K$, such that $\int_{\mathbf{x}_{\mathrm{V}_k}} q_k(\mathbf{x}_{\mathrm{V}_k}) \, \mathrm{d}\mathbf{x}_{\mathrm{V}_k} = 1$, and $\int_{\mathbf{x}_{\mathrm{F}_i}} q_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i}) \, \mathrm{d}\mathbf{x}_{\mathrm{F}_i} = 1$. Moreover $q_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i}) \in \mathcal{Q}$,

$\forall i$, and $q_k(\mathbf{x}_{\mathrm{V}_k}) \in \mathcal{Q}$, $\forall n$, and the following *weak consistency constraints* are satisfied

$$\mathbb{E}_{q_k}\left[\boldsymbol{\phi}_k(\mathbf{x}_{\mathrm{V}_k})\right] = \mathbb{E}_{f_{\mathrm{F}_i}\prod_{j\neq i} q_{\mathrm{F}_j}}\left[\boldsymbol{\phi}_k(\mathbf{x}_{\mathrm{V}_k})\right], \forall \mathbf{x}_{\mathrm{V}_k} \in \mathcal{V}(\mathrm{F}_i). \tag{2.82}$$

The variational Bayes problem still remains as the optimization of *Bethe free energy*, but under these looser constraints, the approximate posteriors of variables need only to share the same sufficient statistics (moments) with the approximate joint PDF. Moreover, the joint approximate posterior and factor node approximations are factorizable into cluster marginals

$$q(\mathbf{x}) \propto \prod_{i=1}^{N_f} q_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i}) \propto \prod_{k=1}^{K_c} q_k(\mathbf{x}_{\mathrm{V}_k}), \qquad q_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i}) \propto \prod_{\mathbf{x}_{\mathrm{V}_k}\in\mathcal{V}(\mathrm{F}_i)} q_{k,i}(\mathbf{x}_{\mathrm{V}_k}). \tag{2.83}$$

This optimization problem's solutions are shown to be the fixed-points of the Expectation Propagation (EP) message passing algorithm, given by [HZ02]

$$m_{\mathrm{F}_i\to\mathbf{x}_{\mathrm{V}_k}}(\mathbf{x}_{\mathrm{V}_k}) \triangleq \frac{1}{m_{\mathbf{x}_{\mathrm{V}_k}\to\mathrm{F}_i}(\mathbf{x}_{\mathrm{V}_k})}\mathrm{proj}_{\mathcal{Q}}\left[\int_{\mathbf{x}_{\mathrm{F}_i}^{\setminus\mathbf{x}_{\mathrm{V}_k}}} f_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i}) \prod_{\mathbf{x}_{\mathrm{V}_m}\in\mathcal{V}(\mathrm{F}_i)} m_{\mathbf{x}_{\mathrm{V}_m}\to\mathrm{F}_i}(\mathbf{x}_{\mathrm{V}_k})\,\mathrm{d}\mathbf{x}_{\mathrm{F}_i}^{\setminus\mathbf{x}_{\mathrm{V}_k}}\right], \tag{2.84}$$

where the argument of the projection operator is the *pre-projection posterior* of $\mathbf{x}_{\mathrm{V}_k}$ (also known as the BP posterior), at factor F_i, denoted

$$\tilde{q}_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{V}_k}) \triangleq \int_{\mathbf{x}_{\mathrm{F}_i}^{\setminus\mathbf{x}_{\mathrm{V}_k}}} f_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i}) \prod_{\mathbf{x}_{\mathrm{V}_m}\in\mathcal{V}(\mathrm{F}_i)} m_{\mathbf{x}_{\mathrm{V}_m}\to\mathrm{F}_i}(\mathbf{x}_{\mathrm{V}_m})\,\mathrm{d}\mathbf{x}_{\mathrm{F}_i}^{\setminus\mathbf{x}_{\mathrm{V}_k}}. \tag{2.85}$$

The projection operation carries out Kullback-Lieber minimization as reverse-information projection (as opposed to the information-projection in the variational free energy), which results in moment matching over the sufficient statistics of the selected PDFs family $\mathcal{Q}$.

From a solely factor node approximation point of view, EP operates by resolving

$$q_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i}) = \underset{q_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i})\in\mathcal{Q}}{\arg\min}\, D_{\mathrm{KL}}\left[f_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i}) \prod_{j=1,j\neq i}^{N_f} q_{\mathrm{F}_j}(\mathbf{x}_{\mathrm{F}_j}) \middle\| q_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i}) \prod_{j=1,j\neq i}^{N_f} q_{\mathrm{F}_j}(\mathbf{x}_{\mathrm{F}_j})\right], \tag{2.86}$$

for all $i = 1, \ldots, N_f$. The Kullback-Leibler divergence in the expression above is also known as the *expectation consistent* approximation of free energy [OW05], which is an alternative variational bound to the Gibbs free energy.

EP is an approximate method, with no convergence guarantee on graphs with cycles, as with BP, but unlike BP, the computational complexity of EP is traded off depending on the choice of the PDF on which the approximate posteriors are projected into. This provides a large selection of performance-complexity trade-off, depending on the size of the sufficient statistics, and it provides a tractable solution to complex graphical models.

The selection of the PDF to which approximations are projected can also be seen as a

choice of expectation consistent free energy to be minimized, and some common variants include [OW05; ÇO18]:

- diagonal free energy (or, diagonal EP), which uses uncorrelated Gaussian distributions, i.e. $q(\mathbf{x}) = \mathcal{CN}(\boldsymbol{\mu}_{\mathbf{x}}, \boldsymbol{\Sigma}_{\mathbf{x}})$, with $\boldsymbol{\Sigma}_{\mathbf{x}} = \mathbf{Diag}([\sigma^2_{x_1}, \dots, \sigma^2_{x_{N_v}}])$.
- diagonal-restricted free energy (or, scalar EP), with "white" IID Gaussian distributions, i.e. $q(\mathbf{x}) = \mathcal{CN}(\boldsymbol{\mu}_{\mathbf{x}}, \boldsymbol{\Sigma}_{\mathbf{x}})$, with $\boldsymbol{\Sigma}_{\mathbf{x}} = \sigma^2_w \mathbf{I}_{N_v}$.

More complex energy choices include correlations between latent variables, likely with a tree-like structure for simplifying message passing complexity [OW05]. A remarkable property of diagonal EP is that under some category of random matrix ensembles it can be shown that variances of the approximated variable distributions asymptotically converge towards the same value [Çak+16]. This suggests that the use of scalar EP, with some adjustments, could be sufficient in some applications.

Generalization of Variational Methods (Power EP) In [Min+05], an alternative perspective to the free energy approach is provided for deriving variational Bayesian inference techniques, by the means of a generalized divergence metric, called α-*divergence*, given by

$$D_\alpha(p(x)\|q(x)) = \frac{1}{\alpha(1-\alpha)} \int_x \left[\alpha p(x) + (1-\alpha)q(x) - p(x)^\alpha q(x)^{1-\alpha}\right] \mathrm{d}x, \tag{2.87}$$

with $\alpha \in \mathbb{R}$. Then generic variational inference method consists in finding $q(\mathbf{x})$ in a family of distributions $\mathcal{Q}$, such that $q(\mathbf{x}) = \arg\min_{q(\mathbf{x})\in\mathcal{Q}} D_\alpha(p(\mathbf{x})\|q(\mathbf{x}))$. In other words, the Kullback-Leibler divergence that appear in the free-energy minimization for classical variational methods is extended to a generalized divergence metric.

When the Bethe approach is selected, by considering $q(\mathbf{x})$ such that its marginals on factors and variables approximate that of the true posterior, a generic message passing algorithm is obtained, also known as the power Expectation Propagation (EP) algorithm. When $\alpha = 0$, this approach coincides with the mean-field approximation, otherwise, for $\alpha \neq 0$ the pre-projection posterior of a variable node is given as

$$\tilde{q}_{\mathrm{F}_i}(x_k) \propto \int_{\mathbf{x}_{\mathrm{F}_i}^{\setminus x_k}} f_{\mathrm{F}_i}(\mathbf{x}_{\mathrm{F}_i})^\alpha \prod_{x_m \in \mathcal{V}(\mathrm{F}_i)} m_{\mathrm{F}_i \to x_m}(x_m)^{1-\alpha} m_{x_m \to \mathrm{F}_i}(x_m) \, \mathrm{d}\mathbf{x}_{\mathrm{F}_i}^{\setminus x_k}, \tag{2.88}$$

and the associated factor messages are

$$m_{\mathrm{F}_i \to x_k}(x_k) \triangleq \frac{1}{m_{x_k \to \mathrm{F}_i}(x_k)} \mathrm{proj}_{\mathcal{Q}}\left[\tilde{q}_{\mathrm{F}_i}(x_k)\right]. \tag{2.89}$$

This perspective allows for relating previously-discussed message passing approaches between each other, indeed, when $\alpha = 1$, we get the EP algorithm. If $\mathcal{Q}$ is selected to be the same family of PDF as the original posterior (ensuring that projection operation is the identity),

then the generic message passing is also known as fractional Belief Propagation (BP) [WH03], and for $\alpha = 1$, we get the classical BP algorithm.

The advantage of this approach lies in the empirical observation that some divergence measures are better-suited to a factor graph, than others depending of the graph structure. Note that $\alpha = 1$ (i.e. EP/BP) is the only value for which the evidence, mean and the variance of the true posterior can be matched by those of the approximate posterior. More generally, it may be more advantageous to use different divergence metrics at each factor of a graph, either to improve performance, or to make computations tractable. An example to this is the hybrid BP and mean-field message passing framework, for complex factor graphs where different message passing rules are attributed to different regions [Rie+13].

Damped Message Passing Algorithms In general, there is no guarantee of convergence on variational inference algorithms based on factor-graph message passing. To circumvent this issue, the use of a certain amount of damping parameters is recommended in order to improve convergence to the global minimum [Min+05; Hes04]. This consists in smoothing the update of either messages, beliefs or their parameters using a certain weight β, which can be considered as a *learning rate* on local approximations of factors.

Indeed, due to Bethe free energy being non convex (due to the concave entropy term), its optimization requires a double-loop concave-convex optimization [Hes04]. While BP/EP (depending on the consistency conditions) has the same fixed points as the Bethe free energy, their iterative dynamics are different, and message passing algorithms are prone to getting stuck on erroneous local extrema. It can be shown that the "convexification" of the Bethe free energy with the double loop solution can be cast into an equivalent single-loop BP/EP algorithm where damping is used on the update rules of approximate posteriors [Pre05].

For instance, when damping extrinsic messages are sent from a factor, either a geometric weight could be used, with

$$m^{\text{new}}_{\text{F}_i \to x_k}(x_k) = m^{\text{old}}_{\text{F}_i \to x_k}(x_k)^{1-\beta} m_{\text{F}_i \to x_k}(x_k)^{\beta}, \tag{2.90}$$

or an arithmetic weight could be used, with

$$m^{\text{new}}_{\text{F}_i \to x_k}(x_k) = (1-\beta) m^{\text{old}}_{\text{F}_i \to x_k}(x_k) + (\beta) m_{\text{F}_i \to x_k}(x_k). \tag{2.91}$$

The geometric-approach can also be called *feature-based damping* as, for the exponential family, it results in arithmetic average of its sufficient statistics, and the arithmetic weight results in the *exponential smoothing* of the message.

When operating with the exponential family, in many practical message passing algorithms, instead of computing messages or probabilities, simply the sufficient statistics of related PDFs are computed. Hence, for simplicity, we could also consider exponential smoothing directly on the sufficient statistics of such PDFs.

While for optimized values of damping, the selection of the damping method itself has a small impact, the *sensitivity* of the inference accuracy to changes in damping parameter is drastically different from one method to another [Pre05].

2.4.2 Self-Iterated SISO Detector Design with Expectation Propagation

We have seen that GaBP-based SISO detector has the structure of linear estimator with IC, where soft symbol estimates based on decoder's extrinsic messages are used for regenerating the interference. However, repeatedly using SISO decoding can have significant computational or latency costs, hence the idea of iterating the detection process on its own can be attractive, if it has affordable computational complexity. This idea of self-iterating the detection process is already present in many fundamental techniques such as the well-known Decision Feedback Equalizer (DFE) structures used for ISI mitigation [Bel+79; Cio08; BT05], or the SIC structures in Multiple-User Detectors (MUDs) [TV05; CVB04; VBL08]. However there exist various approaches to the exact manner of designing the linear symbol detector (MMSE, zero-forcing, matched-filtering) and the interference re-generator (hard decisions, soft estimates). Different possibilities will be compared and discussed in detail in following chapters, and in this subsection, the derivation of a double-loop BICM receiver is discussed, for the linear model with the generic channel, through the use of Expectation Propagation (EP) message passing.

Similarly to GaBP, let us recall the factorization of the model

$$p(\mathbf{x},\mathbf{d}) \propto \underbrace{\exp\left(-\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2_{\mathbf{\Sigma_w}}\right)}_{=f_{\mathrm{EQU}}(\mathbf{x})} \prod_{k=1}^{K} \underbrace{\prod_{q=1}^{Q} \delta(d_{k,q}-\varphi_q^{-1}(x_k))}_{=f_{\mathrm{DEM}_k}(x_k,\mathbf{d}_k)} \underbrace{\exp(-d_{k,q}L_p(d_{k,q}))}_{=f_{\mathrm{DEC}_{k,q}}(d_{k,q})}, \tag{2.92}$$

where $d_{k,q} \in \mathbb{F}_2$, $\forall k,q$ but we assume $x_k \in \mathbb{C}$, unconstrained at first. Given this factor graph, for applying EP, we need first to choose a family of distributions $\mathcal{Q}$ associated to each variable node. For $d_{k,q}$, we keep using Bernoulli-distributed mode, hence messages involving this node will not incorporate projection, and for symbol variables, we consider the *diagonal EP* approach, among the previously discussed two options. To this end, we denote

$$q_{\mathrm{DEM}_k}(x_k) \propto \mathcal{CN}(x_k;\mu^d_{x,k},\gamma^d_{x,k}), \tag{2.93}$$
$$m_{\mathrm{DEM}_k\to x_k}(x_k) \propto \mathcal{CN}(x_k;x^d_k,v^d_{x,k}), \tag{2.94}$$
$$q_{\mathrm{EQU}}(x_k) \propto \mathcal{CN}(x_k;\mu^e_{x,k},\gamma^e_{x,k}), \tag{2.95}$$
$$m_{\mathrm{EQU}\to x_k}(x_k) \propto \mathcal{CN}(x_k;x^e_k,v^e_{x,k}), \tag{2.96}$$

which ensures that the family of symbol variables are de-correlated, and the joint distribution on $\mathbf{x}$ has a diagonal covariance matrix.

At the DEM factor node, the pre-projection posterior on x_k is given by

$$\begin{aligned}\tilde{q}_{\mathrm{DEM}}(x_k) &= \sum_{\mathbf{d}_k} f_{\mathrm{DEM}_k}(x_k, \mathbf{d}_k) m_{\mathrm{EQU}\to x_k}(x_k) \prod_{q=1}^{Q} m_{\mathrm{DEC}\to d_{k,q}}(d_{k,q}), \\ &\propto \sum_{\alpha\in\mathcal{X}} \exp\left(-\frac{|x_k^e-\alpha|^2}{v_{x,k}^e} - \sum_{q'=1}^{Q} \varphi_q'^{-1}(\alpha) L_p(d_{k,q'})\right) \delta(x_k-\alpha) = \sum_{\alpha\in\mathcal{X}} \mathcal{D}_k(\alpha), \end{aligned} \tag{2.97}$$

where $\mathcal{D}_k(\alpha)$ is the posterior PMF defined in Equation (2.53). Then the projection of this PDF onto $q_{\mathrm{DEM}_k}(x_k)$ enforces moment matching with

$$\mu_k^d \triangleq \mathbb{E}_{\mathcal{D}_k}[x_k] = \sum_{\alpha\in\mathcal{X}} \alpha \mathcal{D}_k(\alpha), \quad \gamma_{x,k}^d \triangleq \mathrm{Var}_{\mathcal{D}_k}[x_k] = \sum_{\alpha\in\mathcal{X}} |\alpha - x_k^d|^2 \mathcal{D}_k(\alpha). \tag{2.98}$$

Hence the extrinsic message of DEM is given by the division of the Gaussian PDFs, referred to as "Gaussian division"

$$m_{\mathrm{DEM}_k\to x_k}(x_k) = \frac{q_{\mathrm{DEM}_k}(x_k)}{m_{\mathrm{EQU}\to x_k}(x_k)} = \frac{\mathcal{CN}(x_k; \mu_{x,k}^d, \gamma_{x,k}^d)}{\mathcal{CN}(x_k; \mu_{x,k}^e, \gamma_{x,k}^e)} \propto \mathcal{CN}(x_k; x_k^d, v_{x,k}^d) \tag{2.99}$$

with

$$x_k^d \triangleq \frac{v_{x,k}^e \mu_{x,k}^d - \gamma_{x,k}^d x_k^e}{v_{x,k}^e - \gamma_{x,k}^d}, \quad v_{x,k}^d \triangleq \frac{v_{x,k}^e \gamma_{x,k}^d}{v_{x,k}^e - \gamma_{x,k}^d}. \tag{2.100}$$

At the EQU factor node, similarly, the pre-projection posterior is given by

$$\begin{aligned}\tilde{q}_{\mathrm{EQU}}(x_k) &= \int_{\mathbf{x}^{\setminus x_k}} f_{\mathrm{EQU}}(\mathbf{x}) \prod_{k'} m_{\mathrm{DEM}_{k'}\to x_{k'}}(x_{k'}) \, \mathrm{d}\mathbf{x}^{\setminus x_k}, \\ &\propto \int_{\mathbf{x}^{\setminus x_k}} \exp\left(-\|\mathbf{y}-\mathbf{H}\mathbf{x}\|_{\mathbf{\Sigma_w}}^2 - \sum_{k'} \frac{|x_{k'}-x_{k'}^d|^2}{v_{x,k'}^d}\right) \mathrm{d}\mathbf{x}^{\setminus x_k}, \\ &\propto \mathcal{CN}(x_k; \mu_{x,k}^e, \gamma_{x,k}^e), \end{aligned} \tag{2.101}$$

$$\gamma_{x,k}^e \triangleq \mathbf{e}_k^H \mathbf{\Gamma_x} \mathbf{e}_k = v_{x,k}^d (1 - v_{x,k}^d \mathbf{e}_k^H \mathbf{H}^H (\mathbf{\Sigma_w} + \mathbf{H} \mathbf{V_x^d} \mathbf{H}^H)^{-1} \mathbf{H} \mathbf{e}_k), \tag{2.102}$$

$$\mu_{x,k}^e \triangleq \mathbf{e}_k^H \boldsymbol{\mu}_{\mathbf{x}} = x_k^d + v_{x,k}^d \mathbf{e}_k^H \mathbf{H}^H (\mathbf{\Sigma_w} + \mathbf{H} \mathbf{V_x^d} \mathbf{H}^H)^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}^{\mathbf{d}}). \tag{2.103}$$

Finally, as $\tilde{q}_{\mathrm{EQU}}(x_k)$ already belongs to the diagonal-Gaussian family, $q_{\mathrm{EQU}}(x_k) = \tilde{q}_{\mathrm{EQU}}(x_k)$, and the extrinsic message from EQU is given by

$$m_{\mathrm{EQU}_k\to x_k}(x_k) = \frac{q_{\mathrm{EQU}_k}(x_k)}{m_{\mathrm{DEM}\to x_k}(x_k)} = \frac{\mathcal{CN}(x_k; \mu_{x,k}^e, \gamma_{x,k}^e)}{\mathcal{CN}(x_k; \mu_{x,k}^d, \gamma_{x,k}^d)} \propto \mathcal{CN}(x_k; x_k^e, v_{x,k}^e), \tag{2.104}$$

$$v_{x,k}^e \triangleq 1/\xi_k - v_{x,k}^d, \tag{2.105}$$

$$x_k^e \triangleq x_k^d + \xi_k^{-1} \mathbf{e}_k^H \mathbf{H}^H (\mathbf{\Sigma_w} + \mathbf{H} \mathbf{V_x^d} \mathbf{H}^H)^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}^{\mathbf{d}}). \tag{2.106}$$

where $\xi_k \triangleq \mathbf{e}_k^H \mathbf{H}^H (\mathbf{\Sigma_w} + \mathbf{H} \mathbf{V_x^d} \mathbf{H}^H)^{-1} \mathbf{H} \mathbf{e}_k$. Finally, the extrinsic messages from the demapper is computed with the same approach as GaBP with Equation (2.54).

It can be seen that these expressions bear great resemblance to the SISO detector based on GaBP, and the main difference lies in the soft feedback that is used for interference regeneration and cancellation in Equations (2.105-2.106), which uses DEM's extrinsic estimates $(\mathbf{x}^{\mathbf{d}}, \mathbf{v}_{\mathbf{x}}^{\mathbf{d}})$, from Equation (2.100), rather than DEC's extrinsic estimates (the whole detector's priors) $(\mathbf{x}^{\mathbf{p}}, \mathbf{v}_{\mathbf{x}}^{\mathbf{p}})$, from Equations (2.40).

An important advantage of $(\mathbf{x}^{\mathbf{d}}, \mathbf{v}_{\mathbf{x}}^{\mathbf{d}})$ is their dependence on EQU's outputs $(\mathbf{x}^{\mathbf{e}}, \mathbf{v}_{\mathbf{x}}^{\mathbf{e}})$, unlike DEC's estimates $(\mathbf{x}^{\mathbf{p}}, \mathbf{v}_{\mathbf{x}}^{\mathbf{p}})$. This opens up to possibility to compute EQU's and DEM's messages iteratively to refine symbol estimates, whereas such an inner iteration has no impact on the GaBP SISO detector. Hence, with EP messages, we can derive a turbo receiver where the SISO detector carries out $\mathcal{S}$ self-iterations (with messages between EQU and DEM factor nodes), before computing extrinsic messages towards the SISO decoder. Algorithm 1 illustrates such a detector with the diagonal EP example given above, denoted Double-Loop Diagonal EP (DL-DEP).

Algorithm 1 Double-Loop Diagonal Expectation Propagation

Require: $\mathbf{y}$, $\mathbf{H}$, $L_p(\mathbf{d})$
Ensure: $L_e(\mathbf{d})$
1: $x_k^e = 0$, $v_{x,k}^e = \infty$, for $k = 1, \ldots, K$
2: **for** $s = 0 \ldots \mathcal{S}$ **do**
3: Compute the posterior PMF $\mathcal{D}_k(\alpha)$ with Eq. (2.53), for $\alpha \in \mathcal{X}, k = 1, \ldots, K$
4: Compute $\boldsymbol{\mu}^{\mathbf{d}}$ and $\boldsymbol{\gamma}^{\mathbf{d}}$ with Eq. (2.98)
5: Get symbol estimates $\mathbf{x}^{\mathbf{d}}$ and $\mathbf{v}_{\mathbf{x}}^{\mathbf{d}}$ with Equation (2.100)
6: $\boldsymbol{\Xi} \triangleq \mathbf{Diag}\left(\mathbf{diag}\left(\mathbf{H}^H(\boldsymbol{\Sigma}_{\mathbf{w}} + \mathbf{H}\mathbf{V}_{\mathbf{x}}^{\mathbf{d}}\mathbf{H}^H)^{-1}\mathbf{H}\right)\right)$
7: $\mathbf{x}^{\mathbf{e}} = \mathbf{x}^{\mathbf{d}} + \boldsymbol{\Xi}^{-1}\mathbf{H}^H(\boldsymbol{\Sigma}_{\mathbf{w}} + \mathbf{H}\mathbf{V}_{\mathbf{x}}^{\mathbf{d}}\mathbf{H}^H)^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}^{\mathbf{d}})$
8: $\mathbf{v}_{\mathbf{x}}^{\mathbf{e}} = \mathbf{diag}\left(\boldsymbol{\Xi}^{-1}\right) - \mathbf{v}_{\mathbf{x}}^{\mathbf{d}}$
9: **end for**
10: Compute the posterior PMF $\mathcal{D}_k(\alpha)$ with Eq. (2.53), for $\alpha \in \mathcal{X}, k = 1, \ldots, K$
11: Compute extrinsic LLRs with Eq. (2.54)

In practice, the computation of extrinsic messages from DEM can be numerically unstable, or yield negative values due to poor realizations of noise and channel. To increase EP-based algorithms' robustness to these issues, several heuristic methods have been explored in the literature. When $v_{x,k}^e$ is detected be negative, common heuristics either change its sign [SZW+15], replace DEM extrinsic estimates with DEC's extrinsic estimates [Ć+14] or use DEM's APP estimates [Sen+11].

For ill-conditioned channel matrices, the iterative behavior of this receiver can be oscillatory, and it may converge towards sub-optimal fixed-points, and to overcome this, damping heuristics are useful [Min+05; Ć+14], at the expense of convergence speed.

If scalar EP would have been used, instead of diagonal EP, then approximate joint posteriors and extrinsic messages on $\mathbf{x}$ has a scalar variance, implying that the estimation error is white (uniform power density). This case will be discussed in more detail in Chapter 4.

2.4.3 Probabilistic Data Association

Probabilistic Data Association (PDA) is a filtering technique, originally used in signal-processing sub-fields related to target tracking, where measurements incorporate uncertainties that destabilize the tracking filter (e.g. Kalman filter) [BDH09]. The whole uncertainty is approximated by PDA as a single multi-variate Gaussian signal, and then its parameters are accounted for better estimation of the originally desired measurement. In communications applications, PDA has been used to approximate and then mitigate interference iteratively, and it has led to derivation of several self-iterated SISO detectors [Luo+01; GM08]. Although not explicitly stated, various other iterative algorithms or heuristics that use IC based on APP estimates can be more "formally" derived with the PDA framework, such as the iterative block DFE in [BT05] or time-domain DFE receivers discussed in Section 3.4.1.

The MMSE estimate $\mu^d_{x,k}$ of the transmitted symbol x_k, and the estimation error variance $\gamma^d_{x,k}$ are given by the conditional expectations of the symbol given the observations $\mathbf{y}$, with

$$\mu^d_{x,k} = \mathbb{E}[x_k|\mathbf{y}] = \sum_{\alpha\in\mathcal{X}} \alpha\mathbb{P}(x_k = \alpha|\mathbf{y}), \qquad \gamma^d_{x,k} = \mathbb{E}[|x_k - \mu_{x,k}|^2|\mathbf{y}]. \tag{2.107}$$

These APP estimates are the mean and the variance of the posterior PDF of x_k, given by $p(x_k|\mathbf{y}) = p(\mathbf{y}|x_k)p(x_k)$, and by rewriting the observations in Equation (2.1), the likelihood of x_k is given by

$$p(\mathbf{y}|x_k) \propto \sum_{\mathbf{x}\setminus x_k} \exp\left(-\|\mathbf{y} - \mathbf{h}_k x_k - \boldsymbol{\omega}_k\|^2_{\boldsymbol{\Sigma}_\mathbf{w}}\right), \tag{2.108}$$

with $\boldsymbol{\omega}_k \triangleq \sum_{k'\neq k} \mathbf{h}_{k'} x_{k'}$, and $\mathbf{h}_k = [\mathbf{H}]_{:,k}$. This likelihood has prohibitive computational complexity, due to $\mathbf{x}$ having a high number (M^{K-1}) of discrete values. Previous sections in this chapter discussed the estimation of the posterior through variational inference, but PDA proposes an alternative approach through this likelihood, based on the assumption that

$$\boldsymbol{\omega}_k \sim \mathcal{CN}(\boldsymbol{\mu}^\mathbf{d}_{\boldsymbol{\omega},k}, \boldsymbol{\Sigma}^\mathbf{d}_{\boldsymbol{\omega},k}), \tag{2.109}$$

$$\boldsymbol{\mu}^\mathbf{d}_{\boldsymbol{\omega},k} = \mathbb{E}[\boldsymbol{\omega}_k] \approx \sum_{k'\neq k} \mathbf{h}_{k'}\mu^d_{x,k'}, \qquad \boldsymbol{\Sigma}^\mathbf{d}_{\boldsymbol{\omega},k} = \mathrm{Var}[\boldsymbol{\omega}_k] \approx \sum_{k'\neq k} \gamma^d_{x,k'}\mathbf{h}_{k'}\mathbf{h}^H_{k'}, \tag{2.110}$$

where previously MMSE APP estimates are used when available, otherwise $\mu^d_{x,k} = 0$, and $\gamma^d_{x,k} = \sigma^2_x, \forall k$. This assumption ensures that $\mathbf{y} \sim \mathcal{CN}(\mathbf{h}_k x_k + \boldsymbol{\mu}^\mathbf{d}_{\boldsymbol{\omega},k}, \boldsymbol{\Sigma}_\mathbf{w} + \boldsymbol{\Sigma}^\mathbf{d}_{\boldsymbol{\omega},k})$, and the likelihood $p(x_k|\mathbf{y})$ becomes

$$\begin{aligned} p(\mathbf{y}|x_k) &= \frac{1}{\pi\det(\boldsymbol{\Sigma}_\mathbf{w} + \boldsymbol{\Sigma}^\mathbf{d}_{\boldsymbol{\omega},k})} \exp\left(-\|\mathbf{y} - \mathbf{h}_k x_k - \boldsymbol{\mu}^\mathbf{d}_{\boldsymbol{\omega},k}\|^2_{\boldsymbol{\Sigma}_\mathbf{w}+\boldsymbol{\Sigma}^\mathbf{d}_{\boldsymbol{\omega},k}}\right), \\ &\propto \exp\left(-\frac{|x_k - \mathbf{h}^H_k(\boldsymbol{\Sigma}_\mathbf{w} + \boldsymbol{\Sigma}^\mathbf{d}_{\boldsymbol{\omega},k})^{-1}\mathbf{h}_k\mu^d_{x,k} - \mathbf{h}^H_k(\boldsymbol{\Sigma}_\mathbf{w} + \boldsymbol{\Sigma}^\mathbf{d}_{\boldsymbol{\omega},k})^{-1}(\mathbf{y} - \mathbf{H}_k\boldsymbol{\mu}^\mathbf{d}_\mathbf{x})|^2}{\mathbf{h}^H_k(\boldsymbol{\Sigma}_\mathbf{w} + \boldsymbol{\Sigma}^\mathbf{d}_{\boldsymbol{\omega},k})^{-1}\mathbf{h}_k}\right), \\ &\propto \mathcal{CN}(x_k; x^e_k, v^e_{x,k}), \end{aligned} \tag{2.111}$$

where the squared Mahalanobis distance is expanded, and terms independent of x_k are re-

moved. Moreover, by applying the Woodbury matrix inversion lemma, we have

$$x_k^e = \mu_{x,k}^d + (\mathbf{h}_k^H \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \mathbf{h}_k)^{-1} \mathbf{h}_k^H \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} (\mathbf{y} - \mathbf{H}_k \boldsymbol{\mu}_{\mathbf{x}}^{\mathbf{d}}), \qquad v_{x,k}^e = (\mathbf{h}_k^H \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \mathbf{h}_k)^{-1} - \gamma_{x,k}^d, \tag{2.112}$$

with $\boldsymbol{\Sigma}_{\mathbf{y}} = \boldsymbol{\Sigma}_{\mathbf{w}} + \sum_{k=1}^{K} \gamma_{x,k}^d \mathbf{h}_k \mathbf{h}_k^H$, which has the well-known linear MMSE-IC estimator structure, based on APP estimates.

The next step for the APP MMSE estimation consists in accounting for the prior $p(x_k)$. In the context of BICM receiver with iterative detection, a prior PDF on symbols can be obtained through the extrinsic outputs of the SISO decoder $p_{\text{DET}}^a(d_{k,q})$, described by the detector's prior LLRs $L_p(d_{k,q})$. Then a prior PDF on x_k can be obtained following the same computations carried out for the DEM factor node's message for x_k in the examples provided for BP/GaBP/EP in sections above (Equation (2.39)), with

$$p(x_k) \propto \sum_{\alpha \in \mathcal{X}} \prod_{q=1}^{Q} \exp\left(-\varphi_q^{-1}(\alpha) L_p(d_{k,q})\right) \delta(x_k - \alpha). \tag{2.113}$$

Consequently, the posterior PDF on x_k is

$$p(x_k|\mathbf{y}) \propto \sum_{\alpha \in \mathcal{X}} \exp\left(-\frac{|x_k^e - \alpha|^2}{v_{x,k}^e} - \sum_{q=1}^{Q} \varphi_q^{-1}(\alpha) L_p(d_{k,q})\right) \delta(x_k - \alpha), \tag{2.114}$$

which is proportional to the posterior PMFs $\mathcal{D}_k(\alpha)$ seen before for GaBP and EP in Equations (2.53) and (2.97), but with $(x_k^e, v_{x,k}^e)$ being computed differently to those methods. Consequently, the APP MMSE estimates are given as

$$\mu_{x,k}^d = \mathbb{E}_{\mathcal{D}_k}[x_k] = \sum_{\alpha \in \mathcal{X}} \alpha \mathcal{D}_k(\alpha), \quad \gamma_{x,k}^d = \text{Var}_{\mathcal{D}_k}[x_k] = \sum_{\alpha \in \mathcal{X}} |\alpha - \mu_{x,k}^d|^2 \mathcal{D}_k(\alpha). \tag{2.115}$$

Equations (2.112) and (2.115) constitute fixed point equations for iteratively estimating $\mu_{x,k}^d$ and $\gamma_{x,k}^d$, and when the maximum number of inner iterations are reached (or PDA has converged), the extrinsic LLRs of coded bits can be computed following Equation (2.54).

2.4.4 On the links with Approximate Message Passing Algorithms

Recently, there has been a significant amount of contributions on iterative message passing algorithms for low complexity detection of parsimonious high-dimensional data. This sub-section discusses the application of theses algorithm for BICM detection in digital communications, while alluding to their strong ties to classical variational inference methods.

Approximate Message Passing (AMP) algorithms have been originally designed to estimate $\mathbf{x}$ on a noise-free linear model $\mathbf{y} = \mathbf{H}\mathbf{x}$, with $N \leqslant K$ and a given prior distribution $p_0(\mathbf{x})$. In typical use cases of these methods, $\mathbf{x}$ is sparse, and $p_0(\mathbf{x})$ is a distribution that holds the prior knowledge on the sparsity of the high-dimensional data vector.

AMP-like techniques have been extended to handle the generalized linear model (i.e. any likelihood $p(\mathbf{y}|\mathbf{H}\mathbf{x})$), which bodes well for symbol detection, with the assumption $\mathbf{\Sigma}_{\mathbf{w}} = \sigma_w^2\mathbf{I}$. Their algorithmic structure shares similarities to the self-iteration detection techniques in digital communications, based on variational inference, where data lies in a finite set $\mathcal{X}$, and various recent PHY algorithm proposals exploit these techniques [Guo+13; Wu+14; Zha+15a]. In this section, we summarize AMP-like iterative methods in the context of SISO detector design. In order to provide a clear view of each algorithm, notations have been unified with mostly the hat accent "ˆ" being used for the outputs of a linear estimation component, and the bar "¯" being the non-linear estimator (denoiser or shrinkage function) counterpart.

Iterative Thresholding Earlier approaches for low-complexity estimation of $\mathbf{x}$ consists in iterating a linear estimator with IC, and a non-linear component-wise estimator, in order to reconstruct sparse-solutions for under-determined systems [DDDM04; MD10].

Algorithm 2 Iterative Thresholding

Require: $\mathbf{y}$, $\mathbf{H}$, κ, $\eta(\cdot)$, $p_0(\mathbf{x})$
Ensure: $\bar{\mathbf{x}}$
1: $\bar{\mathbf{x}}^{(-1)} \leftarrow \mathbb{E}_{p_0}[\mathbf{x}]$
2: **for** $s = 0 \dots \mathcal{S}$ **do**
3: $\quad \hat{\mathbf{x}}^{(s)} \leftarrow \bar{\mathbf{x}}^{(s-1)} + \kappa\mathbf{H}^H(\mathbf{y} - \mathbf{H}\bar{\mathbf{x}}^{(s-1)})$
4: $\quad \bar{x}_k^{(s)} \leftarrow \eta(\hat{x}_k^{(s)}; v_{\hat{x}}), \forall k = 1, \dots, K$
5: **end for**

Iterative thresholding is given in Algorithm 2, where $\mathcal{S}$ is the number of iterations, κ a constant parameter and $\eta(\cdot)$ is a non-linear *shrinkage* or *threshold* function. This algorithm consists of a linear estimation step at line 3, which is more specifically a linear interference cancellation scheme with matched-filtering. Indeed estimates are provided by the non-linear function of line 4 are used to regenerate and cancel interference, similar to the use of demodulator or decoder feedback in GaBP/EP or PDA.

In the classical context of parsimonious estimation, the threshold function is given by a regularized least-squares

$$\eta(\hat{x}_k; v_{\hat{x}}) = \arg\min_{x_k \in \mathbb{C}} \left[\frac{\|\hat{x}_k - x_k\|^2}{v_{\hat{x}}} + f_s(x_k) \right], \tag{2.116}$$

where $v_{\hat{x}}$ is a weight, that can be interpreted as an estimated variance of residual error between $\hat{\mathbf{x}}$ and $\mathbf{x}$, and $f_s(\mathbf{x})$ is a regularization function, such as the ℓ_1-norm, or a function related to the prior constraint $p_0(\mathbf{x})$. Iterative thresholding is guaranteed to converge for convex $f_s(\cdot)$ [DDDM04], for some values of $v_{\hat{x}}$.

In a digital communications context, one can consider a shrinkage function related to the finite set $\mathcal{X}$ of x_k, which acts as a decision device. One option is to use

$$f_s(x) = \begin{cases} 0 & x \in \mathcal{X}, \\ +\infty & \text{otherwise}, \end{cases} \tag{2.117}$$

where x_k is projected into the constellation set $\mathcal{X}$, i.e. a hard decision is taken. Alternatively, to prevent loss of information, we can use the multinomial logistic regularization with the prior log-probabilities from the Equation (2.39), i.e.

$$f_s(x) = \begin{cases} \sum_{q=0}^{Q} \varphi_q^{-1}(x) L_p(d_{k,q}) & x \in \mathcal{X}, \\ 0 & \text{otherwise}, \end{cases} \tag{2.118}$$

which carries out Bayesian regularization, and performs the APP MMSE estimation $\eta(\hat{x}_k; v_{\hat{x}}) = \mathbb{E}[x_k|\hat{x}_k, v_{\hat{x}}]$ by assigning prior constraints to $p_0(\mathbf{x})$ from the prior PMF of the Equation (2.39). In the following, we refer to this case as Iterative Soft Thresholding (IST)[2].

Both of these functions require $\hat{x}_k$ to be unbiased, and to this end, $\kappa = 1/\left(\sum_{n=1}^{N} |h_{k,n}|^2\right)$ is enforced, which enables the assumption that $\hat{x}_k^{(s)} = x_k + \nu_k^{(s)}$, where $\nu_k^{(s)}$ is an AWGN, with $v_{\hat{x}} = \mathrm{Var}(\nu_k^{(s)}) \approx \sigma_w^2 \kappa$. The latter assumption neglects the impact of the residual interference, which is the default approach in iterative thresholding. It has been shown that iterative soft thresholding with the APP MMSE estimator, as the threshold function above, has the same fixed points as the mean field energy minimization [Krz+14].

The performance of iterative thresholding can be significantly improved by estimating the noise and residual interference with $v_{\hat{x}} = \mathrm{tr}(\mathbf{H}^H\mathbf{H})^{-1}(\|\mathbf{y}-\mathbf{H}\hat{\mathbf{x}}\|^2)$ [DHD12], where the variance estimation directly depends on the processed samples. Note that iterative thresholding and PDA both use APP estimates to carry out IC, but with different estimates on noise variance and while the former uses the normalized matched filter for the linear estimation, PDA uses an MMSE estimator. Consequently, iterative thresholding is a mean-field method and it is a sub-optimal variant of PDA which avoids matrix inversion by using a matched-filter as the linear estimator, instead of a conditional MMSE filter.

Algorithm 3 Approximate Message Passing

Require: $\mathbf{y}$, $\mathbf{H}$, κ, $\eta(\cdot)$, $p_0(\mathbf{x})$
Ensure: $\bar{\mathbf{x}}$
1: $\bar{x}_k^{(-1)} \leftarrow \mathbb{E}_{p_0}[x_k]$, $\epsilon_k^{(-1)} \leftarrow 0$ for $k = 1, \ldots, K$
2: **for** $s = 0 \ldots \mathcal{S}$ **do**
3: $\quad \hat{\mathbf{x}}^{(s)} \leftarrow \bar{\mathbf{x}}^{(s-1)} + \kappa \mathbf{H}^H(\mathbf{y} - \mathbf{H}\bar{\mathbf{x}}^{(s-1)}) + \boldsymbol{\epsilon}^{(s-1)}$
4: $\quad \bar{x}_k^{(s)} \leftarrow \eta(\hat{x}_k^{(s)}, v_{\hat{x}}), \forall k = 1, \ldots, K$
5: $\quad \epsilon_k^{(s)} \leftarrow \frac{K}{N}\left(\frac{1}{K}\sum_{k'=1}^{K} \frac{\partial \eta}{\partial \hat{x}_{k'}^{(s)}}\left(\hat{x}_{k'}^{(s)}; v_{\hat{x}}\right)\right)\left(\hat{x}_k^{(s)} - \bar{x}_k^{(s)}\right), \forall k = 1, \ldots, K$
6: **end for**

Approximate Message Passing While iterative thresholding has an attractive computational complexity, it fails to operate under ill-conditioned channels and its behavior cannot be accurately predicted. In [DMM09], Approximate Message Passing (AMP) is proposed to

[2] This denomination should not be confused with the convention used in compressed sensing literature, where it refers to the use of a rectified linear unit is used for shrinkage with ℓ_1 norms. For us IST refers to the generic Bayesian thresholding case in [Krz+14].

address these, by using an additive memory (or momentum) term called the *Onsager reaction term*, which was originally used to improve mean field techniques in statistical physics.

AMP is given in Algorithm 3, where the addition of the reaction term vector $\boldsymbol{\epsilon}$ is the main difference compared to iterative thresholding. Indeed, as we expect to have $\hat{x}_k = \bar{x}_k = x_k$ at ideal convergence, this component is an additive quantity that amplifies the estimation error, relative to the local sensitivity (i.e. rate of change) of the thresholding function. For SISO detection with the APP estimator function $\eta(\hat{x}_k; v_{\hat{x},k}) = \mathbb{E}[x_k|\hat{x}_k, v_{\hat{x},k}]$, one still has to enforce $\kappa = 1/\left(\sum_{n=1}^{N} |h_{k,n}|^2\right)$ in order to keep the estimates unbiased. Moreover, the derivative of the threshold function is

$$\frac{\partial \eta(\hat{x}_k; v_{\hat{x}})}{\partial \hat{x}_k} = \frac{1}{v_{\hat{x}}} \mathrm{Var}[x_k|\hat{x}_k, v_{\hat{x}}], \tag{2.119}$$

which can be fairly easily computed through a soft APP demapper, and AMP maintains the advantage of not having any matrix inversion. Similarly to iterative thresholding, the performance AMP can be significantly improved if the thresholding function accounts for the noise estimation $v_{\hat{x}} = \mathrm{tr}(\mathbf{H}^H\mathbf{H})^{-1}(\|\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}\|^2), \forall k$.

AMP has the advantage of being accurately predictable for mildly-conditioned channel matrices. The iterative dynamics can be evaluated through its asymptotic Mean Square Error (MSE) on x_k of $\hat{x}_k^{(s)}$ and $\bar{x}_k^{(s)}$, and yields scalar fixed-point equations which allows finding fixed-points. These aspects will be discussed in the next section.

Generalized Approximate Message Passing GAMP extends AMP to any observation model $p(\mathbf{y}|\mathbf{Hx})$ which can be fully-factorized on y_n, and it is also known as *Bayesian AMP* [Ran11]. It can be derived, with the use of the auxiliary variable $\mathbf{z} = \mathbf{Hx}$, on the factor graph of $p(\mathbf{x}|\mathbf{y}) \propto \prod_n p(y_n|z_n)\delta(\mathbf{z}-\mathbf{Hx})p_0(\mathbf{x})$, through BP message passing, followed by some simplifications. More precisely, messages from the factor $p(y_n|z_n)$ are simplified with the central limit theorem, and messages from variable nodes x_k are simplified through second order Taylor expansion [Ran11]. GAMP is given in Algorithm 4, and unlike AMP it incorporates the prediction of uncertainties, with diagonal covariance matrices $\mathbf{V}_{\hat{\mathbf{x}}} = \mathbf{Diag}([v_{\hat{x}_1}; \ldots; v_{\hat{x}_K}])$, which avoids the need of having a heuristic for $v_{\hat{x}_k}$, and also similarly for $\bar{\mathbf{x}}$, $\hat{\mathbf{z}}$ and $\bar{\mathbf{z}}$.

The MSE-optimal implementation of GAMP uses the threshold functions $\eta_{\text{in}}(\hat{\mathbf{x}}, \mathbf{V}_{\hat{\mathbf{x}}}) = \mathbb{E}[\mathbf{x}|\hat{\mathbf{x}}, \mathbf{V}_{\hat{\mathbf{x}}}]$ and $\eta_{\text{out}}(\hat{\mathbf{z}}, \mathbf{V}_{\hat{\mathbf{z}}}) = \mathbb{E}[\mathbf{z}|\hat{\mathbf{z}}, \mathbf{V}_{\hat{\mathbf{z}}}]$, where $\hat{\mathbf{x}}$ and $\hat{\mathbf{z}}$ are respectively considered to be uncorrelated noisy estimations of $\mathbf{x}$ and $\mathbf{z}$ (e.g. with diagonal covariance matrices). Moreover, $p(\mathbf{y}|\mathbf{z})$ being an Gaussian noise channel in our context, we have

$$\eta_{\text{out}}(\hat{\mathbf{z}}, \mathbf{V}_{\hat{\mathbf{z}}}) = \mathbb{E}[\mathbf{z}|\hat{\mathbf{z}}, \mathbf{V}_{\hat{\mathbf{z}}}] = (\boldsymbol{\Sigma}_{\mathbf{w}} + \mathbf{V}_{\hat{\mathbf{z}}})^{-1}(\boldsymbol{\Sigma}_{\mathbf{w}}\hat{\mathbf{z}} + \mathbf{V}_{\hat{\mathbf{z}}}\mathbf{y}), \tag{2.120}$$

$$\frac{\partial \eta_{\text{out}}}{\partial \hat{\mathbf{z}}}(\hat{\mathbf{z}}, \mathbf{V}_{\hat{\mathbf{z}}}) = \mathbf{Cov}[\mathbf{z}|\hat{\mathbf{z}}, \mathbf{V}_{\hat{\mathbf{z}}}] = (\boldsymbol{\Sigma}_{\mathbf{w}} + \mathbf{V}_{\hat{\mathbf{z}}})^{-1}\boldsymbol{\Sigma}_{\mathbf{w}}\mathbf{V}_{\hat{\mathbf{z}}}, \tag{2.121}$$

$$\eta_{\text{in}}(\hat{\mathbf{x}}, \mathbf{V}_{\hat{\mathbf{x}}}) = \mathbb{E}[\mathbf{x}|\hat{\mathbf{x}}, \mathbf{V}_{\hat{\mathbf{x}}}] = \mathbb{E}_{\mathcal{D}}[\mathbf{x}], \tag{2.122}$$

$$\frac{\partial \eta_{\text{in}}}{\partial \hat{\mathbf{x}}}(\hat{\mathbf{x}}, \mathbf{V}_{\hat{\mathbf{z}}}) = \mathbf{Cov}[\mathbf{x}|\hat{\mathbf{x}}, \mathbf{V}_{\hat{\mathbf{x}}}] = \mathbf{Cov}_{\mathcal{D}}[\mathbf{x}], \tag{2.123}$$

where $\mathcal{D}$ is the posterior PMF on $\mathbf{x}$, from the Equation (2.53). Moreover, in this case GAMP is

Algorithm 4 Generalized Approximate Message Passing

Require: $\mathbf{y}$, $\mathbf{H}$, κ, $\eta(\cdot)$, $p_0(\mathbf{x})$
Ensure: $\bar{\mathbf{x}}$

$\bar{x}_k^{(-1)} \leftarrow \mathbb{E}_{p_0}[x_k]$, $v_{\bar{x}_k}^{(-1)} \leftarrow \mathrm{Var}_{p_0}[x_k]$, $\epsilon_k^{(-1)} \leftarrow 0$ for $k = 1, \ldots, K$

for $s = 0 \ldots \mathcal{S}$ **do**

$$\hat{\mathbf{z}}^{(s)} \leftarrow \mathbf{H}\bar{\mathbf{x}}^{(s-1)} - \mathbf{V}_{\hat{\mathbf{z}}}^{(s)}\boldsymbol{\epsilon}^{(s-1)}, \qquad \mathbf{V}_{\hat{\mathbf{z}}}^{(s)} \leftarrow \mathbf{Diag}(\mathbf{diag}(\mathbf{H}^H\mathbf{V}_{\bar{\mathbf{x}}}^{(s-1)}\mathbf{H}))$$

$$\bar{\mathbf{z}}^{(s)} \leftarrow \eta_{\text{out}}\left(\hat{\mathbf{z}}^{(s)}, \mathbf{V}_{\hat{\mathbf{z}}}^{(s)}\right), \qquad \mathbf{V}_{\bar{\mathbf{z}}}^{(s)} \leftarrow \mathbf{V}_{\hat{\mathbf{z}}}^{(s)}\left(\frac{1}{N}\sum_{n=1}^{N}\frac{\partial\eta_{\text{out}}}{\partial\hat{z}_n^{(s)}}\left(\hat{z}_n^{(s)}, v_{\hat{z}_n}^{(s)}\right)\right)$$

$$\boldsymbol{\epsilon}_k^{(s)} \leftarrow \mathbf{V}_{\hat{\mathbf{z}}}^{(s)-1}(\bar{\mathbf{z}}^{(s)} - \hat{\mathbf{z}}^{(s)}), \qquad \mathbf{V}_{\epsilon}^{(s)} \leftarrow \mathbf{V}_{\hat{\mathbf{z}}}^{(s)-1}(\mathbf{I} - \mathbf{V}_{\bar{\mathbf{z}}}^{(s)}\mathbf{V}_{\hat{\mathbf{z}}}^{(s)-1})$$

$$\hat{\mathbf{x}}^{(s)} \leftarrow \bar{\mathbf{x}}^{(s-1)} - \mathbf{V}_{\hat{\mathbf{x}}}^{(s)}\mathbf{H}^H\boldsymbol{\epsilon}^{(s)}, \qquad \mathbf{V}_{\hat{\mathbf{x}}}^{(s)} \leftarrow \mathbf{Diag}(\mathbf{diag}(\mathbf{H}^H\mathbf{V}_{\boldsymbol{\epsilon}}^{(s)}\mathbf{H}))^{-1}$$

$$\bar{\mathbf{x}}^{(s)} \leftarrow \eta_{\text{in}}\left(\hat{\mathbf{x}}^{(s)}, \mathbf{V}_{\hat{\mathbf{x}}}^{(s)}\right), \qquad \mathbf{V}_{\bar{\mathbf{x}}}^{(s)} \leftarrow \mathbf{V}_{\hat{\mathbf{x}}}^{(s)}\left(\frac{1}{K}\sum_{k=1}^{K}\frac{\partial\eta_{\text{in}}}{\partial\hat{x}_k^{(s)}}\left(\hat{x}_k^{(s)}, v_{\hat{x}_k}^{(s)}\right)\right)$$

end for

shown to have the same-fixed points as the Bethe free energy, in the large-system limit where $N, K \to \infty$, with K/N remaining constant [Krz+14]. Remarkable properties of this algorithm are its provable and fast convergence for IID zero-mean Gaussian channel matrices, and the accurate predictability of its dynamics through state evolution. However the convergence of GAMP in general is strongly dependent on the singular values and conditioning of $\mathbf{H}$. In particular, when $p_0(\mathbf{x})$ and $p(\mathbf{y}|\mathbf{z})$ are Gaussians, and damping is applied on GAMP, through exponential smoothing of $\bar{\mathbf{x}}$ and $\bar{\mathbf{z}}$, there is always a value of damping that guarantees convergence, for any channel [RSF14]. In addition, the amount of necessary damping increases with the peak-to-average ratio of squared singular values of $\mathbf{H}$.

Finally, GAMP is an extension of the previously discussed AMP, which exhibits the same APP estimate-based linear IC structure, but with refined tracking of uncertainties in addition to the Onsager reaction term. While it does not use MMSE filtering as in PDA, it uses an enhanced matched-filter structure where the noise and residual interference uncertainties are accounted for, but by neglecting underlying inter-correlations. The presence of a bias compensation term remains as a considerable advantage for GAMP over PDA, but their behavior should be numerically compared for concluding in a given setting.

Orthogonal Approximate Message Passing In order to understand how AMP manages to outperform iterative thresholding and to alleviate the unpredictability issues for handling more general channel matrices, [MP17] proposes Orthogonal Approximate Message Passing (OAMP). The correlations between the linear and non-linear estimation errors are suggested to be among the underlying causes behind the prediction errors of state evolution for AMP. Thus, OAMP is proposed to ensure orthogonality between these errors, and removes the need for computing an Onsager reaction term, while also enabling predictability and convergence for a wider category of channels.

This technique is shown in Algorithm 5, where $\eta^\star(\cdot)$ and $\mathbf{W}^\star$ are respectively the constrained non-linear and the linear estimators. OAMP ensures the residual error at the output of the linear estimator to be orthogonal to the non-linear estimator's residual error, with the

Algorithm 5 Orthogonal Approximate Message Passing

Require: $\mathbf{y}$, $\mathbf{H}$, $\eta^\star(\cdot)$, $\mathbf{W}^\star$
Ensure: $\bar{\mathbf{x}}$
$\bar{x}_k^{(-1)} = \mathbb{E}_{p_0}[x_k]$, $v_{\bar{x}}^{(-1)} \leftarrow \mathrm{tr}(\mathbf{H}^H\mathbf{H})^{-1}(\|\mathbf{y} - \mathbf{H}\bar{\mathbf{x}}^{(-1)}\|^2 - N\sigma_w^2)$
for $s = 0 \dots \mathcal{S}$ **do**
  $\hat{\mathbf{x}}^{(s)} \leftarrow \bar{\mathbf{x}}^{(s-1)} + \mathbf{W}^\star(\mathbf{y} - \mathbf{H}\bar{\mathbf{x}}^{(s-1)})$
  $v_{\hat{x}}^{(s)} \leftarrow K\mathrm{tr}(\mathbf{W}^\star\mathbf{H})^{-1} - v_{\bar{x}}^{(s-1)}$
  if $s < \mathcal{S}$ **then**
    $\bar{x}_k^{(s)} \leftarrow \eta^\star(\hat{x}_k^{(s)}, v_{\hat{x}|x}^{(s)}), \forall k = 1, \dots, K$
    $v_{\bar{x}}^{(s)} \leftarrow \mathrm{tr}(\mathbf{H}^H\mathbf{H})^{-1}(\|\mathbf{y} - \mathbf{H}\bar{\mathbf{x}}^{(s)}\|^2 - N\sigma_w^2)$
  else
    $\bar{x}_k^{(s)} \leftarrow \eta_{\mathrm{out}}(\hat{x}_k^{(s)}, v_{\hat{x}}^{(s)}), \forall k = 1, \dots, K$
  end if
end for

latter being IID and independent of $\mathbf{A}$ and $\mathbf{w}$. For any given linear filter-bank $\mathbf{W} \in \mathbb{C}^{K\times N}$ and non-linear component-wise function $\eta(\cdot)$, these requirements are satisfied by

- a *"de-correlated"* linear estimator, i.e. $\mathrm{tr}(\mathbf{I}_K - \mathbf{WH}) = 0$,
- a *divergence-free* non-linear estimator, i.e. $\mathbb{E}[\frac{\partial\eta(\hat{x}_k)}{\partial\hat{x}_k}] = 0$, with $\hat{x}_k \sim \mathcal{CN}(x_k, v_{\hat{x}})$, $\mathbf{x} \sim p_0(\mathbf{x})$.

Such uncorrelated estimators $\eta^\star(\cdot)$ and $\mathbf{W}^\star$, can be constructed from any given shrinkage function $\eta(\cdot)$ and linear estimator $\mathbf{W}$ with

$$\mathbf{W}^\star = \frac{K}{\mathrm{tr}(\mathbf{WH})}\mathbf{W}, \qquad \eta^\star(\hat{x}_k) \propto \eta(\hat{x}_k) - \hat{x}_k\mathbb{E}\left[\frac{\partial\eta(\hat{x})}{\partial\hat{x}}\right]. \tag{2.124}$$

The divergence-free constraint ensures that the non-linear estimator incorporates naturally the impact of the Onsager reaction term, and the de-correlated constraint on the linear filter-bank removes the estimation bias. As seen in other AMP-like algorithms, estimators need to use the uncertainty of its priors as a parameter, and OAMP models uncertainty in $\hat{\mathbf{x}}$ and $\bar{\mathbf{x}}$ as AWGN with respectively a variance of $v_{\hat{x}}$ and $v_{\bar{x}}$.

In particular, the MSE-optimal parameters are given by

$$\mathbf{W}^\star = \frac{K}{\mathrm{tr}(\mathbf{H}^H(\sigma_w^2 + v_{\bar{x}}\mathbf{HH}^H)^{-1}\mathbf{H})}\mathbf{H}^H(\sigma_w^2 + v_{\bar{x}}\mathbf{HH}^H)^{-1}, \tag{2.125}$$

$$\eta^\star(\hat{x}_k) = \frac{v_{\hat{x}}\mu_{x,k} - \gamma_x\hat{x}_k}{v_{\hat{x}} - \gamma_x}, \tag{2.126}$$

$$\eta_{\mathrm{out}}^\star(\hat{x}_k) = \mu_{x,k}, \tag{2.127}$$

where $\mu_{x,k} = \mathbb{E}[x_k|\hat{x}_k]$ and $\gamma_x = \mathrm{Var}[x_k|\hat{x}_k]$. It is seen that the MSE-optimal linear and non-linear estimators share similar expressions with respectively the equalizer's and the de-modulator's extrinsic messages in EP algorithm. Main differences lie in the estimation of

the non-linear estimation reliability $v_{\bar{x}}$, and on the fact that OAMP uses scalar covariances. Indeed, while EP computes the variance of estimation noise at the demodulator's output through message-passing with Equation (2.100), OAMP uses least-squares estimation to estimate this variance. Consequently OAMP and the scalar-EP algorithms are very similar when the channel noise is white, and their differences would emerge from the use of damping heuristics in EP, or due to the impact of short block length on the estimation of $v_{\bar{x}}$.

Vector Approximate Message Passing Another AMP-like algorithm that is closely tied to EP is the Vector Approximate Message Passing (VAMP), proposed concomitantly to OAMP in [RSF17]. It can be seen as a more robust but computationally complex variant of GAMP: VAMP is derived directly from the application of scalar-EP message-passing on the linear detection model, along with the use of Singular Value Decomposition (SVD) on the channel matrix. Indeed, as the primary objective of AMP techniques is to perform low-complexity parsimonious detection in large systems, matrix inversions involved in EP is considered to be prohibitive.

Algorithm 6 Vector Approximate Message Passing

Require: $\mathbf{y}$, $\mathbf{H}$, $\eta_{\text{in}}(\cdot)$
Ensure: $\bar{\mathbf{x}}$

1: Perform compact SVD with $\mathbf{U}\mathbf{Diag}(\mathbf{s})\mathbf{V}^H = \mathbf{H}$, with $\mathbf{s} \in \mathbb{R}_+^R$, $R = \text{rank}(\mathbf{H})$, $\mathbf{U}^H\mathbf{U} = \mathbf{I}_R$, $\mathbf{V}^H\mathbf{V} = \mathbf{I}_R$
2: Compute $\tilde{\mathbf{y}} \leftarrow \mathbf{Diag}(\mathbf{s})^{-1}\mathbf{U}^H\mathbf{y}$
3: **for** $s = 0 \dots \mathcal{S}$ **do**
4: $\quad \boldsymbol{\mu}_{\bar{\mathbf{x}}}^{(s)} \leftarrow \eta_{\text{in}}\left(\hat{\mathbf{x}}^{(s)}, v_{\hat{x}}^{(s)}\right), \qquad \gamma_{\bar{x}}^{(s)} = v_{\hat{x}}^{(s)}\left(\frac{1}{K}\sum_{k=1}^{K}\frac{\partial \eta_{\text{in}}}{\partial \hat{x}_k^{(s)}}\left(\hat{x}_k^{(s)}, v_{\hat{x}}^{(s)}\right)\right)$
5: $\quad \bar{\mathbf{x}}^{(s)} \leftarrow (\boldsymbol{\mu}_{\bar{\mathbf{x}}}^{(s)} v_{\hat{x}}^{(s)} - \hat{\mathbf{x}}^{(s)}\gamma_{\bar{x}}^{(s)})/(v_{\hat{x}}^{(s)} - \gamma_{\bar{x}}^{(s)}), \qquad v_{\bar{x}}^{(s)} \leftarrow (v_{\hat{x}}^{(s)}\gamma_{\hat{x}}^{(s)})/(v_{\hat{x}}^{(s)} - \gamma_{\bar{x}}^{(s)})$
6: $\quad \mathbf{f}^{(s)} \leftarrow \mathbf{Diag}(\sigma_w^2 + v_{\hat{x}}^{(s)}\mathbf{s} \odot \mathbf{s})^{-1}(\mathbf{s} \odot \mathbf{s}), \qquad \xi^{(s)} \leftarrow \frac{R}{K}\left(\frac{1}{R}\sum_{k=1}^{R}\mathbf{e}_k^H\mathbf{f}^{(s)}\right)$
7: $\quad \hat{\mathbf{x}}^{(s)} \leftarrow \bar{\mathbf{x}}^{(s)} + \xi^{(s)-1}\mathbf{f}^{(s)} \odot (\tilde{\mathbf{y}} - \mathbf{V}^H\bar{\mathbf{x}}^{(s)})$
8: $\quad v_{\hat{x}}^{(s)} \leftarrow \xi^{(s)-1} - v_{\bar{x}}^{(s)}$
9: **end for**

If SVD can be carried out on $\mathbf{H}$, then linear estimator's output (equivalent to the EQU node's extrinsic message in Equation (2.106)) can be computed with orthogonal transformations and element-wise parallel operations. More specifically, [RSF17] uses the compact SVD where non-positive singular values are removed, and orthogonal transformations are truncated. VAMP therein, given in Algorithm 6, also proposes the use of exponential smoothing on the APP mean of DEM node, in Eq. (2.98), and to the extrinsic variance of EQU node, in Eq. (2.105), with β being the damping factor. To handle potentially negative extrinsic variances of the DEM node, its values are saturated to lie between 10^{-11} and 10^{11}.

2.4.5 Asymptotic Analysis for Self-Iterated Detection and Conclusions

Predictability of AMP-like algorithms has an important role in their analysis and optimization, and single-parameter state-evolution techniques were developed to this end [BM11]. Rigorous formalism for state-evolution is out-of-scope for this thesis, we will simply discuss their usage and implications.

State evolution consists in computing the fixed-point equations of the asymptotic Mean Square Error (MSE) at output of linear and non-linear estimators, in the large system limit, where $K, N \to +\infty$ with K/N remaining constant. More explicitly, the state evolution consists in analyzing the dynamics of $\varepsilon_{\bar{x},k}^{(s)} \triangleq \bar{x}_k^{(s)} - x_k$ and of $\varepsilon_{\hat{x},k}^{(s)} \triangleq \hat{x}_k^{(s)} - x_k$, with $\bar{x}_k^{(s)}$ and $\hat{x}_k^{(s)}$ being the estimates in AMP, VAMP or OAMP algorithms (for GAMP additional metrics are needed), through the characterization of $v_{\bar{x}} \triangleq \mathrm{Var}(\varepsilon_{\bar{x},k}^{(s)})$ and $v_{\hat{x}} \triangleq \mathrm{Var}(\varepsilon_{\hat{x},k}^{(s)})$, as self-iteration index s goes from 0 to $\mathcal{S}$. In the large system limit, components of error vectors are assumed to behave as independently distributed scalar random processes, and yield simple fixed-point equations, but rarely with analytical closed-form expressions. Random matrix theory is often used in such studies in order to remove the dependence of these scalar quantities from the channel matrix $\mathbf{H}$, and obtain equations that only depend on the ratio K/N, σ_w^2 and prior parameters.

Thanks to this asymptotic analysis technique, SISO detectors based on AMP have been used to evaluate the achievable rates of turbo receivers in [Liu+19b]. The resulting transfer functions are used for optimized LDPC code design, and it is seen that state evolution yields relatively accurate analysis and prediction of achievable rates and decoding threshold.

A concomitant work to the contributions in this thesis has outlined these advantages in [Ma+18], where OAMP-based state evolution is shown to be accurate in the context of iterative detection for AWGN MIMO communications systems. Indeed, APP estimate based systems are shown to be poorly predicted, and OAMP is shown to significantly outperform conventional turbo detection based on GaBP.

Besides, OAMP and VAMP are known to remain robust for a wider selection of channel matrices, while remaining accurately predictable with asymptotic analysis, which is enabled by the use of *extrinsic* estimates, instead of APP estimates. OAMP provides a framework proving that Bayes optimal MMSE AMP technique is based on direct application of EP, but it also enables decorrelating and removing the bias of any couple of linear and component-wise separable non-linear estimators [Ma+18]. On the other hand, VAMP enables analyzing techniques that can go beyond MMSE inference, by possibly generalizing the "linear estimator" to a more general shrinkage function, using MAP detectors as shrinkage functions [Pan+19]. This makes them an interesting candidate for predictable SISO detector design, which is one of our main motivations for studying EP-based approximate inference for receiver design throughout this thesis.

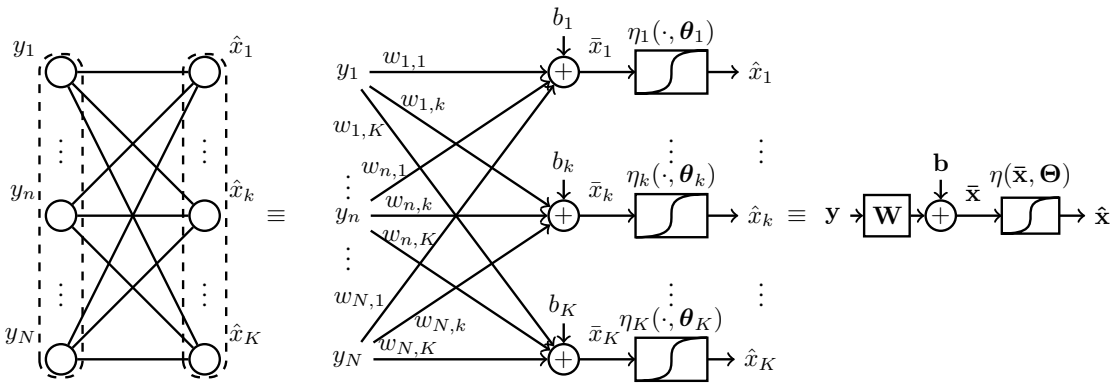


Figure 2.9: Parallels between neural network architectures and iterative detection

2.5 Unfolded Deep Receivers

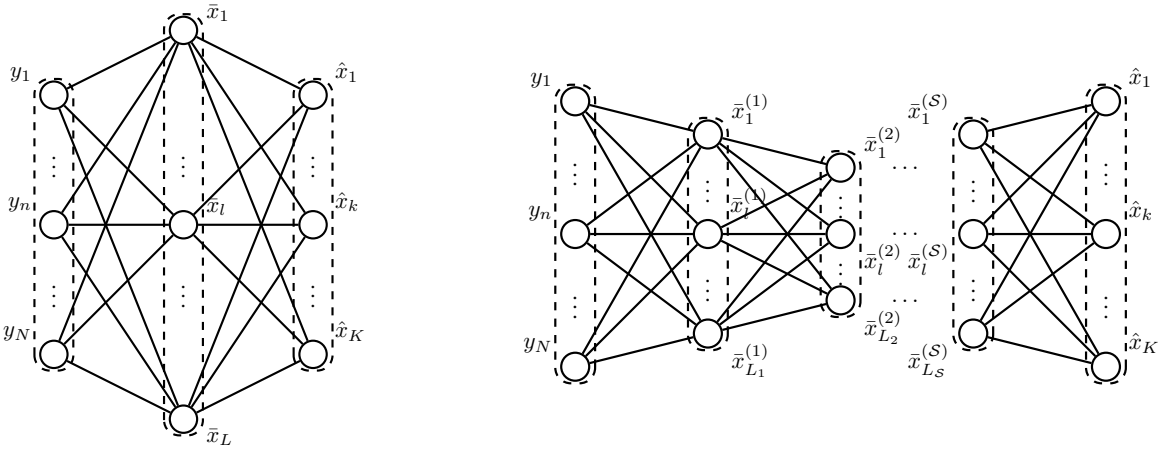
2.5.1 On Deep Learning for PHY Receivers

Wireless communications networks are man-made, model-based systems, and more specifically, as PHY layer transmitters use deterministic encoding and modulation techniques, receivers designed with model-based statistical algorithms are expected to be robust, when the physical channel modeling is accurate.

However as the computational complexity of the optimal receiver is prohibitive in most cases, and as practical transmission medium can incorporate poorly modeled phenomena, due to approximations used in receiver design. The application of emerging machine learning and deep learning techniques for mitigating these aspects gather a lot of attention for communications system design [ZDRD19]. The applications of machine learning in wireless systems are vast, here we will mainly discuss its usage for PHY layer receivers for data transmission [Dö+18].

Purely *data-based* deep learning strategies have been investigated for transceivers through auto-encoders, however practical interest for an end-to-end use remains limited, especially due to the prohibitive training costs required for channel coding [OH17]. Indeed, the required size of the training set for the the encoder-decoder network scales exponentially with the transport block size, hence auto-encoder techniques are more attractive when they are only partially used to handle the mitigation of neglected channel phenomena or system imperfections [CB19].

Alternatively model-oriented learning is more practical, when considered for learning existing algorithms' hyper-parameters, in order to optimize the receiver to account for neglected interference, correlations or other challenges in the underlying system [HLW14; NBB16; SDW19]. In these approaches, a neural detection or decoding networks is typically initialized with parameters from a conventional algorithm, and then optimization methods are used to fine tune some of the hyper-parameters of the network. Alternatively, there is a category of neural architectures that incorporates variational inference principles, but by also updat-



(a) One hidden layer (b) $\mathcal{S}$ hidden layers

Figure 2.10: Multi-layer perceptron as a universal approximator.

ing parameters of the underlying PDFs that characterize the probabilistic model, based on observed data [Zha+18; CB18]. These are out-of-scope of this thesis, but in the following section, the reasons why deep learning has seen such an important success will be discussed and their similarities with iterative algorithms commonly used for receiver design are underlined.

2.5.2 On Neural Networks and Growing Interest for Deep Learning

Early concepts of neural networks used multi-layer perceptrons to perform machine learning, which consists in identifying defining features of the observed process from the training data. The structure of an elementary layer of a perceptron is given in Figure 2.9, where the input signal is denoted $\mathbf{y} = [y_1; \ldots; y_N]$ and the output signal is denoted $\hat{\mathbf{x}} = [\hat{x}_1; \ldots; \hat{x}_K]$, to stay consistent with the inverse problem at hand of estimating transmitted data $\mathbf{x}$ from observations corrupted by noise and interference. This layer consists of a linear filter-bank $\mathbf{W}$, a bias vector $\mathbf{b}$ and a non-linear component-wise activation function η, with an associated parameter vector $\boldsymbol{\Theta} = [\boldsymbol{\theta}_1; \ldots; \boldsymbol{\theta}_K]$. The input of the activation function is an affine transformation, and the non-linearity of the activation functions play a key role in bringing learning capabilities when there are one or more hidden layers.

Indeed, a multi-layer perceptron with one-hidden layer, consisting of D components, is a universal approximator, [Hor91], meaning that the structure in the Figure 2.10-(a) can approximate any given function, with adequately chosen parameters, for a sufficiently large, but finite depth D. However, due to the potentially prohibitive complexity of such structures, early developments in neural networks has led to multi-layer perceptrons with more than one hidden layers, which keep the same approximation capability, but with smaller depths at hidden layers, as illustrated in the Figure 2.10-(b). In this case, filter-banks $\mathbf{W}^{(s)}$, biases $\mathbf{b}^{(s)}$ and parameters $\boldsymbol{\Theta}^{(s)}$ of the activation function $\eta^{(s)}$ for layers[3] $s = 1, \ldots, \mathcal{S}$ need to be

[3]We use the same indexing for layers and inner iterations due to their somewhat similar algorithmic roles.

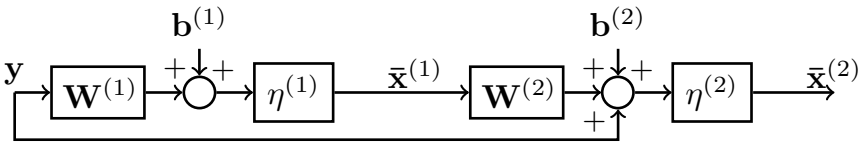


Figure 2.11: A residual layer with 2 inner layers, for building residual networks [He+16].

optimized through training techniques, such as back-propagation

Nevertheless, fully-connected layers suffer from over-fitting problems, and historically the use of *convolutional layers* is preferred for overcoming this issue. In this case $\mathbf{W}^{(s)}$ is forced to be a Toeplitz matrix, and the number of parameters to be trained is significantly reduced. Moreover, a certain local behavior is introduced through the filter-like nature of this affine transform, which increases robustness against some imperfections, such as delays in the observed signal. With the use of a succession of convolutional layers followed by pooling layers (i.e. down-sampling), convolutional neural network architectures have emerged and played a pivotal role in the field of machine learning since their discovery in 1980s [LeC+89]. More recently, convolutional networks have re-emerged in 2012, where deep architectures were shown to significantly overcome feature learning performance, by using a high number of hidden layers [KSH12]. Since then deep learning community is focused on seeking methods that enable further performance gain with additional layers [He+16].

This development has been possible thanks to several improvements on convolutional networks, such as pre-processing is used to *de-correlate*, *whiten*[4] and *normalize* the input of a layer. Moreover, the use of non-saturated activation functions, based on rectified non-linear unit, and the increased use of residual layers, where the input of a previous layer also impacts the activation function of the next layer, as illustrated in Figure 2.11, enabled increasing the number of hidden layers. The residual processing is particularly important as it enabled neural layers to approximate a function that can be close to the identity, which enabled further increasing the number of layers without degrading the performance of deep networks [He+16]. Finally, another major factor behind the revolution of deep learning lies in the emergence of advanced stochastic convex and non-convex optimization techniques that have facilitated learning of more complex architectures [DHS11; KB15].

2.5.3 Deep Unfolding: the Marriage of Iterative Inference and Deep Networks

When considering the developments in deep learning, with the previously discussed self-iterated SISO architectures in mind, one can draw strong parallels. First, the similarity between convolutional networks and self-iterated algorithms is significant; the linear estimator/filtering operation followed by a non-linear estimator is structurally similar to a convolu-

[4]Whitening refers to the pre-processing required to enforce a stationary behavior to the covariance of the concerned signals.

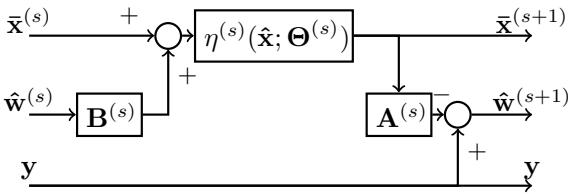


Figure 2.12: Deep unfolding of the iterative thresholding algorithm.

tional layer, as emphasized in Figure 2.9. There is also a strong similarity between the use of interference cancellation and damping operations in self-iterated filtering algorithms, and the use of residual connections in convolutional networks. Moreover the fact that normalization, decorrelation and whitening constraints further improve the capabilities of neural networks is an interesting parallel to the (MMSE) optimal iterative thresholding techniques [MP17]. The importance of having de-correlated and normalized inputs bears similarity to the use of the Onsager term in AMP or the derivation of constrained modules for OAMP.

These similarities are a source of motivation for evaluating deep learning techniques for deriving self-iterated detectors for PHY layer receivers. Indeed, recent results in [SDW19] explored the use of fully-connected networks and a more refined architecture based on residual networks and gradient-descent algorithm, to carry out iterative MIMO detection. This approach significantly outperforms AMP both in an ill-conditioned difficult channels and random mild-channels, and operated fairly close to sphere-detection performance. However, the design of such neural network architectures can require many trial-and-error experimentation, and requires expertise on how to choose right layers or dimensions to effectively resolve the detection problem.

For instance, through the use of a "soft-max" activation function, successive of convolutional layers within a multi-layer perceptron can be seen as an extension of iterative thresholding with a soft demapper. This has in fact gathered attention in the field of sparse signal recovery, where iterative soft thresholding has been used as an inspiration to develop deep networks [GL10], to get more robust detection structures. Recalling the update steps of iterative thresholding in the Algorithm 2, one can define $\mathbf{A}^{(s)} \triangleq \mathbf{H}$, $\mathbf{B}^{(s)} \triangleq \kappa\mathbf{H}^H$ and $\boldsymbol{\Theta}^{(s)} = v_{\hat{x}}$, and rewrite these steps as

$$\hat{\mathbf{w}}^{(s)} \leftarrow \mathbf{y} - \mathbf{A}^{(s-1)}\bar{\mathbf{x}}^{(s)}, \tag{2.128}$$

$$\hat{\mathbf{x}}^{(s)} \leftarrow \bar{\mathbf{x}}^{(s)} + \mathbf{B}^{(s)}\hat{\mathbf{w}}, \tag{2.129}$$

$$\bar{\mathbf{x}}_k^{(s)} \leftarrow \eta^{(s)}(\hat{\mathbf{x}}^{(s)};\boldsymbol{\Theta}^{(s)}). \tag{2.130}$$

Then the principle of *unfolding* is used, by considering $\mathbf{A}^{(s)}$, $\mathbf{B}^{(s)}$, $\boldsymbol{\Theta}^{(s)}$ as unknown parameters to be trained, for a given set of activation functions $\eta^{(s)}$, and thus building a potentially fully-connected neural layer with residual links, as in Figure 2.12, for deep networks.

This concept have been generalized and investigated based on various iterative algorithms for designing deep detection networks for PHY layer is *deep unfolding* [HLW14]. It consists

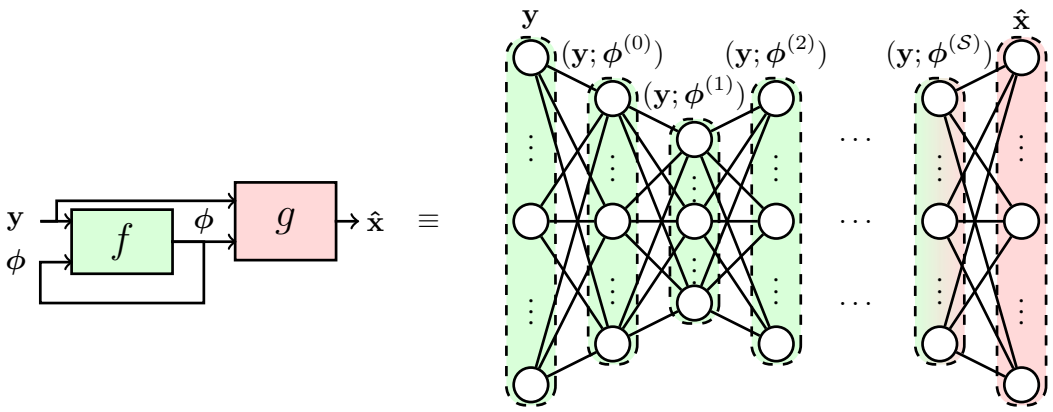


Figure 2.13: Deep unfolding for an iterative algorithm.

in generating a multi-layer neural network architecture from a given iterative algorithm that is used as a baseline for identifying the connections and parameters of the architecture. Then the parameters of the original algorithm can be left "free" for optimization through training. More specifically, we consider the state-space representation of an iterative algorithm with input vector $\mathbf{y} \in \mathbb{C}^N$ and an output vector $\hat{\mathbf{x}} \in \mathbb{C}^K$, such that

$$\hat{x}_k^{(s)} = g_k(\mathbf{y}, \boldsymbol{\phi}^{(s)}, \boldsymbol{\theta}_k^{(s)}), \quad k = 1, \ldots, K \tag{2.131}$$

$$\phi_m^{(s)} = f_m(\mathbf{y}, \boldsymbol{\phi}^{(s-1)}, \boldsymbol{\theta}_k'^{(s)}), \quad m = 1, \ldots, M, \tag{2.132}$$

where $\boldsymbol{\phi}^{(s)} = [\phi_1^{(s)}, \ldots, \phi_M^{(s)}]$ is a state vector with f_m and g_k being respectively state and output functions. Deep unfolding consists in representing such iterative algorithms as multi-layer deep feedforward networks, with parameters $\boldsymbol{\Theta}^{(s)}$ to be *chosen* and then optimized, as illustrated in Figure 2.13.

This strategy has been applied to damped BP for channel decoding [NBB16] and MIMO detection [Tan+18; He+18], where exponential smoothing and other damping parameters are optimized through learning. Alternatively, unfolded OAMP is evaluated for MIMO detection, where an attenuating factor on non-linear estimations is learned. These works have shown performance benefits of fine-tuning iterative algorithms through deep learning, and in this thesis, it will be discussed in Chapter 4, for the optimization of proposed EP-based algorithms.

2.6 Conclusion

Near-optimum iterative detection has seen light for practical Soft Input Soft Output (SISO) detector and decoder design, with the discovery of turbo codes in 1993 [BGT93]. The following decade has seen the derivation of a variety of SISO algorithms for addressing different communications problems through BP and GaBP [DJB+95; WP99; TSK02]. In particular powerful EXIT analysis tools enabled optimizing channel codes for operating at

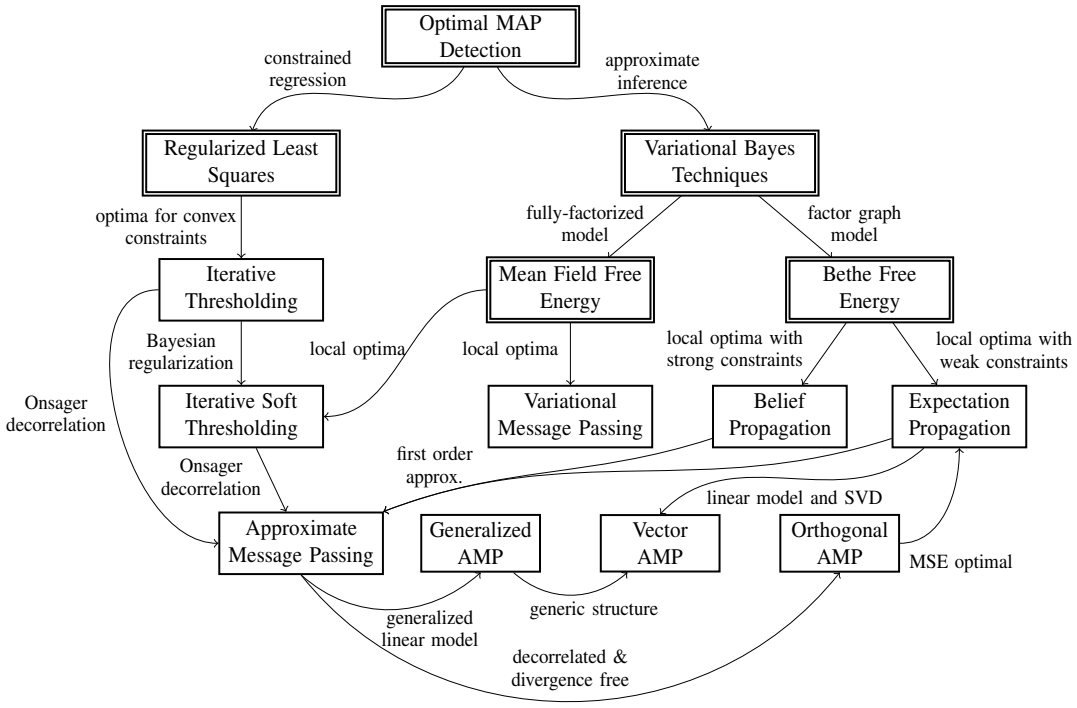


Figure 2.14: Mind map of iterative detection/estimation techniques discussed in this chapter. Double-lined blocks represent optimization criteria, whereas single-lined blocks are iterative solutions to these criteria.

the near-capacity achievable rate bounds of these receivers [TB00; Hag04; Ngu+15].

Meanwhile with the popularization of approximate statistical optimization methods, a wide selection of iterative detection or estimation algorithms [Sen+11; SA11; GM08; Guo+13] were designed based on variational Bayes methods, PDA or AMP algorithms, for addressing the problem of Equation (2.1). On one hand, variational Bayes methods address the issue of approximately computing MAP estimation through message passing algorithms based on graphical models, and provide a wide selection of methods with varying complexity, depending on the concerned problem. On the other hand, iterative thresholding algorithms, originally designed for regularized least-squares problems have given birth to various approximate message passing algorithms, which share strong ties. In Figure 2.14, a schematic provides a summary of relationships between these iterative algorithms, and following conclusions can be drawn about their advantages and their limits:

- BP is optimal when the system has an acyclic factor graph, otherwise it might yield algorithms of appropriate complexity if all the variables of the graph are either categorical (discrete) or continuous, from the exponential family.

- MF methods with variational message passing, enable carrying out message passing on graphs with variables belonging to one or more categories of continuous distributions, but it drastically fails with discrete variables.

- EP provides a means to manually decide which category of distributions each variable belongs to, which is enabled by moment matching for distributions in the exponential family. Although it is obtained from a looser approximation than BP, for loopy graphs it provides an alternate approach to seek fixed points of Bethe free energy.
- IST is a relatively simple algorithm for carrying out constrained regression by selecting a regularization function. Complexity and convergence depend strongly on the constraint at hand, in a turbo detection context, it behaves as a matched filter with APP estimate based interference cancellation.
- AMP/GAMP improves iterative thresholding through the use of a decorrelating Onsager term. They are provably predictable for IID Gaussian channel matrices.
- OAMP shows that Onsager term is not needed when the AMP components are decorrelated and divergence free, and in the MSE optimal case, it yields similar results to scalar EP.
- VAMP simplifies the use of scalar EP on linear models through the use of SVD, and provides a generic framework where shrinkage functions could go beyond the limitation of EP to exponential families.

Among those, EP message passing, and related algorithms such as OAMP or VAMP, have proven to improve detection performance in various scenarios, while having the remarkable property of being asymptotically predictable.

This chapter looks into signal processing techniques from different communities (communications theory, compressed sensing, artificial intelligence) with a common goal of resolving the problem of low-complexity detection of a constrained signal for use in the context of turbo detection of BICM transmissions. Hence it is aimed to be a synthesizing contribution to underline the strong and weak aspects of various emerging techniques.

In the remainder of this thesis, we are interested in investigating turbo detectors that also use an inner feedback from the demodulator. Thus, the above-mentioned strategies provide a means to derive doubly-iterative turbo receiver structures through the use of self-iterated detection algorithms. Moreover, following the emerging trends in using artificial intelligence for digital communications, we have briefly discussed the reasons behind the success of deep learning, their similarities to the model-based algorithms, and the concept of deep unfolding that could be used to optimize previously existing iterative algorithms with structure.

Chapter 3

Turbo Equalization and Closing the Gap on Achievable Rates

Contents

3.1 Introduction

In this chapter, the mitigation of Inter-Symbol Interference (ISI) is discussed with the use of *turbo equalization* techniques. Such receivers incorporate Soft Input Soft Output (SISO) detectors that either exploit the sparsity of the channel impulse response, or that use filtering structures to compensate the ISI. We have come to note that the literature on FIR filter-based turbo equalizers have only been studied for low rate applications. Hence, our broad objective is to investigate the capabilities of turbo FIR equalizers, and propose novel techniques, to push the achievable rates of such structures closer to the capacity bound, while trying to maintain a low computational complexity. In summary, the contributions brought by this chapter are as follows:

- a survey on turbo-equalization, where BP and GaBP message passing algorithms, from Chapter 2, have been used for detection in ISI channels, with
  - a brief discussion on trellis-based detection, underlining its links with the message passing framework, in Section 3.2,
  - followed by a discussion on classical block and FIR structures in Section 3.3.1,
  - and completed by a review of conventional turbo FIR solutions in Section 3.3.2,
- discussion on turbo DFE issues, derivation of novel solutions and their analysis:
  - generalization of turbo FIR equalizer model in Section 3.4.2,
  - proposal of an APP-feedback based novel DFE-IC solutions in Section 3.4.3,
  - the use of a factor graph framework to obtain an original DFE structure with a novel EP-based soft feedback in Section 3.4.4,
  - a novel matrix inversion technique is proposed in Section 3.4.5 to reduce complexity of exact (i.e. with dynamics filters) turbo FIR equalizers, by exploiting Cholesky decomposition and rank-1 updates with Givens rotation matrices [GVL96],
- derivation of *self-iterated* EP-based exact FIR LE and DFE, followed by finite-length performance and complexity analysis, and asymptotic EXIT analysis, in Section 3.5,
- an online-prediction approach for implementing a low-complexity FIR DFE (i.e. DFE with static filters), with APP or EP feedback, where the reliability of feedback symbols are predicted with a novel semi-analytic online prediction technique, in Section 3.6.

To better illustrate our contributions with respect to the literature, the case of *time-domain equalization* is discussed throughout this chapter, for single-user single-carrier transmission, and with a base-band channel model corrupted by time-varying ISI and AWGN, without interference from other blocks.

This model belongs to the general context of this thesis by considering the PHY model presented in Figures 1.5 and 1.6, by using a SC waveform with Zero Padding (ZP), with

$$\mathbf{F}_{\text{ZP}} = \left[\mathbf{0}_{L-1\times K}; \mathbf{I}_K; \mathbf{0}_{L-1\times K}\right], \quad \mathbf{G}_{\text{ZP}} = \left[\mathbf{0}_{(K+L-1)\times L-1}, \mathbf{I}_{K+L-1}\right], \tag{3.1}$$

where the baseband channel memory is $L-1$. The resulting $N = K+L-1$ observed baseband samples are given by

$$y_k = \sum_{l=1}^{L} h_{k,l} x_{k-l+1} + w_k, \qquad \forall k = 1, \ldots, K+L-1, \tag{3.2}$$

by assuming $x_{-L+1} = \cdots = x_0 = 0$ and $x_{K+1} = \cdots = x_{K+L} = 0$. Then, this time-domain ISI model can be cast into the general linear channel model in Equation (2.1) as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}, \tag{3.3}$$

where $\mathbf{y} = [y_1; \ldots; y_{K+L-1}]$, $\mathbf{w} \sim \mathcal{CN}(\mathbf{0}_{N\times 1}, \sigma_w^2 \mathbf{I}_N)$ and $\mathbf{H}$ is a sparse $(K+L-1)\times K$ matrix given by

$$\mathbf{H} = \begin{bmatrix} h_{1,1} & & & & \\ h_{2,2} & h_{2,1} & & & \\ \vdots & \ddots & & & \\ h_{L,L} & \ldots & h_{L,1} & & \\ & \ddots & \ddots & & \\ & h_{K,L} & \ldots & h_{K,1} & \\ & & \ddots & \vdots & \\ & & h_{K+L-2,L} & h_{K+L-2,L-1} & \\ & & & h_{K+L-1,L} & \end{bmatrix}. \tag{3.4}$$

When the channel is static, time-invariant, i.e. $h_{k,l} = h_{k',l}$, for all k, k', l, this channel matrix is a tall Toeplitz matrix.

3.2 On MAP Turbo Equalization

The problem of optimal MAP SISO detector design with Belief Propagation (BP) for a generic linear channel has been discussed in the previous Chapter, Section 2.3.3. That brute-force approach is discouraging due to the prohibitive computational costs of marginalization operations involved in message computations at the detection factor node, denoted EQU.

In this section, we discuss how the flexibility of factor graph formalism enables simplifying these problems thanks to the inherent sparsity of the channel model [Dec01; CB05]. Through this approach we recover well-known trellis-based optimal detection algorithms from the literature, and we then discuss how their complexity can be further reduced, at the expense of poorer performance.

This section shows that a same problem can have multiple solutions of varying complexity, depending on the graph and the schedule for which the message passing derivations are resolved.

3.2.1 Factor Graph Modelling of ISI Channels and BCJR

ISI mitigation on the time-domain observations is a sparse problem where each column of the channel matrix has only L non-zero elements. Hence it is possible to alleviate the complexity bottleneck of this problem, to an extent, by exploiting the sparsity of the problem, and the white noise assumption[1], which leads to the use of a category of techniques such as the Viterbi algorithm for ML detection, or BCJR algorithm for MAP detection [Bah+74; DJB+95]. In the following, a factor graph derivation of BCJR is given to illustrate the links between the general message passing framework and these algorithms.

First, the EQU factor node in Equation (2.27) is factorized thanks to the AWGN properties

$$f_{\mathrm{EQU}}(\mathbf{x}) \propto p(\mathbf{y}|\mathbf{x},\mathbf{H},\sigma_w^2) = \prod_{k=1}^{K+L-1} p(y_k|\mathbf{x},\mathbf{H},\sigma_w^2). \tag{3.5}$$

Next, by rewriting Equation (3.2) to isolate the term on x_k, we have

$$y_k = h_{k,1}x_k + \mathbf{h}_k^{\setminus 1 T}\mathbf{x}_{k-L+1}^{k-1} + w_k, \tag{3.6}$$

where $\mathbf{h}_k^{\setminus 1} = [h_{k,2};\ldots;h_{k,L}]$, and $\mathbf{x}_{k-L+1}^{k-1} = [x_{k-1};\ldots;x_{k-L+1}]$, and the auxiliary variable $\mathfrak{s}_k \triangleq \mathbf{x}_{k-L+1}^{k-1}$ is introduced for tracking the channel memory with $x_k \in \mathcal{X} \cup \{0\}$[2]. $\mathfrak{s}_k$ is in the set $\mathcal{S}_x = \bigtimes_{l=1}^{L-1}(\mathcal{X}\cup\{0\})$, which has a cardinality of $(M+1)^{L-1}$. This *state* variable is subject to a hidden Markov model, such that values of $\mathfrak{s}_{k+1}$ directly depend on $\mathfrak{s}_k$ and x_k, which is a hard constraint on allowed transitions through the state-space model

$$\mathfrak{s}_{k+1} = \begin{bmatrix} \mathbf{0}_{1\times(L-2)} & 0 \\ \mathbf{I}_{L-2} & \mathbf{0}_{(L-2)\times 1} \end{bmatrix}\mathfrak{s}_k + \begin{bmatrix} 1 \\ \mathbf{0}_{L-2} \end{bmatrix} x_k \triangleq T(\mathfrak{s}_k, x_k) \tag{3.7}$$

In the conventional view of the BCJR algorithm, the state transition is modelled on a trellis graph, where this functional $T : \mathcal{S}_x \times (\mathcal{X}\cup\{0\}) \to \mathcal{S}_x$, such that $(\mathfrak{s},x) \mapsto \mathfrak{s}' = T(\mathfrak{s},x)$, indicates the reachable future states on the Trellis from the previous state $\mathfrak{s}$ and the input x is transmitted. Considering this, likelihood on y_k, for $k = 1,\ldots,K$ is rewritten as

$$\begin{aligned} p(y_k|\mathbf{x},\mathbf{H},\sigma_w^2) &= p(y_k|x_k,\mathfrak{s}_k,\mathbf{H},\sigma_w^2), && (3.8)\\ &= p(y_k|\mathfrak{s}_{k+1},\mathfrak{s}_k,\mathbf{H},\sigma_w^2)\delta(\mathfrak{s}_{k+1} = T(\mathfrak{s}_k,x_k)), && (3.9)\\ &\propto \underbrace{\exp(-|y_k - h_{k,1}x_k - \mathbf{h}_k^{\setminus 1 T}\mathfrak{s}_k|^2/\sigma_w^2)\delta(\mathfrak{s}_{k+1} = T(\mathfrak{s}_k,x_k))}_{\triangleq f_{\mathrm{EQU}_k}(\mathfrak{s}_{k+1},\mathfrak{s}_k,x_k)}, && (3.10) \end{aligned}$$

with the initial state $\mathfrak{s}_1 = \mathbf{x}_{k-L+1}^{n-1} = \mathbf{0}_{(L-1)\times 1}$, due to zero-padding assumption, and for

[1] This is ensured by the partial matched-filter assumption, but in general, a Forney observation model would be required [For72].

[2] 0 is an additional value for handling zero-padding boundary effects of the data block.

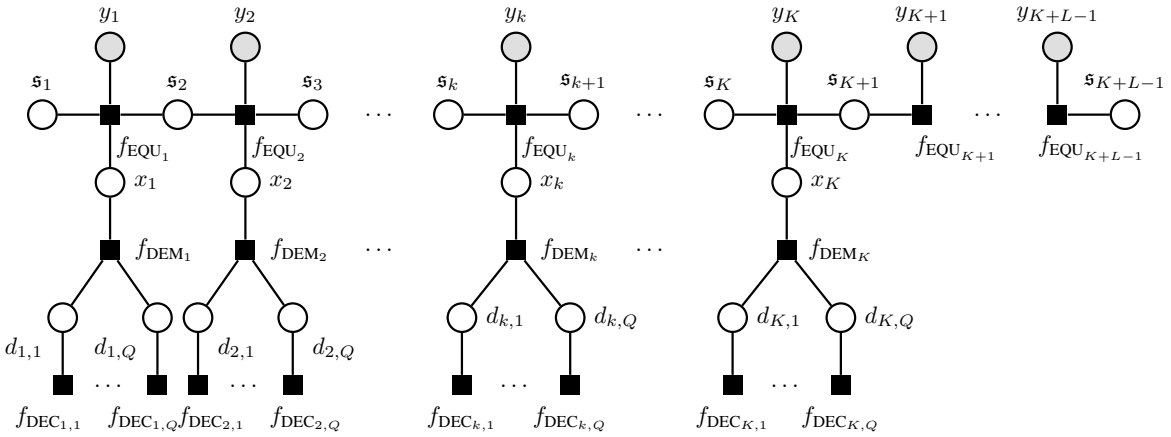


Figure 3.1: The factor graph of equation (3.12).

trailing observations, for $k = K+1, \ldots, K+L$

$$p(y_k|\mathbf{x}, \mathbf{H}, \sigma_w^2) = \underbrace{p(y_k|\mathfrak{s}_{k+1}, \mathfrak{s}_k, \mathbf{H}, \sigma_w^2)\delta(\mathfrak{s}_{k+1} = T(\mathfrak{s}_k, 0))}_{\triangleq f_{\mathrm{EQU}_k}(\mathfrak{s}_{k+1}, \mathfrak{s}_k)}. \tag{3.11}$$

Considering this, a novel factor graph is constructed by involving the intermediate variables $\mathfrak{s}_1, \ldots, \mathfrak{s}_{K+L}$, with

$$p(\mathbf{x}, \mathbf{d}) \propto \left(\prod_{k=1}^{K} f_{\mathrm{EQU}_k}(\mathfrak{s}_{k+1}, \mathfrak{s}_k, x_k) f_{\mathrm{DEM}_k}(x_k, \mathbf{d}_k) \prod_{q=1}^{Q} f_{\mathrm{DEC}_{k,q}}(d_{k,q}) \right) \prod_{k=K+1}^{K+L} f_{\mathrm{EQU}_k}(\mathfrak{s}_{k+1}, \mathfrak{s}_k), \tag{3.12}$$

which results in a Wiberg-type graph, as illustrated in Figure 3.1 [Wib96] .

This factor graph is acyclic, hence a scheduling consisting of a forward pass on $\mathfrak{s}_k$, followed by a backward pass on $\mathfrak{s}_k$ will yield the optimal SISO MAP detection.

The equivalence of this factor graph representation to the BCJR algorithm becomes apparent, when message passing rules are applied. Denoting $\gamma_k(\mathfrak{s}_k, \mathfrak{s}_{k+1}, x_k) \triangleq f_{\mathrm{EQU}_k}(\mathfrak{s}_{k+1}, \mathfrak{s}_k, x_k)$, priors as $p_a(x_k) \triangleq m_{\mathrm{DEM}_k \to x_k}(x_k)$, forward messages as $\alpha_k(\mathfrak{s}_k) \triangleq m_{\mathfrak{s}_k \to \mathrm{EQU}_k}(\mathfrak{s}_k)$, with $m_{\mathfrak{s}_k \to \mathrm{EQU}_k}(\mathfrak{s}_k) = m_{\mathrm{EQU}_{k-1} \to \mathfrak{s}_k}(\mathfrak{s}_k)$, and backward messages as $\beta_{k+1}(\mathfrak{s}_{k+1}) \triangleq m_{\mathfrak{s}_{k+1} \to \mathrm{EQU}_k}(\mathfrak{s}_{k+1})$, with $m_{\mathfrak{s}_{k+1} \to \mathrm{EQU}_k}(\mathfrak{s}_{k+1}) = m_{\mathrm{EQU}_{k+1} \to \mathfrak{s}_{k+1}}(\mathfrak{s}_{k+1})$, BP message passing rules yield

$$\alpha_k(\mathfrak{s}_k) = \sum_{\mathfrak{s}_{k-1}, x_{k-1}} \alpha_{k-1}(\mathfrak{s}_{k-1}) \gamma_{k-1}(\mathfrak{s}_{k-1}, \mathfrak{s}_k, x_{k-1}) p_a(x_{k-1}), \tag{3.13}$$

$$\beta_k(\mathfrak{s}_k) = \sum_{\mathfrak{s}_{k+1}, x_k} \beta_{k+1}(\mathfrak{s}_{k+1}) \gamma_k(\mathfrak{s}_k, \mathfrak{s}_{k+1}, x_k) p_a(x_k), \tag{3.14}$$

and the extrinsic symbol message in Equation (2.30) is simplified into

$$m_{\mathrm{EQU}\to x_k}(x_k) = \sum_{\mathfrak{s}_k,\mathfrak{s}_{k+1}} \alpha_k(\mathfrak{s}_k)\gamma_k(\mathfrak{s}_k,\mathfrak{s}_{k+1},x_k)\beta_{k+1}(\mathfrak{s}_{k+1}). \tag{3.15}$$

Finally, the extrinsic coded bit messages of the MAP detector are given by

$$m_{\mathrm{DEM}_k\to d_{k,q}}(d_{k,q}) = \sum_{\beta=0}^{1}\left[\sum_{\mathfrak{s}_k,\mathfrak{s}_{k+1},x_k} \alpha_k(\mathfrak{s}_k)\gamma_k(\mathfrak{s}_k,\mathfrak{s}_{k+1},x_k)\beta_{k+1}(\mathfrak{s}_{k+1}) \prod_{q'\neq q} e^{-\varphi_{q'}^{-1}(x_k)L_p(d_{k,q'})}\right]\delta(d_{k,q}-\beta), \tag{3.16}$$

which is used to compute extrinsic LLRs $L_e(d_{k,q})$. These equations show that MAP detection for ISI channel requires computing and storing roughly KM^L metrics[3], which are considerably less complex than the metrics of the brute force approach given by Equations (2.30) and (2.31).

Unlike [Bah+74; DJB+95; Dec01; CB05], forward-backward messages are given on the symbol level, i.e. through extrinsic message on x_k, in Equations (3.15), rather than directly marginalizing states over the coded bits $d_{k,q}$. We have made this choice in order to make different detection structures discussed in this thesis easily comparable among each other. As filter-based structures (discussed below) operate on the symbol level, this allows exposing differences and similarities of SISO detectors, through the same factor graph formalism.

This optimal detector has been theoretically investigated thoroughly for characterizing the extent and the limits of its convergence [SRF08; SRF06]. An important property of this forward-backward algorithm is its ability to yield an estimate of the Symmetric Information Rate (SIR) of ISI channels, i.e. the supremum of all rates. While the Equation (1.36) is an analytical a upper bound on ISI channel capacity with Gaussian distributed symbols, computing constellation-constrained capacity of ISI channels (i.e. SIR) is computationally prohibitive when L is greater than a few symbols. Hence the SISO BCJR algorithm alleviates this issue, to an extent, through the use of the area theorem of its EXIT chart, as explained in Section 2.3.6. Moreover, it has been shown that the SIR can be simply estimated by a forward recursion of this receiver's messages over the factor graph in Figure 3.1 [ALV+06].

While SIR is an effective bound for evaluating equalization performance, its computation can be prohibitive on channels with large delay spread. An alternative, but a looser upper bound for turbo detection performance is given by the Matched Filter Bound (MFB), which corresponds to the receiver behaviour, when only a single, interference-free symbol is transmitted, and matched-filter is used at the detector. When the mutual information of LLRs provided by the decoder gets close to 1, the BCJR performance converges towards MFB, and this property is also true for various sub-optimal turbo equalizers.

[3]The additional metrics needed due to $x_k = 0$ at the boundaries is neglected here.

3.2.2 Reducing the Computational Complexity of BCJR

Although BCJR is less complex than the generic MAP detection with brute-force enumerations, its computational and memory complexity still increases exponentially with L and in polynomial scale with M. While this receiver's use might be affordable for non-frequency selective channels, or for channels with memory using BPSK or QPSK constellations, for high-spectral efficiency systems or for channels with high delay spreads, alternative structures with less complexity are needed.

Among existing strategies for reducing the complexity of the BCJR detector, a wide category of receivers keep the Trellis graph structure, and make simplifications [FA98; CFR01; SC05; VAP06; Rad18]. These reduced-complexity structures seek to reduce the cardinality of $\mathcal{S}_x$, by removing the number of possible transition states [FA98], or by ignoring the impact some channel memory components (i.e. reducing L), and thereby reducing the number of possible states [CFR01]. These strategies can also be applied simultaneously, and possibly by self-iterating the detection process, as in [VAP06; Rad18], or alternatively L can be reduced with a channel-shortening filter [HR18; AFR18].

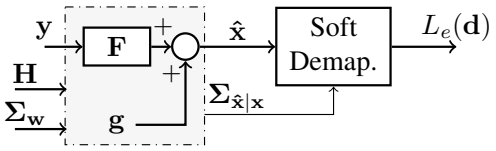
On the other hand, BCJR can be completely replaced by filter-based channel equalizers, such as turbo linear structures in [GLL97; WP99; BC02; Wit+02; TS11]. This is a limit-case of the BCJR simplified by channel-shortening, where the filter is designed in order to completely remove interference (i.e. $L = 1$), which enables the use of memoryless symbol-wise MAP soft demapper [BSY98].

There is also an alternative factor-graph and BP-based receiver, that carry out estimation only through symbol variable nodes x_k, either with a parallel or a sequential message passing schedule [CG05]. This iterative algorithm is sub-optimal due to the cycles of the graph, but with channels that satisfy some sparsity conditions, the cycle-girth becomes large enough, and this approach can approach BCJR performance. Moreover, by omitting some edges on the factor graph, the complexity of this approach is regulated [CFP11].

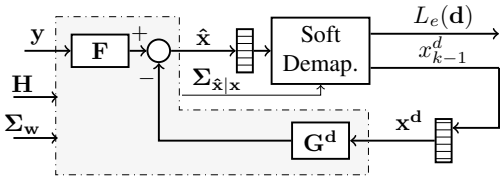
3.3 Filter-Based Turbo Equalization

Equalization of frequency-selective channels through filter-based receiver architectures has been a well-studied topic, as it is often related to the fundamentals of the estimation theory [Kay93]. Such techniques also enable (partially or completely) analytical analysis of receiver performance, which is an attractive property, for prediction and for quantifying properties of communications channels (diversity, spectral decomposition, etc.), through the use of infinite-impulse response channel models [Bel+79; Cio+95].

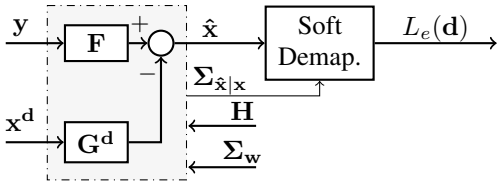
Extension of filtering techniques to finite-length architectures is obtained through optimization of filter-banks (block processing) or Finite Impulse Response (FIR) filters (windowed processing), with an optimization criteria such as least-squares or Minimum Mean Square Error (MMSE), and the latter leads to the optimal equalization performance, when combined



(a) Linear Equalizer



(b) Decision Feedback Equalizer



(c) Interference Canceller

Figure 3.2: Block equalization with filter-banks.

with a MAP symbol-wise demapper [Kal95; AC95b]. An interesting link with MAP detection and these techniques appears when transmitted symbols are assumed to be Gaussian-distributed, for which case, the unbiased MMSE filtering corresponds to the MAP estimation [Kay93], as also previously discussed in Section 2.3.5.

In this section, filter-based TD equalization is investigated by reviewing the classical structures and their extension to turbo equalizers, in order to point out some singularities between the behaviors of the SISO LE and DFEs receivers in the state of the art. This motivates the forthcoming investigations that is carried out in the remainder of this chapter.

3.3.1 On Classical Equalization

In this subsection, classical filter-based equalization architectures are summarized, to outline the technical vocabulary adopted in this thesis, with a focus on block and FIR equalizers.

3.3.1.1 Block Equalization

Block equalization structures process the whole received block simultaneously in order to mitigate the impact of ISI over each symbol [Kal95]. These structures apply filter-banks on observations and on available prior data estimates in order to remove interference, and they have the ability to follow the time-selectivity of the channel, by using the knowledge of $\mathbf{H}$. These structures are illustrated in Figure 3.2.

Block Linear Equalizer The block LE's structure consists of a filter-bank $\mathbf{F} \in \mathbb{C}^{N \times K}$ applied on $\mathbf{y}$ along with an additive bias vector $\mathbf{g} \in \mathbb{C}^K$, with equalized estimates being

$$\hat{\mathbf{x}} \triangleq \mathbf{F}^H \mathbf{y} + \mathbf{g}. \tag{3.17}$$

The *zero forcing* block LE is given by the solution to the *least-square* criterion

$$[\mathbf{F}^{\text{zf}}, \mathbf{g}^{\text{zf}}] = \arg\min_{\mathbf{F},\mathbf{g}} \|\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}\|^2 = \arg\min_{\mathbf{F},\mathbf{g}} \|(\mathbf{I} - \mathbf{H}\mathbf{F}^H)\mathbf{y} + \mathbf{H}\mathbf{g}\|^2, \tag{3.18}$$

which is straightforwardly shown to be $\mathbf{F}^{\text{zf}} = \mathbf{H}(\mathbf{H}^H\mathbf{H})^{-1}$ and $\mathbf{g}^{\text{zf}} = 0$. Alternatively, the zero forcing equalizer can also be derived with the *peak distortion criterion* which seeks $\mathbf{F}$ such that $\mathbf{F}^H\mathbf{H} = \mathbf{I}$, which yields the same pseudo-inverse solution for the tall channel matrix [Kal95]. However, zero forcing approach completely neglects noise, and in channels with spectral nulls the noise is amplified and performance is severely degraded.

In the remainder of this chapter, we will only investigate filters based on MMSE criterion. However, as MMSE solutions are biased by nature, in order to simplify the use of soft demapping, it is more attractive to use unbiaised-MMSE solutions. Following [Kay93; Kal95], the unbiased MMSE block LE is given by

$$\begin{aligned} \mathbf{F}^{\text{ble}} &= \mathbf{\Sigma}^{\text{ble}-1}\mathbf{H}\mathbf{\Xi}^{\text{ble}-1}, \\ \mathbf{g}^{\text{ble}} &= (\mathbf{I}_K - \mathbf{F}^{\text{ble}H}\mathbf{H})\boldsymbol{\mu}_{\mathbf{x}}, \end{aligned} \tag{3.19}$$

with $\mathbf{\Sigma}^{\text{ble}} \triangleq \sigma_w^2\mathbf{I}_N + \mathbf{H}\mathbf{\Sigma}_{\mathbf{x}}\mathbf{H}^H$, $\mathbf{\Xi}^{\text{ble}} = \mathbf{Diag}(\xi_1^{\text{ble}}, \ldots, \xi_K^{\text{ble}})$ and $\xi_k^{\text{ble}} \triangleq [\mathbf{H}^H\mathbf{\Sigma}^{\text{ble}-1}\mathbf{H}]_{k,k}$, where $\boldsymbol{\mu}_{\mathbf{x}} = \mathbb{E}[\mathbf{x}]$, $\mathbf{\Sigma}_{\mathbf{x}} = \mathbf{Diag}([\sigma_{x,1}^2, \ldots, \sigma_{x,K}^2])$, with $\sigma_{x,k}^2 = \text{Var}[x_k]$ by assuming $\mathbf{x}$ is uncorrelated.

The amount residual interference at the equalizer output is given by

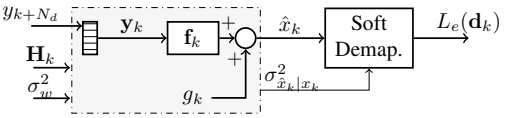
$$\begin{aligned} \sigma_{\hat{x}_k|x_k}^2 \triangleq \text{Var}[\hat{x}_k|x_k = x] &= \mathbf{e}_k^H\mathbf{F}^{\text{ble}H}(\sigma_w^2\mathbf{I}_N + \mathbf{H}\text{Var}[\mathbf{x} - \boldsymbol{\mu}_{\mathbf{x}}|x_k = x]\mathbf{H}^H)\mathbf{F}^{\text{ble}}\mathbf{e}_k \\ &= 1/\xi_k^{\text{ble}} - \sigma_{x,k}^2, \end{aligned} \tag{3.20}$$

and we denote $\mathbf{\Sigma}_{\hat{\mathbf{x}}|\mathbf{x}} = \mathbf{Diag}([\sigma_{\hat{x}_1|x_1}^2; \ldots; \sigma_{\hat{x}_K|x_K}^2])$. This is the optimal solution for block linear structures but the complexity of inverting a $K \times K$ matrix can be prohibitive for real-time implementation, depending on the symbol rate and the block size K. In classical usage, the symbol statistics are selected with IID assumption over the constellation, i.e. $\mu_{x,k} = 0$ and $\sigma_{x,k}^2 = \sigma_x^2 = 1$, for $k = 1, \ldots, K$ [Kal95].

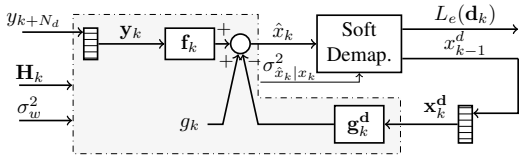
Block Decision Feedback Equalizer The use of a *decision feedback*[4] from the output of a detection device, in order to improve equalization, by removing the interference caused by causal symbols, was initially proposed and thoroughly analysed on infinite impulse response filters [Aus67; FF73; Bel+79]. It has been also used to reduce the complexity of the Maximum Likelihood Sequence Estimation (MLSE) [WH77], and later extended to be applicable to FIR receivers in [AC95b], where the feedforward filter minimizes the impact of ISI from the pre-cursor symbols, and the feedback filter cancels the impact of remaining ISI from the post-cursor symbols.

Block DFE structure is a block extension of the FIR DFE (see below), analogous to the block LE, where in addition to the feedforward filterbank $\mathbf{F} \in \mathbb{C}^{N\times K}$, a feedback filterbank

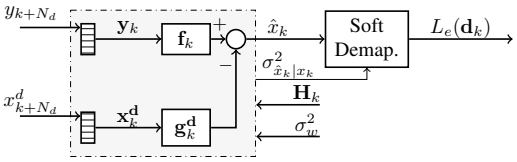
[4] Here the term "decision feedback" refers to a feedback provided by a demodulator, and not to the prior feedback provided to the SISO detector.



(a) Linear Equalizer



(b) Decision Feedback Equalizer



(c) Interference Canceller

Figure 3.3: Sliding window equalization with Finite Impulse Response (FIR) filters.

$\mathbf{G}^{\mathbf{d}} \in \mathbb{C}^{K\times K}$ is used, such that equalized estimates are [Kal95]

$$\hat{\mathbf{x}} \triangleq \mathbf{F}^H\mathbf{y} - \mathbf{G}^{\mathbf{d}H}\mathbf{x}^{\mathbf{d}}, \tag{3.21}$$

with $\mathbf{x}^d = [x_1^d, \dots x_K^d]$, and $\mathbf{G}^{\mathbf{d}}$ is restricted to be a strictly upper triangular matrix. Without the latter constraint, this approach would lead to a block IC structure (discussed below for FIR) where the decision feedback would be needed over the whole block in parallel. However the one of the key reasons behind FIR DFE's success is the successive cancellation of interference, which enables the possibility of removing interference using the ongoing equalization session's (probably) higher quality estimates. Hence to enforce similar characteristics on block structures, this structural constraint is needed, however this also makes the resolution of MMSE criterion non-trivial. Nevertheless, by maintaining the classical perfect decision feedback assumption (see the description of the FIR DFE), an unbiaised MMSE optimal solution is given by a two-step optimization process [Ver98; Rou00],

$$\mathbf{F}^{\text{bdfe}} = (\mathbf{I} - \mathbf{D}_{\boldsymbol{\epsilon}}^{-1}\boldsymbol{\Sigma}_{\mathbf{x}}^{-1})^{-1}\sigma_w^{-2}\mathbf{H}\mathbf{L}_{\boldsymbol{\epsilon}}^{-H}\mathbf{D}_{\boldsymbol{\epsilon}}^{-1}, \tag{3.22}$$

$$\mathbf{G}^{\mathbf{d}\text{-bdfe}} = (\mathbf{I} - \mathbf{D}_{\boldsymbol{\epsilon}}^{-1}\boldsymbol{\Sigma}_{\mathbf{x}}^{-1})^{-1}\mathbf{L}_{\boldsymbol{\epsilon}} - \mathbf{I}_K, \tag{3.23}$$

where $\mathbf{L}_{\boldsymbol{\epsilon}}\mathbf{D}_{\boldsymbol{\epsilon}}\mathbf{L}_{\boldsymbol{\epsilon}}^H \triangleq \boldsymbol{\Sigma}_{\mathbf{x}}^{-1} + \mathbf{H}^H(\sigma_w^{-2})\mathbf{H}$, is the LDL decomposition. The residual ISI variance at the output of this equalizer is

$$\sigma^2_{\hat{x}_k|x_k} \triangleq \text{Var}[\hat{x}_k|x_k = x] = (1 - [\mathbf{D}_{\boldsymbol{\epsilon}}]_{k,k}^{-1}\sigma_{x,k}^{-2})^{-2}[\mathbf{D}_{\boldsymbol{\epsilon}}]_{k,k}^{-1}. \tag{3.24}$$

Remark 1

Block DFE structures can be implemented through a noise-predictive architecture, where the feedback filterbank processes $(\mathbf{x}^{\mathbf{d}} - \mathbf{F}^H\mathbf{y})$ *instead of* $\mathbf{x}^{\mathbf{d}}$ *alone. In this case, the block MMSE DFE feedfoward filter is given by* $\mathbf{F}^{\text{bdfe}} = (\mathbf{I} - \mathbf{D}_{\boldsymbol{\epsilon}}^{-1}\boldsymbol{\Sigma}_{\mathbf{x}}^{-1})^{-1}\boldsymbol{\Sigma}^{\text{ble}-1}\mathbf{H}\boldsymbol{\Sigma}_{\mathbf{x}}$*, which establishes a direct link between block DFE and block LE, up to diagonal matrix product, which enables the use of the same inverted matrix for both block LE and DFE, while the feedback filterbank does*

not change. This technique is also applicable to the FIR DFE [Cio08].

3.3.1.2 FIR Equalization

Finite Impulse Response (FIR) equalizers reduce the complexity of block filter-bank techniques by constraining observations to belong to a limited window around the desired symbol to be detected. In order to be able to formalize this approach, the TD ISI channel model in Equation (3.3) is rewritten as a linear model with a fat Toeplitz-like channel matrix

$$\mathbf{y} = \tilde{\mathbf{H}}\tilde{\mathbf{x}} + \mathbf{w}, \tag{3.25}$$

where $\mathbf{w} \sim \mathcal{C}(\mathbf{0}_{N\times 1}, \sigma_w^2\mathbf{I}_N)$, $\tilde{\mathbf{x}} = [\mathbf{0}_{(L-1)\times 1}; \mathbf{x}; \mathbf{0}_{(L-1)\times 1}]$ and $\tilde{\mathbf{H}}$ is a sparse $(K+L-1)\times(K+2L-2)$ matrix given by

$$\tilde{\mathbf{H}} = \begin{bmatrix} h_{1,L} & h_{1,L-1} & \dots & h_{1,1} & & \\ & h_{2,L} & \dots & h_{2,2} & h_{2,1} & \\ & & \ddots & \ddots & \ddots & \ddots \\ & & h_{N,L} & \dots & h_{N,2} & h_{N,1} \\ & & & \ddots & \ddots & \ddots \\ & & & h_{N+L-1,L} & \dots & h_{N+L-1,1} \end{bmatrix}. \tag{3.26}$$

This model exploits SC ZP transmissions by also accounting for the trailing time instants in order to complete the channel impulse response. Then FIR structures can be modelled by windowed processes; applying a sliding window $[-N_p, N_d]$ on the observation vector $\mathbf{y}$, we define $\mathbf{y}_k = [y_{k-N_p}, \dots, y_{k+N_d}]^T$. N_p and N_d are respectively the number of *pre-cursor* and *post-cursor* observed samples, and we denote $N_w \triangleq N_p + N_d + 1$ to refer to the window length. By using the same window on $\mathbf{w}$, and by applying the window $[-N_p', N_d]$ on $\mathbf{x}$, with $N_p' \triangleq N_p + L - 1$, the channel model, for $k = 1, \dots, K$, becomes

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{w}_k, \tag{3.27}$$

with $\mathbf{H}_k \triangleq [\tilde{\mathbf{H}}]_{k-N_p:k+N_d,\, k-N_p':k+N_d}$. These structures are shown in Figure 3.3.

FIR Linear Equalizer The FIR LE's structure consists of a filter $\mathbf{f}_k \in \mathbb{C}^{N_w\times 1}$, such that $\mathbf{f}_k = [f_{k,N_p}, \dots, f_{k,-N_d}]$, applied on $\mathbf{y}_k$ along with an additive bias $g_k \in \mathbb{C}$, with equalized estimates being [Hay14; Cio08]

$$\hat{x}_k \triangleq \mathbf{f}_k^H\mathbf{y}_k + g_k. \tag{3.28}$$

Then similarly to the block case, the unbiased optimal FIR LE is given by

$$\mathbf{f}_k^{\text{le}} = \boldsymbol{\Sigma}_k^{\text{le}-1}\mathbf{h}_k\xi_k^{\text{le}-1}, \tag{3.29}$$

$$g_k^{\text{le}} = \mu_{x,k} - \mathbf{f}_k^{\text{le}H}\mathbf{H}_k\boldsymbol{\mu}_{\mathbf{x},k}, \tag{3.30}$$

where $\mathbf{\Sigma}_k^{\text{le}} \triangleq \sigma_w^2 \mathbf{I}_{N_w} + \mathbf{H}_k \mathbf{\Sigma}_{\mathbf{x},k} \mathbf{H}_k^H$, $\mathbf{h}_k = [\mathbf{H}_k]_{:,k}$, $\boldsymbol{\mu}_{\mathbf{x},k} = \mathbb{E}[\mathbf{x}_k]$ and $\mathbf{\Sigma}_{\mathbf{x},k} = \mathbf{Cov}[\mathbf{x}_k]$, $\xi_k^{\text{le}} \triangleq \mathbf{h}_k^H \mathbf{\Sigma}_k^{\text{le}-1} \mathbf{h}_k$ and, also similarly to the block LE, the variance of residual ISI and noise on the equalized estimates are

$$\sigma^2_{\hat{x}_k|x_k} \triangleq \text{Var}[\hat{x}_k | x_k = x] = 1/\xi_k^{\text{le}} - \sigma^2_{x,k}. \tag{3.31}$$

This FIR structure replaces the inversion of a $N \times N$ matrix, by K inversions of $N_w \times N_w$ matrices, which might be less or more advantageous, depending on the time-selectivity of the channel and the uniformity of statistics of $\mathbf{x}$. In the traditional usage, where $\mu_{x,k} = 0$ and $\sigma^2_{x,k} = \sigma_x^2 = 1$, for $k = 1, \ldots, K$, FIR equalizer is clearly much more computationally efficient, and the complexity of FIR receivers in more general situation will be discussed in more detail, in paragraphs below related to turbo equalization.

Note that, the number of post-cursor samples N_d causes a delay for this equalizer, and its value could be optimized by evaluating the amount of residual ISI and noise for $N_d = 0, \ldots, N_w$, for a fixed value of N_w. A typical rule of thumb is to use the $N_d = N_p + L - 1$ constraint, along with a window size $N_w = (2m-1)(L-1)+1$, where m is a positive integer. This ensures that the window on $\mathbf{x}$ has a size of $2m(L-1)+1$, such that an equal amount of post-cursor and pre-cursor data symbols are processed. Unless specified otherwise, this is how these equalizers are configured in this thesis.

FIR Decision Feedback Equalizer The FIR DFE architecture incorporates a feedforward filter $\mathbf{f}_k \in \mathbb{C}^{N_w \times 1}$ and a feedback filter $\mathbf{g}_k^{\mathbf{d}} \in \mathbb{C}^{N_b \times 1}$, with $\mathbf{g}_k^{\mathbf{d}} = [g^d_{k,N_b}, \ldots, g^d_{k,1}]$, such that equalized estimates are

$$\hat{x}_k = \mathbf{f}_k^H \mathbf{y}_k - \mathbf{g}_k^{\mathbf{d}H} \mathbf{x}_k^{\mathbf{d}} + g_k, \tag{3.32}$$

where $\mathbf{x}_k^{\mathbf{d}} = [x^d_{k-N_b}, \ldots, x^d_{k-1}]$ is the decision vector. For resolving the MMSE criterion, some assumptions are needed on feedback symbols. Traditionally the feedback consists of hard decisions, i.e. $x_k^d = \hat{\mu}_{x,k} \triangleq \arg\min_x \|\hat{x}_k - x\|$. Moreover, *perfect decision* assumption simplifies these equations, by taking $x_k^d = x_k$ during the resolution of the criterion, which also gives $\mathbb{E}[x_k^d] = \mathbb{E}[x_k] = \boldsymbol{\mu}_{\mathbf{x},k}$ and $\text{Var}[x_k^d] = \text{Var}[x_k] = \sigma^2_{x,k}$. Then the unbiased FIR MMSE DFE is

$$\mathbf{f}_k^{\text{dfe}} = \mathbf{\Sigma}_k^{\text{dfe}-1} \mathbf{h}_k \xi_k^{\text{dfe}-1}, \tag{3.33}$$

$$\mathbf{g}_k^{\mathbf{d}\text{-dfe}} = \mathbf{H}_k^{\mathbf{d}H} \mathbf{f}_k^{\text{dfe}}, \tag{3.34}$$

$$g_k^{\text{dfe}} = \mu_{x,k} - \mathbf{f}_k^{\text{dfe}H} (\mathbf{H}_k \boldsymbol{\mu}_{\mathbf{x},k} - \mathbf{H}_k^{\mathbf{d}} \boldsymbol{\mu}_{\mathbf{x}^{\mathbf{d}},k}). \tag{3.35}$$

where $\xi_k^{\text{dfe}} \triangleq \mathbf{h}_k^H \mathbf{\Sigma}_k^{\text{dfe}-1} \mathbf{h}_k$, with

$$\mathbf{\Sigma}_k^{\text{dfe}} \triangleq \sigma_w^2 \mathbf{I}_{N_w} + \mathbf{H}_k \mathbf{Diag}([\sigma^2_{x,k-N_p'}; \ldots; \sigma^2_{x,k-\min(N_p';N_b)}; \mathbf{0}_{N_b \times 1}; \sigma^2_{x,k}; \ldots; \sigma^2_{x,k+N_d}]) \mathbf{H}_k^H,$$

$\mathbf{H}_k^{\mathbf{d}} = \mathbf{H}_k \mathbf{E}^{\mathbf{d}T}$, and $\mathbf{E}^{\mathbf{d}} = [\mathbf{I}_{N_b}, \mathbf{0}_{N_b,N_w}]$. For DFE with the traditional assumptions ($\mu_{x,k} = 0$ and $\sigma^2_{x,k} = \sigma_x^2 = 1$), for $k = 1, \ldots, K$, the covariance matrix $\mathbf{\Sigma}_k^{\text{dfe}}$ is non trivial to invert, due to the non-uniform diagonal caused by the decisions. However, efficient inversion algorithms exist by exploiting Cholesky decomposition of these matrices with Shur's algorithm [AC95a].

Under these assumptions, the variance of post-equalization residual ISI is

$$\sigma^2_{\hat{x}_k|x_k} \triangleq \mathrm{Var}[\hat{x}_k|x_k = x] = 1/\xi_k^{\mathrm{dfe}} - \sigma^2_{x,k}. \tag{3.36}$$

In addition to the delay parameter N_d, DFE has another degree of freedom on the length N_b of the feedback filter. An analysis on the optimal selection of these values [VLC96] shows that when N_b is large enough to cover the delay spread of the channel, the optimal delay value is $N_d = N - 1$. Hence, for channels with moderate L, $N_b = L - 1$ is a practical choice.

Interference Cancellers Interference Cancellation (IC) structures are a different category of data-aided equalizer, which unlike DFE, does not use a feedback that is computed serially, as equalization goes on, but in a parallelized approach, where estimates of the whole data block or the window (for FIR receivers) are available beforehand. It has been originally proposed in [GL81; MS81], as a technique which regenerates interference from data estimates, and completely cancels it. Unlike DFE, which is unable, in general, to reach the MFB with perfect decision feedback [LB03], equalizers with IC are able to completely remove the interference when given perfect data estimates.

In its FIR instantiation, equalized estimates are given by

$$\hat{x}_k = \mathbf{f}_k^H \mathbf{y}_k - \mathbf{g}_k^{\mathbf{d}H} \mathbf{x}_k^{\mathbf{d}}, \tag{3.37}$$

where $\mathbf{x}_k^{\mathbf{d}} = [x^d_{k-N'_p}, \ldots, x^d_{k+N_d}]$ and $\mathbf{g}_k^{\mathbf{d}} = [g^d_{k,N'_p}, \ldots, g^d_{k,N_d}]$ is a filter of size $N'_p + N_d + 1$, with constraint $g^d_{k,0} = 0$, in order to avoid the trivial solution where the desired data x_k is cancelled. Hence FIR IC can be rewritten to outline mitigation of ISI from anti-causal and causal symbols with

$$\hat{x}_k = x_k^d + \mathbf{f}_k^H \mathbf{y}_k - \mathbf{g}_k^{\mathbf{c}H} \mathbf{x}_k^{\mathbf{c}} - \mathbf{g}_k^{\mathbf{a}H} \mathbf{x}_k^{\mathbf{a}}, \tag{3.38}$$

such that $\mathbf{g}_k^{\mathbf{d}} = [\mathbf{g}_k^{\mathbf{c}}; 0; \mathbf{g}_k^{\mathbf{a}}]$ and $\mathbf{x}_k^{\mathbf{d}} = [\mathbf{x}_k^{\mathbf{c}}, x_k^d, \mathbf{x}_k^{\mathbf{a}}]$. In practice, the prior estimates x_k^d for IC can be provided by the output of the demapper of a previous LE stage, and thus it also denoted as LE-IC. However, to simplify the resolution of this system, the classical approach is to assume prior estimates to be perfect [GL81], as in the DFE case, hence $x_k^a = x_k^c = x_k$, and the unbiased MMSE IC equalizer is given by

$$\begin{aligned} \mathbf{f}_k^{\mathrm{ic}} &= (\sigma_w^2 + \mathbf{h}_k^H \boldsymbol{\Sigma}_{\mathbf{x},k} \mathbf{h}_k)^{-1} \mathbf{h}_k \xi_k^{\mathrm{ic}\,-1}, \\ \mathbf{g}_k^{\mathbf{c}\text{-ic}} &= \mathbf{H}_k^{\mathbf{c}H} \mathbf{f}_k^{\mathrm{ic}}, \qquad \mathbf{g}_k^{\mathbf{a}\text{-ic}} = \mathbf{H}_k^{\mathbf{a}H} \mathbf{f}_k^{\mathrm{ic}}, \end{aligned} \tag{3.39}$$

where $\xi_k^{\mathrm{ic}} \triangleq (\sigma_w^2 + \mathbf{h}_k^H \boldsymbol{\Sigma}_{\mathbf{x},k} \mathbf{h}_k)^{-1} \|\mathbf{h}_k\|^2$ and $\mathbf{H}_k^{\mathbf{c}} \triangleq \mathbf{H}_k \mathbf{E}^{\mathbf{c}T}$ and $\mathbf{H}_k^{\mathbf{a}} \triangleq \mathbf{H}_k \mathbf{E}^{\mathbf{a}T}$, with

$$\mathbf{E}^{\mathbf{c}} \triangleq [\mathbf{I}_{N'_p}, \mathbf{0}_{N'_p, N_d+1}], \quad \mathbf{E}^{\mathbf{a}} \triangleq [\mathbf{0}_{N_d+1, N'_p}, \mathbf{I}_{N_d+1}]. \tag{3.40}$$

The variance of the residual ISI at the output of this equalizer is

$$\sigma^2_{\hat{x}_k|x_k} \triangleq \mathrm{Var}[\hat{x}_k|x_k = x] = 1/\xi_k^{\mathrm{ic}} - \sigma^2_{x,k} = \sigma_w^2 / \|\mathbf{h}_k\|^2. \tag{3.41}$$

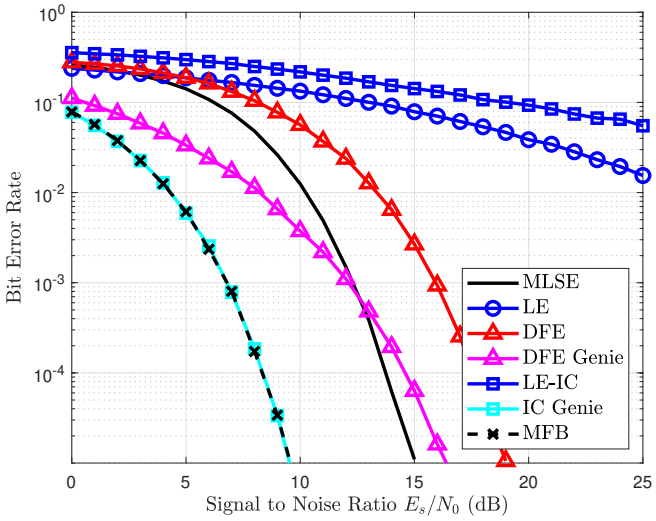


Figure 3.4: Behaviour of FIR equalizers in Proakis C with respect to MLSE and MFB.

This corresponds to the MFB of the ISI channel, but this is only true under the assumption of perfect data estimates at the receiver, which is rarely the case for real-world communications systems. However IC structure plays an important role in the development of turbo equalizers, and it can be used for analysing turbo equalizer performance, as it will be shown below.

Figures 3.3 and 3.2 illustrate the architectures of filter-based equalization techniques described in this section. Separate detection and decoding (see Section 2.2.2) with these filter-based structures have been used as the default approach for communications in ISI channels with affordable receiver costs. However, while the theoretical bounds on the post-equalization SNR of non-linear structures such as the DFE and IC are very attractive, the perfect decisions assumption is only met at high SNR i.e. asymptotically, and the true detection thresholds can be severely degraded in highly selective channels. These aspects have been discussed with more detail in [LB03], and an illustration is reproduced in Figure 3.4 where FIR LE, IC and DFE receivers are used in uncoded system, over the the infamous Proakis C channel [Pro+94] with BPSK, and compared to the MFB and the MLSE, with parameters being taken from [LB03].

To improve upon these performance, the idea of using low complexity joint detection and decoding architectures goes back to the early years of DFE. Indeed, to minimize the issue of error propagation in DFE, while the possibility of including the channel code within the feedback loop is an attractive idea, it is unrealistic, due to the excessive decoding delay needed for each feedback symbol. A solution to this problem is provided in [Eyu88], where the noise-predictive structure of a zero-forcing DFE is exploited, by separating the linear equalization component and the noise-predictive feedback component with an interleaver, and by incorporating the channel decoder in this loop. This is later extended to MMSE DFE [ZPL90], and these techniques were iteratively able to reach the lower bound of DFE performance (i.e. perfect decision feedback). However, DFE with perfect feedback does not reach the MFB, unlike an IC detector, as shown in Figure 3.4.

Table 3.1: Classification of FIR Turbo LE-IC vs. the Usage of Prior Information.

	Linear Structures		
Update Type	TI	IV	TV
References	[GLL97; TK+02]	[Fij+00; Rou00; BC02; TSK02; TH01; Vis+05; LB03]	[TSK02; TK+02; BC02; WP99; DV02; OT04]

Considering the observations above, in the following, we will only consider iterative filter-based receivers based on an Interference Cancellation (IC) structure.

3.3.2 Overview on Turbo Equalization: Linear Structures

Turbo equalization structures have emerged thanks to the concept of probabilistic iterative processing with Soft Input Soft Output (SISO) modules, that was originally used for turbo codes [BGT93], and, in particular, Finite Impulse Response (FIR) filter-based turbo Linear Equalizers (LEs), which have attractive computational complexity trade-off, have been widely investigated from mid-90s up to 2010s [TS11]. There also block turbo equalizers and more general turbo detectors, which will be mentioned briefly in the following, but the focus of this chapter will be on FIR receivers, due to their extensive use in the literature, which makes comparisons easier, and due to their generally lower complexity.

There are numerous variants to the FIR turbo equalizer, and they could be categorized in multiple ways, however, in this thesis, we adopt a categorization based on the nature of the filter updates with prior information, proposed in [Jeo11; LB03]. This nomenclature is attractive as it directly relates to the assumptions involved in the derivation of the equalizer, and it also gives an indication on its computational complexity. Time Invariant (TI) structures update their filters only once at each block reception, using the available channel state information. Iteration Variant (IV) equalizers are updated at each turbo iteration, by additionally using an overall estimate of the prior information on the processed block and its filters remain static for a given iteration. Time Varying (TV) structures update their filters at each symbol within a block, using both symbol-wise prior information and CSI, making them particularly suitable for doubly selective channels, where the impulse response varies over time. The latter approach is also referred to as the *exact* FIR turbo equalizer [TSK02].

FIR Turbo MMSE LE-IC The first filter-based turbo equalizer is derived in [GLL97], through the use of Finite Impulse Response (FIR) filters that use matched-filter based IC [GL81], where soft-symbol estimates, computed by the decoder, are used for regenerating and then removing the ISI. In [LGL01] this TI structure is studied from a frequency-domain infinite-impulse response filter point of view, and in [Fij+00], it is extended to an IV structure, where the filters are optimized with an MMSE criterion through online-estimation of prior soft symbols' statistics. Finally an extension to TV MMSE IC is carried out in [LB03].

An alternative approach to turbo detection is the Wang-Poor's work where a MMSE linear estimator is used with signal statistics computed from the decoder's feedback, and the residual

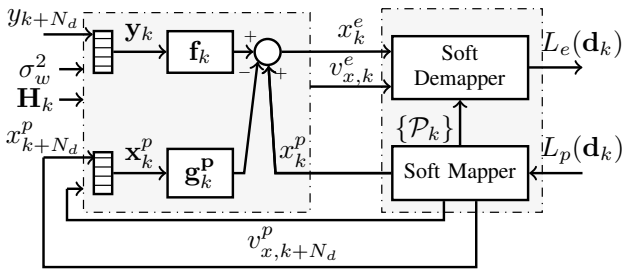


Figure 3.5: Stucture of FIR turbo TV MMSE LE-IC.

estimation error is assumed to be Gaussian distributed [WP99]. This consists in resolving the MMSE criterion with expectations conditioned on the a priori information provided by the decoder, and it has been applied for equalization with FIR filters in [RW01; TSK02] and the resulting TV MMSE LE is further investigated in [TK+02; TS11], which has led to formal derivation of optimal IV and TI LE structures. The equivalence between these approaches and the MMSE IC-based techniques were shown in [LB03], which has led to the denomination "MMSE LE-IC" for these techniques. This FIR structure has been extended to high-order modulations, time-varying channels and to frequency domain structures, and also to MUD for MIMO systems [DV02; TH01; OT04; Vis+05]. The SISO equalizer in the Wang-Poor approach is also known as an "Elementary Symbol Estimator" (ESE), which comes from its equivalence to the jointly Gaussian turbo detection technique in [LLL03]. The classification of the bibliography above, with regards to the filter update schedule is given in Table 3.1.

Following the approach in [WP99; TSK02], the FIR MMSE LE-IC can be obtained from the equations of the classical FIR LE in the previous section, by removing the simplifying assumptions on the signal statistics, and replacing the signal statistics by those indicated by the decoder's extrinsic message, i.e.

$$\mu_{x,k} = x_k^p, \qquad \sigma_{x,k}^2 = v_{x,k}^p, \tag{3.42}$$

where x_k^p and $v_{x,k}^p$ are the mean and the variance of the SISO decoder's extrinsic PMF, defined in Equations (2.39)-(2.40). Then by assuming the output of the equalizer to be Gaussian-distributed, and denoting $x_k^e = \hat{x}_k$ and $v_{x,k}^e = \mathrm{Var}[\hat{x}_k | x_k = x]$, the MAP symbol demapping can be used with Equations (2.53)-(2.54). The resulting turbo equalization algorithm is given in Algorithm 7, for turbo iterations $\tau = 0, \ldots, \mathcal{T}$, and illustrated in Figure 3.5.

This algorithm is the Time Varying (TV) version, and it needs inverting the matrix $\mathbf{\Sigma}_k^{\text{le-ic}(\tau)}$, for each $k = 1, \ldots, K$, and $\tau = 0, \ldots, \mathcal{T}$. Following [TSK02], an Iteration Variant (IV) structure can be obtained, by using the assumption

$$v_{x,1}^{p(\tau)} \approx \cdots \approx v_{x,K}^{p(\tau)} \approx \frac{1}{K} \sum_{k=1}^{K} v_{x,k}^{p(\tau)} \triangleq v_x^{p(\tau)}, \tag{3.43}$$

which reduces equalization complexity significantly, as filters becomes static per iteration (IV), $\mathbf{f}_k^{\text{le-ic}(\tau)} = \mathbf{f}^{\text{le-ic}(\tau)}, \forall k$, if the channel is static. Furthermore, if one uses the Time Invariant (TI)

Algorithm 7 FIR Turbo TV MMSE LE-IC

Input $\mathbf{y}, \mathbf{H}$
Output $\hat{\mathbf{b}}$

$L_p^{(0)}(\mathbf{d}) = \mathbf{0}_{K_c}$
for $\tau = 0 \dots \mathcal{T}$ **do**
  Compute prior PMFs $\mathcal{P}_k^{(\tau)}(\alpha)$ with Eq. (2.39) and $L_p^{(\tau)}(\mathbf{d})$, for $\alpha \in \mathcal{X}, k = 1, \dots, K$.
  Compute $\mathbf{x}^{\mathbf{p}(\tau)}$ and $\mathbf{v}_{\mathbf{x}}^{\mathbf{p}(\tau)}$ with Eq. (2.40).
  for $k = 1 \dots K$ **do**
    $\boldsymbol{\Sigma}_k^{\text{le-ic}(\tau)} \leftarrow \sigma_w^2 \mathbf{I}_{N_w} + \mathbf{H}_k \mathbf{Diag}([v_{x,k-N_p'}^{p(\tau)}; \dots; v_{x,k+N_d}^{p(\tau)}]) \mathbf{H}_k^H$
    $\xi_k^{\text{le-ic}(\tau)} \leftarrow \mathbf{h}_k^H \boldsymbol{\Sigma}_k^{\text{le-ic}(\tau)-1} \mathbf{h}_k$
    $\mathbf{f}_k^{\text{le-ic}(\tau)} \leftarrow \boldsymbol{\Sigma}_k^{\text{le-ic}(\tau)-1} \mathbf{h}_k \xi_k^{\text{le-ic}(\tau)-1}$
    $g_k^{\text{le-ic}(\tau)} \leftarrow x_k^p - \mathbf{f}_k^{\text{le-ic}(\tau)H} \mathbf{H}_k \mathbf{x}_k^{\mathbf{p}}$
    $x_k^e \leftarrow \mathbf{f}_k^{\text{le-ic}(\tau)H} \mathbf{y}_k + g_k^{\text{le-ic}(\tau)}$
    $v_{x,k}^e \leftarrow 1/\xi_k^{\text{le-ic}(\tau)} - v_{x,k}^{p(\tau)}$
  end for
  Compute the posterior PMF $\mathcal{D}_k^{(\tau)}(\alpha)$ with Eq. (2.53), for $\alpha \in \mathcal{X}, k = 1, \dots, K$
  Compute extrinsic LLRs $L_e^{(\tau)}(\mathbf{d})$ with Eq. (2.54)
  Execute de-interleaving, SISO decoding and interleaving to get $L_p^{(\tau+1)}(\mathbf{d})$
end for

approach [TK+02], and needs to keep static filters for all τ, then $v_{x,1}^{p(\tau)} \approx \dots \approx v_{x,K}^{p(\tau)} \approx \sigma_x^2$, and $\mathbf{f}_k^{\text{le-ic}(\tau)} = \mathbf{f}^{\text{le}}, \forall k, \tau$.

Block Turbo MMSE LE-IC In [LL04], an extending-window LE-IC structure is proposed, which is shown to converge faster than the FIR approach (i.e. fixed window) and it asymptotically reaches lower detection error rates, at the cost of increased memory and computational complexity. An alternative formalization of LE-IC is provided in [BC02], in the context of an iterative MUD design, where this method is linked with the use of a MAP criterion on Gaussian-approximated data symbols. This connection is also independently pointed out in [GP08], for BPSK, by establishing links with the GaBP formalism [Loe+07] by deriving a *block-structured* turbo LE-IC.

This framework is later generalized to any constellation for BICM in [GH11], which involves the likelihood over the whole observed block, and thus naturally outperforms both FIR, and extending-window LE-IC, at the cost of increased computational complexity due the observation covariance matrix inversion. Block LE-IC is also provided in [TS11], as an intermediary step for deriving frequency-domain structures (which will be the subject of the next chapter). This turbo receiver can be obtained by the direct application of the generic MMSE-like turbo detector, obtained through GaBP in Section 2.3.5 to the ISI channel model.

Note that block LE-IC can also be implemented in the time-domain as a Kalman smoother [Kay93], which minimizes the computational complexity [QPL05], if a state-space model-based

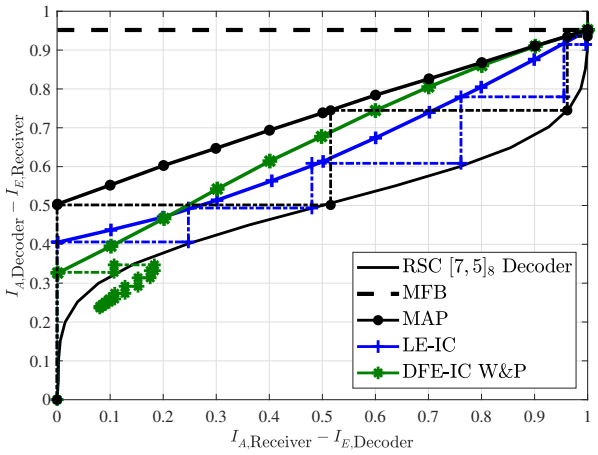


Figure 3.6: Finite-length and asymptotic behaviour mismatch with the DFE-IC [TK+02].

factorization is used (similar to a Wiberg graph) along with the use of GaBP. This is an exact implementation of the block receiver, when used with a forward-backward serial schedule, but with significantly reduced complexity thanks to the sparsity of the channel. It is shown to outperform the sub-optimal BP receiver of [CG05], which underlines the importance of selecting a well-suited factorization for applying message passing algorithms, but at the cost of complexity [CFP11]. Note that alternative parallel, or hybrid schedules can be considered for this Kalman smoother structure, without loss on asymptotically reachable performance, but with slower convergence speed [GP08].

We will not discuss further discuss these structures in this chapter, but they are partially addressed in the next chapter, when discussing Frequency Domain (FD) solutions, and Kalman smoothers are not investigated in detail, in this thesis, due to lack of time.

3.3.3 Limits of Conditional MMSE for Turbo DFE

As DFE is traditionally known as a remarkable improvement over LE , seeking its turbo extensions is hence a topic of interest. However, to point out the difficulty and the complexity of this design problem, the earliest turbo DFE structure of the literature is presented here, with regards to its behaviour versus the turbo LE. Indeed, this FIR DFE-IC is obtained through direct application of the Wang-Poor approach, and other techniques will be discussed afterwards, separately, in Section 3.4.

FIR Turbo MMSE DFE-IC The derivation of a SISO equalizer based on the FIR DFE structure is a complex problem due to the assumptions used in its derivation. Following, the conditional MMSE approach, [TK+02] extended the FIR MMSE DFE to use prior information from the decoder. This results in a DFE-IC structure that uses hard decisions $\hat{\mu}_{x,k}$, as a classical DFE, for mitigating interference caused by causal symbols, and it uses IC with prior

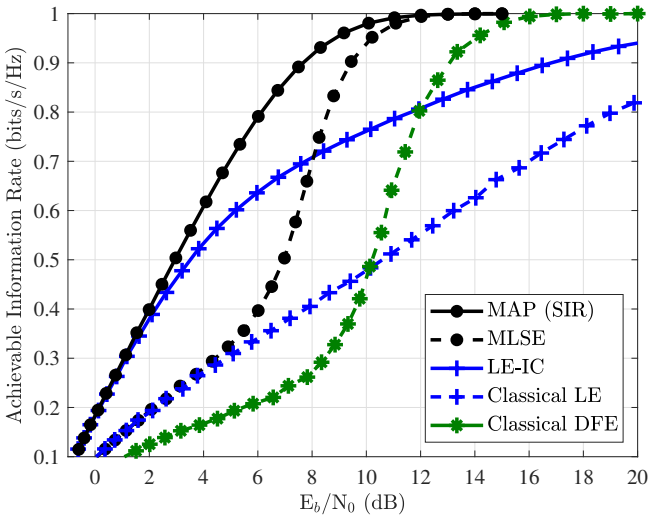


Figure 3.7: Achievable rates of LE-IC and classical LE and DFE in Proakis C with BPSK.

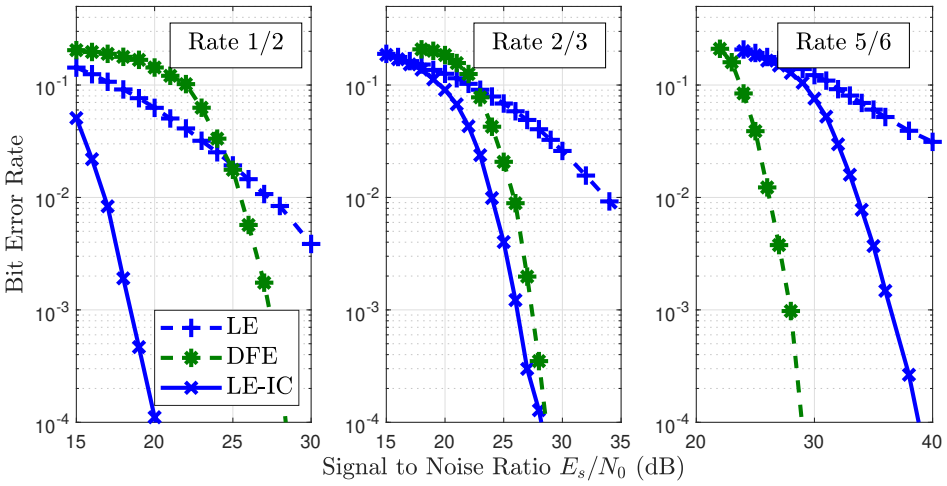


Figure 3.8: Bit error rate performance of LE-IC and DFE in Proakis C with 8-PSK.

information for handling the ISI from anti-causal symbols. In order to derive this structure, Wang-Poor approach is used with perfect decision assumption, by taking $\mathbb{E}[x_k] = \mathbb{E}[x_k^d] = x_k^p$ and $\mathrm{Var}[x_k] = \mathrm{Var}[x_k^d] = v_{x,k}^p$, but while $x_k^d = \hat{\mu}_{x,k}$ being the output of a hard demapper.

Turbo DFE-IC structure is known for its error propagation issues which makes its exact TV implementation even less efficient than TI LE, and its EXIT analysis yields contradictory results [TK+02, Fig. 14]. Indeed, finite-length analysis shows that the MI transfer function predicted obtained with EXIT methodology and the actual MI trajectories do not match. This is reproduced here in Figure 3.6, in the Proakis C channel, with the Recursive Systematic Convolutional (RSC) code of polynomial $[1, 5/7]_8$ with BPSK, and this illustrates that turbo DFE-IC cannot be predicted with EXIT analysis asymptotically.

Another interesting fact appears when one compares the achievable rates of the classical

DFE to the turbo LE-IC, as shown in Figure 3.7. Although LE-IC brings significant improvements over classical filtering and it appears to be capacity achieving at low spectral efficiency operating points (i.e. with strong channel codes, $R_c \ll 1/2$), it falls far behind classical DFE with channel coding, at high spectral efficiency operating points ($R_c \gg 1/2$). Moreover classical DFE falls significantly behind classical LE at low spectral efficiency operating points. These observations can be verified at finite-length simulations, as shown with BER simulations for $K = 256$ symbols, in Proakis C with 8-Phase Shift Keying (PSK), and RSC $[1, 5/7]_8$ punctured optimally up to rates 2/3 and 5/6 [Fre+98].

From these observations, there are various important questions that can arise:

- For non-iterative classical receivers, DFE falls behind LE at low rates. Perfect-decision assumption and/or hard decisions must be behind this drawback, as they incur a loss of information.

- At high spectral efficiency LE-IC falls considerable behind the SIR bound, and even classical DFE outperform it above a certain point. There should be room here for improving achievable rates with filter-based turbo detector, that involves feedback from a demodulation.

To improve the behaviour of DFE-IC, various soft decision feedback techniques were proposed, but their advantages and shortcomings has not been exhaustively investigated in the literature [Jeo11; LB06; Tao16]. In particular, it is unknown if the EXIT functions of these receivers are accurate or not, and their achievable rates are not available. These design issues on FIR DFE-IC will be mainly investigated in the remainder of this chapter.

Block Turbo MMSE DFE-IC The extension of block DFE structure has received very little attention in the literature. Not only the Wang-Poor approach faces the same issues as the FIR DFE-IC, due to perfect decision assumption and hard decisions, but also the structural constraint required by the feedback filterbank is non-trivial to model.

Note however that a block turbo DFE has been proposed in [WZ07; WZ08] where prior information is used, similarly to the Wang-Poor approach, to characterize the statistics of both transmitted symbols and decisions (i.e. perfect decision assumption). This results in a detector with TV filterbanks and a sequence-based symbol detector, which computes soft APP feedback. This receiver has a computational complexity order of $\mathcal{O}(K^4)$, and a reduced complexity IV version, with an order of $\mathcal{O}(K^3)$, is also proposed. Hence, its derivation is similar to that of the FIR DFE-IC of [TK+02], except for the type of decision device (soft APP vs. hard), which is similar to a shortened BCJR. Due to the complexity of this structure remaining too high, it will not be investigated in this thesis, and left for future work.

3.3.4 Discussion and Conclusions

There is a great number of filter-based equalization structures in the literature. These can be classified with respect to their architectures (block filter-banks vs. FIR filters vs. Kalman smoothers), their design criterion (MMSE or least-squares/zero-forcing) and based on whether a serial feedback from the demodulator is used or not (LE vs. DFE).

FIR structures have received particular attention, with regards to their usage as a SISO detector for turbo equalization, and in this chapter we focus on this case of MMSE FIR turbo structures, for easier comparison to the literature. Moreover, with regards to performance issues between the behavior of LE-IC and DFE-IC underlined above, we propose to focus on what has been previously done for DFE-IC with soft feedback, investigate possible improvements, and perform a numerical and analytical analysis on their behavior.

3.4 FIR Turbo MMSE DFE with Soft Feedback

This section focuses on the developments on turbo DFE-IC equalizer, and more specifically those which try to address the aforementioned issues on excessive error propagation and unpredictability. More specifically, we are foremost interested in TV FIR structures, which constitutes the most refined implementation of turbo FIR equalizers, in order to evaluate the performance of turbo DFE with respect to the optimal detector. Such receivers are of interest when doubly-selective channels are involved, such as high frequency (HF) channel [ESN13].

Following a short review of previously existing DFE with soft feedback, this section provides a generic derivation for turbo MMSE FIR DFE and LE structures, with a generalized conditional MSE minimization approach, using the conjugate prior property of Gaussian distributions in Bayesian inference. This enables extending various suboptimal DFE-IC with soft APP feedback into TV structures that operate on any constellation $\mathcal{X}$. Next, considering the flexibility of this approach, and the links between block turbo LE-IC and the GaBP message-passing, an alternative factor graph for the BICM model of Equation (2.27) in Section 2.3.3 is proposed, which enables deriving FIR LE-IC through GaBP message passing. In particular, this connection is exploited by extending the message passing algorithm to use Expectation Propagation (EP), which leads to novel a FIR DFE-IC structure with an original soft feedback. Furthermore, a low-complexity filter-computation approach for TV FIR is provided, which significantly reduces these receivers' computational complexity, by exploiting the Cholesky decomposition of the observation covariance matrix. Finally, the performance of these receivers is analysed theoretically with analytical and numerical methods.

3.4.1 Previous Works on Turbo DFE Structures

DFE proposals mainly differ with the nature of decision feedback, and with the filter updating method. Hence, for clarity, we propose to classify related works in two sub-categories.

Table 3.2: Classification of FIR Turbo DFE-IC vs. the Usage of Prior Information.

	Decision Feedback Structures			
Update Type	Dec. Type	TI	IV	TV
References	Hard	[JM11; JM10]	[JM11; JM10]	[TK+02; JM11; JM10]
	Soft APP	[Bal99]	[Tao16; LB06; LX11; TW+11], **Section 3.6**	**Section 3.4.3**
	Soft EP		**Section 3.6**	**Section 3.4.4**

Iterative Hard DFE-IC In [JM11], the previously discussed DFE-IC structure of [TK+02] is enhanced with a powerful soft demapper that uses the distribution of residual ISI sequences for symbol detection. This modified structure outperforms turbo LE-IC, but the residual ISI distribution is very difficult to derive even in the simple BPSK case. A more practical solution, proposed in [JM10], consists in approximating the residual ISI, at the output of DFE-IC, as an AWG noise, which simplifies the symbol-wise demapper. While this solution challenges TV LE-IC on BPSK, its extension to multilevel modulations has not been explored so far. To the authors' knowledge, until our contributions, this is the only DFE-IC outperforming exact, TV LE-IC in the reference scenario of Proakis-C channel with BPSK symbols. DFE-ICs in [JM10; JM11] were later used as constituent elements for more advanced receivers such as bi-directional DFE, or structures obtained by parallel concatenation of FIRs [JM13].

Iterative Soft DFE-IC Literature on turbo soft DFE-IC is more diverse; although feedback is mostly based on the use of the posterior symbol distribution at the demapper (see Equation (2.53)), there is no common strategy for evaluating its variance [LB06; Tao16; TRK05]. Such an iterative structure is first presented in [Bal99], where various TI DFE with soft feedbacks are evaluated with a perfect decision hypothesis, within a sub-optimal iterative receiver that uses hard decoder feedback. In particular, it is seen that soft demapper feedback mitigates to some extent the error propagation in DFE, despite ignoring decision errors in filter computation. Another notable structure is the IV soft Interference Cancellation (IC) in [LB06]; using both prior and posterior LLRs for filtering, and for IC with BPSK, this scheme significantly outperforms IV LE-IC, but it requires stochastic methods for estimating the correlation properties of posterior LLRs. Several other IV soft feedback structures exist [TRK05; LX11], with alternative heuristics for feedback quality assessment. Structural comparison of IV DFE with posterior feedback is given in [Tao16], extending [LB06] and [LX11] to higher order modulations, but requiring a new heuristic based on a LE-IC pre-equalization step for filter computation. The drawbacks of these approaches are their limitations in usable constellations [LB06; TRK05; LX11], or the sub-optimal heuristics used in filter computation [LB06; LX11; Tao16]. Indeed, IV structures need static statistics of its soft feedback for computing its filters, which naturally requires approximations, as the feedback itself depends on filter coefficients, this issue is addressed in Section 3.6.

Time Varying (TV) soft posterior feedback structures do not have such issues; they can update their filters after each symbol is detected, as it had been done for MIMO receivers in [Cho+08]. For equalization of ISI channels, the structure closest to [Cho+08] is a block-feedback turbo DFE in [TW+11], which updates its filters every P_s symbols. A classification of the references above, and where our contributions will stand, is given in Tables 3.2.

3.4.2 Derivation of FIR MMSE LE and DFE with Conjugate Priors

Below, a generic structure of an unbiased MMSE FIR receiver is given for comparing different structures and their dynamics in the remainder of this chapter. The availability of prior data estimates on $\mathbf{x}$ is assumed, with estimates' values being $\bar{\mathbf{x}}_k^{\text{fir}} \triangleq [\bar{x}_{k-N_p'}^{\text{fir}}, \ldots, \bar{x}_{k+N_d}^{\text{fir}}]$, and variance of estimation noise (uncertainty) being $\bar{\mathbf{v}}_{\mathbf{x},k}^{\text{fir}} \triangleq [\bar{v}_{x,k-N_p'}^{\text{fir}}, \ldots, \bar{v}_{x,k+N_d}^{\text{fir}}]$. Then MMSE FIR receivers which use these prior estimates for IC, provide an output estimate on x_k as x_k^e, with the variance of the residual interference and noise being v_k^e, such that

$$\begin{array}{ll} x_k^e & = \mathbf{f}_k^{\text{fir}H}\mathbf{y}_k + g_k^{\text{fir}} \\ v_{x,k}^e & = 1/\xi_k^{\text{fir}} - \bar{v}_{x,k}^{\text{fir}} \end{array}, \quad \begin{cases} \mathbf{f}_k^{\text{fir}} \triangleq \boldsymbol{\Sigma}_k^{\text{fir}-1}\mathbf{h}_k/\xi_k^{\text{fir}}, \\ g_k^{\text{fir}} \triangleq \bar{x}_k^{\text{fir}} - \mathbf{f}_k^{\text{fir}H}\mathbf{H}_k\bar{\mathbf{x}}_k^{\text{fir}}, \\ \xi_k^{\text{fir}} \triangleq \mathbf{h}_k^H\boldsymbol{\Sigma}_k^{\text{fir}-1}\mathbf{h}_k, \end{cases} \tag{3.44}$$

where $\boldsymbol{\Sigma}_k^{\text{fir}} \triangleq k_w\sigma_w^2\mathbf{I}_N + \mathbf{H}_k\bar{\mathbf{V}}_{\mathbf{x},k}^{\text{fir}}\mathbf{H}_k^H$, $\bar{\mathbf{V}}_{\mathbf{x},k}^{\text{fir}} \triangleq \mathbf{diag}(\bar{\mathbf{v}}_{\mathbf{x},k}^{\text{fir}})$, $\mathbf{h}_k \triangleq \mathbf{H}_k\mathbf{e}_k$, $\mathbf{e}_k = [\mathbf{0}_{N_p',1}; 1; \mathbf{0}_{N_d,1}]$, and $k_w = 1/2$ when signals with one real degree of freedom are used (e.g. $\mathcal{X}$ is BPSK), and otherwise $k_w = 1$ [Cio08]. Note that $\bar{\mathbf{x}}_k^{\text{fir}}$ and $\bar{\mathbf{v}}_{\mathbf{x},k}^{\text{fir}}$ completely characterize such receivers. When $\bar{\mathbf{x}}_{k'}^{\text{fir}}$ and $\bar{\mathbf{v}}_{\mathbf{x},k'}^{\text{fir}}$ are independent of $x_k^e, v_{x,k}^e, \forall k', k$ (within the current turbo iteration), we call this receiver a LE-IC, and when $\bar{x}_{k'}^{\text{fir}}$ and $\bar{v}_{x,k'}^{\text{fir}}$ are dependent on $x_k^e, v_{x,k}^e$, $\forall k' > k$, we refer to it as a DFE-IC.

The proof of this general form resembles to the Wang-Poor approach [WP99], but we make no assumption on the nature, or origin of prior estimates, and show that this approach can be used to model either DFE or LE structures. FIR equalization with IC is derived by minimizing the Bayesian conditional MMSE criterion

$$J = \mathbb{E}_{|p_k^0(\mathbf{x}_k)}[|x_k - x_k^{e\prime}|^2], \tag{3.45}$$

where, $x_k^{e\prime} = \mathbf{f}_k'^H\mathbf{y}_k + g_k'$ is the equalized linear estimate, and the expectation is taken over the observation noise and on non-uniform prior $p_k^0(\mathbf{x}_k)$, a joint multivariate Gaussian prior distribution on $\mathbf{x}_k$ defined with means $\bar{\mathbf{x}}_k^{\mathbf{fir}}$ and variances $\bar{\mathbf{v}}_{\mathbf{x},k}^{\mathbf{fir}}$. The Gaussian assumption on this distribution is needed to exploit the conjugate prior rules on the Gaussian likelihood of observations [Kay93], which is the reason why this criteria is denominated as "Bayesian". Solution to this criterion, is given through inference on Gaussian distributions, by $x_k^{e\prime} = \mathbb{E}_{p_k^0(\mathbf{x}_k)}[x_k|\mathbf{y}_k, \mathbf{H}_k]$, i.e. the expectation over the distribution $p(x_k|\mathbf{y}_k, \mathbf{H}_k)$, the marginal of the joint conjugate Gaussian posterior $p(\mathbf{x}_k|\mathbf{y}_k, \mathbf{H}_k)$, for likelihood $p(\mathbf{y}_k|\mathbf{x}_k, \mathbf{H}_k)$ and prior $p_k^0(\mathbf{x}_k)$, where

$$\begin{aligned} p(\mathbf{x}_k|\mathbf{y}_k, \mathbf{H}_k) &\propto p(\mathbf{y}_k|\mathbf{x}_k, \mathbf{H}_k)p_k^0(\mathbf{x}_k) \\ &\propto \mathcal{CN}(\mathbf{y}_k; \mathbf{H}_k\mathbf{x}_k, \sigma_w^2\mathbf{I}_N)\mathcal{CN}(\mathbf{x}_k; \bar{\mathbf{x}}_k^{\mathbf{fir}}, \bar{\mathbf{V}}_{\mathbf{x},k}^{\text{fir}}) \\ &\propto \mathcal{CN}(\mathbf{x}_k; (\mathbf{H}_k^H\mathbf{H}_k)^{-1}\mathbf{H}_k^H\mathbf{y}_k, \sigma_w^2(\mathbf{H}_k^H\mathbf{H}_k)^{-1})\mathcal{CN}(\mathbf{x}_k; \bar{\mathbf{x}}_k^{\mathbf{fir}}, \bar{\mathbf{V}}_{\mathbf{x},k}^{\text{fir}}) \\ &\propto \mathcal{CN}(\mathbf{x}_k; \boldsymbol{\Gamma}_{\mathbf{x},k}^{\text{fir}}(\sigma_w^{-2}\mathbf{H}_k^H\mathbf{y}_k + \bar{\mathbf{V}}_{\mathbf{x},k}^{\text{fir}\,-1}\bar{\mathbf{x}}_k^{\mathbf{fir}}), \boldsymbol{\Gamma}_{\mathbf{x},k}^{\text{fir}}), \end{aligned} \tag{3.46}$$

where $\bar{\mathbf{V}}^{\text{fir}}_{\mathbf{x},k} = \mathbf{Diag}(\bar{\mathbf{v}}^{\text{fir}}_{\mathbf{x},k})$ and with the use of Gaussian distributions' conjugacy

$$\mathbf{\Gamma}^{\text{fir}}_{\mathbf{x},k} = (\sigma_w^{-2}\mathbf{H}_k^H\mathbf{H}_k + \bar{\mathbf{V}}^{\text{fir}}_{\mathbf{x},k}{}^{-1})^{-1} = (\mathbf{I}_{N_w} - \bar{\mathbf{V}}^{\text{fir}}_{\mathbf{x},k}\mathbf{H}_k^H\mathbf{\Sigma}_k^{\text{fir}-1}\mathbf{H}_k)\bar{\mathbf{V}}^{\text{fir}}_{\mathbf{x},k}, \tag{3.47}$$

where $\mathbf{\Sigma}_k^{\text{fir}} = \sigma_w^2\mathbf{I}_{N_w} + \mathbf{H}_k\bar{\mathbf{V}}^{\text{fir}}_{\mathbf{x},k}\mathbf{H}_k^H$. As $p(\mathbf{x}_k|\mathbf{y}_k,\mathbf{H}_k)$ is a Gaussian PDF, $\mathbb{E}_{p_k^0(\mathbf{x}_k)}[\mathbf{x}_k|\mathbf{y}_k,\mathbf{H}_k] = \mathbf{\Gamma}^{\text{fir}}_{\mathbf{x},k}(\sigma_w^{-2}\mathbf{H}_k^H\mathbf{y}_k + \bar{\mathbf{V}}^{\text{fir}}_{\mathbf{x},k}{}^{-1}\bar{\mathbf{x}}_k^{\mathbf{fir}})$ is the MAP estimate of $\mathbf{x}_k$, and $x_k^{e\prime}$ is then deduced by multiplying the MMSE estimator of $\mathbf{x}_k$ [Kay93], by $\mathbf{e}_k$:

$$\begin{aligned}
x_k^{e\prime} &= \mathbf{e}_k\mathbb{E}_{|p_k^0(\mathbf{x}_k)}[\mathbf{x}_k|\mathbf{y}_k,\mathbf{H}_k] \\
&= \mathbf{e}_k\mathbf{\Gamma}^{\text{fir}}_{\mathbf{x},k}(\sigma_w^{-2}\mathbf{H}_k^H\mathbf{y}_k + \bar{\mathbf{V}}^{\text{fir}}_{\mathbf{x},k}{}^{-1}\bar{\mathbf{x}}_k^{\mathbf{fir}}) \\
&= \mathbf{e}_k(\sigma_w^{-2}\bar{\mathbf{V}}^{\text{fir}}_{\mathbf{x},k}\mathbf{H}_k^H\mathbf{y}_k + \bar{\mathbf{x}}_k^{\mathbf{fir}} - \bar{\mathbf{V}}^{\text{fir}}_{\mathbf{x},k}\mathbf{H}_k^H\mathbf{\Sigma}_k^{\text{fir}-1}\mathbf{H}_k(\sigma_w^{-2}\bar{\mathbf{V}}^{\text{fir}}_{\mathbf{x},k}\mathbf{H}_k^H\mathbf{y}_k + \bar{\mathbf{x}}_k^{\mathbf{fir}})) \\
&= \mathbf{e}_k\bar{\mathbf{x}}_k^{\mathbf{fir}} + \mathbf{e}_k\bar{\mathbf{V}}^{\text{fir}}_{\mathbf{x},k}\mathbf{H}_k^H(\mathbf{y}_k - \mathbf{H}_k\bar{\mathbf{x}}_k^{\mathbf{fir}}) \\
&= \mathbf{f}_k^{\prime H}\mathbf{y}_k + g_k^{\prime}
\end{aligned} \tag{3.48}$$

where the filter coefficients are identified as

$$\mathbf{f}_k^{\prime} = \bar{v}^{\text{fir}}_{x,k}\mathbf{h}_k^H(\mathbf{\Sigma}_k^{\text{fir}})^{-1}, \tag{3.49}$$

$$g_k^{\prime} = \bar{x}_k^{\text{fir}} - \mathbf{f}_k^{\prime H}\mathbf{H}_k\bar{\mathbf{x}}_k^{\text{fir}}, \tag{3.50}$$

with $\mathbf{\Sigma}_k^{\text{fir}} = k_w\sigma_w^2\mathbf{I}_{N_w} + \mathbf{H}_k\bar{\mathbf{V}}^{\text{fir}}_{\mathbf{x},k}\mathbf{H}_k^H$ and $\bar{\mathbf{V}}^{\text{fir}}_{\mathbf{x},k} = \mathbf{Diag}(\bar{\mathbf{v}}^{\text{fir}}_{\mathbf{x},k})$. This receiver is biased, as its MMSE estimators' nature, and the bias terms are given by

$$\mathbb{E}_{|p_k^0(\mathbf{x}_k)}[x_k^{e\prime}|x_k = x] = (1 - \bar{v}^{\text{fir}}_{x,k}\xi_k^{\text{fir}})\bar{x}_k^{\text{fir}} + \bar{v}^{\text{fir}}_{x,k}\xi_k^{\text{fir}}x,$$

with $\xi_k^{\text{fir}} = \mathbf{h}_k^H\mathbf{\Sigma}_k^{\text{fir}-1}\mathbf{h}_k$. Removing additive and multiplicative biases with $x_k^e = (x_k^{e\prime} - (1 - \bar{v}^{\text{fir}}_{x,k}\xi_k^{\text{fir}})\bar{x}_k^{\text{fir}})/(\bar{v}^{\text{fir}}_{x,k}\xi_k^{\text{fir}})$ yields the estimator given in Equation (3.44), which completes the proof.

Let us show how this generalized FIR model can be used to address both LE and DFE structures. Recalling $p_k^0(\mathbf{x}_k) = \mathcal{CN}(\bar{\mathbf{x}}_k^{\text{fir}}, \bar{\mathbf{v}}^{\text{fir}}_{\mathbf{x},k})$, if

- $\bar{\mathbf{x}}_k^{\text{fir}} = \bar{\mathbf{x}}_k^{\text{le}} \triangleq \mathbf{0}_{(N_w+L-1)\times 1}$ and $\bar{\mathbf{v}}^{\text{fir}}_{\mathbf{x},k} = \bar{\mathbf{v}}^{\text{le}}_{\mathbf{x},k} \triangleq \sigma_x^2\mathbf{1}_{(N_w+L-1)\times 1}$, then the resulting structure is the classical MMSE Linear Equalizer (LE),

- $\bar{\mathbf{x}}_k^{\text{fir}} = \bar{\mathbf{x}}_k^{\text{dfe}} \triangleq [\hat{\mu}^d_{x,k-L+1};\ldots;\hat{\mu}^d_{x,k-1};\mathbf{0}_{N_w\times 1}]$ and $\bar{\mathbf{v}}^{\text{fir}}_{\mathbf{x},k} = \bar{\mathbf{v}}^{\text{dfe}}_{\mathbf{x},k} \triangleq [\mathbf{0}_{(L-1)\times 1};\sigma_x^2\mathbf{1}_{N_w\times 1}]$, with $\hat{\mu}^d_{x,k}$ being hard demapper decision, then this structure is the classical MMSE Decision Feedback Equalizer (DFE) with perfect feedback assumption,

- $\bar{\mathbf{x}}_k^{\text{fir}} = \bar{\mathbf{x}}_k^{\text{ic}} \triangleq [x^d_{k-N_p'};\ldots;x^d_{k-1};0;x^d_{k+1};\ldots;x^d_{k+N_d}]$ and $\bar{\mathbf{v}}^{\text{fir}}_{\mathbf{x},k} = \bar{\mathbf{v}}^{\text{ic}}_{\mathbf{x},k} \triangleq [\mathbf{0}_{N_p'\times 1};\sigma^2_{x,k};\mathbf{0}_{N_d\times 1}]$, with x_k^d and $\sigma^2_{x,k}$ being some prior estimates (e.g. decoder feedback), then this structure is the MMSE Interference Cancellation (IC) with perfect feedback assumption. When $x_k^d = x_k^p = \mathbb{E}_{\mathcal{P}_k}[x_k]$ and $\sigma^2_{x,k} = v^p_{x,k} = \text{Var}_{\mathcal{P}_k}[x_k]$, this is the low-complexity turbo matched-filter TI LE-IC structure in [TK+02],

- $\bar{\mathbf{x}}_k^{\text{fir}} = \bar{\mathbf{x}}_k^{\text{le-ic}} \triangleq [x^p_{k-N_p'};\ldots;x^p_{k+N_d}]$ and $\bar{\mathbf{v}}^{\text{fir}}_{\mathbf{x},k} = \bar{\mathbf{v}}^{\text{le-ic}}_{\mathbf{x},k} \triangleq [v^p_{x,k-N_p'};\ldots;v^p_{x,k+N_d}]$, with $x_k^p = \mathbb{E}_{\mathcal{P}_k}[x_k]$ and $v^p_{x,k} = \text{Var}_{\mathcal{P}_k}[x_k]$, then this is the turbo TV MMSE LE-IC [TSK02],

- $\bar{\mathbf{x}}_k^{\text{fir}} = \bar{\mathbf{x}}_k^{\text{le-ic}} = [x_{k-N_p'}^p; \ldots; x_{k+N_d}^p]$ and $\bar{\mathbf{v}}_{\mathbf{x},k}^{\text{fir}} = v_x^p \mathbf{1}_{(N_w+L-1)}$, with $x_k^p = \mathbb{E}_{\mathcal{P}_k}[x_k]$ and $v_x^p = K^{-1}\sum_{k=1}^{K} \text{Var}_{\mathcal{P}_k}[x_k]$, then this structure is the turbo IV MMSE LE-IC [TSK02] (called LC - low complexity, therein),

- $\bar{\mathbf{x}}_k^{\text{fir}} = \bar{\mathbf{x}}_k^{\text{le-ic}} = [x_{k-N_p'}^p; \ldots; x_{k+N_d}^p]$ and $\bar{\mathbf{v}}_{\mathbf{x},k}^{\text{fir}} = \bar{\mathbf{v}}_{\mathbf{x},k}^{\text{le}} = \sigma_x^2 \mathbf{1}_{(N_w+L-1)\times 1}$, with $x_k^p = \mathbb{E}_{\mathcal{P}_k}[x_k]$, then this is the turbo TI MMSE LE-IC [TK+02] (called NA - no a priori, therein),

- $\bar{\mathbf{x}}_k^{\text{fir}} = [\hat{\mu}_{x,k-N_p'}^d; \ldots; \hat{\mu}_{x,k-1}^d; x_k^p; \ldots; x_{k+N_d}^p]$ and $\bar{\mathbf{v}}_{\mathbf{x},k}^{\text{fir}} = [\mathbf{0}_{N_p'\times 1}; v_{x,k}^p; \ldots; v_{x,k+N_d}^p]$, with $x_k^p = \mathbb{E}_{\mathcal{P}_k}[x_k]$, $v_{x,k}^p = \text{Var}_{\mathcal{P}_k}[x_k]$ and $\hat{\mu}_{x,k}^d$ being hard demapper decision, then this is the turbo TV MMSE DFE-IC with perfect feedback assumption in [TK+02],

- $\bar{\mathbf{x}}_k^{\text{fir}} = [\mu_{x,k-N_p'}^d; \ldots; \mu_{x,k-1}^d; x_k^p; \ldots; x_{k+N_d}^p]$ and $\bar{\mathbf{v}}_{\mathbf{x},k}^{\text{fir}} = [\gamma_x^d \mathbf{1}_{N_p'\times 1}; v_x^p \mathbf{1}_{N_D\times 1}]$, with $x_k^p = \mathbb{E}_{\mathcal{P}_k}[x_k]$, $v_x^p = K^{-1}\sum_{k=1}^{K} \text{Var}_{\mathcal{P}_k}[x_k]$, $\mu_{x,k}^d = \mathbb{E}_{\mathcal{D}_k}[x_k]$, with $\mathcal{D}_k$ being the posterior PMF in Equation (2.97), and γ_x^d is an pre-estimation of $K^{-1}\sum_{k=1}^{K} \text{Var}_{\mathcal{D}_k}[x_k]$, making this structure a turbo IV MMSE DFE-IC. Indeed, this value γ_x^d cannot be exactly computed, without already having carried out the detection over the data block, hence [LB06; LX11; Tao16] use different heuristics to estimate it.

The generality of this model remains limited, for instance the TV DFE-IC structures of [JM10; JM11] cannot be directly compared with the one proposed by [TK+02], as it uses an alternate approach for computing the variance of the post-equalization residual ISI $v_{x,k}^e$. However, it still remains a powerful way of addressing FIR turbo receivers, as it allows clear comparison of their underlying differences with few parameters. In particular, in the next subsection, it will allow us to generalize various IV or hard feedback DFE-IC schemes, into various TV soft feedback structures, in merely few paragraphs.

3.4.3 Time-Varying Turbo DFE based on APP Estimates

In the literature review on TV DFE-IC structures, we can note the lack of structures with *soft feedback*, and as hard decisions appear to be degrading the receiver performance with respect to LE-IC [TK+02], we are interested in extending these schemes to soft feedback. Moreover [Bal99] showed improvements with soft posterior feedback on non-turbo, invariant filters, and IV DFE-IC structures with soft APP feedback [LB06; Tao16] outperform IV LE-IC, but do not meet the performance of TV LE-IC for BPSK transmissions in the Proakis C channel. Hence, in this subsection, DFE-IC schemes from the literature [TK+02; JM10; Tao16] are extended to Time Varying (TV) structures that use soft APP feedback, and their characteristics are assessed with analytical, asymptotic and finite-length analysis.

References on TV DFE-IC with soft feedback are limited, and to our knowledge, soft posterior (APP) feedback was the only imperfect decision feedback, with a reasonable complexity in the literature, applicable to any constellation [Bal99; LB06]. Nevertheless, it is not possible to derive a structure using such feedback within the classical Wang-Poor formalism, but here its usage is justified with the generalized FIR model provided in Equation (3.44).

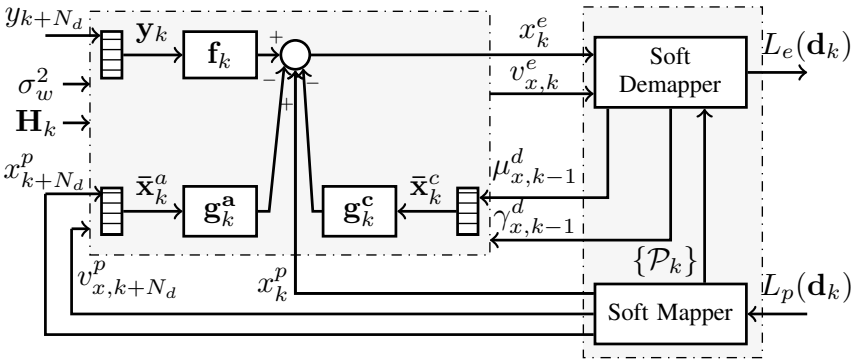


Figure 3.9: Structure of FIR DFE-IC with soft APP feedback.

One can consider the equalization problem within a Bayesian inference framework, as in the Chapter 2, where a particular realization of a random data symbol is estimated through noisy observations. Then the Wang-Poor framework [BC02] can be generalized, with the use of *any* Gaussian-distributed prior $p^0_k(\mathbf{x}_k)$, on data symbols. Hence in continuation of the list, at the end of Section 3.4.2, one could consider the *sequential* MMSE estimator, which uses its previously estimated posteriors to improve the estimations of future data (Section 12.6 in [Kay93]). This can be achieved by using the extrinsic output of the SISO decoder as priors on anti-causal symbols and the desired symbol, and A Posteriori Probability (APP) estimates at the equalizer's output, as priors on causal symbols, i.e.

$$p^{\text{seq}}_k(\mathbf{x}_k) \propto \prod_{l=k}^{k+N_d} \mathcal{CN}(x^p_l, v^p_{x,l}) \prod_{l=k-N'_p}^{k-1} \mathcal{CN}(\mu^d_l, \gamma^d_{x,l}), \tag{3.51}$$

where $x^p_k = \mathbb{E}_{\mathcal{P}_k}[x_k]$, $v^p_{x,k} = \text{Var}_{\mathcal{P}_k}[x_k]$, $\mu^d_{x,k} = \mathbb{E}_{\mathcal{D}_k}[x_k]$, with $\mathcal{D}_k$ being the posterior PMF in Equation (2.97), and $\gamma^d_{x,k} = \text{Var}_{\mathcal{D}_k}[x_k]$. Note that this concept of refining estimates with posteriors, is also the core idea behind the concept of PDA, discussed in Section 2.4.3, but instead of using it in a block-by-block schedule, here it is applied sequentially symbol-wise. In the following, turbo DFEs with soft feedback are derived, by using this sequential estimator with the model of Equation (3.44), with different approaches on the selection of $\bar{\mathbf{v}}^{\mathbf{fir}}_{\mathbf{x},k}$.

Exact TV DFE-IC with APP Feedback This equalizer is obtained by the direct usage of the sequential APP priors, and it is a generalization of invariant schemes originally proposed in [LB06; Tao16] to TV structures. It is derived by using the joint posterior $p^{\text{seq}}_k(\mathbf{x}_k)$ with the model (3.44). The resulting exact TV APP FIR structure is fully defined by

$$\begin{aligned} \bar{\mathbf{x}}^{\text{dfe-app}}_k &= [\mu^d_{k-N'_p}, \ldots, \mu^d_{k-1}, v^p_k, \ldots, x^p_{k+N_d}]^T, \\ \bar{\mathbf{v}}^{\text{dfe-app}}_{\mathbf{x},k} &= [\gamma^d_{x,k-N'_p}, \ldots, \gamma^d_{x,k-1}, v^p_{x,k}, \ldots, v^p_{x,k+N_d}]^T. \end{aligned} \tag{3.52}$$

This structure will be referred to as *DFE-IC APP* in the remainder of this chapter, and it can also be rewritten in the three-filter Interference Cancellation (IC) form of Equations (3.38) and (3.39), with $x^e_k = x^p_k + \mathbf{f}^{\text{dfe-app}\,H}_k \mathbf{y}_k - \mathbf{g}^{\mathbf{c}\text{-dfe-app}\,H}_k \boldsymbol{\mu}^{\mathbf{d}}_k - \mathbf{g}^{\mathbf{a}\text{-dfe-app}\,H}_k \mathbf{x}^{\mathbf{p}}_k$ (where the sizes of

feedback vectors are implied), and

$$\mathbf{f}_k^{\text{dfe-app}} = (\sigma_w^2 + \mathbf{h}_k^H \boldsymbol{\Sigma}_{\mathbf{x},k}^{\text{dfe-app}} \mathbf{h}_k)^{-1} \mathbf{h}_k \xi_k^{\text{dfe-app}-1},$$
$$\mathbf{g}_k^{\mathbf{c}\text{-dfe-app}} = \mathbf{H}_k^{\mathbf{c}H} \mathbf{f}_k^{\text{dfe-app}}, \quad \mathbf{g}_k^{\mathbf{a}\text{-dfe-app}} = \mathbf{H}_k^{\mathbf{a}H} \mathbf{f}_k^{\text{dfe-app}},$$

where $\xi_k^{\text{dfe-app}}$ and $\boldsymbol{\Sigma}_{\mathbf{x},k}^{\text{dfe-app}}$ are given as in Equation (3.44), and it is illustrated in Figure 3.9. Optimal filter parameters for this case appear to be the same as for FIR LE-IC, when N_p' and N_d have the same value by fixing $N_w = (2m-1)(L-1)+1$, for positive integer m.

TV DFE-IC with Perfect APP Feedback Here we propose to generalize [TK+02; Bal99] to APP feedback, with perfect decision hypothesis, which imposes covariances associated to the APP feedback to 0, and focuses the forward filter to the mitigation of anti-causal symbol interference. This case named *DFE-IC PAPP*, differs from the *DFE-IC APP* with the variance estimates:

$$\begin{aligned} \bar{\mathbf{x}}_k^{\text{dfe-papp}} &= \bar{\mathbf{x}}_k^{\text{dfe-app}}, \\ \bar{\mathbf{v}}_{\mathbf{x},k}^{\text{dfe-papp}} &= [\mathbf{0}_{N_p'}^T, v_{x,k}^p, \ldots, v_{x,k+N_d}^p]^T. \end{aligned} \tag{3.53}$$

In this case, similar to the classical DFE, optimal filtering performance appears to be given by $N_p = 0$, which yields $N_p' = L-1$ and $N_d = N_w$.

Hybrid TV DFE-IC with APP Feedback This is an extension of the hard decision TV structure in [JM10] to APP feedback. In [JM10], the DFE-IC with perfect hard decisions from [TK+02] is improved by adding an estimate of the amount of decision errors to the equalizer's output variance $v_{x,k}^e$. This quantity is given by $\text{Var}_{\mathcal{D}_k}[\mathbf{g}_k^{cH}([\mathbf{x} - \boldsymbol{\mu}_{\mathbf{x}}^{\mathbf{d}}]_{k-N_p':k-1})]$, using Equation (2.97). Moreover, this structure checks whether this increase in variance, causes sign changes in extrinsic LLRs, and sets ambiguous LLRs to zero. This receiver is extended to use APP soft feedback, instead of hard decisions, and it is denoted *DFE-IC HAPP*.

3.4.4 Time-Varying Turbo DFE with Expectation Propagation

In Section 3.4.2, we have shown that the rules of conjugacy of Gaussian distributions can be exploited to carry out inference on the windowed observation model in Equation (3.25) for deriving FIR SISO equalization structures. Hence, in the following, we propose an alternative factorization for $p(\mathbf{x}, \mathbf{d})$ to enable deriving FIR receivers with message passing algorithms.

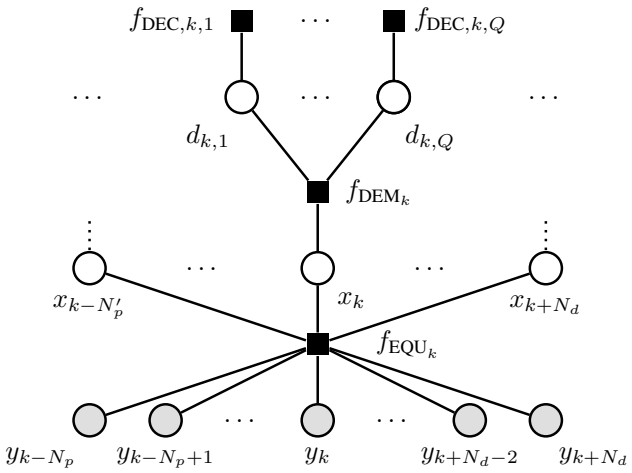


Figure 3.10: Factor graph for the posterior PDF (3.55) on x_k and $\mathbf{d}_k$.

Factor Graph Modelling for FIR Turbo Equalization

Let us recall the factorization of generic detection problem from the Section 2.3.5 in the previous chapter

$$p(\mathbf{x},\mathbf{d}|\mathbf{y}) = \underbrace{p(\mathbf{y}|\mathbf{x})}_{=f_{\text{EQU}}(\mathbf{x})} \prod_{k=1}^{K} \underbrace{p(x_k|\mathbf{d}_k)}_{=f_{\text{DEM}_k}(x_k,\mathbf{d}_k)} \prod_{q=1}^{Q} \underbrace{p(d_{k,q})}_{=f_{\text{DEC}_{k,q}}(d_{k,q})} , \tag{3.54}$$

where GaBP was used to derive a SISO detector based on an MMSE filter-bank. When applied to the specific case of the ISI channel in Equation 3.3, the resulting structure is a block turbo MMSE LE-IC [GH11; TS11].

In order to modify this factorization to design FIR structures, we rely on the sparsity of the channel, which allows us to assume that elements of the vector $\mathbf{y}$ which will impact x_k are likely to be localized in a sub-vector of $\mathbf{y}$, around the index k and with a length that depends on the channel spread L. With this line of thought, the windowed observation model in Equation (3.25) can be used to perform inference, considering that we have

$$p\left(\bar{\mathbf{d}}_k, \mathbf{x}_k|\mathbf{y}_k\right) \propto p(\mathbf{y}_k|\mathbf{x}_k) \prod_{k'=k-N'_p}^{k+N_d} p(x_{k'}|\mathbf{d}_{k'}) \prod_{j=0}^{q-1} p(d_{k',j}), \tag{3.55}$$

where $\bar{\mathbf{d}}_k = \mathbf{d}_{k-N_p-L+1:k+N_d}$ and $p\left(\bar{\mathbf{d}}_k, \mathbf{x}_k|\mathbf{y}_k\right) \approx p\left(\bar{\mathbf{d}}_k, \mathbf{x}_k|\mathbf{y}\right)$ is assumed.

Note that working with $p\left(\bar{\mathbf{d}}_k, \mathbf{x}_k|\mathbf{y}\right) \approx p\left(\bar{\mathbf{d}}_k, \mathbf{x}_k|\mathbf{y}_k\right)$ is not the only option for estimating x_k. Indeed x_k can be estimated through inference on any $\mathbf{x}_{k'}$, with $k' = k - N_d, \ldots, k + N'_p$, i.e. on any window containing x_k, but by selecting $\mathbf{x}_k$ as the default data window, this option is indirectly translated to the choice of window parameters, which is a common aspect of FIR equalizers, as discussed above.

By denoting

$$f_{\mathrm{EQU}_k}(\mathbf{x}_k) \triangleq p(\mathbf{y}_k|\mathbf{x}_k) \propto \exp\left(-\frac{\mathbf{x}_k^H \mathbf{H}_k^H \mathbf{H}_k \mathbf{x}_k}{\sigma_w^2} + 2\mathcal{R}\left(\frac{\mathbf{x}_k^H \mathbf{H}_k^H \mathbf{y}_k}{\sigma_w^2}\right)\right), \tag{3.56}$$

where the dependence on $\mathbf{y}_k$ is omitted, as observations are fixed during the message-passing procedure, the posterior (3.55) yields the factor graph shown in Fig. 3.10. Note that, when considering the factor node $f_{\mathrm{EQU}_k}(\mathbf{x}_k)$, only variable node x_k is to be updated through a message passing algorithm, while $x_{k-N_p'}, \ldots, x_{k-1}, x_{k+1}, \ldots, x_{k+N_d}$ provide prior information.

EP and Serial Scheduling for a Novel Decision Feedback Equalizer

The use of GaBP, as in Section 2.3.5, on this new factor graph yields the FIR LE-IC SISO equalizer, when variable nodes x_k are updated either in parallel or serial schedule. However, as EP's extrinsic messages from the factor node DEM_k to $\mathrm{EQU}_{k'}$, for $k' = k - N_d, \ldots, k + N_p'$, are directly dependent on the extrinsic message from EQU_k to DEM_k, selecting a serial or a parallel structure significantly changes the receiver behavior, which is a feature that appears when loopy graphs are under consideration.

Hence following a similar approach to the *diagonal EP* in Section 2.4.2, for nodes involving data symbol variables, we have

$$q_{\mathrm{DEM}_k}(x_k) \propto \mathcal{CN}(x_k; \mu_{x,k}^d, \gamma_{x,k}^d), \tag{3.57}$$
$$m_{\mathrm{DEM}_k \to x_k}(x_k) \propto \mathcal{CN}(x_k; x_k^d, v_{x,k}^d), \tag{3.58}$$
$$q_{\mathrm{EQU}_k}(x_k) \propto \mathcal{CN}(x_k; \mu_{x,k}^e, \gamma_{x,k}^e), \tag{3.59}$$
$$m_{\mathrm{EQU}_k \to x_k}(x_k) \propto \mathcal{CN}(x_k; x_k^e, v_{x,k}^e), \tag{3.60}$$

which ensures that the family of symbol variables are de-correlated, and the joint distribution on $\mathbf{x}_k$ has a diagonal covariance matrix.

The computation of messages from DEM_k to x_k is identical to the computations in Section 2.4.2, and following the "Gaussian division" in the computation of $m_{\mathrm{DEM}\to x}(x_k)$ from (2.100), its mean and variance are denoted

$$x_k^* = \frac{\mu_{x,k}^d v_{x,k}^e - x_k^e \gamma_{x,k}^d}{v_{x,k}^e - \gamma_{x,k}^d}, \text{ and, } v_{x,k}^* = \frac{v_{x,k}^e \gamma_{x,k}^d}{v_{x,k}^e - \gamma_{x,k}^d}. \tag{3.61}$$

These quantities are the major novelty in using EP for FIR receivers, instead of GaBP, because of their direct dependence on $x_{x,k}^e$ and $v_{x,k}^e$. Moreover, note that we were reluctant to use x_k^d and $v_{x,k}^d$ to denote the statistics of the "Gaussian division" involved for computing $m_{\mathrm{DEM}\to x}(x_k)$, because due to the potential occurrences of contradicting messages between EQU_k's outputs and the mapping constraints, there could be estimates where $v_{x,k}^e \leqslant \gamma_{x,k}^d$. The occurrence of negative variances is a classical aspect of EP-based algorithms which arises from the use of projections towards restrictive family of distributions. To avoid their destructive effect, [C+14] suggests replacing the concerned $(x_k^d, v_{x,k}^d)$ with priors from the decoder

$(x_k^p, v_{x,k}^p)$, and [Sen+11] suggests replacing with posteriors $(\mu_{x,k}^d, \gamma_{x,k}^d)$. From experimentations not exposed here, we have found the latter case to be more advantageous asymptotically, and thus we use

$$(x_k^d, v_{x,k}^d) = \begin{cases} (x_k^*, v_{x,k}^*) & \text{if } v_{x,k}^e > \gamma_{x,k}^d + \varepsilon, \\ (\mu_{x,k}^d, \gamma_{x,k}^d) & \text{otherwise,} \end{cases} \tag{3.62}$$

where ε is a small constant.

At the factor node EQU_k, similar to EQU in Section 2.4.2, the pre-projection posterior is obtained by using Equation (2.85), with

$$q_{\text{EQU}}(x_k) = \int_{\mathbf{x}_k^{\setminus k}} f_{\text{EQU}}(\mathbf{x}_k) \prod_{k'=k-N_p'}^{k+N_d} m_{x \to \text{EQU}}(x_{k'}) \mathbf{dx}_k^{\setminus k}. \tag{3.63}$$

The integrand of the equation above is a multivariate Gaussian distribution $\mathcal{CN}(\boldsymbol{\mu}^{\mathbf{e}}, \boldsymbol{\Gamma}^{\mathbf{e}})$ (whose diagonal elements follow Equation (3.59), due to the moment-matching aspect of EP), hence, using eq. (3.56), its covariance and mean should satisfy

$$\begin{aligned} \boldsymbol{\Gamma}_{\mathbf{x},k}^{\mathbf{e}} &= ({\mathbf{V}_{\mathbf{x},k}^{\mathbf{d}}}^{-1} + \sigma_w^{-2} \mathbf{H}_k^H \mathbf{H}_k)^{-1}, \\ \boldsymbol{\mu}_{\mathbf{x},k}^{\mathbf{e}} &= \boldsymbol{\Gamma}_{\mathbf{x},k}^{\mathbf{e}} ({\mathbf{V}_{\mathbf{x},k}^{\mathbf{d}}}^{-1} \mathbf{x}_k^{\mathbf{d}} + \sigma_w^{-2} \mathbf{H}_k^H \mathbf{y}_k), \end{aligned} \tag{3.64}$$

where $\mathbf{V}_{\mathbf{x},k}^{\mathbf{d}} = \mathbf{Diag}(\mathbf{v}_{\mathbf{x},k}^{\mathbf{d}})$, with $\mathbf{v}_{\mathbf{x},k}^{\mathbf{d}} = [v_{x,k-N_p'}^d, \ldots, v_{x,k+N_d}^d]$, and $\mathbf{x}_k^{\mathbf{d}} = [x_{k-N_p'}^d, \ldots, x_{k+N_d}^d]$. Using some matrix algebra, and Woodbury's identity on $\boldsymbol{\Gamma}_{\mathbf{x},k}^{\mathbf{e}}$ [GVL96], the mean $\mu_{x,k}^e$ and the variance $\gamma_{x,k}^e$ of the marginalized PDF $q_{\text{EQU}_k}(x_k)$, in Equation (3.59), are given by

$$\begin{aligned} \gamma_{x,k}^e &= \mathbf{e}_k^H \boldsymbol{\Gamma}_{\mathbf{x},k}^{\mathbf{e}} \mathbf{e}_k = v_{x,k}^d (1 - v_{x,k}^d \mathbf{h}_k^H {\boldsymbol{\Sigma}_k^{\mathrm{d}}}^{-1} \mathbf{h}_k), \\ \mu_{x,k}^e &= \mathbf{e}_k^H \boldsymbol{\mu}_{\mathbf{x},k}^{\mathbf{e}} = x_k^d + v_k^d \mathbf{h}_k^H {\boldsymbol{\Sigma}_k^{\mathrm{d}}}^{-1} (\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k^{\mathbf{d}}), \end{aligned} \tag{3.65}$$

with $\boldsymbol{\Sigma}_k^{\mathbf{d}} \triangleq k_w \sigma_w^2 \mathbf{I}_N + \mathbf{H}_k \mathbf{V}_k^{\mathbf{d}} \mathbf{H}_k^H$. Message to the demapper, in Equation (3.60), is then computed with the Gaussian density division in Equation (2.84) with

$$v_{x,k}^e = \frac{\gamma_{x,k}^e v_{x,k}^d}{v_{x,k}^d - \gamma_{x,k}^e}, \text{ and, } x_k^e = \frac{v_{x,k}^d \mu_k^e - \gamma_{x,k}^e x_k^d}{v_{x,k}^d - \gamma_{x,k}^e}. \tag{3.66}$$

Developing these equations, one obtains a FIR receiver expression as in Equation (3.44), with $\bar{\mathbf{x}}_k^{\mathbf{d}} \triangleq [x_{k-N_p'}^d, \ldots, x_{k+N_d}^d]$ and $\bar{\mathbf{v}}_{\mathbf{x},k}^{\mathbf{d}} \triangleq [v_{x,k-N_p'}^d, \ldots, v_{x,k+N_d}^d]$ for IC.

Note that in the absence of any information from the EQU_k factor node, i.e. when $m_{x_k \to \text{DEM}}(x_k) \propto 1$, which is equivalent to having $v_{x,k}^e = +\infty$, and $x_k^e = 0$ (or any other placeholder finite constant), we have $\mathcal{D}_k = \mathcal{P}_k$ and consequently $x_k^* = \mu_{x,k}^d = x_k^p$ and $v_{x,k}^* = \gamma_{x,k}^d = v_{x,k}^p$. This is the conventional prior information in GaBP, or the Wang-Poor approach, hence to innovate with EP, a serial update schedule is needed, in order of exploit messages $m_{x \to \text{DEM}}(x_k)$ on causal symbols. Thus, variables x_k have to be updated sequentially from both factor nodes DEM_k and then EQU_k, and this yields the proposed DFE-IC EP, given in Algorithm 8.

Algorithm 8 Proposed *DFE-IC EP* receiver.

Input $\mathbf{y}, \mathbf{H}$
Output $\hat{\mathbf{b}}$

$L_p^{(0)}(\mathbf{d}) \leftarrow \mathbf{0}_{K_c}, v_{x,k}^{d(0)} \leftarrow \sigma_x^2, v_{x,k}^{p(0)} \leftarrow \sigma_x^2, k = 1, \ldots, K.$
for $\tau = 0 \ldots \mathcal{T}$ **do**
  Compute prior PMFs $\mathcal{P}_k^{(\tau)}(\alpha)$ with Eq. (2.39) and $L_p^{(\tau)}(\mathbf{d})$, for $\alpha \in \mathcal{X}, k = 1, \ldots, K$.
  Compute $\mathbf{x}^{\mathbf{p}(\tau)}$ and $\mathbf{v}_{\mathbf{x}}^{\mathbf{p}(\tau)}$ with Eq. (2.40).
  for $k = 1 \ldots K$ **do**
    $\boldsymbol{\Sigma}_k^{\text{dfe-ep}(\tau)} \leftarrow \sigma_w^2 \mathbf{I}_{N_w} + \mathbf{H}_k \mathbf{Diag}([v_{x,k-N_p'}^{d(\tau)}; \ldots; v_{x,k-1}^{d(\tau)}; v_{x,k}^{p(\tau)}; \ldots; v_{x,k+N_d}^{p(\tau)}]) \mathbf{H}_k^H$
    $\xi_k^{\text{dfe-ep}(\tau)} \leftarrow \mathbf{h}_k^H \boldsymbol{\Sigma}_k^{\text{dfe-ep}(\tau)-1} \mathbf{h}_k$
    $\mathbf{f}_k^{\text{dfe-ep}(\tau)} \leftarrow \boldsymbol{\Sigma}_k^{\text{dfe-ep}(\tau)-1} \mathbf{h}_k \xi_k^{\text{dfe-ep}(\tau)-1}$
    $\mathbf{g}_k^{\mathbf{c}\text{-dfe-ep}(\tau)} \leftarrow \mathbf{H}_k^{\mathbf{c}H} \mathbf{f}_k^{\text{dfe-ep}(\tau)}, \quad \mathbf{g}_k^{\mathbf{a}\text{-dfe-ep}(\tau)} \leftarrow \mathbf{H}_k^{\mathbf{a}H} \mathbf{f}_k^{\text{dfe-ep}(\tau)}$
    $x_k^{e(\tau)} \leftarrow x_k^{p(\tau)} + \mathbf{f}_k^{\text{dfe-ep}(\tau)H} \mathbf{y}_k - \mathbf{g}_k^{\mathbf{c}\text{-dfe-ep}(\tau)H} \mathbf{x}_k^{\mathbf{d}(\tau)} - \mathbf{g}_k^{\mathbf{a}\text{-dfe-ep}(\tau)H} \mathbf{x}_k^{\mathbf{p}(\tau)}$
    $v_{x,k}^{e(\tau)} \leftarrow 1/\xi_k^{\text{dfe-ep}(\tau)} - v_{x,k}^{p(\tau)}$
    Compute the posterior PMF $\mathcal{D}_k^{(\tau)}(\alpha)$ with Eq. (2.53), for $\alpha \in \mathcal{X}, k = 1, \ldots, K$
    Generate EP feedback $(x_k^{d(\tau)}, v_{x,k}^{d(\tau)})$ with (3.61)-(3.62).
  end for
  Compute extrinsic LLRs $L_e^{(\tau)}(\mathbf{d})$ with Eq. (2.54)
  Execute de-interleaving, SISO decoding and interleaving to get $L_p^{(\tau+1)}(\mathbf{d})$
end for

To clarify the dynamics of the proposed receiver, $\tau = 0, \ldots, \mathcal{T}$ denotes turbo iterations, i.e. exchanges between the DEM and DEC factor nodes. Exchanges between EQU and DEM factor nodes *sequentially* update the whole block $\mathbf{x}$ as in Algorithm 8, and this acts as a FIR receiver which uses the following means and variances for IC

$$\begin{aligned} \bar{\mathbf{x}}_k^{\text{dfe-ep}} &\triangleq [x_{k-N_p'}^d, \ldots, x_{k-1}^d, x_k^p, \ldots, x_{k+N_d}^p]^T, \\ \bar{\mathbf{v}}_{\mathbf{x},k}^{\text{dfe-ep}} &\triangleq [v_{x,k-N_p'}^d, \ldots, v_{x,k-1}^d, v_{x,k}^p, \ldots, v_{x,k+N_d}^p]^T, \end{aligned} \tag{3.67}$$

for $k = 0, \ldots, K-1$. This layout shows that this structure indeed follows a TV DFE-IC evolution, with decoder's extrinsic messages being used for compensating the ISI caused by anti-causal symbols, and the EP feedback from the demapper handles interference generated by causal symbols. This DFE-IC structure, using a novel kind of soft feedback, is unlike any hard or soft APP feedback previously used in the literature [TK+02; JM10; JM11; LB06; Tao16; TRK05; Bal99; LX11]. The demapper's extrinsic symbol-wise feedback is obtained by jointly using the prior information of the SISO detector, and the past equalized symbol estimates x_k^e (see (2.98), (3.61) and (3.62)). As for DFE-IC APP, this EP-based receiver an also be rewritten in the three-filter Interference Cancellation (IC) form of Equations (3.38) and (3.39), with $x_k^e = x_k^p + \mathbf{f}_k^{\text{dfe-ep}H} \mathbf{y}_k - \mathbf{g}_k^{\mathbf{c}\text{-dfe-ep}H} \mathbf{x}_k^{\mathbf{d}} - \mathbf{g}_k^{\mathbf{a}\text{-dfe-ep}H} \mathbf{x}_k^{\mathbf{p}}$, and

$$\begin{aligned} \mathbf{f}_k^{\text{dfe-ep}} &= (\sigma_w^2 + \mathbf{h}_k^H \boldsymbol{\Sigma}_{\mathbf{x},k}^{\text{dfe-ep}} \mathbf{h}_k)^{-1} \mathbf{h}_k \xi_k^{\text{dfe-ep}-1}, \\ \mathbf{g}_k^{\mathbf{c}\text{-dfe-ep}} &= \mathbf{H}_k^{\mathbf{c}H} \mathbf{f}_k^{\text{dfe-ep}}, \quad \mathbf{g}_k^{\mathbf{a}\text{-dfe-ep}} = \mathbf{H}_k^{\mathbf{a}H} \mathbf{f}_k^{\text{dfe-ep}}, \end{aligned}$$

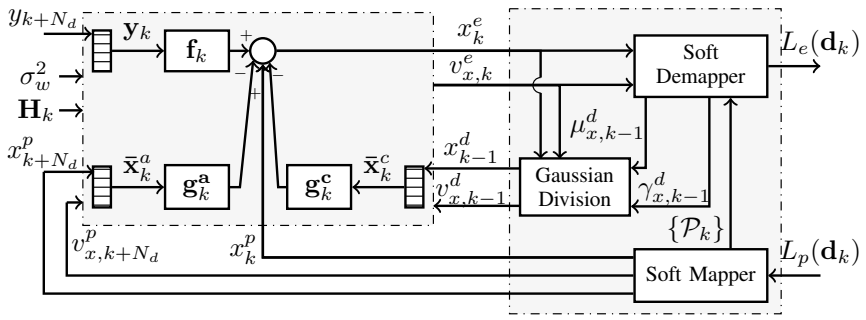


Figure 3.11: Structure of FIR DFE-IC with soft EP feedback.

Algorithm 9 Recursive inverse algorithm for LE-IC [TSK02]

Input $\boldsymbol{\Sigma}_{k-1}^{-1}, \sigma_w^2, \bar{\mathbf{V}}_{\mathbf{x},k}, \mathbf{H}_k$
Output $\boldsymbol{\Sigma}_k^{-1}$

1: {*Separate previous inverse into sub-matrices*}
2: $\begin{bmatrix} \mathbf{U} & \mathbf{u} \\ \mathbf{u}^H & u \end{bmatrix} \leftarrow \boldsymbol{\Sigma}_{k-1}^{-1}$
3: {*Matrix manipulation for removing the previous element and adding the new one*}
4: $\mathbf{U} \leftarrow \mathbf{U} - \mathbf{u}\mathbf{u}^H/u$
5: $\begin{bmatrix} \mathbf{u} \\ u \end{bmatrix} \leftarrow \begin{bmatrix} \mathbf{0}_{N_w-1} \\ \sigma_w^2 \end{bmatrix} + \mathbf{H}_k \bar{\mathbf{V}}_{\mathbf{x},k} \mathbf{H}_k^H \begin{bmatrix} \mathbf{0}_{N_w-1} \\ 1 \end{bmatrix}$
6: $\mathbf{u}' \leftarrow \mathbf{U}\mathbf{u}$
7: $u \leftarrow 1/(u - \mathbf{u}^H\mathbf{u}')$
8: $\mathbf{u} \leftarrow -u\mathbf{u}'$
9: $\mathbf{U} \leftarrow \mathbf{U} + u\mathbf{u}'\mathbf{u}'^H$
10: {*Reassemble sub-matrices into the updated inversed matrix*}
11: $\boldsymbol{\Sigma}_k^{-1} = \begin{bmatrix} u & \mathbf{u}^H \\ \mathbf{u} & \mathbf{U} \end{bmatrix}$.

where $\xi_k^{\text{dfe-ep}}$ and $\boldsymbol{\Sigma}_{\mathbf{x},k}^{\text{dfe-ep}}$ are given as in Equation (3.44), and it is illustrated in Figure 3.11.

3.4.5 Reducing the Complexity of Exact FIR Equalizers

TV FIR turbo equalizers, as in Algorithm 7, or 8 have excessive computational complexity due to the symbol-wise filter updates. Before going further into their analysis, in this section we investigate on a low complexity implementation. Symbol-wise filter updates require recursive matrix inversion methods to more efficiently carry out the computation of the filter $\mathbf{f}_{k+1}$. This is done by exploiting the previous filter $\mathbf{f}_k$, considering the filters are of the form $\mathbf{f}_k = \boldsymbol{\Sigma}_k^{-1}\mathbf{h}_k$, where $\boldsymbol{\Sigma}_k$ is the conditional covariance of windowed observations.

Shortcomings of Existing Approaches

In [TSK02], *Tüchler et al.* propose for turbo FIR MMSE LE-IC, a recursive matrix inversion algorithm, based on common sub-matrices between successive inverses, similar to Levinson's algorithm [GVL96]. This procedure requires the computation of an initial inverse (Gauss-Jordan inversion) with a complexity order[5] of $4N_w^3/3$, but further recursions' complexity is $2N_w^2$. Detailed description of the update procedure is given in Algorithm 9.

A shortcoming of the approach above, is then need to carry out the non-sparse product $\mathbf{\Sigma}_k^{-1}\mathbf{h}_k$ for each symbol and iteration. On the other hand, practical matrix inversion implementations avoid direct inversion by solving the system $\mathbf{\Sigma}_k\mathbf{f}_k = \mathbf{h}_k$ for $\mathbf{f}_k$ with triangular factorizations [SF+11] and using forward/backward substitutions. This approach is even more advantageous in equalization where the system is sparse. However, the matrix decomposition can be as expensive, if it needs to be carried out for each symbol.

In this section a novel recursive inversion strategies are proposed for LE-IC and DFE-IC, based on an initial Cholesky decomposition, and followed by sparse rank-1 updates/downdates of the Cholesky factors of following filter computations. Unlike [SF+11], our algorithm is able to deal with channel matrices evolving in time, making it more efficient for turbo TV FIR. For LE-IC the complexity order is of N_w^2, hence, asymptotically around 50% less complex than [TSK02].

Proposed Cholesky Update Algorithm for FIR MMSE LE-IC

We consider a LE-IC with priors variances $\bar{v}_{x,k}$, let $\mathbf{L}_{k-1}$ be the lower triangular Cholesky decomposition of the covariance matrix $\mathbf{\Sigma}_{k-1}$, i.e. $\mathbf{L}_{k-1}\mathbf{L}_{k-1}^H = \mathbf{\Sigma}_{k-1}$. Algorithm 10 uses $\mathbf{L}_{k-1}$ and the latest values $\bar{v}_{x,k+N_d}$ and $\mathbf{e}_{k+N_d}^H\mathbf{H}_k$ (new row with $\mathbf{h}[k+N_d]$) appended to the sliding window, to compute $\mathbf{L}_k$. Impact of latest generated value is appended to the decomposition using $\mathbf{l}_{\mathbf{12}}^H$ and l_{22}, then past data is removed. The resulting updated Cholesky decomposition is a rank-1 update [GVL96] of $\mathbf{L}_{\mathbf{22}}$, defined within algorithm 10. These steps, followed by forward/backward substitutions $\mathbf{f}_k = \mathbf{L}_k^{-H}\mathbf{L}_k^{-1}\mathbf{h}_k$, allow fpr low complexity filter computation.

Proposed Cholesky Update Algorithm for FIR MMSE DFE-IC

In the case of DFE-IC, the diagonal of the covariance matrix $\bar{\mathbf{V}}_{\mathbf{x}}^{\text{tdfe}}$ is composed of two independently sliding parts: one for causal symbols $\bar{v}_{x,k}^c$, between symbols $k-N_p'$ and $k-1$, the other for anti-causal $\bar{v}_{x,k}^a$, between symbols k and $k+N_d$. The LE-IC update procedure above handles the addition of $\bar{v}_{x,k+N_d}^a$ and the removal of $\bar{v}_{x,k-N_p'-1}^c$, but the change in $(k-1)^{\text{th}}$ symbol remains to be updated.

[5] "Order" means asymptotic expansion as $N_w \to +\infty$, assuming $N_w \propto 3L$, i.e. sliding window operating on $4L$ symbols.

Algorithm 10 Cholesky update algorithm for LE-IC.

Input $\mathbf{L}_{k-1}, \sigma_w^2, \bar{v}_{x,k+N_d}, \mathbf{H}_{k-1}, \mathbf{H}_k, \bar{\mathbf{V}}_{\mathbf{x},k-1}$
Output $\mathbf{L}_k$

$\{$*Add a row and a column*$\}$
$[\mathbf{h_1}_k, h_{2k}] \leftarrow [0, \mathbf{e}_{k+N_d}^H \mathbf{H}_k]$
$\mathbf{w} \leftarrow \mathbf{H}_{k-1} \bar{\mathbf{V}}_{\mathbf{x},k-1} \mathbf{h_1}_k$
$\mathbf{l_{12}} \leftarrow \mathbf{L}_{k-1}^{-1} \mathbf{w}$
$l_{22} \leftarrow \sqrt{\mathbf{h_1}_k^H \bar{\mathbf{V}}_{\mathbf{x},k-1} \mathbf{h_1}_k + \bar{v}_{x,k+N_d} |h_{2k}|^2 - \mathbf{l_{12}}^H \mathbf{l_{12}} + \sigma_w^2}$
$\{$*Build augmented matrix and remove row & column*$\}$
$\begin{bmatrix} \times & \mathbf{0}_{1,N} \\ \mathbf{l_{21}} & \mathbf{L_{22}} \end{bmatrix} \leftarrow \begin{bmatrix} \mathbf{L}_{k-1} & \mathbf{0}_{N,1} \\ \mathbf{l_{12}}^H & l_{22} \end{bmatrix}$
$\{$*Rank-1 update* $\mathbf{L}_k \mathbf{L}_k^H = \mathbf{L_{22}} \mathbf{L_{22}}^H + \mathbf{l_{21}} \mathbf{l_{21}}^H\}$
for $l = 1$ to N **do**
  $r \leftarrow \sqrt{[\mathbf{L_{22}}]_{l,l}^2 + |[\mathbf{l}_{21}]_l|^2}$, $c \leftarrow \frac{[\mathbf{L_{22}}]_{l,l}}{r}$, $s \leftarrow \frac{[\mathbf{l}_{21}]_l^*}{r}$
  $[\mathbf{L_{22}}]_{l:N,l} \leftarrow c[\mathbf{L_{22}}]_{l:N,l} + s[\mathbf{l}_{21}]_{l:N}$
  $[\mathbf{l}_{21}]_{l:N} \leftarrow c[\mathbf{l}_{21}]_{l:N} - s^*[\mathbf{L_{22}}]_{l:N,l}$
end for
$\mathbf{L}_k \leftarrow \mathbf{L_{22}}$

Algorithm 11 Cholesky update algorithm for DFE-IC.

Input $\tilde{\mathbf{L}}_k, \bar{v}_{x,k-1}^a, \bar{v}_{x,k-1}^c, [\mathbf{H}_k]_{:,-1}$
Output $\mathbf{L}_k$

$\mathbf{w} \leftarrow \sqrt{|\bar{v}_{x,k-1}^a - \bar{v}_{x,k-1}^c|}[\mathbf{H}_k]_{:,-1}$
for $l = N_p$ to N **do**
  if $\bar{v}_{x,k-1}^c < \bar{v}_{x,k-1}^a$ **then**
    $\{$*Rank-1 downdate* $\mathbf{L}_k \mathbf{L}_k^H = \tilde{\mathbf{L}}_k \tilde{\mathbf{L}}_k^H - \mathbf{w}\mathbf{w}^H\}$
    $r \leftarrow \sqrt{[\tilde{\mathbf{L}}_k]_{l,l}^2 - |[\mathbf{w}]_l|^2}$, $c \leftarrow \frac{[\tilde{\mathbf{L}}_k]_{l,l}}{r}$, $s \leftarrow \frac{[\mathbf{w}]_l^*}{r}$
    $[\tilde{\mathbf{L}}_k]_{l:N,l} \leftarrow c[\tilde{\mathbf{L}}_k]_{l:N,l} - s[\mathbf{w}]_{l:N}$
  else if $\bar{v}_{x,k-1}^c > \bar{v}_{x,k-1}^a$ **then**
    $\{$*Rank-1 update* $\mathbf{L}_k \mathbf{L}_k^H = \tilde{\mathbf{L}}_k \tilde{\mathbf{L}}_k^H + \mathbf{w}\mathbf{w}^H\}$
    $r \leftarrow \sqrt{[\tilde{\mathbf{L}}_k]_{l,l}^2 + |[\mathbf{w}]_l|^2}$, $c \leftarrow \frac{[\tilde{\mathbf{L}}_k]_{l,l}}{r}$, $s \leftarrow \frac{[\mathbf{w}]_l^*}{r}$
    $[\tilde{\mathbf{L}}_k]_{l:N,l} \leftarrow c[\tilde{\mathbf{L}}_k]_{l:N,l} + s[\mathbf{w}]_{l:N}$
  end if
  $[\mathbf{w}]_{l:N} \leftarrow c[\mathbf{w}]_{l:N} - s^*[\tilde{\mathbf{L}}_k]_{l:N,l}$
end for
$\mathbf{L}_k \leftarrow \tilde{\mathbf{L}}_k$

Algorithm 11 gives this update procedure for DFE-IC, by applying either a rank-1 update or downdate on $\tilde{\mathbf{L}}_k$, the Cholesky factor who has already been updated by algorithm 10, depending on the sign of $\bar{v}_{x,k-1}^c - \bar{v}_{x,k-1}^a$. Such updates are carried out using Givens plane rotations [GVL96].

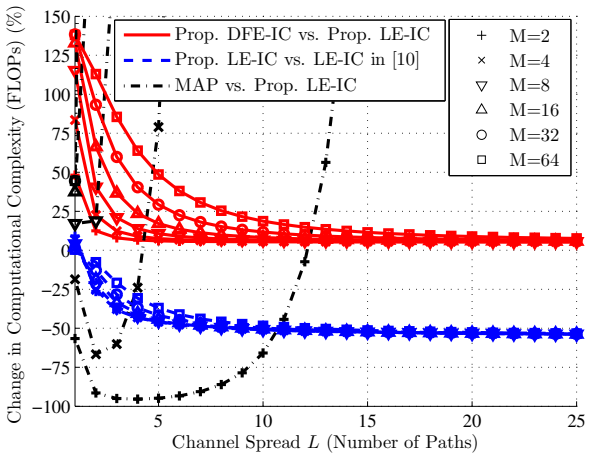


Figure 3.12: Complexity of LE-IC and DFE-IC with the proposed matrix inversion algorithm.

Theoretical Computational Complexity Analysis The computational complexity of the proposed algorithm is evaluated with the number of required multiply and accumulate units, estimated by the number of real additions and multiplications, amounting to half a Floating Point Operations (FLOPs) each.

FLOP counting ratios between different FIR implementations are plotted in Figure 3.12, depending on the channel spread, with a block length $K = 2048$ and a FIR window given by $N_w = 3L + 2$, $N_d = 2L$. The blue dashed curves show the FLOP count ratio of a LE-IC using our strategy relative to using the algorithm in [TSK02], for different constellation orders. Up to 50% saving is observed as channel spread increases.

DFE-IC FLOP count is compared to LE-IC, both using the proposed inversion strategies, with red solid lines. This ratio is high for a low number of channel taps, but decreases to 7% as L increases, more or less quickly depending on the modulation order M. Finally, MAP detector is seen to be an interesting alternative to FIR receivers for BPSK/QPSK signalling, in channels with very short channel spreads.

3.4.6 Asymptotic and Finite-length Analysis of FIR Turbo Equalizers

Following the proposed technique to implement FIR turbo receivers with lower complexity, we will now perform a performance comparison of TV DFE-IC with soft feedback to the previously established TV FIR techniques in the literature.

Analytic Comparison of DFE-IC vs. LE-IC

This paragraph semi-analytically assesses the behaviour of a DFE-IC relative to a LE-IC to underline the interest in jointly using decision feedback and a generic prior information for IC as in Equation (3.44).

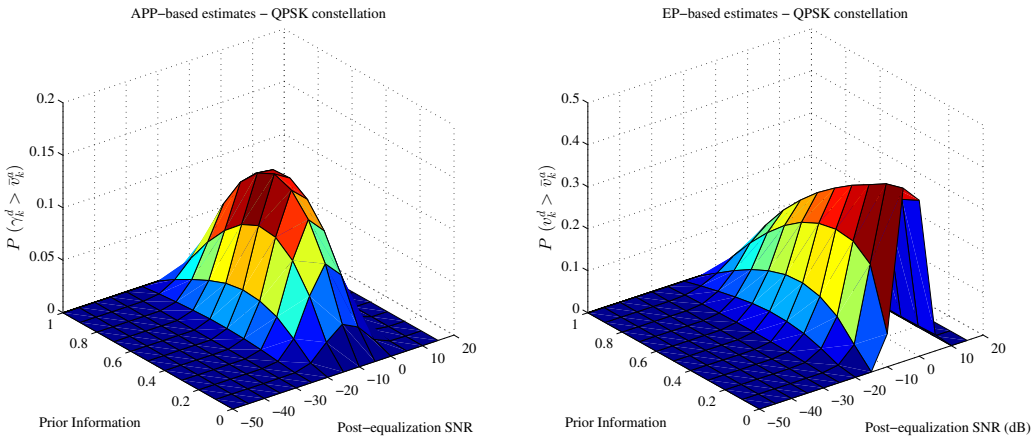


Figure 3.13: $\mathbb{P}[\bar{v}^c_{x,k} > \bar{v}_{x,k}] <$ for DFE-IC APP and EP with random channels and QPSK.

In fact, we investigate whether LE-IC operating with priors $(\bar{x}_k, \bar{v}_{x,k})$ provides a lower bound for the achievable information rate of a DFE-IC structure, using the same prior information for its anti-causal symbols $(\bar{x}^{\mathrm{a}}_k, \bar{v}^{\mathrm{a}}_{x,k}) = (\bar{x}_k, \bar{v}_{x,k})$, alongside decision feedback estimates $(\bar{x}^{\mathrm{c}}_k, \bar{v}^{\mathrm{c}}_{x,k})$. The mean post-equalization SNR of a FIR receiver is given by $\mathrm{SNR}^{\mathrm{fir}}_{\mathrm{out}} \triangleq \sigma_x^2/\mathbb{E}[v^{e(\mathrm{fir})}_{x,k}]$, where the priors $\mathbf{v}^{\mathrm{fir}}_{\mathbf{x},k} = \bar{\mathbf{v}} = [\bar{v}^a_{x,k-N'_p}; \ldots; \bar{v}^a_{x,k-1}; \bar{v}^c_{x,k}; \ldots; \bar{v}^c_{x,k+N_d}]$ are used as the basis of the FIR equalizer. By exploiting the structural similarities between DFE-IC and LE-IC, the causal feedback's impact is characterized by a ratio of mean post-equalization SNRs[6]

$$G \triangleq \frac{\mathrm{SNR}^{\mathrm{dfe}}_{\mathrm{out}}}{\mathrm{SNR}^{\mathrm{le}}_{\mathrm{out}}} = \frac{\sigma_x^2}{\mathbb{E}[v^{e(\mathrm{dfe})}_{x,k}]} \frac{\mathbb{E}[v^{e(\mathrm{le})}_{x,k}]}{\sigma_x^2} = \frac{\xi^{\mathrm{dfe}}}{\xi^{\mathrm{le}}} \frac{1-\bar{v}_x\xi^{\mathrm{le}}}{1-\bar{v}_x\xi^{\mathrm{dfe}}} \tag{3.68}$$

where $\bar{v}_x = \mathbb{E}[\bar{v}_{x,k}]$ and $\xi^{\mathrm{XX}} = \mathbb{E}[\xi^{\mathrm{XX}}_k]$, where XX stands for "le" or "dfe". This gain is greater than unity if and only if $\xi^{\mathrm{dfe}} \geqslant \xi^{\mathrm{le}}$, or equivalently if and only if $\mathbb{E}[\bar{\mathbf{V}}^{\mathbf{le}}_{\mathbf{x},k} - \bar{\mathbf{V}}^{\mathbf{dfe}}_{\mathbf{x},k}]$ is positive semi-definite. Hence having $\bar{v}_x > \bar{v}^c_x = \mathbb{E}[\bar{v}^c_{x,k}]$ is required for achieving improvements with a DFE-IC. Based on empirical and experimental evaluations, the conjecture $\mathbb{P}[\bar{v}^c_{x,k} > \bar{v}_{x,k}] < 0.5$ has been verified over a wide range of channels, SNRs, and for different constellations, for $\bar{v}^c_{x,k} = v^d_{x,k}$ (DFE-IC EP) and for $\bar{v}^c_{x,k} = \gamma^d_{x,k}$ (DFE-IC APP) (see Figure 3.13). This ensures $\bar{v}_x > \bar{v}^c_x$ asymptotically and thus, LE-IC output SNR is a lower bound on DFE-IC EP/APP.

G is plotted in Fig. 3.14, with $N_w = 17, N_d = 10$ and $\sigma_x^2 = 1$ for the static Proakis C channel, $\mathbf{h} = [1,2,3,2,1]/\sqrt{19}$; when decisions are more reliable than priors, G increases, otherwise DFE-IC brings small improvements. When $\bar{v}^a_x \to 1$, there is no prior information, and decisions bring a significant gain. Oppositely, when $\bar{v}^a_x \to 0$, prior information is already close to the ideal, and DFE-IC cannot improve further. This indicates boosted performance at initial turbo-iterations, especially at low SNR.

[6] $\mathrm{SNR}^{\mathrm{XX}}_{\mathrm{out}} = \sigma_x^2/\mathbb{E}[v^{e(\mathrm{XX})}_{x,k}]$ is the post-equalization SNR, where XX is "dfe" or "le", (see Equation (3.44) for $v^e_{x,k}$). Superscript "le" refers to the use of $(\bar{x}_k, \bar{v}_{x,k})$ for IC, and "dfe" refers to the use of $(\bar{x}^{\mathrm{a}}_k, \bar{v}^{\mathrm{a}}_{x,k})$ and $(\bar{x}^{\mathrm{c}}_k, \bar{v}^{\mathrm{c}}_{x,k})$ for IC, as in Equations (3.38) and (3.39).

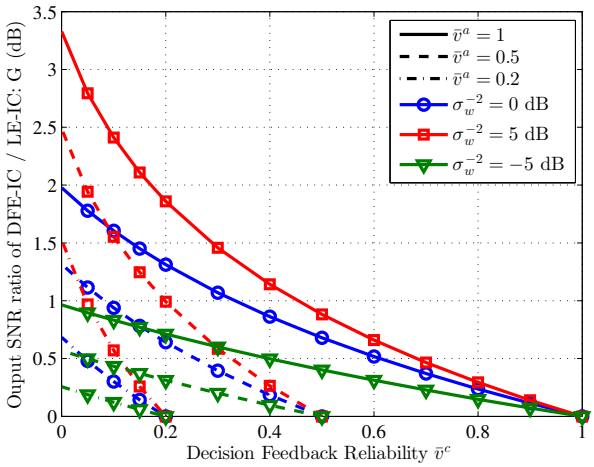
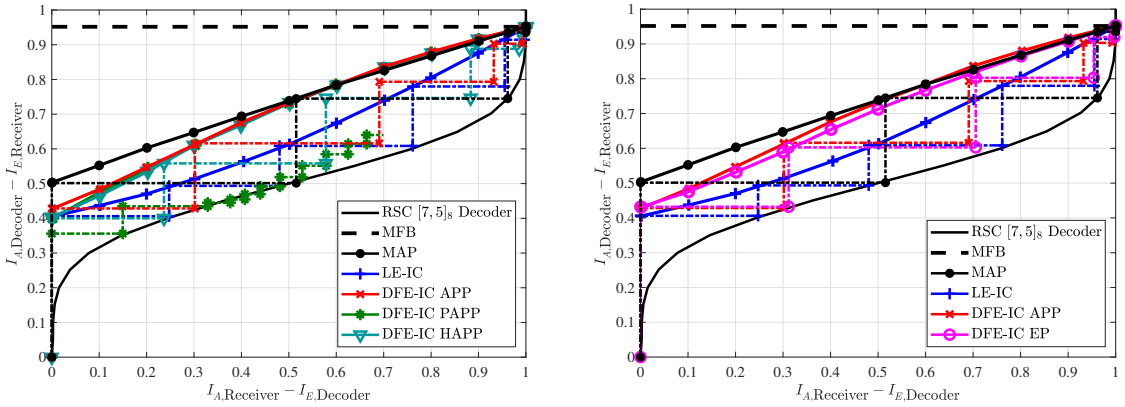


Figure 3.14: Post-equalization SNR ratio G depending on channel noise precision σ_w^{-2}, prior reliability $\bar{v}_x^a$ and "decision" reliability $\bar{v}_x^c$.



(a) Comparison of APP-based DFE-IC (b) Comparison with EP-based DFE-IC

Figure 3.15: EXIT curves and average MI trajectories of FIR equalizers with BPSK in Proakis C channel at $E_b/N_0 = 7$dB.

Asymptotic Analysis and Performance Prediction

To assess the full potential of DFE-IC, asymptotic analysis is used to evaluate its achievable rates. Extrinsic Information Transfer (EXIT) analysis [TB00], summarized in Section 2.3.6, is used to characterize the asymptotic limits of this receiver.

In Figure 3.15, transfer curves $\mathcal{T}_R$ are plotted in solid lines for considered receivers along with the reverse transfer $\mathcal{T}_D^{-1}$ of the BCJR decoder of a Recursive Systematic Convolutional (RSC) code. DFE-IC APP yields a higher I_E than LE-IC for all I_A, unsurprisingly given the posterior feedback, and there is little difference with DFE-IC EP, which has slightly lower rates at low prior information. In particular, the improvement at $I_A = 0$ lets us conjecture a

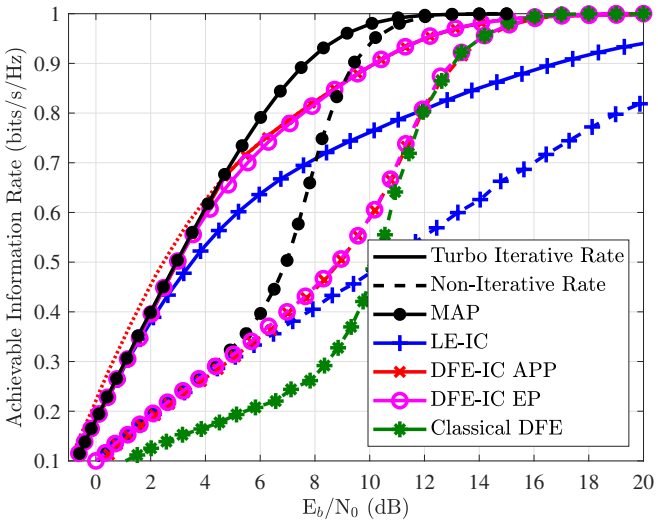


Figure 3.16: Achievable spectral efficiency on deterministic Proakis C channel with BPSK.

lower waterfall threshold in BPSK, and the higher slope of the $\mathcal{T}_R$ curve at low I_A hints an improved convergence speed for DFE-IC across turbo iterations, relative to LE-IC.

Another use of EXIT analysis is performance prediction, however this involves strong assumptions on prior inputs, that often cannot be met for FIR turbo equalizers (Assumption 2 of Section 2.3.6 is needed). Hence, EXIT curves merely provide an upper-bound on information rates, for receivers other than the MAP detector. In this respect, it is then interesting to compare transfer curves, with actual finite-length mutual trajectories (in dashed lines in Figure 3.15), in order to assess the accuracy of EXIT prediction for a given receiver.

It had been noted in [TK+02], that trajectories of DFE-IC with hard feeback, and with filters designed with "perfect" decision assumption, do not follow EXIT curves (see Figure 3.6). This issue remains with DFE-IC PAPP, although less severely, indicating that the perfect decisions assumption causes a severe information loss, despite using a soft decision feedback. Other FIRs' trajectories overall follow receiver and decoder curves and reach the MFB. Nevertheless, following a number of iterations, these trajectories no longer fits with transfer curves, and lose convergence speed. This is a common disadvantage of FIR equalizers, attributed to short cycles in the associated factor graph, which causes neglected correlations with the neighbouring symbols, as shown in Figure 16 in [TK+02]. However note that among different DFE-IC receivers, EP feedback yields trajectories that remains closest to EXIT curves, making it relatively easier to predict.

As discussed previously, in Section 2.3.6, the achievable spectral efficiency for a given receiver can be measured with the help of the area theorem for EXIT charts. In Figure 3.16, achievable rates for BPSK constellation are plotted, alongside the SIR of the channel, given by the SISO MAP detector (which is accurate). As non-iterative FIR do not depend on prior inputs, their achievable rates are also accurately computed. For turbo FIR, upper bounds are obtained by combining results of area theorem with the channel SIR; indeed, for APP based receivers, the area theorem predicts that these receivers can surpass the channel SIR, which

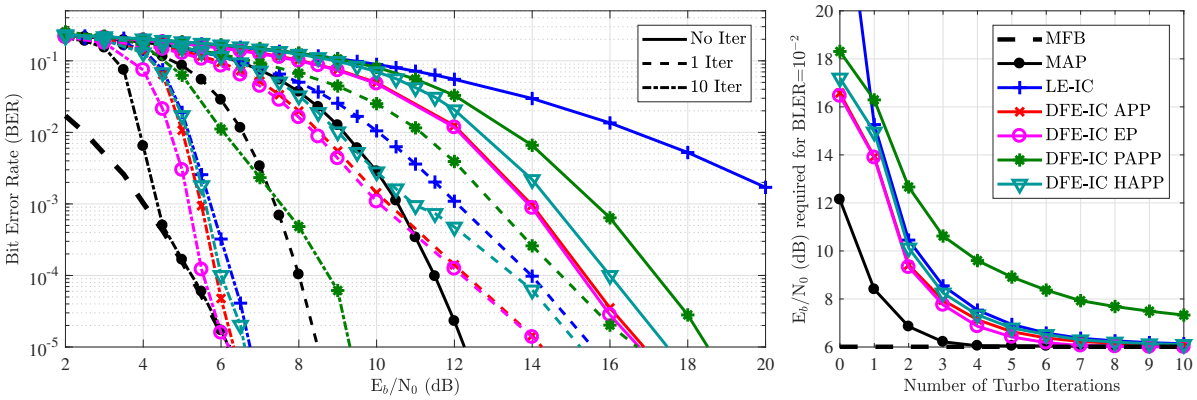


Figure 3.17: BER and convergence performance of the proposed DFE-IC in Proakis C channel with BPSK constellation.

is not physically plausible (this is shown in red dotted lines in Figure 3.16). In this case, the achievable rates obtained with EXIT must be limited to that of MAP, but this significantly puts into question the accuracy of these bounds for DFE-IC APP.

Tightness of these bounds depend on the closeness of true MI trajectories to EXIT charts in Figure 3.15, so APP feedback's asymptotic performance is most likely overestimated compared to EP feedback. This observation is in accordance with emerging theoretical work on the analysis of detectors using APP estimates or EP techniques (e.g. OAMP) [Ma+18].

Finite-Length Comparison with Existing Schemes

Monte Carlo based performance analysis remains the most reliable analysis approach, hence finite-length evaluation of joint detection and decoding for BPSK symbols is considered with parameters in section 3.4.6, and $K_b = 2048$, coded with a terminated $[7, 5]_8$ RSC code. Bit Error Rate (BER) of various receivers are plotted in Figure 3.17. For the reported iterations, the DFE-IC APP outperforms other APP feedback DFE structures, and their convergence speeds are compared on the right side of the figure, at a Packet Error Rate (PER) of 10^{-2}. EP-based feedback provides further improvement of the threshold by 0.5 dB relative to APP, and it is shown to reach MFB limit within 7 iterations, earlier than DFE-IC APP.

Assessing DFE-IC performance at low spectral efficiency conditions, as above, is of interest, to remedy the poor behaviour of classical DFE at those operating points (see Figure 3.16). Indeed, well-designed turbo processing helps DFE structures to outperform LE at all rates.

A higher spectral efficiency case is plotted on the left side of the Figure 3.18, with 8-PSK constellation in the same configuration; DFE-IC APP is shown to improve LE-IC waterfall threshold by 2 dB, and DFE-IC EP asymptotically provides an additional 1.2 dB. On the right side of the Figure 3.18, 16-QAM is considered; showing that DFE-IC EP provides further performance enhancements for one or more iterations.

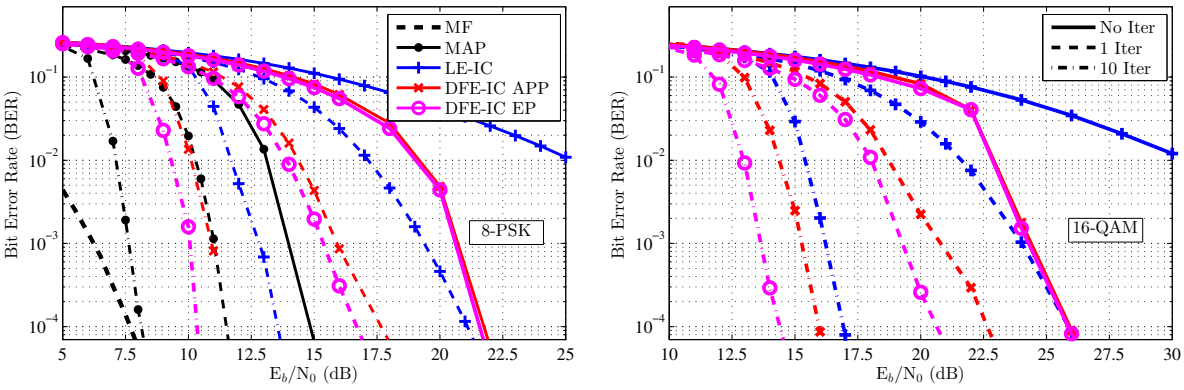


Figure 3.18: BER performance of the proposed DFE-IC in Proakis-C with 8-PSK and 16-QAM constellations.

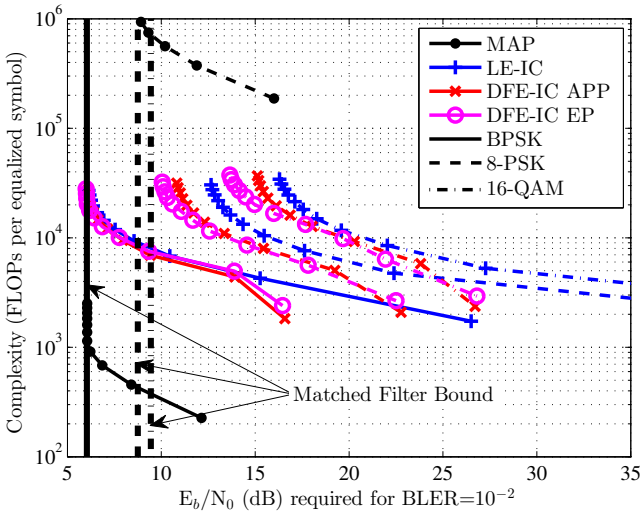


Figure 3.19: Performance complexity trade-off in Proakis C.

Finally, the coded performance of DFE-IC is balanced with complexity considerations. In Figure 3.19, the receiver computational complexity (Floating Point Operation (FLOP) per symbol) required to decode with a PER of 10^{-2} is plotted as a function of E_b/N_0. These values are computed, assuming the use of the proposed matrix inversion algorithm in section 3.4.5, and by accounting for the equalization, the demapping and the decoding costs. A curve represents the evolution of PER and the complexity of a receiver across turbo iterations.

DFE-IC provides a better trade-off than LE-IC; at any given amount of computationally complexity, it is more energy efficient, especially at initial iterations, and the asymptotic E_b/N_0 gap between LE-IC and DFE-IC increases with the modulation order M. The use of EP feedback is more advantageous at higher iterations, for higher order constellations, while APP is more cost-effective for non-iterative receivers.

In conclusion, DFE-IC outperforms LE-IC in various aspects: it converges faster towards

MFB, has a lower decoding threshold than LE-IC, especially at higher spectral efficiencies. Among DFE-IC with APP feedback, exact derivation DFE-IC APP is superior according to both finite-length and asymptotic analysis. Although EXIT charts show little difference between DFE-IC EP and APP, in practical simulations EP feedback tends to outperform APP. This is justified by the tightness of EP-based receiver's MI trajectories to corresponding EXIT curves; APP is overestimated. Although DFE-IC EP appears to be able to reach channel SIR at low to medium spectral efficiencies, there is still a gap to MAP performance qt very high rate.

Performance Over a Doubly Selective Standardized Channel

In this paragraph, a PHY layer inspired by the LTE uplink is discussed, with single carrier transmissions, but with time-domain receivers and perfect channel knowledge. Symbol rate is 1.08 Mbauds, i.e. 6 physical resource blocks in LTE, corresponding to an occupied bandwidth of 1.35MHz with shaping filters of 0.35 roll-off factor. The 1 ms subframe format from [3gpa] is used for the uplink data only, and 15 Modulation and Coding Schemes (MCSs) with QPSK, 16-QAM and 64-QAM constellations are used, with rate matching on the LTE turbo code. Transport block size has been selected from [3gpa], in order to satisfy code rate requirements per subframe.

The achievable throughput defined as

$$\eta_{\text{throughput}} = (1 - \text{PER})\eta_{\text{eff-MCS}}, \tag{3.69}$$

where $\eta_{\text{eff-MCS}}$ is spectral efficiency of an MCS, and PER is obtained through Monte Carlo simulations in the extended typical urban (ETU) channel [3gpb]. Its independent paths experience a Jakes Doppler spectrum, with a maximum Doppler shift of 300Hz, corresponding to a velocity of 135km/h for a carrier at 2.4GHz. This causes strong time variations over the subframe length of 1ms. The scheduling of joint detection and decoding at the receiver is one turbo code iteration (inner code iteration with memory) for each turbo equalization recursion. For both equalizers we use $N_w = 34, N_d = 22$, corresponding to a sliding window span over four times the channel spread. Fig. 3.20 shows the maximum achievable throughput (best MCS is selected) as a function of E_s/N_0. Operation region limits of the fifteen MCS are delimited with dash-dotted lines.

With non-iterative receivers, DFE-IC EP and APP are equivalent, and provide less than 20% energy saving w.r.t. LE-IC, for early QPSK MCS (1-6, up to 1.28Mbps). For 16-QAM MCS (7-9), this steadily increases from 15% up to 65%, and LE-IC can barely operate above MCS 10 due to the limited filter span. Finally, with five turbo iterations, DFE-IC APP provides less than 10% energy savings for QPSK schemes, around 15% for MCS 11, up to 60% for MCS 12, and LE-IC fails to follow further than 4.5Mbps. DFE-IC EP brings an additional improvement of around 0.5-1dB. Computational burden of DFE-IC with a time-varying channel is increased by 0.12%, 0.51% and 2.0% w.r.t to LE-IC, for respectively QPSK, 16-QAM and 64-QAM.

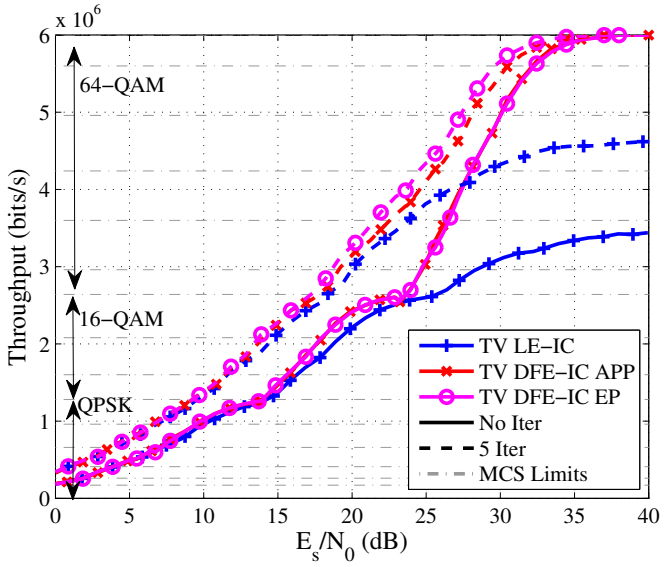


Figure 3.20: Achievable throughput in ETU300 with LTE-like coding and turbo FIR at PHY.

3.4.7 Conclusions

This section addresses the irregularities of the behaviour of turbo FIR DFE-IC based on soft feedback, and proposed alternative solutions to mitigate these problems.

To achieve this, a generic model for turbo FIR is given, through conjugate priors on Gaussian distributions, and APP feedback schemes from the literature are generalized to TV exact structures. Moreover, a factor graph view of FIR receivers is proposed, which enabled deriving a novel DFE-IC with soft feedback based on EP. This receiver is later extended to also cover Fractionally-Spaced (FS) DFE [Cas+97], by taking $m_{\text{up}} > 1$, in the sampling model of Figure 1.6, and removing the matched-pulse-shaping-filter at the receiver [Pet+19].

Numerical results show that the soft feedback and the TV modelling of prior estimates' uncertainty allows to solve previously encountered issues. Next, a matrix inversion strategy is proposed for addressing the time-varying filtering dynamics of these equalizers, and numerical and analytical results indicate that APP based techniques remain poorly predicted and difficult to analyse unlike EP. Indeed, this section provides another example to the importance of using extrinsic estimates in iterative systems, which is well suited for prediction and outperforms other alternatives in very selective channel configurations. Expectation Propagation (EP) framework enables computing such feedback in this problem.

3.5 Self-Iterated FIR Receivers with Expectation Propagation

3.5.1 On Self-Iterated Turbo Equalizers

As mentioned in the heart of the previous chapter, in Section 2.4.1, self-iterated detection structures have been gathering interest lately, thanks to either well-designed information combining techniques, or through advanced iterative algorithms, such as EP, which enables iteratively improving extrinsic messages, when mixed discrete/continuous PDFs are involved in a factor. A more general introduction to the concept of inner detection iterations is provided in Section 2.4.2, and further examples to this will be clear when discussion Frequency Domain (FD) strucutres in the next chapter.

Regarding DFEs, the idea of inner detection loop, *in addition* to the serial symbol-wise decsion feedback, was discussed in [JM13] for SISO DFE-IC, where parallel concatenation of previously discussed FIRs [JM10; JM11], is shown to significantly improve performance, or it is used to build bi-directional DFE, which itself can be used in concatenation. These structures have become possible thanks to the improve DFE-IC of [JM10], which is the sub-obtimal hard decision version of DFE-IC HAPP of the previous section, where the amount of residual ISI is more realistically assessed, compared to the classical DFE-IC [TK+02], enablling accurate MMSE combining of different FIR receivers.

As initially evoked in Section 2.4, EP has been introduced for communications systems with various applications such as Kalman smoothers for detection in time-varying channels [QM07], or in frequency-selective channels [SZW+15]. In particular, self-iterations through EP, which has been previously exploited for MIMO detection in [Sen+11; C$\acute{+}$14], has also been shown to considerably improve performance with block LE-IC [SMF+17a], and with Kalman smoothers [SZW+15; SMF+17b]. In particular, a concomitant work [SMF+18] has recently extended the schemes in [C$\acute{+}$14; SMF+17a; SMF+17b] to use prior information, and also propose a Self-Iterated (SI) FIR LE-IC[7]. The receivers above use EP in a parallel interference cancellation scheduling through self-iterations, i.e. the whole data block is detected, and then detection process is repeated using EP feedback from the demapper.

These structures are not decision feedback structures as in [Bel+79], which are natural successive interference cancellers. In this section, as in [Sen+11], we present a factor-graph based derivation of self-iterated FIR receivers with EP, which provides an alternative view of the simultaneously developed SILE-IC in [SMF+18], and which also allows derivation of a novel SIDFE-IC. The asymptotic and finite-length performance of this new structure is analysed, and it is shown to significantly approach the SIR of the channel, and outperforming even block SILE-IC structures (or their equivalent Kalman smoothing implementations) [SMF+17a; SMF+17b; SMF+18] when communicating at the high spectral efficiency operating points of the channel.

[7] This work was unavailable in any format to the public, when the contents of this chapter were being written as a journal article, and submitted for publication.

3.5.2 FIR Self Iterated DFE-IC with EP

A Doubly-Iterative Message-Passing Schedule for SISO Detection

It has been argued, with the derivation of a generic diagonal EP detection in Section 2.4.2, that self-iterated detection process can provide attractive performance-complexity trade-offs in turbo receiver design. Originally, a Soft Input Soft Output (SISO) MIMO detector based on this approach with one self-iteration has been proposed in [Sen+11], and it has been generalized to multiple self-iterations in [C$\acute{+}$14], where the authors note importance of using damped EP feedback for IC, for achieving better convergence. In the following, this concept is investigated for FIR receiver design, and as the demapping process is computationally less intensive than channel decoding, such structures have significant practical interest.

Following the recommendations on *damped* message passing in Section 2.4.1, and considering the successful use of feature-based damping in [C$\acute{+}$14], the EP-based feedback in Equation (3.62) is replaced with

$$\begin{aligned} v_{x,k}^{d(\text{next})} &= \left[(1-\beta) \frac{1}{v_{x,k}^{\text{curr}}} + \beta \frac{1}{v_{x,k}^{d(\text{prev})}} \right]^{-1}, \\ x_k^{d(\text{next})} &= v_k^{d(\text{next})} \left[(1-\beta) \frac{x_k^{d(\text{curr})}}{v_{x,k}^{d(\text{curr})}} + \beta \frac{x_k^{d(\text{prev})}}{v_{x,k}^{d(\text{prev})}} \right], \end{aligned} \tag{3.70}$$

where $0 \leqslant \beta \leqslant 1$ configures the damping and $x_{x,k}^{d(\text{curr})}$ and $v_{x,k}^{d(\text{curr})}$ are given by Equation (3.62). EP message passing algorithm consists in minimizing global divergence through iterative minimization of simpler local divergences, and it might lock on undesirable fixed points, thus a damping heuristic, as recommended in [Min+05, eq. (17)], is used to improve accuracy, and its effectiveness has been verified in [SMF+18].

In this section, we consider the use of self-iterations with FIR receivers, through the FIR EQU_k factor node in Equation 3.56, illustrated in the factor graph of Figure 3.10. To clarify the dynamics of the proposed receivers, $\tau = 0, \ldots, \mathcal{T}$ denotes turbo iterations (TI), i.e. exchanges between the DEM_k and DEC factor nodes for updating extrinsic LLRs, for $k = 1, \ldots, K$, with a parallel schedule. Each TI consists of $s = 0, \ldots, \mathcal{S}_\tau$ self-iterations (SI) (may vary with τ), i.e. exchanges between EQU_k and DEM_k factor nodes. In the following, EQU_k and DEM_k messages, derived previously in Section 3.4.4, are appended a superscript (τ, s) for denoting iterative dynamics.

Self-Iterated LE-IC with Expectation Propagation

When following a parallel message passing schedule between factor nodes EQU_k and DEM_k, messages to variable nodes $m_{\text{EQU}_k \to x_k}(x_k)$ are sequentially or simultaneously computed, for $k = 1, \ldots, K$, first, and then and only then, DEM_k factor nodes can start computing their messages. This structure corresponds to a LE-IC, within each self-iteration, when con-

Algorithm 12 Proposed *Self-Iterated DFE-IC EP* receiver.

Input $\mathbf{y}$, $\mathbf{H}$
Output $\hat{\mathbf{b}}$

$L_p^{(0)}(\mathbf{d}) = \mathbf{0}_{K_c}$
for $\tau = 0 \dots \mathcal{T}$ **do**
  Compute prior PMFs $\mathcal{P}_k^{(\tau)}(\alpha)$ with Eq. (2.39) and $L_p^{(\tau)}(\mathbf{d})$, for $\alpha \in \mathcal{X}, k = 1, \dots, K$.
  Compute $\mathbf{x}^{\mathbf{p}(\tau)}$ and $\mathbf{v}_{\mathbf{x}}^{\mathbf{p}(\tau)}$ with Eq. (2.40), and set $(x_k^{d(\tau,0)}, v_{x,k}^{d}{}^{(\tau,0)}) \leftarrow (x_k^{p(\tau)}, v_{x,k}^{p}{}^{(\tau)})$.
  for $s = 0$ to $\mathcal{S}_\tau$ **do**
    for $k = 1 \dots K$ **do**
      $\mathbf{\Sigma}_k^{\text{dfe-ep}(\tau,s)} \leftarrow \sigma_w^2 \mathbf{I}_{N_w} + \mathbf{H}_k \mathbf{Diag}(\bar{\mathbf{v}}_{\mathbf{x},k}^{\text{dfe-ep}(\tau,s)}) \mathbf{H}_k^H$
      $\xi_k^{\text{dfe-ep}(\tau,s)} \leftarrow \mathbf{h}_k^H \mathbf{\Sigma}_k^{\text{dfe-ep}(\tau,s)-1} \mathbf{h}_k$
      $\mathbf{f}_k^{\text{dfe-ep}(\tau,s)} \leftarrow \mathbf{\Sigma}_k^{\text{dfe-ep}(\tau,s)-1} \mathbf{h}_k \xi_k^{\text{dfe-ep}(\tau,s)-1}$
      $\mathbf{g}_k^{\mathbf{c}\text{-dfe-ep}(\tau,s)} \leftarrow \mathbf{H}_k^{\mathbf{c}H} \mathbf{f}_k^{\text{dfe-ep}(\tau,s)}$, $\mathbf{g}_k^{\mathbf{a}\text{-dfe-ep}(\tau,s)} \leftarrow \mathbf{H}_k^{\mathbf{a}H} \mathbf{f}_k^{\text{dfe-ep}(\tau,s)}$
      $x_k^{e(\tau,s)} \leftarrow x_k^{p(\tau,s)} + \mathbf{f}_k^{\text{dfe-ep}(\tau)H} \mathbf{y}_k - \mathbf{g}_k^{\mathbf{c}\text{-dfe-ep}(\tau,s)H} \mathbf{x}_k^{\mathbf{d}(\tau,s+1)} - \mathbf{g}_k^{\mathbf{a}\text{-dfe-ep}(\tau,s)H} \mathbf{x}_k^{\mathbf{d}(\tau,s)}$
      $v_{x,k}^{e(\tau,s)} \leftarrow 1/\xi_k^{\text{dfe-ep}(\tau,s)} - v_{x,k}^{p(\tau,s)}$
      Compute the posterior PMF $\mathcal{D}_k^{(\tau)}(\alpha)$ with Eq. (2.53), for $\alpha \in \mathcal{X}, k = 1, \dots, K$
      Generate EP feedback $(x_k^{d(\tau,s+1)}, v_{x,k}^{d(\tau,s+1)})$ with Eqs. (3.61)-(3.62), and if $v_{x,k}^{*(\tau,s+1)} \leqslant 0$, store k in the set $\mathcal{I}_{\text{err}}^{(s)}$, and then apply damping with Eq. (3.70).
    end for
    $\forall k \in \mathcal{I}_{\text{err}}^{(s)}$, $(x_k^{d(s+1)}, v_{x,k}^{d(s+1)}) \leftarrow (x_k^{*(s)}, v_{x,k}^{*(s)})$.
  end for
  Compute extrinsic LLRs $L_e^{(\tau)}(\mathbf{d})$ with Eq. (2.54)
  Execute de-interleaving, SISO decoding and interleaving to get $L_p^{(\tau+1)}(\mathbf{d})$
end for

sidering the generic FIR model in Equation (3.44), and the associated prior estimates and their reliabilities for IC are

$$\begin{aligned} \bar{\mathbf{x}}_k^{\text{le-ep}(\tau,s)} &= [x_{k-N_p'}^{d(\tau,s)}, \dots, x_{k+N_d}^{d(\tau,s)}]^T, \\ \bar{\mathbf{v}}_{\mathbf{x},k}^{\text{le-ep}(\tau,s)} &= [v_{x,k-N_p'}^{d(\tau,s)}, \dots, v_{x,k+N_d}^{d(\tau,s)}]^T. \end{aligned} \tag{3.71}$$

Independently of our work, this EP-based FIR structure is derived in the concomitant work [SMF+18], and unlike the message passing formalism used in Section 2.4.1, structure in [SMF+18] is obtained by approximating a self-iterated block receiver. If the computations of messages on EQU is carried out only once ($\mathcal{S}_\tau = 0$), this SILE-IC receiver yields the same result as the conventional turbo LE-IC [TSK02], as $(x_k^{d(\tau,0)}, v_{x,k}^{d(\tau,0)}) = (x_k^{p(\tau)}, v_{x,k}^{p(\tau)})$.

Self-Iterated DFE-IC with Expectation Propagation

In this paragraph, a serial scheduling across variable nodes x_k, $k = 1, \dots, K$ is considered. In detail, when EQU_k updates a variable node x_k, the factor node DEM_k is immediately

activated in order to provide its own extrinsic estimation of x_k, jointly using prior information from the decoder and the equalizer's extrinsic output. When detection across the whole block is completed, this serial scheduling can be repeated by keeping the previously updated DEM_k messages, yielding a SIDFE-IC structure.

The proposed scheduling, given in Algorithm 12, generates an EP FIR receiver which uses the following means and variances for interference cancellation

$$\begin{aligned}\bar{\mathbf{x}}_k^{\text{dfe-ep}(\tau,s)} &\triangleq [x_{k-N_p'}^{d(s+1)}, \ldots, x_{k-1}^{d(\tau,s+1)}, x_k^{d(s)}, \ldots, x_{k+N_d}^{d(s)}]^T, \\ \bar{\mathbf{v}}_{\mathbf{x},k}^{\text{dfe-ep}(s)} &\triangleq [v_{x,k-N_p'}^{d(s+1)}, \ldots, v_{x,k-1}^{d(s+1)}, v_{x,k}^{d(s)}, \ldots, v_{x,k+N_d}^{d(s)}]^T,\end{aligned} \tag{3.72}$$

for $k = 0, \ldots, K-1$. This layout shows that this structure indeed follows a TV DFE-IC evolution, with anti-causal symbols using demapper's output from the previous self-iteration, and causal symbols using current EP feedback from the demapper. The extrinsic feedback from the demapper is obtained by using jointly the prior information of the SISO decoder, the extrinsic outputs of the demapper from the previous self-iteration, and the equalizer's outputs on causal symbols of the current self iteration (see (3.62) and (3.70)). The Algorithm 12 also incorporates a mechanism to deal with EP-based feedback's infamous negative variances [Sen+11; SMF+17a], with the set $\mathcal{I}_{\text{err}}^{(s)}$ which stores their indexes. In the case of the self-iterated receiver, a hybrid approach is used where these erroneous values are replaced with APP-based variances in the current self-iteration, and then replaced again with their previous values for the next self-iteration.

This self-iterated receiver can be rewritten in the three-filter IC form of Equations (3.38) and (3.39), with

$$x_k^{e(\tau,s)} = x_k^{p(\tau,s)} + \mathbf{f}_k^{\text{sidfe-ep}(\tau,s)H}\mathbf{y}_k - \mathbf{g}_k^{\mathbf{c}\text{-dfe-ep}(\tau,s)H}\mathbf{x}_k^{\mathbf{d}(\tau,s+1)} - \mathbf{g}_k^{\mathbf{a}\text{-dfe-ep}(\tau,s)H}\mathbf{x}_k^{\mathbf{d}(\tau,s)}, \tag{3.73}$$

such that the filters are given by

$$\begin{aligned}\mathbf{f}_k^{\text{dfe-ep}(\tau,s)} &= (\sigma_w^2 + \mathbf{h}_k^H \boldsymbol{\Sigma}_{\mathbf{x},k}^{\text{dfe-ep}(\tau,s)}\mathbf{h}_k)^{-1}\mathbf{h}_k\xi_k^{\text{dfe-ep}(\tau,s)-1}, \\ \mathbf{g}_k^{\mathbf{c}\text{-dfe-ep}(\tau,s)} &= \mathbf{H}_k^{\mathbf{c}H}\mathbf{f}_k^{\text{dfe-ep}(\tau,s)}, \quad \mathbf{g}_k^{\mathbf{a}\text{-dfe-ep}(\tau,s)} = \mathbf{H}_k^{\mathbf{a}H}\mathbf{f}_k^{\text{dfe-ep}(\tau,s)},\end{aligned}$$

where $\xi_k^{\text{dfe-ep}(\tau,s)}$ and $\boldsymbol{\Sigma}_{\mathbf{x},k}^{\text{dfe-ep}(\tau,s)}$ are given as in Equation (3.44).

3.5.3 Asymptotic and Finite-Length Analysis on Self-Iterations

Asymptotic Comparison

First, we look into the achievable rates of SILE-IC and SIDFE-IC EP to identify operating points where self-iterations have an advantage.

We consider 8-PSK signalling on the Proakis C channel, and use the area theorem to obtain an upper bound on asymptotic achievable rates (i.e. $\tau \to \infty$), plotted on the left side of the

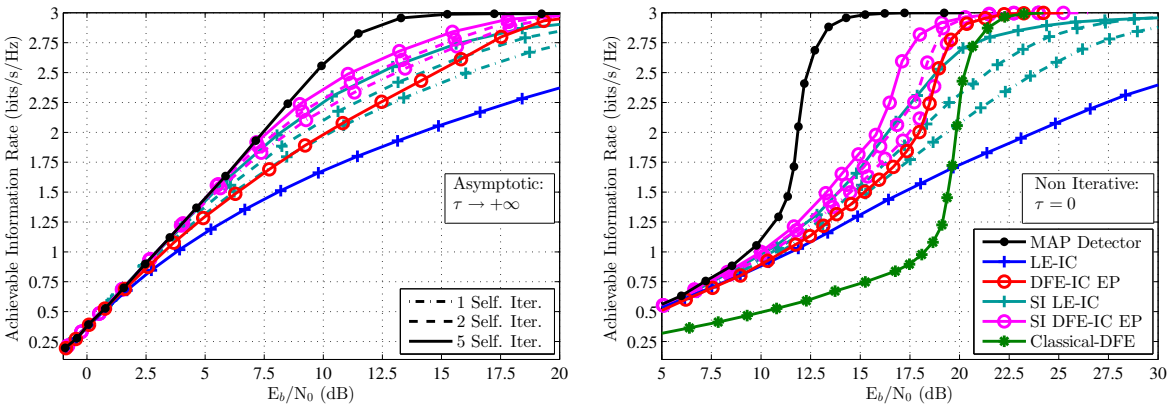


Figure 3.21: Achievable Rates of Self-iterated LE-IC and DFE-IC in Proakis-C with 8-PSK.

Figure 3.21. Information rates of the optimal MAP detector, LE-IC and DFE-IC EP without self iterations, and SILE-IC and SIDFE-IC are considered. For self-iterated receivers, a static damping with $\beta = 0.6$ is used. Numerical results show that self iterations are no needed for LE-IC up to 0.75 bits/s/Hz (i.e. using a code rate less than 1/4), as LE-IC is already close to the channel SIR, whereas DFE-IC EP reaches MAP detector's rates up to 1 bit/s/Hz (up to a code rate of 1/3). On the other hand, when using 5 self iterations, DFE-IC EP follows SIR within 0.5 dB up to 2.25 bits/s/Hz, while LE-IC follows it up to 1.85 bits/s/Hz. It is also interesting to note that DFE-IC EP with 2 self iterations outperforms LE-IC with 5 self iterations, at all rates, indicating at faster convergence of DFE-IC EP towards asymptotic limits of FIR IC structures.

At the right side of the Figure 3.21, non-turbo iterative achievable rates of these receivers, and those of the classical DFE [Bel+79], are compared. These rates are accurate, and not an upper bound, unlike asymptotic rates, and note that SISO MAP detector is a mere ML detector in this case. Although self iterations significantly improve LE-IC performance, at rates above 2.75 bits/s/Hz, classical DFE still outperforms these receivers. DFE-IC EP on the other hand outperforms alternative FIRs at any given self iteration.

Note that the gap to capacity still remains significant for non turbo iterative rates, and to some extent, for asymptotic rates. Hence with the objective of deriving capacity achieving practical receivers in mind, future work should explore the usage of the proposed DFE-IC EP as a constituent element for the bidirectional DFE of [JM11], or for paralellely concatenated FIRs [JM13] receivers.

Finite-Length Comparison

In this section, numerical finite-length results complete the previous analysis. In addition to receivers above, the SI block LE-IC (SI BLE-IC), denoted nuBEP in [SMF+18], is considered. Without self iterations, this receiver is equivalent to turbo block LE-IC[TS11], and it outperforms the self iterated block receiver and Kalman smoother in [SMF+17a; SMF+17b].

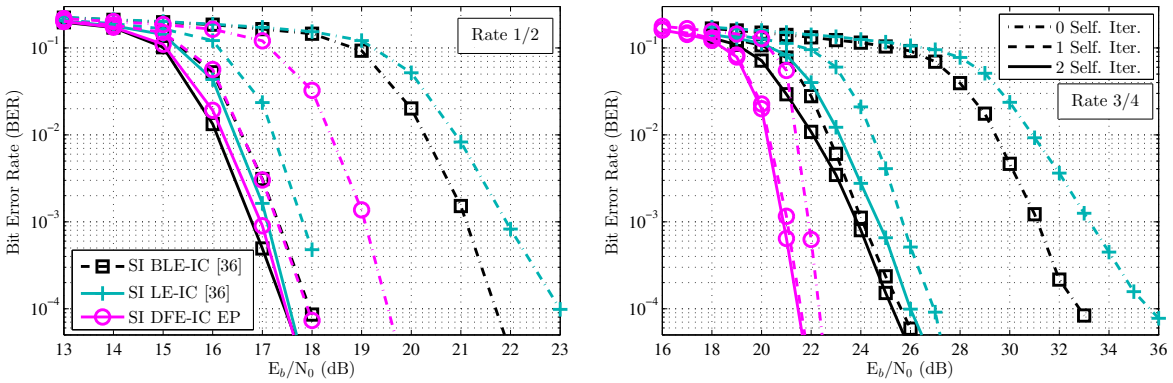


Figure 3.22: SILE-IC and SIDFE-IC with LDPC coded 16-QAM for 5 turbo iterations.

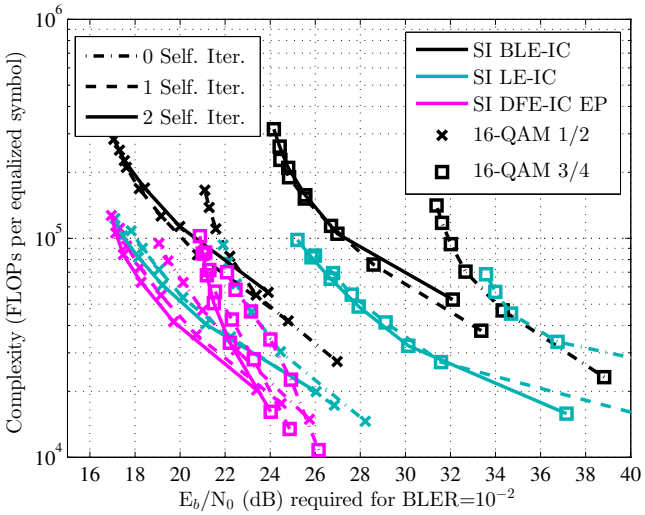


Figure 3.23: Performance complexity trade-off for self-iterations in LDPC coded Proakis C.

SIBLE-IC provides a lower bound to the BER performance of SILE-IC.

A LDPC coded 16-QAM transmissions over the Proakis C channel, with rate 1/2 and 3/4 encoding of $K_b = 2048$ bits (Figure 3.22). The proposed SIDFE-IC EP uses respectively $\beta = \min(0.5, 1 - e^{\tau/2.5}/10)$ and $\beta = \min(0.1, 1 - e^{\tau/1.5}/10)$ for damping, in these two cases, whereas the optimized damping reported in [SMF+18] is kept for SIBLE-IC and SILE-IC. The LDPC codes are obtained by Progressive Edge Growth methods, and a BP decoder up to a 100 iterations is used. The low rate case, with a (3,6) regular LDPC, shows that while all self-iterated receivers reach the same asymptotic performance as $\mathcal{S}_\tau$ increases, DFE-IC converges much faster at intermediary iterations. On the other hand, at the high rate configuration, with a (3,12) regular LDPC, DFE-IC is strictly superior to LE-IC, even without self-iterations. Asymptotically even the exact SIBLE-IC is 3.8 dB behind the proposed SIDFE-IC.

These numerical performance results are completed with computational complexity considerations in Figure 3.23, where decoding threshold for BLER $= 10^{-2}$ is evaluated for

$\tau = 0, \ldots, 5$, for each receiver. In the medium rate (2 bits/s/Hz: 16-QAM with rate 1/2 code) case the three considered receivers converges to the same asymptotic limit near 17 dB, but DFE-IC offers lower complexity at intermediary iterations. At 3 bits/s/Hz configuration (16-QAM with rate 3/4 code), with 5 turbo iterations and 3 self iterations, DFE-IC requires 3 dB less energy, and 3 times less computational resources than BLE-IC. With $\tau = s = 0$, LE-IC is unable to decode, BLE-IC decodes around 39 dB, and DFE-IC decodes with 13 dB less energy.

These numerical results confirms conclusions drawn by the asymptotic analysis; the proposed SIDFE-IC is of a significant interest for high data rate applications where linear structures are less efficient. Using the efficient implementation method of section 3.4.5, DFE-IC outperforms prior work in terms of both complexity and performance.

3.5.4 Discussion

This section evaluated the principle of double-loop turbo equalization with EP based self-iterations with FIR DFE-IC structures. Numerical results attest that self-iterations significantly boost equalization performance and achievable rates of the system further get closer to the channel SIR.

The process of detection being often less computationally intensive than decoding, self-iteration are a mean to accelerate the turbo process, and possibly achieve required performance behaviour with lower number of turbo iterations. Moreover, DFE-IC based on EP was already an impressive receiver, which outperformed non-self iterated alternatives, and through this extension it also outperforms, especially at high spectral efficiency operating points, or behaves similarly to various powerful equalization structures. The further reinforces the conjecture that the impact of serial demodulation feedback is very significant for high-spectral efficiency near-capacity systems. Thanks to finite-length and asymptotic analysis, DFE-IC EP, with SIs or not, sets new upper limits in achievable performance among FIR turbo receivers. At high data rates, even self-iterated block linear receivers fall over 3 dB behind the proposal.

Finally, the gap of achievable rates by turbo DFE-IC to the channel capacity remains still significant at very high spectral efficiencies. Mayhaps the proposed technique could further be improved by extending it as a bidirectional DFE, or by concatenating it with other FIR structures [JM13].

3.6 Low Complexity FIR Turbo DFE with Online Prediction

Up to this point, we have discussed novel techniques for improving the performance of Time Varying (TV) turbo FIR DFE structures, and in the remaining sections, we will discuss the design of an improved Iteration Variant (IV) FIR DFE, by extension of the proposed exact TV DFE-IV with APP and EP feedback.

3.6.1 Previous Works on FIR turbo DFE with Static Filters

The literature on turbo DFE-IC has been discussed in Section 3.4.1, but here we would like to recall some developments, and focus on the sub-optimal heuristics that has been used for IV filter design [LB06; LX14; Tao16]. The derivation of optimum turbo IV DFE-IC receivers is non-trivial, as static filters should depend on the decision feedback reliability, and the decision feedback naturally depends on the filters. As a consequence, there does not exist a closed-form expression of the optimal filter due to this non-linear 'chicken-and-egg' inter-dependence.

For DFE with hard decisions, the conventional approach is to assume a perfect feedback, but this causes error propagation and performance degradation with tentative decisions [TK+02]. The use of soft APP "decisions", while still assuming perfect feedback, partially mitigates error propagation, as soft symbols' magnitudes scale down with unreliability [Bal99]. In this case, an FIR receiver corresponds to the previously evoked IV DFE-IC with APP feedback, with $\bar{\mathbf{x}}_k^{\text{fir}} = [\mu_{x,k-N_p'}^d; \dots; \mu_{x,k-1}^d; x_k^p; \dots; x_{k+N_d}^p]$ and $\bar{\mathbf{v}}_{\mathbf{x},k}^{\text{fir}} = [\gamma_x^d \mathbf{1}_{N_p' \times 1}; v_x^p \mathbf{1}_{N_D \times 1}]$, with regards to the model in Equation (3.44) and with $x_k^p = \mathbb{E}_{\mathcal{P}_k}[x_k]$, $v_x^p = K^{-1} \sum_{k=1}^K \text{Var}_{\mathcal{P}_k}[x_k]$ and $\mu_{x,k}^d = \mathbb{E}_{\mathcal{D}_k}[x_k]$, with $\mathcal{D}_k$ being the posterior PMF in Equation (2.97), and $\gamma_x^d = 0$. However such a perfect feedback assumption, for the TV case, still resulted in poor and unpredictable iterative behaviour, as seen in Figure 3.15-(a). If it were possible to replace γ_x^d with a predicted estimation of the average posterior variance $K^{-1} \sum_{k=1}^K \text{Var}_{\mathcal{D}_k}[x_k]$, then optimal MMSE DFE-IC FIR with APP-based feedback could be derived.

The first reference to incorporate APP soft feedback reliability γ_x^d, in filter computations, is the receiver proposed in [LB06], for the special case of BPSK modulation. The direct dependencies between soft symbols and LLRs for the BPSK constellation enable the use of a tractable density evolution on the APP LLR distribution, given a prior LLR distribution from the decoder. This property is used by the BPSK receiver of [LB06] to estimate the decision feedback reliability, before equalization, to compute its filter coefficients.

However, this scheme cannot be directly generalized to high-order constellations; hence [LX14] proposed a two-tiers receiver, that implements an IV LE-IC at the first turbo iteration, and then it uses previous turbo iteration's demapper's APP LLRs to estimate soft-symbol statistics (i.e. γ_x^d) for IV DFE-IC. Note that this approach is only possible with Gray mapped constellations.

More recently, [Tao16] proposed to perform pre-equalization with LE-IC over a few symbols, and then to compute APP probability mass function, $\mathcal{D}_k$ of these symbols, to estimate APP soft-feedback reliability γ_x^d, to be used for the IV DFE-IC filter computation, due to the underlying bit-wise independent mapping model.

In the remainder, we provide a general view on how IV DFE-IC structures could be computed with an accurate estimation of the APP covariance γ_x^d, or more generally the covariance of causal symbol feedback $\bar{v}_x^c$, and we propose solutions based on semi-analytic receiver prediction techniques. To this end, the channel is assumed to be static, such that

$\mathbf{h}_k = \mathbf{h}_0$, for all $k = 1, \ldots, K$.

3.6.2 Predictive FIR DFE with Soft Feedback

In the case of generic DFE-IC, with anti-causal symbol priors $\bar{\mathbf{x}}_k^{\mathbf{a}} = [\bar{x}_k^a; \ldots; \bar{x}_{k+N_d}^a]$, and causal estimates $\bar{\mathbf{x}}_k^{\mathbf{c}} = [\bar{x}_{k-N_p'}^c; \ldots; \bar{x}_{k-1}^c]$, with associated variances $\bar{v}_{x,k}^a$ and $\bar{v}_{x,k}^c$ for $k = 1, \ldots, K$, when the IV DFE-IC is rewritten in the three-filter IC form, with all filters, $\mathbf{f}$, $\mathbf{g}^{\mathbf{c}}$ and $\mathbf{g}^{\mathbf{a}}$, being invariant per iteration, we have

$$\begin{array}{rcl} x_k^e & = & \bar{x}_k^a + \mathbf{f}^{\text{iv-dfe}\,H}\mathbf{y}_k - \mathbf{g}^{\mathbf{c}-\text{iv-dfe}\,H}\bar{\mathbf{x}}_k^{\mathbf{c}} - \mathbf{g}^{\mathbf{a}-\text{iv-dfe}\,H}\bar{\mathbf{x}}_k^{\mathbf{a}}, \\ v_x^e & = & 1/\xi - \bar{v}_x^a, \end{array} \quad , \quad \begin{cases} \mathbf{f}^{\text{iv-dfe}} \triangleq \boldsymbol{\Sigma}^{\text{iv-dfe}\,-1}\mathbf{h}_0\xi^{\text{iv-dfe}\,-1}, \\ \mathbf{g}^{\mathbf{c}-\text{iv-dfe}} \triangleq \mathbf{H}_k^{\mathbf{c}\,H}\mathbf{f}^{\text{iv-dfe}}, \\ \mathbf{g}^{\mathbf{a}-\text{iv-dfe}} \triangleq \mathbf{H}_k^{\mathbf{a}\,H}\mathbf{f}^{\text{iv-dfe}}, \\ \xi^{\text{iv-dfe}} \triangleq \mathbf{h}_0^H\boldsymbol{\Sigma}^{\text{iv-dfe}\,-1}\mathbf{h}_0, \end{cases} \tag{3.74}$$

where $\boldsymbol{\Sigma}^{\text{iv-dfe}} \triangleq k_w\sigma_w^2\mathbf{I}_N + \mathbf{H}_k\mathbf{Diag}\left(\bar{\mathbf{v}}_{\mathbf{x}}^{\text{iv-dfe}}\right)\mathbf{H}_k^H$, and where the variances of soft interference cancellation estimates are

$$\bar{\mathbf{v}}_{\mathbf{x}}^{\text{iv-dfe}} = [\bar{v}_x^c\mathbf{1}_{N_p',1}, \bar{v}_x^a\mathbf{1}_{N_d+1,1}], \tag{3.75}$$

where $\bar{v}_x^a$ and $\bar{v}_x^c$ are respectively the overall reliability of anti-causal and causal estimates.

For interference cancellation, the set of anti-causal estimates are available before equalization, and an accurate value of their reliability is given by the least-squares estimate

$$\bar{v}_x^a = \frac{1}{K}\sum_{k=1}^{K}\bar{v}_{x,k}^a. \tag{3.76}$$

In most SISO DFE-IC structures, the anti-causal estimates are the prior estimates given by the decoder, $x_k^p \triangleq \mathbb{E}_{\mathcal{P}_k}[x_k], v_k^p \triangleq \text{Var}_{\mathcal{P}_k}[x_k]$. However, note that in the case of the self-iterated DFE-IC EP of Section 3.5, the anti-causal estimates are the causal estimates of previous iterations.

The core of the problem lies in the computation of $\bar{v}_x^c$. A simple, but inaccurate solution is the "perfect decision assumption" : $\bar{x}_k^c$ are all assumed to be equal to x_k, yielding $\bar{v}_x^c = 0$. This approach is sufficient at very high SNR operating points, but as shown in [TK+02], and in previous sections, it degrades performance in moderately or severely selective channels. Hence, in the remainder of this section, we aim to use novel prediction methods to compute it, for enabling enhancement of the the IV DFE performance.

APP Soft Feedback Computation Let us recall that for APP-based soft feedback, estimates are given by the mean and the variance of the posterior symbol PMF $\mathcal{D}_k$, with Equations (2.97)-(2.98). For the IV DFE-IC APP filter computation, an invariant variance γ^d is needed, as the predicted causal reliability $\bar{v}^c$. Unlike the anti-causal reliability measures, γ^d cannot be estimated using the causal estimates $(\bar{x}_k^c, \bar{v}_k^c) = (\mu_k^d, \gamma_k^d)$, as these can only be generated once the filter is computed. Thus, a predictive estimation is required, and the

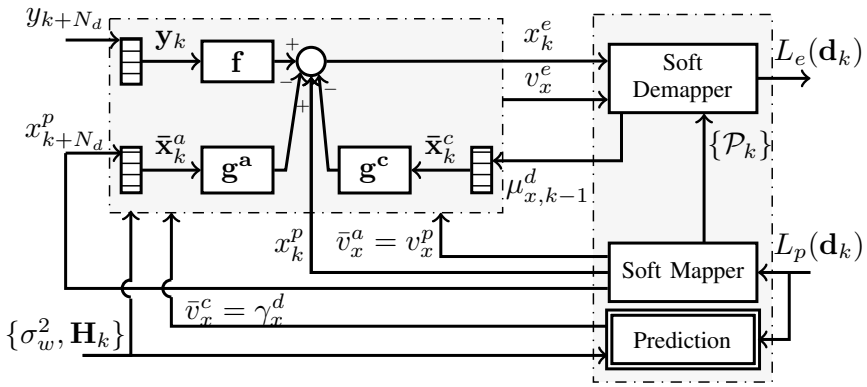


Figure 3.24: Soft-Input Soft-Output IV DFE APP receiver structure.

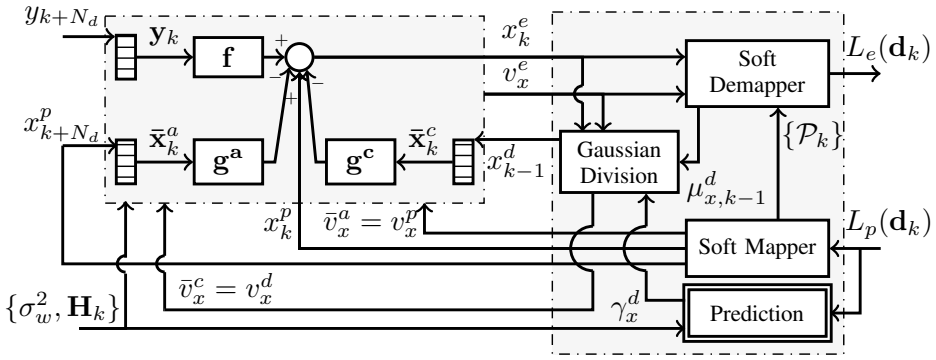


Figure 3.25: Soft-Input Soft-Output IV DFE EP receiver structure.

corresponding predictive DFE is illustrated in the Figure 3.24.

Predictive EP-based Soft Feedback Computation Unlike APP estimates, EP-based estimates carry only the extrinsic information brought by the demapper, and prevents DFE-IC from relying on its own bias. For TV DFE-IC EP, these estimates are obtained by a "Gaussian division", following Equations (3.61) and (3.62). For an IV DFE-IC EP structure, this feedback is not well-suited, as the invariant filter is unable to vary its coefficients to handle the strong variations of $v^d_{x,k}$, which depends on instantaneous APP variance $\gamma^d_{x,k}$. Hence it is preferable to use scalar EP, rather than diagonal EP, with extrinsic estimates

$$x^d_k = \frac{\mu^d_{x,k} v^e_x - x^e_k \gamma^d_x}{v^e_x - \gamma^d_x}, \text{ and, } v^d_{x,k} = v^d_x \triangleq \frac{v^e_x \gamma^d_x}{v^e_x - \gamma^d_x}, \tag{3.77}$$

where a predicted invariant APP variance γ^d_x is used to generate the feedback, as shown on Figure 3.25. Moreover EP-based estimates have an invariant variance $v^d_x = v^d_{x,k}, \forall k$ as the causal reliability, directly related to the predicted APP variance.

3.6.3 Semi-Analytic Abstraction of FIR DFE

In this section, a prediction model for the turbo DFE-IC structure in Equation (3.74) is exposed, without loss of generality, for the case where anti-causal estimates are given by decoder's extrinsic LLRs, i.e. $\bar{v}_x^a = v_x^p$. Such models are conventionally used for taking link adaptation decisions with low computational complexity, but here it will be exploited for online estimation of the reliability of causal estimates.

General Structure and Analytical Equalizer Model

SISO DFE-IC is modelled with two independent components; an analytical model for the equalizer and a numerical model for the soft demapper. Unlike asymptotic models ($K \to +\infty$) used in EXIT analysis, finite-length models are used for the demapper, as prior works on performance prediction noted their impact on accuracy [SF09; VB+10].

Following Equation (3.74), the IV DFE-IC reliability v_x^e is modelled by ϕ_{REC} as

$$\begin{aligned} v_x^e &= \phi_{\text{REC}}(\sigma_w^2, \mathbf{h}, v_x^p, \bar{v}_x^c) \\ &\triangleq (\mathbf{h}_0^H[\sigma_w^2 \mathbf{I}_N + \mathbf{H}\mathbf{diag}(\bar{\mathbf{v}}_{\mathbf{x}}^{\text{dfe}})\mathbf{H}^H]^{-1}\mathbf{h}_0)^{-1} - v_x^p, \end{aligned} \tag{3.78}$$

where $\bar{\mathbf{v}}_{\mathbf{x}}^{\text{dfe}}$ is given by Equation (3.75). This function is strictly increasing with $\bar{v}_x^c \in [0, \sigma_x^2]$.

The demapper is modelled with a Lookup Table (LUT) referred to as ϕ_{DEM}

$$\bar{v}_x^c = \phi_{\text{DEM}}(v_x^e, \cdot), \tag{3.79}$$

where $\bar{v}_x^c$ is the expected value of causal estimates' variance, taken over realizations of the channel noise, the equalizer outputs and the prior LLRs. The second argument '$\cdot$' in Equation (3.79) models prior information, and its exact nature depends on the selected prediction approach and improvements proposed in the upcoming subsections concern this module.

Since the equalizer and the demapper iteratively exchange reliabilities, the two functions representing their model must be composed to yield a recursive equation on $\bar{v}_x^c$:

$$\bar{v}_x^c[n+1] = \phi_{\text{DEM}}(\phi_{\text{REC}}(\sigma_w^2, \mathbf{h}, v_x^p, \bar{v}_x^c[n]), \cdot) \triangleq f_{\text{pred}}(\bar{v}_x^c[n]), \tag{3.80}$$

where $n = 0, \ldots, N_{\text{pred}}$ denotes iterations for the fixed point equations. If f_{pred} admits a unique fixed-point on $\bar{v}_x^c$, then the desired predicted reliability estimate is this fixed-point. Moreover, the optimality of IV DFE-IC strongly depends on $\bar{v}_x^c$ and hence on the accuracy of Φ_{DEM}. Figure 3.26 illustrates causal reliability estimation structures using two semi-analytical models that will be introduced below.

Numerical Demapping Models for APP/EP

Modelling the demapper with prior information is challenging due to the high amount of non-linearity it incorporates, and due to strong simplifying assumptions. The main focus

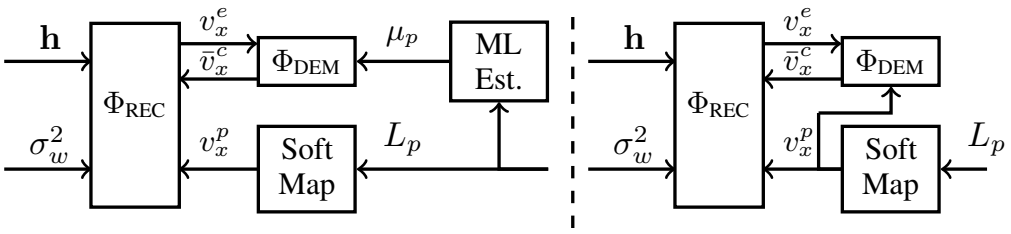


Figure 3.26: Binary (left) and symbol-wise (right) causal reliability prediction.

of the numerical model will be the characterization of APP variance γ_x^d required for using APP feedback, in the predictive receiver of Section 3.6.2. Similarly, as the variance v_x^d of the proposed predictive EP feedback in Section 3.6.2 is analytically linked to γ_x^d, numerical assessment of APP estimates is a common aspect required for both types of feedback.

Mutual Information based Prediction (Binary) In the BPSK receiver of [LB06], a prediction scheme is considered, assuming input/output LLRs of the demapper to be consistent Gaussian, i.e. $L_{(\cdot)}(d_{k,j}) \sim \mathcal{N}(\bar{d}_{k,j}\mu_{(\cdot)}, 2\mu_{(\cdot)})$, where $\bar{d}_{k,j} = 1 - 2d_{k,j}$, and where $(\cdot)$ is p, e or void, depending on concerned LLRs. Using a semi-analytical density evolution, parameter μ_e of extrinsic LLRs is predicted using μ_p. The parameter μ_p is bijectively linked to the average prior mutual information I_A between prior LLRs and the associated coded bits, as defined in Section 2.3.6, with the function $J(\cdot)$, which is usable for binary prediction, as shown in [VB+10]. Hence using such formalism, the approach of [LB06] can be extended to any constellation and mapping.

More specifically, the demapper behaviour is numerically integrated for each $\gamma_{x,k}^d$, $k = 1, \ldots, K$, over realizations of consistent Gaussian prior LLRs $L_p(d_{k,j})$ given μ_p, and of Gaussian-approximated equalized symbols $x_k^e \sim \mathcal{CN}(x_k, v_x^e)$, and a LUT on μ_p and v_x^e is built with

$$\bar{v}_x^c = \phi_{\text{DEM}}(v_x^e, \mu_p) \triangleq \frac{1}{K}\sum_{k=1}^{K} \mathbb{E}_{L_p, x^e}[\bar{v}_k^c], \tag{3.81}$$

where $\bar{v}^c$ is the APP/EP variance of soft symbols and μ_p represents prior information, which can be measured with a ML estimator (see Fig. 3.26, left)

$$\mu_p \approx \sqrt{1 + \sum_{k=1}^{K}\sum_{j=0}^{q-1} |L_p(d_{k,j})|^2} - 1. \tag{3.82}$$

Prior Variance based Prediction (Symbol-wise) The binary prediction scheme above appeared to yield too optimistic estimates in [LB06], and they instead decided to obtain μ_e and μ_p through BPSK channel estimators, which circumvents consistent Gaussian LLR approximation.

More specifically, this problem ensues from well known issues with regards to performance

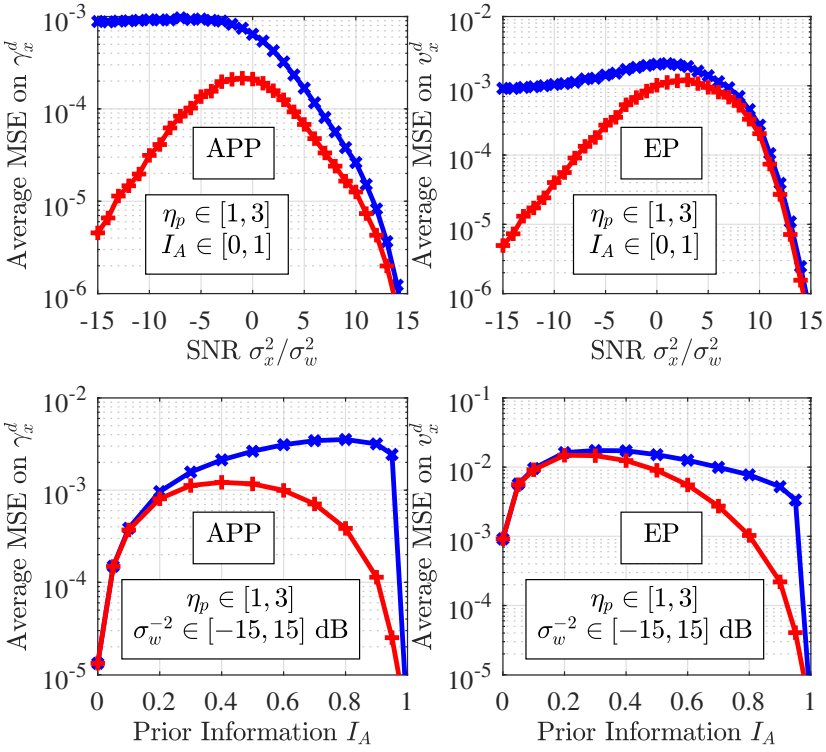


Figure 3.27: Binary (blue, $\times$) and symbol-wise (red, $+$) prediction.

prediction of turbo iterative systems, for which the consistent Gaussian approximation of LLRs was shown to be inaccurate across turbo iterations due to the internal non-linear dynamics of channel decoding [Fu05]. To overcome this prediction bias, performance prediction based on a two-parameter LLRs' model has been shown to be much more accurate [IB10]. Such models consider $L_{(\cdot)}(d_{k,j}) \sim \mathcal{N}(\bar{d}_{k,j}\mu_{(\cdot)}, \eta_{(\cdot)}\mu_{(\cdot)})$, where $\eta_{(\cdot)}$ is no longer 2. The ML estimator used for measuring μ_p in the binary prediction is very sensitive to η_p, which is the reason why the binary prediction is not enough robust in practice.

For our context, two-parameters models are too complex as they require expensive online parameter estimators to get both μ_p and η_p. Hence, a single-parameter demapper model with reasonable estimation complexity has been preferred. We searched for the parameter which is the most robust to the variations of prior LLRs' variance-to-mean ratio η_p.

Following a thorough and almost exhaustive study of the different alternative parameters for tracking evolution of $\bar{v}_x^c$, anti-causal variance v_x^p has been found to be *less sensitive* to the changes on η_p, very similarly to $\bar{v}_x^c$, with the advantage of v_x^p being directly computable online using a simple least-squares estimation. Hence, we propose the following LUT

$$\bar{v}_x^c = \phi_{\text{DEM}}(v_x^e, v_x^p), \quad \begin{cases} \bar{v}_x^c \triangleq K^{-1} \sum_{k=1}^{K} \mathbb{E}_{L_p, x^e}[\bar{v}_{x,k}^c], \\ v_x^p \triangleq K^{-1} \sum_{k=1}^{K} \mathbb{E}_{L_p}[v_{x,k}^p], \end{cases} \tag{3.83}$$

where both input v_x^p and output $\bar{v}_x^c$ are numerically integrated using prior LLRs generated

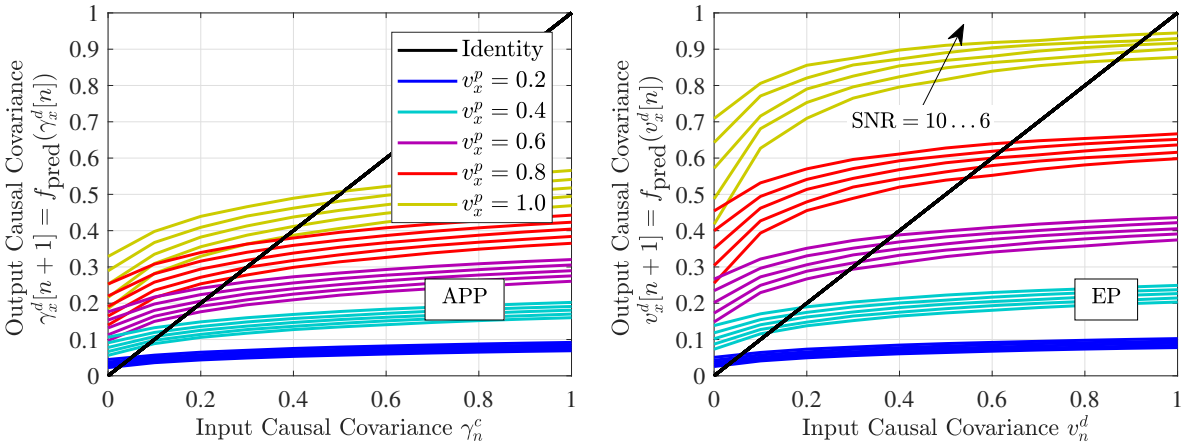


Figure 3.28: Numerically assessed fixed-points of the symbol-wise f_{pred} for SNR varying from 6 to 10 dB, for each value of prior reliability.

for a fixed value of η_p. Indeed, as η_p cannot be accurately measured online, the conventional consistent approximation [VB+10] is kept with $\eta_p = 2$, and in the following, we will assess its impact on the prediction accuracy.

Robustness of demapper prediction The sensitivity of the considered prediction schemes to variations in η_p is evaluated. This aspect is important for characterizing the robustness of iterative receiver prediction schemes, as the hypothesis $\eta_p = 2$, used for LUT generation, is only true at the initial turbo-iteration and then it varies [Fu05].

An AWGN channel is simulated with blocks of 16-QAM symbols with $K = 1024$, to emulate the output x^e of the equalizer, for v_x^e varying from -15 to 15 dB, along with Gaussian-distributed prior LLRs generated with I_A varying from 0 to 1 bit, with η_p varying from 1 to 3. The average MSE between the predicted causal covariance and true causal covariance is measured, and plotted in Figure 3.27. The left side of the figure provides results for APP feedback, and the right side for EP-based feedback. The binary approach is seen to be severely impacted by the changes in η_p, whereas the symbol-wise approach, although not perfect, remains more robust. Considerable differences are seen at low to medium SNR for high prior information, which suggests that symbol-wise schemes would have an advantage at the decoding threshold in asymptotic behaviour, i.e. when a high number of turbo-iterations are used. Oppositely, without any turbo-iteration, both schemes would perform identically.

Convergence Analysis

The convergence of the proposed iterative semi-analytical prediction schemes could be assessed formally through fixed-point analysis of Equation (3.80). However, due to the untractable non-linear expression of Φ_{DEM}, an analytic approach is not possible, and one has to resort to numerical methods.

Numerical evaluations of the proposed f_{pred} show that we can reasonably conjecture that

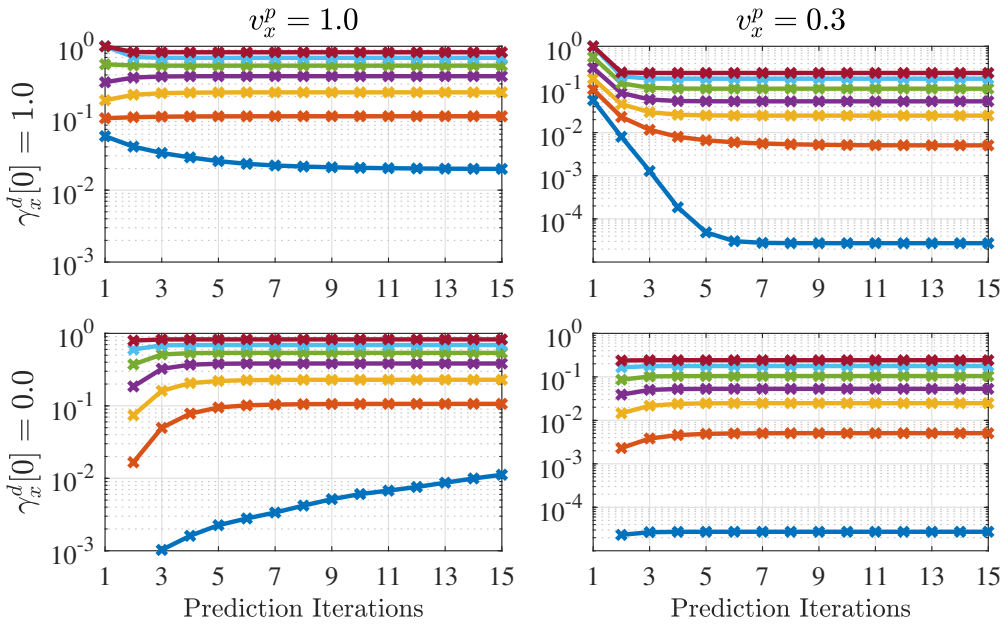


Figure 3.29: Evolution of APP covariance prediction $\gamma_x^d[n+1] = f_{\text{pred}}(\gamma_x^d[n])$ for SNR varying from -5 to 25 dB, with 5 dB steps (towards lower plots).

this function is continuous on the interval $[0, +\infty[$, with a Lipschitz constant strictly less than 1, for all $\sigma_w^2 \geqslant 0$ and $0 \leqslant v_x^p \leqslant \sigma_x^2$. This ensures that eq. (3.80) reaches a unique fixed-point $\bar{v}_x^c \in [0, +\infty[$ for any initial guess. This conjecture has been checked for various common channels $\mathbf{h}$ and Figure 3.28 plots f_{pred} for the Proakis C channel ($\mathbf{h} = [1, 2, 3, 2, 1]/\sqrt{19}$), using the symbol-wise demapper model for Gray-mapped 16-QAM.

The convergence speed of the prediction scheme is also evaluated numerically. The fixed-point $\bar{v}_x^c = \bar{v}_x^c[\infty]$ is reached more or less quickly depending on if the initial value $\bar{v}_x^c[0]$ is close to $\bar{v}_x^c[\infty]$. In particular, due to the near flat evolution of f_{pred} for v_x^c close to $\sigma_x^2 = 1$, initializing with $\bar{v}_x^c[0] = 1$ results in fast convergence at low SNR, and high anti-causal covariance, but slow convergence otherwise. Oppositely with $\bar{v}_x^c[0] = 0$ faster convergence is achieved for high SNRs and low anti-causal covariance. This behaviour is illustrated for Proakis-C 16-QAM APP covariance in Figure 3.29.

We propose to use the heuristic $\bar{v}_x^c[0] = \min(1, \sigma_w)$, when $v_x^p > 0.5$, where the standard deviation of the channel noise is experimentally shown to serve as a convergence accelerating heuristic. Otherwise using $\bar{v}_x^c[0] = 0$ is preferable for faster convergence.

3.6.4 Numerical Results

Uncoded equalization behaviour

In this paragraph, the uncoded finite-length behaviour of the proposed IV DFE-IC with online prediction is evaluated. Exact TV DFE-IC counterparts are used as lower-bound

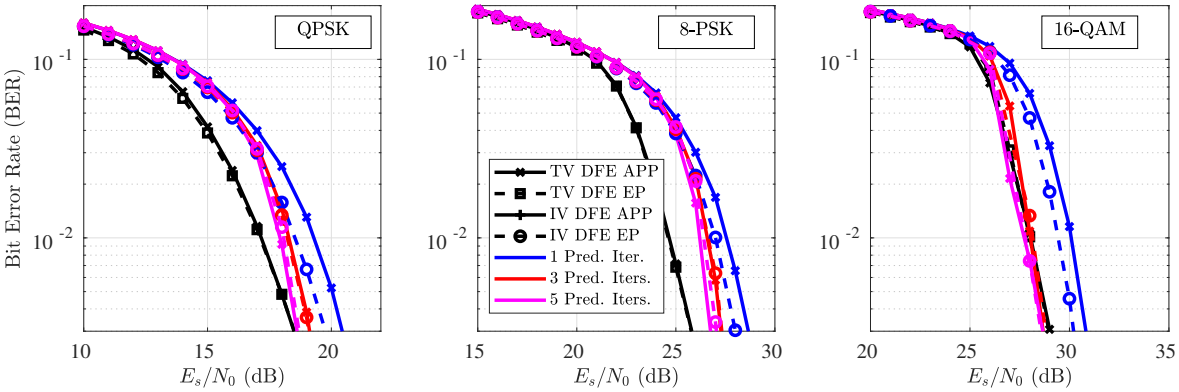


Figure 3.30: Uncoded BER performance of proposed predictive IV DFEs.

references on BER, to assess the prediction accuracy.

Block transmission in Proakis C channel is considered with $K = 256$ and with QPSK, 8-PSK and 16-QAM constellations. In Figure 3.30, BER of TV DFE-IC with APP and EP feedback are compared to the proposed predictive IV implementations. IV DFE-IC converges towards the curve of TV counterparts, especially at high SNR, but it is seen that a gap remains at medium BER for some constellations, due to dynamic filtering capabilities of TV receivers. EP feedback is shown to be mostly equivalent to APP feedback in this uncoded use case, but at high BER, EP-based feedback has an advantage over APP-based one, for both TV and IV receivers, which suggests that improved decoding thresholds can be obtained with channel coding.

On the Operating Regions of FIR Receivers

A previous work on TV FIR turbo equalizers concluded that TV DFE-IC significantly outperforms TV LE-IC at high data rates, as seen in Section 3.4, whereas TV LE-IC remains preferable at very low rates, as it achieves same performance with less complexity. In the following, the asymptotic behaviour and the computational complexity of the proposed receiver is evaluated in a similar manner, along with those of other IV FIR receivers.

Considering that for filter-based receivers which violate the extrinsic message principle of turbo detection, the rates predicted by EXIT can be too optimistic. This has been already observed with the TV APP feedback based receivers in Section 3.4, but the TV EP-based DFE did not suffer from this phenomenon. Hence in the following, the proposed predictive EP-based IV DFE-IC is evaluated.

IV DFE-IC EP, with symbol-wise prediction scheme, is used for 8-PSK transmissions in the Proakis C channel, and numerically obtained achievable rates are plotted in solid lines in Figure 3.31. Dotted plots illustrate the achievable rates without turbo-iterations, for each receiver. IV receivers are shown to follow the behaviour of their TV counterpart within a gap of about 0.1 bits/s/Hz for both LE-IC and DFE-IC, but IV DFE-IC still keep a significant

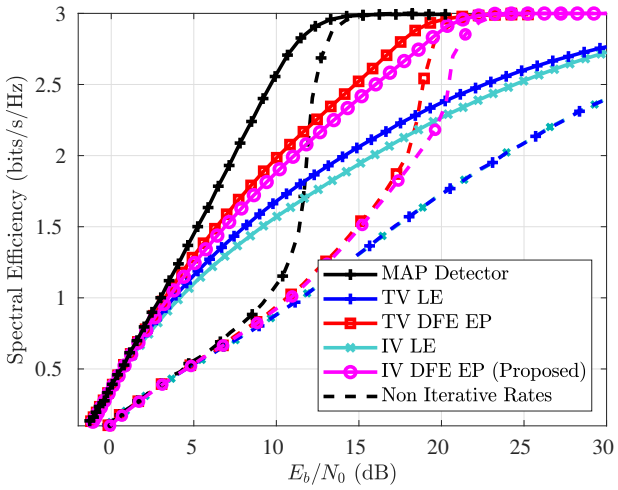


Figure 3.31: Achievable rates of FIR receivers for 8-PSK in Proakis C channel.

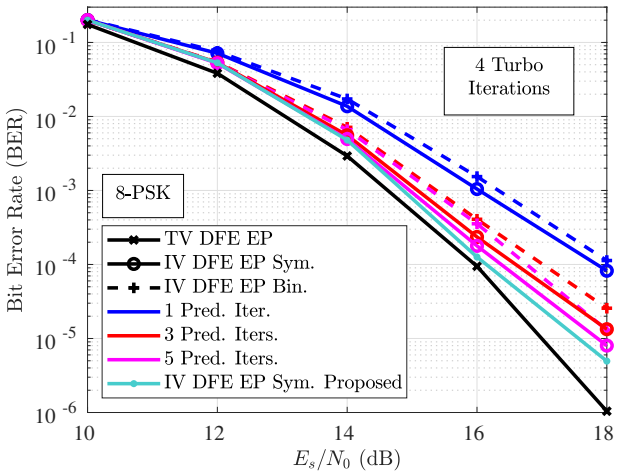


Figure 3.32: Rate-1/2 coded BER with proposed binary and symbol-wise prediction.

upper hand over TV LE-IC at medium and high spectral efficiency operating points. Using IV FIR receivers to operate at a given rate requires about 1.5 dB more energy than TV FIR, but with significant complexity savings.

Table 3.3: Computational Complexity of FIR Receivers

Structure	Filter Computation	Filtering and Detection
TV LE	$\mathcal{T}K(5L^3+56L^2)$	$\mathcal{T}K(25L+(11+3q)M)$
TV DFE	$\mathcal{T}K(5L^3+71L^2)$	$\mathcal{T}K(25L+(18+3q)M)$
IV LE	$\mathcal{T}(6L^3+28L^2)$	$\mathcal{T}K(25L+(11+3q)M)$
IV DFE Pred.	$\mathcal{T}(N_{\text{pred}}+1)(6L^3+34L^2)$	$\mathcal{T}K(25L+(18+3q)M)$

Approximate computational complexity of considered FIR receivers is given in the Table 3.3, where $\mathcal{T}$ denotes the number of turbo-iterations, and N_{pred} denotes the number of prediction iterations. TV LE-IC and DFE-IC receivers use the reduced-complexity TV matrix

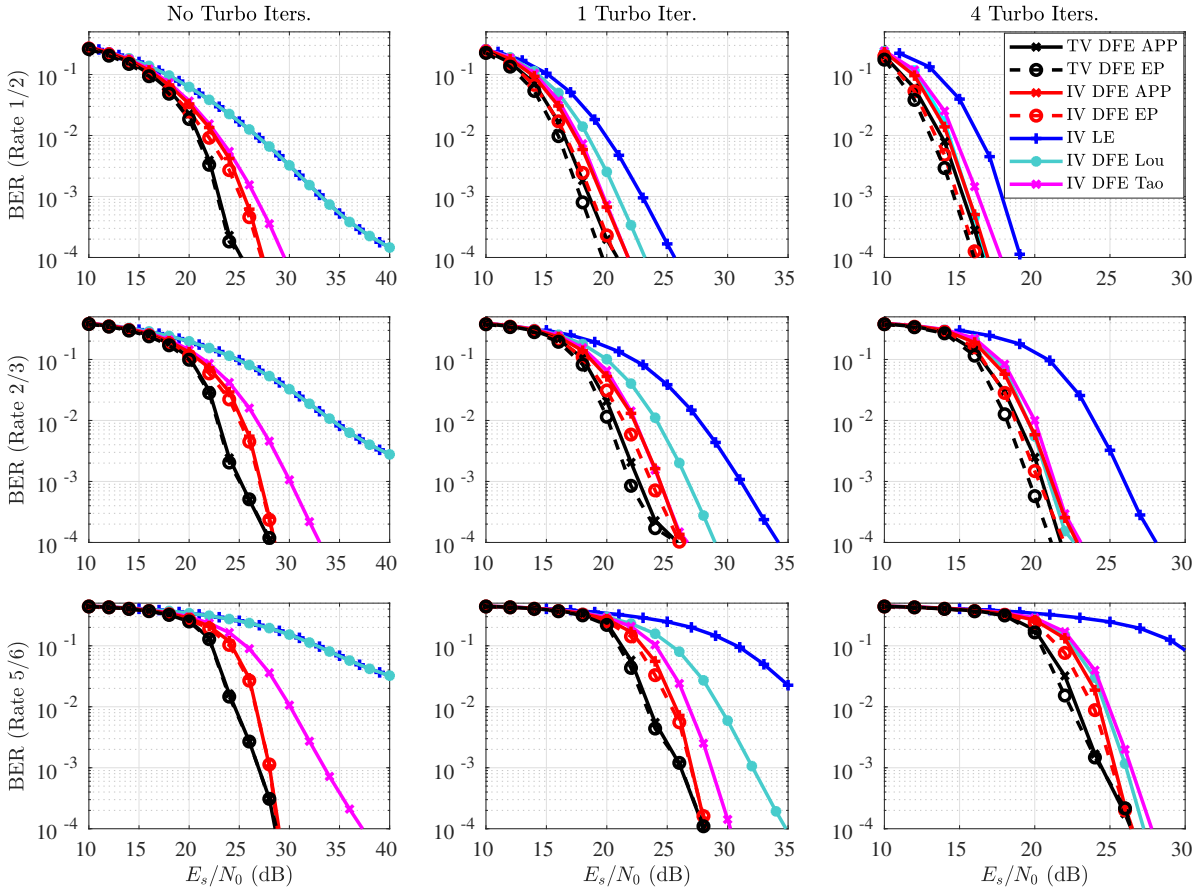


Figure 3.33: Coded 8-PSK BER performance comparison of turbo FIR receivers across turbo-iterations for different code rates.

inversion algorithms in Section 3.4.5, and IV receivers exploit a single Cholesky decomposition for matrix inversion.

Finite-Length Turbo-Equalization Performance

In this section, the prediction accuracy is assessed for transmissions encoded with Non-Recursive Non-Systematic Convolutional (NRNSC) code with polynomials $[7, 5]_8$. Note that IV FIR receivers might outperform TV FIR structures, in some cases [JM13], as the latter are more sensitive to the decoder errors.

First, the impact of using symbol-wise or binary prediction is assessed through finite-length BER evaluations. The block size is kept at $K = 256$, similarly to the uncoded case, and a MAP decoder based on BCJR algorithm is used as a SISO decoder. Fig. 3.32 shows the case of the EP-based feedback with 8-PSK, and the use of symbol-wise prediction is shown to accelerate convergence of the IV DFE-IC performance towards TV DFE-IC.

However, despite the improvements brought by the symbol-wise prediction, covariance

estimations tend to be too optimistic for high prior information at high SNR (following, 1 or 2 turbo iterations), and degrade BER performance. A similar observation has been made for the semi-analytic prediction of turbo linear MMSE receivers in [NVB12], where a calibration step is applied to correct the predicted prior covariance with a multiplicative factor. After some ad hoc damping, this yields more smoothed predictions that improve BER performance.

Here, this mechanism is adapted to the proposed online prediction. To avoid over-estimation of the causal covariance, the anti-causal covariance can be exploited to derive a "lower-bound" to estimated causal covariances. Empirically, turbo detection systems bring most of the improvements at the initial iterations, hence the improvements after a certain number of iterations can no longer be substantial. Thus, the predicted causal covariance $\bar{v}_x^c$ of the current turbo iteration is modified with $\bar{v}_x^c = \max(\bar{v}_x^c, \beta\bar{v}_x^a)$, with $\beta < 1$. The proposed heuristic is integrated with the symbol-wise prediction, with 3 prediction iterations and $\beta = 0.2$, and the IV DFE-IC with EP is displayed in turquoise in Figure 3.32.

Finally, to compare our proposal to the prior work and to evaluate its behavior in different operating regimes, the previously used rate-1/2 encoding with NRNSC code $[7, 5]_8$ is punctured to get rate-2/3 encoding with $[11; 01]$ puncturing pattern and rate-5/6 encoding with $[10001; 01111]$ puncturing pattern. The BER performance of the proposed IV DFE-IC APP and IV DFE-IC EP receivers are shown in red in Figure 3.33, for 8-PSK transmissions in Proakis C channel, with above mentioned codes of rate 1/2, 2/3 and 5/6, and for 0, 1 and 4 turbo-iterations. Proposed predictive IV DFE-IC receivers use symbol-wise prediction with 3 iterations, and the heuristic parameter is $\beta = 0.2$. IV DFE-IC APP significantly outperforms other APP-based DFE-IC receivers when there are no turbo iterations, as this is the operating point where the prediction scheme is the most accurate. In Figure 3.34, the evolution of BER is plotted as the number of turbo-iterations increases. During intermediary iterations of the rate 2/3 system, previous works of *Tao et al.* [Tao16] and *Lou et al.* [LX14] close most of the gap of iteration zero, with the receiver of [LX14] slowly converging to the same limit as the proposed receiver. At high rate systems (rate 5/6) the gap between them and our proposal increases, even for 4 turbo-iterations, and from Figure 3.34 it is seen that the receiver of [Tao16] cannot converge to the same asymptotic limits, probably due to the usage of only a few samples for covariance estimation heuristic. The use of EP-feedback instead of APP does not bring significant improvement for high-rates, or without turbo-iterations, but at medium and low rates, it allows for an additional asymptotic gain over 0.5 dB. However, the predictability of the EP feedback over a wider set of configurations (see Section 2.4.4) makes it a more attractive solution.

3.6.5 Discussion

This section discussed an original approach to the design of turbo DFE receivers with static filters, through the use of online prediction, based on semi-analytic performance prediction techniques as used in physical layer abstraction methods. Due to the lack of a closed-form solution for such receivers, various heuristics are used throughout the literature, but here we raise the question on the accuracy or optimality of such approaches.

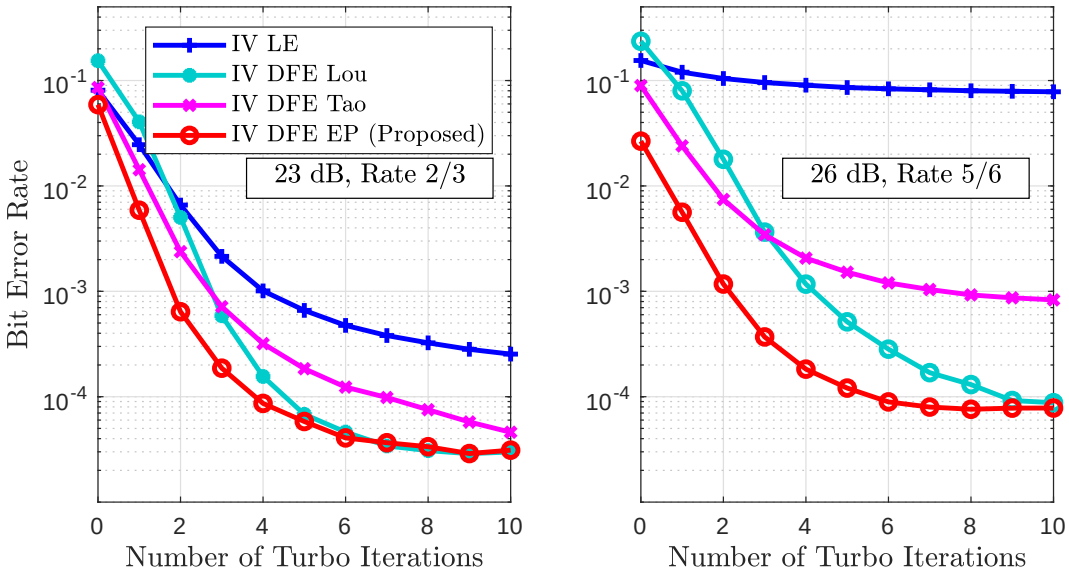


Figure 3.34: Coded 8-PSK BER performance across turbo-iterations.

Here, semi-analytical performance prediction of exact TV turbo DFE with dynamic filters is exploited to derive static DFE filters. This approach has been carried out for DFE with APP-based or EP-based soft feedback and their detection performance has been evaluated in various configurations. This framework could also be applied to self-iterated FIR DFE from Section 3.5 for further improved performance, by updating anti-causal variances with causal EP variance of the last self-iteration.

This receiver brings significant complexity savings with respect to TV DFE, while offering reasonably close performance. Moreover, our method is compatible with any constellation, and spectrally efficient on a large interval of coding rates and with or without turbo-iterations.

3.7 Conclusion

This chapter provides a very brief survey of the literature on turbo equalizers, by starting with the optimal SISO detector for ISI channel based on MAP criterion [BF98], equivalent to the originally derived ML turbo equalizer [DJB+95], which is an extension of the soft-output Viterbi algorithm. Next, the complexity reduction of the optimum detector is discussed, with a focus on filter-based MMSE turbo equalizers. In particular, we point out the importance of equalization architecture for the complexity-performance trade-off, and expose the limits of widely used Finite Impulse Response (FIR) linear MMSE equalizer, with Interference Cancellation (IC) with a priori information [TSK02]. Indeed, this equalizer and its extensions based on linear MMSE criterion are widely acclaimed as the main reference for turbo-equalization, even including 3GPP standardization performance evaluation reports [3gpc], and their DFE extensions did not receive a lot of attention. In particular, it is shown that the performance of the MMSE LE-IC is rather poor for at high spectral efficiency applications, remaining behind

the capabilities of the 40 year-old classical DFE [Bel+79].

Hence, to address this problem, the state-of-the art in FIR turbo-equalizers is investigated, and based on observed improvements and shortcomings, various novel DFE turbo-equalizers are proposed by incorporating the use of Expectation Propagation (EP). The performance-complexity trade-off brought by the proposed DFE receivers and their asymptotic performance are analysed, which favours solutions with EP-based soft feedback. Finally, we conclude that FIR DFE-IC structures are of interest when the target operating region on the the spectral-efficiency versus energy-efficiency plane corresponds to the use of high code rates. Otherwise, for robust codes, FIR LE-IC is sufficient to reach capacity achieving performance, but if faster convergence is needed, or the number of turbo-iterations has to be reduced, DFE-IC could become more attractive.

Chapter 4

A Framework for Frequency Domain Receiver Design with EP

Contents

4.1 Introduction

Although the detection complexity in ISI channels is significantly reduced when using FIR or block equalization, relative to the optimal detectors, such filtering architectures are still at the heart of receiver computational bottleneck. Among those structures, block receivers offer better performance but with a computational cost scaling at best quadratically in block length, and approximate FIR receivers have quadratic complexity in channel spread. Moreover, when considering DFEs, even with static filters, there would be a considerable amount of latency on the feedback from the soft decision device, which would require large buffers for handling this sequential process in the hardware. For exact turbo equalizers or TV channels, this latency would be higher when considering the underlying matrix inversion updates for filter computation, even with the low cost methods presented in Section 3.4.5.

To alleviate these issues, Frequency Domain (FD) receivers are used to transform the TD convolution operation into a single-tap filtering (element-wise multiplication of Channel Frequency Response (CFR)) in the FD, through the use of Fast Fourier Transform (FFT) and Inverse Fast Fourier Transform (IFFT), enabling low complexity receiver design. In particular in the 3GPP's LTE uplink, sidelink, D2D or vehicule-to-vehicle communications standards for 4G, SC-FDE, or its multi-user extension, Single-Carrier Frequency Division Multiple Access (SC-FDMA) have been successfully used for mitigating ISI in quasi-static wideband channels [SKJ94; Baz+17]. Moreover, various investigations for upcoming 5G uplink NOMA strategy with low-cost terminals also involve the use of SC-FDE [Liu+19a].

In this chapter, we discuss an original use of scalar EP message passing, along with self-iterated schedule and damped message computation principles for FD receiver design. This approach yields a doubly iterative receiver architecture which is shown to yield attractive performance-complexity trade-off, and we believe this to be the contribution of this thesis having the highest potential for impact in practical applications. Indeed, as single-carrier transmissions using FD receiving algorithms are widely used in technical specifications for various wireless systems standards such as LTE, proposing plausible extensions of those is an interesting development axis in the PHY layer. Existing low-cost linear MMSE receivers are a particular case of this structure, and self-iterations incorporate lower computational costs and lower latency compared to turbo-iterations, which makes proposed receivers a practical extension of the baseline. In summary, the contributions brought by this chapter are:

- a review on existing single-tap iterative FDE structures, in Section 4.2,
- proposal of a scalar EP framework for FD receiver design, in Section 4.3.2, explained through the derivation of an elementary FDE, and the analysis of its finite-length and

asymptotic performance, and complexity,

- optimization of the proposed structure through deep learning techniques, through the use of deep unfolding, in Section 4.4,
- extension and application of this receiver to various common SC waveforms, with Cyclic Prefix (CP), Zero Padding (ZP) or Unique Word (UW), also with upsampled waveforms with Fractionally-Spaced (FS) equalization, Frequency Domain Oversampling (FDO) or SC-FDMA, in Sections 4.5.1 and 4.5.2,
- FD equalization of time-varying channels is discussed in Section 4.5.3,
- extension of the framework to a multi-user and multi-antenna NOMA system in Section 4.6, through the use of Space-Time Bit Interleaved Coded Modulation (STBICM),
- discussion on the impact of practical channel estimation on the performance of these algorithms in Section 4.7.

4.2 On Iterative Frequency Domain Receivers

In this section frequency-domain filtering based receiver architectures for the linear channel model in Equation (2.1) is discussed, with a particular focus for the case of ISI mitigation, as in Equation (3.3). Extensions for handling Inter-Block Interference (IBI), Inter-Carrier Interference (ICI), Multiple-Antenna Interference (MAI) and Multiple-User Interference (MUI) will be discussed in the later sections of the chapter.

When the channel is quasi-static relative to a data block (i.e. $h_{k,l} = h_{k',l} = h_l$), the transmission is said to be *circular* if the Equation (3.2) can be rewritten as a circular convolution

$$y_k = \sum_{l=1}^{L} h_l x_{\langle k-l \rangle_K} + w_k, \qquad \forall k = 1, \ldots, K, \tag{4.1}$$

with $\mathbf{w} \sim \mathcal{CN}(\mathbf{0}_K, \mathbf{\Sigma}_\mathbf{w})$ and $\langle k \rangle_K \triangleq 1 + [k-1 \mod K]$, for performing modulo operation between 1 and K. Practical examples of how circular transmissions take place is discussed in the Section 4.5.1 of this chapter.

In this situation, if the statistics of prior estimates on data symbols is white (i.e. the reliability of prior estimates is static over the TD block, with $\mathbf{\Sigma}_\mathbf{x}$ in Equation (3.19) being a scaled identity matrix), then block MMSE LE is efficiently implemented via FFTs as the FD LE [WS73; SKJ94], with the so-called "one-tap" filters where frequency bins (also called sub-carriers) are processed in parallel. The major benefit in doing so is the avoidance of block matrix inversion of block filter-banks and reduce the computational complexity from the quadratic scaling (e.g. $\mathcal{O}(K^2)$) (at best), to quasi-linear scaling (e.g. $\mathcal{O}(K \log K)$). However, this comes with a limitation on the channel delay spread to the block length, i.e. $L \leqslant K$.

Frequency Domain Linear Equalizer In detail, when the circular time-domain ISI model of Equation (4.1) can be cast into the general linear channel model in Equation (2.1), the block observation model of Equation (3.3) becomes

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}, \tag{4.2}$$

where $\mathbf{y} \in \mathbb{C}^K$, $\mathbf{w} \sim \mathcal{C}(\mathbf{0}_{K\times 1}, \mathbf{\Sigma}_\mathbf{w})$ and by denoting $\mathbf{h} = [h_1, \ldots, h_L]$, the sparse $K \times K$ circulant matrix $\mathbf{H}$ is given by

$$\mathbf{H} = \mathbf{Circ}([\mathbf{h}; \mathbf{0}_{(K-L)\times 1}]) \triangleq \begin{bmatrix} h_1 & 0 & \ldots & h_L & \ldots & h_3 & h_2 \\ h_2 & h_1 & 0 & \ldots & h_L & \ldots & h_3 \\ h_3 & h_2 & h_1 & 0 & \ldots & h_L & \vdots \\ & & \ddots & \ddots & \ddots & & \\ & h_L & \ldots & h_3 & h_2 & h_1 & 0 \\ 0 & \ldots & h_L & \ldots & h_3 & h_2 & h_1 \end{bmatrix}. \tag{4.3}$$

Consequently, the channel matrix has some attractive algebraic properties through the use of the K-point DFT matrix $\mathcal{F}_K$, with $\underline{\mathbf{H}} \triangleq \mathcal{F}_K \mathbf{H} \mathcal{F}_K^H = \mathbf{Diag}(\underline{\mathbf{h}})$ where $\underline{\mathbf{h}} = [\underline{h}_1; \ldots; \underline{h}_K] = \sqrt{K}\mathcal{F}_K[\mathbf{h}; \mathbf{0}_{K-L\times 1}]$ is the K-point CFR with

$$\underline{h}_k = \sum_{l=1}^{L} h_l \exp(-2j\pi(k-1)(l-1)/K), \quad k = 1, \ldots, K. \tag{4.4}$$

Then an equivalent Frequency Domain (FD) transmission model is given by

$$\underline{\mathbf{y}} = \underline{\mathbf{H}}\underline{\mathbf{x}} + \underline{\mathbf{w}}, \tag{4.5}$$

with $\underline{\mathbf{x}} = \mathcal{F}_K \mathbf{x}$, $\underline{\mathbf{y}} = \mathcal{F}_K \mathbf{y}$, $\underline{\mathbf{w}} = \mathcal{F}_K \mathbf{w}$.

Recalling the MMSE block LE, given in Equation (3.19), the estimated data symbols are

$$\hat{\mathbf{x}} \triangleq \mathbf{F}^{\text{ble}H}\mathbf{y} + \mathbf{g}^{\text{ble}} = \boldsymbol{\mu}_\mathbf{x} + \mathbf{\Xi}^{\text{ble}-1}\mathbf{H}^H\mathbf{\Sigma}^{\text{ble}-1}(\mathbf{y} - \mathbf{H}\boldsymbol{\mu}_\mathbf{x}). \tag{4.6}$$

where $\boldsymbol{\mu}_\mathbf{x} = \mathbb{E}[\mathbf{x}]$, $\mathbf{\Sigma}_\mathbf{x} = \sigma_x^2 \mathbf{I}_K$, with $\sigma_x^2 = \text{Var}[x_k], \forall k$ (i.e. white/uniform statistics) and with the assumption of $\mathbf{x}$ being uncorrelated, $\mathbf{\Sigma}^{\text{ble}} = \mathbf{\Sigma}_\mathbf{w} + \sigma_x^2 \mathbf{H}\mathbf{H}^H$ and $\mathbf{\Xi}^{\text{ble}} = \mathbf{Diag}(\xi_1^{\text{ble}}, \ldots, \xi_K^{\text{ble}})$ with $\xi_k^{\text{ble}} = [\mathbf{H}^H\mathbf{\Sigma}^{\text{ble}-1}\mathbf{H}]_{k,k}$. These equations are simplified through the use of the FD channel matrix's properties, and by neglecting non-diagonal elements of $\mathbf{\Sigma}_\mathbf{w}$ and $\mathbf{\Xi}^{\text{ble}}$

$$\begin{aligned} \hat{\mathbf{x}} &= \mathcal{F}_K^H \left(\mathcal{F}_K \boldsymbol{\mu}_\mathbf{x} + \mathcal{F}_K \mathbf{\Xi}^{\text{ble}-1} \mathcal{F}_K^H \mathcal{F}_K \mathbf{H}^H \mathcal{F}_K^H \mathcal{F}_K \mathbf{\Sigma}^{\text{ble}-1} \mathcal{F}_K^H (\mathcal{F}_K \mathbf{y} - \mathcal{F}_K \mathbf{H} \mathcal{F}_K^H \mathcal{F}_K \boldsymbol{\mu}_\mathbf{x}) \right), \\ &\approx \mathcal{F}_K^H \left(\underline{\boldsymbol{\mu}}_\mathbf{x} + \xi^{\text{fde}-1} \underline{\mathbf{H}}^H (\bar{\sigma}_w^2 \mathbf{I}_N + \sigma_x^2 \underline{\mathbf{H}}\underline{\mathbf{H}}^H)^{-1} (\underline{\mathbf{y}} - \underline{\mathbf{H}}\underline{\boldsymbol{\mu}}_\mathbf{x}) \right), \end{aligned} \tag{4.7}$$

where $\underline{\boldsymbol{\mu}}_\mathbf{x} \triangleq \mathcal{F}_K \boldsymbol{\mu}_\mathbf{x}$ are FD prior data estimates, $\bar{\sigma}_w^2 = K^{-1} \sum_{k=1}^K \sigma_{w,k}^2$ is the variance of noise and interference in the frequency domain, $\xi^{\text{fde}} = K^{-1} \sum_{k=1}^K |\underline{h}_k|^2 / (\bar{\sigma}_w^2 + \sigma_x^2 |\underline{h}_k|^2)$. Equivalently, this equalization process can be written in a parallelized form for the FD estimates

$$\underline{\hat{x}}_k = \underline{\mu}_{x,k} + \xi^{\text{fde}-1} \underline{h}_k^* (\bar{\sigma}_w^2 + \sigma_x^2 |\underline{h}_k|^2)^{-1} (\underline{y}_k - \underline{h}_k \underline{\mu}_{x,k}). \tag{4.8}$$

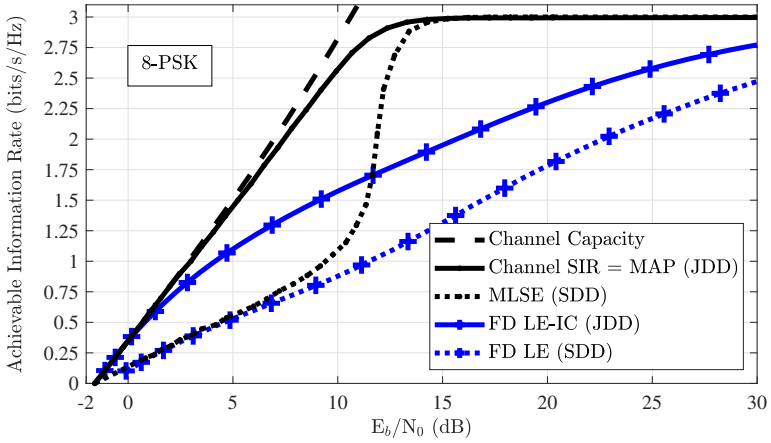


Figure 4.1: Achievable rates of BCJR, FD LE-IC and classical FD-LE in Proakis C.

This structure replaces large matrix inversions of the time-domain block receiver to element-wise scalar affine operations in the frequency domain, which reduces the computational bottleneck to the use of FFTs.

There is a long research track on FDE, starting from very the low-complexity linear FDE above, up to non-linear FD turbo equalizers. Table 4.1 lists chronological milestones on developments regarding how Interference Cancellation (IC) with either decoder or decision (demapper) feedback is used. The "schedule" column indicates in which manner the decoder/demapping feedback is used by the equalizer, with "parallel" meaning block-wise feedback, and "serial" referring to a symbol-wise feedback (like FIR DFE-IC). The position of the novel FDE that will be derived in Section 4.3.2 is also shown in the table.

4.2.1 From Block LE-IC to Single-Tap Turbo Equalization

The FD receiver structure has been naturally extended to turbo equalization following the Wang-Poor approach [WP99; TSK02], discussed in Chapter 3, by whitening the estimates used for block LE-IC (following the Equation (3.43) for computing v_x^p, as for IV FIR structure) [TH01; TS11]. The resulting receiver is an EXT feedback based receiver, denoted FD LE-EXTIC, and it is given by

$$\underline{x}_k^e = \underline{x}_k^p + \xi^{\text{fdle-ic}\,-1}\,\underline{f}_k^{\text{fdle-ic}*}(\underline{y}_k - \underline{h}_k\underline{x}_k^p), \tag{4.9}$$

where $\underline{f}_k^{\text{fdle-ic}*} = \xi^{\text{fdle-ic}\,-1}(\bar{\sigma}_w^2 + v_x^p|\underline{h}_k|^2)^{-1}\underline{h}_k$, with $\xi^{\text{fdle-ic}} = K^{-1}\sum_{k=1}^{K}|\underline{h}_k|^2/(\sigma_{w,k}^2 + v_x^p|\underline{h}_k|^2)$, and v_x^p is the variance of the SISO decoder's extrinsic PMF, defined in Equations (2.39)-(2.40). This is a simplified variant of block turbo MMSE LE-IC, where the exact prior data covariance matrix $\mathbf{Diag}([v_{x,1}^p;\ldots;v_{x,K}^p])$ is replaced by $v_x^p\mathbf{I}_K$. As stated for IV FIR equalizers, although this simplification slightly reduces the achievable rates of the receiver, it does not necessarily mean a loss in finite length performance, as exact structures are more sensible to errors when the SISO decoder erroneously converges towards a bad codeword.

Table 4.1: Developments Related to Iterative Single-tap FDE.

Reference	Contribution on single-tap FDE	Decoder Feedback	Decision Feedback	Schedule
[WS73]	Initially proposed frequency domain equalization.	-	-	-
[SKJ94]	Revived FDE through the comparison of SC-FDE relative to multicarrier signalling.	-	-	-
[TH01]	Derived the turbo iterated FD LE-IC. Denoted as FD LE-Extrinsic (EXT)IC.	EXT	-	Parallel/-
[Fal+02] [BT02]	A hybrid DFE with a frequency domain feedforward filter and a time domain feedback filter is proposed.	-	Hard	-/Serial
[KSW03]	Simplifies hybrid DFE by using noise prediction.	-	Hard	-/Serial
[BT05]	Proposed a non-linear receiver with FD feedforward and feedback filters, called iterative block DFE IBDFE.	-	Hard/APP	-/Parallel
[VBC06]	Proposed a turbo FD MIMO receiver. It uses APP estimates from the decoder instead of extrinsic. Denoted as FD LE-APPIC.	APP	-	Parallel/-
[NLF07]	Compared FDE with FD feedforward and TD/FD feedback filters. TD/FD are equivalent (parallel schedule). Soft better than hard.	EXT	Hard/APP	Parallel/- or -/Parallel
[GM08]	Proposed a self-iterated BPSK turbo receiver with APP estimates, based on PDA.	EXT	APP	Par./Par.
[Guo+13]	Derived a self-iterated turbo receiver based on GAMP, that exploits APP estimates at each turbo iteration.	EXT	APP	Par./Par.
[Che+15]	Equivalence of coded IBDFE to FD LE-EXTIC is shown.	-	-	-
[Tao15]	Extended results in [TH01; BT05; GM08] to a turbo FDE with a APP-based self-iteration, denoted as FD SILE-APPIC.	EXT	APP	Par./Par.
Section 4.3.2	Proposes a self-iterated FD LE-IC with EP.	EXT	EXT	Par./Par.

When quantifying the achievable rates of FD LE-EXTIC, it is seen that this receiver achieves the channel SIR for low coding rates, as its FIR counterpart in Section 3.3.3. However, despite the improvements brought by the turbo-iterations, there is a significant gap between its achievable rates and the channel SIR, especially in moderately or highly selective channels, at high spectral efficiency operating points. These observations are illustrated in Figure 4.1, where the achievable rates for 8-PSK are compared between optimal turbo receiver (MAP), optimal one-shot detector (MLSE) and turbo and non-turbo FD-LE, for the Proakis C channel. Note that the gap of the baseline turbo FDE receiver to the SIR is very significant in this channel, as the spectral efficiency increases. Indeed, for rates lower than

0.75 bits/s/Hz (which translates into code rates lower than 1/4), conventional FD-LE is very close to the optimal receiver, however at 2 bits/s/Hz (code rate of 2/3) there is a gap over 7 dB for the turbo FD-LE to SIR, and about 10 dB for the non-iterative FDE to the MLSE rates.

Consequently, we will discuss below non-linear extensions of the FD LE that have been explored to improve performance for high-spectral efficiency operating points [Fal+02; BT02; KSW03; BT05; VBC06; NLF07; GM08; Tao15].

An improvement upon this well-known turbo FDE reference is given in [VBC06], which uses APP LLRs as feedback from the decoder, instead of EXT LLRs. This is based on the observed improvement in turbo detection in [Wit+02] by the use of APP feedback, and this FDE structure is denoted FD LE-APPIC. Nevertheless, APP feedback violates the independence principle of turbo iterative systems [DJB+95], so theoretical background for such structures is absent, and its asymptotic performance is non-trivial to predict [BC02].

4.2.2 On Decision Feedback in the Frequency Domain

Independently of the emerging turbo equalization literature, given that TD block DFE structures outperformed block LE [Kal95], derivation of non-linear FDE was of interest. In particular, a hybrid implementation of block DFE was carried out in [Fal+02; BT02]. This structure uses a FD feedforward filter and a TD feedback filterbank, which carries out symbol-wise, i.e. serial, IC with hard decisions. The use of noise prediction in [KSW03], simplifies the computation of hybrid DFE, by forcing the feedforward filter to be the same as the FD LE filter, while the overall structure remained equivalent to block DFE, as briefly mentioned in Chapter 3.

In [BT05], the frequency domain feedback concept was introduced, and denoted Iterative Block DFE (IBDFE). This structure uses decision feedback in a blockwise, parallel schedule, allowing the use of FFTs over feedback symbol block, and significantly reducing complexity. Despite its name, this structure is a LE-IC, with the decision feedback being used for interference cancellation, and it is not related to the TD block DFE in [Kal95]. Indeed, the TD block DFE of [Kal95] uses serial symbol-wise hard decision feedback via a fairly complicated feedback filterbank, and thus it is unrelated to the linear IC scheme of [BT05]. In [NLF07], variations of IBDFE were evaluated with hard or soft APP, and TD or FD feedback. It is noted in [NLF07; BDF+10] that when used with forward error correction in the feedback loop, this structure is equivalent to FD LE-EXTIC.

In [GM08], PDA is used to derive a non-linear FDE for BPSK, through a self-iterated MMSE LE-IC using APP feedback from previous detections, before computing extrinsic LLRs for decoding. This structure, and IBDFE [BT05; NLF07; Che+15] were later extended to generalized constellations in [Tao15]. In the latter work, non-linear block FDEs were evaluated, using APP decision feedback with serial and parallel schedules. These results are then used to derive a single-tap FD self-iterated LE-IC with an initial IC carried out with EXT feedback from the decoder, followed by a second round of IC carried out with APP feedback

from the detector. Here, this structure is denoted as FD SILE-APPIC.

4.2.3 Approximate Inference for Frequency Domain Receivers

Following the development of FD SILE-IC with APP estimates, thanks to PDA in [GM08], an alternative APP-based iterative FDE were derived through GAMP [Guo+13]. This structure benefits from the use of an Onsager reaction term (see Section 2.4.1) for compensating the impact of APP estimates' correlations, which improves the convergence of the algorithm.

Another category of iterative FD receivers, based on approximate inference techniques, is given in [Zha+15a] with a hybrid BP-MF framework [Rie+13]. In these receivers MF is used for handling channel observations, and BP for demapping, in an iterative way, and this yields an APP-based IC receiver, with similarities to iterative thresholding techniques, with the use of a matched-filter as the linear component. When it is used in a parallel schedule it is shown to remain behind the capabilities of the GAMP-based FDE, but with a serial scheduling it can close the gap or even outperform it if a noise estimation step is included.

Finally, Expectation Propagation (EP) has been used in FD turbo detection in [Zha+16; WRM+17] mainly for the mitigation of inter-band interference, between multiple users. The former reference uses it for a generalized frequency division multiplexing receiver, as an iterative block receiver, with cubic complexity in block length. That structure is extended for SC-FDMA in [WRM+17] under the acronym of joint-EP (J-EP). The latter reference also includes a single-tap simplification of that receiver, denoted distributed-EP (D-EP), which was however obtained through a zero-forcing type derivation, which makes it severely vulnerable to spectral nulls [WRM+17, eq.(48)]. These turbo receivers do not perform multiple inner iterations for the detection step.

4.3 Scalar EP for Doubly Iterative FD Turbo Equalization

The literature review above shows that although several doubly-iterative turbo FDE attempts have been explored in the literature, it has been only limited to APP feedback related techniques (PDA, GAMP, BP-MF). These approaches are suboptimal and harder to predict than techniques based on extrinsic feedback, which makes the exploration of EP-based approaches an attractive direction for improvement, based on conclusions in Section 2.4.1. Moreover, in the previous chapter, the derivation of a self-iterated FIR receiver enabled improving the reachable rates of filter-based turbo equalizers with respect to the channel SIR, with the use of the diagonal EP technique (Section 2.4.2) for FIR, in Section 3.5.

In this section we will discuss the use of scalar EP message passing for deriving a similar receiver but with a FD single-tap equalization architecture.

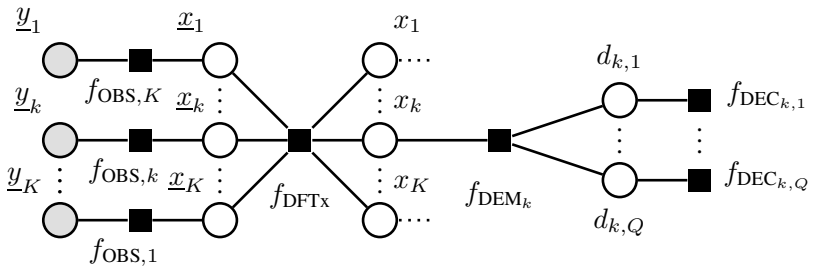


Figure 4.2: Factor graph for the posterior density function in Equation (4.10).

Table 4.2: Posteriors and messages of variables nodes.

Description	Notation	PDF
Posterior at OBS_k	$q_{\text{OBS}_k}(\underline{x}_k)$	$\mathcal{CN}(\underline{x}_k; \mu^o_{\underline{x},k}, \gamma^o_{\underline{x},k})$
Extrinsic from OBS_k	$m_{\text{OBS}_k \to \underline{x}_k}(\underline{x}_k)$	$\mathcal{CN}(\underline{x}_k; \underline{x}^o_k, v^o_{\underline{x},k})$
Posterior at DFTx	$q_{\text{DFTx}}(\mathbf{x})$	$\mathcal{CN}(\mathbf{x}; \boldsymbol{\mu}^{\mathbf{e}}_{\mathbf{x}}, \gamma^e_x \mathbf{I}_K)$
	$q_{\text{DFTx}}(\underline{\mathbf{x}})$	$\mathcal{CN}(\underline{\mathbf{x}}; \boldsymbol{\mu}^{\mathbf{e}}_{\underline{\mathbf{x}}}, \boldsymbol{\Gamma}^{\mathbf{e}}_{\underline{\mathbf{x}}})$
Extrinsic from DFTx	$m_{\text{DFTx} \to \mathbf{x}}(\mathbf{x})$	$\mathcal{CN}(\mathbf{x}; \mathbf{x}^{\mathbf{e}}, v^e_x \mathbf{I}_K)$
	$m_{\text{DFTx} \to \underline{\mathbf{x}}}(\underline{\mathbf{x}})$	$\mathcal{CN}(\underline{\mathbf{x}}; \underline{\mathbf{x}}^{\mathbf{e}}, \mathbf{V}^{\mathbf{e}}_{\underline{\mathbf{x}}})$
Posterior at DEM_k	$q_{\text{DEM}_k}(x_k)$	$\mathcal{CN}(x_k; \mu^d_{x,k}, \gamma^d_x)$
Extrinsic from DEM_k	$m_{\text{DEM}_k \to x_k}(x_k)$	$\mathcal{CN}(x_k; x^d_k, v^d_x)$

4.3.1 On Factor and Variable Node Assumptions for Single-Tap FDE

By reconsidering the factor graph of $p(\mathbf{x}, \mathbf{d}|\underline{\mathbf{y}})$, for the generic detection problem in the Section 2.3.5, the EQU factor node that uses time-domain observations can be replaced by FD observations following Equation (4.5). Then a new graph is obtained with

$$p(\mathbf{x}, \mathbf{d}|\underline{\mathbf{y}}) \propto \left(\prod_{k=1}^{K} \underbrace{p(\underline{y}_k|\underline{x}_k)}_{=f_{\text{OBS}_k}(\underline{x}_k)} \right) \underbrace{p(\underline{\mathbf{x}}|\mathbf{x})}_{=f_{\text{DFTx}}(\mathbf{x},\underline{\mathbf{x}})} \prod_{k=1}^{K} \underbrace{p(x_k|\mathbf{d}_k)}_{=f_{\text{DEM}_k}(x_k,\mathbf{d}_k)} \prod_{q=1}^{Q} \underbrace{p(d_{k,q})}_{=f_{\text{DEC}_{k,q}}(d_{k,q})}, \tag{4.10}$$

where the FD symbols $\underline{x}_k$ are introduced as intermediary unconstrained variables in $\mathbb{C}$, as illustrated in Figure 4.2. More specifically the EQU factor node considered in the generic case is split into K symbolwise FD observation factors $\{\text{OBS}_k\}_{k=1}^K$, with

$$f_{\text{OBS}_k}(\underline{x}_k) \propto \exp\left(-|\underline{y}_k - \underline{h}_k \underline{x}_k|^2 / \bar{\sigma}^2_w\right), \tag{4.11}$$

and a transform factor node DFTx bring hard constraints on data vectors

$$f_{\text{DFTx}}(\mathbf{x}, \underline{\mathbf{x}}) = \delta(\underline{\mathbf{x}} - \mathcal{F}_K \mathbf{x}) = \delta(\mathbf{x} - \mathcal{F}^H_K \underline{\mathbf{x}}). \tag{4.12}$$

Diagonal EP is used for handling Frequency Domain (FD) data symbols $\underline{\mathbf{x}}$, similar to

the Section 2.4.2, and *scalar* EP message passing is used for handling TD data symbols $\mathbf{x}$, to benefit from the low complexity of single-tap FDE structures. To this end, for the proposed framework, our simplifying assumption is that variable nodes $\underline{\mathbf{x}}$ lie in multivariate white Gaussian distributions. Hence, a message involving these variable nodes is fully characterized by a *vector mean* and a *scalar variance*. Then along with the remainder of the self-iterated approach in Section 2.4.2, these assumptions yield the distributions in Table 4.2, where covariance matrices $\underline{\mathbf{\Gamma}}$ and $\underline{\mathbf{V}}$ are diagonals with components $\underline{v}_{\cdot,k}$ which ensures that the covariance matrices of variables $\underline{\mathbf{x}}$ and $\mathbf{x}$ have respectively an unconstrained diagonal structure, and a matrix proportional to the identity.

Following Expectation Propagation (EP) message-passing procedure in Section 2.4.1, at the OBS_k factor node, the pre-projection posterior on variable $\underline{x}_k$ is given by

$$\begin{aligned}\tilde{q}_{\mathrm{OBS}_k}(\underline{x}_k) &= f_{\mathrm{OBS}_k}(\underline{x}_k) m_{\mathrm{DFTx}_k \to \underline{x}_k}(\underline{x}_k), \\ &\propto \exp\left(-\frac{|\underline{y}_k - \underline{h}_k \underline{x}_k|^2}{\bar{\sigma}_w^2} - \frac{|\underline{x}_k - \underline{x}_k^e|^2}{v_{\underline{x},k}^e}\right) \propto \mathcal{CN}(\underline{x}_k; \mu_{\underline{x},k}^o, \gamma_{\underline{x},k}^o), \end{aligned} \tag{4.13}$$

$$\gamma_{\underline{x},k}^o = \left(\frac{|\underline{h}_k|^2}{\bar{\sigma}_w^2} + \frac{1}{v_{\underline{x},k}^e}\right)^{-1} = \frac{\bar{\sigma}_w^2 v_{\underline{x},k}^e}{\bar{\sigma}_w^2 + v_{\underline{x},k}^e |\underline{h}_k|^2}, \tag{4.14}$$

$$\mu_{\underline{x},k}^o = \frac{\bar{\sigma}_w^2 v_{\underline{x},k}^e}{\bar{\sigma}_w^2 + v_{\underline{x},k}^e |\underline{h}_k|^2} \left(\frac{\underline{h}_k^* \underline{y}_k}{\bar{\sigma}_w^2} + \frac{\underline{x}_k^e}{v_{\underline{x},k}^e}\right) = \frac{\bar{\sigma}_w^2 \underline{x}_k^e + v_{\underline{x},k}^e \underline{h}_k^* \underline{y}_k}{\bar{\sigma}_w^2 + v_{\underline{x},k}^e |\underline{h}_k|^2}. \tag{4.15}$$

This distribution on $\underline{\mathbf{x}}$ already belongs to the family of Gaussian distributions with diagonal covariance matrices, hence diagonal EP can be applied with $q_{\mathrm{OBS}_k}(\underline{x}_k) = \tilde{q}_{\mathrm{OBS}_k}(\underline{x}_k)$, and the extrinsic messages from OBS_k are

$$m_{\mathrm{OBS}_k \to \underline{x}_k}(\underline{x}_k) = \frac{q_{\mathrm{OBS}_k}(\underline{x}_k)}{m_{\mathrm{DFTx}_k \to \underline{x}_k}(\underline{x}_k)} \propto \frac{\mathcal{CN}(\underline{x}_k; \mu_{\underline{x},k}^o, \gamma_{\underline{x},k}^o)}{\mathcal{CN}(\underline{x}_k; \underline{x}_k^e, v_{\underline{x},k}^e)} \propto \mathcal{CN}(\underline{x}_k; \underline{x}_k^o, v_{\underline{x},k}^o), \tag{4.16}$$

$$v_{\underline{x},k}^o = \frac{\bar{\sigma}_w^2}{|\underline{h}_k|^2}, \tag{4.17}$$

$$\underline{x}_k^o = \frac{\underline{h}_k^* \underline{y}_k}{|\underline{h}_k|^2}. \tag{4.18}$$

At the factor node DFTx, the pre-projection posterior on variable $\underline{\mathbf{x}}$ is given by

$$\begin{aligned}\tilde{q}_{\mathrm{DFTx}}(\underline{\mathbf{x}}) &= \int_{\mathbf{x}} f_{\mathrm{DFTx}}(\mathbf{x}, \underline{\mathbf{x}}) \prod_{k=1}^{K} m_{\mathrm{DEM}_k \to x_k}(x_k) m_{\mathrm{OBS}_k \to \underline{x}_k}(\underline{x}_k) \, \mathrm{d}\mathbf{x}, \\ &\propto \mathcal{CN}(\mathcal{F}_K^H \underline{\mathbf{x}}; \mathbf{x}^{\mathbf{d}}, v_x^d \mathbf{I}_K) \mathcal{CN}(\underline{\mathbf{x}}; \underline{\mathbf{x}}^{\mathbf{o}}, \mathbf{V}_{\underline{\mathbf{x}}}^{\mathbf{o}}), \\ &\propto \mathcal{CN}(\underline{\mathbf{x}}; \boldsymbol{\mu}_{\underline{\mathbf{x}}}^{\mathbf{e}}, \mathbf{\Gamma}_{\underline{\mathbf{x}}}^{\mathbf{e}}), \end{aligned} \tag{4.19}$$

$$\mathbf{\Gamma}_{\underline{\mathbf{x}}}^{\mathbf{e}} = v_x^d \mathbf{V}_{\underline{\mathbf{x}}}^{\mathbf{o}} (v_x^d \mathbf{I}_K + \mathbf{V}_{\underline{\mathbf{x}}}^{\mathbf{o}})^{-1}, \tag{4.20}$$

$$\boldsymbol{\mu}_{\underline{\mathbf{x}}}^{\mathbf{e}} = (v_x^d \mathbf{I}_K + \mathbf{V}_{\underline{\mathbf{x}}}^{\mathbf{o}})^{-1} \left(\mathbf{V}_{\underline{\mathbf{x}}}^{\mathbf{o}} \underline{\mathbf{x}}^{\mathbf{d}} + v_x^d \underline{\mathbf{x}}^{\mathbf{o}}\right), \tag{4.21}$$

where $\boldsymbol{\mu}^{\mathbf{e}}_{\underline{\mathbf{x}}} = [\mu^e_{\underline{x},1}; \ldots; \mu^e_{\underline{x},K}]$ and $\boldsymbol{\Gamma}^{\mathbf{e}}_{\underline{\mathbf{x}}} = [\gamma^e_{\underline{x},1}; \ldots; \gamma^e_{\underline{x},K}]$ and $\underline{\mathbf{x}}^{\mathbf{d}} \triangleq \mathcal{F}_K \mathbf{x}^{\mathbf{d}}$. This distribution readily belongs to the family of Gaussian distributions with diagonal covariance matrices, hence diagonal EP directly yields the extrinsic messages from DFTx to $\underline{x}_k$ with

$$m_{\mathrm{DFTx}\to\underline{x}_k}(\underline{x}_k) \propto \mathcal{CN}(\underline{x}_k; \underline{x}^e_k, v^e_{\underline{x}}), \tag{4.22}$$

$$v^e_{\underline{x},k} = v^d_x, \tag{4.23}$$

$$\underline{x}^e_k = \underline{x}^d_k = \mathbf{e}_k \mathcal{F}_K \mathbf{x}^{\mathbf{d}}. \tag{4.24}$$

Still considering the same factor node, the pre-projection posterior on variable $\mathbf{x}$ is similarly obtained by

$$\begin{aligned}
\tilde{q}_{\mathrm{DFTx}}(\mathbf{x}) &= \int_{\underline{\mathbf{x}}} f_{\mathrm{DFTx}}(\mathbf{x}, \underline{\mathbf{x}}) \prod_{k=1}^{K} m_{\mathrm{DEM}_k\to x_k}(x_k) m_{\mathrm{OBS}_k\to\underline{x}_k}(\underline{x}_k)\, \mathrm{d}\underline{\mathbf{x}}, \\
&\propto \mathcal{CN}(\mathbf{x}; \mathbf{x}^{\mathbf{d}}, v^d_x \mathbf{I}_K) \mathcal{CN}(\mathcal{F}_K \mathbf{x}; \underline{\mathbf{x}}^{\mathbf{o}}, \mathbf{V}^{\mathbf{o}}_{\underline{\mathbf{x}}}), \\
&\propto \mathcal{CN}(\mathbf{x}; \boldsymbol{\mu}^{\mathbf{e}}_{\mathbf{x}}, \boldsymbol{\Gamma}^{\mathbf{e}}_{\mathbf{x}}),
\end{aligned} \tag{4.25}$$

$$\boldsymbol{\Gamma}^{\mathbf{e}}_{\mathbf{x}} = v^d_x \mathbf{V}^{\mathbf{o}}_{\underline{\mathbf{x}}} (v^d_x \mathbf{I}_K + \mathbf{V}^{\mathbf{o}}_{\underline{\mathbf{x}}})^{-1}, \tag{4.26}$$

$$\boldsymbol{\mu}^{\mathbf{e}}_{\mathbf{x}} = \mathcal{F}^H_K (v^d_x \mathbf{I}_K + \mathbf{V}^{\mathbf{o}}_{\underline{\mathbf{x}}})^{-1} \left(\mathbf{V}^{\mathbf{o}}_{\underline{\mathbf{x}}} \underline{\mathbf{x}}^{\mathbf{d}} + v^d_x \underline{\mathbf{x}}^{\mathbf{o}} \right), \tag{4.27}$$

$\boldsymbol{\mu}^{\mathbf{e}}_{\mathbf{x}} = [\mu^e_{x,1}; \ldots; \mu^e_{x,K}]$ and $\boldsymbol{\Gamma}^{\mathbf{e}}_{\mathbf{x}} = [\gamma^e_{x,1}; \ldots; \gamma^e_{x,K}]$. This distribution readily belongs to the family of Gaussian distributions with diagonal covariance matrices and it needs to be projected into the family of Gaussian distributions with diagonal-restricted ("whitened") covariance matrices. To this end, it is needed to find $\gamma^e_x = \gamma^e_{x,k}$, for all $k = 1, \ldots, K$, and an approximate solutions to this overdetermined system of equations is given by the least-squares solution, which coincides with sample average

$$\gamma^e_x = \frac{1}{K} \sum_{k=1}^{K} \gamma^e_{x,k} = \frac{1}{K} \sum_{k=1}^{K} \frac{v^d_x v^o_{x,k}}{v^o_{x,k} + v^d_x} = v^d_x (1 - \xi_x v^d_x), \tag{4.28}$$

where $\xi_x \triangleq K^{-1} \sum_{k=1}^{K} (v^o_{x,k} + v^d_x)^{-1}$. Then $q_{\mathrm{DFTx}}(\mathbf{x}) \propto \mathcal{CN}(\mathbf{x}; \boldsymbol{\mu}^{\mathbf{e}}_{\mathbf{x}}, \gamma^e_x \mathbf{I}_K)$, and scalar EP is used to compute the extrinsic messages from DFTx to x_k with

$$m_{\mathrm{DFTx}\to x_k}(x_k) \propto \frac{\mathcal{CN}(x_k; \mu^e_{x,k}, \gamma^e_x)}{\mathcal{CN}(x_k; x^d_k, v^d_{x,k})} \propto \mathcal{CN}(x_k; x^e_k, v^e_x), \tag{4.29}$$

$$v^e_x = 1/\xi_x - v^d_x, \tag{4.30}$$

$$x^e_k = \mathbf{e}_k \mathcal{F}^H_K \left[\underline{\mathbf{x}}^{\mathbf{d}} + \xi^{-1}_x (\mathbf{V}^{\mathbf{o}}_{\underline{\mathbf{x}}} + v^d_x \mathbf{I}_K)^{-1} (\underline{\mathbf{x}}^{\mathbf{o}} - \underline{\mathbf{x}}^{\mathbf{d}}) \right]. \tag{4.31}$$

It can be seen that these expressions have some similarities with a MMSE FDE receiver, in the next section this will be shown explicitly by defining a message passing schedule for deriving a receiver with this framework.

Finally, the computation of messages from DEM_k to x_k is considerably different than the computations in Section 2.4.2, as the projection of messages has to be made through scalar

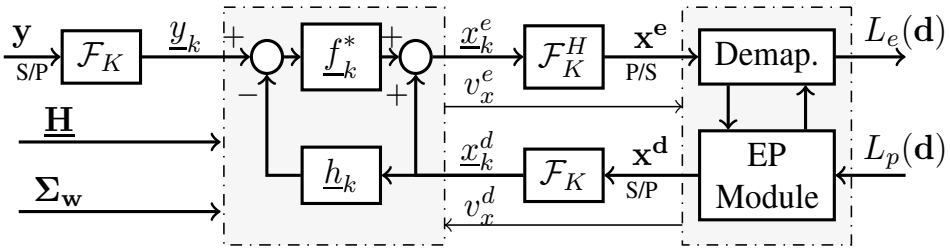


Figure 4.3: Proposed turbo FD SILE-EPIC structure.

EP. This involves projecting $\mathcal{CN}(\mathbf{x}; \boldsymbol{\mu}_{\mathbf{x}}^{\mathbf{d}}, \boldsymbol{\Gamma}_{\mathbf{x}}^{\mathbf{d}})$, from Equation (2.98), with $\boldsymbol{\mu}_{\mathbf{x}}^{\mathbf{d}} = [\mu_{x,1}^d; \ldots; \mu_{x,K}^d]$ and $\boldsymbol{\Gamma}_{\mathbf{x}}^{\mathbf{d}} = \mathbf{Diag}([\gamma_{x,1}^d; \ldots; \gamma_{x,K}^d])$, and the resulting distribution is $\mathcal{CN}(\mathbf{x}; \boldsymbol{\mu}_{\mathbf{x}}^{\mathbf{d}}, \gamma_x^d \mathbf{I}_K)$, which is, similarly to the Equation(4.28), given by

$$\gamma_x^d = \frac{1}{K} \sum_{k=1}^{K} \gamma_{x,k}^d. \tag{4.32}$$

Then, unlike in the Equation (2.100) for diagonal EP, the "Gaussian division" in scalar EP for the computation of mean and the variance of extrinsic messages $m_{\mathrm{DEM}_k \to x}(x_k)$ yields

$$x_k^* = \frac{\mu_{x,k}^d v_x^e - x_k^e \gamma_x^d}{v_x^e - \gamma_x^d}, \text{ and, } v_x^* = \frac{v_x^e \gamma_x^d}{v_x^e - \gamma_x^d}, \tag{4.33}$$

with the heuristic of using APP estimates for negative variance occurrences

$$(x_k^d, v_x^d) = \begin{cases} (x_k^*, v_x^*) & \text{if } v_x^e > \gamma_x^d + \varepsilon, \\ (\mu_{x,k}^d, \gamma_x^d) & \text{otherwise,} \end{cases} \tag{4.34}$$

where ε is a small constant, as recommended by [Sen+11].

This extrinsic message is one of the major aspects of our contribution; as the estimation variance is assumed to be static over a data block, these estimates are computationally less intensive. Moreover, due to the averaged variance, the occurrences of $v_x^e \leqslant \gamma_x^d$ is less likely than the occurrences of $v_{x,k}^e \leqslant \gamma_{x,k}^d$ (from diagonal EP), which improves the detector robustness to erroneous feedback, and there are improvements in decoding performance due to the reduced sensitivity to decoder error propagation (especially for weak codes).

4.3.2 FD Self-Iterated LE-IC based on EP (FD SILE-EPIC)

As the considered factor graph has cycles, and it is a finite graph, the scheduling matters, the convergence speed of the iterative algorithm and the possible error floors will depend on the schedule which coordinates the update of variable and factor nodes. To keep the equalization complexity reasonable, a parallel scheduling across variables nodes $\underline{x}_k$ and x_k is considered, in line with conventional FD LE or block LE receivers. Note that the use of a serial schedule would yield a DFE-like structure as in Section 3.4.4.

The parallel schedule across variables $\underline{x}_k$ and x_k enable us to simplify messages by combining the updates of DFTx node and $\{\text{OBS}_k\}_{k=1}^K$ nodes[1]. In the remainder of this section, for the sake of simplifying the readability of our proposal, we *redefine* $\underline{x}_k^e$ as the k^{th} subcarrier of corresponding to the frequency domain response of the variable node x_k^e, whereas it has been used as a separate variable node above, hence $\underline{x}_k^e \triangleq \mathbf{e}_k^H \mathcal{F}_K \mathbf{x}^{\mathbf{e}}$. Indeed when inserting Equations (4.17)-(4.18) into Equations (4.30)-(4.31), we get a LE-IC structure given by

$$\underline{x}_k^e = \underline{x}_k^d + \underline{f}_k^{\text{ep}*}(\underline{y}_k - \underline{h}_k \underline{x}_k^d), \tag{4.35}$$

$$v_x^e = \frac{1}{\xi_x^{\text{ep}}} - v_x^d, \tag{4.36}$$

where the FD filter and normalization parameters are given by

$$\underline{f}_k^{\text{ep}} = \frac{1}{\xi_x^{\text{ep}}} \frac{\underline{h}_k}{\bar{\sigma}_w^2 + v_x^d |\underline{h}_k|^2}, \tag{4.37}$$

$$\xi_x^{\text{ep}} = \frac{1}{K} \sum_{k=1}^{K} \frac{|\underline{h}_k|^2}{\bar{\sigma}_w^2 + v_x^d |\underline{h}_k|^2}. \tag{4.38}$$

These equations bear great similarity to the conventional MMSE FD LE-IC structure, with interference cancellation being carried out using extrinsic EP feedback x_k^d.

To fully exploit the benefits of the feedback computed by the demapper, as for the self-iterated DFE in Section 3.5, a flexible double-loop FDE structure is proposed. The first loop refers to the exchange of extrinsic information between the decoder and the demapper in a turbo-iteration, while the second loop refers to the message exchange in a self-iteration between the demapper and the equalizer.

As for FIR DFE with SIs in Section 3.5, damped message passing recommendations from the Section 2.4.1 are used. In this case, along with the feature-based damping in [Min+05; C+14], the EP-based feedback in Equation (3.62) is replaced with

$$\begin{aligned} v_x^{d(\text{next})} &= \left[(1-\beta) \frac{1}{v_x^{\text{curr}}} + \beta \frac{1}{v_x^{d(\text{prev})}} \right]^{-1}, \\ x_k^{d(\text{next})} &= v_x^{d(\text{next})} \left[(1-\beta) \frac{x_x^{d(\text{curr})}}{v^{d(\text{curr})}} + \beta \frac{x_k^{d(\text{prev})}}{v^{d(\text{prev})}} \right], \end{aligned} \tag{4.39}$$

where $0 \leqslant \beta \leqslant 1$ configures the damping and $x_{x,k}^{d(\text{curr})}$ and $v_x^{d(\text{curr})}$ are given by Equation (4.33). Alternatively, as evoked in Section 2.4.1, exponential smoothing of message parameters (linear

[1] In this previously published journal and conference article versions of this contribution [Şa+18a; Şa+18b; Şa+18e], this receiver is exposed more succinctly, due to lack of space, by having a EQU factor node between FD observations and TD symbols. The interest in having $\{\text{OBS}_k\}_{k=1}^K$ and DFTx nodes separate is to enable an easier and clearer generalization of this framework to more complex systems.

Algorithm 13 Proposed *Self-Iterated FD LE-EPIC* receiver.

Input $\mathbf{y}, \mathbf{H}$
Output $\hat{\mathbf{b}}$

$L_a^{(0)}(\mathbf{d}) = \mathbf{0}_{K_c}$
for $\tau = 0 \dots \mathcal{T}$ **do**
  Compute prior PMFs $\mathcal{P}_k^{(\tau)}(\alpha)$ with Eq. (2.39) and $L_a^{(\tau)}(\mathbf{d})$, for $\alpha \in \mathcal{X}, k = 1, \dots, K$.
  Compute $\mathbf{x}^{\mathbf{p}(\tau)}$ and $\mathbf{v}_{\mathbf{x}}^{\mathbf{p}(\tau)}$ with Eq. (2.40), and set $(\mathbf{x}^{\mathbf{d}(\tau,0)}, v_x^{d(\tau,0)}) \leftarrow (\mathbf{x}^{\mathbf{p}(\tau)}, v_x^{p(\tau)})$.
  for $s = 0$ to $\mathcal{S}_\tau$ **do**
    Apply FFT on $\mathbf{x}^{\mathbf{d}(\tau,s)}$ to get $\underline{\mathbf{x}}^{\mathbf{d}(\tau,s)}$.
    $\xi_x^{\mathrm{ep}(\tau,s)} \leftarrow K^{-1} \sum_{k=1}^{K} |\underline{h}_k|^2 / (\bar{\sigma}_w^2 + v_x^{d(\tau,s)} |\underline{h}_k|^2)$
    $v_x^{e(\tau,s)} \leftarrow \xi_x^{\mathrm{ep}(\tau,s)-1} - v_x^{d(\tau,s)}$
    for $k = 1 \dots K$ **do**
      $\underline{f}_k^{\mathrm{ep}(\tau,s)} \leftarrow \xi_x^{\mathrm{ep}}(\tau,s)^{-1} \underline{h}_k / (\bar{\sigma}_w^2 + v_x^{d(\tau,s)} |\underline{h}_k|^2)$
      $\underline{x}_k^{e(\tau,s)} \leftarrow \underline{x}_k^{d(\tau,s)} + \underline{f}_k^{\mathrm{ep}(\tau,s)*} (\underline{y}_k - \underline{h}_k \underline{x}_k^{d(\tau,s)})$
    end for
    Apply IFFT on $\underline{\mathbf{x}}^{\mathbf{e}(\tau,s)}$ to get $\mathbf{x}^{\mathbf{e}(\tau,s)}$.
    for $k = 1 \dots K$ **do**
      Compute the posterior PMF $\mathcal{D}_k^{(\tau)}(\alpha)$ with Eq. (2.53), for $\alpha \in \mathcal{X}$, and APP estimates $(\mu_{x,k}^{d(\tau,s+1)}, \gamma_{x,k}^{d(\tau,s+1)})$ with Eq. (2.98).
    end for
    Generate EP feedback $(\mathbf{x}^{\mathbf{d}(\tau,s+1)}, v_x^{d(\tau,s+1)})$ with Eqs. (4.32)-(4.34).
    Apply damping with either Eq. (4.39) or (4.40), for $s > 0$.
  end for
  Compute extrinsic LLRs $L_e^{(\tau)}(\mathbf{d})$ with Eq. (2.54)
  Execute de-interleaving, SISO decoding and interleaving to get $L_a^{(\tau+1)}(\mathbf{d})$
end for

combination with memory) is considered with

$$\begin{aligned} v_x^{d(\text{next})} &= (1-\beta) v_x^{\text{curr}} + \beta v_x^{d(\text{prev})}, \\ x_k^{d(\text{next})} &= (1-\beta) x_x^{d(\text{curr})} + \beta x_k^{d(\text{prev})}. \end{aligned} \tag{4.40}$$

Similarly to the Section 3.5, each turbo-iteration $\tau = 0, \dots, \mathcal{T}$ consists of $\mathcal{S}_\tau$ self-iterations (may depend on τ), where DFTx and DEM factor nodes are updated in parallel schedule, for $s = 0, \dots, \mathcal{S}_\tau$, and then the DEC factor nodes are updated with a selected SISO decoder. To clarify this, Algorithm 13 above explicitly describes the proposed scheduling, where involved quantities are indexed by (τ, s) in the superscript.

The iterative FDE derived in this section, by applying the EP framework in the FD, with the family of white Gaussian distributions, yields the low-complexity single-tap receiver structure shown in Figure 4.3. In the next section, the behaviour of this receiver will be assessed with achievable rate analysis and comparisons with structures from the prior work. A the scalar EP modelling of extrinsic messages over data symbols, and the considered double loop

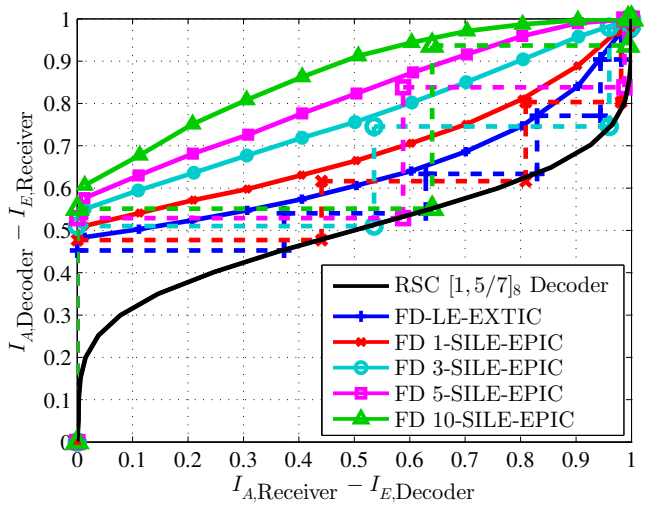


Figure 4.4: EXIT curves and finite-length average MI trajectories of the proposed equalizer with 8-PSK in Proakis C channel at $E_b/N_0 = 15$dB.

schedule are at the core of this FD SILE-EPIC detector, and its extensions, this framework is also denoted as Double Loop - Scalar EP (DL-SEP).

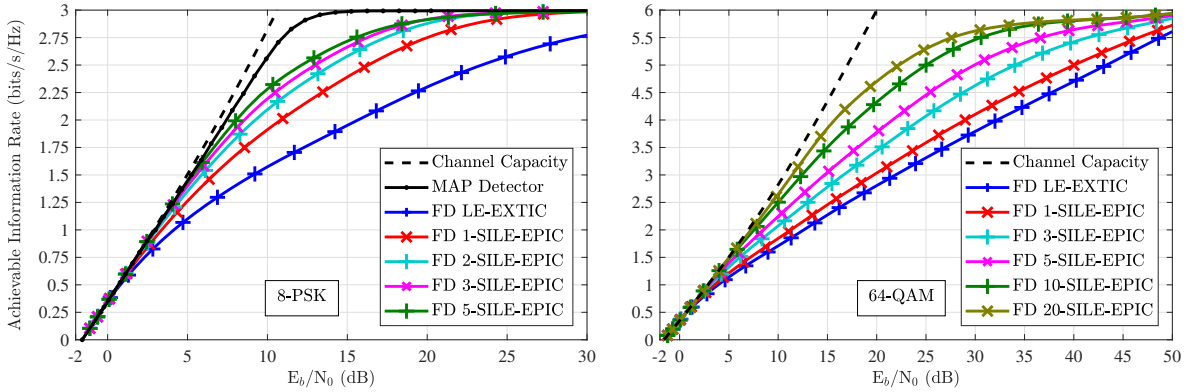
4.3.3 Asymptotic and Finite-Length Analysis

Here, the receiver derived in the previous section is used with a fixed number of SIs $\mathcal{S}_\tau = \mathcal{S}$ per turbo-iteration, and this structure is referred as the $\mathcal{S}$-self-iterated FD LE-IC with EP (*FD $\mathcal{S}$-SILE-EPIC*).

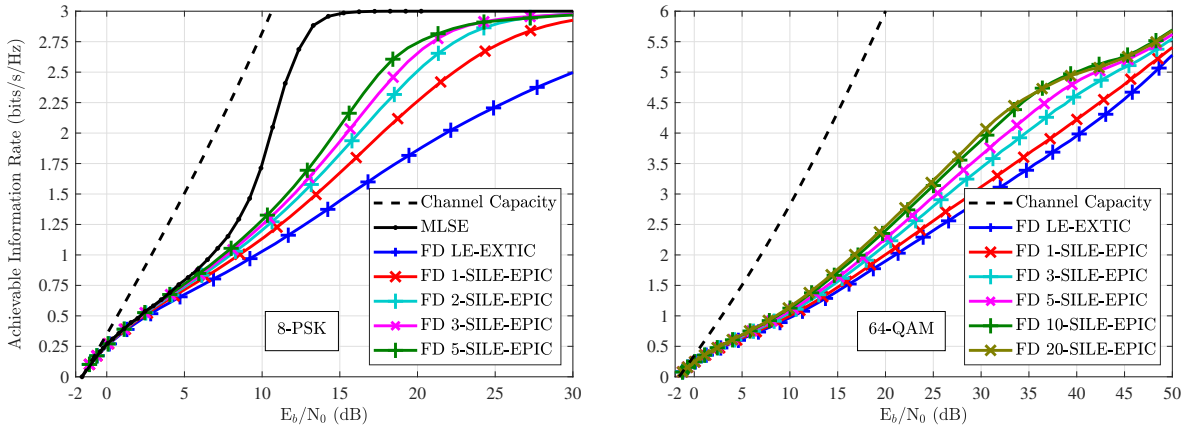
Asymptotic Analysis In order to evaluate asymptotic behaviour ($\tau \to \infty$) of the proposed receiver, Extrinsic Information Transfer (EXIT) analysis [TB00], summarized in Section 2.3.6, is used to characterize the asymptotic limits of this receiver.

In Figure 4.4, EXIT charts of the proposed receiver, for $\mathcal{S} = 0, 1, 3, 5, 10$, using a fixed damping with exponential smoothing (see eq. (4.40)), with $\beta = 0.75$, is provided in solid curves, within the highly selective Proakis C channel, $\mathbf{h} = [1, 2, 3, 2, 1]/\sqrt{19}$, using the Gray-mapped 8-PSK constellation. Self-iterations are seen to significantly improve the extrinsic mutual information for high I_A, which indicates a boosted convergence speed and an improved achievable rate. However, improvements for $I_A = 0$ is relatively small, thus, the finite-length performance improvement will strongly depend on the EXIT chart of the decoder.

This figure also shows the reverse transfer curve of the RSC code $[1, 5/7]_8$. Moreover, in dashed curves, the finite-length MI trajectories of this receiver with data blocks of length $K_b = 768$ bits, using this channel decoder is plotted. The trajectories of the proposed EP-based receiver appears to follow the predicted transfer function fairly well, despite the short packet length, unlike the APP-based receivers as observed in Section 3.4. This suggests that



(a) Asymptotic achievable information rates ($\mathcal{T} \to +\infty$).



(b) Non-turbo iterative achievable rates ($\mathcal{T} = 0$).

Figure 4.5: Achievable rates of the proposed receiver in Proakis C with 8-PSK and 64-QAM.

this receiver's EXIT analysis reflects its practical behaviour.

Achievable rates of the FD LE-EXTIC and the proposed receiver are given in Figure 4.5, for the Proakis C channel with 8-PSK and 64-QAM constellations. The Gaussian capacity of this channel, without transmit power optimization, is also plotted in dashed lines, it is computed using Equation (1.36) with the vector-input AWGN channel capacity. Channel SIR with 8-PSK is given by the MAP detector curve in 8-PSK, but it is not plotted with 64-QAM due to the excessive computational resources it requires [ALV+06]. A feature-based damping (see eq. (4.39)) with $\beta = 0.7 \times 0.9^s$ is used for 8-PSK, whereas a linear damping (see eq. (4.39)) with $\beta = 0.8$ is used for 64-QAM.

For 8-PSK, while the conventional FD LE-EXTIC [TH01] follows the SIR limit within 0.5 dB up to 0.75 bits/s/Hz, proposed EP-based self-iterations increase this range up to 2 bits/s/Hz. In the 64-QAM case, FD LE-EXTIC follows the channel capacity within 1 dB up to 1 bit/s/Hz and 3.33 bits/s/Hz becomes achievable with 20 self-iterations. For a rate-1/2 coded usage, the proposed receiver with $s \to +\infty$ brings over 3.9 dB and 10.7 dB improvement, over the conventional turbo FD LE in this channel, for respectively 8-PSK and 64-QAM

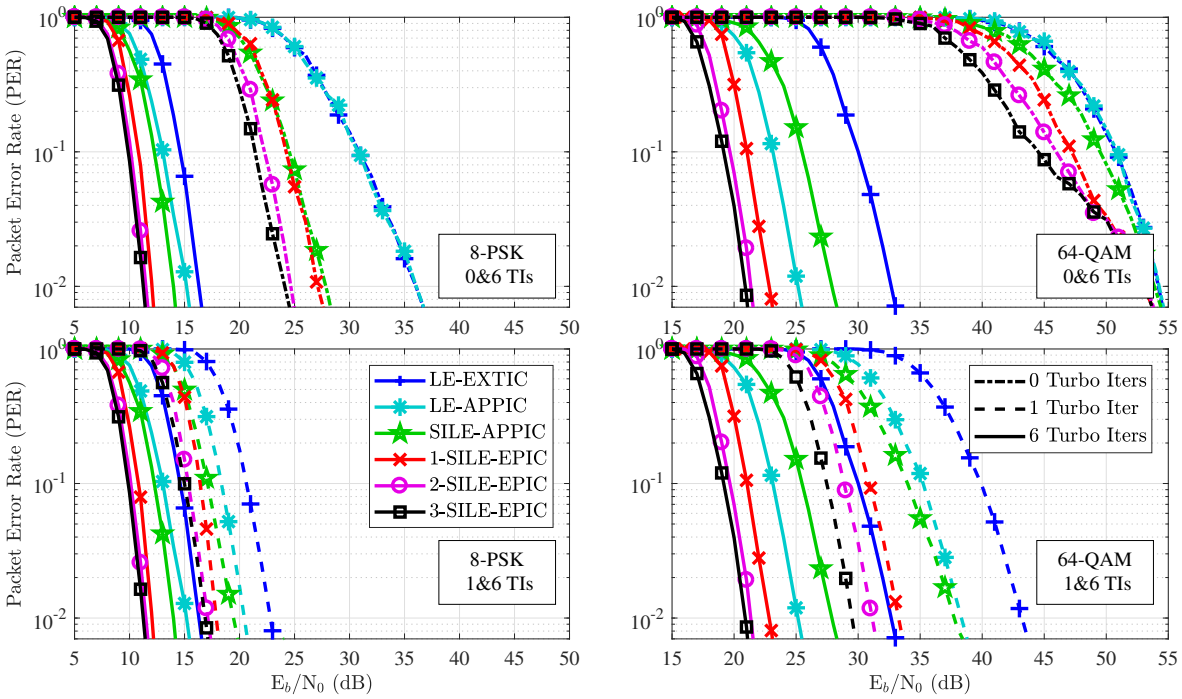


Figure 4.6: PER comparison of single-tap FD equalizers in Proakis C channel, with $K = 256$ coded with rate-1/2 RSC $[1, 5/7]_8$.

constellations. In the non-turbo iterative case, the proposed receiver does not reach the MLSE performance, but there are still non-negligible performance improvements. These rates are achievable with properly designed coding schemes.

Comparison with Single Tap FDE in Prior Work In this paragraph, observations in the previous section are completed with finite-length results within the same channel with a RSC code with soft MAP decoder. In particular, as APP feedback based receiver cannot be reliably analysed with EXIT analysis, this is the main approach for assessing their performance.

PER is obtained by Monte-Carlo simulations, with 30000 sent packets per point. Unlike in asymptotic analysis, here we use dynamic damping that also depends on turbo-iterations, τ, and accelerates convergence. A feature-based damping with $\beta_{\tau,s} = 0.7 \times 0.9^{s+\tau}$ is used for 8-PSK, and a hybrid damping, consisting of a linear smoothing in the first turbo-iteration, and feature-based damping afterwards, is applied with $\beta_{\tau,s} = 0.85^{1+s+\tau}$, for 64-QAM. Several single-tap FD equalizers are compared to our proposal in Figure 4.6: the conventional linear equalizer [TH01] (*LE-EXTIC*), the LE-IC with APP feedback [VBC06; Wit+02] (*LE-APPIC*), and the self-iterated LE-IC of [GM08; Tao15] (*SILE-APPIC*), based on PDA. The equalization complexity of these receivers has the same order of computational complexity scaling of $K \log_2 K$, at a given SI and turbo-iteration, with slight differences underlying in the feedback computation.

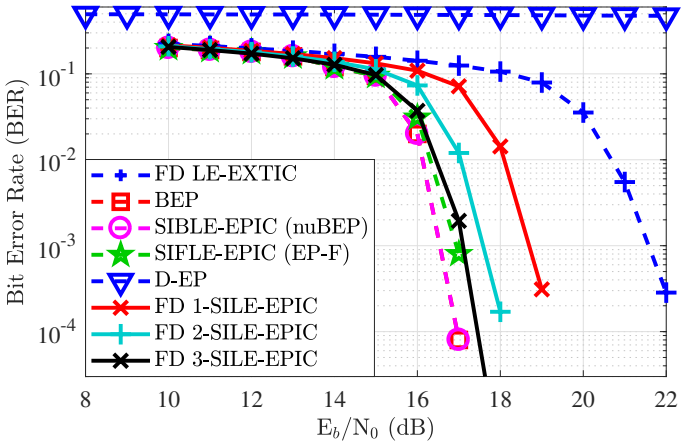
Results in Figure 4.6 show that our proposal brings significant improvement on the decoding threshold, that grows with the number of SIs, at all turbo-iterations. On the contrary, multiple SIs with APP feedback degrades this threshold (not shown here due to lack of space). Without turbo-iterations, 3 SIs bring respectively 9 dB and 6 dB gains for 8-PSK and 64-QAM, compared to LE-EXTIC, at PER = 10^{-1}. Performance in 64-QAM is limited at low PER without turbo-iterations, but our proposal with a single turbo-iteration and 3 SIs reaches PER the prior work reach with 6 turbo-iterations, e.g. with six times lower decoding complexity. Besides, "asymptotically" (6 turbo-iterations), SIs with EP bring over 8 dB gain with respect to SILE-APPIC, and about 5 dB gain over LE-APPIC, for 64-QAM, at BLER = 10^{-2}. Compared to FD LE-EXTIC, 3 self-iterations bring around 4 dB and 11.5 dB gain, respectively for 8-PSK and 64-QAM, which are close to the 1/2-rate gains observed in the asymptotic analysis above.

These results encourage replacing turbo-iterations with self-iterations as demapping complexity is often insignificant relative to decoding.

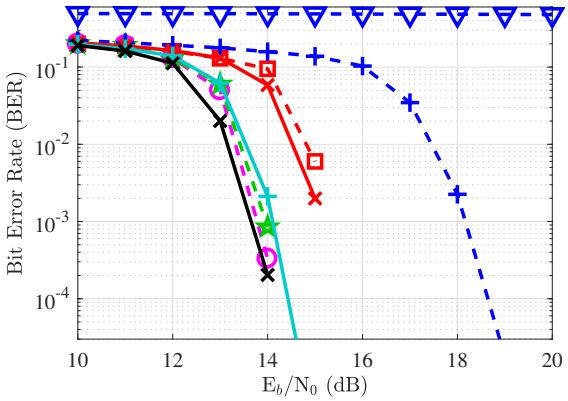
Comparison with EP-based Receivers in Prior Work There are numerous emerging EP-based receivers in the literature, as stressed in the Chapter 2, and in previous sections. In this paragraph the proposed FDE is compared with self-iterated time-domain linear block (SIBLE-EPIC, denoted nuBEP in [SMF+18]) and filter (SIFLE-EPIC, denoted EP-F in [SMF+18]) receivers and to the single-tap FD receiver D-EP in [WRM+17]. The proposed receiver is not compared to the exact FD receiver, J-EP in [WRM+17; Zha+16], as that receiver is equivalent to the SIBLE-EPIC with a single SI, without damping, making it sub-optimal compared to the SIBLE-EPIC. The block receiver in [SMF+17a], denoted BEP, is a sub-optimal receiver, where the demapper ignores prior information from the decoder, when computing EP-based feedback at each SI.

In Figure 4.7, the BER of the proposed receiver is compared with alternatives listed above. We consider 8-PSK constellation, and the LDPC coded Proakis C scenario from [SMF+18]. The regular $(3, 6)$ LDPC code is obtained by Progressive-Edge Growth (PEG) algorithm, and the decoder uses BP algorithm up to 100 iterations. The FD receiver, D-EP, cannot decode in Proakis-C channel, up to very high signal to noise ratios due to its sensitivity to channel nulls [WRM+17, eq. (48)]. Our FD proposal performs nearly as good as the TD EP-based block filterbank receivers, with an order of computational complexity of $(\mathcal{S}+1)K\log_2 K$ instead of $3LK^2$ (SIBLE-EPIC, 2 SIs) and or $27KL^2$ (SIFLE-EPIC, 2 SIs). For $\tau = 5$, block and filter TD receivers have around 0.2 dB gain over FD 3-SILE EPIC, but they are respectively around 500 and 16 times more complex.

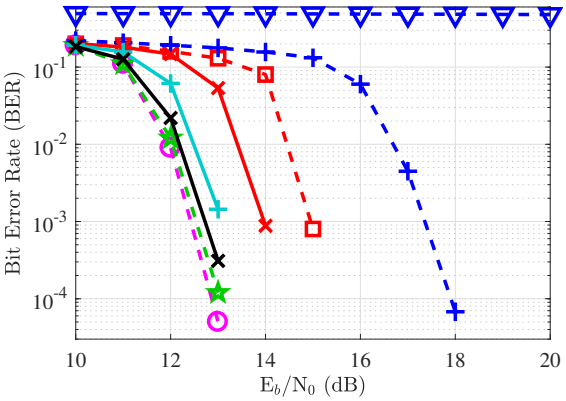
Another LDPC-coded scenario in the Proakis C, with 16-QAM and with rate 1/2 and 3/4 encoding over $K_b = 2048$ bits is reported in Figure 4.8. All receivers use feature-based damping with the optimized parameters in [SMF+18], i.e. $\beta = \min(0.3, 1 - e^{\tau/1.5}/10)$. The regular $(3, 12)$ LDPC code is also obtained by the PEG algorithm. In the rate-1/2 case, the proposed FDE is lower-bounded in BER by the block receiver, and following one SI, the difference between FD SILE-EPIC and SIFLE-EPIC is negligible. For the high rate case, at the right side of the figure, filter receiver's performance is over 1 dB worse for BER $< 10^{-3}$,



(a) No Turbo Iterations



(b) 2 Turbo Iterations



(c) 5 Turbo Iterations

Figure 4.7: BER comparison in Proakis C with 8-PSK, $K_d = 4096$ and $(3, 6)$ LDPC code.

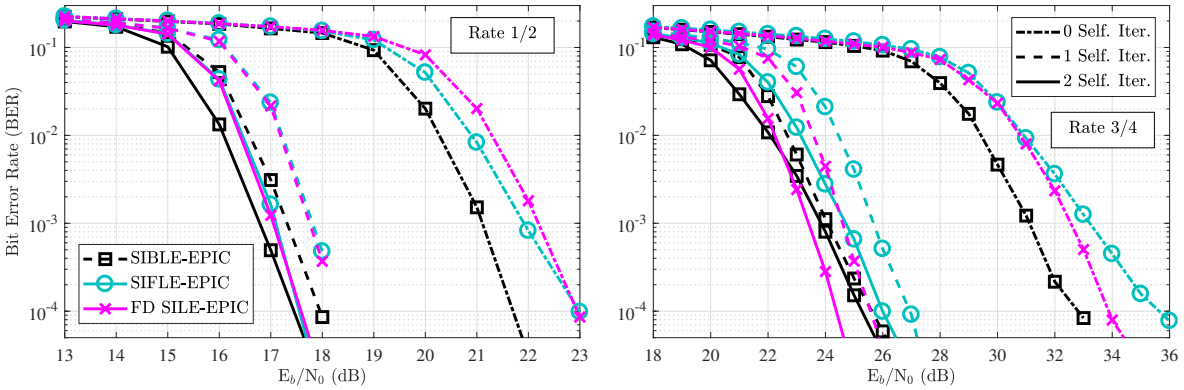


Figure 4.8: Comparison of block, FIR and single-tap FD equalization structures in Proakis C with LDPC coded 16-QAM, with 5 turbo iterations.

and although SIBLE-EPIC still has a better decoding threshold, it recovers less diversity than the proposed FD SILE-EPIC. This phenomenon should not be surprising, as exact receivers

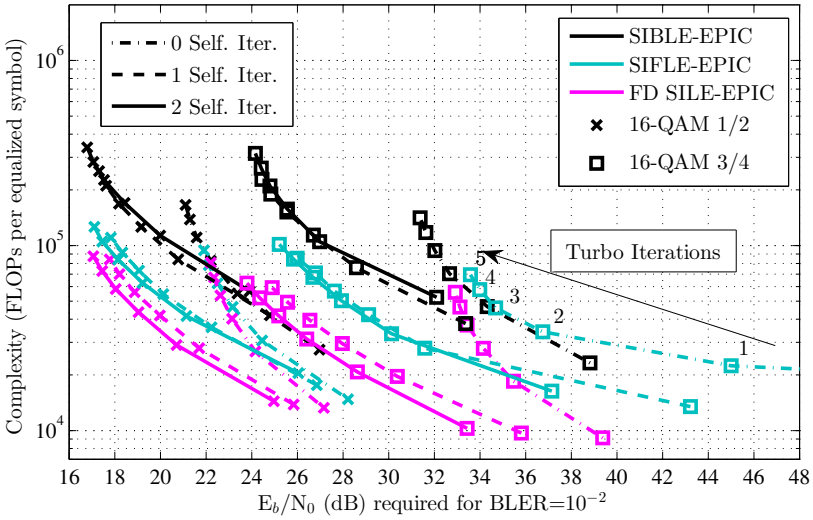


Figure 4.9: Performance complexity trade-off for self-iterations in LDPC coded Proakis C.

can be more prone to error propagation when decoder provides erroneous feedback, as also observed in filter receivers [JM13].

These error rate results are completed with detailed computational complexity estimations in Figure 4.9. This is evaluated with the number of multiply and accumulate units required to implement the receiver, estimated by the number of real additions and multiplications, amounting to half a floating point operation (0.5 FLOPs) each. Complexity is plotted versus the required $\mathcal{E}_b/N_0$ to decode transmitted packets with PER = 10^{-2}, for $\tau = 0, \ldots, 5$. These FLOP-counts also include the decoder complexity, which is considerably higher than equalizer complexity. The proposed receiver performs overall efficiently, both complexity and energy-wise, compared to the SIBLE-EPIC, with respectively 2.5, 4 and 5.4 times lower complexity for $\mathcal{S} = 0, 1$ and 2 in the rate 3/4 case, and with respectively 2, 3.1 and 4.1 times lower complexity for the rate 1/2 case. This ratio is around ten times bigger, if the decoding complexity is not accounted for.

4.3.4 Comparison with work on Approximate Message Passing

There is a great number of contributions on iterative message-passing algorithms for low complexity parsimonious detection. In this section the extension of these algorithms for SC-FDE BICM detection is discussed, in order to establish their theoretical links to the proposed EP-based self-iterated inference methods. These algorithms have all roughly the same computational cost order of $\mathcal{O}(\mathcal{S}K \log_2 K)$.

AMP based algorithms originally do not address (de)mapping aspects of turbo detection, and they output APP symbol estimates, as they are intended for providing estimates for other signal processing fields such as compressed sensing or data classification. These estimates can be directly fed to a soft-demapper, but the decoding performance is significantly degraded

Algorithm 14 Turbo FDE with *Iterative Thresholding*.

Input $\mathbf{y}$, $\mathbf{H}$
Output $\hat{\mathbf{b}}$

$L_a^{(0)}(\mathbf{d}) = \mathbf{0}_{K_c}$.
for $\tau = 0 \dots \mathcal{T}$ **do**
  Compute prior PMFs $\mathcal{P}_k^{(\tau)}(\alpha)$ with Eq. (2.39) and $L_a^{(\tau)}(\mathbf{d})$, for $\alpha \in \mathcal{X}, k = 1, \dots, K$.
  Compute $\mathbf{x}^{\mathbf{p}(\tau)}$ and $\mathbf{v}_{\mathbf{x}}^{\mathbf{p}(\tau)}$ with Eq. (2.40), and set $\boldsymbol{\mu}_{\mathbf{x}}^{\mathbf{d}(\tau,0)} \leftarrow \mathbf{x}^{\mathbf{p}(\tau)}$.
  for $s = 0$ to $\mathcal{S}_\tau$ **do**
    Apply FFT on $\boldsymbol{\mu}_{\mathbf{x}}^{\mathbf{d}(\tau,s)}$ to get $\underline{\boldsymbol{\mu}}_{\mathbf{x}}^{\mathbf{d}(\tau,s)}$.
    for $k = 1 \dots K$ **do**
      $\underline{f}_k^{\mathrm{it}(\tau,s)} \leftarrow \left(\sum_{k=1}^{K} |\underline{h}_k|^2\right)^{-1} \underline{h}_k$
      $\underline{x}_k^{e(\tau,s)} \leftarrow \underline{\mu}_{x,k}^{d(\tau,s)} + \underline{f}_k^{\mathrm{it}(\tau,s)*}(\underline{y}_k - \underline{h}_k \underline{\mu}_{x,k}^{d(\tau,s)})$
    end for
    Apply IFFT on $\underline{\mathbf{x}}^{\mathbf{e}(\tau,s)}$ to get $\mathbf{x}^{\mathbf{e}(\tau,s)}$.
    $v_x^{e(\tau,s)} \leftarrow \max\left([\sum_k |\underline{y}_k - \underline{h}_k \underline{x}_k^{e(\tau,s)}|^2 - K\bar{\sigma}_w^2]/\sum_k |\underline{h}_k|^2, K\bar{\sigma}_w^2/\sum_k |\underline{h}_k|^2\right)$.
    for $k = 1 \dots K$ **do**
      Compute the posterior PMF $\mathcal{D}_k^{(\tau)}(\alpha)$ with Eq. (2.53), for $\alpha \in \mathcal{X}$, and APP estimates $\mu_{x,k}^{d(\tau,s+1)}$ with Eq. (2.98).
    end for
  end for
  Compute extrinsic LLRs $L_e^{(\tau)}(\mathbf{d})$ with Eq. (2.54).
  Execute de-interleaving, SISO decoding and interleaving to get $L_a^{(\tau+1)}(\mathbf{d})$.
end for

due to the processing of biased APP estimates[2]. As an alternative to the direct application of AMP algorithms for turbo FDE, a better thought approach is to replace the final APP estimation step with extrinsic bit LLR demapping as done in [Guo+13] for GAMP. Here we discuss the application of the same method to iterative thresholding, AMP, and the recently proposed EP-based VAMP and OAMP algorithms.

Iterative Soft Thresholding (IST) When IST is applied to turbo FDE, with single-tap, normalized matched-filtering and it generates APP-based feedback for IC. As mentioned in Section 2.4.1 and in [DHD12], instead of following this algorithm's original unknown noise variance assumption, for fair comparison with other FDE, a maximum-likelihood sample variance estimator is used for improving the convergence

However, as previously mentioned, this technique is a suboptimal, but low-complexity variant of PDA, hence FDE based on iterative thresholding should not be expected to outperform FD SILE-APPIC [GM08; Tao15], and for the FD implementation, the gain in computational complexity is small. As shown in Algorithm 14, there is no need to compute the covariance associated to the soft feedback for this receiver, and lines corresponding to the adaptation of AMP algorithms to a BICM system are faded out with a gray color.

[2]This fact illustrated below in Figure 4.10.

Algorithm 15 Turbo FDE with *Approximate Message Passing (AMP)*

Input $\mathbf{y}$, $\mathbf{H}$
Output $\hat{\mathbf{b}}$

$L_a^{(0)}(\mathbf{d}) = \mathbf{0}_{Kc}, \underline{\boldsymbol{\epsilon}}^{\mathbf{d}(\tau,0)} = \mathbf{0}_K, \forall \tau.$
for $\tau = 0 \dots \mathcal{T}$ **do**
  Compute prior PMFs $\mathcal{P}_k^{(\tau)}(\alpha)$ with Eq. (2.39) and $L_a^{(\tau)}(\mathbf{d})$, for $\alpha \in \mathcal{X}, k = 1, \dots, K$.
  Compute $\mathbf{x}^{\mathbf{p}(\tau)}$ and $\mathbf{v}_{\mathbf{x}}^{\mathbf{p}(\tau)}$ with Eq. (2.40), and set $(\boldsymbol{\mu}_{\mathbf{x}}^{\mathbf{d}(\tau,0)}, \gamma_x^{d(\tau,0)}) \leftarrow (\mathbf{x}^{\mathbf{p}(\tau)}, v_x^{p(\tau)})$.
  for $s = 0$ to $\mathcal{S}_\tau$ **do**
    Apply FFT on $\boldsymbol{\mu}_{\mathbf{x}}^{\mathbf{d}(\tau,s)}$ to get $\underline{\boldsymbol{\mu}}_{\mathbf{x}}^{\mathbf{d}(\tau,s)}$.
    for $k = 1 \dots K$ **do**
      $\underline{f}_k^{\mathrm{amp}(\tau,s)} \leftarrow \left(\sum_{k=1}^K |\underline{h}_k|^2\right)^{-1} \underline{h}_k$
      $\underline{x}_k^{e(\tau,s)} \leftarrow \underline{\mu}_{x,k}^{d(\tau,s)} + \underline{f}_k^{\mathrm{amp}(\tau,s)*}(\underline{y}_k - \underline{h}_k \underline{\mu}_{x,k}^{d(\tau,s)}) + \underline{\epsilon}_k^{d(\tau,s)}$
    end for
    Apply IFFT on $\underline{\mathbf{x}}^{\mathbf{e}(\tau,s)}$ to get $\mathbf{x}^{\mathbf{e}(\tau,s)}$.
    $v_x^{e(\tau,s)} \leftarrow \max\left([\sum_k |\underline{y}_k - \underline{h}_k \underline{x}_k^{e(\tau,s)}|^2 - K\bar{\sigma}_w^2]/\sum_k |\underline{h}_k|^2, K\bar{\sigma}_w^2/\sum_k |\underline{h}_k|^2\right)$.
    for $k = 1 \dots K$ **do**
      Compute the posterior PMF $\mathcal{D}_k^{(\tau)}(\alpha)$ with Eq. (2.53), for $\alpha \in \mathcal{X}$, and APP estimates $(\mu_{x,k}^{d(\tau,s+1)}, \gamma_{x,k}^{d(\tau,s+1)})$ with Eq. (2.98).
      $\underline{\epsilon}_k^{d(\tau,s+1)} \leftarrow [\gamma^{d(\tau,s+1)}/v^{e(\tau,s)}]\left(\underline{x}_k^{e(\tau,s)} - \underline{\mu}_{x,k}^{d(\tau,s+1)}\right)$.
    end for
  end for
  Compute extrinsic LLRs $L_e^{(\tau)}(\mathbf{d})$ with Eq. (2.54).
  Execute de-interleaving, SISO decoding and interleaving to get $L_a^{(\tau+1)}(\mathbf{d})$.
end for

Approximate Message Passing (AMP) AMP improves iterative thresholding by compensating the impact of neglected correlations through the use of an Onsager reaction term (see Section 2.4.1). When applied in the context of a turbo FDE, this results in the computation of the estimation error between the equalizer's output and the demodulator's soft APP estimate, which is scaled by the ratio of associated estimates' variances.

Algorithm 15 gives the details of this technique for turbo FDE, and as the computation of the Onsager term involves the use of the covariance of soft APP estimates, the computational complexity of this receiver is higher than iterative thresholding, and it is close to that of FD SILE-APPIC or FD SILE-EPIC.

The major drawbacks of this technique lie in the use of a sub-optimal matched-filter as an equalizer, and the use of a heuristic estimator for computing the covariance of equalized estimates v_x^e. Although among APP-based receivers, FD SILE-APPIC, based on PDA, seems to be a more robust solution, with filters designed with MMSE criterion, the decorrelated iterations of AMP could improve the performance of self-iterations.

Algorithm 16 Turbo FDE with *Generalized Approximate Message Passing (GAMP)*

Input $\mathbf{y}$, $\mathbf{H}$
Output $\hat{\mathbf{b}}$

$L_a^{(0)}(\mathbf{d}) = \mathbf{0}_{K_c}, \underline{\boldsymbol{\epsilon}}^{\mathbf{d}(\tau,0)} = \mathbf{0}_K, \forall \tau.$
for $\tau = 0 \dots \mathcal{T}$ **do**
  Compute prior PMFs $\mathcal{P}_k^{(\tau)}(\alpha)$ with Eq. (2.39) and $L_a^{(\tau)}(\mathbf{d})$, for $\alpha \in \mathcal{X}, k = 1, \dots, K$.
  Compute $\mathbf{x}^{\mathbf{p}(\tau)}$ and $\mathbf{v}_{\mathbf{x}}^{\mathbf{p}(\tau)}$ with Eq. (2.40), and set $(\boldsymbol{\mu}_{\mathbf{x}}^{\mathbf{d}(\tau,0)}, \gamma_x^{d(\tau,0)}) \leftarrow (\mathbf{x}^{\mathbf{p}(\tau)}, v_x^{p(\tau)})$.
  for $s = 0$ to $\mathcal{S}_\tau$ **do**
    Apply FFT on $\boldsymbol{\mu}_{\mathbf{x}}^{\mathbf{d}(\tau,s)}$ to get $\underline{\boldsymbol{\mu}}_{\mathbf{x}}^{\mathbf{d}(\tau,s)}$.
    $\xi_x^{\text{gamp}(\tau,s)} \leftarrow K^{-1} \sum_{k=1}^{K} |\underline{h}_k|^2 / (\bar{\sigma}_w^2 + \gamma_{x,k}^{d(\tau,s)} |\underline{h}_k|^2)$
    $v^{e(\tau,s)} \leftarrow \xi_x^{\text{gamp}(\tau,s)-1}$
    for $k = 1 \dots K$ **do**
      $\underline{f}_k^{\text{gamp}(\tau,s)} \leftarrow \xi_x^{\text{gamp}(\tau,s)-1} \underline{h}_k / (\bar{\sigma}_w^2 + \gamma_{x,k}^{d(\tau,s)} |\underline{h}_k|^2)$
      $\underline{x}_k^{e(\tau,s)} \leftarrow \underline{\mu}_{x,k}^{d(\tau,s)} + \underline{f}_k^{\text{gamp}(\tau,s)*} (\underline{y}_k - \underline{h}_k (\underline{\mu}_{x,k}^{d(\tau,s)} - \underline{\epsilon}_k^{d(\tau,s)}))$
    end for
    Apply IFFT on $\underline{\mathbf{x}}^{\mathbf{e}}$ to get $\mathbf{x}^{\mathbf{e}}$.
    for $k = 1 \dots K$ **do**
      Compute the posterior PMF $\mathcal{D}_k^{(\tau)}(\alpha)$ with Eq. (2.53), for $\alpha \in \mathcal{X}$, and APP estimates $(\mu_{x,k}^{d(\tau,s+1)}, \gamma_{x,k}^{d(\tau,s+1)})$ with Eq. (2.98).
      $\underline{\epsilon}_k^{d(\tau,s+1)} \leftarrow [\gamma^{d(\tau,s+1)} / v^{e(\tau,s)}] \left(\underline{x}_k^{e(\tau,s)} - \underline{\mu}_{x,k}^{d(\tau,s+1)} \right)$
    end for
  end for
  Compute extrinsic LLRs $L_e^{(\tau)}(\mathbf{d})$ with Eq. (2.54).
  Execute de-interleaving, SISO decoding and interleaving to get $L_a^{(\tau+1)}(\mathbf{d})$.
end for

Generalized Approximate Message Passing (GAMP) A GAMP based turbo FDE has been formulated in [Guo+13], which combines the advantages of AMP and PDA, by yielding a structure with both MMSE filters and decorrelating Onsager reaction term. The details of this technique is provided in Algorithm 16, and it will be shown to yield the most robust equalization performance among any receivers that use APP-based estimates.

Orthogonal Approximate Message Passing (OAMP) MMSE-optimal solution of OAMP yields an iterative algorithm that uses EP-like soft feedback for Interference Cancellation (IC). Indeed such a feedback is found to be divergence-free [MP17] and by coupling it with a heuristic for covariance estimation, an alternative self-iterated algorithm is obtained.

Algorithm 17 shows this technique's instantiation for turbo FDE has multiple similarities to the proposed FD SILE-EPIC, and the main differences are the use of an exact computation of covariance and damping in our method. Lack of a damping parameter can be seen as a simplifying advantage, and the heuristic ML estimation of the covariance is more robust, after a certain number of iterations, to poor noise realizations on short or medium length packets.

Algorithm 17 Turbo FDE with *Orthogonal Approximate Message Passing (OAMP)*

Input $\mathbf{y}$, $\mathbf{H}$
Output $\hat{\mathbf{b}}$

1: $L_a^{(0)}(\mathbf{d}) = \mathbf{0}_{K_c}$.
2: **for** $\tau = 0 \dots \mathcal{T}$ **do**
3: Compute prior PMFs $\mathcal{P}_k^{(\tau)}(\alpha)$ with Eq. (2.39) and $L_a^{(\tau)}(\mathbf{d})$, for $\alpha \in \mathcal{X}, k = 1, \dots, K$.
4: Compute $\mathbf{x}^{\mathbf{p}(\tau)}$ and $\mathbf{v}_{\mathbf{x}}^{\mathbf{p}(\tau)}$ with Eq. (2.40), and set $(\mathbf{x}^{*(\tau,0)}, v_x^{*(\tau,0)}) \leftarrow (\mathbf{x}^{\mathbf{p}(\tau)}, v_x^{p(\tau)})$.
5: **for** $s = 0$ to $\mathcal{S}_\tau$ **do**
6: Apply FFT on $\mathbf{x}^{*(\tau,s)}$ to get $\underline{\mathbf{x}}^{*(\tau,s)}$.
7: $\xi_x^{\text{oamp}(\tau,s)} \leftarrow K^{-1} \sum_{k=1}^{K} |\underline{h}_k|^2 / (\bar{\sigma}_w^2 + v_x^{*(\tau,s)} |\underline{h}_k|^2)$
8: $v_x^{e(\tau,s)} \leftarrow \xi_x^{\text{oamp}(\tau,s)-1} - v_x^{*(\tau,s)}$
9: **for** $k = 1 \dots K$ **do**
10: $\underline{f}_k^{\text{oamp}(\tau,s)} \leftarrow \xi_x^{\text{oamp}(\tau,s)-1} \underline{h}_k / (\bar{\sigma}_w^2 + v_x^{*(\tau,s)} |\underline{h}_k|^2)$
11: $\underline{x}_k^{e(\tau,s)} \leftarrow \underline{x}_k^{*(\tau,s)} + \underline{f}_k^{\text{oamp}(\tau,s)*} (\underline{y}_k - \underline{h}_k \underline{x}_k^{*(\tau,s)})$
12: **end for**
13: Apply IFFT on $\underline{\mathbf{x}}^{\mathbf{e}(\tau,s)}$ to get $\mathbf{x}^{\mathbf{e}(\tau,s)}$.
14: **for** $k = 1 \dots K$ **do**
15: Compute the posterior PMF $\mathcal{D}_k^{(\tau)}(\alpha)$ with Eq. (2.53), for $\alpha \in \mathcal{X}$, and APP estimates $(\mu_{x,k}^{d(\tau,s+1)}, \gamma_{x,k}^{d(\tau,s+1)})$ with Eq. (2.98).
16: **end for**
17: Generate the raw EP feedback $\mathbf{x}^{*(\tau,s+1)}$ with Eqs. (4.32)-(4.33), and estimate its reliability with $\hat{v}^{*(\tau,s)} \leftarrow [\sum_k |\underline{y}_k - \underline{h}_k \underline{x}_k^{*(\tau,s)}|^2 - K\sigma_w^2] / \sum_k |\underline{h}_k|^2$.
18: **end for**
19: Compute extrinsic LLRs $L_e^{(\tau)}(\mathbf{d})$ with Eq. (2.54).
20: Execute de-interleaving, SISO decoding and interleaving to get $L_a^{(\tau+1)}(\mathbf{d})$.
21: **end for**

Vector Approximate Message Passing (VAMP) Another EP-related AMP technique is VAMP [RSF17], whose main distinguishing aspect is its SVD-based implementation and its damping heuristics [RSF17].

When VAMP is cast into the turbo FDE problem, the compact SVD decomposition in Algorithm 18 is roughly equivalent to DFT, with the singular values $\{s_r\}_{r=1}^{R}$ corresponding to non-zero absolute values of the channel frequency response $\{|\underline{h}_k|\}_{k=1}^{K}$, ordered with a decreasing norm. Hence the main difference with the proposed double-loop EP based technique is the damping. VAMP damping procedure (eqs. (26)-(27) in [RSF17]) applies exponential smoothing on the non-linear APP estimate μ_k^d, and on the linear estimator's precision $1/v^e$ (lines 11 and 20 of Algorithm 18), unlike our proposed approach which smooths the extrinsic non-linear estimator's mean and variance (line 18 of Algorithm 13).

While, in a sense, VAMP generalizes the proposed framework of using scalar EP for TD variables and diagonal for FD variables to a general transform domain, its scope is limited to the generalized linear model, while EP message passing is applicable to more complex graphical models. In [Li+19], the proposed FD SILE-EPIC receiver is studied with the damping

Algorithm 18 Turbo FDE with *Vector Approximate Message Passing (VAMP)*

Input $\mathbf{y}$, $\mathbf{H}$
Output $\hat{\mathbf{b}}$

$L_a^{(0)}(\mathbf{d}) = \mathbf{0}_{K_c}$.
Perform the compact SVD $[\mathbf{U}, \mathbf{V}, \mathbf{s}] \leftarrow \mathbf{H}$, such that $\mathbf{H} = \mathbf{U}\mathbf{Diag}(\mathbf{s})\mathbf{V}^H$ is of rank R.
Compute the transform of observations $\tilde{\mathbf{y}} = \mathbf{U}^H\mathbf{y}$.
for $\tau = 0 \dots \mathcal{T}$ **do**
  Compute prior PMFs $\mathcal{P}_k^{(\tau)}(\alpha)$ with Eq. (2.39) and $L_a^{(\tau)}(\mathbf{d})$, for $\alpha \in \mathcal{X}, k = 1, \dots, K$.
  Compute $\mathbf{x}^{\mathbf{p}(\tau)}$ and $\mathbf{v}_{\mathbf{x}}^{\mathbf{p}(\tau)}$ with Eq. (2.40), and set $(\mathbf{x}^{\mathbf{d}(\tau,0)}, v_x^{d(\tau,0)}) \leftarrow (\mathbf{x}^{\mathbf{p}(\tau)}, v_x^{p(\tau)})$.
  for $s = 0$ to $\mathcal{S}_\tau$ **do**
    Apply the transform on data estimates with $\tilde{\mathbf{x}}^{\mathbf{d}} = \mathbf{V}^H\mathbf{x}^{\mathbf{d}}$.
    $\xi_x^{\mathrm{vamp}(\tau,s)} \leftarrow K^{-1}\sum_{r=1}^{R} s_r^2/(\bar{\sigma}_w^2 + v_x^{*(\tau,s)}s_r^2)$
    $v_x^{e(\tau,s)} \leftarrow \xi_x^{\mathrm{vamp}(\tau,s)-1} - v_x^{d(\tau,s)}$
    Apply damping with $v_x^{e(\tau,s)} \leftarrow (\beta/v_x^{e(\tau,s-1)} + (1-\beta)/v_x^{e(\tau,s)})^{-1}$, for $s > 0$.
    for $r = 1 \dots R$ **do**
      $\underline{f}_r^{\mathrm{vamp}(\tau,s)} \leftarrow \xi_x^{\mathrm{vamp}(\tau,s)-1} s_r/(\bar{\sigma}_w^2 + v_x^{d(\tau,s)}s_r^2)$
      $\tilde{x}_r^{e(\tau,s)} \leftarrow \tilde{x}_r^{d(\tau,s)} + \underline{f}_r^{\mathrm{vamp}(\tau,s)*}(\tilde{y}_r - s_r\tilde{x}_r^{d(\tau,s)})$
    end for
    Apply the inverse transfrom on data with $\mathbf{x}^{\mathbf{e}} = \mathbf{V}\tilde{\mathbf{x}}^{\mathbf{e}}$.
    for $k = 1 \dots K$ **do**
      Compute the posterior PMF $\mathcal{D}_k^{(\tau)}(\alpha)$ with Eq. (2.53), for $\alpha \in \mathcal{X}$, and APP estimates $(\mu_{x,k}^{d(\tau,s+1)}, \gamma_{x,k}^{d(\tau,s+1)})$ with Eq. (2.98).
    end for
    Apply exp. smoothing on $\mu_{x,k}^{d(\tau,s+1)} \leftarrow \beta\mu_{x,k}^{d(\tau,s)} + (1-\beta)\mu_{x,k}^{d(\tau,s+1)}$.
    Set $(\mathbf{x}^{\mathbf{d}(\tau,s+1)}, v_x^{d(\tau,s+1)}) \leftarrow (\mathbf{x}^{*(\tau,s+1)}, v_x^{*(\tau,s+1)})$ with the raw EP feedback computed with Eqs. (4.32)-(4.33).
  end for
  Compute extrinsic LLRs $L_e^{(\tau)}(\mathbf{d})$ with Eq. (2.54).
  Execute de-interleaving, SISO decoding and interleaving to get $L_a^{(\tau+1)}(\mathbf{d})$.
end for

heuristics of VAMP, with a focus on proposing an adaptive damping strategy. However in the following it is shown that the considered damping schemes enable reaching lower error rates than VAMP, for which a learning-based optimization method is given in the next section.

Conclusions on AMP-like algorithms AMP-like methods have been derived to reduce the complexity of original inference algorithms, but for the considered communications problem, inference methods have similar complexity. BER performance of turbo receivers based on AMP-like algorithms, PDA, and those derived with approximate inference methods such as GaBP and EP are provided in Figures 4.10 and 4.11, with a block length $K = 256$ and RSC channel code $[1, 5/7]_8$, in the Proakis C channel.

In Figure 4.10, FDEs based on AMP methods are compared with the proposed FD SILE-

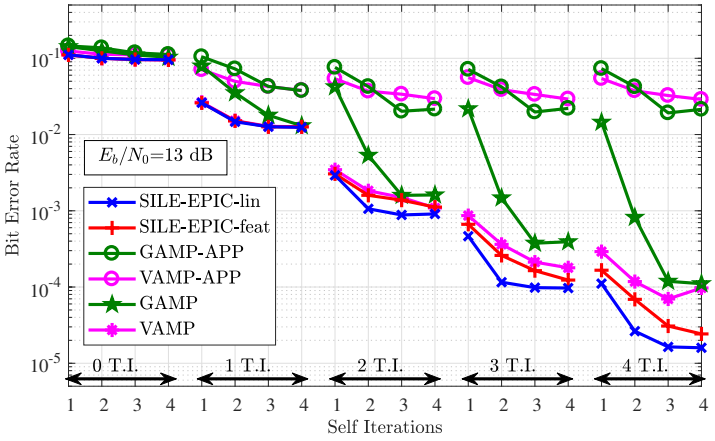


Figure 4.10: BER for coded 8-PSK, with static BER-optimized damping over iterations.

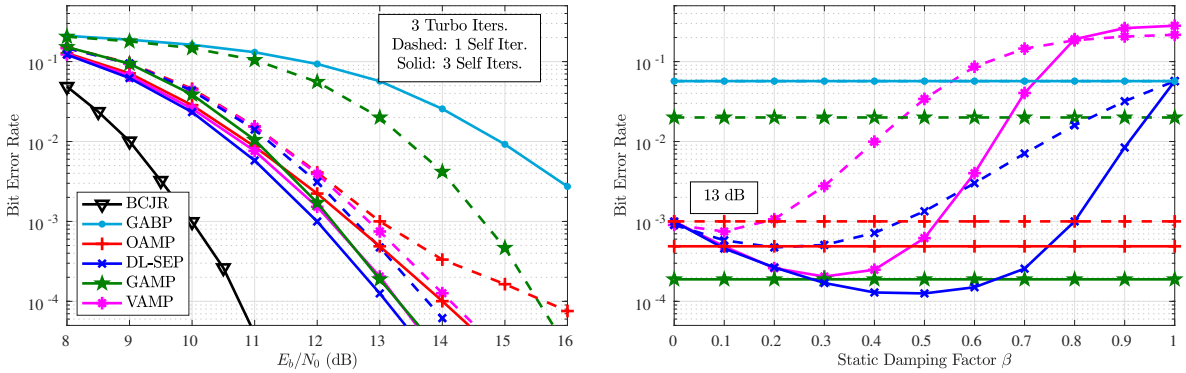


Figure 4.11: Coded 8-PSK optimized BER (at left) and damping sensitivity (at right).

EPIC, obtained with DL-SEP, with both exponential smoothing (linear combination, denoted "lin") damping from Equation (4.40) and feature-based (geometric combination, denoted "feat") damping from Equation (4.39). Receivers that involve damping parameters are optimized, by brute-force search, for each value of SNR, $\mathcal{S}$ and $\mathcal{T}$. Numerical results show that our original proposal FD SILE-EPIC converges to further lower error rates than AMP-based alternatives, and the decrease is BER is faster. It is seen that the exponential smoothing-based damping achieves lower error rates, but feature-based damping still remains attractive for a few turbo iterations. Moreover, the original formulations of GAMP and VAMP with the APP outputs are compared (denoted with "-APP") and numerical results indicate that they are not the best suited for turbo-detection use-case originally (but rather minimizing the MSE of estimates).

In Figure 4.10, at left, the same situation is evaluated for varying SNR, and GaBP (i.e. FD LE-EXTIC) and BCJR algorithms respectively provide an upper and a lower bound on achievable BER performance as conventional methods, and AMP methods and DL-SEP (with exponential smoothing) are compared. While GAMP, OAMP and VAMP approach DL-SEP performance as self-iterations increase, GAMP has a slower convergence speed, and OAMP has a diversity loss at high SNR, due to the sub-optimal feedback variance estimation. The

right side of Figure 4.11 shows the sensitivity of BER to the changes in a static damping parameter β, showing that there are locally robust optimum damping values.

4.3.5 Discussion

Finite-length error rate performance and the asymptotic analysis show that the proposed SC-FDE receiver, obtained by the considered EP-based message passing framework, DL-SEP, is able to bring significant performance improvements for the mitigation of ISI for circular single carrier transmissions. This receiver outperforms similar alternative receivers (single-tap FDE), and performs almost identically to the exact TD block receivers while having a significantly lower computational complexity. Similarities between EP-based approximate-inference and AMP-like algorithms are laid out, and for the considered FDE problem, EP-based inference is shown to reach lower error rates among other self-iterated turbo-receivers.

DL-SEP framework, which instantiates EP with multivariate white Gaussian distributions for TD variables, has been exposed in this elementary SC-FDE system to improve readability and to simplify the analysis. In the following sections, the extensions and applications of this framework to more complex communication systems is discussed.

4.4 Optimizing FD SILE-EPIC with Deep Unfolding

4.4.1 Optimality of EP-based Detectors for Deep Networks

In Section 2.5.3, the concept of deep unfolding and the improvements it had brought upon iterative thresholding and AMP algorithms was briefly discussed. Indeed, when the whole matrix parameters are considered as trainable parameters, such structures are shown to behave like multi-layer perceptrons with residual feedforward connections [GL10; BS16]. These networks are well-suited for detection, and in particular for the learned-AMP network, the feedforward connection that corresponds to the Onsager reaction term is shown to "Gaussianize" residual error at the output of the each layer.

In [BS16], this concept has been extended to VAMP, with it being unfolded as a layer that consists of a fully-connected layer with a residual feedforward connection from the input followed by a *decoupling* function, an activation function and another decoupling function, as shown in Figure 4.12. The decoupling functions compute extrinsic estimates by performing division of a Gaussian PDF corresponding to the two previous outputs of fully-connected layer and the activation function. This structure outperforms previous deep networks based on AMP or iterative thresholding. When trained with well-selected activation functions it also has the remarkable property of yielding the performance of the VAMP algorithm itself (with genie channel knowledge) [BSR17]. This indeed implies that the EP-like VAMP algorithm naturally yields the optimal parameters of the structure in Figure 4.12, consisting of an affine

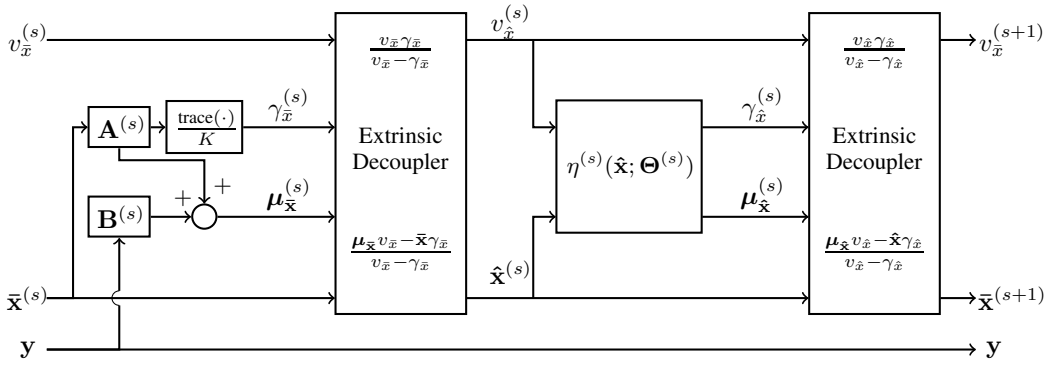


Figure 4.12: Deep unfolding of Vector Approximate Message Passing (VAMP).

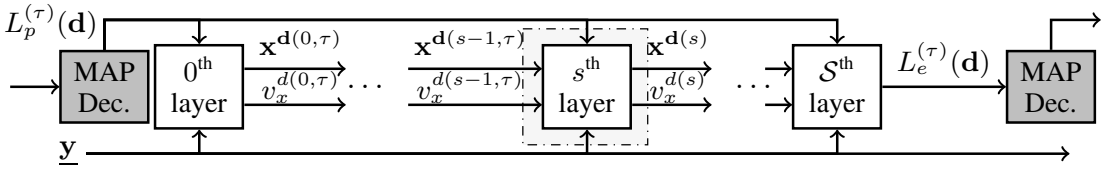


Figure 4.13: "Learned-DL-SEP": Unfolded deep EP network at the τ^{th} turbo iteration.

layer, an extrinsic decoupler, a non-linear function, and another decoupler[3].

Returning into an iterative detection context, this observation suggests that when channel parameters are available, VAMP/OAMP/EP-like algorithms (with no damping) that use extrinsic estimates, can perform as good as deep networks that would require exhaustive training for learning underlying fully-connected layers. Hence these iterative algorithms can be used to determine an architecture for a neural network, by prescribing involved connections, and initializing such detection networks' main parameters, and deep learning can be used to learn remaining additional parameters in order to fine-tune these architectures.

Note that this observation is for raw VAMP/OAMP/Scalar EP-like algorithms ($\beta = 0$), but as Figure 4.11 attests, exponential smoothing improves these algorithms, and the BER sensitivity is a smooth function with a local extrema on β. Hence, unfolding the detection graph of DL-SEP is investigated, by considering each self-iteration as a neural layer, where most parameters are computed through EP, as they should be near-optimal, but damping coefficient become learning parameters $\boldsymbol{\Theta} = [\beta^{(1)}, \ldots, \beta^{(\mathcal{S})}]$ as shown in Figures 4.13 and 4.14. The corresponding neural structure is akin to a network of convolutional layers with residual connections and where layer outputs are linearly mixed with its inputs, with weight β.

[3]Note that this has been numerically verified in [BSR17] for Bernoulli-Gaussian symbols, and we base our investigations on the conjecture that this would be same for zero-mean constellations.

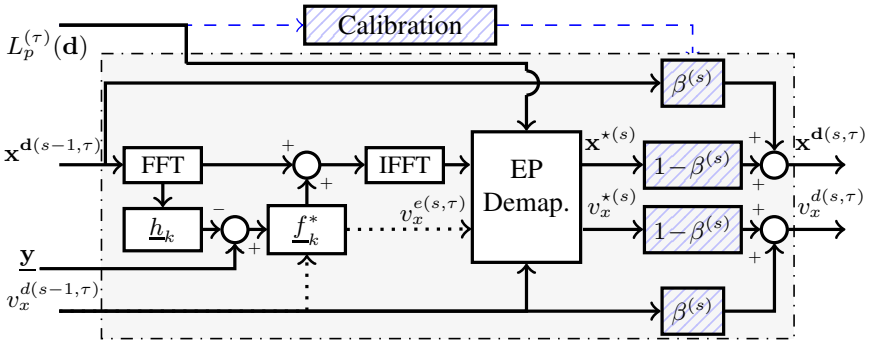


Figure 4.14: "Learned-DL-SEP": s^{th} neural equalization layer.

4.4.2 An MI-based Cost Function for Unfolding Turbo Receivers

To optimize DL-SEP, a loss function $\mathcal{L}$ is proposed, to track the turbo detection and decoding dynamics within the SISO detector, by taking into consideration the decoder outputs. The main idea is to use a loss metric that is correlated to the output BER that also accounts for a quality indicator on a priori LLRs fed from the decoder.

The binary-cross entropy between soft-bit outputs and the transmitted bits is a commonly used bit-wise measure; hence to access SISO capabilities of DL-SEP, it can be used to assess soft-bits on the *extrinsic* LLRs of the module, with

$$\ell(d_{k,q}, \hat{d}_{k,q}) \triangleq -\log\left((\hat{d}_{k,q})^{(1-d_{k,q})}(1-\hat{d}_{k,q})^{d_{k,q}}\right),$$

where $\hat{d}_{k,q} \triangleq 1/(1+\exp(-L_e(d_{k,q}))$ is the extrinsic soft bit. This loss function corresponds to the Kullback-Leibler divergence between transmitted and estimated bits.

For the complete loss metric, inspired from EXIT function synthesis methodology [TB00] (also see Figure 2.6), the neural network is fed with a set of N_{s} sample codewords of Gaussian-distributed prior LLRs, corresponding to a priori information I_A, $L_p(d_{k,q}, I_A)[n] \sim \mathcal{N}((1-2d_{k,q})\mu_p, 2\mu_p)$, $n = 1, \ldots, N_{\text{s}}$, with $\mu_p = J^{-1}(I_A)$ where J is the binary MI function, defined in Property 1 of Section 2.3.6. Then the detector's extrinsic LLRs, for these samples, are $L_e(\mathbf{d})[I_A, n]$, and the learning cost function is

$$\mathcal{L}(\mathbf{d}, \hat{\mathbf{d}}, I_A) \triangleq \frac{1}{QKN_{\text{s}}} \sum_{k,q} \sum_{n} \ell(d_{k,q}, \hat{d}_{k,q}[I_A, n]), \tag{4.41}$$

where $\hat{d}_{k,q}[I_A, n]$ is the soft-bit related to $L_e(\mathbf{d})[I_A, n]$.

This loss function enables learning optimal values of $\boldsymbol{\Theta}$ for a given I_A, and there is a bijective mapping between I_A and the prior variance $v^{a(\tau)} \triangleq K^{-1}\sum_k \text{Var}_{\mathcal{P}_k^{(\tau)}}[x_k]$, where $\mathcal{P}_k^{(\tau)}$ is the PMF in Equation (2.39). Thus, trained parameters are tabulated as a function of $v^a \in [0, \sigma_x^2]$, and the receiver adjusts its weights $\boldsymbol{\Theta}^{(\tau)}$, with the measured $v^{a(\tau)}$ and linear interpolation, at the ongoing turbo-iteration τ.

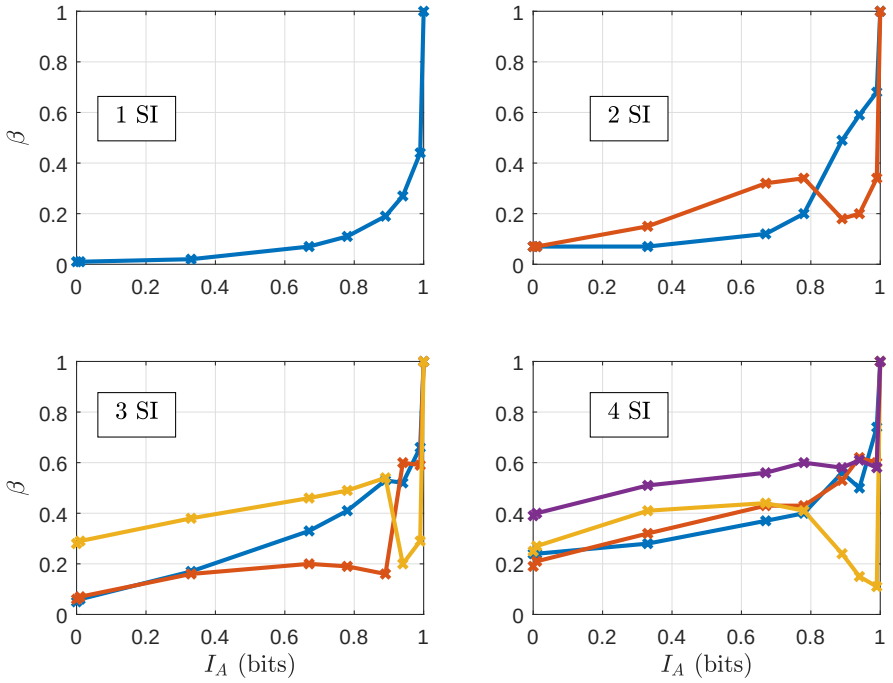


Figure 4.15: Values of β optimized through learning with respect to I_A for $\mathcal{S} = 1, \ldots, 4$.

4.4.3 Learning for the Deep EP Network

Training is carried out with the ADAM optimizer [KB15], with an initial learning rate of 0.025 and mini-batches consisting of 200 samples of:

- a value of σ_w^2, from uniformly distributed $\mathrm{SNR}_c = 20 \log_{10} \sigma_x / \sigma_w$, on an interval of SNR of interest,
- a "dummy codeword" $\mathbf{d}$ from 2^{K_d} i.i.d. possible sequences,
- a noise vector $\mathbf{w}$ and a channel realization $\mathbf{H}$ according to σ_w^2 and the channel profile,
- prior LLRs realizations on the codeword $L_a(\mathbf{d}, I_a)[n]$, with $n = 1, \ldots, N_\mathrm{s}$ samples.

This learning strategy enables fine optimization of the DL-SEP algorithm, when considered as a deep network.

Hence, we investigate the proposed unfolded DL-SEP, to check whether deep learning can automatically optimize DL-SEP parameters to predict its optimum behaviour. For considering a highly-selective situation, training is considered in the fixed Proakis C channel, with $K = 256$, RSC code with polynomial $[1, 5/7]_8$ and $\mathrm{SNR}_c \in [5, 20]$ dB. Prior LLRs realization samples of $N = 25$ is found to be sufficient, with 150 training iterations, to have a precision within 0.05 on β. Weights are learned for $I_A \in \{0, 0.33, 0.67, 0.78, 0.89, 0.94, 0.99, 1\}$, and the learned parameters are plotted in Figure 4.15.

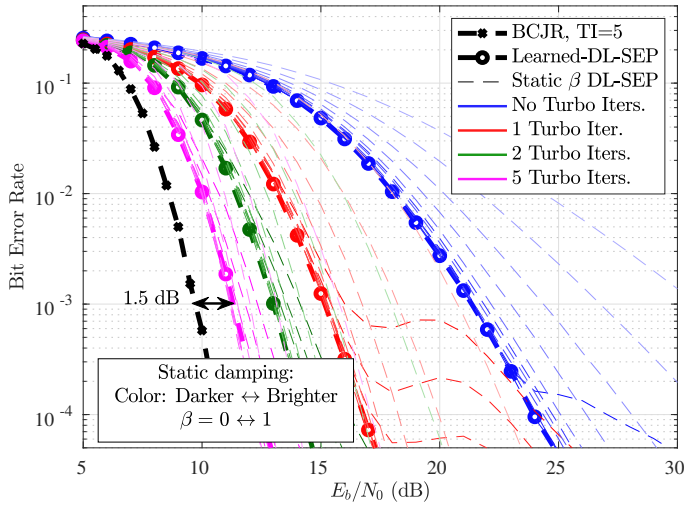


Figure 4.16: Comparison, in Proakis C with coded 8-PSK, of DL-SEP with static damping and 3 self-iterations. and Learned-DL-SEP with 3 layers.

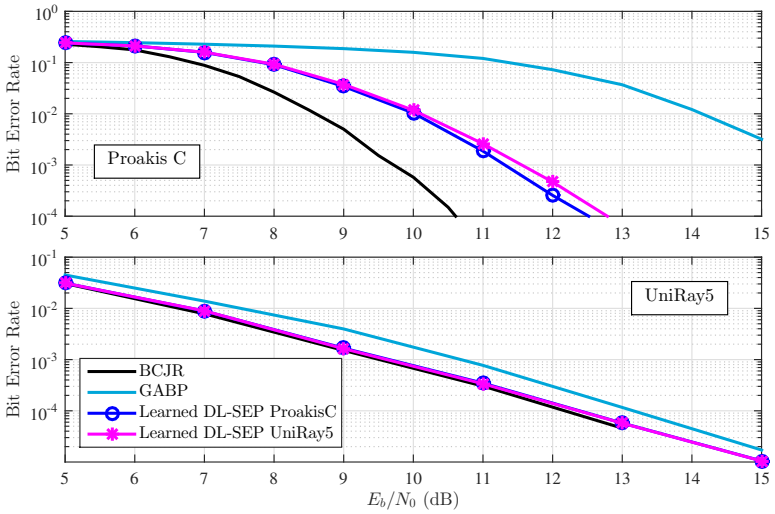


Figure 4.17: Comparison of Learned-DL-SEP with 3 layers for 8-PSK, trained and evaluated in two different channels, at 5 turbo-iterations.

The performance of "Learned-DL-SEP" is shown in Fig. 4.16 along with DL-SEP with static damping β (across self-iterations), with β varying between 0 and 1 with 0.1 steps. While DL-SEP with low damping has good detection threshold, it suffers from error propagation at high SNR, oppositely high damping slows down convergence. "Learned-DL-SEP" manages thus to dynamically adapt to the situation, as deep learning allows us to find optimal values of $\beta^{(s,\tau)}$ as a function of $v^{a(\tau)}$ (dynamic damping). In Fig. 4.16, DL-SEP appears to reach the convex-hull of its feasible set of BER performance. In the end, the "learned-DL-SEP" with 3 layers, is within 1.5 dB of BCJR, at BER = 10^{-3} for $\mathcal{T} = 5$.

To pursue the analysis of the Deep EP network, training is now carried out on a Rayleigh fading channel with uniform power profile with $L = 5$ (denoted UniRay5), and hence, the

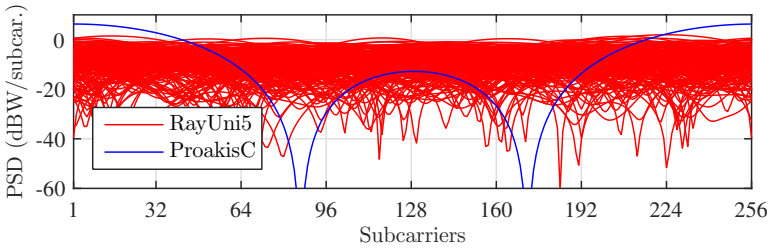


Figure 4.18: Power spectral density for Proakis C and UniRay5 channels.

learning process accounts for many different channel realizations. In previous works, it was noted that training in variable channels still ensures good detection performance for difficult channels (ill-conditioned channel matrix) [SDW17]. In this work, we are also interested in seeing how would a network, trained in a difficult channel, would perform in random channels.

To evaluate this, in Figure 4.17, the performance of Proakis-C-trained network and UniRay5-trained network are compared in both Proakis C and UniRay5 validation sets. It is shown that while UniRay5-trained network performs within 0.3-0.5 dB of the Proakis C trained network's BER, in the Proakis C channel, the Proakis-C-trained network performs identically to the UniRay5-trained network. Fig. 4.18 shows the power spectral density of 200 random UniRay5 channels, and the Proakis C channel; the latter has a significant spectral null region. This suggests that training sets with highly selective, difficult channels should possibly enable a learned receiver to perform near-optimally also in less selective channels.

4.4.4 Discussion

Deep unfolding is shown to be a means to optimize the performance of FD SILE-EPIC with relative ease, and a reasonable complexity. This is enabled by the proposed turbo-oriented learning loss function, whose utility goes beyond the scope of this specific application and receiver, to any soft-input soft-output detector. Finally, the impact of choosing the training set in more or less mild conditions is shown to impact the scope of optimality of such receivers.

4.5 Extensions and Applications of SC-FDE with EP

4.5.1 Direct Applications of FD-SILE-EPIC for CP or ZP Insertion

As the proposed receiver is derived with a generic circular channel model, it can be applied to numerous signalling schemes related to SC-FDE which we will discuss in this section. These strategies that consist in pre-processing and post-processing a transmitted data block in order to transform the impact of the typical channel convolution into a circular convolution, for enabling the underlying hypotheses used for FD processing.

CP SC-FDE The use of a Cyclic Prefix (CP) or a cyclic suffix is a widely used approach for obtaining and ideal circular channel model at the base-band, with uncorrelated noise. Thus the proposed receiver in Algorithm 13 can be applied, as it is, on any SC schemes involving CP. This approach belongs to the model of the general context of this thesis, by considering the PHY model presented in Figures 1.5 and 1.6, by using a SC waveform with

$$\mathbf{F}_{\text{CP}} = [\mathbf{0}_{K_{\text{CP}} \times K}; \mathbf{I}_K], \quad \mathbf{G}_{\text{CP}} = [\mathbf{0}_{K \times K_{\text{CP}}}, \mathbf{I}_K], \tag{4.42}$$

where the CP length is $K_{\text{CP}} \geqslant L$, and the number of baseband samples in Equation (1.11) is $N' = K + K_{\text{CP}}$.

ZP SC-FDE An alternative to CP is the use of Zero Padding (ZP) transmissions, as in Chapter 3, with the transmitter side using

$$\mathbf{F}_{\text{ZP}} = [\mathbf{0}_{K_{\text{ZP}} \times K}; \mathbf{I}_K; \mathbf{0}_{K_{\text{ZP}} \times K}], \tag{4.43}$$

for waveform generation, with K_{ZP} being the number of zeroes padded before and after the data block. In this case, there are three options for the receiver:

1. First approach is to operate on $(K + K_{\text{ZP}})$-point FFT blocks, considering the latter K_{ZP} zeros as "symbols" to be equalized, and apply Algorithm 13 on the extended data block, by considering that priors on zeroes have zero mean and IID variance σ_x^2. Hence

$$\mathbf{G}_{\text{ZP}} = \left[\mathbf{0}_{(K+K_{\text{ZP}}) \times K_{\text{ZP}}}, \mathbf{I}_{K+K_{\text{ZP}}}\right], \tag{4.44}$$

2. An alternative is to keep operating on $K + K_{\text{ZP}}$ symbols, with the same $\mathbf{G}_{\text{ZP}}$ as above, but by including the prior knowledge on these zeros, i.e. zero-mean and null variance, which modifies the overall priors in Equation (4.32), with $\gamma_x^d \leftarrow K(K + K_{\text{ZP}})^{-1}\gamma_x^d$, and it is more suitable when $K \gg K_{\text{ZP}}$. This is an extension of the "FAST" method in [MWG+02] to iterative equalization.

3. Final approach for ZP SC-FDE, is to use an Overlap and Add (OLA) at the receiver, to artificially circularize the channel over the data blocks with

$$[\mathbf{y}]_{1:K_{\text{ZP}}} \leftarrow [\mathbf{y}]_{1:K_{\text{ZP}}} + [\mathbf{y}]_{K+1:K+K_{\text{ZP}}}, \tag{4.45}$$

   and then use K-point FFT to operate on the data block, which will yield the model in Equation (4.1) with correlated noise. This can be obtained with the operator

$$\mathbf{G}_{\text{ZP}} = [\mathbf{0}_{K \times K_{\text{ZP}}}, \mathbf{I}_K, \mathbf{0}_{K \times K_{\text{ZP}}}] + \left[\mathbf{0}_{K_{\text{ZP}} \times (K+K_{\text{ZP}})}, \mathbf{I}_{K_{\text{ZP}}}; \mathbf{0}_{(K-K_{\text{ZP}}) \times (K+2K_{\text{ZP}})}\right], \tag{4.46}$$

   and receiver can still treat the noise as uncorrelated, and yield similar performance to CP SC-FDE [MWG+02], or alternatively $\bar{\sigma}_w^2$ is replaced by $(K+K_{\text{ZP}})K^{-1}\bar{\sigma}_w^2$ to account for the increase in noise power at early symbols (while still neglecting correlations).

When an exact ZP SC-FDE receiver is used (a block structure without any approximations

on the covariance), then CP SC-FDE is almost always outperformed, as K_{zp} additional zeroes reduce the probability of detection errors [MWG+02]. But the three low-complexity schemes described above have comparable performance to CP without clear-cut conclusions on one of them outperforming the other.

Common block transmission schemes, such as UW SC or pseudo-noise extended SC insert respectively fixed and varying pilot sequences between data-blocks, to avoid IBI and provide means for channel estimation. These schemes can be effectively detected by the ZP SC-FDE receivers above, when pilots and the ISI generated by them are removed as pre-processing. The case of UW transmission will be discussed in Section 4.7.1.

4.5.2 FD-SILE-EPIC for Upsampled Single Carrier Waveforms

Fractionally-Spaced (FS) SC-FDE While the work in this thesis is focused in symbol-spaced detection, by taking $m_{\text{up}} = 1$, in the sampling model of Figure 1.6. In this paragraph, the extension of the proposed FD method to the case where $m_{\text{up}} > 1$ is discussed, with Fractionally-Spaced (FS) equalization [Cio08; BCT15]. In this case, for integer m_{up}, with ideal synchronization, the sampling operation in Equation 1.20 provides multiple observations of data symbols, with

$$y_{k,m} = \sum_{l=1}^{L} h_{l,m} x_{\langle k-l \rangle_K} + w_{k,m}, \qquad \forall k = 1, \ldots, K, \; \forall m = 1, \ldots, m_{\text{up}} \tag{4.47}$$

and by vectorizing over time with $\mathbf{y}_m = [y_{1,m}; \ldots; y_{K,m}]$ and $\mathbf{w}_m = [w_{1,m}; \ldots; w_{K,m}]$, there exists a circulant channel model between $\mathbf{y}_m$ and $\mathbf{x}$, which can be diagonalized through DFT. Then by denoting CFR components as $\underline{h}_{k,m}$ for $k = 1, \ldots, K$ and $m = 1, \ldots, m_{\text{up}}$, and then vectorizing FD observations over fractional samples with $\underline{\mathbf{y}}_k = [\underline{y}_{k,1}; \ldots; \underline{y}_{k,m_{\text{up}}}]$, $\underline{\mathbf{h}}_k = [\underline{h}_{k,1}; \ldots; \underline{h}_{k,m_{\text{up}}}]$ and $\underline{\mathbf{w}}_k = [\underline{w}_{k,1}; \ldots; \underline{w}_{k,m_{\text{up}}}]$, we have

$$\underline{\mathbf{y}}_k = \underline{\mathbf{h}}_k \underline{x}_k + \underline{\mathbf{w}}_k. \tag{4.48}$$

Then FS receiver design is carried out with the following factorization of the joint posterior density of Equation (4.10), but symbolwise FD observation factors are replaced by

$$f_{\text{OBS}_k}(\underline{x}_k) \triangleq p(\underline{x}_k | \underline{\mathbf{y}}_k) \propto \exp\left(- \frac{\| \underline{\mathbf{y}}_k - \underline{\mathbf{h}}_k \underline{x}_k \|^2}{\bar{\sigma}_w^2} \right), \tag{4.49}$$

where $\mathbf{Cov}(\underline{\mathbf{w}}_k)$ is assumed to be $\bar{\sigma}_w^2 \mathbf{I}_{m_{\text{up}}}$. Note that, in the time domain, noise samples can be strongly correlated due to the (partial) matched filter in Figure 1.6, and it could be better in FS systems to avoid matched-filtering if $m_{\text{up}} \geqslant 2$, and let the detector handle ISI. Here we detail the case where noise samples are decorrelated for exposing low-complexity detector expressions, but this framework is applicable to the case where $\mathbf{Cov}(\underline{\mathbf{w}}_k)$ is a non-diagonal matrix, by incorporating a whitening filterbank at the equalizer through its inverse.

Thus, the computations at the OBS_k factor node are updated for FS observations, which

replaces Equations (4.14)-(4.15) with

$$\gamma^{o}_{\underline{x},k} = \frac{\bar{\sigma}^2_w v^{e}_{\underline{x},k}}{\bar{\sigma}^2_w + v^{e}_{\underline{x},k} \underline{\mathbf{h}}^H_k \underline{\mathbf{h}}_k}, \quad \mu^{o}_{\underline{x},k} = \frac{\bar{\sigma}^2_w \underline{x}^{e}_k + v^{e}_{\underline{x},k} \underline{\mathbf{h}}^H_k \underline{\mathbf{y}}_k}{\bar{\sigma}^2_w + v^{e}_{\underline{x},k} \underline{\mathbf{h}}^H_k \underline{\mathbf{h}}_k}. \tag{4.50}$$

and the parameters of the extrinsic messages from OBS_k in Equations (4.17)-(4.18) are replaced by

$$v^{o}_{\underline{x},k} = \frac{\bar{\sigma}^2_w}{\underline{\mathbf{h}}^H_k \underline{\mathbf{h}}_k}, \quad \underline{x}^{o}_k = \frac{\underline{\mathbf{h}}^H_k \underline{\mathbf{y}}_k}{\underline{\mathbf{h}}^H_k \underline{\mathbf{h}}_k}. \tag{4.51}$$

Finally, by inserting these new parameters into Equations (4.30)-(4.31), the equalization structure of Equations (4.35)-(4.38) become

$$\underline{x}^{e}_k = \underline{x}^{d}_k + \sum_{m=1}^{m_{\text{up}}} \underline{f}^{\text{ep}}_{k,m} {}^{*} (\underline{y}_{k,m} - \underline{h}_{k,m} \underline{x}^{d}_k), \quad v^{e}_x = 1/\xi^{\text{ep}}_x - v^{d}_x, \tag{4.52}$$

with

$$\underline{f}^{\text{ep}}_{k,m} = \frac{1}{\xi^{\text{ep}}_x} \frac{\underline{h}_{k,m}}{\bar{\sigma}^2_w + v^{d}_x \sum_{m'=1}^{m_{\text{up}}} |\underline{h}_{k,m'}|^2}, \tag{4.53}$$

$$\xi^{\text{ep}}_x = \frac{1}{K} \sum_{k=0}^{K-1} \frac{\sum_{m'=1}^{m_{\text{up}}} |\underline{h}_{k,m'}|^2}{\bar{\sigma}^2_w + v^{d}_x \sum_{m'=1}^{m_{\text{up}}} |\underline{h}_{k,m'}|^2}. \tag{4.54}$$

This passage shows the extension of the proposed DL-SEP framework to FS equalization, which has an important role to play for closing the gap to optimal BCJR detection, and it can be of used for low complexity detection of emerging Faster-than Nyquist (FTN) waveforms [TPB16; Rad18]. FS equalization is also widely used for detection non-linear modulations such as the continuous phase modulation signals [Cha19]. However FS processing poses difficult problems in channel estimation (in a larger band than the baseband), synchronization and overall spectral efficiency (neighbouring frequency bands should not be occupied).

Remark 2

Note that at the baseband, FS observation model shares a lot of similarities to Single-Input Multiple-Output (SIMO) channel model. Hence this receiver can be, for instance, used for single transmit-antenna and multiple receive-antenna node systems, with m_{up} being the number of receive-antennas.

SC-FDE with Frequency Domain Oversampling (FDO) Frequency Domain Oversampling (FDO) is an extension of ZP SC-FDE from the previous section, where the first approach is used at the receiver, with zeroes being considered as zero-mean IID data symbols. It proposes to append an additional K'_{ZP} zeroes to the received block $\mathbf{y}$ in order to increase the resolution of the FFT and better mitigate channel ISI. In [BY16] it is shown that, under perfect CSI, increasing the size of the FFT can further improve the performance of

ZP SC-FDE compared to CP SC-FDE. Moreover the validity of the assumption that padded zeroes are zero-mean IID, used for the first approach of ZP detection, is verified by using a receiver that appends IID unit-variance Gaussian variables instead of zeroes, which causes small degradation.

The proposed FD SILE-EPIC can also be extended to cover this case, by increasing the resolution of FFT at the factor node DFTx to work with $K + K_{\text{ZP}} + K'_{\text{ZP}}$ TD and FD symbols. Padded FD symbols are connected to the factor nodes OBS_k to use zero-padded and transformed FD observations, and padded TD symbols are not connected to DEM_k, but to a novel ZP_k factor node such that $f_{\text{ZP}_k}(x_k) \triangleq \mathcal{CN}(x_k; 0, \sigma_x^2)$.

Single-Carrier Frequency Division Multiple Access (SC-FDMA) The SC-FDMA modulation enables transmitting data on different frequency bands, possibly for different users or services, by mapping K symbols, of a user or a service, to K sub-carriers among a total of $N > K$, through the use of

- a K-point DFT,
- a sub-band mapping matrix $\mathbf{F_{SB}}$, of size $N \times K$,
- an N-point inverse DFT for going back into the time domain.

The sub-band mapping matrix satisfies $[\mathbf{F_{SB}}]_{n,k} = \delta(n - f_{\text{SB}}(k))$, where f_{SB} is an injective function from $\{1, \ldots, K\}$ to $\{1, \ldots, N\}$. In this case, the waveform processing operation at the transmitter (Figure 1.5) becomes

$$\mathbf{F}_{\text{SC-FDMA}} = \mathcal{F}_N^H \mathbf{F_{SB}} \mathcal{F}_K, \tag{4.55}$$

which generates an up-sampled, or spread variant of the data block, depending on the sub-band mapping scheme [Zha+15b], and a CP of K_{CP} samples is appended to it. Hence an SC-FDMA data block occupies $N' = N + N_{\text{CP}}$ channel uses, without accounting for other control signals and pilot sequences.

At the receiver (Figure 1.6), the baseband SC-FDMA demodulator applies a N-point DFT, then the sub-band demapper $\mathbf{F}_{\mathbf{SB}}^H$ provides a FD observation model, if FD detection is performed, as in this chapter, with

$$\mathbf{G}_{\text{SC-FDMA}} = \mathbf{F}_{\mathbf{SB}}^H \mathcal{F}_N. \tag{4.56}$$

Otherwise if a time-domain detection will be carried out, as in Chapter 3, then an additional K-point inverse DFT is performed

$$\mathbf{G}_{\text{SC-FDMA}} = \mathcal{F}_K^H \mathbf{F}_{\mathbf{SB}}^H \mathcal{F}_N. \tag{4.57}$$

Note that the use of a sub-band mapping can significantly change the perception of a physical channel at the receiver, especially in the FD. For instance, considering a wideband

CIR of h'_l, for $l = 1, \ldots, L$, (over the whole band with N subcarriers) generating a circulant channel matrix $\mathbf{H}'$, the baseband SC-FDMA channel at the FD is given by

$$\underline{\mathbf{H}} = \mathcal{F}_K \mathbf{H}' \mathcal{F}_K^H = \mathbf{F}_{\mathbf{SB}}^H \mathcal{F}_N \mathbf{H}' \mathcal{F}_N^H \mathbf{F}_{\mathbf{SB}} = \mathbf{Diag}(\underline{\mathbf{h}}), \tag{4.58}$$

with the channel frequency response component at the k^{th} subcarrier being

$$\underline{h}_k = \sum_{l=1}^{L} h'_l \exp\left(-2j\pi f_{\text{SB}}(k) l/N\right), \tag{4.59}$$

which can be quite different from the CFR in Equation (4.4), depending on N, and the sub-band mapping method.

Numerical examples with this waveform will be given, when discussing the Multi-User Multiple Input Multiple Output (MU-MIMO) problem in Section 4.6.

4.5.3 Time-Varying Channel Equalization with SC-FDE and EP

A notable issue of FDE is its inability to mitigate time-varying channels whose coherence time is shorter than the processing block duration. In this case the FD channel matrix is no longer diagonal, and inter-carrier interference is generated.

On TV Equalization in the Frequency Domain When the channel is time-varying the FD channel matrix has multiple non-zero diagonals, and EP-based receiver design would result in a block linear MMSE structure as in [SMF+18].

A possible, intermediary complexity solution is to use *banded FDE*, where only J upper and lower diagonals of the FD channel matrix $\underline{\mathbf{H}}$ are accounted for (in addition to the main diagonal), and the remaining channel components are neglected [Sch04; RBL05; RBL06; Cha19]. In this case low complexity matrix inversion techniques that can be used to extend DL-SEP framework above for TV channels, through LDL decomposition.

To achieve this, observation factor nodes OBS_k have to be replaced by a single factor node OBS which links $\underline{\mathbf{x}}$ to $\underline{\mathbf{y}}$ through $f_{\text{OBS}}(\underline{\mathbf{x}}) \propto \mathcal{CN}(\underline{\mathbf{y}}; \underline{\mathbf{H}}\underline{\mathbf{x}}, \sigma_w^2 \mathbf{I}_K)$, assuming AWGN. In this case, the posterior parameters over the FD data vector is

$$\begin{aligned} \tilde{q}_{\text{OBS}}(\underline{\mathbf{x}}) &= f_{\text{OBS}}(\underline{\mathbf{x}}) \prod_{k=1}^{K} m_{\text{DFTx}_k \to \underline{x}_k}(\underline{x}_k), \\ &\propto \exp\left(-\frac{\|\underline{\mathbf{y}} - \underline{\mathbf{H}}\underline{\mathbf{x}}\|^2}{\bar{\sigma}_w^2} - \sum_{k=1}^{K} \frac{|\underline{x}_k - \underline{x}_k^e|^2}{v_{\underline{x},k}^e} \right) \propto \mathcal{CN}(\underline{\mathbf{x}}; \boldsymbol{\mu}_{\underline{x}}^o, \boldsymbol{\Gamma}_{\underline{x}}^o). \end{aligned} \tag{4.60}$$

Then following projection onto fully-factorized FD symbol data set, posteriors estimates of

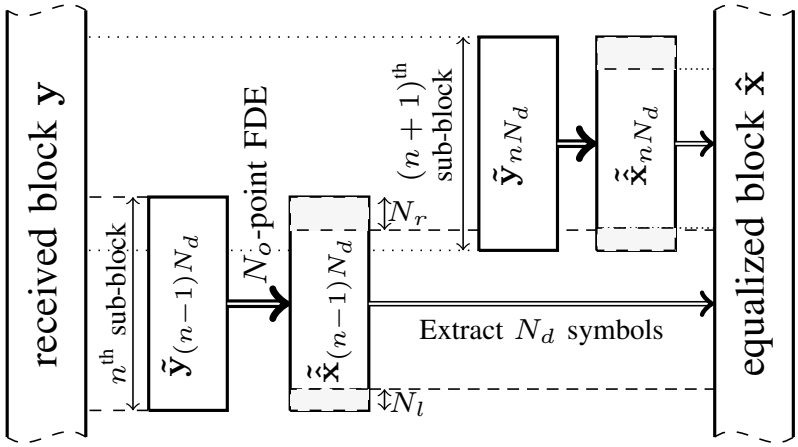


Figure 4.19: Overlap FDE processing scheme with sub-blocks.

factor node OBS on node x_k is $q_{\text{OBS}}(\underline{x}_k) \propto \mathcal{CN}(\underline{x}; \mu^o_{\underline{x},k}, \gamma^o_{\underline{x},k})$, where

$$\gamma^o_{\underline{x},k} = \mathbf{e}_k^H \left(\frac{\underline{\mathbf{H}}^H \underline{\mathbf{H}}}{\bar{\sigma}_w^2} + \mathbf{V}_{\underline{x}}^{\mathbf{e}\,-1} \right)^{-1} \mathbf{e}_k = (1 - v^e_{\underline{x},k} \mathbf{e}_k^H \boldsymbol{\Sigma}_{\underline{\mathbf{y}}}^{-1} \mathbf{e}_k) v^e_{\underline{x},k}, \tag{4.61}$$

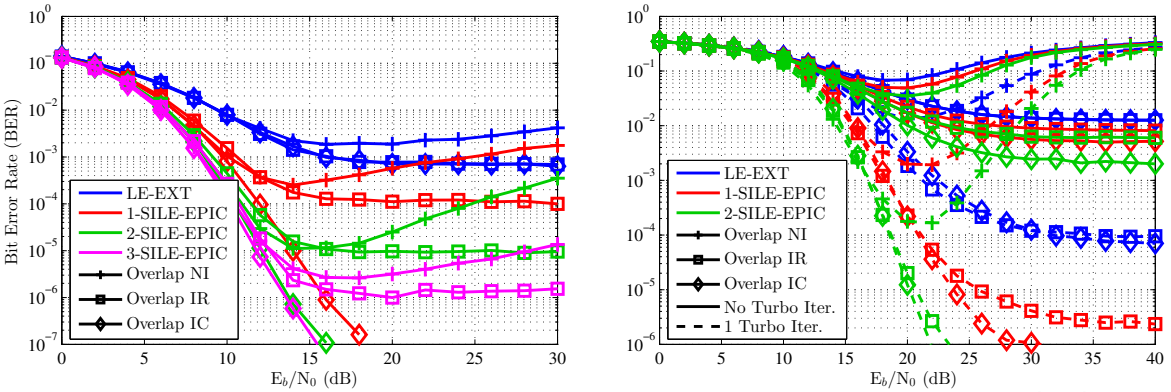
$$\mu^o_{\underline{x},k} = \mathbf{e}_k^H \left(\frac{\underline{\mathbf{H}}^H \underline{\mathbf{H}}}{\bar{\sigma}_w^2} + \mathbf{V}_{\underline{x}}^{\mathbf{e}\,-1} \right)^{-1} \left(\frac{\underline{\mathbf{H}}^H \underline{\mathbf{y}}}{\bar{\sigma}_w^2} + \mathbf{V}_{\underline{x}}^{\mathbf{e}\,-1} \underline{\mathbf{x}}^{\mathbf{e}} \right) = \underline{x}^e_k + v^e_{\underline{x},k} \underline{\mathbf{H}}^H \boldsymbol{\Sigma}_{\underline{\mathbf{y}}}^{-1} (\underline{\mathbf{y}} - \underline{\mathbf{H}} \underline{\mathbf{x}}^{\mathbf{e}}), \tag{4.62}$$

where $\boldsymbol{\Sigma}_{\underline{\mathbf{y}}} = \bar{\sigma}_w^2 \mathbf{I}_K + \underline{\mathbf{H}} \mathbf{V}_{\underline{x}}^{\mathbf{e}} \underline{\mathbf{H}}^H$, and then these Equations have to be inserted into Equation (4.16) and then Equations (4.30)-(4.31) to compute equalized estimates. With banded processing, $\boldsymbol{\Sigma}_{\mathbf{y}}$ is limited to J upper and lower diagonals also, and its inversion through LDL decomposition can be carried out with an order of complexity of $\mathcal{O}(J^2 K)$ [Cha19].

Finally, *overlap FDE* is another possible approach for mitigating problems above, with quasi-linear receiver complexity in the FD, by keeping the use of single-tap equalizers. This technique consists in using N-point FFTs, with $N < K$, to carry out baseband processing, on *virtual* overlapping sub-blocks of received samples [VHG01; Mar+03]. This strategy inherently generates IBI between sub-blocks, which is mitigated either by selecting an appropriate sub-block length N, or by using additional signal processing. Some recent usage examples include its usage with faster-than-Nyquist signalling [FH15], and with doubly selective channels [RNM17].

In the remainder of this section, various EP-based overlap FDE receivers are derived and evaluated.

Conventional Overlap FDE with "No-Interference" Overlap FDE, also called FDE with overlap-and-save or overlap-and-cut, carries out a linear deconvolution with multiple circular convolutions. Given a signal block $\mathbf{v} \in \mathbb{C}^K$, its N_o-point sub-blocks are denoted $\tilde{\mathbf{v}}_k = [v_k, \ldots, v_{k+N_o-1}]^T$, with $v_k = 0$, for all $k < 0$ or $k \geqslant K$. SC-FDE model with sub-blocks



(a) QPSK in Rayleigh fading EQU16 channel, uncoded

(b) 16-QAM in AWGN EQU 7 channel, with RSC $[1, 5/7]_8$ encoding

Figure 4.20: IBI mitigation capabilities of overlap FDE schemes.

is written as

$$\tilde{\mathbf{y}}_k = \mathbf{H}_k \tilde{\mathbf{x}}_k + \mathbf{G}_k (\tilde{\mathbf{x}}_{k-N} - \tilde{\mathbf{x}}_k) + \tilde{\mathbf{w}}_k, \tag{4.63}$$

where $\mathbf{H}_k$ is a $N_o \times N_o$ circular channel matrix as in Equation (4.1), and $\mathbf{G}_k$ is an $N_o \times N_o$ matrix, whose $L-1$ upper diagonals are equal to those of $\mathbf{H}_k$, and other elements are zeros. Unlike the channel model in Equation (4.1), here the channel may quasi-statically vary between sub-blocks. Hence with a small enough N_o, a time-varying frequency selective channel can be approximated by this model.

SC-FDE is used on sub-blocks, by *ignoring the IBI term*, and N_l symbols from the head and N_r symbols from the tail of the equalized sub-block are thrown away. $N_l + N_r$ symbols are overlapping between two successive sub-blocks, as shown in Figure 4.19, and by extracting the remaining $N_d = N_o - N_l - N_r$ symbols, this procedure is repeated for $N_b = \lceil K/N_d \rceil$ sub-blocks in parallel. For extending this scheme to use the proposed EP-based framework from the previous sections; one could implement each equalizer of length N_o using FD SILE-EPIC. Hence each sub-equalizer would have its own self-iteration loops, and independently evolving estimate variances. But as BICM is used across all sub-blocks, differences of estimate variances between sub-blocks is small, hence, for simplicity, all the sub-equalizers (FD SILE-EPIC) are assumed to use a common SI loop, with the common output variance denoted $v_x^e \triangleq N_b^{-1} \sum_{n=1}^{N_b} v_{x,n}^e$, where $v_{x,n}^e$ is the n^{th} sub-equalizer's output variance, and the common demodulator feedback variance is v_x^d.

We denote this overlap FDE scheme, no-interference (NI), its performance at a given SNR can be close to that of SC-FDE with K-point FFT, if N_o, N_l and N_r are sufficient to remove all residual IBI over N_d extracted symbols [VHG01; Mar+03; FH15]. However, for moderately or highly selective channels, the IBI spread can be very large, requiring $N_o \gg N_d$. Otherwise, residual IBI is present, and causes detection errors, whose occurrence increases with the SNR, due to interference enhancement caused by mismatched filter weights.

Overlap FDE with Interference Rejection Interference enhancement caused by overlap FDE NI causes prohibitive constraints for selecting N_o, N_l and N_r, in order to avoid residual

IBI. Moreover, for channel with severe spectral nulls, IBI spread can be as large as N_o, making overlap FDE unusable with any parameters. If the channel is also time-varying, coherence time constraints on N_o are imposed, which may cause overlap FDE NI to have no viable solution, due to excessive complexity costs, or increased latency.

In this paragraph, the interference rejection (IR) strategy, which mitigates interference enhancement, is exposed by designing filters that account for the presence of IBI. The equivalent noise which also includes the IBI is

$$\tilde{\mathbf{w}}'_k \triangleq \mathbf{G}_k(\tilde{\mathbf{x}}_{k-N} - \tilde{\mathbf{x}}_k) + \tilde{\mathbf{w}}_k. \tag{4.64}$$

Considering the noise model (see eq. (4.1)) used in the FDE design, in section 4.3.2, one can compute a SC-FDE equalizer, as in Equations (4.35)-(4.38), by using the equivalent noise covariance

$$\mathbf{\Sigma}_{\underline{\tilde{\mathbf{w}}}'_k} = \mathbf{\Sigma}_{\underline{\tilde{\mathbf{w}}}_k} + 2\sigma_x^2 \mathcal{F}_N \mathbf{G}_k \mathbf{G}_k^H \mathcal{F}_N^H, \tag{4.65}$$

assuming IID transmitted symbols. The equalizer neglects noise correlations between different subcarriers, but accounts for the FD colored noise with diagonals of matrix (4.65). IR was applied using the whitened covariance of the IBI in [TTA06], however using a colored representation, as in this Chapter, was shown to significantly improve performance [OA11]. This strategy does not suffer from error enhancement at high SNR, and produce steady error-floors. Nevertheless IR can perform slightly worse than NI at low SNR, due to pessimistic representation of IBI covariance.

Overlap FDE with Interference Cancellation To completely remove residual IBI in overlap FDE with limited overlap interval, IC is needed, especially for highly selective channels, where equalization filter has time response of length comparable to FFT, and spreads IBI over all symbols.

There are various approaches to IBI cancellation in overlap FDE, either with serial decision feedback for joint ISI/IBI cancellation [Tom05], with hard decision feedback for successive IBI cancellation [WLL08], or with hybrid turbo and hard successive decision feedback [RNM17]. Unlike these references, which uses decisions on previously processed sub-blocks, here we focus on parallel IBI cancellation, using solely a feedback generated from the previous self or turbo iteration, for ensuring parallel processing of sub-blocks in practical implementations. Moreover, EP-based feedback is used, as its overall superiority compared to EXT or APP feedback was shown in the previous section.

At $\tau = s = 0$, IR is used via (4.65), then IBI is removed before the N_o-point FFTs with

$$\tilde{\mathbf{y}}'_k \triangleq \tilde{\mathbf{y}}_k - \mathbf{G}_k \left(\tilde{\mathbf{x}}_{k-N}^{d(\tau',s')} - \tilde{\mathbf{x}}_k^{d(\tau',s')} \right), \tag{4.66}$$

where τ' and s' denote the previous turbo and self iteration indexes.

Moreover, unlike prior work on overlap FDE-IC, we use adaptive IR, by accounting for

the *residual* IBI in filter weight computations with

$$\mathbf{\Sigma}_{\underline{\tilde{\mathbf{w}}}'_k} = \mathbf{\Sigma}_{\underline{\tilde{\mathbf{w}}}_k} + 2v^{d(\tau',s')}\mathcal{F}_N\mathbf{G}_k\mathbf{G}_k^H\mathcal{F}_N^H. \tag{4.67}$$

As in overlap FDE NI/IR strategies above, N_b parallel equalizers are operated concurrently for detecting all sub-blocks. Finally, it is possible, depending on the channel coherence time, to set $N_l = N_r = 0$, for $\tau > 0$, to reduce N_b, as in [RNM17], to reduce the receiver complexity.

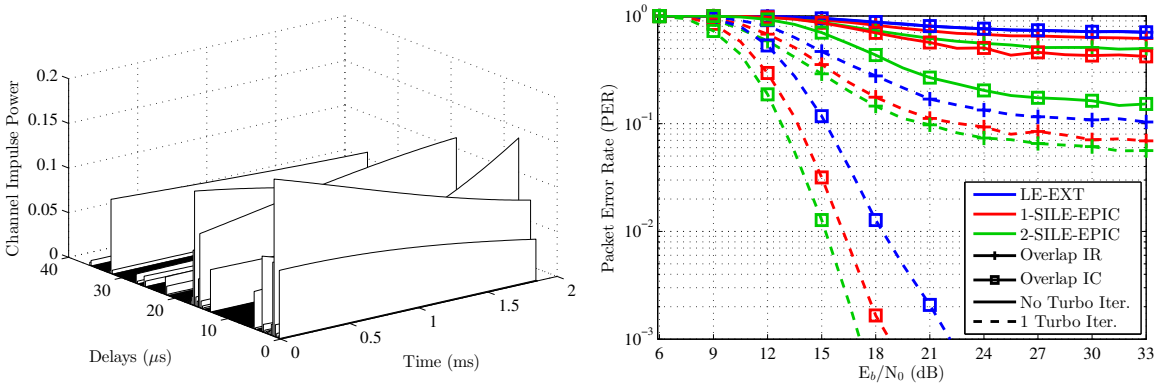
Inter-block interference mitigation performance In this section, K-block quasi-static channels are considered, to focus on the EP-based overlap FDEs' IBI mitigation capabilities. The benefits of self-iterations are compared to the conventional FD LE-EXTIC (i.e. $\mathcal{S} = 0$), for overlap FDE, possibly equipped with IR and/or IC. The IC strategy of setting $N_l = N_r = 0$ for $\tau > 0$ for overlap FDE IC is used for these simulations.

First we consider an uncoded scenario, similar to the benchmark [OA11], with QPSK constellation in a quasi-static Rayleigh fading frequency-selective channel with symbol spaced 16-path uniform power delay profile (EQU16). Transmission parameters are $K = 2048$, $N_o = 256$ and $N_l = N_r = 16$, and 80000 block transmissions per SNR are used to numerically approximate the BER for $\mathcal{S} = 0 \dots 3$ in Figure 4.20-(a). The conventional scheme (NI) is unusable, as the overlap interval is insufficient to contain all IBI, and SIs ($\beta = 0.25 \times 0.5^{s+\tau}$) do not resist to IBI amplifications. But IR significantly benefits from SIs, as it further reduces the error floor. Finally, overlap IC with SIs removes most of the interference, even with a single SI.

A more extreme case, with strong IBI, is considered in Figure 4.20-(b) (16-QAM, RSC $[1, 5/7]_8$), within a 7-path static AWGN channel, with uniform power delay profile. We consider 50000 block transmissions with $K = 1024$, $N_o = 128$, and $N_l = N_r = 7$, to evaluate the BER. In this case, SIs ($\beta = 0.75 \times 0.9^{s+\tau}$) alone cannot remove error floors even with IC and channel coding, but with the help of a single turbo-iteration, even IR's error floor, with EP-based SI, becomes at least two order of magnitudes smaller than traditional FD LE-EXTIC.

Performance in a doubly-selective channel The behaviour of the overlap FDE with the proposed EP-based self-iterations is evaluated within a MANET scenario where mobile-to-mobile communications between two high-speed vehicles is considered in a harsh environment. The mountainous channel model from [Fis16, Tab. 5.10] is used. Vehicles are assumed to move at 130 km/h each, in opposing directions, hence generating a maximum Doppler shift of 96 Hz, assuming the use of a carrier frequency at 400 MHz. A snapshot of a random channel realization is plotted in Figure 4.21-(a).

SC transmissions with 1/2-rate-coded 16-QAM constellation is considered, with a baud-rate of 1 Mbauds/s, and a root raised-cosine pulse-shaping with a roll-off factor of 0.35. In this case, the base-band channel spread is $L = 45$ symbols. $N_o = 256$ symbol is chosen to ensure that the channel remains static on each sub-block. We consider $K = 1536$ and $N_l = N_r = 18$,



(a) A snapshot of the Mountainous channel power delay profile

(b) 16-QAM in the Mountainous channel, with RSC $[1, 5/7]_8$ encoding

Figure 4.21: EP-based overlap FDE performance in the doubly-selective mountainous channel.

an ovelap length of 18 symbols is chosen as most significant paths of the mountainous channel, (and other urban, hilly or rural channels in [Fis16]) are contained within 18 μs.

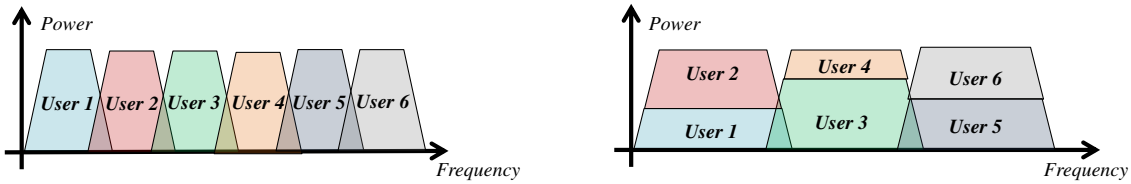
In Figure 4.21-(b), the PER of overlap FDE IR/IC are plotted. It can be seen that IR cannot get rid of the error floor but using overlap IC and one turbo-iteration, robust transmissions are possible. In this case, one and two SIs ($\beta = \min(0.5, 0.7^{1+s+\tau})$) respectively bring 2.7 dB and 3.9 dB improvements, at PER $= 3.10^{-3}$. The use of SC-FDE with six block transmissions of $K = 256$, with cyclic prefix and guard intervals to avoid IBI, instead of using the considered overlap FDE, would have required 90 additional symbol slots per block, and would have caused a loss of throughput and energy-efficiency of respectively 12 % and 0.6 dB.

4.5.4 Discussion and Conclusions

This section has illustrated the proposed receiver design technique various telecommunications applications where the proposed framework could either be applied directly, or be extended, through the message passing framework, to handle detection problem.

4.6 Extension to Multi-antenna and Multiple Access Systems

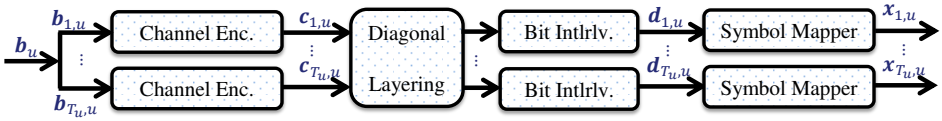
In this section, the extension of the FD detection framework for handling Multiple-User Detector (MUD) is discussed, with also the possibility of incorporating multiple transmit or receive antennas. To this end, first, to cover a practical use case of transmit and receive antennas, transmission schemes that extend BICM to multiple-antennas and multiple users for spatial multiplexing is discussed, and then receiver design for SC-FDMA in a MU-MIMO context is investigated, with the use of the proposed DL-SEP framework.



(a) Orthogonal User Access (b) Non-orthogonal User Access

Figure 4.22: Non-orthogonal versus orthogonal power allocation with SC-FDMA.

(a) Space-Time Bit Interleaved Coded Modulation (STBICM)



(b) Turbo Bell-Labs Layered Space-Time (BLAST)

Figure 4.23: Non-orthogonal transmission mechanisms based on BICM.

4.6.1 On NOMA based on BICM and SIC

As already discussed in broad terms, in the general introduction and the Chapter 1, embracing interference has become a necessity to increase throughput for future wireless communications systems, due to resource scarcity. Fundamental techniques to increase the system load include the use of spatial multiplexing with multiple transmit and receive antennas [TV05], and generalizing this idea, the non-orthogonal sharing of resources by multiple users [Hig+15]. This concept has been known as spatial division multiple-access [TV05] and it is today somewhat related to power-domain NOMA. An illustration of this concept for a multi-user system SC-FDMA is provided in Figure 4.22, where the non-orthogonal stacking of users enable using larger spectral resources, but this also requires handling the ISI caused by frequency-selective channels in the MUD, in addition to the MAI and MUI.

Computationally-efficient FD receivers designed with the MMSE criterion and Successive Interference Cancellation (SIC) can perform quite efficiently in such situations, when received powers between users are sufficiently different [TV05]. But more generally, the resolution of MAI, and handling of NOMA when users are received with similar SNRs push the research community towards the exploration of more advanced techniques. In particular, to enable a more efficient separation of users and antennas at the MUD, the concept of Interleave Division Multiple Access (IDMA) [Pin+06; PWW07; Wan+19] extends BICM scheme to encoding mechanisms over users and antennas. One such technique is STBICM [Ton00], which directly

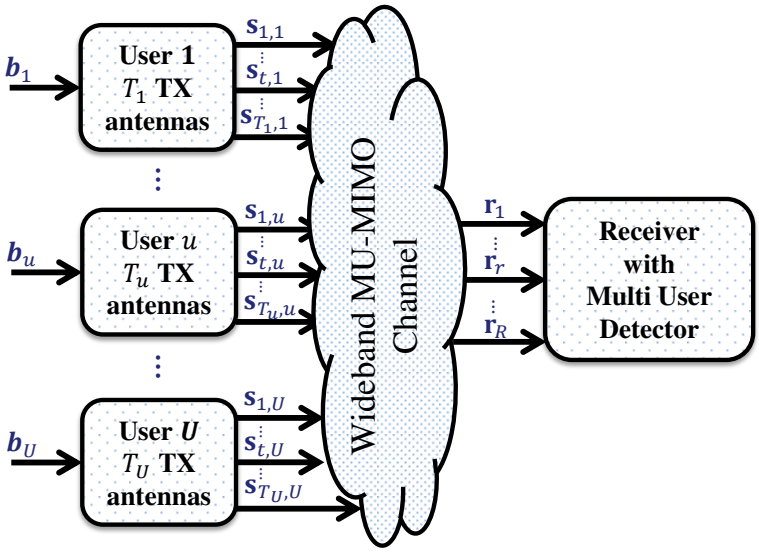


Figure 4.24: MU-MIMO transmissions over a frequency-selective wideband channel.

extends single-antenna single-user receiver design techniques for BICM to higher dimensions, by interleaving coded bits over antennas, and using different interleavers at each user, as shown in Figure 4.23-(a). Alternatively, conventional MIMO signalling schemes based on BLAST architectures are extended to Turbo-BLAST [SH00], which is rather destined for spreading different codewords over antennas through diagonal layering [TV05], as illustrated in Figure 4.23-(b), but it can be considered as a specific code construction approach for STBICM.

With the use of iterative receiver design methods, such as EXIT charts, high performance MUD were proposed in [Yua+08; VBC06]. Similarly to the single-user single-antenna turbo FDE case, discussed earlier in this chapter, various turbo receivers based on decoder's EXT feedback [VBC06], decoder's APP feedback [Wit+02; VBC06], or demodulator's APP feedback [Tao15] were proposed as FD MUD.

In the following the derivation of a low-complexity MUD with DL-SEP framework is discussed for NOMA in the presence of ISI and MAI.

4.6.2 SC-FDMA MU-MIMO System Model

Let us consider a synchronous, open-loop, multi-user spatial multiplexing system with U independent users, possibly with multiple antennas, each employing a Space-Time Bit Interleaved Coded Modulation (STBICM) SC-FDMA modulation to transmit a single data block ($P = 1$) in a multiple-access scenario, simultaneously over a shared frequency-selective channel to a multi-antenna receiver node, equipped with a MUD, as in Figure 4.24, to decode all of the U users.

Each user u in $\{1, \ldots, U\}$ has T_u transmit antennas, with the total number of antennas being $T \triangleq \sum_{u=1}^{U} T_u$, and all users share K sub-carriers and N time samples for the SC-FDMA.

The STBICM scheme for the u^{th} produces the user's data block $\mathbf{x}_{t,u}$, and the associated coded blocks $\mathbf{d}_{t,u}$, for the antennas $t = 1, \ldots, T_u$. In this scheme, users' interleavers may be different (which results in IDMA [Pin+06; PWW07; Wan+19] and helps performing SIC more efficiently), but it is not necessary in general.

The receiver is equipped with R antennas, it is ideally synchronized both in time and frequency and perfect channel knowledge is available. Using a CP of sufficient length for the SC-FDMA, the received signal at the r^{th} antenna, following SC-FDMA FD demodulation at the baseband

$$\underline{\mathbf{y}}_r = \sum_{u=1}^{U} \sum_{t=1}^{T_u} \mathbf{F}_{\mathbf{SB}}^T \mathcal{F}_N \mathbf{H}_{r,t,u} \mathcal{F}_N^H \mathbf{F}_{\mathbf{SB}} \mathcal{F}_K \mathbf{x}_{t,u} + \underline{\mathbf{w}}_r, \tag{4.68}$$

where the $N \times N$ matrix $\mathbf{H}_{r,t,u}$ is the circulant matrix generated by the L-tap channel impulse response $[h_{1,r,t,u}, \ldots, h_{L,r,t,u}]$ between u^{th} user's t^{th} antenna, and the receiver's r^{th} antenna, $\mathbf{F}_{\mathbf{SB}}$ is the sub-band mapping matrix from the Section 4.5.2 and the noise $\underline{\mathbf{w}}_r \sim \mathcal{CN}(0, \sigma_w^2 \mathbf{I}_N)$, i.e. a circularly symmetric AWGN with covariance $\sigma_w^2 \mathbf{I}_N$.

Denoting $\underline{\mathbf{H}}_{r,t,u} = \mathbf{F}_{\mathbf{SB}}^T \mathcal{F}_N \mathbf{H}_{r,t,u} \mathcal{F}_N^H \mathbf{F}_{\mathbf{SB}}$ as the equivalent FD baseband channel, and by stacking transmitted, received and noise FD samples by antennas and by user, on a time instant as $\underline{\mathbf{y}}_k = [\underline{y}_{k,1}; \ldots; \underline{y}_{k,R}]$, $\underline{\mathbf{x}}_k = [\underline{x}_{k,1,1}; \ldots; \underline{x}_{k,T_1,1}; \underline{x}_{k,1,2}; \ldots; \underline{x}_{k,T_u,u}]$ and $\underline{\mathbf{w}}_k = [\underline{w}_{k,1}; \ldots; \underline{w}_{k,R}]$, then we have

$$\underline{\mathbf{y}}_k = \underline{\mathbf{H}}_k \underline{\mathbf{x}}_k + \underline{\mathbf{w}}, \tag{4.69}$$

where $\underline{\mathbf{H}}_k$ is the $R \times T$ MU-MIMO matrix such that $[\underline{\mathbf{H}}_k]_{r,t'} = \underline{h}_{k,r,t,u}$, such that $t' = \sum_{u'=1}^{u-1} T_{u'} + t$.

Remark 3
Note that this MU-MIMO model is equivalent to that of our contribution in [Şa+18e], where we have preferred to work with time-stacked vectors to obtain block partitioned receivers. That approach enabled to identify an exact diagonal EP solution and then carry out approximations to get the receiver we will describe below, whereas here, the DL-SEP framework is applied to directly derive the low-complexity solution.

4.6.3 Multi-User Detection with EP

Factor Graph Model and Message Passing Resolution This section aims to approximate the optimum MAP FD joint detection and decoding for this MUD design problem by using the DL-SEP framework.

Firstly, the factor graph of this system is obtained by factorizing the joint posterior PDF

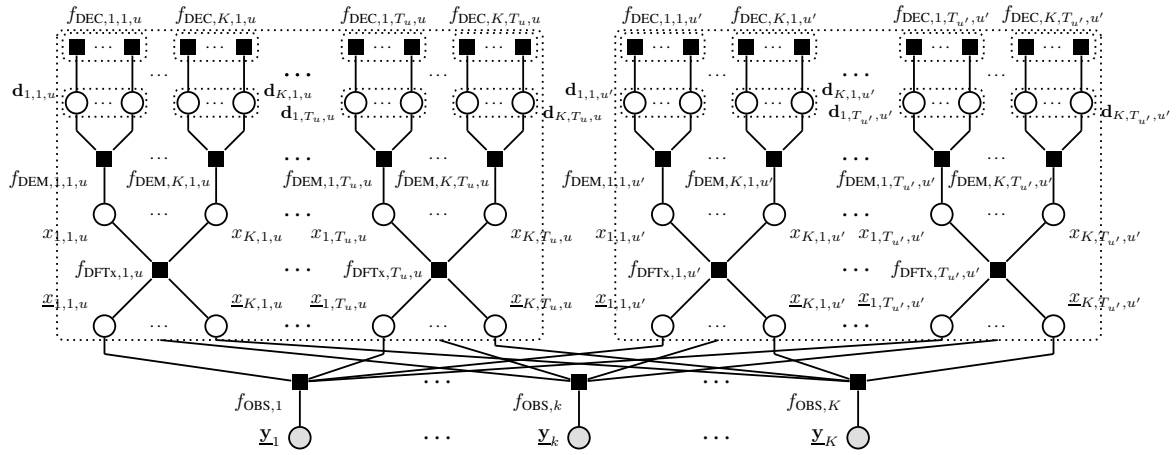


Figure 4.25: MU-MIMO STBICM factor graph with users u and u'.

of this system, with

$$p\left(\{\mathbf{x}_{t,u},\mathbf{d}_{t,u}\}_{u,t}|\{\underline{\mathbf{y}}_r\}_{u,t}\right) \propto \left(\prod_{k=1}^{K}\underbrace{p(\underline{\mathbf{y}}_k|\underline{\mathbf{x}}_k)}_{=f_{\mathrm{OBS}_k}(\underline{\mathbf{x}}_k)}\right)\prod_{u=1}^{U}\prod_{t=1}^{T_u}\underbrace{p(\underline{\mathbf{x}}_{t,u}|\mathbf{x}_{t,u})}_{=f_{\mathrm{DFTx}_{t,u}}(\mathbf{x}_{t,u},\underline{\mathbf{x}}_{t,u})}$$
$$\prod_{k=1}^{K}\underbrace{p(x_{k,t,u}|\mathbf{d}_{k,t,u})}_{=f_{\mathrm{DEM}_{k,t,u}}(x_{k,t,u},\mathbf{d}_{k,t,u})}\prod_{q=1}^{Q}\underbrace{p(d_{k,q,t,u})}_{=f_{\mathrm{DEC}_{k,q,t,u}}(d_{k,q,t,u})}, \tag{4.70}$$

where the observation factor node OBS_k operates on antenna and user stacked data symbols $\underline{\mathbf{x}}_k$ of size $T\times 1$, for $k=1,\ldots,K$, and the transform nodes $\mathrm{DFTx}_{t,u}$ operates on time-stacked data symbols $\mathbf{x}_{t,u}$ of size $K\times 1$ and frequency-stacked data symbols $\underline{\mathbf{x}}_{t,u}$ of size $K\times 1$, for $t=\sum_{u'=1}^{u-1}T_{u'}+t'$, for $u=1,\ldots,U$, and $t'=1,\ldots,T_u$. The factor graph associated to this model is given Figure 4.25 and following Expectation Propagation (EP) message-passing procedure in Section 2.4.1 with these factor nodes, the distributions on variables of interest are updated.

Factor nodes $\mathrm{DFTx}_{t,u}$, $\mathrm{DEM}_{k,t,u}$ and $\mathrm{DEC}_{q,k,t,u}$ (regrouped over bits of each symbol, in the figure, for readability as $\mathrm{DEC}_{k,t,u}$) keep the same message passing parameters as in the Section 4.3.2. However, at the OBS_k factor node, the observations are subject to MAI and MUI and the pre-projection posterior on variable $\underline{\mathbf{x}}_k$ is given by

$$\tilde{q}_{\mathrm{OBS}}(\underline{\mathbf{x}}_k) = f_{\mathrm{OBS}}(\underline{\mathbf{x}}_k)\prod_{u=1}^{U}\prod_{t=1}^{T_u} m_{\mathrm{DFTx}_k\to\underline{x}_{k,t,u}}(\underline{x}_{k,t,u}),$$
$$\propto \exp\left(-\frac{\|\underline{\mathbf{y}}_k-\underline{\mathbf{H}}_k\underline{\mathbf{x}}_k\|^2}{\bar{\sigma}_w^2}-\sum_{u=1}^{U}\sum_{t=1}^{T_u}\frac{|\underline{x}_{k,t,u}-\underline{x}^e_{k,t,u}|^2}{v^e_{\underline{x},k,t,u}}\right)\propto\mathcal{CN}(\underline{\mathbf{x}};\boldsymbol{\mu}^o_{\underline{x},k},\boldsymbol{\Gamma}^o_{\underline{x},k}). \tag{4.71}$$

Then following projection onto fully-factorized FD symbol data set, posteriors estimates of

factor node OBS on node x_k is $q_{\text{OBS}}(\underline{x}_k) \propto \mathcal{CN}(\underline{x}; \mu^o_{\underline{x},k}, \gamma^o_{\underline{x},k})$, where

$$\gamma^o_{\underline{x},k,t,u} = \mathbf{e}^H_{t,u}\left(\frac{\underline{\mathbf{H}}^H_k \underline{\mathbf{H}}_k}{\bar{\sigma}^2_w} + {\mathbf{V}^{\mathbf{e}}_{\underline{x},k}}^{-1}\right)^{-1} \mathbf{e}_{t,u} = (1 - v^e_{\underline{x},k,t,u}\xi_{\underline{x},k,t,u})v^e_{\underline{x},k,t,u}, \tag{4.72}$$

$$\begin{aligned}\mu^o_{\underline{x},k,t,u} &= \mathbf{e}^H_{t,u}\left(\frac{\underline{\mathbf{H}}^H_k \underline{\mathbf{H}}_k}{\bar{\sigma}^2_w} + {\mathbf{V}^{\mathbf{e}}_{\underline{x},k}}^{-1}\right)^{-1}\left(\frac{\underline{\mathbf{H}}^H_k \underline{\mathbf{y}}_k}{\bar{\sigma}^2_w} + {\mathbf{V}^{\mathbf{e}}_{\underline{x},k}}^{-1}\underline{\mathbf{x}}^{\mathbf{e}}_k\right)\\ &= \underline{x}^e_{k,t,u} + v^e_{\underline{x},k,t,u}\mathbf{e}^H_{t,u}\underline{\mathbf{H}}^H_k \underline{\mathbf{\Sigma}}^{-1}_{\underline{\mathbf{y}},k}(\underline{\mathbf{y}}_k - \underline{\mathbf{H}}_k\underline{\mathbf{x}}^{\mathbf{e}}_k),\end{aligned} \tag{4.73}$$

with $\underline{\mathbf{\Sigma}}_{\underline{\mathbf{y}},k} \triangleq \bar{\sigma}^2_w \mathbf{I}_K + \underline{\mathbf{H}}_k \mathbf{V}^{\mathbf{e}}_{\underline{x},k}\underline{\mathbf{H}}^H_k$ and $\xi_{\underline{x},k,t,u} \triangleq \mathbf{e}^H_{t,u}\underline{\mathbf{H}}^H_k \underline{\mathbf{\Sigma}}^{-1}_{\underline{\mathbf{y}},k}\underline{\mathbf{H}}_k\mathbf{e}_{t,u}$. Then the extrinsic messages from the observation nodes can be computed as

$$m_{\text{OBS}_k \to \underline{x}_k}(\underline{x}_k) = \frac{q_{\text{OBS}_k}(\underline{x}_k)}{m_{\text{DFTx}_k \to \underline{x}_k}(\underline{x}_k)} \propto \frac{\mathcal{CN}(\underline{x}_k; \mu^o_{\underline{x},k}, \gamma^o_{\underline{x},k})}{\mathcal{CN}(\underline{x}_k; \underline{x}^e_k, v^e_{\underline{x},k})} \propto \mathcal{CN}(\underline{x}_k; \underline{x}^o_k, v^o_{\underline{x},k}), \tag{4.74}$$

$$v^o_{\underline{x},k,t,u} = \left(\frac{1}{\gamma^o_{\underline{x},k,t,u}} - \frac{1}{v^e_{\underline{x},k,t,u}}\right)^{-1} = 1/\xi_{\underline{x},k,t,u} - v^e_{\underline{x},k,t,u}, \tag{4.75}$$

$$\underline{x}^o_{k,t,u} = v^o_{\underline{x},k,t,u}\left(\frac{\mu^o_{\underline{x},k,t,u}}{\gamma^o_{\underline{x},k,t,u}} - \frac{\underline{x}^e_{k,t,u}}{v^e_{\underline{x},k,t,u}}\right) = \underline{x}^e_{k,t,u} + \xi^{-1}_{\underline{x},k,t,u}\mathbf{e}^H_{t,u}\underline{\mathbf{H}}^H_k \underline{\mathbf{\Sigma}}^{-1}_{\underline{\mathbf{y}},k}(\underline{\mathbf{y}}_k - \underline{\mathbf{H}}_k\underline{\mathbf{x}}^{\mathbf{e}}_k). \tag{4.76}$$

These computations over spatial dimensions are explicitly written to reveal the underlying MMSE receive beamformer, as in Figure 4.26. Let use denote the residual observations

$$\tilde{\underline{y}}_{k,r} \triangleq \underline{y}_{k,r} - \sum_{u=1}^{U}\sum_{t=1}^{T_u} \underline{h}_{k,r,t,u}\underline{x}^d_{k,t,u}, \tag{4.77}$$

obtained from IC, then we have the beamformer filter-bank

$$\underline{x}^e_{k,t,u} \triangleq \mathbf{e}^H_k \mathcal{F}_K \mathbf{x}^e_{t,u} = \underline{x}^d_{k,t,u} + \sum_{r=1}^{R} f^*_{k,r,t,u}\tilde{\underline{y}}_{k,r}, \tag{4.78}$$

where the MMSE filter-bank coefficients are

$$f_{k,r,t,u} = \xi^{-1}_{x,t,u}\sum_{r'=1}^{R} \lambda_{k,r,r'}\underline{h}_{k,r',t,u}, \tag{4.79}$$

$$\xi_{x,t,u} = \frac{1}{K}\sum_{k=1}^{K}\xi_{\underline{x},k,t,u} = \frac{1}{K}\sum_{k=1}^{K}\sum_{r=1}^{R}\underline{h}_{k,r,t,u}\sum_{r'=1}^{R}\lambda_{k,r,r'}\underline{h}_{k,r',t,u}, \tag{4.80}$$

with $\lambda_{k,r,r'} \triangleq [\underline{\mathbf{\Sigma}}^{-1}_{\underline{\mathbf{y}},k}]_{r,r'}$ the covariance of FD observations between r^{th} and r'^{th} antennas.

Finally, K-point IFFT is performed over the data blocks of each user and antenna to recover the TD data blocks and perform EP-based demodulation and decoding. DL-SEP framework assigns a scalar parameter for the reliability of each data block $\mathbf{x}_{t,u}$, with $v^d_{x,t,u} = v^e_{\underline{x},t,u}$ and $v^e_{x,t,u}$, being respectively the prior and extrinsic estimates' reliabilities for $\text{DFTx}_{t,u}$.

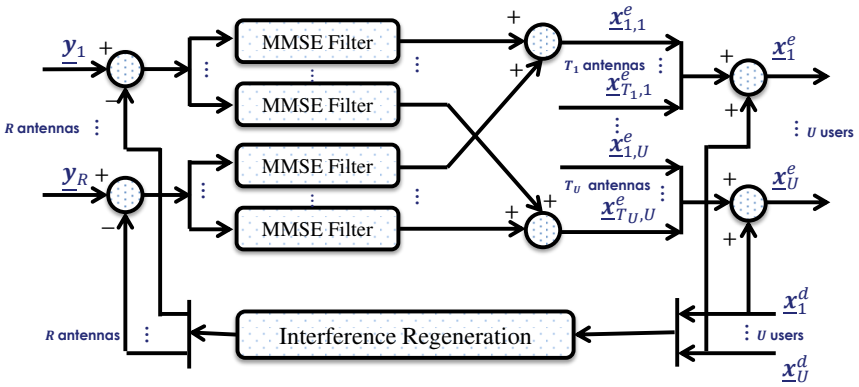


Figure 4.26: A constituent MMSE beamformer for Successive Interference Cancellation (SIC).

4.6.3.1 Computationally-Efficient Multi-user Covariance Matrix Inversion

In the expression of the novel factor node OBS_k's messages, there is an inversion of a $R \times R$ covariance matrix, for decorrelating observations corrupted by MAI and MUI. If typical Gauss-Jordan inversion is performed, filter-bank computation involves a computational complexity order of $\mathcal{O}(KR^3)$, and if LDL decomposition based inversion is used while the inversion complexity can be reduced to the order of $\mathcal{O}(KR^2)$, there would be K decompositions of $R \times R$ matrices to perform.

To alleviate this computational problem, the inversion of the covariance matrix can be simplified by exploiting the nature of DL-SEP based Interference Cancellation (IC). Indeed, as MUD decoding procedure characterizes each antenna and user with a single scalar parameter, the update operation on $\boldsymbol{\Sigma}_{\underline{\mathbf{y}},k}^{-1}$ can be investigated when $v_{\underline{x},t,u}^{e}$ is updated. This is inspired from a method originally used with Code Division Multiple Access (CDMA) receivers in Table 1 of [WP99], and considered for FD MUD in [Yua+08], we propose a recursive update strategy for handling turbo or self iterations.

Let $\boldsymbol{\Sigma}_{\underline{\mathbf{y}},k}^{(\mathrm{old})} = \bar{\sigma}_w^2 \mathbf{I}_K + \underline{\mathbf{H}}_k \mathbf{V}_{\underline{x},k}^{\mathbf{e}(\mathrm{old})} \underline{\mathbf{H}}_k^H$ denote the previous value of the covariance matrix, and let $\Delta v_{\underline{x},t,u}^{e} \triangleq v_{\underline{x},t,u}^{e(\mathrm{new})} - v_{\underline{x},t,u}^{e(\mathrm{old})}$. Then Woodbury's matrix inversion lemma states

$$\left(\boldsymbol{\Sigma}_{\underline{\mathbf{y}},k}^{(\mathrm{old})} + \Delta v_{\underline{x},t,u}^{e} \underline{\mathbf{h}}_{k,t,u} \underline{\mathbf{h}}_{k,t,u}^H\right)^{-1} = \boldsymbol{\Sigma}_{\underline{\mathbf{y}},k}^{(\mathrm{old})-1} - \delta v_{k,t,u}^{e} \boldsymbol{\phi}_{k,t,u} \boldsymbol{\phi}_{k,t,u}^H \tag{4.81}$$

where $\underline{\mathbf{h}}_{k,t,u}$ denotes the $(\sum_{u'=1}^{u-1} T_{u'} + t)^{\mathrm{th}}$ column of $\underline{\mathbf{H}}_k$, $\delta v_{k,t,u}^{e} = \left(\Delta v_{\underline{x},t,u}^{e}{}^{-1} + \underline{\mathbf{h}}_{k,t,u}^H \boldsymbol{\phi}_{k,t,u}\right)^{-1}$ and $\boldsymbol{\phi}_{k,t,u} \triangleq \boldsymbol{\Sigma}_{\underline{\mathbf{y}},k}^{(\mathrm{old})-1} \underline{\mathbf{h}}_{k,t,u}$. Hence, by denoting the inverse $\boldsymbol{\Lambda}_k \triangleq \boldsymbol{\Sigma}_{\underline{\mathbf{y}},k}^{-1}$, we have

$$\boldsymbol{\Lambda}_k^{(\mathrm{new})} = \boldsymbol{\Lambda}_k^{(\mathrm{old})} - \delta v_{k,t,u}^{e} \boldsymbol{\phi}_{k,t,u} \boldsymbol{\phi}_{k,t,u}^H, \tag{4.82}$$

and this algorithms could be initialized with $\boldsymbol{\Lambda}_k = \sigma_w^{-2} \mathbf{I}_R$, for all k, and then updated with $\Delta v_{\underline{x},t,u}^{e} = \sigma_x^2$, for all u and t, at the zeroth iteration of the receiver. Then, on upcoming turbo and inner iterations, $\boldsymbol{\Lambda}_k$ is updated either sequentially on u and t at each iteration, or

updated right away on all u and t, depending on whether a serial or a Parallel Interference Cancellation (PIC) scheduling is used. Existing schemes [WP99] did not store updated values of the covariance inverse between iterations, and had to recompute the whole procedure for all antennas/users at each iteration. Hence for Successive Interference Cancellation (SIC) or hybrid schedules, this update brings significant complexity savings, with a single matrix-vector product, two inner products, three scalar products and two scalar sums.

4.6.3.2 Scheduling

Although the messages exchanged between the factor nodes are defined, due to the multiple cycles present on the graph a scheduling is required. A robust MUD is proposed via a flexible structure, achieved by exploiting the presence of two loops: the first one refers to turbo-iterations between the $\text{DEC}_{q,k,t,u}$ and the $\text{DEM}_{k,t,u}$ factor nodes, while the second one refers to self-iterations between the $\text{DEM}_{k,t,u}$ and the OBS_k nodes.

Within each turbo iteration $\tau = 0, \ldots, \mathcal{T}$, each user u's data is decoded *successively*, with the information provided by variable nodes $\mathbf{d}_{k,t,u}$ to the factor node $\text{DEC}_{q,k,t,u}$ (i.e. $L_e(d_{q,k,t,u})$), for $q = 1, \ldots, Q$, $k = 1, \ldots, K$ and $t = 1, \ldots, T_u$. A natural ascending decoding order is used, for notation convenience, but in practice, a reliability-based ordering could be used [TV05].

In detail, while decoding a user u, the factor nodes OBS_k, $\text{DFTx}_{t,u}$ and $\text{DEM}_{k,t,u}$ are updated in a *parallel* scheduling, for joint ISI, MAI and MUI cancellation, and detection of data symbols $x_{k,t,u}$. This is enabled by factor node $\text{DEM}_{k,t,u}$'s extrinsic outputs $x^d_{k,t,u}$, computed with prior inputs from $\text{DEC}_{q,k,t,u}$ of users $u' < u$ already decoded in the current turbo iterations, and prior inputs of users $u' \geqslant u$, decoded in the previous turbo iteration. Messages between OBS_k, $\text{DFTx}_{t,u}$ factor nodes and between $\text{DFTx}_{t,u}$ and $\text{DEM}_{k,t,u}$ factor nodes are self-iterated, for all k, t, u, for improving u^{th} user's detection by processing all users, with $s = 0, \ldots, \mathcal{S}$ denoting self-iterations. At initial self iteration, $s = 0$, OBS_k, $\text{DFTx}_{t,u}$ and $\text{DEM}_{k,t,u}$ factor nodes' messages are reset with, in particular $x^{e(s=0)}_{k,t,u} = 0$ and $v^{e(s=0)}_{k,t,u} = +\infty$. When $s = \mathcal{S}$, current self iterations messages on users other than u are discarded, and their variable nodes are reinitialized to their previous values from factor node $\text{DEC}_{q,k,t,u}$, whereas the user u's data are given to $\text{DEC}_{q,k,t,u}$ by $\text{DEM}_{k,t,u}$.

Selection of $\mathcal{S}$ and $\mathcal{T}$ allow for a control over the processing latency and the required computational complexity. As self iterations of a FD MUD consist of detection (with single-tap filters) and demapping, they are computationally less complex than decoding, and this makes turbo iterations less desirable.

Combination of the simplified messages from the previous subsection with this schedule, yields the proposed low-complexity MUD with hybrid IC, displayed on Figure 4.27, which extends FD SILE-EPIC to MU-MIMO spatial multiplexing.

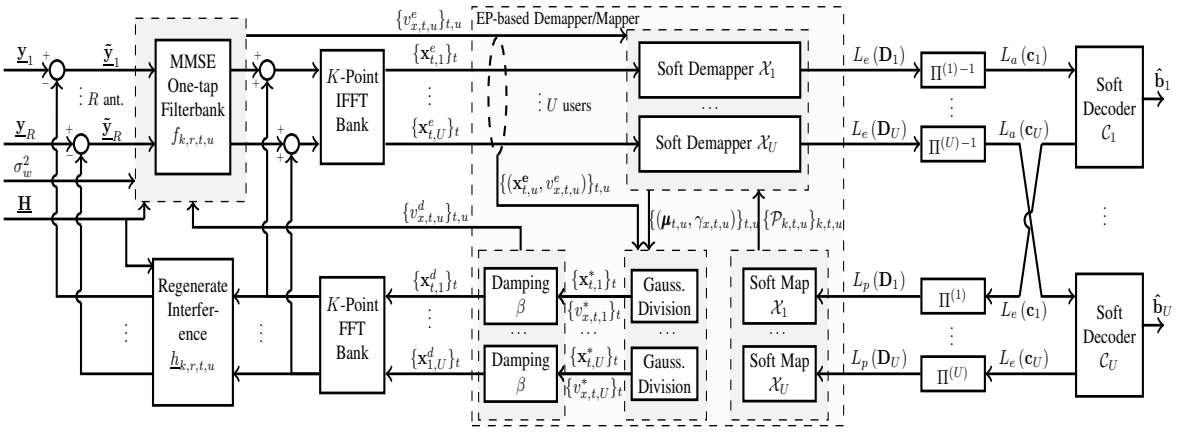


Figure 4.27: Proposed iterative MU-MIMO SC-FDMA receiver with bin-wise filterbank and interference cancellation based on EP.

4.6.3.3 Performance Analysis

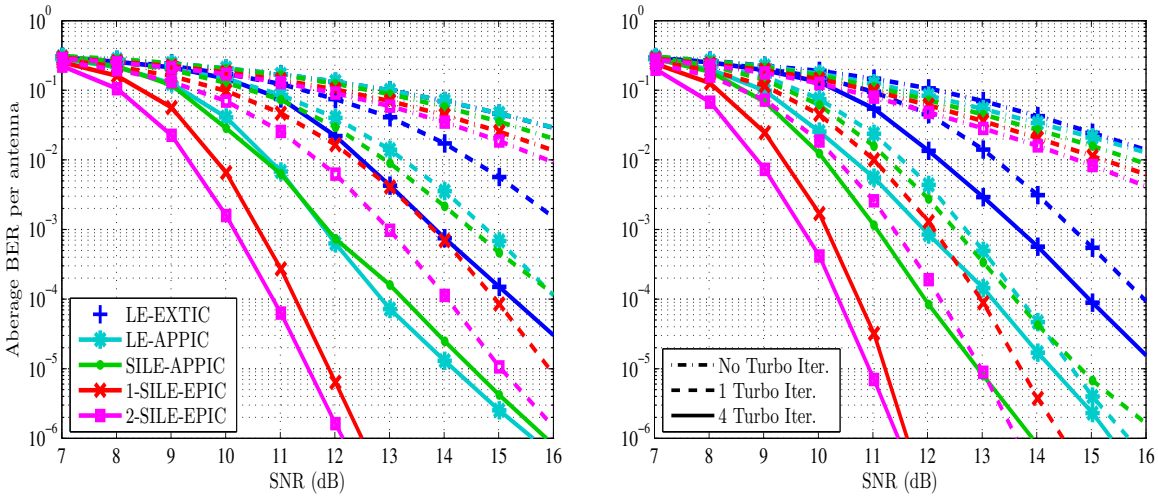
In this subsection, the proposed MUD extension of FD SILE-EPIC is compared in terms of finite-length BER performance with similar bin-wise FD receivers from the literature, which all the same order of complexity for filter computation (per self and turbo iteration).

Comparison with Prior Works and the Impact of Scheduling As references, LE-EXTIC, and its APP-based variant LE-APPIC [Wit+02; VBC06] are considered. Among recent structures, we consider the self-iterated LE with APP feedback SILE-APPIC [Tao15; BT05] (denoted "Subopt SD-FDE-SIC-III" in [Tao15]). Note that, *for fair comparison, SILE-APPIC is extended* to multi-antenna users, and all receivers use our proposed hybrid scheduling with random interleavers.

In this paragraph we consider SC-FDE signalling with $K = N$, with the scenario from [Tao15], with two single-antenna users, where a 2-antenna MUD, and a non-systematic convolutional encoder $[17, 13]_8$ are used, in a generalized AWGN Proakis B channel. Up to two self-iterations are considered with the exponential smoothing damping parameter $\beta = \max(0.3, 0.5 \times (0.8)^{s+\tau})$, and average BER performance per user is plotted in Figure 4.28.

In Figure 4.28-(a), interference is mitigated with a PIC schedule, as in [Tao15], i.e. with simultaneous detection over antennas in each SI, and simultaneous decoding of all antennas in each turbo-iteration. Our proposal displays remarkable gains over APP-based prior work, with over 2 dB and 2.5 dB gains at 4 TIs, at BER = 10^{-5}, with respectively 1 and 2 SIs.

In Fig. 4.28-(b), the MAI is mitigated with the proposed hybrid PIC/SIC schedule, which is expected to converge faster. Our proposal outperforms concurrent structures for all TI, with over 1.5 dB margin for BER = 10^{-5}. Moreover, SILE-APPIC with either SIC or PIC, at 4 turbo-iterations, is outperformed by either 1-SILE-EPIC with $\mathcal{T} = 1$ with PIC or SIC.



(a) Parallel Interference Cancellation schedule

(b) Successive Interference Cancellation schedule

Figure 4.28: BER in AWGN Proakis B with 16-QAM, $U = 2$, $R = 2$ and $T_1 = T_2 = 1$ (0 TIs: dash-dotted, 1 TI: dashed, 4 TIs: plain).

Table 4.3: Computational costs for the QPSK scenario in Figure 4.29

Receiver	Costs per TI	SNR (dB)		Total Costs	
		1 TI	3 TIs	1 TI	3 TIs
LE-EXTIC	4.739e^5	10.8	6.7	2.1e^6	4.0e^6
LE-APPIC	4.746e^5	9.1	6.1	2.1e^6	4.0e^6
SILE-APPIC	9.070e^5	7.8	5.6	2.9e^6	5.8e^6
0-SILE-EPIC	4.452e^5	10.8	6.7	2.0e^6	3.9e^6
1-SILE-EPIC	8.559e^5	7.2	5.1	2.8e^6	5.6e^6
2-SILE-EPIC	12.665e^5	6.8	4.8	3.7e^6	7.2e^6

Asymptotically ($\mathcal{T} = 4$), SIC improves our proposal's BER around 0.5 dB over PIC, but SIC with 1 turbo-iteration is shown to significantly outperform alternatives, which provides an attractive compromise of fewer decoder iterations, but increased detector iterations, to provide interesting complexity-performance options, especially when using powerful decoders.

Spatial Multiplexing for MU-MIMO with SC-FDMA Finally, we evaluate the throughput of two MU-MIMO spatial multiplexing schemes. Two users, with two spatially uncorrelated antennas each, transmit over the EQU4 channel, which has four equal power taps following Rayleigh statistics. Localized-FDMA is used, with $K = 256$ and $N = 512$.

In the first scenario, an under-determined system with a MUD with $R = 2$ is considered. Thanks to IC with STBICM, such situations become resolvable [Ton00]. User throughput, using QPSK with a $[1, 5/7]_8$ encoder, is plotted on the left side of the Figure 4.29. Our

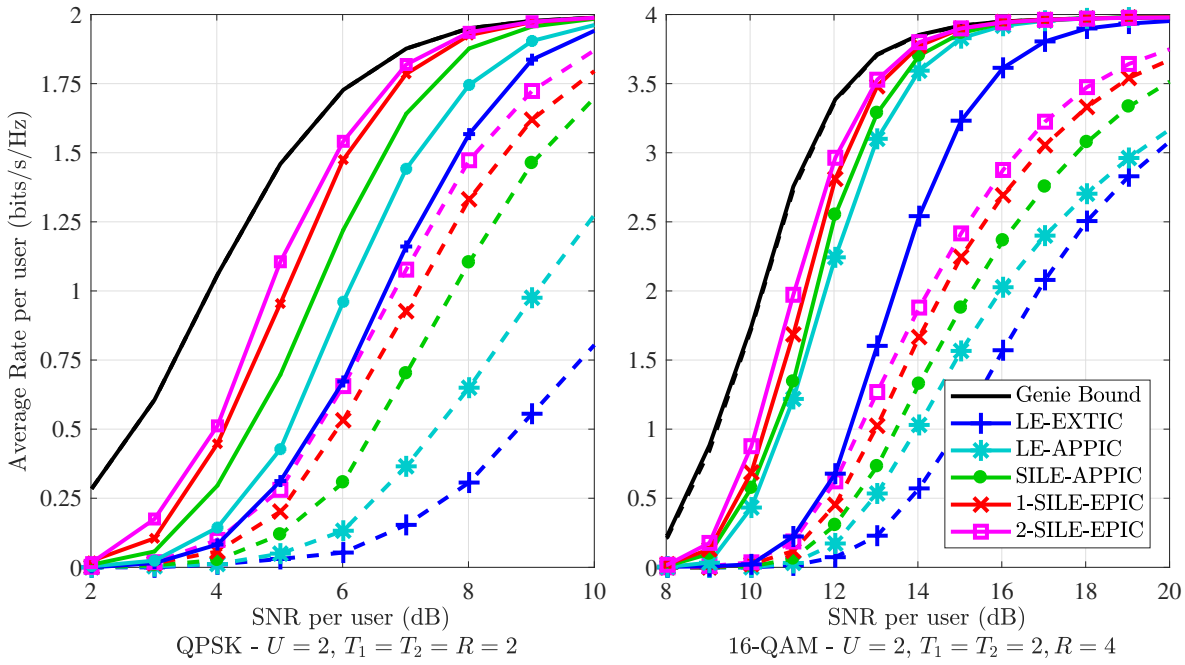


Figure 4.29: Throughput in Rayleigh EQU4, $U = 2$ (for QPSK, 1 turbo-iteration: dashed, 3 turbo-iterations: plain; for 16-QAM 0 turbo-iterations: dashed, 1 turbo-iteration: plain).

proposal, with $\beta = 0$, requires 0.5 - 1 dB less energy than FD-SILE-APPIC for achieving 1 bit/s/Hz.

At the right side of the figure, a more conventional situation with $R = 4$, and 16-QAM constellation, using the same channel encoder is considered. In this case, the interest of our receiver is more prevalent for $\mathcal{T} = 0$ with over 1 dB gain at 2 bits/s/Hz. Note that, in both schemes, MUDs without SIs have a significant disadvantage.

In Table 4.3, the computational costs are given in the number of required FLOPs. The required SNR for reaching 1 bits/s/Hz, and the total detection and decoding FLOPs at the reported turbo-iteration are given in the latter columns. It is assumed the proposed MUD uses the matrix-inversion scheme of Section 4.6.3.1, whereas prior works use the one of [Yua+08]. First, 0-SILE-EPIC, equivalent to LE-EXTIC, shows 6% complexity savings thanks to our update scheme alone. Moreover, it is seen that self-iterations do not cost much (relative to decoding), and 2-SILE-EPIC with 1 turbo-iteration, is 10% less expensive but offers similar performance as LE-EXTIC with 3 turbo-iterations.

4.6.4 Discussion

A novel multi-user detector is proposed that exploits the expectation propagation framework, to jointly mitigate MAI and ISI. By writing the factor-graph model of this MIMO spatial multiplexing system and applying the DL-SEP framework, a low-complex MUD is obtained. Indeed, the key element of our proposal, is enabled again by the EP-based mes-

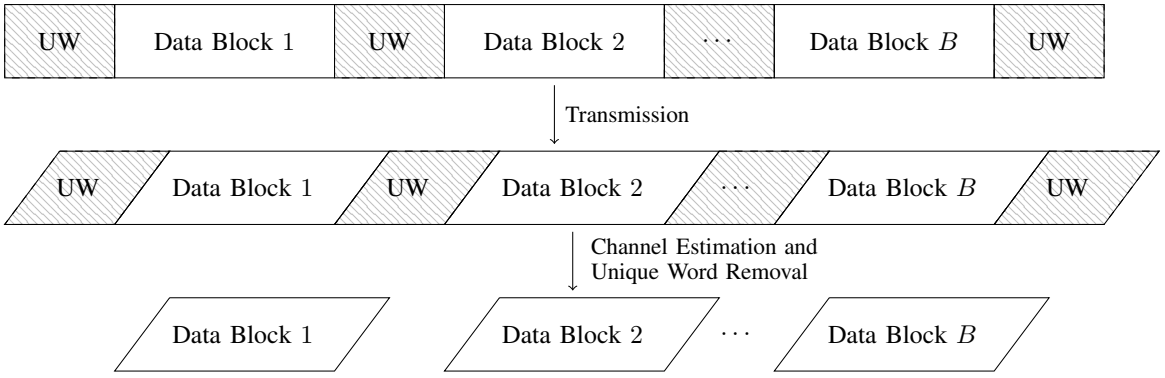


Figure 4.30: Equivalence of UW SC-FDE to SC-FDE with Zero Padding (ZP).

sage passing with statistically whitened Gaussian distributions, which yields simple bin-wise equalization filterbanks. Moreover, an efficient matrix inversion scheme from the literature is optimized for iterative MUDs, in order to further reduce the amount of computations for filter design.

Through finite-length and asymptotic analysis, it is seen that the proposed structure outperforms other bin-wise FD MUD, and attests the importance of a feedback computed by using EP. Improvements are seen to be remarkable in especially highly selective channels. It is shown that exploiting a self-iteration loop between demapping device and the MUD can improve performance at lesser costs than decoding, and provides novel options for MUD design. Moreover EP-based SIs are seen to be more efficient than APP-based alternatives.

4.7 Impact of Channel and SNR Estimation

In this section, the performance of the proposed FD SILE-EPIC with imperfect channel estimation is evaluated. These aspects will be considered through the use of Unique Word (UW)-padded SC-FDE transmissions, which enable the use of practical CIR and Signal to Interference and Noise Ratio (SINR) estimation algorithms at the receiver, among others, as shown in Figure 4.30. In particular, a FD CFR and SINR estimation method is presented and the concepts of mismatched and robust receivers are discussed for the case of the proposed doubly-iterative turbo equalizer. The impact of CIR and SINR estimation on self-iterations is also discussed.

4.7.1 On Channel Estimation for SC-FDE with Unique Word

UW sequences are often chosen to have good deterministic auto-correlation properties in order to reduce the complexity of channel estimation. Hence we assume a K_P-symbol training sequence $\mathbf{x_P}$ is available, such that its circular auto-correlation with its k-cyclic shifted sequence is $K_P\delta[k]$, where $\delta[\cdot]$ is the Kronecker's delta function. This can be satisfied

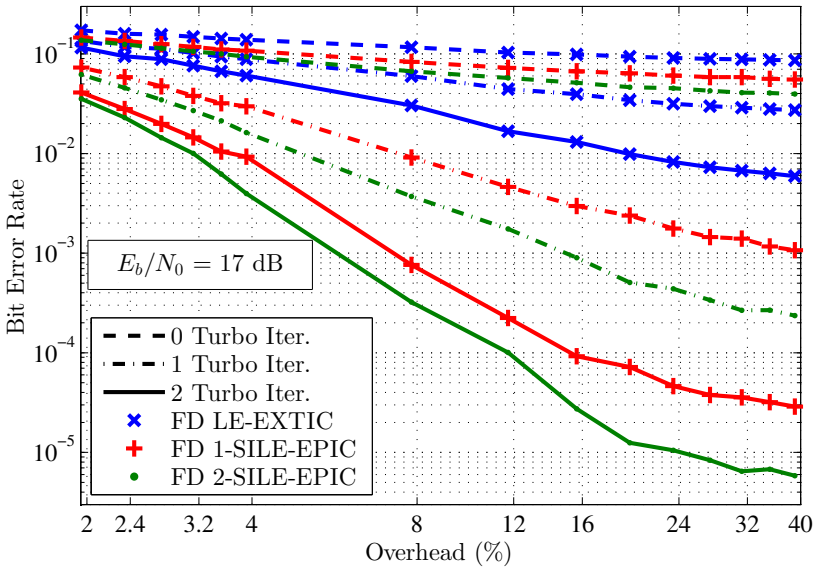


Figure 4.31: Channel estimation overhead.

with Zadoff-Chu sequences for instance.

Thus, the baseband FD observations for pilot symbols are $\mathbf{y_P} = \mathbf{X_P h_P} + \mathbf{w_P}$, where $\mathbf{X_P} = \mathbf{Circ}(\mathbf{x_P}^T)$, is the circulant matrix whose first column is $\mathbf{x_P}^T$, $\mathbf{w_P}$, and $\mathbf{h_P} = [\mathbf{h}; \mathbf{0}_{K_P-L,1}]$ and with the K_P-point DFT, frequency domain observations are

$$\underline{\mathbf{y}}_\mathbf{P} = \underline{\mathbf{X}}_\mathbf{P}\underline{\mathbf{h}}_\mathbf{P} + \underline{\mathbf{w}}_\mathbf{P}, \tag{4.83}$$

with $\underline{\mathbf{X}}_\mathbf{P} = \mathcal{F}_{K_P}\mathbf{X_P}\mathcal{F}_{K_P}^H$, $\underline{\mathbf{h}}_\mathbf{P} = \mathcal{F}_{K_P}\mathbf{h_P}$, and $\underline{\mathbf{w}}_\mathbf{P} = \mathcal{F}_{K_P}\mathbf{w_P}$, which is assumed to be AWGN. The auto-correlation properties of the training sequence satisfy $\underline{\mathbf{X}}_\mathbf{P}\underline{\mathbf{X}}_\mathbf{P}^H = \underline{\mathbf{X}}_\mathbf{P}^H\underline{\mathbf{X}}_\mathbf{P} = K_P\mathbf{I}_{K_P}$.

In this system, the CIR is interpolated respectively to K_P and to K point CFRs for pilot and data observations. To model this more succinctly, we denote the truncated DFT matrix $\mathcal{F}'_K = \mathcal{F}_K[\mathbf{I}_L; \mathbf{0}_{K-L,L}]$, of size $K \times L$ such that

$$\underline{\mathbf{h}}_\mathbf{P} = \mathcal{F}'_{K_P}\mathbf{h}, \quad \underline{\tilde{\mathbf{h}}} = \mathcal{F}'_K\mathbf{h}, \tag{4.84}$$

where $\mathbf{h}$ is TD CIR. Note that while ${\mathcal{F}'_K}^H\mathcal{F}'_K = \mathbf{I}_L$, $\mathcal{F}'_K{\mathcal{F}'_K}^H$ is a non-diagonal matrix. Then pilot and data observations are rewritten as

$$\underline{\mathbf{y}}_\mathbf{P} = \mathbf{A_P}\underline{\tilde{\mathbf{h}}} + \underline{\mathbf{w}}_\mathbf{P}, \quad \underline{\mathbf{y}} = \sqrt{K}\underline{\tilde{\mathbf{h}}} \odot \underline{\mathbf{x}} + \underline{\mathbf{w}}, \tag{4.85}$$

where $\mathbf{A_P} \triangleq \underline{\mathbf{X}}_\mathbf{P}\mathcal{F}'_{K_P}{\mathcal{F}'_K}^H$. In this section, we consider the use of a least-squares criterion on channel estimation, by assuming a complete lack of prior knowledge on $\mathbf{h}$. Hence the

frequency-domain channel estimate $\underline{\tilde{\mathbf{h}}}^{\mathrm{LS}}$ is given by

$$\underline{\tilde{\mathbf{h}}}^{\mathrm{LS}} \triangleq (\mathbf{A}_{\mathbf{P}}^H \mathbf{A}_{\mathbf{P}})^{-1} \mathbf{A}_{\mathbf{P}}^H \underline{\mathbf{y}}_{\mathbf{P}}, \tag{4.86}$$

$$v_{\underline{h}}^{\mathrm{LS}} \triangleq (\mathbf{A}_{\mathbf{P}}^H \mathbf{A}_{\mathbf{P}})^{-1} \bar{\sigma}_w^2, \tag{4.87}$$

with $v_{\underline{h}}^{\mathrm{LS}}$ being the estimation error variance caused by the channel noise. By neglecting correlations caused by DFT interpolation, i.e. with $(\mathbf{A}_{\mathbf{P}}^H \mathbf{A}_{\mathbf{P}})^{-1} \approx K_P^{-1} \mathbf{I}_K$.

4.7.2 Mistmatched FD SILE-EPIC

The first case for receiver design would be to consider $\mathbf{h}$ as a parameter whose true value is the least-square estimate used above. In this case, one could consider the PDF of $\mathbf{h}$ as $p(\underline{\mathbf{h}}) = \delta(\underline{\mathbf{h}} - \underline{\tilde{\mathbf{h}}}^{\mathrm{LS}})$ which enforces the value of the estimate for receiver design, which ends-up becoming *mismatched* to the true channel.

In Figure 4.31, the BER performance of a mismatched receiver is considered for 8-PSK transmissions with $K = 256$, for $\mathcal{E}_b/N_0 = 17$ dB, in the Proakis C channel, as a function of the pilot transmission overhead. In this case, channel estimation errors are artificially generated with the Gaussian process $\mathcal{CN}(\mathbf{0}_{L\times 1}, v_{\underline{h}}^{\mathrm{LS}} \mathbf{I}_K)$ being added to the true channel at each Monte-Carlo step. More specifically, the overhead is defined as K_P/K, and the proposed FDE proposed in Section 4.3.2 is evaluated for different values of turbo and self iterations. It is seen that, for a target BER of 10^{-2}, a significant reduction of overhead is achieved thanks to EP-based self-iterations; while baseline FDE with 2 turbo-iterations requires around 19% overhead for channel estimation, using our proposal, one turbo and one self iteration requires only 8% overhead, and one self and 2 turbo iterations requires 4%. Thus PHY data frames with shorter number of pilot symbols could be designed to increase spectral efficiency, thanks to the use of double loop scalar EP based framework.

4.7.3 Robust FD SILE-EPIC

An alternative approach tries to account for the channel estimate uncertainty in filter design, by modelling the PDF of $\underline{\mathbf{h}}$ as $p(\underline{\mathbf{h}}) \propto \mathcal{CN}(\underline{\tilde{\mathbf{h}}}^{\mathrm{LS}}, v_{\underline{h}}^{\mathrm{LS}} \mathbf{I}_K)$, which causes changes in the filter-computations by replacing the noise variance by $\bar{\sigma}_w^2 + \sigma_x^2 v_{\underline{h}}^{\mathrm{LS}}$. This technique is known as the *robust* implementation of the equalizer.

In Figure 4.32, the performance of mismatched and robust receivers are compared for BPSK transmissions with $K = 2048$. In this case CAZAC sequences are used to carry out channel estimation for $K_P = 16$ and $K_P = 50$ symbols (including 5 symbol CP, not used in estimation), over the Proakis C channel. It is seen that robust structures have little to no effect when $K_P = 50$ as the variance of the channel estimates is not very high.

While self-iterations improve the PER performance, especially, when turbo-iterations are unavailable, when comparing the use of perfect CSI performance (genie estimate) with the

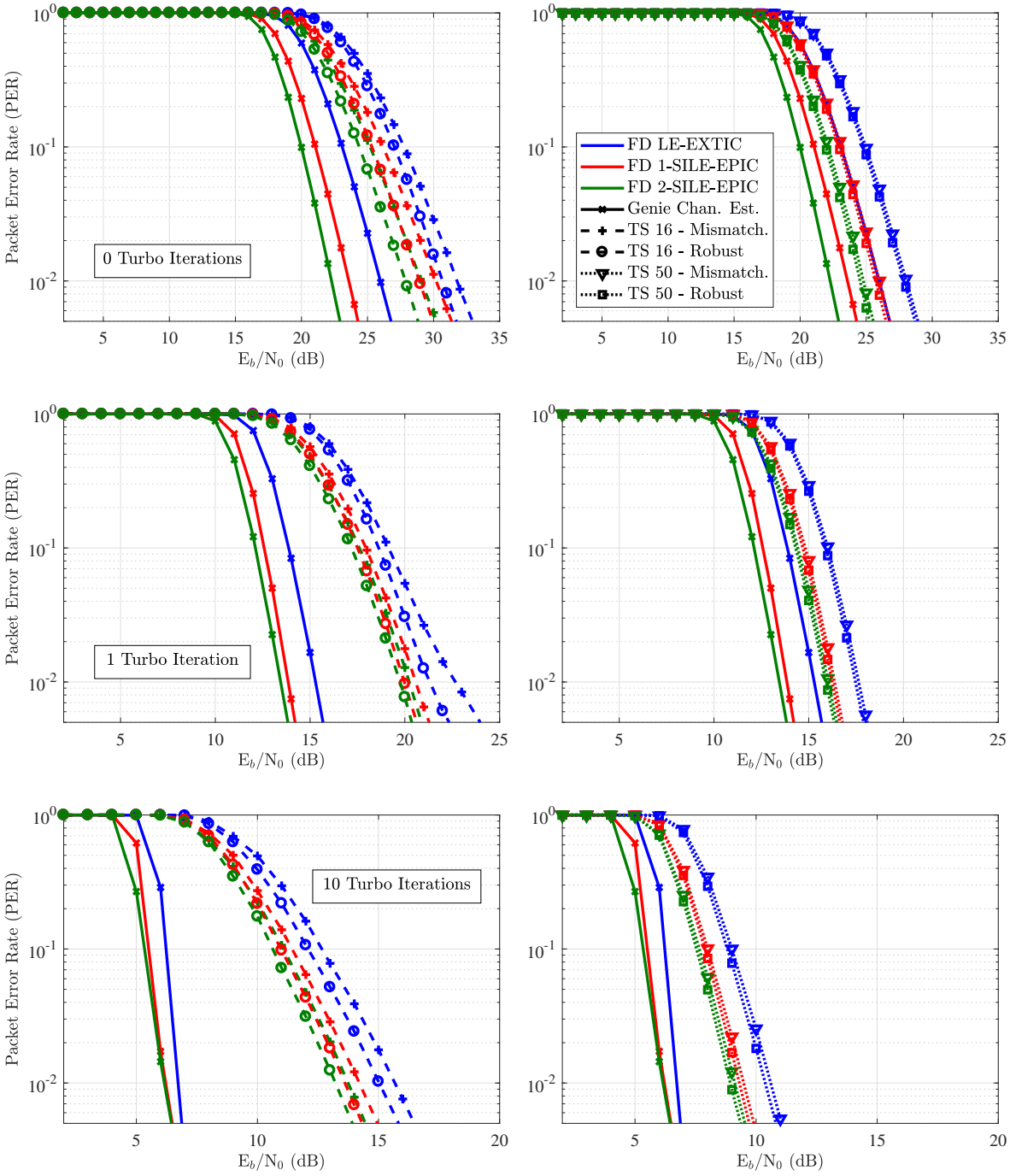


Figure 4.32: BPSK channel estimation with $K_P = 16$ pilots at left, $K_P = 50$ pilots at right.

EP-based receiver, there is still a considerable gap that needs to be to be reduced, which is left for future work on channel estimation.

4.8 Conclusion

This chapter starts by outlining the computational benefits of FD receivers and then moves on to investigate a message passing technique; based on EP [Min01]. Indeed, having discussed in Chapter 2 the advantages of using scalar EP message passing for doubly-iterative turbo equalization with improved performance and predictability, here we theoretically investigate on their benefits for single-carrier FDE and its extensions.

It is shown that, thanks to the better stability of scalar EP messages, with regards to extrinsic parameter computation, receivers derived with this framework can nearly perform as efficiently as their TD block filter-banks [SMF+17a]. Similarly to the previous chapter, the use of EP-based self iterations bring significant performance improvements at the high data rate applications. Moreover, based on the previously established behaviour of AMP based algorithms and on emerging techniques in deep learning, it is shown that this FD receiver can be further optimized by casting into a deep networks and learning damping parameters through an original turbo-oriented cost function.

Furthermore, extensions of this framework for more complex SC waveforms is discussed, such as SC-FDMA or fractionally-spaced equalization, handling time-varying channels with overlap FDE and a multi-user MIMO non-orthogonal access use-case, with spatial multiplexing. And finally, the impact of channel estimation is investigated on the behaviour of this category of receivers.

Chapter 5

System-Level Simulations through PHY Abstraction for MANETs

Contents

5.1 Introduction

In this final chapter of this thesis, we will discuss methods for the system-level evaluation of advanced receivers' impact on a cooperative MANET. As previously addressed, in the introductory context chapter, although there are various cross-layer cooperative protocols for MANETs, their fair comparison would require precisely evaluating the impact of protocol overhead and receiver limitations on a realistic channel model. However, carrying this out with a complete PHY simulator would cause unpalatable simulation costs, in time or computational resources, which creates the need for lightweight but relatively accurate PHY layer and channel abstraction techniques.

To address this problem, this chapter brings various contributions, summarized as follows:

- a cross-layer simulation framework is proposed, that accounts for the unconventional characteristics of distributed cooperative transmissions in MANETs by modelling the positive and negative impacts of non-orthogonal access on the physical channel,
- existing techniques for the semi-analytical prediction of turbo equalizers [VB+10] are briefly reviewed and applied to the doubly-iterative Frequency Domain (FD) EP based receiver of the Chapter 4 receiver, by extending state-evolution techniques used in approximate inference (see Chapter 2),
- an original idea is developed for improving the finite-length prediction of iterative BICM receivers, by analysing the impact of soft feedback's dispersion on this receiver.

The contributions on the semi-analytical prediction of FD SILE-EPIC constitute the major innovative contributions of this chapter, and the simulation framework provides a means for illustrating its value for communications system design.

More explicitly, the value of prediction techniques is illustrated for distributed cooperation technique that is very attractive for MANETs: cooperative broadcast. In this context, the advantages of this cross-layer protocol, and the ensuing challenges are exposed by analysing its impact on the physical channel, through a novel stochastic model on the frequency-selective channel properties of cooperative broadcasting. Finally, the importance of advanced PHY receivers for such protocols is illustrated within some tactical MANET context, with some off-the-shelf error correction codes.

5.2 A Cross-layer Simulator for Cooperative MANETs

The level of abstraction used by conventional system-level simulators is often limited to assimilating a link between nodes u and v to a Packet Error Rate (PER) metric $\mathrm{PER}_{u,v}$, that is estimated by taking in consideration the instantaneous received SINR, denoted $\mathrm{SINR}_{u,v}$. In particular, when multiple nodes simultaneously transmit in the same frequency resources, typically, either signals from all interfering nodes are assimilated to noise, or, if the receiver posses a MUD, SIC decoding model is used.

However, these cases do not enable system-level evaluation of cooperative protocols at the PHY layer, where relaying with coherent signals can take place, if relay links get activated, depending on the ongoing network configuration and channel state. To this end, system-level simulators need to be re-thinked, by not only keeping track of codeword-based performance metrics, in order to be able to detect and model coherent signal combining at the PHY. Moreover, simultaneous transmissions in real-world channels only results in coherent combining, if the signal is sufficiently narrowband. Otherwise, signals relayed over frequency-selective channels, and with propagation delays, create intra-signal interference such as ISI, IBI or ICI, which takes place on the "cooperative link" [Ram05]. In this case, PHY receiver may be able to recover some amount of interference, by mitigating the composite Channel Impulse Response (CIR) of the cooperative link, which is another behavior that needs to be modelled.

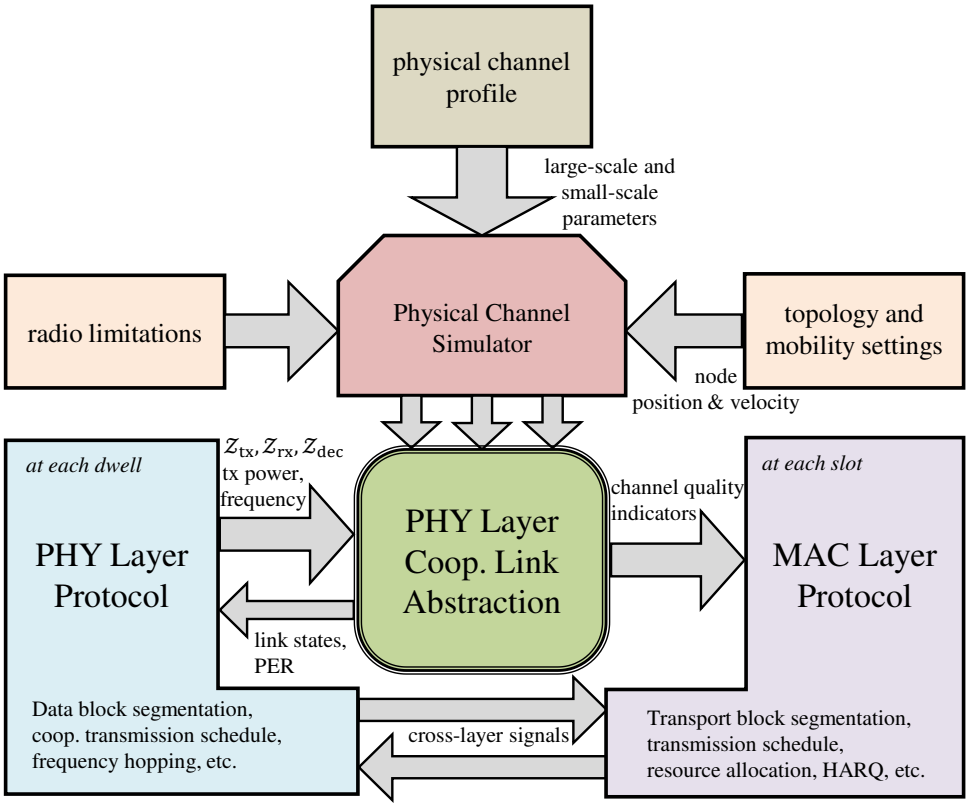


Figure 5.1: Simulation framework for cross-layer cooperative protocols.

Recalling the vocabulary used for describing the network model in Figure 1.3, we consider the simulation framework illustrated in Figure 5.1. The metrics exchanged between PHY and MAC protocols, and the channel simulator, enable tracking the states of each payload, transport block and data block, and keep the underlying links between them (segmentation, network or channel coding). Moreover, the PHY layer cooperative link abstraction module

- identifies transmitted data blocks that combine coherently, and those that remain as a noise-like interference, at each dwell, for each receiving radio,
- computes the composite or "artificial" Channel Impulse Response (CIR), corresponding to the data that coherently combine, seen at each receiving radio,
- tracks baseband channel state for each received data block, at least for the duration of a MAC slot, in order to be able to perform code combining or SIC, in upcoming dwells.

5.2.1 Protocol and Network Model

Reconsidering the network hypotheses of Section 1.2.3, let us recall that a time-frequency resource grid is available (see Figure 1.3), where

- T_{frame} is the duration of a MAC frame, during which access operations required by the protocol are periodically carries out,
- T_{slot} is the duration of a MAC slot, which is the minimum duration of an access operation, such as transmitting, receiving or remaining idle, such that $T_{\text{frame}} = N_{\text{slot}}^{\text{frame}} T_{\text{slot}}$,
- T_{dwell} is the lowest granularity of time resource, from the protocol's point of view, during which PHY layer carries out its elementary operations, such that $T_{\text{slot}} = N_{\text{dwell}}^{\text{slot}} T_{\text{dwell}}$.

Hence the simulator operates as a three-loop algorithm, with the loop on frames being the main loop, followed by a loop on slots during which protocol access operations are sequentially scheduled, and finally a loop on dwells, for simulating the PHY layer protocol, along side a realistic channel and mobility model.

Let us recall that, at any slot ℓ_s, at the MAC layer there is a set of transmitting node set $\mathcal{Z}_{\text{tx}}[\ell_s]$ (sources), receiving node set $\mathcal{Z}_{\text{rx}}[\ell_s]$ (destinations), such that for all $u \in \mathcal{Z}_{\text{rx}}[\ell_s]$, there exists the set of radios whose messages are to be decoded by u, $\mathcal{Z}_{u,\text{dec}}[\ell_s]$. In this regard, MAC assigns users to these sets depending on the protocol, and generates transport blocks at the transmitting nodes from available payloads, and assigns transmit powers. For simplicity, in this thesis we do not address resource allocation problems, and we simply assume that frequency resources correspond to different Frequency Hopping (FH) bands, which change in a pseudo-random manner from one slot to another.

At the PHY level, the protocol operates similarly, such that at each dwell ℓ_d of each slot ℓ_s, PHY also assigns radios to transmit and receive sets, depending on the operating mode of the cooperative protocol. PHY also assigns the channel, depending on the FH state, for each data block $\mathbf{x}_u^{(b)}$ to be transmitted.

5.2.2 Channel Model and Cooperative Link Abstraction

Considering the protocol model above, N_{FH} channels are simulated, for each user-destination link, under the assumptions that frequency hopping bands are decorrelated. For each frequency band, and each dwell, a statistical quasi-static channel model is simulated, with small-scale aspects being generated following the guidelines of Section 1.3, and Appendix A, by keeping track of each node's velocity, for temporal correlations. Large scale features are simulated by keeping track of nodes positions, in order to compute path-loss, and eventually use the model of [Fis+13a] for incorporating shadowing.

In the end, for each pair of nodes u and v, a Channel Impulse Response (CIR) $\mathbf{h}_{u,v}$ is obtained at each dwell, and if a same data block is transmitted by multiple nodes, then their CIRs are combined by accounting for the propagation delays between cooperating nodes. For simplicity this operation is performed over the samples of the upsampled channel model, which is later filtered by shaping filters to obtain the equivalent "useful" baseband channel model. Similarly, for each interfering data block, interference CIRs are computed by the CIR of different transmitting nodes.

Thanks to this identification of useful and interfering links, cooperative transmissions are modelled with better fidelity, by accounting for the composite useful CIR obtained by simultaneously transmitting nodes, and similarly obtained interfering CIR, caused by the transmission of data blocks sent for other users. Instead of tracking SINR, tracking the couple of useful and interfering CIR enable using semi-analytical PHY abstraction models for prediction.

5.3 FD SILE-EPIC Abstraction for System-Level Simulations

PHY receiver and physical link abstraction play a major role in wireless communication system design, whether it be for system-evaluation [Bru+05; RV05], or for link-adaptation in radio-resource management [Jen+10]. In system-level simulators there can be a large number of simultaneously active links, which is further increased when a cooperative protocol is involved. Hence to avoid evaluation of each individual link with a complete PHY layer simulator, it is preferable to predict the deployed radio behavior given an instantaneous topology, a transmission protocol and physical channel states. In this perspective, the physical-layer link abstraction is of crucial importance, and it has led to a significant amount of research from the academia [Rou+01], the industry [Bru+05] and the standardization institutions [3gpc; Ieea; Ieeb]

In this section, we seek an abstraction model for the low-complexity double-loop scalar EP based Frequency Domain (FD) turbo receiver from the Chapter 4. To this end, we first briefly overview the broad developments on links abstraction in order to identify key elements for successful models. Following this step, we conclude that an asymptotic semi-analytic evolution model of FD SILE-EPIC is needed to be used along with the joint demapping and decoding model of [VB+10]. An asymptotic model is obtained by revisiting the EXIT model (see Section 2.3.6) of this receiver, along with considerations from the state-evolution techniques of AMP-like algorithms (see Section 2.4.4), and finally, we discuss the accuracy of the resulting finite-length abstraction scheme.

5.3.1 On PHY Link Abstraction for Classical Receivers

Link abstraction strategies for classical non iterative receivers consists in determining the post-equalization Signal to Interference and Noise Ratio (SINR) from available channel parameters and to assimilate the equalizer output to an AWGN channel. Moreover, the Packet Error Rate (PER) versus SNR of the channel decoder in an AWGN is evaluated off-line to generate a Lookup Table (LUT), which can be then combined with the post-equalization SINR estimation technique to predict the overall receiver behavior.

However, in many wireless systems, a codeword is transmitted over different dwell periods, frequency bands or antennas, for increasing the link diversity [Bru+05; Ieea]. This would be the case in this thesis with the considered TDMA MANET where $B > 1$ data-blocks are

transmitted per codeword. In this case, it is non-trivial to propose a simple semi-analytical model, unlike the $B = 1$ case described in the previous paragraph. Hence, widespread prediction approaches rely on an approximation based on SINR "compression", which consists in combining signal quality indicators of data blocks from different diversity sources (time, space, frequency etc.) into a scalar codeword quality indicator. This is called the Effective Signal-to-noise Metric (ESM) compression [Bru+05; Jen+10; Van+15], and it is then used to identify the corresponding PER of the channel decoder, thanks to off-line generated LUTs.

Exponential ESM (EESM) [Lam+03] and Mutual Information ESM (MIESM) [Bru+05] figure amount the most common approaches. While EESM is more adapted to optimization problems in resource management (as it has analytical closed-form expressions), it requires a significant amount of calibration [CVS08]. MIESM is more accurate and requires less calibration heuristics, thus it is better-suited for prediction within a system-level simulator.

5.3.2 On PHY Link Abstraction for Iterative Receivers

The predictability of the iterative detection dynamics is of significant importance both for facilitating algorithm design of advanced receivers, and for physical layer link abstraction. For instance, as previously evoked in Section 2.3.6, in channel coding, density evolution was proposed for tracking the dynamics of PDFs of exchanged bit LLRs between SISO modules [RSU01; BC02]. We have seen that EXIT analysis [TB00] simplifies this problem to an asymptotic single-parameter tracking problem, and it enabled tracking the achievable rates of SISO detectors and it also enables the asymptotic prediction of decoders BER or PER performance across turbo-iterations.

When considering the use of EXIT for predicting the behavior of the considered MANET model, where a codeword is subject to P independent channel realizations due to fading and frequency-hopping [KH00], the receiver's EXIT function becomes an at least B-dimensional function. Analyzing receiver behavior under such circumstances requires numerical Monte Carlo integration over each dimension, during the EXIT function synthesis. Moreover, when the receiver has configurable parameters, the dimensionality of the EXIT function further increases. Considering the computations involved during the EXIT function synthesis (see Section 2.3.6), this curse of dimensionality can hinder their efficiency for predicting the behavior of turbo receivers in MANETs.

This issue has been addressed for SISO MMSE LE-IC [TS11], by evaluating many analytical and numerical alternatives to detector's complete EXIT function, by using analytical approximations and channel-independent average-MI based EXIT models of its sub-blocks [KM07; YGP07; VB+10].

In particular [LSS03] proposes a semi-analytical method where an affine analytical approximation of the receiver's EXIT function is shown to predict accurately linear turbo equalizer asymptotically (for large blocks) for BPSK constellation. [RV05] captures equalization model analytically up to the SINR of extrinsic LLRs, then uses only numerical EXIT model to handle the decoder behavior at the bit-level. [KM07] considers the prediction of turbo FD LE-EXTIC

with general QAM signals, by modeling the equalizer and the demapper output LLRs analytically, through an approximation of the constrained BICM capacity, the decoder is modeled with EXIT and a MI-to-variance mapping model is used for the turbo feedback. Alternatively, [SF09] aims for the finite-length prediction of turbo FD LE-EXTIC and proposes analytical models for Gray-mapped QAM constellations, with finite-length BER transfer charts.

Although these methods semi-analytically simplify the SISO detector (equalizer) model, they still suffer from computationally prohibitive multi-dimensional EXIT models at the decoder. This aspect is handled by using an MIESM technique in [YGP07], which studies turbo FDE and FD MIMO detectors. Receiver behavior is captured by a completely analytical modeling the equalized BPSK/QPSK symbols' behavior, and the decoder is modeled through a SINR-variance transfer function.

The contributions above do not use prior information appropriately at the symbol-wise demapper, as devised in [BSY98], which is primordial for exact MAP symbol detection on non-Gray mapped constellations. [VB+10] includes this aspect and unifies the contributions above, through formal modeling of the STBICM joint detection and decoding with binary and symbol-wise perspectives, using a finite-length EXIT at the decoder. Their model is applicable on linear turbo detectors that could either use extrinsic decoder covariance (LE-EXTIC), or APP decoder covariance LE-APPIC, with the latter requiring a calibration procedure for reasonable accuracy.

An important contribution of this reference is the accentuation of the importance of *symbol-wise joint demapping and decoding* model for accurate tracking of asymptotic turbo detection dynamics. Indeed, their binary method carries out MIESM on coded bits' MI, and allows decoupling of mapping from channel coding, while their symbol-wise method jointly models of the demapper and the decoder, and uses symbol-wise MIESM compression with constrained coded-modulation capacity. This approach is more accurate as it captures aspects of finite-length constellations and mapping that does not fit with the assumption required for prediction through EXIT functions (see Assumption 2). This method has later been extended to provide link abstraction for studies carried out on 3GPP standard [NVB13; 3gpc].

5.3.3 FD SILE-EPIC with Finite-Length Prediction

In the remainder of this section, the baseline prediction techniques will be extended for the FD SILE-EPIC turbo receiver, illustrated in Figure 4.3. To this end, first, an asymptotic evolution analysis model for this equalizer is developed, to enable semi-analytical prediction of the SISO detector, with non-analytical LUTs being independent of the channel, and then the joint demapping and decoding approach of [VB+10] is extended to this receiver.

Indeed, as it has been observed in Chapters 3 and 4, that these turbo-iterative receivers that exploit APP soft estimates as feedback for IC suffer from various predictability and significant error-propagation issues. Similar observations have also been made in [Ma+18; Li+19] very recently. These disadvantages are less of a concern with EP-based receivers. Indeed, as evoked in Section 2.4.4, message-passing with EP is shown to be asymptotically

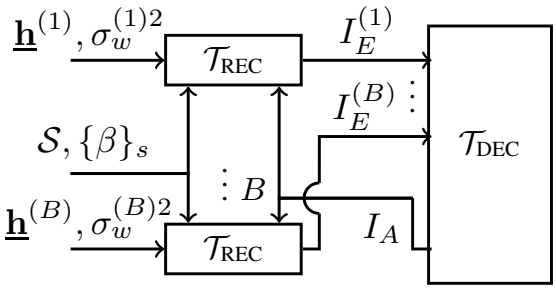


Figure 5.2: EXIT evolution analysis model for the receiver.

predictable through asymptotic MSE state-evolution functions [Tak17; ÇO18], thanks to its divergence-free "decision device" [Ma+18]. However these proofs are limited to Gaussian IID channels, and predictability of EP in general is only an experimental conjecture, as seen in previous chapters with EXIT charts and asymptotic MI trajectory prediction.

5.3.3.1 Asymptotic Evolution Analysis for FD SILE-EPIC

In this subsection, a novel extrinsic information evolution analysis method is proposed, which simplifies the receiver's EXIT function into independent inner transfer functions. The core idea is to track state-evolution dynamics of EP through numerically stable Extrinsic Variance to Information Transfer (EXVIT) functions. These symbol-wise transfer functions, along with an analytical receiver model constitutes a novel evolution analysis method which depends only on channel independent LUT, and analytical closed-form expressions. In the case of BPSK and Gray-mapped QPSK constellations, a full analytical model can be derived and remove the need for generating LUTs. The asymptotic prediction capabilities of this method is illustrated with numerical results.

The Limitations of EXIT Analysis The detection dynamics of EP-based BICM receivers have already been partially investigated in the previous chapters, through extrinsic message evolution analysis, in order to deduce achievable rates in the case where $B = 1$. In the PHY model considered in this chapter with multi-block transmissions with FD SILE-EPIC receiver, the EXIT functions of the receiver and the decoder in Section 2.3.6 are replaced as follows

$$I_E^{(b,\tau)} = \mathcal{T}_{\text{REC}}(I_A^{(\tau)}; \underline{\mathbf{h}}^{(b)}, \sigma_w^{(b)2}, \beta^{(\tau,0)}, \ldots, \beta^{(\tau,\mathcal{S}_\tau)}), \quad b = 1, \ldots, B, \tag{5.1}$$

$$I_A^{(\tau+1)} = \mathcal{T}_{\text{DEC}}(I_E^{(1,\tau)}, \ldots, I_E^{(b,\tau)}, \ldots, I_E^{(B,\tau)}), \tag{5.2}$$

and Figure 5.2 illustrates this model. However as $\mathcal{T}_{\text{REC}}$ depends on inner iterations and its damping parameters, an EXIT function has to be computed for each possible configuration.

Synthesis of EXIT functions for evolution analysis of SISO receivers can be impractical as numerical integrations need to be carried out for each possible channel realization. As seen above, for FD SILE-EPIC, the number of self-iterations and the damping add additional

static parameters which further increase the dimensionality of the EXIT generation procedure. To alleviate this issue, we aim to decompose $\mathcal{T}_{\mathrm{REC}}$ into a set of functions with a symbol-wise evolution analysis method is proposed here, removing the dependence of the synthesis process on channel and on inner loop parameters.

The asymptotic behavior of some EP-based algorithms has previously been investigated through state evolution where MSE on extrinsic estimates are tracked, but here, the BICM context with the SISO demodulator with prior information, and damping brings additional modelling complexity. To address this issue, a symbol-wise extrinsic parameter transfer model is proposed, which consists in characterizing

- the equalizer output (demodulator input) with the modulation constrained, mutual information $I_L^{(b)}$ of the b^{th} data block,
- the equalizer input (demodulator output) with the covariance $v_x^{d(b)}$ of the soft symbol feedback over the b^{th} data block.

The choice of these measures aims to ensure evolution analysis' ability to track evolution dynamics with sufficient accuracy, and to keep them numerically stable with values on finite intervals. Various candidate measures were evaluated for extrinsic evolution dynamics in [TBH02]. Mutual information and second-order statistics of the extrinsic PDFs (called "fidelity" therein) are measures that maintain accuracy, by avoiding restrictive assumptions on the measured extrinsic PDFs. Regarding the numerical stability, we have $I_L^{(b)} \in [0, Q]$, and $v_x^{d(b)} \in [0, \sigma_x^2]$, where $\sigma_x^2 = 1$ is the average symbol power, which enable limited interval representation of dynamic parameters, similarly to conventional binary EXIT functions.

Transfer Function Synthesis for EXVIT Analysis As equalizer's extrinsic estimates are unbiased and decorrelated [Ma+18], and EP is asymptotically Bayes optimal [Tak17], the following assumption is considered.

Assumption 3
The extrinsic symbol outputs of the b^{th} data block's equalizer are IID and for $k = 1, \ldots, K$, we have $x_k^{e(b)} = x_k^{(b)} + w_k^{e(b)}$ with $w_k^{e(b)} \sim \mathcal{CN}(0, v_x^{e(b)})$.

Property 2
If Assumption 3 holds, then the modulation constrained mutual information between the equalized symbols $x_k^{e(b)}$ and the transmitted symbols is the coded modulation AWGN with

$$\psi_{\mathcal{X}}(v_x^{e(b)}) \triangleq I(x_k^{e(b)}; x_k^{(b)}) = Q - \int_{x^e} \log_2 \left(\frac{\sum_{x \in \mathcal{X}} \mathcal{CN}(x^e; x, v_x^{e(b)})}{\mathcal{CN}(x^e; x, v_x^{e(b)})} \right) \mathcal{CN}(x^e; x, v_x^{e(b)}) \, \mathrm{d}x^e.$$

$\psi_{\mathcal{X}}(\cdot)$ depends on $\mathcal{X}$, and it does not have a closed form in general. It can be obtained by using Monte Carlo integration.

The transfer function of the equalizer satisfies

$$I_L^{(b,\tau,s)} = \mathcal{T}_{\text{EQU}}(v_x^{d(b,\tau,s)}; \underline{\mathbf{h}}^{(b)}, \sigma_w^{(b)2}), \tag{5.3}$$

and, it is defined under Assumption 3, with the analytical transfer function

$$\mathcal{T}_{\text{EQU}}(v_x^{d(b,\tau,s)}; \mathbf{h}^{(b)}, \sigma_w^{(b)2}) \triangleq \psi_{\mathcal{X}} \left(\left(\frac{1}{K} \sum_{k=0}^{K-1} \frac{|\underline{h}_k^{(b)}|^2}{\sigma_w^{(b)2} + v_x^{d(b,\tau,s)} |\underline{h}_k^{(b)}|^2} \right)^{-1} - v_x^{d(b,\tau,s)}, \right), \tag{5.4}$$

through the use of Equations (4.36) and (4.38). This constitutes a symbol-wise variance to MI transfer function, as desired, and it is an extension of the semi-analytical equalization prediction methods from the literature [YGP07; VB+10] which processes soft reliability from the EP-based SISO demodulator.

To capture the non-linear behavior of demapping and EP-feedback computation with an extrinsic mutual information to variance transfer function, the dynamics of the extrinsic symbol PDF $\mathcal{CN}(x_k^{d(b,\tau,s)}, v_x^{d(b,\tau,s)})$ is tracked with

$$v_x^{d(b,\tau,s+1)} = \mathcal{T}_{\text{EP}}(I_L^{(b,\tau,s)}, I_A^{(\tau)}, v_x^{d(b,\tau,s)}; \beta^{(\tau,s)}), \tag{5.5}$$

given a value for prior information $I_A^{(\tau)}$ and on extrinsic equalizer output $I_L^{(b,\tau,s)}$. In particular the computation of the feedback and the damping procedure can be explicitly separated with the composition $\mathcal{T}_{\text{EP}} \triangleq \mathcal{T}_{\text{Damp}} \circ \mathcal{T}_{\text{EP}*}$, with

$$v_x^{d(b,\tau,s+1)} = \mathcal{T}_{\text{Damp}} \left(\mathcal{T}_{\text{EP}*}(I_L^{(b,\tau,s)}, I_A^{(\tau)}), v_x^{d(b,\tau,s)}; \beta^{(\tau,s)} \right), \tag{5.6}$$

such that $\mathcal{T}_{\text{Damp}}$ analytically models damping heuristics, and $\mathcal{T}_{\text{EP}*}$ characterizes the behavior of $\mathcal{CN}(x_{x,k}^{\star(b,\tau,s)}, v^{\star(b,\tau,s)})$, obtained through a Gaussian division with Equation (4.33).

In the case where damping is carried out with an exponential smoothing, the analytical damping model is given by

$$\mathcal{T}_{\text{Damp}}(v^{\star(b,\tau,s)}, v^{d(b,\tau,s)}; \beta^{(\tau,s)}) \triangleq (1 - \beta^{(\tau,s)}) v^{\star(b,\tau,s)} + \beta v^{d(b,\tau,s)}, \tag{5.7}$$

and the transfer function $\mathcal{T}_{\text{EP}*}$, of the raw extrinsic EP estimates is obtained with

$$v^{\star(b,\tau,s)} = \mathcal{T}_{\text{EP}*}(I_L^{(b,\tau,s)}, I_A^{(\tau)}) \tag{5.8}$$

such that $\mathcal{T}_{\text{EP}*}$ models the extrinsic covariance $v^{\star(b,\tau,s)}$ of Equation (4.33), with

$$\mathcal{T}_{\text{EP}*}(I_L, I_A) \triangleq \left[\left(\frac{1}{K} \sum_{k=0}^{K-1} \mathcal{T}_{\text{APP},k}(I_L, I_A) \right)^{-1} - \frac{1}{\psi_{\mathcal{X}}^{-1}(I_L)} \right]^{-1}. \tag{5.9}$$

In the function above, $\mathcal{T}_{\text{APP},k}(I_L^{(b,\tau,s)}, I_A^{(\tau)})$ is the expected value of APP covariance of $x_k^{(b)}$, given the equalized estimate $x_k^{e(b,\tau)}$ and the prior LLRs $L(\mathbf{d}_k^{(b,\tau)})$ satisfying Assumptions 1

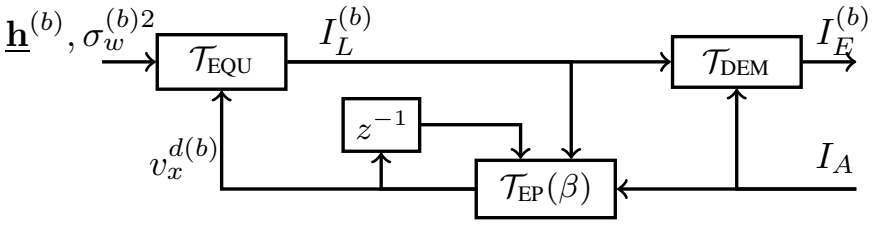


Figure 5.3: Proposed evolution analysis model for $\mathcal{T}_{\text{REC}}$.

and 3. Hence, given I_L and I_A, we have

$$\mathcal{T}_{\text{APP}}(I_L, I_A) \triangleq \frac{1}{M} \sum_{x_k \in \mathcal{X}} \int_{x_k^e, \mathbf{L}_{\mathbf{p},k}} \gamma_k^d(x_k, x_k^e, \mathbf{L}_{\mathbf{p},k}) f_\gamma(x_k, x_k^e, \mathbf{L}_{\mathbf{p},k}) \, \mathrm{d}\mathbf{L}_{\mathbf{p},k} \, \mathrm{d}x_k^e, \tag{5.10}$$

$$f_\gamma(x_k, x_k^e, \mathbf{L}_{\mathbf{p},k}) \triangleq \mathcal{CN}(x_k^e; x_k, v^e) \prod_{q=1}^{Q} \mathcal{CN}\left(L_{p,k,q}; (1 - 2\varphi_q^{-1}(x_k)) J^{-1}(I_A), 2J^{-1}(I_A)\right), \tag{5.11}$$

$$\gamma_k^d(x_k, x_k^e, \mathbf{L}_{\mathbf{p},k}) \triangleq \sum_{\alpha \in \mathcal{X}} \left[\left| \alpha - \sum_{\alpha' \in \mathcal{X}} \alpha' \frac{\mathcal{CN}(x_k^e; \alpha', v^e) \prod_{q=0}^{Q-1} e^{-\varphi_q^{-1}(\alpha') L_{p,k,q}}}{\sum_{\alpha" \in \mathcal{X}} \mathcal{CN}(x_k^e; \alpha", v^e) \prod_{q=0}^{Q-1} e^{-\varphi_q^{-1}(\alpha") L_{p,k,q}}} \right|^2 \right.$$
$$\left. \frac{\mathcal{CN}(x_k^e; \alpha, v^e) \prod_{q=0}^{Q-1} e^{-\varphi_q^{-1}(\alpha) L_{p,k,q}}}{\sum_{\alpha" \in \mathcal{X}} \mathcal{CN}(x_k^e; \alpha", v^e) \prod_{q=0}^{Q-1} e^{-\varphi_q^{-1}(\alpha") L_{p,k,q}}} \right], \tag{5.12}$$

with $\mathbf{L}_{\mathbf{p},k} = [L_{p,k,0}, \ldots, L_{p,k,Q-1}]$. While $\mathcal{T}_{\text{EP}*}$ has no closed form expression in general, it exclusively depends on $\mathcal{X}$ and on K[1].

Finally, to complete the characterization of FD SILE-EPIC, SISO demapper's extrinsic LLR output needs to be characterized with respect to the estimates of the final self-iteration. Following Assumption 3, the equalizer's output is seen as an AWGN channel, thus we use the EXIT function of a SISO MAP symbol demapper for AWGN, with

$$I_E^{(b,\tau)} = \mathcal{T}_{\text{DEM}}(I_A^{(\tau)}, I_L^{(b,\tau,\mathcal{S}_\tau)}). \tag{5.13}$$

In conclusion, symbol-wise EXVIT analysis consists in applying transfer functions $\mathcal{T}_{\text{EP}}$ and $\mathcal{T}_{\text{EQU}}$ successively, on each block b to obtain $\{v^{d(b,\tau,\mathcal{S}_\tau)}\}_{p=1}^{B}$ and then $\{I_L^{(b,\tau,\mathcal{S}_\tau)}\}_{p=1}^{B}$, starting with $v^{d(b,\tau,-1)} = 1, \forall b$. Then, the extrinsic output of the overall receiver is $\{I_E^{(b,\tau)}\}_{p=1}^{B}$, computed with $\mathcal{T}_{\text{DEM}}$.

Unlike EXIT analysis, EXVIT depends analytically on the channel $\{\underline{\mathbf{h}}^{(b)}, \sigma_w^{(b)2}\}_{b=1}^{B}$ and on $\{\beta^{(\tau,s)}\}_{s=0}^{\mathcal{S}_\tau}$. Numerical integrations are only needed for the two-dimensional $\mathcal{T}_{\text{EP}*}$ and $\mathcal{T}_{\text{DEM}}$ functions, which only depend on $\mathcal{X}$ and on K.

[1]The dependence on K has no effect at this point, as $\mathcal{T}_{\text{APP},k}$ is independent on k, but in the next section we will get back to this point.

Simplifying the decoder's EXIT function The decoder's EXIT function in Equation (5.2) requires as many inputs as B, the number of channel realizations over the codeword, which elevates synthesis complexity. A similar issue is noted for evolution analysis of BPSK MIMO systems in [YGP07]. The problem is identical from the MI point of view, hence the following assumption is made to address it.

Assumption 4
The decoder behavior is identical for inputs constituted from a LLRs of B data blocks of K symbols, i.e. $\{L_e(\mathbf{d}_k^{(b)})\}_{k=1}^{K}$, $b = 1, \ldots, B$, and for LLRs of a single data block of KB symbols, e.g. $\{\tilde{L}_e(\mathbf{d}_{k'})\}_{k'=1}^{KB}$, if they have the same average MI with respect to $\mathbf{d}$.

Thus the SISO decoder's EXIT function is reduced to single dimensional mapping with the effective mutual information

$$I_A^{(\tau+1)} = \mathcal{T}_{\mathrm{DEC}}\left(\frac{1}{B}\sum_{b=1}^{B} I_E^{(b,\tau)}\right). \tag{5.14}$$

This operation is equivalent to performing MIESM, considering the bijective relationship between the codeword MI and the post-equalization variance through $\psi_{\mathcal{X}}$.

Analytic evolution analysis for BPSK/QPSK This section provides analytical derivation of $\mathcal{T}_{\mathrm{EP}^*}$ for BPSK constellation (extension to Gray-mapped QPSK is straightforward), in order to provide an analytical approximation to the computationally intensive numerical integrations within $\mathcal{T}_{\mathrm{APP}}$. Due the symmetry of the constellation, and IID LLRs, $\gamma_k^d(x, x_k^e, L_{p,k}) = \gamma_k^d(1, x^e, L_p)$, $x^e = 1 + w^e$, $w^e \sim \mathcal{N}(0, v^e)$, and $L_p = \mu_p + w^p$, with $w^p \sim \mathcal{N}(0, 2\mu^p)$. Then

$$\gamma^d(w^e, w^p) \triangleq \gamma_k^d(1, x^e, L_p) = 1 - \left|\tanh\left(\frac{1}{2}(\mu_d + w^d)\right)\right|^2, \tag{5.15}$$

$$\mu_d \triangleq \mu_p + 2/v^e, \qquad w^d \triangleq w^p + 2w^e/v^e, \tag{5.16}$$

with $\mu_p = J^{-1}(I_A)$, $v^e = \psi_{\mathcal{X}}^{-1}(I_L) = 2/J^{-1}(I_L)$. Thus, we have $w^d \sim \mathcal{N}(0, 2\mu^d)$ and

$$\mathcal{T}_{\mathrm{APP}}(I_L, I_A) = 1 - \int_{w^d} \left|\tanh\left(\frac{1}{2}(\mu_d + w^d)\right)\right|^2 \mathcal{N}(w^d; 0, 2\mu_d)\, \mathrm{d}w^d.$$

and the Gaussian integrals therein can be approximated analytically with an n_q^{th} order Gauss-Hermite quadrature [SBG69], with

$$\mathcal{T}_{\mathrm{APP}}(I_L, I_A) \approx 1 - \sum_{i=1}^{n_q} \frac{2^{i-1} i!}{i^2 \mathcal{H}_{n_q-1}^2(x_i)} \left|\tanh\left(\frac{1}{2}(\mu_d + 2\sqrt{\mu_d} x_i)\right)\right|^2, \tag{5.17}$$

where x_i is the i^{th} root of the Hermite polynomial $\mathcal{H}_{n_q}(X)$. Experimental analysis indicates that using a small value for n_q creates issues at operating points close to $(v^d, I_L, I_A) = (0, 1, 0)$, and that $n_q \leqslant 20$ is precise for $I_L \leqslant 0.95$.

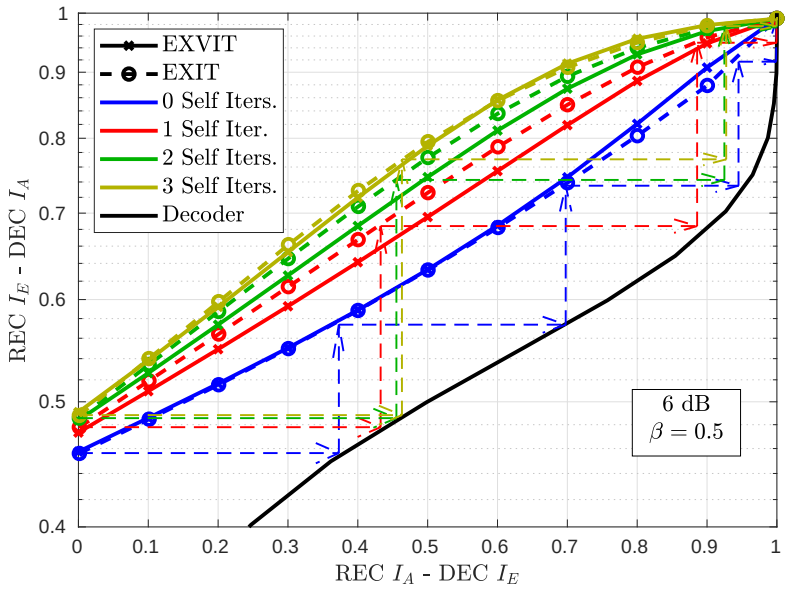


Figure 5.4: EXIT & EXVIT functions with $[1, 5/7]_8$ recursive convolutional code.

Numerical Results and Discussion In the following, BPSK and QPSK EXVIT functions are evaluated analytically, with $n_q = 15$, and 16-QAM EXVIT, and all EXIT functions are given by Monte-Carlo evaluations.

First, the accuracy of the proposed EXVIT model is evaluated by comparison with the measured MI trajectories of a coded BPSK system with $K_b = 16384$ and $B = 1$. The closeness of transfer functions to the measured MI trajectories indicates the accuracy of an analysis method. Equivalent EXIT functions in Proakis C channel, $[0.23, 0.46, 0.69, 0.46, 0.23]$, are plotted in Figure 5.4, for a SNR of 5 dB. Both EXVIT and EXIT transfer functions appear to be close the MI trajectories for the first few turbo-iterations, hence both methods manage to predict initial turbo-iterations fairly well. Besides, there is a relatively slight difference between both curves, hence they would predict similar decoding thresholds or achievable rates.

Next, to assess the accuracy of EXVIT for predicting more quantitatively, transmissions of $K_b = 16384$ bits over a 10-tap Rayleigh fading channel with uniform power profile is considered with $\beta = 0.3$. The codeword, obtained with the recursive convolutional code $[1, 5/7]_8$ is subject to $B = 8$ channel realizations, and the decoder's BER and PER are quantified with regards to its input prior information. Finite-length simulations are plotted with solid lines on Figure 5.5, and EXVIT is used to track the evolution of the extrinsic information of the equalizer, which is fed into the decoder, and to predict the corresponding error rate performance. Predictions are carried out on each channel realization on which that actual PHY layer is simulated, and the average predicted BER and PER are plotted. Predictions appear to be close to the actual Monte-Carlo simulations, for QPSK and 16-QAM constellations, at the zeroeth and the first turbo-iterations. These results attest to the accuracy of the proposed method.

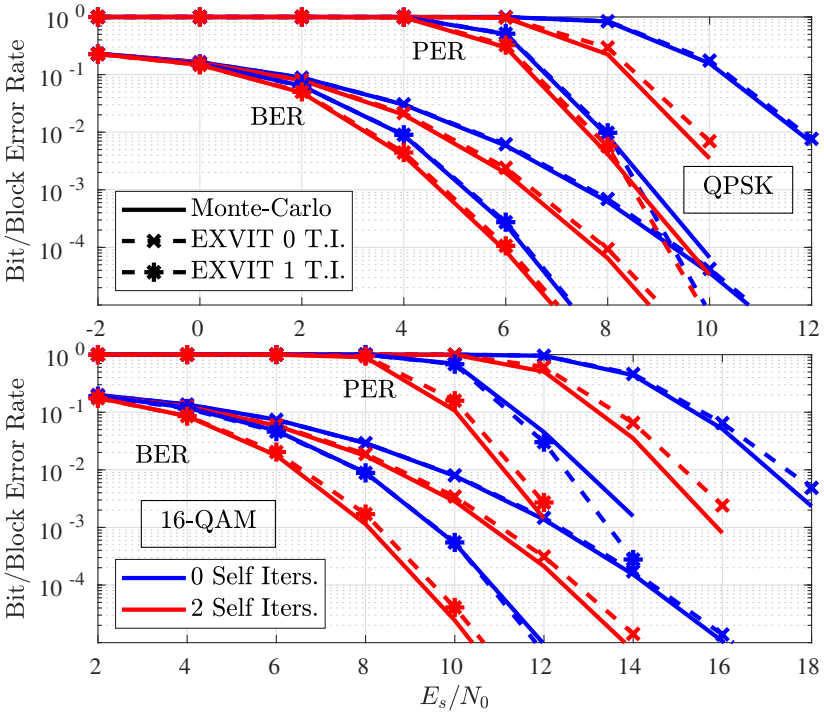


Figure 5.5: BER and PER prediction with analytical EXVIT.

In conclusion, a symbolwise extrinsic analysis method for an EP-based, highly non-linear receiver is proposed. This approach replaces the EXIT function of the receiver, with inner EXVIT functions having analytical dependence on the channel and the receiver parameters. The receiver is shown to be predicted with adequate accuracy over multi-block channels only with a few numerical integrations and independently of the modulation format.

5.3.3.2 Semi-Analytic FD SILE-EPIC Prediction with Joint Demapping and Decoding

In this subsection, the prediction scheme of [VB+10] is extended to handle double-loop scalar EP dynamics, through the use of the asymptotic analysis method proposed in the previous subsection. The schematic of this semi-analytical finite-length prediction scheme is given in Figure 5.6.

Indeed, asymptotic evolution analysis for filter-based turbo receivers can only accurately predict the decoding performance when $\tau = 0$ and $\tau \to +\infty$. Otherwise, as mentioned in Section 2.3.6, the Assumption 2 is needed for carrying out predictions through asymptotic extrinsic analysis. However it has been observed in [Fu05] that we cannot consider the conditional distributions of extrinsic LLRs as symmetric Gaussian distribution during intermediary turbo-iterations. As a consequence, asymptotic prediction is only accurate for the first few turbo-iterations, and alternative methods are needed for alleviating this issue.

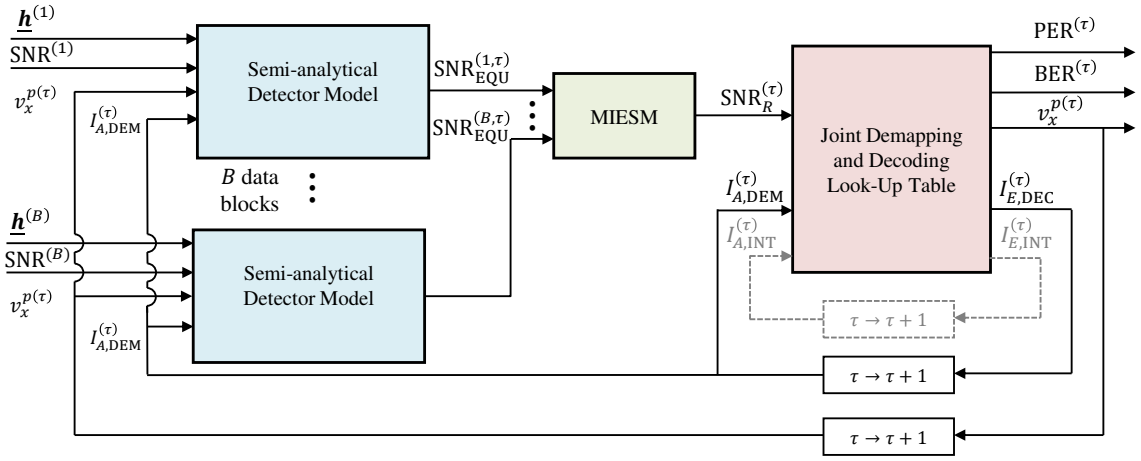


Figure 5.6: Symbolwise semi-analytical turbo-receiver prediction scheme based on [VB+10].

The symbolwise prediction method proposed in [VB+10], for LE-EXTIC and LE-APPIC, exploits a joint demapping and decoding model. This model does not directly characterize the behavior of extrinsic LLRs used during turbo-iterations, in order to avoid making strong assumptions on the conditional LLR distributions. It makes a symbol-wise abstraction of demapping and decoding behavior, by taking as input an effective received SNR on the equivalent AWGN channel (see Assumption 3) at the soft demapper's input, SNR_R, and the prior information at the LLR input of the soft demapper, $I_{A,\mathrm{DEM}}$. At the output, the covariance v_x^p of the extrinsic soft symbols produced by the SISO decoder and soft mapping (see Equation (2.40) is used for keeping track of symbol-level turbo dynamics, along with desired performance metrics such as the predicted BER or PER at the ongoing iteration. Binary dynamics can also followed by measuring the MI of extrinsic LLRs, denoted $I_{E,\mathrm{DEC}}$.

The effective SNR or SINR is obtained by MIESM [Bru+05] with

$$\mathrm{SNR}_R = \alpha_{\mathrm{ESM}} I^{-1}\left(\frac{1}{B}\sum_{b=1}^{B} I\left(\frac{\mathrm{SNR}_{\mathrm{EQU}}^{(b)}}{\alpha_{\mathrm{ESM}}}\right)\right), \tag{5.18}$$

where $\mathrm{SNR}_{\mathrm{EQU}}^{(b)}$ is the post-equalization SNR of the b^{th} data block, α is a calibration parameter and $I(\cdot)$ is a constrained-capacity function, such as the coded-modulation capacity of Equation (1.37), i.e. $I(\mathrm{SNR}) = \psi_{\mathcal{X}}(\sigma_x^2/\mathrm{SNR})$, or the BICM capacity of Equation (1.38).

This LUT is synthesized by generating samples from the finite-length BICM encoder of the PHY layer, from information bits up to data symbols. While this procedure is computationally intensive, the resulting LUT depends only on the channel code $\mathcal{C}$, the interleaver Π, the soft mapper ϕ of constellation $\mathcal{X}$, and the block length K. For a well-defined wireless system, required LUTs can be generated only once for all available MCS. This model can also track the iterative dynamics of constituent SISO modules of the channel decoder, such as parallel-concatenated turbo codes [NVB13], by using a prior information input $I_{A,\mathrm{INT}}$ for the inner code, and measuring the extrinsic information $I_{E,\mathrm{INT}}$ at the output.

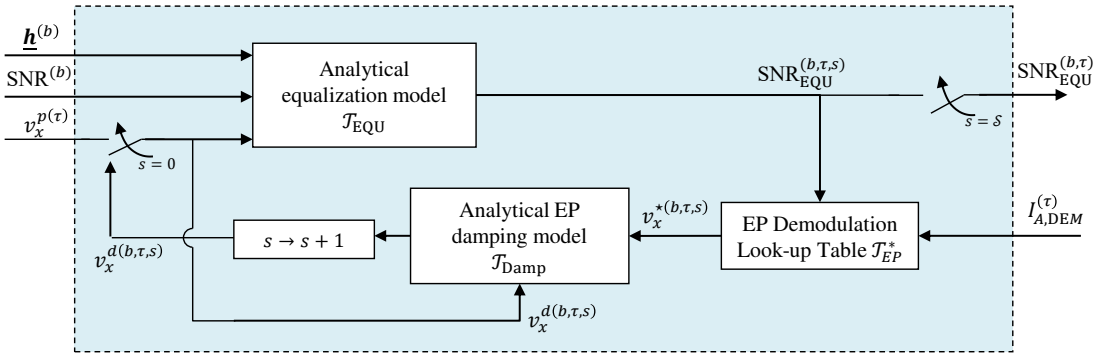


Figure 5.7: Proposed semi-analytical detector model for FD SILE-EPIC.

Although this "finite-length" method predicts considerably more accurately than the pure asymptotic analysis approach, its accuracy seems to depend on the channel parameters and the link configuration (number of antennas, constellation, code rate etc.). In particular, it has been observed to significantly overestimate the decoding threshold for very selective channels, such as the Proakis C, or fading channels with high Rice factors, whereas for Rayleigh fading the accuracy remains adequate. Moreover, when feeding APP covariances to the analytical equalization model, this technique loses a considerable amount of precision. To remedy this, the reference [NVB12] has proposed to use a calibration method on the soft covariance feedback, as previously discussed in Chapter 3, in the context of the FIR DFE with online prediction. Following calibration, this method accurately predicts a vast among of configurations and it has been proposed for reviewing turbo receiver performance for 3GPP standard [3gpc].

We now propose to extend the previously derived asymptotic analysis scheme for FD SILE-EPIC for operating within this semi-analytic prediction framework. To this end, the following changes are made to the EXVIT model of this receiver:

- when $s = 0$, due to the nature of EP-based feedback, we have $v_x^{d(b,\tau,0)} = v_x^{p(\tau)}$, this has to be explicitly reflected for the input of the equalization transfer function $\mathcal{T}_{\text{EQU}}$, with a switch, as $v_x^{p(\tau)}$ is estimated by the joint demapping and decoding model,

- the output of $\mathcal{T}_{\text{EQU}}$ is the extrinsic symbol-wise mutual information $I_L^{(b,\tau,s)}$, but this output constraints the MIESM to not use its calibration parameter α_{ESM}, hence for more flexibility, SNR$_{\text{EQU}}^{(b,\tau,s)}$ is used as this function's output, by redefining

$$\mathcal{T}_{\text{EQU}}(v_x^{(b)}; \underline{\mathbf{h}}^{(b)}, \sigma_w^{(b)2}) \triangleq \sigma_x^2 \left(\left(\frac{1}{K} \sum_{k=0}^{K-1} \frac{|\underline{h}_k^{(b)}|^2}{\sigma_w^{(b)2} + v_x^{(b)} |\underline{h}_k^{(b)}|^2} \right)^{-1} - v_x^{(b)} \right)^{-1}, \tag{5.19}$$

- a switch has to be inserted, in order to affect the post-equalization SINR value at the final self-iteration, to the output, i.e. SNR$_{\text{EQU}}^{(b,\tau)} = $ SNR$_{\text{EQU}}^{(b,\tau,\mathcal{S}_\tau)}$.

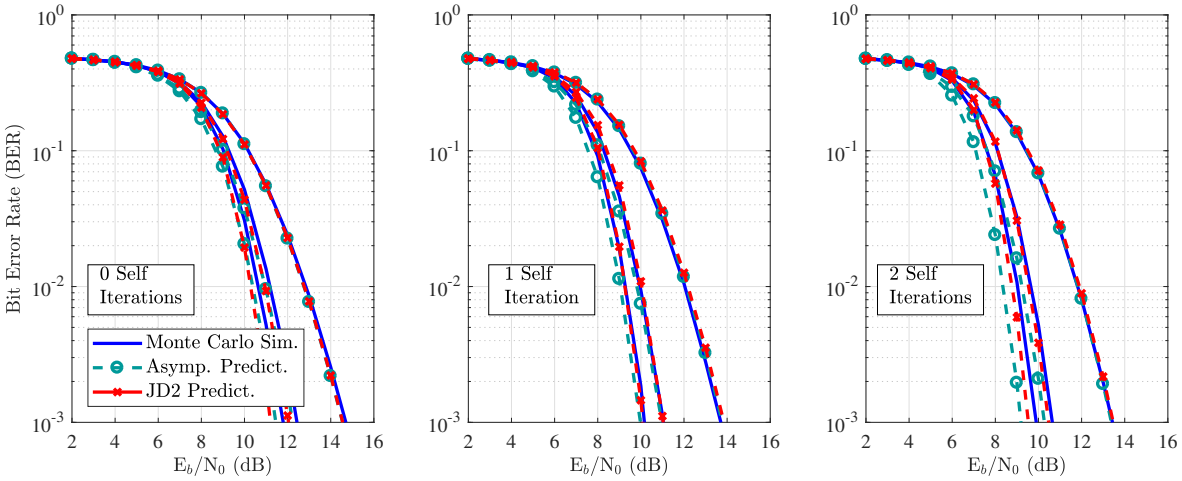


Figure 5.8: Comparison of the asymptotic prediction scheme and the finite-length JD2 methods, as turbo-iterations go from 0 to 2.

This proposed semi-analytical model is illustrated in Figure 5.7, and it is abbreviated as the joint demapping and decoding (JD2) prediction scheme. In Figure 5.8, we consider the transmission of four data blocks ($B = 4$) over a uniform, symbol spaced power-delay profile, such that each tap is Ricean distributed with $K_l^{\mathrm{SS}} = 10$, for $l = 1, \ldots, 10$, and the 16-QAM constellation is used along with the convolutional code of polynomial $[133, 165, 171]_8$, optimally punctured by $[1, 0, 1; 1, 0, 1]$ to obtain a rate 1/2 code [Fre+98], with $K = 128$. Asymptotic schemes typically manage to predict finite-length transmission on fading Rayleigh channels, but as line of sight components become prominent the prediction quality decreases. In the considered scenario, although the asymptotic prediction mechanisms copes fairly well with the Rice factor and short blocks, it starts providing too optimistic predictions as the number self-iterations increase, in this case, JD2 becomes helpful for maintaining accuracy.

5.4 Pseudorandom Finite-Length PHY Layer Prediction

While the symbol-wise semi-analytic finite-length prediction scheme above can yield satisfying accuracy, it may require intensive calibration, which is undesirable for a system-level simulator to be used in protocol design, where the definitive MCS configurations may not have been established.

To alleviate this drawback, in this section we investigate a method for naturally incorporating the impact of data block and codeword lengths, without exhaustive calibration. This idea is based on the observations of [Rib+15], where a heuristic is characterized for fitting the predicted performance of the iterative block DFE receiver (FD SILE-APPIC without turbo-iterations) to the actual finite-length performance. This is achieved scaling the predicted post-equalization SNR by a compensation factor that is dependent on the multi-path channel's delay spread.

However, the idea of a channel dependent heuristic is unattractive to us, especially when considering cooperative MANET where the CIR may abruptly change from one block to another, if a protocol such as cooperative broadcasting is used. To this end, by keeping the general idea of compensating the post-equalization SNR, that ideally depends on the block length, we have looked into the works on finite-length coding and short-packet communications communities for inspirations on heuristics. Following this, the behavior of FD SILE-EPIC is investigated, when the soft feedback covariance is considered as a random process whose statistical dispersion depends on the block length. Finally, we propose a novel finite-length prediction scheme that uses pseudo-random behavior of the soft-feedback in order to take in account for the performance degradation caused by the lengths of data blocks and codewords.

5.4.1 On Finite-Length Information Theory and Channel Coding

There is increased interest on short-packet communications, in recent years, due to the emerging ultra-reliable communications (URLLC) and machine-to-machine (M2M) use-cases for 5G networks [DKP16]. This has gathered significant interest in PHY design for short blocks through the developments on finite block length information theory [PPV10].

Let us consider x and y as respectively the input and the output of a channel, then let us recall that according to the noisy channel coding theorem, given in Equation (1.33), there exists *asymptotically* a channel code of rate $R \leqslant C = \max_{p(x)} I(x;y)$ which achieves *error-free* communications over the channel defined by $p(y|x)$, of Shannon capacity C. An alternative way to look at this problem is to define $R^*(n,\varepsilon)$ as the maximum achievable rate for the channel defined by this channel, by considering that a sequence $\mathbf{x}$ of length n is transmitted, and $\mathbf{y}$ of length n is received, with a packet error probability of ε. Then by relying on the asymptotically error-free definition of capacity, and using Equation (1.31), we have

$$C = \lim_{n\to+\infty} \lim_{\varepsilon\to 0} R^*(n,\varepsilon) = \mathbb{E}_{p(\mathbf{x},\mathbf{y})}\left[\log_2 \frac{p(\mathbf{y}|\mathbf{x})}{p(\mathbf{y})}\right]. \tag{5.20}$$

Considering that the capacity appears to be the expected value of what we can assimilate to a random variable, the *information density* is defined as $i(\mathbf{x};\mathbf{y}) = \log_2 \frac{p(\mathbf{y}|\mathbf{x})}{p(\mathbf{y})}$. This distribution's behavior as a function of n constitutes a fundamental challenge for finite-length code design, as it enables the characterization of upper and lower bounds as a function of n, R^* and ε. However this is not simple, as the complexity of determining $R^*(n,\varepsilon)$ appears to be doubly exponential in n.

An attractive idea emerged in [PPV10], where a Gaussian approximation is made for the distribution $i(\mathbf{x};\mathbf{y})$, such that the mean $C = \mathbb{E}[i(\mathbf{x};\mathbf{y})]$ is the capacity, as before, and the variance $V = \mathrm{Var}[i(\mathbf{x};\mathbf{y})]$ is defined as the *channel dispersion*. Through this approximation, an explicit expression of $R^*(n,\varepsilon)$ simplified the derivation of upper or lower bounds, and the channel characterization problem is reduced to the computation of the dispersion V.

5.4.2 On the Impact of Soft Feedback Reliability Dispersion

Following a similar line of thought in the context of turbo detection with FD SILE-EPIC, the expression of extrinsic variances in computations are exact when block lengths are asymptotically large. Following the treatment of information density, one can treat the soft feedback reliability as a random variable whose mean is the asymptotic value v_x^d.

First, let $\tilde{v}_x^d \triangleq \mathrm{Var}[\mathbf{x^d} - \mathbf{x}]$ be the "instantaneous" covariance of soft estimates computed at the EP-based demapper, and let the covariance error $\epsilon_{v_x^d}$ be defined as $\epsilon_{v_x^d} \triangleq \tilde{v}_x^d - v_x^d$, such that the turbo equalizer is computed by assuming $\tilde{v}_x^d \approx v_x^d$, and the value of $\epsilon_{v_x^d}$ is unknown for filter computation. In this setting, the filters are mismatched to the true value of prior estimate covariance, and we then investigate the impact of a covariance error at the equalizer's output. Indeed, let us denote $\boldsymbol{\epsilon}_{\mathbf{x^e}} \triangleq \mathbf{x}^e - \mathbf{x}$, then following Equations (4.35)-(4.38) we have

$$\begin{aligned}\boldsymbol{\epsilon}_{\mathbf{x^e}} &= \mathbf{x}^e - \mathbf{x} \\ &= \mathbf{x}^d + \mathcal{F}_K^H \underline{\mathbf{F}}^H \mathcal{F}_K (\mathbf{H}(\mathbf{x} - \mathbf{x}^d) + \mathbf{w}) - \mathbf{x} \\ &= \left(\mathbf{I}_K - \mathcal{F}_K^H \underline{\mathbf{F}}^H \underline{\mathbf{H}} \mathcal{F}_K\right)(\mathbf{x}^e - \mathbf{x}) + \mathcal{F}_K^H \underline{\mathbf{F}}^H \mathcal{F}_K \mathbf{w}, \end{aligned} \tag{5.21}$$

where $\underline{\mathbf{F}} = \mathbf{Diag}([\underline{f}_1^{\mathrm{ep}}; \ldots; \underline{f}_K^{\mathrm{ep}}])$. By computing the covariance of this vector, we have

$$\begin{aligned}\mathbf{Cov}\left(\boldsymbol{\epsilon}_{\mathbf{x^e}}\right) &= \left(\mathbf{I}_K - \mathcal{F}_K^H \underline{\mathbf{F}}^H \underline{\mathbf{H}} \mathcal{F}_K\right) \mathbf{Cov}(\mathbf{x}^e - \mathbf{x}) \left(\mathbf{I}_K - \mathcal{F}_K^H \underline{\mathbf{F}}^H \underline{\mathbf{H}} \mathcal{F}_K\right)^H \\ &\quad + \mathcal{F}_K^H \underline{\mathbf{F}}^H \mathcal{F}_K \mathbf{Cov}(\mathbf{w}) \mathcal{F}_K^H \underline{\mathbf{F}} \mathcal{F}_K \\ &= \left(\mathbf{I}_K - \mathcal{F}_K^H \underline{\mathbf{F}}^H \underline{\mathbf{H}} \mathcal{F}_K\right) \left(\mathbf{I}_K - \mathcal{F}_K^H \underline{\mathbf{F}}^H \underline{\mathbf{H}} \mathcal{F}_K\right)^H (v_x^d + \epsilon_{v_x^d}) + \mathcal{F}_K^H \underline{\mathbf{F}}^H \underline{\mathbf{F}} \mathcal{F}_K \sigma_w^2 \\ &= \left(\mathbf{I}_K + \mathcal{F}_K^H \underline{\mathbf{F}}^H \underline{\mathbf{H}} \underline{\mathbf{H}}^H \underline{\mathbf{F}} \mathcal{F}_K - 2\Re(\mathcal{F}_K^H \underline{\mathbf{F}}^H \underline{\mathbf{H}} \mathcal{F}_K)\right)^H \epsilon_{v_x^d} \\ &\quad + \left(\mathbf{I}_K - 2\Re(\mathcal{F}_K^H \underline{\mathbf{F}}^H \underline{\mathbf{H}} \mathcal{F}_K)\right) v_x^d + \xi_x^{\mathrm{ep}\,-1} \mathcal{F}_K^H \underline{\mathbf{F}}^H \underline{\mathbf{H}} \mathcal{F}_K. \end{aligned} \tag{5.22}$$

Considering that $\mathbf{e}_k^H \mathcal{F}_K^H \underline{\mathbf{F}}^H \underline{\mathbf{H}} \mathcal{F}_K \mathbf{e}_k = 1$, by denoting $\chi_x^{\mathrm{ep}} \triangleq \mathbf{e}_k^H \mathcal{F}_K^H \underline{\mathbf{F}}^H \underline{\mathbf{H}} \underline{\mathbf{H}}^H \underline{\mathbf{F}} \mathcal{F}_K \mathbf{e}_k$, we have

$$\begin{aligned}\mathrm{Var}\left(\epsilon_{x^e,k}\right) = \mathbf{e}_k^H \mathbf{Cov}\left(\boldsymbol{\epsilon}_{\mathbf{x^e}}\right) \mathbf{e}_k &= 1/\xi_x^{\mathrm{ep}} - v_x^d + (\chi_x^{\mathrm{ep}} - 1)\epsilon_{v_x^d} \\ &= v_x^e + (\chi_x^{\mathrm{ep}} - 1)\epsilon_{v_x^d}. \end{aligned} \tag{5.23}$$

This expression reveals that if there is a mismatch between the value of prior variance and the actual variance (or reliabilities) of prior estimates, then the post-equalization variance v_x^e is affected by an additive term proportional to this mismatch. In particular, the quantity χ_x^{ep} is the power of the channel-after-equalization (i.e. the channel representing the residual ISI), which can be considered also as a measure of frequency-selectivity of the physical channel $\mathbf{H}$. Indeed, when the CFR $\underline{\mathbf{H}}$ does not have spectral nulls and the filter $\underline{\mathbf{F}}$ is able to invert the channel response, $\underline{\mathbf{H}}^H \underline{\mathbf{F}}$ is close to the identity matrix $\mathbf{I}_K$, and χ_x^{ep} is close to 1. Otherwise, our numerical experimentations indicate that χ_x^{ep} increases significantly when the channel cannot be completely inverted (which naturally also depend on the SNR).

The behavior of χ_x^{ep} with respect to channel selectivity also coincides with the situations where the behavior of the un-calibrated semi-analytical prediction method of the previous

section, which loses accuracy on more selective channels, while having no issues in "simple" channels. These observations have motivated investigating the use of a positive penalty term $\epsilon_{v_x^d}$ on the predicted post-equalization residual noise and variance, in order to attempt to account for the negative impact of short-blocks on the detection performance.

In this regard, we can consider $\epsilon_{v_x^d}$ to be a random variable such that $\mathrm{Var}[\epsilon_{v_x^d}] = \sigma^2_{v_x^d}$ is the prior covariance dispersion, or "reliability dispersion", caused by the finite block length. Ideally, we would need to impose $\mathbb{E}[\tilde{v}_x^d] = v_x^d$ to maintain accuracy asymptotically, but it is preferable to release this constraint, and consider $\epsilon_{v_x^d}$ as a positive random value, to avoid optimistic predictions, at the expense of inducing some amount of pessimistic bias to the average value of prior covariance. Hence, if the value of $\sigma^2_{v_x^d}$ is known, then it is possible to describe the post-equalization SINR as a random variable, and perform prediction over a number of pseudorandom trials. In the next sub-section, we propose an original prediction scheme based on this idea, which describes an effective method for estimating $\sigma^2_{v_x^d}$, and tracking its value across iterations, in order to perform semi-analytical prediction over pseudorandom trials of the prior covariance error.

5.4.3 Finite-Length Prediction with Pseudorandom Feedback Dispersion

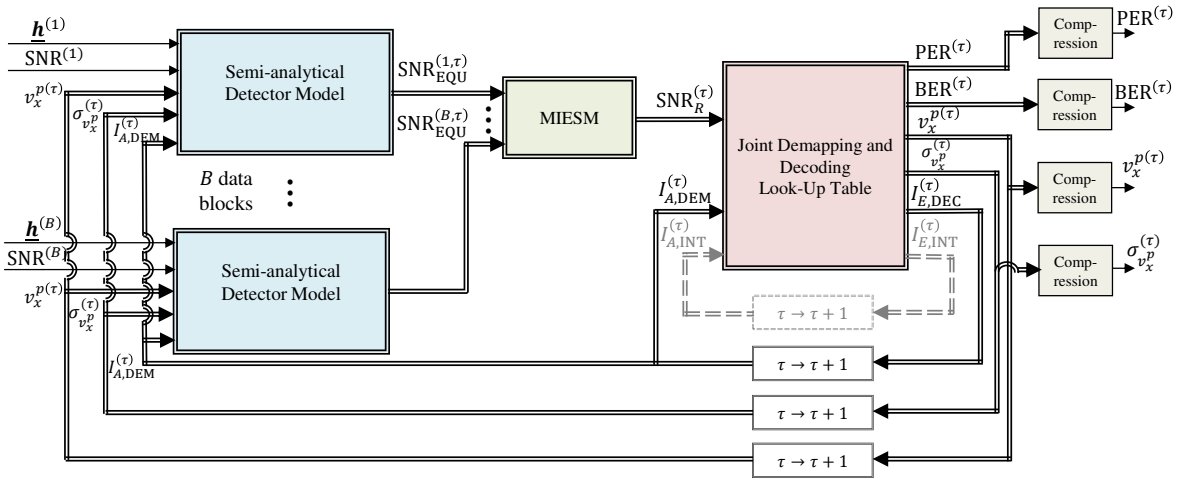
The proposed prediction scheme is illustrated in Figure 5.9, where double lines indicate that there are N_{try} copies of the concerned modules and connections, such than an independent prediction scheme is carried out for $n = 1, \ldots, N_{\mathrm{try}}$, corresponding to different realizations of the random reliability error $\epsilon_{v_x^d}[n]$.

First, the proposed scheme requires revising the LUT synthesis procedure for the joint demapping and decoding model, in order to measure the standard deviation of finite-length (codeword length of KBQ bits) extrinsic soft estimate covariances, in order to characterize the dispersion $\sigma_{v_x^p}$ of v_x^p as a function of its inputs.

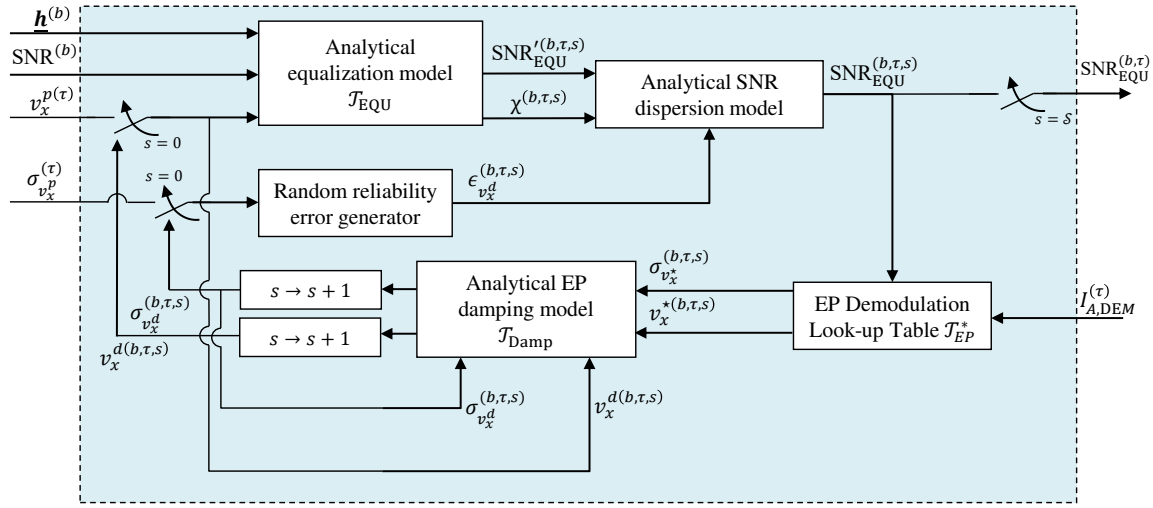
Within the semi-analytical detector model for FD SILE-EPIC, the output SNR of the analytical equalization model is renamed as $\mathrm{SNR}_{\mathrm{EQU}}^{\prime(b,\tau,s)}$, and a new output is added with $\chi_x^{(b,\tau,s)}$. And finally, a new module is added for analytically modelling the penalties incurred by the covariance error on the post-equalization SINR, which operates as follows

$$\mathrm{SNR}_{\mathrm{EQU}}^{(b,\tau,s)} = \begin{cases} \left[\frac{1}{\mathrm{SNR}_{\mathrm{EQU}}^{\prime(b,\tau,s)}} + (1-\chi_x^{(b,\tau,s)})\epsilon_{v_x^d}^{(b,\tau,s)}\right]^{-1}, & \text{if } \frac{1}{\mathrm{SNR}_{\mathrm{EQU}}^{\prime(b,\tau,s)}} \geqslant (1-\chi_x^{(b,\tau,s)})\epsilon_{v_x^d}^{(b,\tau,s)}, \\ \left[\frac{1}{\mathrm{SNR}_{\mathrm{EQU}}^{\prime(b,\tau,s)}} - (1-\chi_x^{(b,\tau,s)})\epsilon_{v_x^d}^{(b,\tau,s)}\right]^{-1}, & \text{otherwise,} \end{cases} \tag{5.24}$$

where $\epsilon_{v_x^d}^{(b,\tau,s)}$ is the pseudorandom covariance error, of standard deviation $\sigma_{v_x^d}^{(b,\tau,s)}$, and a heuristic is used to avoid negative covariances that may occur due to of $\epsilon_{v_x^d}^{(b,\tau,s)}$. In this context, it may be attractive to look into family of PDF that are defined only for positive values, such that, when the variance of the distribution is set to $\sigma_{v_x^d}^{(b,\tau,s)2}$, the expected value $\mathbb{E}[\epsilon_{v_x^d}^{(b,\tau,s)}]$ remains relatively close to zero. In this regard, exponential and gamma distributions have



(a) Proposed pseudo-random semi-analytical turbo-receiver prediction scheme.



(b) Proposed pseudo-random semi-analytical detector model for FD SILE-EPIC.

Figure 5.9: Pseudo-random prediction of with joint demapping and decoding LUT.

attractive properties, with the latter having more degrees of freedom.

Next, the EP demodulation LUT measures the standard deviation of finite-length (data block length of K symbols) extrinsic soft estimate covariances, similarly to the new joint demapping and decoding model, in order to characterize the dispersion $\sigma_{v_x^\star}$ of the raw EP estimate covariance $v_x^\star$, as a function of its inputs.

Furthermore, the analytical EP damping model is modified to update the dispersion as follows

$$\sigma_{v_x^d}^{(b,\tau,s+1)} = \sqrt{(1-\beta^{(\tau,s)})^2 \sigma_{v_x^\star}^{(b,\tau,s)2} + \beta^{(\tau,s)2} \sigma_{v_x^d}^{(b,\tau,s)2} + (1-\beta^{(\tau,s)})^2 (\chi_x^{(b,\tau,s)} - 1)^2 \sigma_{v_x^d}^{(b,\tau,s)2}}, \tag{5.25}$$

where the first two terms are naturally obtained by assessing the impact of random variables for the damping based on the exponential smoothing, given in Equation (4.40). The latter term, weighed by $(1-\beta^{(\tau,s)})^2(\chi_x^{(b,\tau,s)}-1)^2$ is a heuristic, that captures the error propagation from the dispersion of post-equalization covariance on the freshly computed raw-estimate covariance dispersion $\sigma_{v_x^\star}$. Indeed, $\sigma_{v_x^\star}$ is directly obtained with the EP demodulation LUT, which expects an asymptotic SINR value at its input, but instead it is fed with the post-equalization SINR with dispersion penalty. Hence potential finite-length block errors are neglected, and the additive term aims to compensate this negligence.

Considering these modified blocks, it is seen that the output of the prediction scheme of Figure 5.9-(b) depends on the values of the random reliability errors $\{\epsilon_{v_x^d}^{(b,\tau,s)}\}_{s=1}^{\mathcal{S}}$. Hence if N_{try} different error realizations are considered following the first self or turbo iteration, then the semi-analytical detector model will predict a post-equalization SINR for each trial. To his end, we consider the simulation of N_{try} general prediction schemes to obtain N_{try} copies of the metrics of interest, which needs to be compressed into a single value, for instance bt taking their mean with $\text{PER}^{(\tau)} = N_{\text{try}}^{-1}\sum_{n=1}^{N_{\text{try}}}\text{PER}^{(\tau)}[n]$, with $\text{PER}^{(\tau)}[n]$ being the predicted PER for the n^{th} trial.

This prediction method is denoted as Pseudo-Random Joint Demapping and Decoding (PRJD2) prediction, and in Figure 5.10, we have compared the prediction performance of the conventional JD2 method, with this new method with $N_{\text{try}} = 100$. In particular the exponential distribution (no degree of freedoms) is considered, and the Gamma distribution of shape 0.05, and scale $\sigma_{v_x^d}/\sqrt{0.05}$, is evaluated, which enables having a relatively smaller mean (i.e. bias). Transmission of QPSK symbols coded at rate 1/2 with the convolutional code of polynomial $[133, 165, 171]_8$, punctured by $[1, 0, 1; 1, 0, 1]$, over the Proakis C channel is considered, with $K = 512$ and $B = 1$. As this deterministic channel has spectral nulls, it is challenging for prediction algorithms to perform accurately with it, as the detector creates non negligible correlations between symbols. In Figure 5.10, the previously discussed JD2 scheme becomes less and less accurate as the number of turbo or self iterations increases, and proposed pseudo-random methods appears to compensate this issue to some extent, by making more pessimistic predictions according to the dispersion of the soft feedback reliability.

5.5 Application: Cooperative Broadcast in MANETs

In order to illustrate the potential impact of our work in a fundamental distributed cooperation technique, here we discuss cooperative broadcasting technique. Indeed this an extreme example of cooperation where the broadcasting capabilities of these receivers are used at its full potential, and because of this, this method has been analyzed with different levels of abstraction over the years.

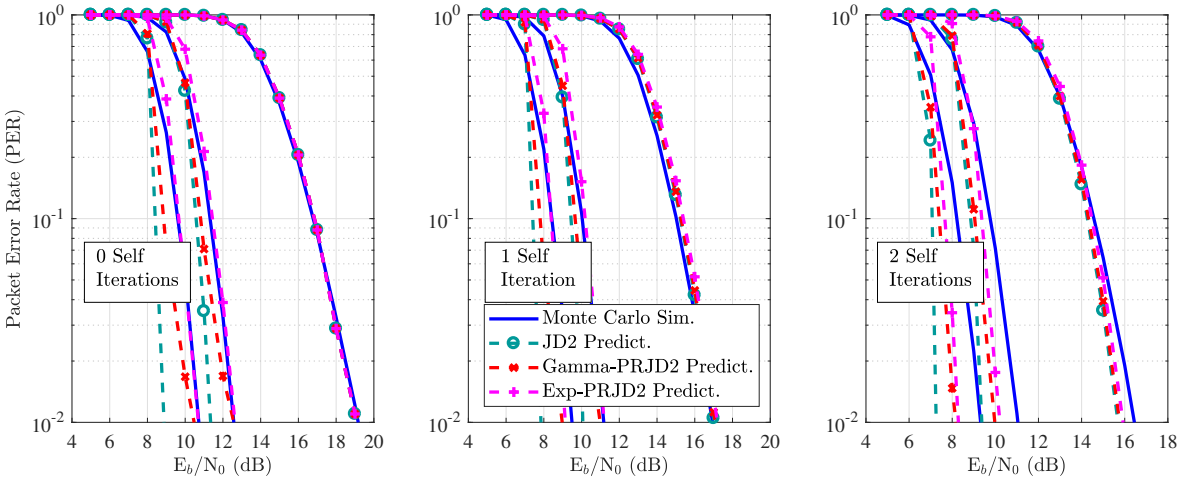


Figure 5.10: Comparison of the pseudo-random finite-length prediction PRJD2 scheme with the finite-length JD2 scheme, for 0 to 2 turbo-iterations.

5.5.1 Impact of Cooperative Broadcasting on the Physical Link

As previously discussed in the epilogue of the Chapter 1, cooperative broadcasting is a highly robust technique where simultaneous transmission of the same data by multiple transmitters generate an artificial multi-path channel at the receiver. The resulting cooperative diversity is supposed to the increase link robustness and also the range of the transmission.

Stochastic behavior of multi-hop networks with cooperative broadcast is evaluated with the assumption of infinite nodes with finite power per area in [SSM06]. Recent models for finite node densities [ÇGT13; HI12; RGW18] investigate on inter-node distance distributions and path-loss to evaluate range improvements brought by the cooperative broadcast. These works assume that transmitted signals coherently combine at destinations, ignoring selective channels caused by propagation delays, clock offsets and oscillator drifts.

In [WGV06], impact of propagation delays and delay dithering are studied for harvesting cooperative diversity as frequency diversity with a time-domain DFE. [HH15] considers multi-hop cooperative broadcasting without frame resynchronization at each hop, and it analyzes the evolution of time synchronization errors across hops. But this work does not consider the impact of path-loss, nor the equalization of the artificial channel.

The design of PHY layer receivers for handling cooperative broadcast detection has been addressed in [VG06; BTV07; WXY09]. A major concern common to these works is the mitigation of Carrier Frequency Offset (CFO) caused by clock synchronization issues. Nevertheless, these works either use time-domain serial DFE for single-carrier signalling [VG06; WXY09], or frequency-domain detection followed by serial DFE for multi-carrier signalling [BTV07]. In all cases, equalization complexity is at least quadratic in block length due to DFE filter computation, and large-delay spreads would further complicate the design.

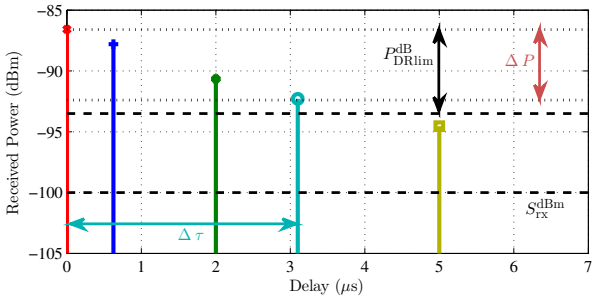


Figure 5.11: A cooperative channel profile with 5 users (each color is a user).

5.5.1.1 General Model

We consider a scenario where radios $1,\ldots,U$ cooperatively broadcast a data block to a single destination (with selective Decode & Forward (DF) relaying), and here we aim to show how the cooperative transmission is perceived by the receiver.

Let us assume that, following a network-wide synchronization procedure, each node has its internal clock shifted by a residual offset of t_u^{off} seconds with respect to an ideal global reference clock. Moreover, oscillator imperfections cause a frequency drift of f_u^{off} Hertz with respect to the carrier frequency f_c (see Appendix A). Considering these parameters, the channel $\mathrm{h}_u(t)$ between u^{th} node and the destination (see Equation (1.9)), and accounting for the transmit power, we have

$$\mathrm{h}_u(t) = \sqrt{\mathcal{E}_{\mathrm{rx},u}}e^{j2\pi t\phi_u}\sum_{l=1}^{L_{\mathrm{ss}}} a_{u,l}(t)\delta(t-\tau_{u,l}), \tag{5.26}$$

where the total delay from the node u, at the l^{th} tap is $\tau_{u,l} \triangleq \tau_l^{\mathrm{ss}} + \tau_u^{\mathrm{clk}} + \tau_u^{\mathrm{prop}}$, with the component due to clock offsets being $\tau_u^{\mathrm{clk}} \triangleq t_u^{\mathrm{off}} - t_0^{\mathrm{off}}$, and the carrier offset frequency being $\phi_u \triangleq f_u^{\mathrm{off}} - f_0^{\mathrm{off}}$. The received power from u^{th} user is $\mathcal{E}_{\mathrm{rx},u} \triangleq \mathcal{E}_{\mathrm{tx}}h_u^{\mathrm{PL}}$.

Then, by combining the U cooperating transmitters, the cooperative broadcast on the observed physical channel results in

$$\mathrm{h}(t) = \sum_{l=1}^{L} a_{l_u,l_l}(t)\delta(t-\tau_{l_u,l_l}), \tag{5.27}$$

where $l_u \triangleq \lceil (l)/L_{\mathrm{SS}} \rceil$, $l_l \triangleq ((l-1) \bmod L_{\mathrm{SS}}) + 1$, with $L = UL_{\mathrm{SS}}$ and

$$a_{l,u}(t) \triangleq 10^{(G^{\mathrm{rx}}_{\mathrm{dB}})/20}\sqrt{\mathcal{E}_{\mathrm{tx},u}g_u^{\mathrm{PL}}g_u^{\mathrm{sh}}(t-\tau_l^{ss})}a_{u,l}^{\mathrm{ss}}(t)e^{j2\pi t\phi_u}. \tag{5.28}$$

The cooperative transmission diversity presents itself in Equation (5.27) as supplementary frequency diversity, and, if paths are resolvable, it can to be harvested through equalization. However, equalization success depends on the frequency selectivity of the channel, which can

be assessed with the delay spread

$$\Delta\tau = \Delta\tau_{\mathrm{SS}} + \Delta\tau_{\mathrm{prp}} + \Delta\tau_{\mathrm{clk}}, \tag{5.29}$$

where the small-scale delay spread is $\Delta\tau_{\mathrm{SS}} \triangleq \max_l \tau_l^{\mathrm{ss}} - \min_l \tau_l^{\mathrm{ss}}$ and the delay spread component caused by large-scale propagation is $\Delta\tau_{\mathrm{prp}} \triangleq \max_u \tau_u^{\mathrm{prop}} - \min_u \tau_u^{\mathrm{prop}}$ and the one due to clock effects is $\Delta\tau_{\mathrm{clk}} \triangleq \max_u \tau_u^{\mathrm{clk}} - \min_u \tau_u^{\mathrm{clk}}$.

The ability to equalize the channel also strongly depends on the dynamic power range of the channel, i.e. expected value of power differences between taps, which is defined in dB as

$$\Delta P \triangleq \Delta P_{\mathrm{SS}} + \Delta P_{\mathrm{prp}}, \tag{5.30}$$

with $\Delta P_{\mathrm{prp}} \triangleq \max_u 10\log_{10}(\mathcal{E}_{\mathrm{rx},u}) - \min_u 10\log_{10}(\mathcal{E}_{\mathrm{rx},u})$, and $\Delta P_{\mathrm{SS}} \triangleq \max_l 10\log_{10}(\mathcal{E}_l^{\mathrm{ss}}) - \min_l 10\log_{10}(\mathcal{E}_l^{\mathrm{ss}})$.

As $\Delta\tau$ increases, and ΔP decreases, the Inter-Symbol Interference (ISI) caused by the channel become more severe. ΔP and $\Delta\tau$ above describe the frequency-selectivity of the cooperative channel with dependence on topology and radios, but some practical limitations at the receiver were omitted.

Indeed, strongly attenuated taps do not impact the frequency-selectivity, hence denoting the received power in dBm $\mathcal{E}_{\mathrm{rx},u}^{\mathrm{dBm}}$, only taps received from users such that

$$\mathcal{E}_{\mathrm{rx},u}^{\mathrm{dBm}} \geqslant S_{\mathrm{rx}}^{\mathrm{dBm}}, \tag{5.31}$$

will be relevant for computing ΔP and $\Delta\tau$. The receiver sensitivity in dBm is $S_{\mathrm{rx}}^{\mathrm{dBm}} \triangleq N_0 - 10\log_{10}(T_s) + L_{\mathrm{r}}$, where N_0 is the noise power spectral density at receiver antenna in dBm/Hz, T_s is the symbol period, L_{r} is a constant in dB, including effects of antenna gains, noise figure and the detection threshold at the receiver.

Moreover, in the presence of strong taps, smaller channel components lose their impact and can even become neglected in receiver channel estimation algorithms. To account for such issues, a constraint on the dynamic range is added

$$\Delta P \leqslant P_{\mathrm{DRlim}}^{\mathrm{dB}}, \tag{5.32}$$

for evaluating the channel selectivity.

Figure 5.11 illustrates these quantities on an instance of a cooperative broadcast channel with non-frequency-selective small-scale components. Alternatively, Figure 5.11 can be understood as the representation of the influence of the delay spread component $\Delta\tau_{\mathrm{prp}}$ caused by large-scale propagation. However this case is not restrictive, as the small-scale profile is independent, and its impact can be incorporated separately.

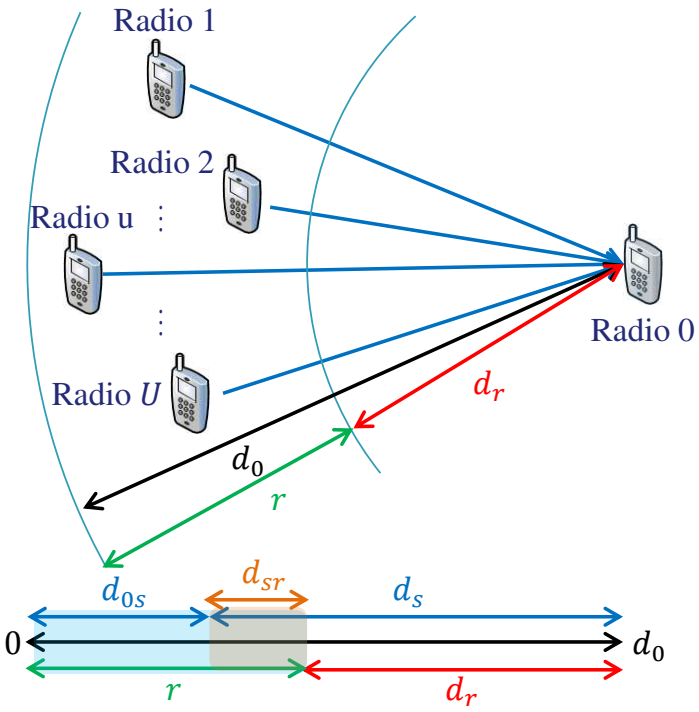


Figure 5.12: Topology model illustration.

5.5.1.2 Stochastic Model on Channel Selectivity

In this subsection, statistical characteristics of the cooperative channel in Equation (5.27), is assessed by assuming emitting node positions and radio imperfections are randomly distributed.

Emitting nodes $1, \ldots, U$ are uniformly distributed on an annulus, centered on the destination node $u = 0$, with an outer radius of d_0 meters and the width of the annulus is given by $r < d_0$, i.e. the inner radius is $d_r \triangleq d_0 - r$. This is equivalent to a one-dimensional model where the destination is located at d_0 and emitters are uniformly distributed on the segment $[0, r]$, shown in Figure 5.12. Moreover t_u^{off} is uniformly distributed on $[-\tau_{\text{clk}}/2, \tau_{\text{clk}}/2]$, with τ_{clk} setting the maximum absolute value of $\Delta\tau_{\text{clk}}$.

Here the PDFs of the $\Delta\tau$ and ΔP are exposed, without including small-scale effects. Detailed derivations are not given due to lack of space, but a sketch of proof is provided.

$\Delta\tau_{\text{prp}}$ and ΔP, tied to radio distance distributions, and the clock offset component $\Delta\tau_{\text{clk}}$ can be modeled separately, using change of variables on uniform and general Beta distributed random variables. However this approach does not consider the practical constraints in Equations (5.31)-(5.32).

The dynamic range constraint (Equation (5.32)) is relevant on short distances, which is out of focus in this paper, hence only the impact of receiver sensitivity in Equation (5.31) is

considered. The latter is equivalent to ignoring radios farther than d_s to radio 0, with

$$d_s = d_{\text{ref}} 10^{\frac{\mathcal{E}_{\text{tx}} - S_{\text{rx}} - P_{\text{ref}} - 20\log_{10}(f_c)}{10\alpha}}. \tag{5.33}$$

Thus, $\Delta\tau_{\text{prp}}$ and ΔP_{prp} are defined with respect to constrained minimum and maximum distances $d_m \triangleq \min_{u,d_u<d_s} d_u$ and $d_M \triangleq \max_{u,d_u<d_s} d_u$. When $d_s < d_0$, $\Delta\tau_{\text{prp}}$ and ΔP_{prp} are non-zero iff at least two radios are in $[d_{0s}, r]$, $d_{0s} = d_0 - d_s$. The probability of having less than two users in this segment is $p_{0s} = d_{0s}^{U-1}(r + (U-1)d_{sr})/r^U$, with $d_{sr} = d_s - d_r$.

Finally, $\Delta\tau_{\text{prp}}$ and $\Delta\tau_{\text{clk}}$ are combined to model the PDF of the total delay spread $\Delta\tau$. To this end, the method used in [PGT98] is generalized here with truncated general Beta random variables to obtain analytical expression of $p(\Delta\tau)$.

Following computations, the distribution of ΔP is given by

$$p(\Delta P) = \frac{\log 10}{10\alpha} \frac{(U-1)}{r^U} \frac{d_0^U - d_r^U 10^{\frac{U\Delta P}{10\alpha}}}{10^{\frac{\Delta P}{10\alpha}} \left(1 - 10^{\frac{-\Delta P}{10\alpha}}\right)^{2-U}}, \tag{5.34}$$

for $0 < \Delta P \leqslant 10\alpha \log_{10}(d_0/d_r)$, when $d_s \geqslant d_0$, and when $d_r < d_s < d_0$, ΔP follows

$$\begin{aligned} p(\Delta P) = p_{0s}\delta(0) + \frac{\log 10}{10\alpha} \frac{10^{\frac{\Delta P}{10\alpha}}}{\left(10^{\frac{\Delta P}{10\alpha}} - 1\right)^2} \frac{1}{r^U} \times \\ [(U-1)\varphi(\Delta P, U) - U d_{0s} \varphi(\Delta P, U-1)], \end{aligned} \tag{5.35}$$

for $0 < \Delta P \leqslant 10\alpha \log_{10}(d_s/d_r)$, with

$$\varphi(x,u) = \left(d_0 - d_s 10^{\frac{-x}{10\alpha}}\right)^u - \left(r - d_s + d_r 10^{\frac{x}{10\alpha}}\right)^u.$$

The analytical expression of $\Delta\tau$'s PDF is given in Equations (5.36)-(5.37), on the next page. The small-scale effects can be incorporated by translating these PDFs by ΔP_{SS} and by $\Delta\tau_{\text{SS}}$.

5.5.1.3 Numerical results on frequency-selectivity

In this section the statistical model above is used to evaluate the behavior of a subset of a MANET with $U = 5$ nodes and $r = 4$ km. We use typical wireless radio parameters $\mathcal{E}_{\text{tx}}^{\text{dBm}} = 45.5$ dBm, $S_{\text{rx}}^{\text{dBm}} = -100$ dBm, within the UHF band with $f_c = 400$ MHz, using a symbol period of $T_s = 1$ μs. Path-loss parameters are in part based on ITU-R P.1546-1 recommendations with $P_{\text{ref}} = -60$ dB, $d_{\text{ref}} = 1$ km.

The previously derived characteristic PDFs of the composite channel are plotted with the proposed analytical model and with Monte Carlo simulation results in the Figure 5.13, which attests to the accuracy the proposed model in Equations (5.36)-(5.37). Moreover, using these PDFs, the mean value, the 5% and 95% quantiles of the delay spread and the dynamic range are evaluated for varying d_0 and path-loss exponent α, without any clock imperfection. The

For $d_s > d_0$, the delay spread follows

$$p(\Delta\tau) = \begin{cases} \dfrac{\Delta\tau^{2U-3}(\tau_m-\Delta\tau)B(U-1,U-1)}{\tau_m^U\tau_M^{U-1}B(U-1,2)^2}F_1\left(\dfrac{\Delta\tau}{\Delta\tau-\tau_m},\dfrac{\Delta\tau}{\tau_M};U-1;-1,-1:2U-2\right) \\ \qquad\qquad \text{for } 0 \leqslant \Delta\tau \leqslant \tau_m, \\ \dfrac{(\Delta\tau-\tau_m)^{U-2}(\tau_t-\Delta\tau)}{\tau_M^U B(U-1,2)}F_1\left(\dfrac{\tau_m}{\tau_m-\Delta\tau},\dfrac{\tau_m}{\tau_t-\Delta\tau};2;2-U,-1;U+1\right) \\ \qquad\qquad \text{for } \tau_m < \Delta\tau \leqslant \tau_M, \\ \dfrac{(\tau_t-\Delta\tau)^3(\Delta\tau-\tau_m)^{U-2}}{6\tau_m^2\tau_M^U B(U-1,2)^2}F_1\left(\dfrac{\tau_t-\Delta\tau}{\tau_m},\dfrac{\Delta\tau-\tau_t}{\Delta\tau-\tau_m};2;2-U,2-U;4\right) \\ \qquad\qquad \text{for } \tau_M < \Delta\tau \leqslant \tau_t, \end{cases} \tag{5.36}$$

with $\tau_m = \min(\tau_r, \tau_{\text{clk}})$, $\tau_M = \max(\tau_r, \tau_{\text{clk}})$, $\tau_r = r/c$, $\tau_t = \tau_m + \tau_M$, and $F_1(x, y; a; b_1, b_2; c)$ is the Appell hypergeometric series of the first kind, and $B(\alpha, \beta)$ is the beta function, and for $d_r < d_s < d_0$, $p(\Delta\tau) = p_{0s}\delta(0) + p'(\Delta\tau)$ with

$$p'(\Delta\tau) = \begin{cases} \dfrac{\tau_{0s}^{U-2}\tau_{sr}\Delta\tau^{U-1}}{\tau_r^U\tau_{\text{clk}}^U B(U-1,2)}\left[\begin{array}{l}\Delta\tau F_1\left(\dfrac{-\Delta\tau}{\tau_{0s}},\dfrac{\Delta\tau}{\tau_{sr}};2;2-U,-1:U+1\right)+ \\ U(\tau_{\text{clk}}-\Delta\tau)F_1\left(\dfrac{-\Delta\tau}{\tau_{0s}},\dfrac{\Delta\tau}{\tau_{sr}};1;2-U,-1:U\right)\end{array}\right] \\ \qquad\qquad \text{for } 0 \leqslant \Delta\tau \leqslant \min(\tau_{\text{clk}}, \tau_{sr}) \\ \dfrac{(\Delta\tau-\tau_{ms})^{U-2}(\tau_{ts}-\Delta\tau)}{\tau_r^U B(U-1,2)}F_1\left(\dfrac{\tau_{\text{clk}}}{\tau_{ms}-\Delta\tau},\dfrac{\tau_{\text{clk}}}{\tau_{ts}-\Delta\tau};2;2-U,-1;U+1\right) \\ \qquad\qquad \text{for } \tau_{\text{clk}} < \Delta\tau \leqslant \tau_{sr} \\ \dfrac{\tau_{sr}^2(\Delta\tau-\tau_{ts})^{U-2}}{6\tau_r^2\tau_{\text{clk}}^U B(U-1,2)^2}\left[\begin{array}{l}\tau_{sr}F_1\left(\dfrac{\tau_{sr}}{\tau_{sr}-\Delta\tau},\dfrac{\tau_{sr}}{\tau_r};2;2-U,2-U:4\right)+ \\ 3(\tau_{\text{clk}}-\Delta\tau)F_1\left(\dfrac{\tau_{sr}}{\tau_{sr}-\Delta\tau},\dfrac{\tau_{sr}}{\tau_r};2;2-U,2-U:3\right)\end{array}\right] \\ \qquad\qquad \text{for } \tau_{sr} < \Delta\tau \leqslant \tau_{\text{clk}} \\ \dfrac{(\tau_{ts}-\Delta\tau)^3(\Delta\tau-t_{ms})^{U-2}}{6\tau_{\text{clk}}^2\tau_r^U B(U-1,2)^2}F_1\left(\dfrac{\tau_{ts}-\Delta\tau}{\tau_{\text{clk}}},\dfrac{\Delta\tau-\tau_{ts}}{\Delta\tau-\tau_{ms}};2;2-U,2-U;4\right) \\ \qquad\qquad \text{for } \max(\tau_{\text{clk}}, \tau_{sr}) < \Delta\tau \leqslant \tau_{ts} \end{cases} \tag{5.37}$$

with $\tau_{ts} = \tau_{\text{clk}} + \tau_{sr}$, $\tau_{ms} = \tau_{\text{clk}} - \tau_{0s}$, $\tau_{sr} = d_{sr}/c$, $\tau_{0s} = d_{0s}/c$. For $d_s < d_r$, the delay spread is zero, i.e. $p(\Delta\tau) = \delta(0)$.

---

results are plotted in Figure 5.14-(a) with solid, dashed and dotted lines, respectively for the mean, the 95% and 5% quantiles. Analytical predictions are illustrated with solid lines and Monte-Carlo simulation results with respectively cross, upward and downward triangle markers. Furthermore, for $\alpha = 4$, the impact of the clock offsets $t_u^{\text{off}} \in [-\tau_{\text{clk}}/2, \tau_{\text{clk}}/2]$ on the mean value of the delay spread is shown in Fig. 5.14-(b).

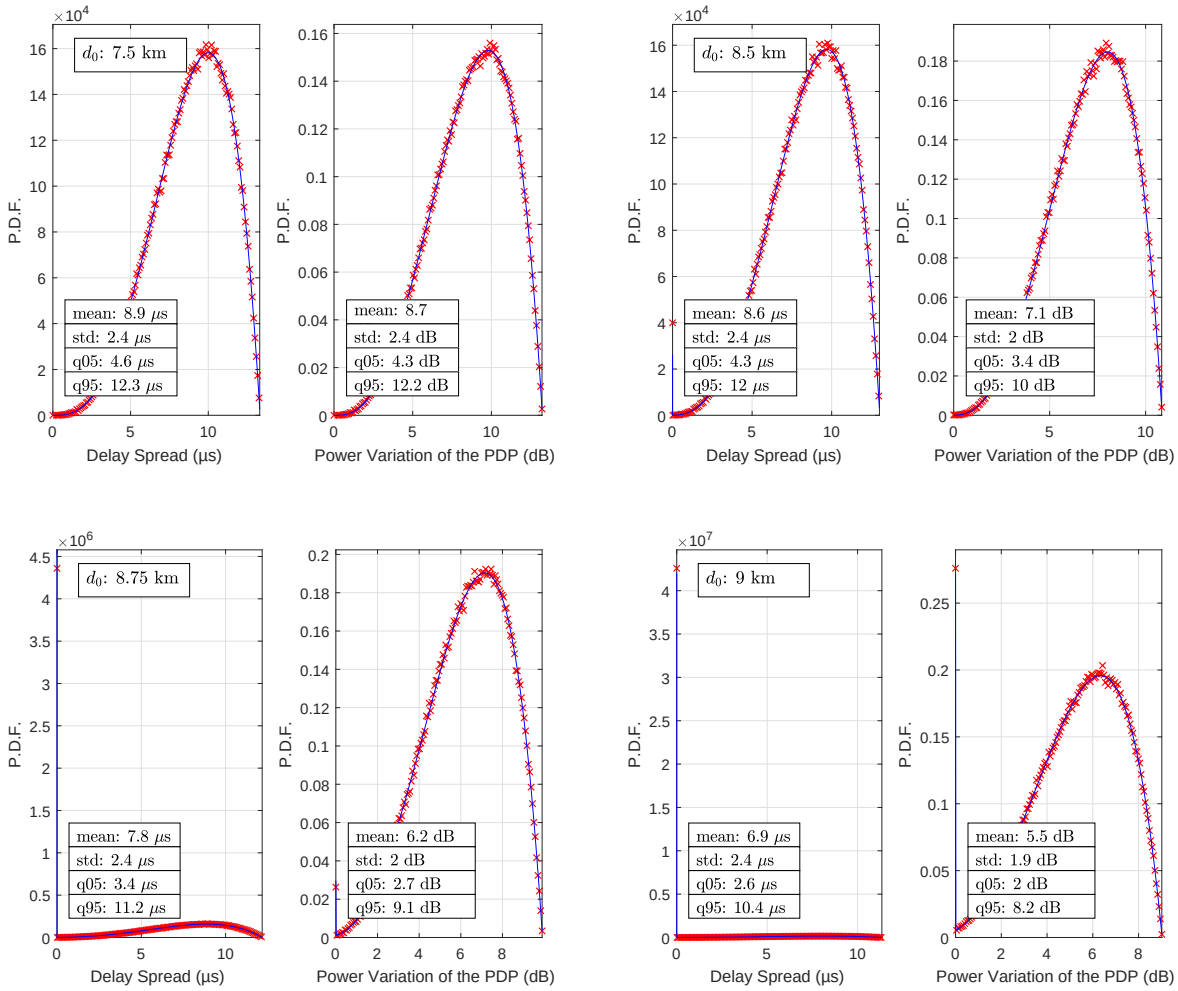
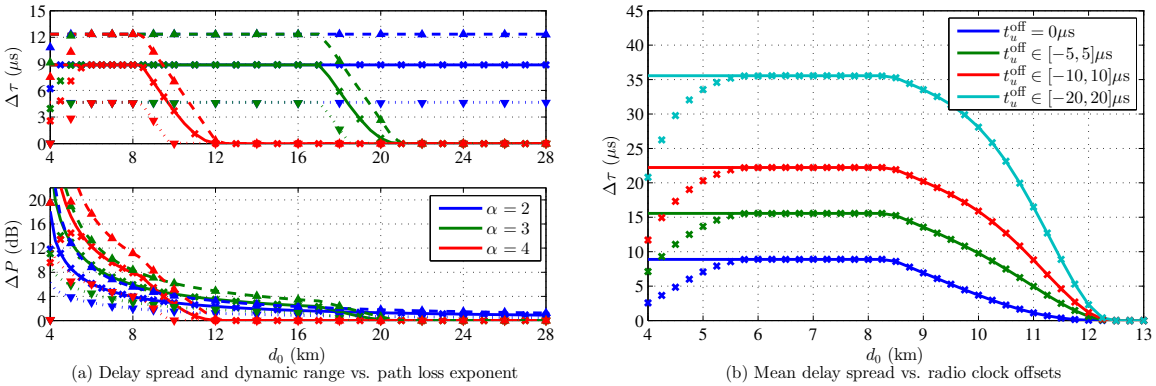


Figure 5.13: Simulated and predicted PDFs of the delay spread and the dynamic range of the composite channel for $\alpha = 4$ and $\tau_{\text{clk}} = 0$ μs (lines: predicted, markers: experimental).



(a) Delay spread and dynamic range vs. path loss exponent

(b) Mean delay spread vs. radio clock offsets

Figure 5.14: Mean values and quantiles of the delay spread and the dynamic range versus destination distance d_0 (lines: predicted, markers: experimental).

Proposed delay spread prediction model is accurate for medium to high distances, but experimental and predicted data diverge for small d_0, due to neglected constraint in Equation (5.32). Indeed, at low distances, the clipping effect of this constraint is seen on ΔP, as the 95% quantiles saturate near $P^{\text{dB}}_{\text{DRlim}} = 20$ dB. In medium distances, the delay spread reaches its topological maximum when it is neither constrained by Equation (5.31) nor by Equation (5.32), and reaches the mean of the delay spread distribution, given by $\Delta\tau = (\tau_{\text{clk}} + \tau_r)(U-1)/(U+1)$, which yields $\Delta\tau = 8.89$ μs for $\tau_{\text{clk}} = 0$ μs. Finally at high distances, the delay spread decreases due to the constraint of Equation (5.31), which allows to neglect paths with received power below receiver sensitivity, i.e. without a significant impact on detection performance.

Although most cooperative broadcast analysis carried out in the literature are based on flat-fading assumptions, results above indicate that frequency selectivity caused by such transmissions can be severe, as $\Delta\tau$ increases and ΔP decreases. In particular, conclusions drawn on range-extension capabilities are likely far from reality, for medium to high data rate applications, as severe ISI is present.

In the following, we discuss low-complexity detection of cooperative broadcast transmissions in MANETs, with frequency domain equalization and off-the-shelf error correction codes.

5.5.1.4 Impact on PHY layer with FD SILE-EPIC

Considering the numerical results above indicating large delay spreads in Figure 5.14, traditional time-domain strategies for mitigating ISI can have excessive computational costs [WGV06; VG06]. Usually in the context of large delays spreads, frequency domain equalization is preferable, and thus we propose to investigate apply the FD EP-based receiver in Chapter 4 strategy in MANETs for cooperative broadcasting.

Considering potential oscillator drifts of cooperating nodes, caused by clock synchronization issues, an encoding strategy across multiple short data blocks is needed, for improving the equalizer's robustness against time variations of the channel.

In the following, we consider the transmission by $U = 5$ relays, with $r = 4$ km, of $B = 3$ blocks of $K = 128$ with 8-PSK and 16-APSK constellations, using root-raised-cosine shaping filters with a roll-off of 0.35, $T_s = 1$ μs, and using an off-the-shelf 3GPP LTE channel coding and rate matching strategy. Indeed, thanks to the use of LTE turbo-code, and its rate-matching, we can also investigate the impact of code rate on the detection performance. The FD SILE-EPIC receiver exponential smoothing with a damping factor of 0.33 (see Equation (4.40)). The use of iterative detection and decoding is avoided at this stage, to solely focus on the benefits of self-iterations, with a widespread channel decoder. In Figure 5.15, the operating points of this PHY in a 5-tap Rayleigh fading channel is provided, for 8-PSK at rates 2/3 and 3/4, and then for 16-APSK at rates 2/3, 3/4 and 5/6, Indeed, a single self-iteration is sufficient to bring between 2 to 3 dBs improvement in energy-efficiency. This change can also be considered as an improvement of the considered modulation and coding scheme, by switching to higher rate operating modes, and benefiting from a rate improvement

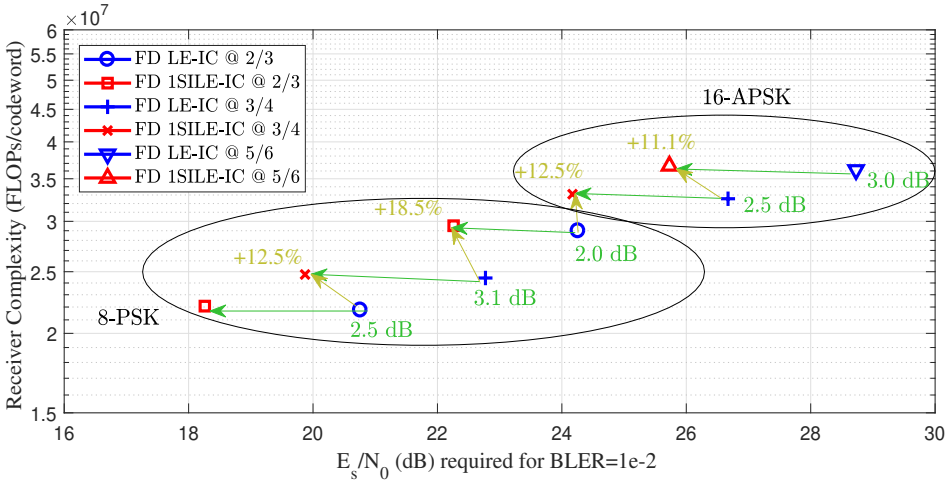


Figure 5.15: FD SILE-EPIC with a turbo-coder at different code rates. Orange arrows show gain in spectral efficiency whereas green arrows show gains in energy efficiency.

of about 15%.

In the following, we focus on some test-bench channels which can affect the PHY layer link robustness, by using the results observed in Figure 5.14.

Ground-to-ground tactical communications scenario First, a ground-to-ground tactical MANET is considered with small-scale channel being a single-tap Rayleigh variable, and the path-loss exponent being $\alpha = 4$. For a destination located at $d_0 = 8$ km, the considered channel power profile $[0, -3.4, -6.2, -8.6, -10.8]$ in dB corresponds to the average topology yielding the 95%-quantile of delay spread ($\Delta\tau_{\text{prop}} = 12$ μs), i.e. equally 900m-spaced nodes. This exponentially decreasing channel is fairly easy to equalize, but radio clock offsets t_u^{off} increase the channel selectivity as observed in Figure 5.14-b, and they can cause a loss of frequency diversity, if delayed signals become un-resolvable.

In Figure 5.16, the average PER performance of FD LE and FD SILE-EP, with $R_c = 2/3$, is displayed in solid lines, by averaging over 150 realizations of uniformly distributed clock offsets, between $[-\tau_{\text{clk}}/2, \tau_{\text{clk}}/2]$, as in Figure 5.14-(b). Some delay realizations which cause independent taps to become unresolvable cause significant diversity loss, but this only slightly degrades the *average* PER.

Nevertheless, for considering the impact of clock offset realizations on the robustness of the average PER, the outage probability $\mathbb{P}[\text{PER} > 10^{-2}]$ is evaluated and it is displayed in dashed lines. For $\tau_{\text{clk}} = 0$ μs, the outage only occurs when the average PER is higher than 10^{-2}, which yields a SNR threshold-like behavior. It is seen that $\tau_{\text{clk}} = 10$ μs cause a significant loss of diversity, but this loss is smaller for $\tau_{\text{clk}} = 20$ μs. This behavior is natural, as realizations with unresolvable taps become more unlikely as τ_{clk} increases too much.

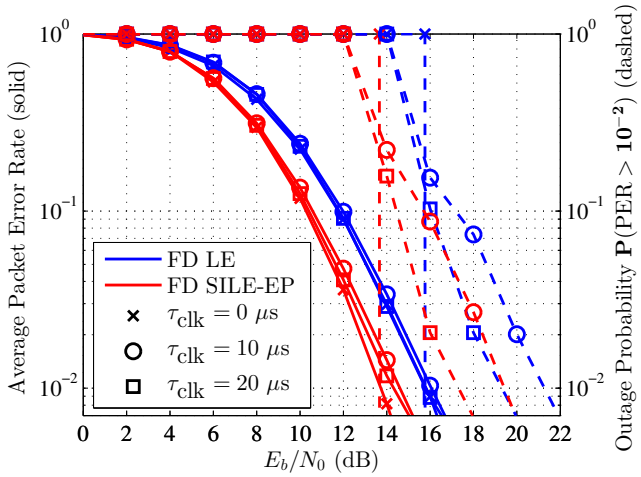


Figure 5.16: Impact of clock offset in the tactical scenario with $R_c = 2/3$.

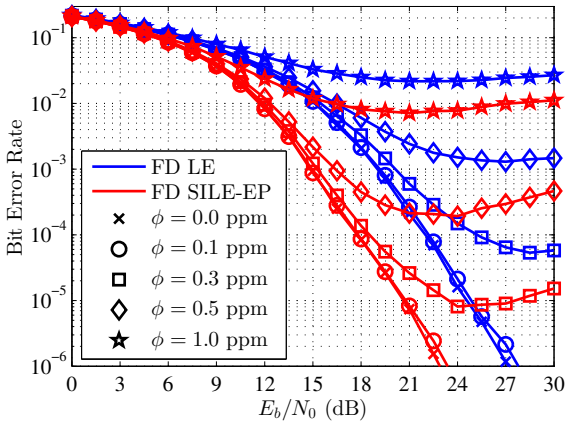


Figure 5.17: Impact of CFO on the uncoded BER in the UAV scenario.

Air-to-air UAV communications scenario We now consider an intra-UAV communication scenario where the small-scale channel is single-tap with a Rice factor of 10 dB [GW15b] and the path-loss exponent is $\alpha = 2$. As UAVs need to be equipped with precise localization systems (e.g. Global Positioning System - GPS), clock offset issues can be greatly reduced and enable good network-wide synchronization. Thus CFOs can realistically be controlled to remain less than a ppm. However, challenging receiving conditions may arise when non-negligible relays signals have significantly different CFOs, thus increasing the time-selectivity of the received signal. Considering a scenario with $d_{1:5} = [4.0, 2.5, 2.2, 0.7, 0.4]$ km, corresponding to a realization of the 95%-quantile of $\Delta\tau_{\text{prop}}$, and $d_0 = 20$ km which yields the near-uniform power profile $[0, -0.78, -0.93, -1.6, -1.8]$ in dB, which is rich in diversity but difficult to equalize. For testing, we assume a worst case situation in which close nodes have very different CFOs, i.e. $\phi_{1:5} = [\phi, -\phi, \phi, -\phi, \phi]$.

In Figure 5.17, uncoded BER performance of considered equalizers is plotted. Increase in CFO is shown to create an error floor, which is then enhanced as the signal-to-noise ratio

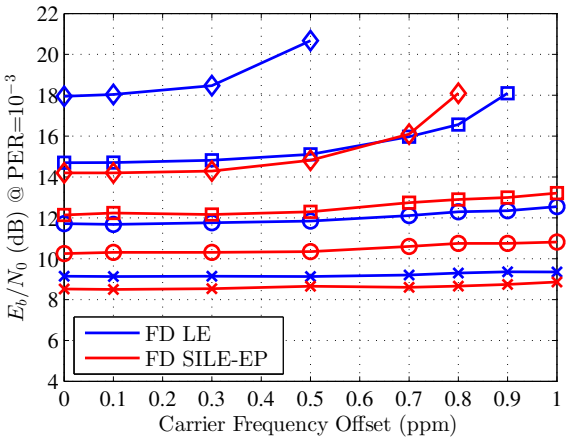


Figure 5.18: Impact of CFO on the coded performance of the UAV scenario. $R_c = 1/3, 1/2, 2/3, 5/6$ are respectively denoted by $\times$, $\circ$, $\square$, $\diamond$.

(SNR) increases, due to the channel estimate mismatch at the equalizer. In practice, the error-floor can be kept at its minimum, by accounting for the channel mismatch errors within the equalizer filters. It is seen that without CFO ($\phi = 0$ ppm), the FD SILE-EP brings around 4.2 dB gain over FD LE at BER=10^{-3}, and regardless of the CFO, FD SILE-EP has lower error floors and notable SNR gains. The $\phi = 1$ ppm case might be unrealistic for UAVs, but it allows assessing the limits of considered receiver.

Considering the same scenario with LTE channel coding for $R_c = [1/3, 1/2, 2/3, 5/6]$, the trade-off between higher throughput or a more powerful code is assessed in Fig. 5.18. The E_b/N_0 required to decode with PER=10^{-3} is plotted as a function of the CFO ϕ, for these code rates and the considered equalizers. For $\phi > 0.8$ ppm, both FDE cannot decode $R_c = 5/6$, and FD LE can no longer decode $R_c = 2/3$ for $\phi > 0.9$ ppm. Strong codes manage to cope with typical values of CFO, and FD SILE-EP bring small improvements, but with higher code rates the performance is severely degraded, and the benefits of using an iterative receiver, such as FD SILE-EP, is more significant. In particular, FD SILE-EP considerably improves spectral efficiency, as it decodes at $R_c = 5/6$ with better energy efficiency than FD LE operating at $R_c = 2/3$, up to $\phi = 0.7$ ppm.

Discussion An analytical model is provided on the distribution of the delay spread and the dynamic range of cooperative broadcast channels that appear in MANETs with imperfect radio clock synchronization. This provides a means to assess the frequency selectivity of the artificial channels generated by such transmissions, and to design the PHY layer accordingly.

We have evaluated the performance of frequency-domain equalization for handling such transmissions in some scenarios of interest. Numerical results showed how radio imperfections could impact the link quality and that modern iterative frequency domain receivers could become viable solutions with significant advantages over conventional FDE, especially when dealing with high data rate transmissions.

Our results assume the use of appropriately-sized cyclic prefixes, and this could lead to some loss of efficiency in certain scenarios. In such cases, the use of an iterative overlap FDE could be preferred (see Section 4.5.3 of the previous chapter).

Previous works on the analysis of cooperative broadcasting in MANETs often ignore the impact of the underlying artificial frequency-selective channel. This can potentially overestimate the performance prediction for practical applications if realistic low-cost radios are to be used. This contribution aims to regain awareness in these issues, and future works will focus on assessing the impact of these PHY layer considerations on higher layer quality-of-service metrics of MANETs.

5.6 Conclusion

With this final chapter, we address the problem of cooperative protocol design for MANETs, by putting in place a system-level simulation framework with a realistic PHY model. In this regard, a considerable part of this chapter is dedicated to identifying successful approaches to link abstraction and extending them for our proposed FD SILE-EPIC receiver.

This has been achieved by first establishing an asymptotic evolution analysis model for this receiver, by combining notions from binary EXIT function techniques and symbol-wise state-evolution methodology, used by AMP-based algorithms. Then two finite-length prediction mechanisms are proposed, one of them being based on the symbolwise abstraction method of [VB+10], where demapping and decoding operations are jointly modelled. The second method we proposed further extends these approaches by analysing the dispersion of the covariance of the soft feedback, which is a method loosely inspired from finite-length coding literature.

In the final part of this chapter, a fundamental distributed cooperation technique, cooperative broadcast, has been on the spotlight, first, in order to show the physical layer challenges it artificially creates, and then to show that the proposed receiver with an off-the-shelf turbo-code can achieve significant improvements in this context.

Conclusion and Perspectives

In this thesis, the design challenges for future MANETs with distributed cooperative protocols is studied, mainly from a PHY-layer receiver design point of view. Communication protocols in MANETs inherently require the use of multi-hop relaying functionalities for coping with mobility and the lack of infrastructure, at the cost of increased latency and jitter. The use of cooperation at the PHY layer can become a driving force for increasing the robustness or the spectral efficiency of such networks.

In order to better understand the underlying design challenges and roadblocks for enabling cooperative relaying, we have provided an overview on major developments involving cooperative transmissions in wireless networks, and in particular MANETs. Firstly, it is noted that most efficient schemes exploit non-orthogonal access techniques, and do not shy away from interference, as the PHY layer is assumed to use state-of-the-art coding strategies and advanced interference mitigation algorithms at the receiver. Secondly, as most of these protocols come with a certain amount of signalling burden, there is a lack of guidance on how to *fairly* assess these techniques, and compare them among each other. Without accounting for overhead penalties, the use of sub-optimal PHY techniques or the use of unrealistic channel models, it is not possible to determine if a practical implementation of a considered cooperative protocol is worth the effort.

Considering the aforementioned challenges for enabling the design of cooperative MANETs, the main problematic of this thesis is tackled by addressing the following objectives:

- identify from the state-of-the-art of PHY receiver design techniques, the strategies that are likely to be adopted for near-future implementation in real-world platforms, by considering the BICM scheme with single-carrier transmissions,
- ensure that selected receivers' behavior can be predicted with link abstraction techniques, for enabling system-level simulators to assess the receiver's impact within a realistic implementation of the cooperative protocol.

In order to address the first objective, we have carried out a study that aims to shed light into various categories of emerging signal processing algorithms, by considering a generic linear model and an iterative BICM receiver structure. Then various Soft Input Soft Output (SISO) detectors are designed using approximate Bayesian inference techniques and algorithms based on AMP for this generic "toy" example, in order to identify their strengths and weaknesses. To this end we have reviewed over 25 years of developments in algorithms related to turbo detection, and we have focused our attention to the emerging approximate statistical estimation methods such as BP, EP, MF, PDA and (·/G/V/O)AMP algorithms. In particular, we mainly focus on techniques that enabled designing double-loop turbo detectors, and in order to comply with the second goal of this thesis, we have sought to understand asymptotic limits and predictability of these methods, with a unified view. The motivation behind

the investigation of double-loop receivers stems from the fact that SISO decoder can have excessive complexity and latency, and we have hoped to be able to replace iterations with the decoder by a simpler demodulation process. We have also investigated emerging deep learning techniques for PHY design, with a particular forcus on the method of deep unfolding which enables optimizing iterative algorithms by assimilating them to deep networks.

The main conclusions drawn from this study is that SISO detectors with inner iterations is a promising category of algorithms, and in particular those designed with EP message passing, or the related VAMP or OAMP techniques possess the property of being predictable.

Considering this, in order to specifically address the ISI mitigation problem at hand for single-carrier transmissions, we have started our investigations for designing an equalizer. First, the state-of-the-art techniques on turbo-equalization are reviewed, with a focus on Time Domain (TD) Finite Impulse Response (FIR) structure, and then to address their shortcomings, novel design techniques for FIR DFE turbo-equalizers were proposed by incorporating the use of Expectation Propagation (EP)-based, and APP-based feedback. Thanks to a novel strategy for efficient matrix inverse update computation, the performance-complexity trade-off analysis on these receiver strongly favour FIR DFE receivers with EP-based soft feedback, thanks to their impressive performance at high spectral efficiency operation points. We have also explored the design of self-iterated FIR DFE to further push the achievable rate bounds close to the capacity. These studies were carried out with exact FIR structures which possess dynamic filters, and in order to reduce computational costs, we have addressed the problem of turbo DFE design with static filters. While this is a challenging non-linear design problem, where the optimal filter depends on the outputs of the equalizer, through semi-analytical prediction, we have proposed an effective method for designing filters with close-to optimal performance.

Although the proposed FIR DFE structures brought significant performance improvement, their implementability in real-world platforms was doubtful, due to the serial nature of the decision feedback causing a significant amount of latency with the demodulation step. To alleviate this we have started investigations on Frequency Domain (FD) structures with decisions being taken over the whole data block. To this end, we proposed a framework based on using scalar EP message passing for designing doubly-iterative turbo equalizers with FD signal processing. Similarly to the previous chapter, the use of EP-based self iterations bring significant performance improvements at the high data rate applications. These receivers are shown to achieve significant improvements over traditional FD structures, while achieving nearly the same performance as its exact TD counterpart. Strengths and weaknesses of these algorithms have been theoretically analyzed based on finite-length Monte Carlo simulations and asymptotic EXIT analysis. This most likely is the technical contribution with the highest potential of having a significant impact, due to this receiver having a simple and low-complexity structure, and due to the widespread use of FD equalizers in low-cost radios. The optimization of this receiver's parameters through an original use of deep unfolding and learning techniques has been explored, and promising results are obtained through numerical analysis. Finally, the proposed framework has been extended to handle complex SC problems, such as SC-FDMA, time-varying channel equalization with overlap-and-save, and

multi-user MIMO spatial multiplexing. The impact of channel estimation is also evaluated for this receiver.

Finally, having found a suitable candidate for PHY receiver of future MANETs, the second objective set for this thesis is addressed, by putting in place a system-level simulation framework with a realistic PHY model. In this regard, novel link abstraction approaches are proposed for the self-iterated FD receiver, through the use of asymptotic analysis techniques, enhanced by some heuristics for increasing finite-length prediction accuracy. In particular, a method is proposed for tracking the dispersion of the soft feedback and to perform pseudo-random generation of soft reliability errors, in order to perform semi-analytical prediction without any calibration. In order to gather attention to the importance of having such simulation tools, a fundamental distributed cooperation technique, called cooperative broadcast, is explored. This protocol is seen to create significant amount of ISI, along with time-selectivity at the receiver, and with a simple, off-the-shelf PHY we have illustrated potential improvements that state-of-the-art receivers can bring to cooperative MANETs.

Perspectives

Further improvements on FIRs

In the works of *Jeong et. al.* [JM13; Jeo11], concatenation of hard DFE FIR was carried out in order to improve detection performance significantly. In particular concatenation of two DFEs with reversed time axis yields the bi-directional DFE structure. Extension of these methods with EP and self-iterations should be explored.

These structures can also be improved through the use of FS equalization [Pet+19], where challenges on complexity and channel estimation have to be addressed.

Kalman smoothers

We have discussed the links between TD block receivers and Kalman smoothers, and various types of Kalman smoothers were proposed for EP in [SZW+15; SMF+17b]. In particular the latter structure appears to yields the same performance as block filter-banks, hence it would be interesting to perform a joint complexity-performance analysis and seek potential improvements for the equalization of time-varying channels.

FD joint channel estimation and detection

Although most part of this thesis is focused on detection, message passing techniques are also widely using for channel estimation [Han+18]. Considering the practicality of the self-iterated FD equalizer, it would be interesting to extend it in order to also refine channel

estimates through self-iterations and approximate inference.

Deep unfolding

In this thesis the use of deep unfolding was justified through the observation that EP-like VAMP algorithm is already optimal, when perfect CSI is available, for a limited set of structures and for Bernoulli-Gaussian signals [BSR17]. It would be interesting to re-iterate this experiment with signals from PSK and QAM constellations in order to ensure that for digital communications problems, this type of structures do not get improved by learning.

Next, the incorporation of loops and learnable parameters, other than damping, can be explored, to attempt to further optimize unrolled FD detection networks. The benefits of deep unfolding should also be explored for channel estimation.

EP for superposition coding and NOMA

In this thesis, BICM systems were under spotlight as they enable using conventional turbo-detection strategies between a SISO detector and a SISO decoder. However for emerging NOMA techniques, superposition coding techniques are being explored. Through the flexibility that EP brings for casting messages into practical family of PDFs one could provide low-complexity but effective techniques for such modulations.

Evaluate the system-level gains for using advanced receivers in MANETs

A natural continuation of the last chapter of this thesis is the use of the proposed cooperative simulation framework and the performance prediction scheme for evaluating system-level gains of using self-iterated FDE in MANETs. These evaluations should consider the characteristics of MANET channels, by accounting for the impact of shadowing and mobility.

Multiple-multicast/broadcast protocols for MANETs

Finally, cooperative protocols explicitly mentioned in this thesis either consisted in unicast transmissions through relay selection, or broadcasting/multicasting through cooperative broadcast. PHY layer abstraction methods for emerging protocols based on NOMA or Physical Layer Network Coding [Nav18; Liu+19a] should be explored, alongside realistic PHY design. Such non-orthogonal protocols could enable combining multiple payloads for PHY-layer transmissions and create simultaneous cooperative links.

Conclusion et Perspectives (Français)

Dans cette thèse on étudie les défis de conception pour des MANETs futurs ayant des protocoles coopératifs distribués, principalement de point de vue des problématiques récepteurs en couche PHY. Les protocoles de communications pour les MANETs nécessitent naturellement l'utilisation de fonctionnalités relayage multi-bonds afin d'affronter la mobilité et le manque d'infrastructure, au prix d'une latence et une gigue plus élevées. La coopération au niveau de la couche PHY peut alors devenir un point clé pour augmenter la robustesse ou l'efficacité spectrale de tels réseaux.

Afin de mieux comprendre les défis techniques et les obstacles pour mettre en oeuvre le relayage coopératif, nous avons fourni une synthèse sur les développements majeurs qui impliquent des transmissions coopératives dans des réseaux sans-fil, et en particulier, les MANETs. D'abord, on remarque que les schémas les plus efficaces exploitent des techniques d'accès non-orthogonal, et ne cherchent pas à éviter l'interférence, puisque la couche PHY est supposé d'être conçue avec des techniques de codage de pointe et avec des algorithmes de mitigation d'interférence avancés. Deuxièmement, comme la plupart de ces protocoles nécessitent de la signalisation, il y a un certain manque de méthodologie afin de pouvoir évaluer l'apport de ces techniques *équitablement*, et pour pouvoir les comparer entre eux. Sans prendre en compte les pénalités de signalisation, ou l'utilisation de techniques couche PHY sous-optimaux avec des modèles de canal réalistes, il n'est pas possible de déterminer si l'implémentation pratique d'un protocole coopératif vaut la peine d'être envisagé.

En vue des défis mentionnés au-dessus pour rendre la conception des MANETs coopératives plus favorable, les problèmes énoncés avec cette thèse sont abordés en poursuivant les deux objectives suivantes:

- identifier dans l'état de l'art sur les techniques de conception de récepteur PHY les stratégies qui semblent être les plus vraisemblables d'être choisis pour implémentation sur une plateforme réel dans le futur proche, dans le cadre des transmissions monoporteuses avec la BICM,
- assurer que le comportement des récepteurs choisis peut être prédit avec des techniques d'abstraction du lien, afin de permettre aux simulateurs système d'évaluer l'impact du récepteur dans le contexte d'une implémentation réaliste du protocole coopérative.

Pour aborder la première objective, nous avons mené une étude qui cherche à éclaircir divers catégories d'algorithmes de traitement du signal, en considérant le problème de régression sur une modèle linéaire générique sur un signal BICM avec une structure de récepteur itérative. Ensuite, plusieurs détecteurs SISO sont conçus avec des techniques d'inférence bayésienne approximée et des algorithmes basés sur l'AMP pour ce cas d'école, afin d'identifier

leurs points forts et leurs faiblesses. À cet égard, nous avons effectué une revue sur les 25 années de développements liées aux algorithmes de turbo-détection, et nous avons focalisé notre attention sur les techniques d'estimation statistique approximés émergents tels que le BP, EP, MF, PDA et les algorithmes (·/G/V/O)AMP. En particulier, nous nous focalisons principalement sur les approches permettant de concevoir des turbo-détecteurs doublement itératives, et afin de prendre en considération le second objectif, nous avons cherché à comprendre leurs limites asymptotiques et la prédictibilité de ces méthodes, avec une analyse unifiée. La motivation derrière l'étude des récepteurs à deux-boucles provient du fait que le décodage SISO peut avoir une complexité et une latence prohibitive, et nous espérions pouvoir remplacer des itérations avec le décodeur par des itérations avec un processus de démodulation (ou de décision) plus simple. Nous avons aussi brièvement investigué les méthodes d'apprentissage profond pour la couche PHY, avec une attention particulière donnée au concept du dépliement profond qui permet d'optimiser des algorithmes itératives en les assimilant à des réseaux profonds.

La conclusion principale tiré de cette étude est que les détecteurs SISO avec des itérations internes constituent une catégorie d'algorithme prometteuse et en particulier, ceux conçus avec le passage de messages à base de EP, ou liées aux techniques comme le VAMP/OAMP, qui ont la propriété d'être prédictibles.

Compte tenu de cette conclusion, afin d'aborder spécifiquement le problème de mitigation de l'ISI pour les transmissions mono-porteuses considérées, nous avons investigué la conception d'un égaliseur. D'abord, nous avons balayé l'état de l'art sur les techniques de turbo-égalisation, avec un accent sur les structures domaine temporelles (TD) avec des filtres à réponse impulsionnelle fini (FIR), et afin de remédier à leur faiblesses, nous avons proposé des nouvelles turbo-égaliseurs FIR DFE, en exploitant du retour souple à base de EP et des estimés APP. Grâce à une nouvelle stratégie de calcul efficace de l'inverse de la matrice de covariance, le compromis performance-complexité est en faveur des structures FIR DFE à base d'EP, en vue de leurs performances impressionnants à des points de fonctionnement à haute efficacité spectrale. Nous avons aussi exploré la conception des FIR DFE auto-itérées afin de davantage pousser les débits atteignables vers la capacité du canal. Ces études ont été menées avec des structures FIR exactes qui utilisent des filtres dynamiques, et pour réduire la complexité calculatoire, nous avons aussi abordé la conception de turbo DFE avec des filtres statiques. Alors que cela soit un problème de conception non-linéaire complexe, avec les coefficients du filtre optimal qui dépendent de la sortie de l'égaliseur elle-même, à travers la prédiction semi-analytique, nous avons proposé une méthode efficace pour obtenir des filtres permettant d'atteindre une performance quasi-optimale.

Même si les structures FIR DFE proposés apportent une amélioration significative de performances, leur capacité à être implémenté dans des plateformes réelles était douteux, à cause de la mode opérateur de retour de décision série qui cause beaucoup de latence avec l'étape de démodulation. Pour remédier à cela nous avons mené une étude avec des structures domaine fréquentielle (FD) avec un retour de décision prise sur l'ensemble de bloc de données. À cet égard, nous avons proposé une méthode de conception qui consiste à utiliser le passage de message EP scalaire pour obtenir des turbo-égaliseurs doublement-itératifs avec traitement

dans la FD. Similairement aux observations sur les FIR DFE, les auto-itérations apportent des améliorations significatives de performance pour des applications haut-débit. Ces récepteurs ont montré des gains significatifs comparés à des structures FD plus conventionnels, toute en atteignant les mêmes limites asymptotiques que les structures domaine temporelles (TD) exactes. Les points forts et les faiblesses de ces algorithmes ont été analysés théoriquement et avec des simulations Monte-Carlo à taille fini, et de l'analyse asymptotique de transfert de l'information extrinsèque (EXIT). Ceci est vraisemblablement la contribution technique qui a le plus grand potentiel d'avoir un impact significative, puisque ce récepteur a une structure simple et avec une complexité réduite, adapté à l'usage répandu des structures FD dans des radios à bas coût. Nous avons aussi exploré l'optimisation des paramètres de ce récepteur avec un usage original du dépliement et de l'apprentissage profond, et nous avons obtenu des résultats prometteurs à travers une analyse numérique. Pour finir, l'approche proposé a été étendu pour traiter des problèmes SC plus complexes, tel que le SC-FDMA, l'égalisation des canaux temps variants et multiplexage spatiale MIMO multi-utilisateurs. L'impact de l'estimation du canal a été aussi discuté pour ce récepteur.

Pour finir, ayant trouvé un candidat convenable pour le récepteur PHY des MANETs futurs, on aborde le second objectif de cette thèse en mettant en place une approche simulation système avec une modèle de PHY réaliste. À cet égard, des nouvelles approches d'abstraction du lien sont proposées pour le récepteur FD auto-itérés, en utilisant des techniques d'analyse asymptotique, renforcés avec quelques heuristiques pour raffiner la précision de la prédiction à taille-fini. En particulier, une méthode est proposée pour suivre la dispersion statistique du retour souple afin de générer des erreurs pseudo-aléatoires sur la fiabilité souple afin de mener une prédiction semi-analytique ne nécessitant pas de calibration. Pour attirer l'attention sur l'importance de disposer de tels outils de simulation, nous avons exploré une méthode de coopération distribuée fondamentale, la diffusion coopérative. Nous avons montré que ce protocole était à l'origine d'une quantité importante de l'ISI, mais aussi de la sélectivité temporelle au niveau de récepteur, et avec une modèle de PHY simple nous avons illustré des gains importants de performance que des récepteurs avancés de l'état de l'art peuvent apporter à des MANETs coopératives.

Perspectives

Améliorations supplémentaires sur les FIRs

Dans les travaux de *Jeong et. al.* [JM13; Jeo11], ils effectuent la concaténation parallèle de plusieurs DFEs FIR afin d'améliorer significativement la performance de détection. En particulier, la concaténation de deux DFEs en renversant l'axe de temps permet d'obtenir la structure de DFE bidirectionnelle. L'extension de ces approches avec EP et des auto-itérations mérite d'être explorée.

Ces structures peuvent être aussi améliorées avec l'usage de l'égalisation fractionnaire (FS) [Pet+19], où de nombreux défis sur la complexité et l'estimation du canal doivent être résolus.

Lisseurs de Kalman

Nous avons évoqué les liens entre les égaliseurs bloc en domaine temporelle et les lisseurs de Kalman, et divers variantes de lisseurs de Kalman ont été étudies avec EP dans [SZW+15; SMF+17b]. En particulier, la dernière structure possède la même performance que des bancs de filtres temporels, ainsi cela sera intéressant de mener une étude du compromis complexité-performance afin d'évaluer l'intérêt de ces structures pour des canaux temps-variants.

Estimation du canal et détection conjointe en domaine fréquentiel

Même si la plupart de cette thèse est focalisé sur le problème de détection, les techniques de passage de messages sont aussi utilisées de manière répandue pour l'estimation du canal [Han+18]. En vue de la commodité offerte par l'égaliseur FD auto-itérée, cela serait intéressant de l'étendre afin d'aussi raffiner les estimés du canal avec des auto-itérations et de l'inférence approximée.

Dépliement profond

Dans cette thèse, l'utilisation du dépliement profond a été justifiée suivant l'observation que l'algorithme VAMP, liée à l'EP, est déjà optimal par rapport à sa structure, quand la connaissance parfaite du CSI est disponible et pour des signaux bernoulli-gaussiennes [BSR17]. Cela serait intéressant de refaire cette expérience avec signaux issues des constellations PSK et QAM afin d'assurer qu'il n'y a pas améliorations à travers l'apprentissage sur cette structure pour des problèmes de communications numérique. Puis, nous pourrons explorer l'inclusion d'autres boucles et des paramètres entrainables, autre que le lissage du retour, afin de davantage optimiser les réseaux de détection domaine fréquentiels dépliées. Les bénéfices du dépliement profond devraient être aussi explorés pour l'estimation du canal.

EP pour le codage par superposition et NOMA

Dans cette thèse, les systèmes BICM ont été au coeur des études, comme ils permettent l'utilisation des stratégies de turbo-détection entre un détecteur SISO et un décodeur SISO. Néanmoins des techniques de codage par superposition sont en train d'être étudiés pour des techniques NOMA émergents. À travers la flexibilité que l'EP apporte pour projeter des messages dans des familles de PDF pratiques, on pourrait explorer des techniques efficaces mais à faible complexité pour de tels stratégies de modulation.

Évaluer les gains au niveau système pour l'utilisation des récepteurs avancés dans les MANETs

Une suite logique au dernier chapitre de cette thèse est l'utilisation de l'approche de simulation coopérative proposée et de la méthode de prédiction de performance pour évaluer les gains au niveau-système de l'utilisation des auto-itérations dans les MANETs. Ces évaluations devraient prendre en compte les caractéristiques des canaux MANETs en modélisant le masquage et la mobilité.

Protocoles de diffusion/multidiffusion multiples pour les MANETs

Pour finir, les protocoles coopératives explicitement mentionnés dans cette thèse consistent soient des transmissions monodiffusion avec sélection de relai, ou du diffusion/multidiffusion avec de la diffusion coopérative. Méthodes d'abstraction de couche PHY pour des protocoles émergents basé sur NOMA ou le codage réseau en couche PHY [Nav18; Liu+19a] devront être exploré avec un PHY réaliste. Tels protocoles non-orthogonaux pourraient permettre la mise en commun de plusieurs charges utiles pour une transmission PHY et donc permettre la création de plusieurs liens coopératifs simultanément.

Appendix A

On Physical Channel Simulation

A.1 Equivalent Upsampled Discrete Channel Model

In practical transceivers, or for numerical simulation, the time-continuous channel and pulse shaping filtering operations are often replaced by equivalent discrete models, by operating at a sampling frequency f_s higher than the signal bandwidth. In this regard, we need to convert continuous convolutions to equivalent discrete-time convolutions, as per Equation (1.20), for generating the impact of $\mathrm{h}_{u,v}^{\prime(p)}(t)$ on the user signals and the impact of $\mathrm{h}_{\mathrm{RRC}}^{(\alpha)}(t)$ on the AWGN. For clarity, let us consider a simpler noise-free single-user transmission, which yields

$$\mathrm{y}_u^{\prime(p)}(t) = \mathrm{h}_{u,v}^{\prime(p)}(\tau, t) * \mathrm{x}_v^{\prime(p)}(t), \tag{A.1}$$

before sampling. The user signal $\mathrm{x}_v^{\prime(p)}(t)$ has infinite-bandwidth, but as the shape filters are used, the useful signal will be band-limited to B, which is at most $2W$, when roll-off factor is $\alpha = 1$. Hence, the sampling frequency of the equivalent discrete model has to be upsampled by a factor $n_s \geqslant 2$ greater than the symbol rate, i.e. $f_s = n_s W$. Moreover, the equally-spaced tapped-delay line method presented in [Isk08] is used to discretize the channel, by taking the cardinal sine expansion of $\mathrm{x}_v^{\prime(p)}(t)$ at rate the f_s. We have

$$\mathrm{y}_u^{\prime(p)}(t) = \int_{-\infty}^{\infty} \mathrm{h}_{u,v}^{\prime(p)}(\tau, t)\mathrm{x}_v^{\prime(p)}(t-\tau)\,\mathrm{d}\tau, \tag{A.2}$$

and by taking the Whittaker–Shannon interpolation expansion of $\mathrm{x}_v^{\prime(p)}(t-\tau)$ on the delay domain, we have

$$\mathrm{x}_v^{\prime(p)}(t-\tau) = \sum_{m=-\infty}^{\infty} \mathrm{x}_v^{\prime(p)}\left(t - \frac{m}{f_s}\right) \mathrm{sinc}(f_s\tau - m). \tag{A.3}$$

Moreover, by writing the channel explicitly as, $\mathrm{h}_{u,v}^{\prime(p)}(\tau, t) \triangleq \sqrt{\mathcal{E}_{\mathrm{rx},u}}\mathrm{h}_{\mathrm{RRC}}^{(\alpha)*}(-\tau) * \mathrm{h}_{u,v}^{(p)}(\tau, t) * \mathrm{h}_{\mathrm{RRC}}^{(\alpha)}(\tau)$ with

$$\mathrm{h}_{u,v}^{\prime(p)}(\tau, t) = \sqrt{\mathcal{E}_{\mathrm{rx},u}} \sum_{l=1}^{L_{\mathrm{ss}}} a_{u,v,l}^{(p)}(t)\mathrm{h}_{\mathrm{RC}}^{(\alpha)}(\tau - \tau_{u,v,l}^{(p)}), \tag{A.4}$$

where $\mathrm{h}_{\mathrm{RC}}^{(\alpha)}$ is the raised-cosine filter, with roll-off α, we have

$$\mathrm{y}_u'^{(p)}(t) = \sqrt{\mathcal{E}_{\mathrm{rx},u}} \sum_{m=-\infty}^{\infty} \mathrm{x}_v'^{(p)}\left(t - \frac{m}{f_s}\right) \sum_{l=1}^{L_{\mathrm{ss}}} a_{u,v,l}^{(p)}(t) \underbrace{\int_{-\infty}^{\infty} \mathrm{sinc}(f_s\tau - m)\mathrm{h}_{\mathrm{RC}}^{(\alpha)}(\tau - \tau_{u,v,l}^{(p)})\,\mathrm{d}\tau}_{\triangleq g_{u,v,l,m}^{(p,\alpha)}}. \quad \text{(A.5)}$$

An important point is about the pulse shaping filters; in practice the are often replaced by truncated, normalized, finite-impulse response filters of size L_{ps}, that are causal and discretized at the sampling frequency f_s, i.e.

$$\mathrm{h}_{\mathrm{RC}}^{(\alpha)}(\tau) = \sum_{l=1}^{L_{\mathrm{ps}}} h_{\mathrm{RC},l}^{(\alpha)} \delta\left(\tau - \frac{l}{f_s}\right). \quad \text{(A.6)}$$

Finally, by using the expression of circularized transmitted sequence, we have

$$\mathrm{y}_u'^{(p)}(t) = \sum_{m=-\infty}^{\infty} \sum_{n'=1}^{N'} x_{v,n'}'^{(p)} \delta\left(t - \frac{m}{f_s} - \frac{n'-1}{W}\right) \underbrace{\sqrt{\mathcal{E}_{\mathrm{rx},u}} \sum_{l=1}^{L_{\mathrm{ss}}} a_{u,v,l}^{(p)}(t) g_{u,v,l,m}^{(p,\alpha)}}_{\triangleq \mathrm{h}_{u,v,m}'^{(p,\alpha)}(t)}, \quad \text{(A.7)}$$

and using this model, we can investigate the computation of sampled observations $\mathrm{y}_u'^{(p)}(t_0 + n/W)$. With a upsampled discrete observations, the synchronization problem of choosing t_0 needs to be replaced by choosing the correct downsampling phase m_0 among $m = 0, ..., n_s - 1$ possibilities, such that $t_0 = m_0/f_s$, and we have

$$y_{u,n}'^{(p)} \triangleq \mathrm{y}_u'^{(p)}\left(t_0 + \frac{n}{W}\right) = \sum_{m=-\infty}^{\infty} \sum_{n'=1}^{N'} x_{v,n'}'^{(p)} \mathrm{h}_{u,v,m}'^{(p,\alpha)}\left(\frac{n}{W} + \frac{m_0}{f_s}\right) \delta\left(\frac{n - n' + 1}{W} - \frac{m - m_0}{f_s}\right), \quad \text{(A.8)}$$

where the Dirac delta function enforces $n' = (m_0 - m)/n_s + n + 1$. Moreover as n' is an integer, only non-zero components of this sum are given for $m = m_0 + n_s l$, with l being any integer, thus

$$y_{u,n}'^{(p)} \triangleq \mathrm{y}_u'^{(p)}\left(t_0 + \frac{n}{W}\right) = \sum_{l=-\infty}^{\infty} \underbrace{\mathrm{h}_{u,v,(m_0+n_s l)}'^{(p,\alpha)}\left(\frac{n}{W} + \frac{m_0}{f_s}\right)}_{\triangleq h_{u,v,n,l}'^{(p,\alpha)}} x_{v,n-l+1}'^{(p)}. \quad \text{(A.9)}$$

By truncating the taps $h_{u,v,n,l}'^{(p,\alpha)}$ that have an average power below a certain threshold, and by considering the channel and pulse shaping filters to be causal, we finally obtain the discrete baseband model

$$y_{u,n}'^{(p)} = \sum_{l=1}^{L} h_{u,v,n,l}'^{(p)} x_{v,n-l+1}'^{(p)}. \quad \text{(A.10)}$$

Note that the discrete channel model is not unique, there can be n_s different impulse responses depending on the choice of m_0, this will be further discussed in the next section.

As outlined in the SNR formulas in the previous section, the channel power $\|\mathrm{h}_{u,v}^{\prime(p)}\|^2$ plays a significant role in it. It is desirable to set the expected value of the small-scale components to be a unitary constant, such that $\mathbb{E}\left[\|\mathrm{h}_{u,v}^{\prime(p)}\|^2\right] = \mathcal{E}_{\mathrm{rx},u} g_{u,v}^{\mathrm{PL}} g_{u,v}^{\mathrm{sh}(p)}$, in order to tie the impact of large-scale path-loss and shadowing directly to the SNR. To this end, it is not sufficient to normalize the power-delay-profile of the small-scale channel component, because the corresponding discrete channel model are correlated, and they have an un-normalized impulse-response due to pulse-shaping, up-sampling factor, and the synchronization instant. To circumvent this issue, in this thesis, we precompute the norm of $\mathbf{h}_{u,v,n}^{\prime(p)} = [h_{u,v,n,1}^{\prime(p)}, \ldots, h_{u,v,n,L}^{\prime(p)}]$, and normalize the small-scale channel profile by this factor, in order to control the value of the received SNR.

A.2 Synchronization and Radio Imperfections

As mentioned in the previous sections, a synchronization algorithm is needed in order to find the sampling instant t_0. Moreover, as the PHY layer front-end operates at a sampling frequency of f_s, it is rather needed to find the ideal sampling phase $m_0 = 0, \ldots, n_s - 1$.

In this thesis, realistic synchronization techniques are out of scope and instead some sub-optimal genie techniques will be used. Moreover, an ideal synchronisation would consist in choosing m_0 to minimize the Packet Error Rate (PER), but that would require carrying out n_s decoding operations in order to select the one with optimal performance. Due to the complexity of this operation, we consider two sub-optimal "genie" approaches (i.e. techniques that use ideal channel knowledge).

First, in the absence of a frequency-selective channel ($L = 1$), m_0 is selected to maximize the SNR at the input of the AWGN symbol demodulator, which corresponds to sampling $\mathrm{y}_u^{\prime(p)}(t)$ when the corresponding eye diagram is at its maximum. Secondly, in the general case where the channel is selective, there is no longer a direct correspondence between the channel and the demodulation (due to underlying equalization or detection procedure), hence m_0 is simply selected to maximize the channel power $\|\mathrm{h}_{u,v}^{\prime(p)}\|^2$ at the output of the matched filter.

Another important PHY aspect is synchronization errors. Indeed, even when using a "genie" method, the resolution of the synchronization instant is limited by the upsampling factor n_s. Indeed, ideal synchronization phase is related to the ideal instant with $m_0 = \mathrm{round}(t_0 f_s)$, with $t_0 < 1/W$, hence small values of n_s result in a quantification error on the sampling instant. Moreover, when considering the use of real-world sub-optimal synchronization algorithm and imperfect local oscillators (especially for low-cost embedded radios), it becomes necessary to model imperfect synchronization, whose impact on cooperative and/or non-orthogonal transmissions can be severe.

We assume that local oscillator is calibrated for reference time and carrier frequency synthesis with a distributed network-wide coarse synchronization algorithm, and the sampling phase selection at the output of the matched-filter is operated by a fine time-synchronization algorithm. Following these, relative to an ideal global clock t, each node u has time-offset

of t_u^{off} seconds, and residual frequency drift of f_u^{off}. Moreover, we will assume, at each MAC frame, that t_u^{off} is uniformly distributed in $[-\tau_{\text{clk}}/2, \tau_{\text{clk}}/2]$, and we define $\tau_{u,v}^{\text{clk}} \triangleq t_u^{\text{off}} - t_v^{\text{off}}$, and $\phi_{u,v} \triangleq f_u^{\text{off}} - f_v^{\text{off}}$. Then these imperfections can be incorporated into the propagation channel model, as

$$\mathrm{h}_{u,v}^{(p)}(\tau, t) = \sum_{l=1}^{L_{\text{ss}}} e^{j2\pi t \phi_{u,v}} a_{u,v,l}^{(p)}(t) \delta(\tau - \tau_{u,v,l}^{(p)} - \tau_{u,v}^{\text{clk}}). \tag{A.11}$$

Finally, to conclude on radio imperfections, for accounting for possible jamming or radio quantification noise or any other source of interference, the covariance matrix of the considered interference has to be added to the Equation (1.23), and SNR expressions has to be accounted for accordingly.

Appendix B

On Receiver Computational Complexity Estimation

B.1 Introduction

Digital receivers can be implemented on practical platforms either through software in embedded processors, digital signal processors, or through hardware architectures such as Digital Signal Processing (DSP), Field Programmable Gate Array (FPGA) or Application Specific Integrated Circuit (ASIC). Hence, computational, area or latency efficiency of receiver algorithms is a crucial metric and its accurate estimation is difficult, due to differences in parallelized implementation possibilities depending on platforms and the algorithm's type. Hence, to get a broad idea of complexity in this thesis, we seek a simple approach for computational complexity estimation.

Among different metrics for such evaluations in numerical analysis, the number of FLOPs and the number of memory accesses are commonly used. Although such metrics ignore the possibility of concurrency in algorithms, they can model respectively processing frequency and memory bandwidth requirements.

Here we consider the operation frequency and the occupied area to be main issues, and thus we use FLOP counts, which we define with regards to an elementary real addition and real multiplication in order to also evaluation the number of multiply-and-accumulate units/operations that would be needed for implementation. Hence the Table B.1 below is built by considering a real addition and a real multiplication to amount to half a floating point operation (0.5 FLOPs) each.

Receiver algorithms will be analyzed by the means of operations above and memory accesses will be ignored. Finally, *logical tests* or *maximum/minimum of two real-valued elements* will be each counted as an elementary addition A_r, i.e. 0.5 FLOPs.

B.2 Demapping Operations/Decision Devices

In this section, we are interested in quantifying the complexity of various constellation related operations that are used in different stages of a receiver.

Table B.1: Complexity of Basic Operations

Operation	Symbol	Mult.	Add.	FLOPs
Real Addition/Substraction	A_r	0	1	0.5
Real Multiplication/Division	M_r	1	0	0.5
Complex-Real Mult./Div.	M_{rc}	2	0	1.0
Complex Addition	A_c	0	2	1.0
Complex Mult.	M_c	4	2	3.0
Complex Div.	D_c	8	3	5.5
Real Square Root	S_r	-	-	3.0
Squared Complex Norm	P_c	2	1	1.5
Real Linear Interpolation	I_r	1	2	1.5
Binary Operations	B_o	-	-	0.5
Memory Access	M_a	-	-	0

First, when using prior bitwise information, a **soft mapper** is also needed on the feedback path for computing prior distribution. Soft mapping uses prior LLRs $L_a(d_{k,j})$ on coded bits $d_{k,j}$, provided from a decoder, to estimate a prior distribution on $x_k = \alpha$, $\forall \alpha \in \mathcal{X}$

$$\mathcal{P}_k(\alpha) \propto \prod_{j=1}^{Q} e^{-\varphi_j^{-1}(\alpha) L_a(d_{k,j})}, \tag{B.1}$$

where $\mathcal{X}$ is the symbol constellation with M elements. This distribution could either be computed in un-normalized log-domain ($\ln \mathcal{P}_k(\alpha) + \text{Cte}$) for further processing for demapping, or in the linear domain for the computation of soft symbol estimates. In the latter case, a non-linear operation is involved, which is approximated by interpolators, and for normalizing the distribution $M-1$ sums and M products are needed.

A **soft demapper** is a required in any receiver with a modern, soft-input error correction code, and it is needed in turbo-equalization to produce soft output extrinsic information, using symbolwise AWGN MAP criterion. Soft demapper **estimates APP symbol distribution**

$$\mathcal{D}_k(\alpha) \propto \exp\left(-|\alpha - x_k^e|^2 / v_{x,k}^e\right) \mathcal{P}_k(\alpha), \ \forall \alpha \in \mathcal{X}, \tag{B.2}$$

and then this APP distribution is **bitwise marginalized** in order to compute extrinsic LLRs

$$L_e(d_{k,j}) = \ln \textstyle\sum_{\alpha \in \mathcal{X}_j^0} \mathcal{D}_k(\alpha) - \ln \sum_{\alpha \in \mathcal{X}_j^1} \mathcal{D}_k(\alpha) - L_a(d_{k,j}), \tag{B.3}$$

with $\mathcal{X}_j^p = \{\alpha \in \mathcal{X} : \varphi_j^{-1}(x) = p\}$ where $p \in \mathbb{F}_2$. First two terms above have each $M/2-1$ log-MAP operations, which we approximate with a maximum (logical operation), an interpolator on the error term and their sum, i.e. $(M/2-1)(I_r + 2A_r)$ FLOPs.

When considering the computation of **hard detection estimate** $\tilde{x}_k$, the maximum of $\mathcal{D}_k(\alpha)$ is computed, hence $M-1$ logic operations are needed.

Prior or APP distributions can be used to compute **soft estimates and their reliabilities** of symbols to be detected. For the prior estimates we have,

$$\begin{aligned} x_k^p \triangleq \mathbb{E}_{\mathcal{P}_k}[x_k] &= \sum_{\alpha\in\mathcal{X}} \alpha \mathcal{P}_k(\alpha), \\ v_k^p \triangleq \mathrm{Var}_{\mathcal{P}_k}[x_k] &= \sum_{\alpha\in\mathcal{X}} |\alpha|^2 \mathcal{P}_k(\alpha) - |x_k^p|^2, \end{aligned} \tag{B.4}$$

and for APP estimates we use

$$\begin{aligned} \mu_k^d \triangleq \mathbb{E}_{\mathcal{D}_k}[x_k] &= \sum_{\alpha\in\mathcal{X}} \alpha \mathcal{D}_k(\alpha), \\ \gamma_k^d \triangleq \mathrm{Var}_{\mathcal{D}_k}[x_k] &= \sum_{\alpha\in\mathcal{X}} |\alpha|^2 \mathcal{D}_k(\alpha) - |\mu_k^d|^2. \end{aligned} \tag{B.5}$$

Usually for APP soft estimate computation, soft demapper's internal log domain APP distribution is used, hence soft estimate computation should account for the cost of going into the linear domain in this case.

The computational complexity of the elementary operations above, using the Table B.1 is given below in Table B.2. Note that there different variants of some operations depending on the availability of priors, or depending on the use of log-probabilities.

Table B.2: Elementary Mapping, Demapping and Decision Operations

Operation	Description	Details	FLOPs
$\{L_a(d_{k,j})\}_j \rightarrow \ln \mathcal{P}_k(\alpha) + \mathrm{Cte}$	Soft Mapping	$M(q-1)A_r$	$0.5M(q-1)$
$\{L_a(d_{k,j})\}_j \rightarrow \mathcal{P}_k(\alpha)$	Soft Mapping	$M[I_r + qA_r + M_r]$ $-A_r$	$M(0.5q+2)-0.5$
$x_k^e \rightarrow \ln \mathcal{D}_k(\alpha) + \mathrm{Cte}$ (no priors)	APP Estimation	$M(A_c + P_c + M_r)$	$3M$
$(x_k^e, \{L_a(d_{k,j})\}_j)$ $\rightarrow \ln \mathcal{D}_k(\alpha) + \mathrm{Cte}$	APP Estimation	$M(A_c + P_c + M_r +$ $qA_r)$	$3M + 0.5Mq$
$\ln \mathcal{D}_k(\alpha) \rightarrow \{L_e(d_{k,j})\}_j$ (no priors)	Bitwise Marginalization	$q((I_r + 2A_r)M$ $-2I_r - A_r)$	$q(2.5M - 3.5)$
$(\ln \mathcal{D}_k(\alpha), \{L_a(d_{k,j})\}_j)$ $\rightarrow \{L_e(d_{k,j})\}_j$	Bitwise Marginalization	$q((I_r + 2A_r)M$ $-2I_r)$	$q(2.5M - 3)$
$\ln \mathcal{D}_k(\alpha) \rightarrow \tilde{x}_k$	Hard Detection	$(M-1)B_o$	$0.5(M-1)$
$\mathcal{P}_k(\alpha)) \rightarrow (x_k^p, v_k^p)$ or $\mathcal{D}_k(\alpha) \rightarrow (\mu_k^d, \gamma_k^d)$	Soft Estimates and Reliabilities	$M(3A_r + 2M_{rc} + P_c)$ $-2A_r$	$5M - 1$
$\ln \mathcal{P}_k(\alpha)) \rightarrow (x_k^p, v_k^p)$ or $\ln \mathcal{D}_k(\alpha) \rightarrow (\mu_k^d, \gamma_k^d)$	Soft Estimates and Reliabilities	$M(4A_r + 2M_{rc} + P_c$ $+I_r + M_r) - 3A_r$	$7.5M - 1.5$

Finally, combining these different elementary operations, the computational complexity of common demapping/mapping or decision devices can be computed. Table B.3 presents a list of such devices.

Soft mapper is a device which combines the estimation of linear prior probability distribution, with the soft estimator which computes the mean and the variance of that distribution.

Soft demapper combines APP estimator with the bitwise marginalization. Moreover, APP estimator is also used for **Hard or Soft detectors**, by combining it with the hard

detection or the log-domain soft estimation.

Finally, in the context of turbo equalization, a **turbo demapper** module which both computes prior estimates for interference cancellation, and handles the demapping is considered. A variant of such devices which also provide soft APP feedback to the equalizer is also modelled.

Table B.3: Mappers, Demappers and Decision Devices

Operation	Description	Details	FLOPs
$\{L_a(d_{k,j})\}_j \rightarrow (x_k^p, v_k^p)$	Soft Mapper	$M[(q+3)A_r + I_r + M_r + 2M_r c + P_c] - 3A_r$	$M(0.5q+7) - 1.5$
$x_k^e \rightarrow \{L_e(d_{k,j})\}_j$	Soft Demapper	$M(A_c + P_c + M_r) + qM(I_r + 2A_r) - q(2I_r + A_r)$	$2.5qM + 3M - 3.5q$
$(x_k^e, \{L_a(d_{k,j})\}_j) \rightarrow \{L_e(d_{k,j})\}_j$	Soft Demapper with Priors	$M(A_c + P_c + M_r) + qM(I_r + 3A_r) - 2qI_r$	$3qM + 3M - 3q$
$x_k^e \rightarrow \tilde{x}_k$	Hard Detector	$M(A_c + P_c + M_r + A_r) - A_r$	$3.5M - 0.5$
$(x_k^e, \{L_a(d_{k,j})\}_j) \rightarrow \tilde{x}_k$	Hard Detector with Priors	$M(A_c + P_c + M_r) + qMA_r + (M-1)A_r$	$0.5qM + 3.5M - 0.5$
$(x_k^e, \{L_a(d_{k,j})\}_j) \rightarrow (\mu_n^d, \gamma_n^d)$	Soft APP Detector with Priors	$M((q+4)A_r + A_c + 2P_c + 2M_r + 2M_{rc} + I_r) - 3A_r$	$0.5qM + 10.5M - 1.5$
$(x_k^e, \{L_a(d_{k,j})\}_j) \rightarrow (\{L_e(d_{k,j})\}_j, x_k^p, v_k^p)$	Turbo Demapper	$M(4A_r + A_c + 2P_c + 2M_r + 2M_{rc} + I_r) + qM(I_r + 3A_r) - 2qI_r - 3A_r$	$10.5M + 3qM - 3q - 1.5$
$(x_k^e, \{L_a(d_{k,j})\}_j) \rightarrow (\{L_e(d_{k,j})\}_j, x_k^p, v_k^p, \mu_k^d, \gamma_k^d)$	Turbo Demapper with APP feedback	$M(8A_r + A_c + 3P_c + 3M_r + 4M_{rc} + 2I_r) + qM(I_r + 3A_r) - 2qI_r - 6A_r$	$18M + 3qM - 3q - 3$

B.3 Trellis and Factor Graph Operations

In this section, the complexity of Trellis-based algorithm BCJR for soft-input soft-output decoding, and soft-input soft-output detection in frequency selective channels is discussed.

Firstly, on a general purpose implementation of the BCJR algorithm, three types of messages are computed:

- **observation metrics** $\gamma(s, s')$ correspond to the probability of transition from a state s to state s', for states s, s' being among N_{state} values.
- **forward metrics** $\alpha(s') = \sum_{s \in \mathcal{P}_{\text{state}}(s')} \alpha(s)\gamma(s, s')$, with $\mathcal{P}_{\text{state}}(s')$ being the set of states that can reach to s',
- **backward metrics** $\beta(s) = \sum_{s' \in \mathcal{N}_{\text{state}}(s)} \gamma(s, s')\beta(s')$, with $\mathcal{N}_{\text{state}}(s)$ being the set of states that can be reached from s.

When computing the computational complexity of BCJR algorithm, we will neglect transitory effects of the Trellis, and consider the cardinalities $|\mathcal{P}_{\text{state}}(s')|$ and $|\mathcal{N}_{\text{state}}(s)|$ to be constant, and equal to N_{in}, i.e. number of inputs that cause a transition. Unless the closure of initial or final symbols of the Trellis do not involve important computations, they are neglected. Hence, BCJR complexity is dominated by the complexity of computing $\gamma(s, s')$. Moreover, as BCJR implementation is made in the log-domain sum operations will be implemented as log-MAP operations, which further increase the complexity.

Outputs of the algorithm are computed through the **APP probability of transition** given by $p(s, s') = \alpha(s)\gamma(s, s')\beta(s')$, and the **APP probability of state** given by $p(s) = \alpha(s)\beta(s)$. In particular, when seeking the probability of a specific input x which triggers any transition $(s, s') \in \mathcal{T}_{\text{input}}(x)$, with $|\mathcal{T}_{\text{input}}(x)| = N_{\text{state}}$ we have $p(x) = \sum_{(s,s')\in\mathcal{T}_{\text{input}}(x)} p(s, s')$.

Table B.4: BCJR General Operations

Operation	Description	Details	FLOPs
$\mathcal{C}_{\text{bcjr-fw/bw}}$	Forward/Backward Recursion	$N_{\text{state}}N_{\text{in}}(I_r + 3A_r)$	$3N_{\text{in}}N_{\text{state}}$
$\mathcal{C}_{\text{bcjr-app}}$	APP Input Estimation	$N_{\text{in}}N_{\text{state}}(I_r + 4A_r)$	$3.5N_{\text{in}}N_{\text{state}}$

Next, we look in detail the computational costs for SISO decoding and SISO frequency-selective channel detection.

B.3.1 SISO Decoding of Convolutional Codes

For decoding a convolutional code of constraint length C_L, with k input bits and n output bits, we have $N_{\text{input}} = 2^k$, $N_{\text{state}} = 2^{C_L-1}$. The observation metric is built using prior information on codewod bits c_n and on information bits u_k.

The prior information on q-bit words is computed with a similar formula to soft mapping(in log-domain) of Table B.2, but summed over 2^q elements, i.e $\mathcal{C}_{\text{bit-map}}(q) = 2^q(q-1)A_r$.

Here we consider a two step schedule for the BCJR algorithm: an initial backward recursion, followed by a joint forward and APP transition probability computation. The output of BCJR for decoding needs to be bitwise LLRs, hence a final step of marginalization over q-bits is required at the end, for $q = k$ for information bits, and for $q = n$ for code bits. This can be carried out alongside APP computation, for and additional cost of $\mathcal{C}_{\text{llr-marg}}(q) = N_{\text{state}}N_{\text{in}}(q-1)(I_r + 2A_r) + 2qA_r$.

The table B.5 below lists the complexity of different BCJR variants for decoding a K-symbol input (i.e Kk input bits, and Kn code bits).

First variant is the full BCJR decoder, to be used in a turbo decoder, which has both code and information bit inputs and outputs, such that $\mathcal{C}_{\text{bcjr-full}} = K(2\mathcal{C}_{\text{bcjr-fwbw}} + \mathcal{C}_{\text{bcjr-app}} + \mathcal{C}_{\text{bit-map}}(k) + \mathcal{C}_{\text{bit-map}}(n) + \mathcal{C}_{\text{llr-marg}}(k) + \mathcal{C}_{\text{llr-marg}}(n))$.

Another variant for turbo detection is the BCJR decoder, with a code input only, which

Table B.5: BCJR SISO Decoding

Operation	Description	Details
$\mathcal{C}_{\text{bcjr-full}}$	Code and Info. Input SISO Decoder	$K(N_{\text{state}}N_{\text{in}}((k+n+1)I_r+(2k+2n+6)A_r)$ $+(k(2^k+2)+n(2^n+2)-2^k-2^n)A_r)$
$\mathcal{C}_{\text{bcjr-cin}}$	Code Input Only SISO Decoder	$K(N_{\text{state}}N_{\text{in}}((k+n+1)I_r+(2k+2n+6)A_r)$ $+(k(2^k+2)+2n-2^k)A_r)$
$\mathcal{C}_{\text{bcjr-io}}$	Info. Output Only SISO Decoder	$K(N_{\text{state}}N_{\text{in}}((k+2)I_r+(2k+8)A_r)$ $+(k(2^k+2)+n2^n-2^k-2^n)A_r)$
$\mathcal{C}_{\text{bcjr-ciio}}$	Code In, Info. Out SISO Decoder	$K(N_{\text{state}}N_{\text{in}}((k+2)I_r+(2k+8)A_r)$ $+(k(2^k+2)-2^k)A_r)$

has both code and information bit outputs, such that $\mathcal{C}_{\text{bcjr-cin}} = K(2\mathcal{C}_{\text{bcjr-fwbw}} + \mathcal{C}_{\text{bcjr-app}} + \mathcal{C}_{\text{bit-map}}(n) + \mathcal{C}_{\text{llr-marg}}(k) + \mathcal{C}_{\text{llr-marg}}(n))$.

Finally, for non-turbo soft-input decoding of convolutional codes, a variant with code bit input and information bit output is needed. In detail, $\mathcal{C}_{\text{bcjr-ciio}} = K(2\mathcal{C}_{\text{bcjr-fwbw}} + \mathcal{C}_{\text{bcjr-app}} + \mathcal{C}_{\text{bit-map}}(n) + \mathcal{C}_{\text{llr-marg}}(k))$.

B.3.2 SISO Decoding of Turbo Codes

In this section we consider turbo codes with N_c constituent codes, all of constraint length C_L, with k input bits and n output bits.

A turbo code based on parallel concatenated convolutional codes (PCCC) requires both code and information inputs and code bit outputs of SISO convolutional decoders. Hence, $\mathcal{C}_{\text{pccc}} = N_c\mathcal{C}_{\text{bcjr-io}}$.

For turbo codes based on serial concatenated convolutional codes (SCCC) with $N_c = 2$, the outer decoder has to have code input only, and the inner decoder needs code output only hence $\mathcal{C}_{\text{sccc}} = \mathcal{C}_{\text{bcjr-cin}} + \mathcal{C}_{\text{bcjr-co}}$.

These complexity values need to be multiplied by the number of inner turbo code iterations.

B.3.3 SISO Decoding of LDPC Codes

For regular LDPC codes of codeword length N, with $M = N - K$ parity checks, of degree density (d_v, d_c), belief-propagation based sum-product decoding complexity is evaluated.

For variable nodes, the complexity of message per code bit is $d_v(A_r + A_r + (I_r + A_r))$, which respectively correspond to the computation of APP, extrinsic and soft-bit (tanh).

For check nodes, the complexity of message per parity-check bit is $d_c(A_r + A_r)) + (d_c -$

$1)(A_r + I_r + A_r)$ which respectively correspond to APP estimation, sign estimation, extrinsic computation and log-likelihood message.

Hence total complexity of LDPC sum-product decoding per inner iteration is $\mathcal{C}_{\text{ldpc-sp}} = Nd_v(I_r + 3A_r) + M(d_c(I_r + 4A_r) - (I_r + 2A_r))$.

B.4 Matrix-Vector Operations for Filtering

In filter based equalizers we commonly find similar sparse matrix operations, hence in this section we will compute the complexity of common operations with channel matrices, covariance matrices and equalizer filters. We will consider a transmitted block of K symbols, with a sliding window size of N with $N = N_p + N_d + 1$, where N_p is the number of causal elements in the window, and N_d is the number of anti-causal symbols.

The channel matrix is the time-varying convolution operator which, in its most general form, has a "Toeplitz-like" structure. When considered with a sliding window of size N, its size is $N \times (N + L - 1)$, it has L non zero elements per row, holding the values of the taps of the channel at the time index given by the row. When considered on the whole transmitted block, we have a $(K + L - 1) \times (K + 2L - 2)$ matrix for including both IBI from the previous block the residual ISI at the end. In some cases where IBI is cancelled, this can be reduced to a $(K + L - 1) \times K$ matrix which has L non zero elements per column. Here is an example of the two cases with $L = 3, K = 5$,

$$\begin{bmatrix} h_{0,2} & h_{0,1} & h_{0,0} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & h_{1,2} & h_{1,1} & h_{1,0} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & h_{2,2} & h_{2,1} & h_{2,0} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & h_{3,2} & h_{3,1} & h_{3,0} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & h_{4,2} & h_{4,1} & h_{4,0} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & h_{5,2} & h_{5,1} & h_{5,0} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & h_{6,2} & h_{6,1} & h_{6,0} \end{bmatrix}, \quad \begin{bmatrix} h_{0,0} & 0 & 0 & 0 & 0 \\ h_{1,1} & h_{1,0} & 0 & 0 & 0 \\ h_{2,2} & h_{2,1} & h_{2,0} & 0 & 0 \\ 0 & h_{3,2} & h_{3,1} & h_{3,0} & 0 \\ 0 & 0 & h_{4,2} & h_{4,1} & h_{4,0} \\ 0 & 0 & 0 & h_{5,2} & h_{5,1} \\ 0 & 0 & 0 & 0 & h_{6,2} \end{bmatrix},$$

the first matrix is full rank in rows, and the second matrix is full rank in columns, and these imply different properties for the related covariance matrix. In fact to obtain numerically stable covariance inverses, we use the first representation for computing $\mathbf{H}\mathbf{H}^H$ and the second for computing $\mathbf{H}^H\mathbf{H}$ but the latter case requires clearing the received symbols for IBI.

The rows of the channel matrices represent different channel states, i.e. "time", whereas the columns are channel taps, i.e. "delay". With time invariant channels, values on each row are identical but we can also encounter segmented channel matrices, where the channel state changes every D symbols.

When dealing with sliding window equalizers or the block linear equalizer, the observation covariance matrix $\mathbf{\Sigma}$ is of the form $\mathbf{\Sigma}_n = \sigma_w^2\mathbf{I}_N + \mathbf{H}_n\mathbf{\Sigma}_{\mathbf{x}_n}\mathbf{H}_n^H$, where $\mathbf{\Sigma}_{\mathbf{x}_n}$ is a diagonal variance matrix and where $\mathbf{H}_n$ is full rank in rows. This matrix is hermitian and in order to ensure this property despite numerical uncertainties we need to compute as the sum of the noise covariance (diagonal) with a Gramian product ($\mathbf{A}\mathbf{A}^H$), hence we define the augmented

channel matrix $\mathbf{G}_n = \mathbf{H}_n\sqrt{\mathbf{\Sigma}_{\mathbf{x}_n}}$, where the square root operation is unambiguous assuming the transmitted symbol covariance matrix is diagonal. Then the complexity of computing the covariance matrix depends on whether $\mathbf{G}_n$ is invariant, segmented or time-varying.

Table B.6: Sparse Covariance Matrix Computation

Operation	Description	Details
$\mathbf{G}_n^r\mathbf{G}_n^{r\,H}$ or $\mathbf{G}_n^{c\,H}\mathbf{G}_n^c$	Variant Gram	$NLP_c + N(L-1)A_r + (\frac{NL^2}{2} - \frac{L^3}{6} - \frac{NL}{2} + \frac{L}{6})M_c +$ $(\frac{NL^2}{2} - \frac{L^3}{5} - \frac{3NL}{2} + \frac{L^2}{2} + N - \frac{L}{3})A_c$
$\mathbf{G}^r\mathbf{G}^{r\,H}$ or $\mathbf{G}^{c\,H}\mathbf{G}^c$	Invariant Gram	$LP_c + (L-1)A_r + \frac{L(L-1)}{2}M_c + \frac{(L-1)(L-2)}{2}A_c$
$\mathbf{G}^r\mathbf{G}^{r\,H}$	Segmented Gram (full rank in row)	$N_{\text{seg}}(LP_c + (L-1)A_r) +$ $(2N_{\text{seg}} - 1)(\frac{L(L-1)}{2}M_c + \frac{(L-1)(L-2)}{2}A_c)$
$\mathbf{G}^{c\,H}\mathbf{G}^c$	Segmented Gram (full rank in column)	$((N_{\text{seg}} - 1)L + 1)LP_c +$ $((N_{\text{seg}} - 1)L + 1)(L-1)A_r +$ $L(L-1)\frac{N\text{seg}(2L-1)-2L+4}{6}M_c +$ $(L-1)(L-2)\frac{N_{\text{seg}}(2L+3)-2L}{6}A_c$

Using the full rank in row structure, the complexity of this equalizer in time varying, time invariant and segmented cases is given in the table B.6. Meanwhile, when dealing with equalizers such as block DFE, where a Gramian such as $\mathbf{A}^H\mathbf{A}$ is required for covariance computation, if the channel matrix is in full rank in row format, then the covariance cannot be inverted with numerical stability, hence we use the full rank in column format and compute its complexity for variant, invariant and segmented cases. We find that for variant and invariant cases, the complexity is identical to the full rank in row $\mathbf{A}\mathbf{A}^H$ product, but its complexity increases for the segmented computation.

Another matrix operation, common in sliding window equalizer design is the computation of the feedforward filter $\mathbf{c}_n$, from which other filters are derived. This operation is commonly revolves around $\mathbf{\Sigma}_n^{-1}\mathbf{h}_n$ where the previously computed sparse covariance matrix's inverse is multiplied with the current channel state vector, which has L non zero elements. Especially when computing this quantity successively, for a time-varying equalizer, we need to consider practical recursive methods. There are multiple approaches for practical computation of this quantity:

- Naive method: we invert the covariance matrix and multiply it with the channel vector,
- Linear system solving: $\mathbf{\Sigma}_n\mathbf{c}_n = \mathbf{h}_n$ is solved using Gauss-Jordan elimination,
- Simplified system solving: the covariance matrix is decomposed into triangular matrices with Cholesky decomposition $\mathbf{\Sigma}_n \triangleq \mathbf{L}_n\mathbf{L}_n^H$, then systems $\mathbf{L}_n\mathbf{r}_n = \mathbf{h}_n$ and $\mathbf{L}_n^H\mathbf{c}_n = \mathbf{r}_n$ are solved for $\mathbf{r}_n$ and $\mathbf{c}_n$, using forward and backward substitutions respectively.

The complexity of different filter computation approaches are given on the table B.7, we can compare roughly the complexity of these operations with looking at dominant multiplication terms; inversion follows $\mathcal{O}(\frac{4}{3}N^3)$, Gauss-Jordan elimination follows $\mathcal{O}(\frac{1}{3}N^3)$ and the

Table B.7: Sparse Filter Computation

Operation	Description	Details
$\mathbf{\Sigma}_n^{-1} = \text{inv}(\mathbf{\Sigma}_n)$, $\mathbf{c}_n = \mathbf{\Sigma}_n^{-1}\mathbf{h}_n$	Matrix Inversion + Matrix-Vector Product	$\frac{8N^3+4N}{6}M_{rc} + \frac{8N^3-9N^2+N}{6}A_c$ $+NLM_c + N(L-1)A_c$
$\mathbf{c}_n = \text{solve}(\mathbf{\Sigma}_n\mathbf{c}_n = \mathbf{h}_n)$	Gauss-Jordan elimination	$\frac{2N^3+6N^2+4N}{6}M_{rc} + \frac{2N^3+3N^2-5N}{6}A_c$
$\mathbf{L}_n = \text{chol}(\mathbf{\Sigma}_n)$, $\mathbf{r}_n = \text{solve}(\mathbf{L}_n\mathbf{r}_n = \mathbf{h}_n)$, $\mathbf{c}_n = \text{solve}(\mathbf{L}_n^H\mathbf{c}_n = \mathbf{r}_n)$	Cholesky Decomposition + substitutions	$(\frac{NL^2}{2} - \frac{L^3}{3} + \frac{NL}{2} - \frac{3L^2}{2} - N + \frac{29L}{6} - 4)M_c$ $+(\frac{NL^2}{2} - \frac{L^3}{3} + \frac{NL}{2} - \frac{3L^2}{2} - N + \frac{35L}{6} - 4)A_c$ $+(NL - \frac{L^2}{2} + \frac{3L}{2} - N - 1)M_{rc} + NS_r$ $+2(N - \frac{L}{2})(L-1)(M_c + A_c) + 2NM_{rc}$

Cholesky approach follwos $\mathcal{O}(\frac{1}{2}NL^2)$. The Cholesky decomposition is computed using the sparsity of the covariance matrix; it has a $L-1$ bandwidth, i.e. it has $2L-1$ non zero diagonals. The sparsity effect cannot be included into matrix inversion or Gauss-Jordan elimination easily because these operations change the positions of zero elements as they are carried out, hence they provide worst case values.

These methods can be used then to estimate the complexity of filter-based equalizers and detectors.

Bibliography

[3gpa] *3GPP; TS Group Radio Access Network; E-UTRA; Multiplexing and channel coding (Release 15)*. Standard. Valbonne, FR: 3rd Generation Partnership Project, Mar. 2019 (cit. on p. 135).

[3gpb] *3GPP; TS Group Radio Access Network; E-UTRA; User Equipment (UE) radio transmission and reception (Release 14)*. Standard. Valbonne, FR: 3rd Generation Partnership Project, Jan. 2017 (cit. on pp. 30, 135).

[3gpc] *3GPP; TS Group Radio Access Network; Study on Network-Assisted Inference Cancellation and Suppression (NAIC) for LTE (Release 12)*. Standard. Valbonne, FR: 3rd Generation Partnership Project, Mar. 2014 (cit. on pp. 156, 221, 223, 232).

[3gpd] *3GPP; TS Group Radio Access Network; Study on Non-Orthogonal Multiple Access (NOMA) for NR (Release 16)*. Standard. Valbonne, FR: 3rd Generation Partnership Project, Dec. 2018 (cit. on p. 49).

[Şa+a] S. Şahin, A. M. Cipriano, C. Poulliat, and M.-L. Boucheret. *Méthode de prédiction des performances d'un récepteur itératif basé sur la propagation d'espérance*. FR Patent, time-stamped: Oct. 2019 (cit. on pp. 7, 15, 18).

[Şa+b] S. Şahin, A. M. Cipriano, C. Poulliat, and M.-L. Boucheret. *Procédé pour calculer une estimation d'un singal numérique modulé et de sa fiabilité*. FR+EU Patent, time-stamped: Feb. 2018 (cit. on pp. 6, 15, 18).

[Şa+18a] S. Şahin, A. M. Cipriano, C. Poulliat, and M.-L. Boucheret. "A Framework for Iterative Frequency Domain EP-Based Receiver Design". In: *IEEE Transactions on Communications* 66.12 (Dec. 2018), pp. 6478–6493 (cit. on pp. 6, 15, 17, 171).

[Şa+18b] S. Şahin, A. M. Cipriano, C. Poulliat, and M.-L. Boucheret. "Iterative Equalization Based on Expectation Propagation: A Frequency Domain Approach". In: *26th European Signal Processing Conference (EUSIPCO)*. Roma, IT, Sept. 2018, pp. 932–936 (cit. on pp. 6, 14, 17, 171).

[Şa+18c] S. Şahin, A. M. Cipriano, C. Poulliat, and M.-L. Boucheret. "Iterative Equalization With Decision Feedback Based on Expectation Propagation". In: *IEEE Transactions on Communications* 66.10 (Oct. 2018), pp. 4473–4487 (cit. on pp. 6, 14, 17).

[Şa+18d] S. Şahin, A. M. Cipriano, C. Poulliat, and M.-L. Boucheret. "On Cooperative Broadcast in MANETs with Imperfect Clock Synchronization". In: *IEEE Military Communications Conference (MILCOM)*. Los Angeles, CA, USA, Oct. 2018, pp. 1–7 (cit. on pp. 6, 7, 15, 17).

[Şa+18e] S. Şahin, C. Poulliat, A. M. Cipriano, and M.-L. Boucheret. "Spectrally Efficient Iterative MU-MIMO Receiver for SC-FDMA Based on EP". In: *IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*. Bologna, IT, Sept. 2018, pp. 1–7 (cit. on pp. 6, 15, 17, 171, 203).

[Şa+19a] S. Şahin, C. Poulliat, A. M. Cipriano, and M.-L. Boucheret. "Doubly Iterative Turbo Equalization: Optimization Through Deep Unfolding". In: *IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*. İstanbul, TR, Sept. 2019, pp. 1–6 (cit. on pp. 5, 6, 14, 15, 17).

[Şa+19b] S. Şahin, A. M. Cipriano, C. Poulliat, and M.-L. Boucheret. "Evolution Analysis of Iterative BICM Receivers with Expectation Propagation over ISI Channels". In: *IEEE International Symposium on information Theory (ISIT)*. Paris, FR, July 2019, pp. 1–4 (cit. on pp. 5, 7, 14, 15, 17).

[Şa+19c] S. Şahin, A. M. Cipriano, C. Poulliat, and M.-L. Boucheret. "Récepteurs Itératifs par Propagation de l'Espérance : Optimisation par Dépliage Profond". In: *27ème Colloque GRETSI sur le Traitement du Signal et des Images (GRETSI 2019)*. Lille, FR: GRETSI CNRS, 2019, pp. 1–4 (cit. on pp. 6, 15, 18).

[Şa+20a] S. Şahin, A. M. Cipriano, C. Poulliat, and M.-L. Boucheret. "Iterative Decision Feedback Equalization Using Online Prediction". In: *IEEE Access* 8 (Jan. 2020), pp. 23638–23649 (cit. on pp. 6, 14, 17).

[Şa+20b] S. Şahin, A. M. Cipriano, C. Poulliat, and M.-L. Boucheret. "Joint Frequency Domain Channel Estimation and Equalization based on Expectation Propagation for Single Carrier Transmissions". In: *45th IEEE 45th International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. Barcelona, ES, May 2020, pp. 1–5 (cit. on pp. 6, 15, 17).

[AC95a] N. Al-Dhahir and J. M. Cioffi. "Fast computation of channel-estimate based equalizers in packet data transmission". In: *IEEE Transactions on Signal Processing* 43.11 (Nov. 1995), pp. 2462–2473 (cit. on p. 106).

[AC95b] N. Al-Dhahir and J. M. Cioffi. "MMSE decision-feedback equalizers: finite-length results". In: *IEEE Transactions on Information Theory* 41.4 (July 1995), pp. 961–975 (cit. on pp. 102, 103).

[ADT11] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse. "Wireless Network Information Flow: A Deterministic Approach". In: *IEEE Transactions on Information Theory* 57.4 (Apr. 2011), pp. 1872–1905 (cit. on pp. 21, 41).

[AFR18] A. Abelló, J. Freixe, and D. Roque. "Closed-Form Expressions for Channel Shortening Receivers Using *APriori* Information". In: *IEEE Communications Letters* 22.7 (July 2018), pp. 1394–1397 (cit. on p. 101).

[AKt02] A. Ashikhmin, G. Kramer, and S. ten Brink. "Code rate and the area under extrinsic information transfer curves". In: *Proceedings IEEE International Symposium on Information Theory,* June 2002, pp. 115– (cit. on p. 65).

[ALV+06] D. M. Arnold, H. A. Loeliger, P. O. Vontobel, et al. "Simulation-Based Computation of Information Rates for Channels With Memory". In: *IEEE Transactions on Information Theory* 52.8 (Aug. 2006), pp. 3498–3508 (cit. on pp. 100, 174).

[ALY00] R. Ahlswede, S.-Y. R. Li, and R. W. Yeung. "Network information flow". In: *IEEE Transactions on Information Theory* 46.4 (July 2000), pp. 1204–1216 (cit. on p. 41).

[AM00] S. M. Aji and R. J. McEliece. "The generalized distributive law". In: *IEEE Transactions on Information Theory* 46.2 (Mar. 2000), pp. 325–343 (cit. on p. 54).

[And+08] J. Andrews et al. "Rethinking information theory for mobile ad hoc networks". In: *IEEE Communications Magazine* 46.12 (Dec. 2008), pp. 94–101 (cit. on pp. 2, 10, 21, 43).

[Aus67] Martin Edward Austin. "Decision-feedback equalization for digital communication over dispersive channels." In: *Technical Report 461* (Aug. 1967) (cit. on p. 103).

[AWH07] J. G. Andrews, S. Weber, and M. Haenggi. "Ad Hoc Networks: To Spread or Not to Spread? [Ad Hoc and Sensor Networks]". In: *IEEE Communications Magazine* 45.12 (Dec. 2007), pp. 84–91 (cit. on p. 23).

[AWW05] Ian F. Akyildiz, Xudong Wang, and Weilin Wang. "Wireless mesh networks: a survey". In: *Computer Networks* 47.4 (2005), pp. 445 –487 (cit. on p. 20).

[Bah+74] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv. "Optimal decoding of linear codes for minimizing symbol error rate". In: *IEEE Transactions on Information Theory* 20.2 (Mar. 1974), pp. 284–287 (cit. on pp. 51, 54, 58, 98, 100).

[Bal99] Jaiganesh Balakrishnan. "Mitigation of Error Propagation in Decision Feedback Equalization". Master's Thesis. Cornell Univ., 1999 (cit. on pp. 116, 119, 121, 125, 144).

[Baz+17] A. Bazzi, B. M. Masini, A. Zanella, and I. Thibault. "On the Performance of IEEE 802.11p and LTE-V2V for the Cooperative Awareness of Connected Vehicles". In: *IEEE Transactions on Vehicular Technology* 66.11 (Nov. 2017), pp. 10419–10432 (cit. on pp. 20, 160).

[BC02] J. Boutros and G. Caire. "Iterative multiuser joint decoding: unified framework and asymptotic analysis". In: *IEEE Transactions on Information Theory* 48.7 (July 2002), pp. 1772–1793 (cit. on pp. 48, 62, 66, 101, 109, 111, 120, 165, 222).

[BCT15] N. Benvenuto, S. Ciccotosto, and S. Tomasin. "Iterative Block Fractionally Spaced Nonlinear Equalization for Wideband Channels". In: *IEEE Wireless Communications Letters* 4.5 (Oct. 2015), pp. 489–492 (cit. on p. 192).

[BDF+10] N. Benvenuto, R. Dinis, D. Falconer, et al. "Single Carrier Modulation With Nonlinear Frequency Domain Equalization: An Idea Whose Time Has Come Again". In: *Proc. of the IEEE* 98.1 (Jan. 2010), pp. 69–96 (cit. on p. 165).

[BDH09] Y. Bar-Shalom, F. Daum, and J. Huang. "The probabilistic data association filter". In: *IEEE Control Systems Magazine* 29.6 (Dec. 2009), pp. 82–100 (cit. on p. 78).

[Bel+79] C. A. Belfiore et al. "Decision feedback equalization". In: *Proc. of the IEEE* 67.8 (Aug. 1979), pp. 1143–1156 (cit. on pp. 75, 101, 103, 137, 141, 157).

[BF98] Gerhard Bauch and Volker Franz. "A comparison of soft-in/soft-out algorithms for turbo detection". In: *Proc. Int. Conf. on Telecomm.* 1998, pp. 259–263 (cit. on p. 156).

[BFY04] J. Boyer, D. D. Falconer, and H. Yanikomeroglu. "Multihop diversity in wireless relaying channels". In: *IEEE Transactions on Communications* 52.10 (Oct. 2004), pp. 1820–1830 (cit. on pp. 21, 22).

[BGT93] C. Berrou, A. Glavieux, and P. Thitimajshima. "Near Shannon limit error-correcting coding and decoding: Turbo-codes. 1". In: *Proceedings of ICC '93 - IEEE International Conference on Communications*. Vol. 2. May 1993, 1064–1070 vol.2 (cit. on pp. 2, 11, 21, 48, 91, 109).

[Bla+07] A. Blair, T. Brown, K. M. Chugg, and M. Johnson. "Tactical Mobile Mesh Network System Design". In: *MILCOM 2007 - IEEE Military Communications Conference*. Oct. 2007, pp. 1–7 (cit. on p. 21).

[BM11] M. Bayati and A. Montanari. "The Dynamics of Message Passing on Dense Graphs, with Applications to Compressed Sensing". In: *IEEE Transactions on Information Theory* 57.2 (Feb. 2011), pp. 764–785 (cit. on p. 86).

[Bri99] S. ten Brink. "Convergence of iterative decoding". In: *Electronics Letters* 35.10 (May 1999), pp. 806–808 (cit. on p. 63).

[Bru+05] K. Brueninghaus, D. Astely, T. Salzer, S. Visuri, A. Alexiou, S. Karger, and G. A. Seraji. "Link performance models for system level simulations of broadband radio access systems". In: *IEEE PIMRC'05*. Vol. 4. Sept. 2005, 2306–2311 Vol. 4 (cit. on pp. 4, 12, 221, 222, 231).

[BS08] C. Bae and W. E. Stark. "Energy-bandwidth tradeoff with spatial reuse in wireless multi-hop networks". In: *IEEE Proceedings of MILCOM 2008*. Nov. 2008, pp. 1–7 (cit. on p. 24).

[BS16] M. Borgerding and P. Schniter. "Onsager-corrected deep learning for sparse linear inverse problems". In: *2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. Dec. 2016, pp. 227–231 (cit. on p. 185).

[BSR17] M. Borgerding, P. Schniter, and S. Rangan. "AMP-Inspired Deep Networks for Sparse Linear Inverse Problems". In: *IEEE Transactions on Signal Processing* 65.16 (Aug. 2017), pp. 4293–4308 (cit. on pp. 185, 186, 254, 258).

[BSY98] S. ten Brink, J. Speidel, and Ran-Hong Yan. "Iterative demapping and decoding for multilevel modulation". In: *IEEE GLOBECOM 1998*. Vol. 1. Nov. 1998, 579–584 vol.1 (cit. on pp. 48, 62, 65, 101, 223).

[BT02] N. Benvenuto and S. Tomasin. "On the comparison between OFDM and single carrier modulation with a DFE using a frequency-domain feedforward filter". In: *IEEE Transactions on Communications* 50.6 (June 2002), pp. 947–955 (cit. on pp. 164, 165).

[BT05] N. Benvenuto and S. Tomasin. "Iterative design and detection of a DFE in the frequency domain". In: *IEEE Transactions on Communications* 53.11 (Nov. 2005), pp. 1867–1875 (cit. on pp. 4, 12, 75, 78, 164, 165, 208).

[BTV07] N. Benvenuto, S. Tomasin, and D. Veronesi. "Multiple Frequency Offsets Estimation and Compensation for Cooperative Networks". In: *2007 IEEE Wireless Communications and Networking Conference*. Mar. 2007, pp. 891–895 (cit. on p. 239).

[Bur+06] J. L. Burbank, P. F. Chimento, B. K. Haberman, and W. T. Kasch. "Key Challenges of Military Tactical Networking and the Elusive Promise of MANET Technology". In: *IEEE Communications Magazine* 44.11 (Nov. 2006), pp. 39–45 (cit. on pp. 2, 10, 20).

[BY16] E. Balevi and A. Ö. Yilmaz. "Analysis of Frequency Domain Oversampled MMSE SC-FDE". In: *IEEE Communications Letters* 20.2 (Feb. 2016), pp. 232–235 (cit. on p. 193).

[C+14] J. Céspedes, P. M. Olmos, M. Sánchez-Fernández, and F. Perez-Cruz. "Expectation Propagation Detection for High-Order High-Dimensional MIMO Systems". In: *IEEE Transactions on Communications* 62.8 (Aug. 2014), pp. 2840–2849 (cit. on pp. 77, 123, 137, 138, 171).

[Çak+16] B. Çakmak, M. Opper, B. H. Fleury, and O. Winther. "Self-averaging expectation propagation". In: *arXiv preprint arXiv:1608.06602* (2016) (cit. on p. 73).

[Cas+97] R. A. Casas, F. Lopez de Victoria, I. Fijalkow, P. Schniter, and T. J. Endrea. "On MMSE fractionally-spaced equalizer design". In: *Proceedings of 13th International Conference on Digital Signal Processing*. Vol. 1. July 1997, 395–398 vol.1 (cit. on p. 136).

[CB05] G. Colavolpe and A. Barbieri. "On MAP symbol detection for ISI channels using the Ungerboeck observation model". In: *IEEE Communications Letters* 9.8 (Aug. 2005), pp. 720–722 (cit. on pp. 58, 97, 100).

[CB18] A. Caciularu and D. Burshtein. "Blind Channel Equalization Using Variational Autoencoders". In: *2018 IEEE International Conference on Communications Workshops (ICC Workshops)*. May 2018, pp. 1–6 (cit. on p. 88).

[CB19] Avi Caciularu and David Burshtein. *Unsupervised Linear and Nonlinear Channel Equalization and Decoding using Variational Autoencoders*. 2019. arXiv: 1905.08795 [cs.LG] (cit. on p. 87).

[CFP11] G. Colavolpe, D. Fertonani, and A. Piemontese. "SISO Detection Over Linear Channels With Linear Complexity in the Number of Interferers". In: *IEEE Journal of Selected Topics in Signal Processing* 5.8 (Dec. 2011), pp. 1475–1485 (cit. on pp. 101, 112).

[CFR01] G. Colavolpe, G. Ferrari, and R. Raheli. "Reduced-state BCJR-type algorithms". In: *IEEE Journal on Selected Areas in Communications* 19.5 (May 2001), pp. 848–859 (cit. on p. 101).

[CG05] G. Colavolpe and G. Germi. "On the application of factor graphs and the sum-product algorithm to ISI channels". In: *IEEE Transactions on Communications* 53.5 (May 2005), pp. 818–825 (cit. on pp. 101, 112).

[CG79] T. Cover and A. E. Gamal. "Capacity theorems for the relay channel". In: *IEEE Transactions on Information Theory* 25.5 (Sept. 1979), pp. 572–584 (cit. on p. 41).

[ÇGT13] Ç. Çapar, D. Goeckel, and D. Towsley. "Broadcast Analysis for Extended Cooperative Wireless Networks". In: *IEEE Transactions on Information Theory* 59.9 (Sept. 2013), pp. 5805–5810 (cit. on p. 239).

[Cha+19] T. Chang, T. Watteyne, X. Vilajosana, and P. H. Gomes. "Constructive Interference in 802.15.4: A Tutorial". In: *IEEE Communications Surveys Tutorials* 21.1 (2019), pp. 217–237 (cit. on pp. 2, 10, 21, 45).

[Cha19] Romain Chayot. "Synchronisation, détection et égalisation de modulation à phase continue dans des canaux sélectifs en temps et en fréquence". Thèse de doctorat dirigée par Poulliat, Charly et Boucheret, Marie-Laure Réseaux, Télécommunications, Systèmes et Architecture Toulouse, INPT 2018. PhD thesis. 2019 (cit. on pp. 193, 195, 196).

[Che+15] Z. Chen, Y. R. Zheng, J. Tao, J. Wang, and J. Song. "Frequency Domain Turbo Equalization under MMSE Criterion for Single Carrier MIMO Systems". In: *GLOBECOM'15*. Dec. 2015, pp. 1–5 (cit. on pp. 164, 165).

[Cho+08] Jun Won Choi, A. C. Singer, Jung Woo Lee, and Nam Ik Cho. "An improved soft feedback V-BLAST detection technique for TURBO-MIMO systems". In: *Proc. IEEE ICASSP'08*. Mar. 2008, pp. 3181–3184 (cit. on p. 116).

[Cio08] J. M. Cioffi. *EE379A - Chapter 3. Equalization*. Stanford University, 2008 (cit. on pp. 75, 105, 117, 192).

[Cio+95] J. M. Cioffi, G. P. Dudevoir, M. Vedat Eyuboglu, and G. D. Forney. "MMSE decision-feedback equalizers and coding. I. Equalization results". In: *IEEE Transactions on Communications* 43.10 (Oct. 1995), pp. 2582–2594 (cit. on p. 101).

[Cip+12] A. M. Cipriano, P. Agostini, A. Blad, and R. Knopp. "Cooperative communications with HARQ in a wireless mesh network based on 3GPP LTE". In: *2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*. Aug. 2012, pp. 1004–1008 (cit. on p. 21).

[Cla+03] Thomas Clausen, Philippe Jacquet, Cédric Adjih, Anis Laouiti, Pascale Minet, Paul Muhlethaler, Amir Qayyum, and Laurent Viennot. *Optimized Link State Routing Protocol (OLSR)*. Network Working Group. 2003 (cit. on pp. 23, 43).

[Cle+16] B. Clerckx, H. Joudeh, C. Hao, M. Dai, and B. Rassouli. "Rate splitting for MIMO wireless networks: a promising PHY-layer strategy for LTE evolution". In: *IEEE Communications Magazine* 54.5 (May 2016), pp. 98–105 (cit. on p. 40).

[ÇO18] B. Çakmak and M. Opper. "Expectation Propagation for Approximate Inference: Free Probability Framework". In: *2018 IEEE International Symposium on Information Theory (ISIT)*. June 2018, pp. 1276–1280 (cit. on pp. 73, 224).

[Cos83] M. Costa. "Writing on dirty paper (Corresp.)" In: *IEEE Transactions on Information Theory* 29.3 (May 1983), pp. 439–441 (cit. on p. 40).

[CPC02] J. L. Cromwell, G. Paparisto, and K. M. Chugg. "On the design and hardware demonstration of a robust, high-speed frequency-hopped radio for severe battlefield channels". In: *MILCOM 2002. Proceedings*. Vol. 2. Oct. 2002, 900–904 vol.2 (cit. on pp. 3, 11, 21).

[CR14] A. Chockalingam and B. Sundar Rajan. *Large MIMO Systems*. Cambridge University Press, 2014 (cit. on p. 54).

[CT12] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. John Wiley & Sons, 2012 (cit. on pp. 37, 38).

[CTB98] G. Caire, G. Taricco, and E. Biglieri. "Bit-interleaved coded modulation". In: *IEEE Transactions on Information Theory* 44.3 (May 1998), pp. 927–946 (cit. on pp. 3, 11, 21, 31, 39, 49, 51).

[CTK16] M. Cardone, D. Tuninetti, and R. Knopp. "On the Optimality of Simple Schedules for Networks With Multiple Half-Duplex Relays". In: *IEEE Transactions on Information Theory* 62.7 (July 2016), pp. 4120–4134 (cit. on p. 42).

[CVB04] S. Chtourou, R. Visoz, and A. O Berthet. "A class of low complexity iterative equalizers for space-time BICM over MIMO block fading multipath AWGN channel". In: *Proceedings of the IEEE ICC'04*. Vol. 1. 2004, pp. 618–624 (cit. on pp. 48, 75).

[CVS08] A. M. Cipriano, R. Visoz, and T. Salzer. "Calibration Issues of PHY Layer Abstractions for Wireless Broadband Systems". In: *IEEE Proc. in VTC'08*. Sept. 2008, pp. 1–5 (cit. on p. 222).

[Dö+18] S. Dörner, S. Cammerer, J. Hoydis, and S. t. Brink. "Deep Learning Based Communication Over the Air". In: *IEEE Journal of Selected Topics in Signal Processing* 12.1 (Feb. 2018), pp. 132–143 (cit. on p. 87).

[DD95] S Dolinar and D Divsalar. "Weight Distributions for Turbo Codes Using Random and Nonrandom Permutations". In: *Telecommunications and Data Acquisition Progress Report* 122 (Aug. 1995), pp. 56–65 (cit. on p. 32).

[DDDM04] I. Daubechies, M. Defrise, and C. De Mol. "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint". In: *Communications on Pure and Applied Mathematics* 57.11 (2004), pp. 1413–1457. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cpa.20042 (cit. on p. 80).

[Dec01] D. Declercq. "Comparison of structures for joint equalization and decoding". In: *IEEE International Conference on Acoustics, Speech, and Signal Processing*. Vol. 4. May 2001, 2585–2588 vol.4 (cit. on pp. 97, 100).

[DHD12] A. Dremeau, C. Herzet, and L. Daudet. "Boltzmann Machine and Mean-Field Approximation for Structured Sparse Decompositions". In: *IEEE Transactions on Signal Processing* 60.7 (July 2012), pp. 3425–3438 (cit. on pp. 81, 179).

[DHS11] John Duchi, Elad Hazan, and Yoram Singer. "Adaptive subgradient methods for online learning and stochastic optimization". In: *Journal of Machine Learning Research* 12.07 (2011), pp. 2121–2159 (cit. on p. 89).

[DJB+95] Catherine Douillard, Michel Jézéquel, Claude Berrou, et al. "Iterative correction of intersymbol interference: Turbo-equalization". en. In: *Eur. Trans. Telecomm.* 6.5 (Sept. 1995), pp. 507–511 (cit. on pp. 58, 91, 98, 100, 156, 165).

[DKP16] G. Durisi, T. Koch, and P. Popovski. "Toward Massive, Ultrareliable, and Low-Latency Wireless Communication With Short Packets". In: *Proceedings of the IEEE* 104.9 (Sept. 2016), pp. 1711–1726 (cit. on p. 234).

[DL10] Mischa Dohler and Yonghui Li. *Cooperative communications: hardware, channel and PHY*. John Wiley & Sons, 2010 (cit. on pp. 21, 25–27).

[DMM09] David L. Donoho, Arian Maleki, and Andrea Montanari. "Message-passing algorithms for compressed sensing". In: *Proceedings of the National Academy of Sciences* 106.45 (2009), pp. 18914–18919 (cit. on pp. 49, 81).

[DPB13] Nikos Dimitriou, Andreas Polydoros, and Ahmed Barnawi. "Cooperative Schemes for Path Establishment in Mobile Ad-hoc Networks Under Shadow-fading". In: *Ad Hoc Netw.* 11.8 (Nov. 2013), pp. 2556–2566 (cit. on pp. 21, 28).

[DPP15] Z. Ding, M. Peng, and H. V. Poor. "Cooperative Non-Orthogonal Multiple Access in 5G Systems". In: *IEEE Communications Letters* 19.8 (Aug. 2015), pp. 1462–1465 (cit. on pp. 2, 10, 42).

[DV02] A. Dejonghe and L. Vandendorpe. "Turbo-equalization for multilevel modulation: an efficient low-complexity scheme". In: *Proc. IEEE Int. Conf. on Commun.* Vol. 3. Apr. 2002, pp. 1863–1867 (cit. on pp. 109, 110).

[Elm12] George F Elmasry. *Tactical wireless communications and networks: design concepts and challenges*. John Wiley & Sons, 2012 (cit. on pp. 20, 23).

[Eri+14] G. Eriksson, K. Fors, K. Wiklundh, and U. Sterner. "Stochastic channel model for simulation of mobile ad hoc networks". In: *The Journal of Engineering* 2014.12 (2014), pp. 691–698 (cit. on pp. 26, 29, 30).

[ESN13] M. A. Elgenedy, E. Sourour, and M. Nafie. "Iterative MMSE-DFE equalizer for the high data rates HF waveforms in the HF channel". In: *Proceedings of the Asilomar'13*. Nov. 2013, pp. 1243–1247 (cit. on p. 115).

[Eyu88] M. V. Eyuboglu. "Detection of coded modulation signals on linear, severely distorted channels using decision-feedback noise prediction with interleaving". In: *IEEE Transactions on Communications* 36.4 (Apr. 1988), pp. 401–409 (cit. on p. 108).

[FA98] V. Franz and J. B. Anderson. "Concatenated decoding with a reduced-search BCJR algorithm". In: *IEEE Journal on Selected Areas in Communications* 16.2 (Feb. 1998), pp. 186–195 (cit. on p. 101).

[Fal+02] D. Falconer, S. L. Ariyavisitakul, A. Benyamin-Seeyar, and B. Eidson. "Frequency domain equalization for single-carrier broadband wireless systems". In: *IEEE Communications Magazine* 40.4 (Apr. 2002), pp. 58–66 (cit. on pp. 164, 165).

[FF73] D. D. Falconer and G. J. Foschini. "Theory of minimum mean-square-error QAM systems employing decision feedback equalization". In: *The Bell System Technical Journal* 52.10 (Dec. 1973), pp. 1821–1849 (cit. on p. 103).

[FH15] Hiroyuki Fukumoto and Kazunori Hayashi. "Overlap frequency domain equalization for faster-than-Nyquist signaling". In: *arXiv preprint arXiv:1509.00562* (2015) (cit. on pp. 196, 197).

[Fij+00] I. Fijalkow, A. Roumy, S. Ronger, D. Pirez, and P. Vila. "Improved interference cancellation for turbo-equalization". In: *Proc. IEEE Int. Conf. on Acoust., Speech, and Signal Process.* Vol. 1. June 2000, pp. 416–419 (cit. on pp. 48, 109).

[Fis+13a] J. Fischer, F. Beer, H. Lieske, J. Robert, J. Thielecke, and A. Heuberger. "A stochastic geometry approach to model the shadow fading for mobile ad-hoc networks". In: *2013 10th Annual Conference on Wireless On-demand Network Systems and Services (WONS)*. Mar. 2013, pp. 119–121 (cit. on pp. 29, 220).

[Fis+13b] J. Fischer, F. Beer, H. Lieske, J. Robert, J. Thielecke, and A. Heuberger. "On the Importance of Realistic Correlated Shadow Fading Modeling for Mobile Ad-Hoc Networks". In: *2013 First International Symposium on Computing and Networking*. Dec. 2013, pp. 468–471 (cit. on pp. 25, 29).

[Fis16] J. Fischer. "Physical Layer Link Modeling for Mobile Ad-Hoc Networks (MANET)". Ph.D. dissertation. Nürnberg: Friedrich-Alexander-Universität, June 2016 (cit. on pp. 20, 26–31, 199, 200).

[FL15] H. Farès and C. Langlais. "Finite-signal-to-noise ratio diversity-multiplexing-delay tradeoff in half-duplex hybrid automatic repeat request relay channels". In: *IET Communications* 9.6 (2015), pp. 872–879 (cit. on p. 40).

[For72] G. Forney. "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference". In: *IEEE Transactions on Information Theory* 18.3 (May 1972), pp. 363–378 (cit. on p. 98).

[Fre+98] Pål Frenger et al. *Multi-rate Convolutional Codes*. Tech. rep. Göteborg, Sweden: Chalmers University of Technology, 1998 (cit. on pp. 114, 233).

[FSA04] Ingo Forkel, Marc Schinnenburg, and Markus Ang. "Generation of two-dimensional correlated shadowing for mobile radio network simulation". In: *Wireless Personal Multimedia Communication*. Vol. 21. Sept. 2004, p. 43 (cit. on p. 29).

[Fu05] M. Fu. "Stochastic analysis of turbo decoding". In: *IEEE Transactions on Information Theory* 51.1 (Jan. 2005), pp. 81–100 (cit. on pp. 66, 149, 150, 230).

[FXC13] H. Feng, Y. Xiao, and L. J. Cimini. "Spectral Efficiency of Centralized and Decentralized Cooperative Networks with Relay Selection". In: *MILCOM 2013 - 2013 IEEE Military Communications Conference*. Nov. 2013, pp. 7–12 (cit. on p. 43).

[Gal62] R. Gallager. “Low-density parity-check codes”. In: *IRE Transactions on Information Theory* 8.1 (Jan. 1962), pp. 21–28 (cit. on pp. 48, 54).

[Gav18] Jérôme Gaveau. “Allocation des Ressources pour la Gestion Dynamique du Spectre dans les Réseaux Ad hoc Clustérisés”. Ph.D. dissertation. Paris: Réseaux, information et communications, Paris Saclay, 2018 (cit. on p. 21).

[GH11] Q. Guo and D. D. Huang. “A Concise Representation for the Soft-in Soft-out LMMSE Detector”. In: *IEEE Communications Letters* 15.5 (May 2011), pp. 566–568 (cit. on pp. 111, 122).

[GHK16] Jimmi Grönkvist, Anders Hansson, and Arwid Komulainen. “Performance analysis of reuse distance in cooperative broadcasting”. In: *Journal of Computer Networks and Communications* 2016 (2016) (cit. on p. 44).

[GK00] P. Gupta and P. R. Kumar. “The capacity of wireless networks”. In: *IEEE Transactions on Information Theory* 46.2 (Mar. 2000), pp. 388–404 (cit. on pp. 21, 22).

[GL10] Karol Gregor and Yann LeCun. “Learning Fast Approximations of Sparse Coding”. In: *Proceedings of the 27th International Conference on International Conference on Machine Learning*. ICML’10. Haifa, Israel: Omnipress, 2010, pp. 399–406 (cit. on pp. 90, 185).

[GL81] A. Gersho and T. L. Lim. “Adaptive cancellation of intersymbol interference for data transmission”. In: *The Bell System Technical Journal* 60.9 (Nov. 1981), pp. 1997–2021 (cit. on pp. 107, 109).

[GLL97] Alain Glavieux, Christophe Laot, and Joël Labat. “Turbo equalization over a frequency selective channel”. In: *Proc. 1st Symp. Turbo Codes*. 1997, pp. 96–102 (cit. on pp. 3, 11, 48, 101, 109).

[GM08] M. Grossmann and T. Matsumoto. “Nonlinear frequency domain MMSE turbo equalization using probabilistic data association”. In: *IEEE Communications Letters* 12.4 (Apr. 2008), pp. 295–297 (cit. on pp. 78, 92, 164–166, 175, 179).

[GP08] Q. Guo and L. Ping. “LMMSE turbo equalization based on factor graphs”. In: *IEEE Journal on Selected Areas in Communications* 26.2 (Feb. 2008), pp. 311–319 (cit. on pp. 111, 112).

[GRW12] N. Goddemeier, S. Rohde, and C. Wietfeld. “Experimental validation of RSS driven UAV mobility behaviors in IEEE 802.11s networks”. In: *2012 IEEE Globecom Workshops*. Dec. 2012, pp. 1550–1555 (cit. on p. 27).

[GT01] M. Grossglauser and D. Tse. “Mobility increases the capacity of ad-hoc wireless networks”. In: *Proceedings IEEE INFOCOM 2001. Conference on Computer Communications. Twentieth Annual Joint Conference of the IEEE Computer and Communications Society (Cat. No.01CH37213)*. Vol. 3. Apr. 2001, 1360–1369 vol.3 (cit. on p. 21).

[Guo+13] Q. Guo, D. (. Huang, S. Nordholm, J. Xi, and Y. Yu. "Iterative Frequency Domain Equalization With Generalized Approximate Message Passing". In: *IEEE Signal Processing Letters* 20.6 (June 2013), pp. 559–562 (cit. on pp. 3, 11, 80, 92, 164, 166, 179, 181).

[GVL96] Gene H. Golub and Charles F. Van Loan. *Matrix computations*. Vol. 3. JHU Press, 1996 (cit. on pp. 96, 124, 127, 128).

[GW02] A. J. Goldsmith and S. B. Wicker. "Design challenges for energy-constrained ad hoc wireless networks". In: *IEEE Wireless Communications* 9.4 (Aug. 2002), pp. 8–27 (cit. on pp. 2, 10, 20).

[GW15a] N. Goddemeier and C. Wietfeld. "Investigation of Air-to-Air Channel Characteristics and a UAV Specific Extension to the Rice Model". In: *2015 IEEE Globecom Workshops (GC Wkshps)*. Dec. 2015, pp. 1–5 (cit. on p. 30).

[GW15b] N. Goddemeier and C. Wietfeld. "Investigation of Air-to-Air Channel Characteristics and a UAV Specific Extension to the Rice Model". In: *IEEE Globecom Workshops*. Dec. 2015, pp. 1–5 (cit. on p. 248).

[Haa00] J. C. Haartsen. "The Bluetooth radio system". In: *IEEE Personal Communications* 7.1 (Feb. 2000), pp. 28–36 (cit. on pp. 1, 9).

[Haa02] E. Haas. "Aeronautical channel modeling". In: *IEEE Transactions on Vehicular Technology* 51.2 (Mar. 2002), pp. 254–264 (cit. on p. 30).

[Hag04] J. Hagenauer. "The EXIT chart - introduction to extrinsic information transfer in iterative processing". In: *Proc. IEEE 12th European Signal Process. Conf.* Sept. 2004, pp. 1541–1548 (cit. on pp. 49, 52, 92).

[Han+18] T. L. Hansen, P. B. Jørgensen, M. Badiu, and B. H. Fleury. "An Iterative Receiver for OFDM With Sparsity-Based Parametric Channel Estimation". In: *IEEE Transactions on Signal Processing* 66.20 (Oct. 2018), pp. 5454–5469 (cit. on pp. 3, 11, 66, 253, 258).

[Hay14] Simon Haykin. *Adaptive Filter Theory*. 5th edition. England: Pearson Education, 2014. 907 pp. (cit. on p. 105).

[HC10] T. R. Halford and K. M. Chugg. "Barrage Relay Networks". In: *2010 Information Theory and Applications Workshop (ITA)*. Jan. 2010, pp. 1–8 (cit. on pp. 2, 10, 44).

[He+16] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep Residual Learning for Image Recognition". In: *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2016 (cit. on p. 89).

[He+18] H. He et al. "A Model-Driven Deep Learning Network for MIMO Detection". In: *Proc. IEEE GlobalSIP*. Nov. 2018, pp. 584–588 (cit. on p. 91).

[HEA11] L. Hanzo, M. El-Hajjar, and O. Alamri. "Near-Capacity Wireless Transceivers and Cooperative Communications in the MIMO Era: Evolution of Standards, Waveform Design, and Future Perspectives". In: *Proceedings of the IEEE* 99.8 (Aug. 2011), pp. 1343–1385 (cit. on pp. 2, 11, 21, 54).

[Hes04] Tom Heskes. "On the Uniqueness of Loopy Belief Propagation Fixed Points". In: *Neural Computation* 16.11 (2004), pp. 2379–2413. eprint: https://doi.org/10.1162/0899766041941943 (cit. on p. 74).

[HH15] M. Hussain and S. A. Hassan. "Performance of Multi-Hop Cooperative Networks Subject to Timing Synchronization Errors". In: *IEEE Transactions on Communications* 63.3 (Mar. 2015), pp. 655–666 (cit. on p. 239).

[HH89] J. Hagenauer and P. Hoeher. "A Viterbi algorithm with soft-decision outputs and its applications". In: *IEEE Global Telecommunications Conference*. Nov. 1989, 1680–1686 vol.3 (cit. on p. 51).

[HHK05] A. Hafslund, T. T. Hoang, and Ø. Kure. "Push-to-talk applications in mobile ad hoc networks". In: *2005 IEEE 61st Vehicular Technology Conference*. Vol. 4. May 2005, 2410–2414 Vol. 4 (cit. on pp. 20, 22).

[HI12] S. A. Hassan and M. A. Ingram. "On the modeling of randomized distributed cooperation for linear multi-hop networks". In: *Proceedings of the IEEE ICC'12*. June 2012, pp. 366–370 (cit. on p. 239).

[Hig+15] Kenichi Higuchi et al. "Non-orthogonal multiple access (NOMA) with successive interference cancellation for future radio access". In: *IEICE Trans. on Commun.* 98.3 (2015), pp. 403–414 (cit. on p. 201).

[HKW04] I. Hammerstrom, M. Kuhn, and A. Wittneben. "Cooperative diversity by relay phase rotations in block fading environments". In: *IEEE 5th Workshop on Signal Processing Advances in Wireless Communications, 2004.* July 2004, pp. 293–297 (cit. on p. 44).

[HL06] R. Hu and J. Li. "Practical Compress-Forward in User Cooperation: Wyner-Ziv Cooperation". In: *2006 IEEE International Symposium on Information Theory*. July 2006, pp. 489–493 (cit. on p. 42).

[HLW14] J. R. Hershey, J. Le Roux, and F. Weninger. *Deep unfolding: Model-based inspiration of novel deep architectures*. Tech. Rep. TR2014-117. Version 12.0.1. Mitsubishi Electric Research Labs, 2014 (cit. on pp. 3, 11, 49, 87, 90).

[HMH16] S. Hong, I. Marić, and D. Hui. "Short Message Noisy Network Coding With Sliding-Window Decoding for Half-Duplex Multihop Relay Networks". In: *IEEE Transactions on Wireless Communications* 15.10 (Oct. 2016), pp. 6676–6689 (cit. on pp. 2, 10, 21, 42).

[HN02] T. E. Hunter and A. Nosratinia. "Cooperation diversity through coding". In: *Proceedings IEEE International Symposium on Information Theory,* June 2002, pp. 220– (cit. on p. 42).

[Hor91] Kurt Hornik. "Approximation capabilities of multilayer feedforward networks". In: *Neural Networks* 4.2 (1991), pp. 251 –257 (cit. on p. 88).

[HR18] S. Hu and F. Rusek. "On the Design of Channel Shortening Demodulators for Iterative Receivers in Linear Vector Channels". In: *IEEE Access* 6 (2018), pp. 48339–48359 (cit. on p. 101).

[HZ02] Tom Heskes and Onno Zoeter. "Expectation Propagation for Approximate Inference in Dynamic Bayesian Networks". In: *Proceedings of the Eighteenth Conference on Uncertainty in Artificial Intelligence*. UAI'02. Alberta, Canada: Morgan Kaufmann Publishers Inc., 2002, pp. 216–223 (cit. on p. 72).

[IB10] A. Ibing and H. Boche. "On Predicting Convergence of Iterative MIMO Detection-Decoding With Concatenated Codes". In: *IEEE Transactions on Vehicular Technology* 59.8 (Oct. 2010), pp. 4134–4139 (cit. on pp. 66, 149).

[Ieea] *IEEE 802.16 Broadband Wireless Access Working Group; MBS Frame Structure Considerations; Prorposal by W. Lee, D. Kim, J. Koo*. Standard. IEEE, Jan. 2008 (cit. on p. 221).

[Ieeb] *IEEE 802.16 Evaluation Methodology Document; Sources by R. Srinivasan, J. Zhuang, L. Jalloul, R. Novak, J. Park*. Standard. IEEE, July 2008 (cit. on p. 221).

[Isk08] Cyril-Daniel Iskander. "A MATLAB-based object-oriented approach to multipath fading channel simulation". In: *Hi-Tek Multisystems* (2008) (cit. on p. 261).

[Itua] *Method for point-to-area predictions for terrestrial services in the frequency range 30 MHz to 3 000 MHz*. Standard. Geneva, CH: International Telecommunication Union, Sept. 2013 (cit. on pp. 27, 28).

[Itub] *Propagation data and prediction methods for the planning of short-range outdoor radiocommunication systems and radio local area networks in the frequency range 300 MHz to 100 GHz*. Standard. Geneva, CH: International Telecommunication Union, June 2017 (cit. on p. 27).

[Jak+06] G. Jakllari, S. V. Krishnamurthy, M. Faloutsos, P. V. Krishnamurthy, and O. Ercetin. "A Framework for Distributed Spatio-Temporal Communications in Mobile Ad Hoc Networks". In: *Proceedings IEEE INFOCOM 2006. 25TH IEEE International Conference on Computer Communications*. Apr. 2006, pp. 1–13 (cit. on pp. 21, 44).

[Jen+10] T. L. Jensen, S. Kant, J. Wehinger, and B. H. Fleury. "Fast Link Adaptation for MIMO OFDM". In: *IEEE Transactions on Vehicular Technology* 59.8 (2010), pp. 3766–3778 (cit. on pp. 221, 222).

[Jeo11] Seongwook Jeong. "Low complexity turbo equalizations and lower bounds on information rate for intersymbol interference channels". Ph.D. dissertation. Minnesota: University of Minnesota, Oct. 2011 (cit. on pp. 109, 114, 253, 257).

[JM10] S. Jeong and J. Moon. "Turbo Equalization Based on Bi-Directional DFE". In: *IEEE Int. Conf. on Commun.* May 2010, pp. 1–6 (cit. on pp. 116, 119, 121, 125, 137).

[JM11] S. Jeong and J. Moon. "Soft-In Soft-Out DFE and Bi-Directional DFE". In: *IEEE Transactions on Communications* 59.10 (Oct. 2011), pp. 2729–2741 (cit. on pp. 3, 12, 116, 119, 125, 137, 141).

[JM13] S. Jeong and J. Moon. "Self-Iterating Soft Equalizer". In: *IEEE Transactions on Communications* 61.9 (Sept. 2013), pp. 3697–3709 (cit. on pp. 116, 137, 141, 143, 154, 178, 253, 257).

[Kal95] G. K. Kaleh. "Channel equalization for block transmission systems". In: *IEEE Journal on Selected Areas of Communications* 13.1 (Jan. 1995), pp. 110–121 (cit. on pp. 4, 12, 102–104, 165).

[Kay93] Steven M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993 (cit. on pp. 48, 59, 101–103, 111, 117, 118, 120).

[KB15] D. P. Kingma and J. Ba. "Adam: a method for stochastic optimization". In: *Proc. Int. Conf. Learn. Represent.* May 2015 (cit. on pp. 89, 188).

[KFL01a] F. R. Kschischang, B. J. Frey, and H. . Loeliger. "Factor graphs and the sum-product algorithm". In: *IEEE Transactions on Information Theory* 47.2 (Feb. 2001), pp. 498–519 (cit. on pp. 49, 54, 68).

[KFL01b] F. R. Kschischang, B. J. Frey, and H. . Loeliger. "Factor graphs and the sum-product algorithm". In: *IEEE Transactions on Information Theory* 47.2 (Feb. 2001), pp. 498–519 (cit. on p. 54).

[KGG05] G. Kramer, M. Gastpar, and P. Gupta. "Cooperative strategies and capacity theorems for relay networks". In: *IEEE Transactions on Information Theory* 51.9 (Sept. 2005), pp. 3037–3063 (cit. on p. 41).

[KH00] R. Knopp and P. A. Humblet. "On coding for block fading channels". In: *IEEE Transactions on Information Theory* 46.1 (Jan. 2000), pp. 189–205 (cit. on pp. 39, 222).

[KH11] G. Kramer and J. Hou. "On message lengths for noisy network coding". In: *2011 IEEE Information Theory Workshop*. Oct. 2011, pp. 430–431 (cit. on p. 41).

[KHRR06] Alireza Keshavarz-Haddad, Vinay Ribeiro, and Rudolf Riedi. "Broadcast Capacity in Multihop Wireless Networks". In: *Proceedings of the 12th Annual International Conference on Mobile Computing and Networking*. MobiCom '06. Los Angeles, CA, USA: ACM, 2006, pp. 239–250 (cit. on p. 21).

[KK14] R. A. M. Khan and H. Karl. "MAC Protocols for Cooperative Diversity in Wireless LANs and Wireless Sensor Networks". In: *IEEE Communications Surveys Tutorials* 16.1 (2014), pp. 46–63 (cit. on pp. 20, 23).

[KM07] K. Kansanen and T. Matsumoto. "An analytical method for MMSE MIMO turbo equalizer EXIT chart computation". In: *IEEE Transactions on Wireless Communications* 6.1 (Jan. 2007), pp. 59–63 (cit. on p. 222).

[KMY+07] Gerhard Kramer, Ivana Marić, Roy D Yates, et al. "Cooperative communications". In: *Foundations and Trends® in Networking* 1.3–4 (2007), pp. 271–425 (cit. on pp. 21, 39, 41).

[Krz+14] F. Krzakala, A. Manoel, E. W. Tramel, and L. Zdeborová. "Variational free energies for compressed sensing". In: *2014 IEEE International Symposium on Information Theory*. June 2014, pp. 1499–1503 (cit. on pp. 81, 83).

[KSH12] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. "ImageNet Classification with Deep Convolutional Neural Networks". In: *Advances in Neural Information Processing Systems 25*. Ed. by F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger. Curran Associates, Inc., 2012, pp. 1097–1105 (cit. on p. 89).

[KSW03] A. Koppler, A. Spronger, and R. Weigel. "Combined frequency domain feedforward and turbo decision feedback equalization for single carrier W-LAN systems". In: *Proc. IEEE ICC'03*. Vol. 3. May 2003, 2119–2123 vol.3 (cit. on pp. 164, 165).

[Lam+03] M. Lampe, T. Giebel, H. Rohling, and W. Zirwas. "PER prediction for PHY mode selection in OFDM communication systems". In: *GLOBECOM '03. IEEE Global Telecommunications Conference*. Vol. 1. Dec. 2003, 25–29 Vol.1 (cit. on p. 222).

[Lan02] J Nicholas Laneman. "Cooperative diversity in wireless networks: Algorithms and architectures". PhD thesis. Massachusetts Institute of Technology, 2002 (cit. on p. 21).

[Lar+10] E. Larsen, L. Landmark, V. Pham, Ø. Kure, and P. E. Engelstad. "Optimized group communication for tactical military networks". In: *IEEE Military Communications Conference (MILCOM 2010)*. Oct. 2010, pp. 1905–1911 (cit. on pp. 20, 22).

[Lat98] Bhagwandas Pannalal Lathi. *Modern Digital and Analog Communication Systems 3e Osece*. Oxford University Press, Inc., 1998 (cit. on p. 34).

[LB03] Raphaël Le Bidan. "Turbo-equalization for bandwidth-efficient digital communications over frequency-selective channels". PhD thesis. INSA de Rennes, Jan. 2003 (cit. on pp. 48, 107–110).

[LB06] R. R. Lopes and J. R. Barry. "The soft-feedback equalizer for turbo equalization of highly dispersive channels". In: *IEEE Transactions on Communications* 54.5 (May 2006), pp. 783–788 (cit. on pp. 3, 12, 114, 116, 119, 120, 125, 144, 148).

[LC15a] S. Lee and S. Chung. "A unified approach for network information theory". In: *2015 IEEE International Symposium on Information Theory (ISIT)*. June 2015, pp. 1277–1281 (cit. on pp. 21, 41).

[LC15b] S. Lee and S. Chung. "Noisy network coding with partial DF". In: *2015 IEEE International Symposium on Information Theory (ISIT)*. June 2015, pp. 2870–2874 (cit. on p. 41).

[LeC+89] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel. "Backpropagation Applied to Handwritten Zip Code Recognition". In: *Neural Computation* 1.4 (1989), pp. 541–551. eprint: https://doi.org/10.1162/neco.1989.1.4.541 (cit. on p. 89).

[Let18] Xavier Leturc. "Resource allocation for HARQ in mobile ad hoc networks". Ph.D. dissertation. Paris: Réseaux, information et communications, Paris Saclay, 2018 (cit. on p. 23).

[LGL01] C. Laot, A. Glavieux, and J. Labat. "Turbo equalization: adaptive equalization and channel decoding jointly optimized". In: *IEEE Journal on Selected Areas in Communications* 19.9 (Sept. 2001), pp. 1744–1752 (cit. on pp. 48, 109).

[Li+12] Q. Li, R. Q. Hu, Y. Qian, and G. Wu. "Cooperative communications for wireless networks: techniques and applications in LTE-advanced systems". In: *IEEE Wireless Communications* 19.2 (Apr. 2012), pp. 22–29 (cit. on p. 21).

[Li+19] D. Li, Y. Wu, J. Tao, and M. Zhu. "Performance Analysis and Improvement for VAMP Soft Frequency-Domain Equalizers". In: *IEEE Access* 7 (Mar. 2019), pp. 42495–42506 (cit. on pp. 182, 223).

[Lim+11] S. H. Lim, Y. Kim, A. El Gamal, and S. Chung. "Noisy Network Coding". In: *IEEE Transactions on Information Theory* 57.5 (May 2011), pp. 3132–3152 (cit. on pp. 21, 41).

[Liu+19a] Hongwu Liu, Theodoros Tsiftsis, Kyeong Jin Kim, Kyung Sup Kwak, and H. Vincent Poor. "Rate Splitting for Asynchronous Uplink NOMA Systems with Cyclic Prefixed Single Carrier". In: *IEEE ICC, 5th Intrnational Workshop on NOMA5G*. May 2019 (cit. on pp. 40, 160, 254, 259).

[Liu+19b] Lei Liu, Chulong Liang, Junjie Ma, and Li Ping. "Capacity Optimality of AMP in Coded Systems". In: *CoRR* abs/1901.09559 (2019). arXiv: 1901.09559 (cit. on p. 86).

[LKK17] S. H. Lim, K. T. Kim, and Y. Kim. "Distributed Decode–Forward for Relay Networks". In: *IEEE Transactions on Information Theory* 63.7 (July 2017), pp. 4103–4118 (cit. on pp. 21, 41).

[LL04] Lihai Liu and Li Ping. "An extending window MMSE turbo equalization algorithm". In: *IEEE Signal Processing Letters* 11.11 (Nov. 2004), pp. 891–894 (cit. on p. 111).

[LLL03] Lihai Liu, W. K. Leung, and Li Ping. "Simple iterative chip-by-chip multiuser detection for CDMA systems". In: *The 57th IEEE Semiannual Vehicular Technology Conference, 2003. VTC 2003-Spring.* Vol. 3. Apr. 2003, 2157–2161 vol.3 (cit. on p. 110).

[LN09] N. H. Lu and C. Nelson. "Cooperative Team Communications". In: *MILCOM 2009 - 2009 IEEE Military Communications Conference*. Oct. 2009, pp. 1–6 (cit. on p. 20).

[LN+14] A. Le Naour, F. Pipon, B. Nahoum, D. Mérel, and B. Obrist. "Performance of the FASTNET Narrowband VHF Frequency Hopping Tactical Radio in Challenging Multipath Environments". In: *2014 IEEE Military Communications Conference*. Oct. 2014, pp. 649–656 (cit. on p. 25).

[Loe+07] H. Loeliger, J. Dauwels, J. Hu, S. Korl, L. Ping, and F. R. Kschischang. "The Factor Graph Approach to Model-Based Signal Processing". In: *Proceedings of the IEEE* 95.6 (June 2007), pp. 1295–1322 (cit. on pp. 54, 60, 111).

[LSS03] Seok-Jun Lee, A. C. Singer, and N. R. Shanbhag. "Analysis of linear turbo equalizer via EXIT chart". In: *IEEE Proc. in GLOBECOM'03*. Vol. 4. Dec. 2003, 2237–2242 vol.4 (cit. on p. 222).

[LTW04] J. N. Laneman, D. N. C. Tse, and G. W. Wornell. "Cooperative diversity in wireless networks: Efficient protocols and outage behavior". In: *IEEE Transactions on Information Theory* 50.12 (Dec. 2004), pp. 3062–3080 (cit. on pp. 21, 40, 42, 44).

[Lu+07] W. Lu, W. K. G. Seah, E. W. C. Peh, and Y. Ge. "Communications Support for Disaster Recovery Operations using Hybrid Mobile Ad-Hoc Networks". In: *32nd IEEE Conference on Local Computer Networks (LCN 2007)*. Oct. 2007, pp. 763–770 (cit. on pp. 1, 9).

[Luo+01] J. Luo, K. R. Pattipati, P. K. Willett, and F. Hasegawa. "Near-optimal multiuser detection in synchronous CDMA using probabilistic data association". In: *IEEE Communications Letters* 5.9 (Sept. 2001), pp. 361–363 (cit. on p. 78).

[LW03] J. N. Laneman and G. W. Wornell. "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks". In: *IEEE Transactions on Information Theory* 49.10 (Oct. 2003), pp. 2415–2425 (cit. on pp. 42, 45).

[LX11] H. Lou and C. Xiao. "Soft-Decision Feedback Turbo Equalization for Multilevel Modulations". In: *IEEE Transactions on Signal Processing* 59.1 (Jan. 2011), pp. 186–195 (cit. on pp. 3, 12, 116, 119, 125).

[LX14] Huang Lou and Chengshan Xiao. "The Soft-Feedback ISI Canceller-Based Turbo Equalizer for Multilevel Modulations". In: *Int. J. of Wireless Inform. Networks* 21.1 (Mar. 2014), pp. 68–73 (cit. on pp. 144, 155).

[Ma+15] X. Ma, C. Liang, K. Huang, and Q. Zhuang. "Block Markov Superposition Transmission: Construction of Big Convolutional Codes From Short Codes". In: *IEEE Transactions on Information Theory* 61.6 (June 2015), pp. 3150–3163 (cit. on p. 42).

[Ma+18] J. Ma, L. Liu, X. Yuan, and L. Ping. "Iterative Detection in Coded Linear Systems Based on Orthogonal AMP". In: *2018 IEEE 10th International Symposium on Turbo Codes Iterative Information Processing (ISTC)*. Dec. 2018, pp. 1–5 (cit. on pp. 86, 133, 223–225).

[Mar+03] L. Martoyo, T. Weiss, F. Capar, and F. K. Jondral. "Low complexity CDMA downlink receiver based on frequency domain equalization". In: *Proceedings of the IEEE VTC'03-Fall*. Vol. 2. Oct. 2003, pp. 987–991 (cit. on pp. 196, 197).

[MD10] A. Maleki and D. L. Donoho. "Optimally Tuned Iterative Reconstruction Algorithms for Compressed Sensing". In: *IEEE Journal of Selected Topics in Signal Processing* 4.2 (Apr. 2010), pp. 330–341 (cit. on p. 80).

[Meu71] Edward C. Van Der Meulen. "Three-Terminal Communication Channels". In: *Advances in Applied Probability* 3.1 (1971), pp. 120–154 (cit. on p. 41).

[Min01] T. P. Minka. "A family of algorithms for approximate Bayesian inference". Ph.D. dissertation. Cambridge: M.I.T., Jan. 2001 (cit. on pp. 49, 215).

[Min+05] Tom Minka et al. *Divergence measures and message passing*. Tech. rep. Microsoft Research, 2005 (cit. on pp. 37, 49, 67–69, 73, 74, 77, 138, 171).

[MLC17] R. Massin, C. J. Le Martret, and P. Ciblat. "A Coalition Formation Game for Distributed Node Clustering in Mobile Ad Hoc Networks". In: *IEEE Transactions on Wireless Communications* 16.6 (June 2017), pp. 3940–3952 (cit. on p. 21).

[MN96] D. J. C. MacKay and R. M. Neal. "Near Shannon limit performance of low density parity check codes". In: *Electronics Letters* 32.18 (Aug. 1996), pp. 1645– (cit. on p. 48).

[MP17] J. Ma and L. Ping. "Orthogonal AMP". In: *IEEE Access* 5 (Jan. 2017), pp. 2020–2033 (cit. on pp. 49, 83, 90, 181).

[MS81] M. S. Mueller and J. Salz. "A unified theory of data-aided equalization". In: *The Bell System Technical Journal* 60.9 (Nov. 1981), pp. 2023–2038 (cit. on p. 107).

[MVB16] A. Mohamad, R. Visoz, and A. O. Berthet. "Cooperative Incremental Redundancy Hybrid Automatic Repeat Request Strategies for Multi-Source Multi-Relay Wireless Networks". In: *IEEE Communications Letters* 20.9 (Sept. 2016), pp. 1808–1811 (cit. on p. 42).

[MWG+02] B. Muquet, Zhengdao Wang, G. B. Giannakis, et al. "Cyclic prefixing or zero padding for wireless multicarrier transmissions?" In: *IEEE Transactions on Communications* 50.12 (Dec. 2002), pp. 2136–2148 (cit. on pp. 191, 192).

[MY04] I. Maric and R. D. Yates. "Cooperative multihop broadcast for wireless networks". In: *IEEE Journal on Selected Areas in Communications* 22.6 (Aug. 2004), pp. 1080–1088 (cit. on pp. 2, 10, 44).

[Naj11] Ziad Naja. "Interprétation et amélioration d'une procédure de démodulation itérative". Thèse de doctorat dirigée par Duhamel, Pierre et Alberge, Florence Physique Paris 11 2011. PhD thesis. 2011 (cit. on pp. 3, 11).

[Nav18] Raphaël Naves. "Apport de la gestion des interférences aux réseaux sans-fil multi-sauts. Le cas du Physical-Layer Network Coding". Ph.D. dissertation. Toulouse: Réseaux, Télécommunications, Systèmes et Architecture - INPT, 2018 (cit. on pp. 254, 259).

[NBB16] E. Nachmani, Y. Be'ery, and D. Burshtein. "Learning to decode linear codes using deep learning". In: *54th Annual Allerton Conf. on Commun., Control, and Computing*. Sept. 2016, pp. 341–346 (cit. on pp. 87, 91).

[Ngu+15] H. V. Nguyen, C. Xu, S. X. Ng, and L. Hanzo. "Near-Capacity Wireless System Design Principles". In: *IEEE Communications Surveys Tutorials* 17.4 (2015), pp. 1806–1833 (cit. on pp. 2, 11, 21, 92).

[NLF07] B. Ng, C. T. Lam, and D. Falconer. "Turbo frequency domain equalization for single-carrier broadband wireless systems". In: *IEEE Transactions on Wireless Communications* 6.2 (Feb. 2007), pp. 759–767 (cit. on pp. 164, 165).

[NS12] J. Nilsson and U. Sterner. "Robust MPR-based flooding in mobile ad-hoc networks". In: *MILCOM 2012 - 2012 IEEE Military Communications Conference*. Oct. 2012, pp. 1–6 (cit. on p. 43).

[NS15] J. Nilsson and U. Sterner. "Performance of MPR-based broadcast methods in mobile ad hoc networks for rural terrains". In: *2015 International Conference on Military Communications and Information Systems (ICMCIS)*. May 2015, pp. 1–7 (cit. on p. 43).

[NVB12] B. Ning, R. Visoz, and A. O. Berthet. "Extrinsic versus a posteriori probability based iterative LMMSE-IC algorithms for coded MIMO communications: Performance and analysis". In: *Proc. IEEE Int. Symp. on Wireless Commun. Systems*. Aug. 2012, pp. 386–390 (cit. on pp. 66, 155, 232).

[NVB13] B. Ning, R. Visoz, and A. O. Berthet. "Physical layer abstraction for turbo coded MIMO systems with LMMSE-IC based turbo equalization". In: *Proc. IEEE Globecom Workshops*. Dec. 2013, pp. 310–315 (cit. on pp. 223, 231).

[OA11] T. Obara and F. Adachi. "MMSE weight for single-carrier overlap FDE". In: *Proceedings of the IEEE 17th Asia Pacific Conference on Communications*. Oct. 2011, pp. 168–172 (cit. on pp. 198, 199).

[OD13] A. Özgür and S. N. Diggavi. "Approximately Achieving Gaussian Relay Network Capacity With Lattice-Based QMF Codes". In: *IEEE Transactions on Information Theory* 59.12 (Dec. 2013), pp. 8275–8294 (cit. on p. 41).

[OH17] T. O'Shea and J. Hoydis. "An Introduction to Deep Learning for the Physical Layer". In: *IEEE Transactions on Cognitive Communications and Networking* 3.4 (Dec. 2017), pp. 563–575 (cit. on pp. 3, 11, 49, 87).

[OLT07] A. Ozgur, O. Leveque, and D. N. C. Tse. "Hierarchical Cooperation Achieves Optimal Capacity Scaling in Ad Hoc Networks". In: *IEEE Transactions on Information Theory* 53.10 (Oct. 2007), pp. 3549–3572 (cit. on pp. 21, 22).

[OT02] R. Otnes and M. Tuchler. "Low-complexity turbo equalization for time-varying channels". In: *IEEE 55th Vehicular Technology Conference. VTC Spring 2002*. Vol. 1. May 2002, 140–144 vol.1 (cit. on p. 60).

[OT04] R. Otnes and M. Tüchler. "Iterative channel estimation for turbo equalization of time-varying frequency-selective channels". In: *IEEE Transactions on Wireless Communications* 3.6 (Nov. 2004), pp. 1918–1923 (cit. on pp. 109, 110).

[OW05] Manfred Opper and Ole Winther. "Expectation consistent approximate inference". In: *Journal of Machine Learning Research* 6.12 (2005), pp. 2177–2204 (cit. on pp. 67, 72, 73).

[Pan+19] P. Pandit, M. Sahraee, S. Rangan, and A. K. Fletcher. "Asymptotics of MAP Inference in Deep Networks". In: *IEEE International Symposium on Information Theory (ISIT)*. June 2019 (cit. on p. 86).

[Pea88] Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1988 (cit. on p. 54).

[Per+15] Marcelo Pereyra, Philip Schniter, Emilie Chouzenoux, Jean-Christophe Pesquet, Jean-Yves Tourneret, Alfred O Hero, and Steve McLaughlin. "A survey of stochastic simulation and optimization methods in signal processing". In: *IEEE Journal of Selected Topics in Signal Processing* 10.2 (2015), pp. 224–241 (cit. on pp. 3, 11, 66, 68, 69, 71).

[Pet+19] Titouan Petitpied, Romain Tajan, Guillaume Ferré, Pascal Chevalier, and Sylvain Traverso. "Algorithme d'égalisation fractionnée fondé sur l'EP-MMSE-DFE". In: *27ème Colloque GRETSI sur le Traitement du Signal et des Images (GRETSI 2019)*. Lille, FR: GRETSI CNRS, 2019, pp. 1–4 (cit. on pp. 136, 253, 257).

[PGT98] T. Pham-Gia and N. Turkkan. "Distribution of the linear combination of two general beta variables and applications". In: *Communications in Statistics - Theory and Methods* 27:7 (1998), pp. 1851–1869 (cit. on p. 243).

[Pin+06] Li Ping, Lihai Liu, Keying Wu, and W. K. Leung. "Interleave division multiple-access". In: *IEEE Transactions on Wireless Communications* 5.4 (Apr. 2006), pp. 938–947 (cit. on pp. 201, 203).

[PPV10] Y. Polyanskiy, H. V. Poor, and S. Verdu. "Channel Coding Rate in the Finite Blocklength Regime". In: *IEEE Transactions on Information Theory* 56.5 (May 2010), pp. 2307–2359 (cit. on p. 234).

[Pre05] Marco Pretti. "A message-passing algorithm with damping". In: *Journal of Statistical Mechanics: Theory and Experiment* 2005.11 (Nov. 2005), P11008–P11008 (cit. on pp. 74, 75).

[Pro+94] John G Proakis, Masoud Salehi, Ning Zhou, and Xiaofeng Li. *Communication systems engineering*. Vol. 2. Prentice Hall New Jersey, 1994 (cit. on p. 108).

[PWW07] Li Ping, Peng Wang, and Xiaodong Wang. "Recent progress in interleave-division multiple-access (IDMA)". In: *MILCOM 2007-IEEE Military Communications Conference*. IEEE. 2007, pp. 1–7 (cit. on pp. 201, 203).

[QM07] Y. Qi and T. P. Minka. "Window-based expectation propagation for adaptive signal detection in flat-fading channels". In: *IEEE Transactions on Wireless Communications* 6.1 (Jan. 2007), pp. 348–355 (cit. on pp. 3, 11, 66, 137).

[QPL05] Qinghua Guo, Ping Li, and H. . Loeliger. "Turbo equalization based on factor graphs". In: *Proceedings. International Symposium on Information Theory, 2005. ISIT 2005.* Sept. 2005, pp. 2021–2025 (cit. on p. 111).

[Rad18] Bilel Raddadi. "Récepteurs avancés et nouvelles formes d'ondes pour les communications aéronautiques". Thèse de doctorat dirigée par Poulliat, Charly et Boucheret, Marie-Laure Réseaux, Télécommunications, Systèmes et Architecture Toulouse, INPT 2018. PhD thesis. 2018 (cit. on pp. 101, 193).

[Ram05] Ram Ramanathan. "Challenges: A Radically New Architecture for Next Generation Mobile Ad Hoc Networks". In: *Proceedings of the 11th Annual International Conference on Mobile Computing and Networking*. MobiCom '05. Cologne, Germany: ACM, 2005, pp. 132–139 (cit. on pp. 21, 23, 44, 218).

[Ran11] S. Rangan. "Generalized approximate message passing for estimation with random linear mixing". In: *IEEE Proc. in Int. Symp. on Information Theory (ISIT)*. July 2011, pp. 2168–2172 (cit. on pp. 49, 82).

[Rau13] R. D. Raut. "Characterizing the Capacity of Wireless Ad Hoc Networks". In: *2013 Third International Conference on Advanced Computing and Communication Technologies (ACCT)*. Apr. 2013, pp. 272–275 (cit. on p. 43).

[RBL05] L. Rugini, P. Banelli, and G. Leus. "Simple equalization of time-varying channels for OFDM". In: *IEEE Communications Letters* 9.7 (July 2005), pp. 619–621 (cit. on p. 195).

[RBL06] Luca Rugini, Paolo Banelli, and Geert Leus. "Low-Complexity Banded Equalizers for OFDM Systems in Doppler Spread Channels". In: *EURASIP Journal on Advances in Signal Processing* 2006.1 (2006), p. 067404 (cit. on p. 195).

[RGW18] Tim Rüegg, Filippo Gentile, and Armin Wittneben. "Cooperative Broadcast Performance Prediction Based on Inter-Node Distance Distributions". In: *Proc. of the IEEE WCNC'18*. Apr. 2018 (cit. on p. 239).

[Rib+15] F. C. Ribeiro, R. Dinis, F. Cercas, and A. Silva. "On the Theoretical BER Performance of SC-FDE Schemes with IB-DFE Receivers". In: *IEEE 81st Vehicular Technology Conference (VTC Spring)*. May 2015, pp. 1–5 (cit. on p. 233).

[Rie+13] E. Riegler, G. E. Kirkelund, C. N. Manchon, M. Badiu, and B. H. Fleury. "Merging Belief Propagation and the Mean Field Approximation: A Free Energy Approach". In: *IEEE Transactions on Information Theory* 59.1 (Jan. 2013), pp. 588–602 (cit. on pp. 3, 11, 66, 74, 166).

[RK18] T. Richardson and S. Kudekar. "Design of Low-Density Parity Check Codes for 5G New Radio". In: *IEEE Communications Magazine* 56.3 (Mar. 2018), pp. 28–34 (cit. on p. 32).

[RL09] William Ryan and Shu Lin. *Channel codes: classical and modern*. Cambridge university press, 2009 (cit. on pp. 32, 52, 57, 65).

[RMO14] Jack Raymond, Andre Manoel, and Manfred Opper. "Expectation propagation". In: *arXiv preprint arXiv:1409.6179* (2014) (cit. on p. 71).

[RNM17] Y. Ravaei, M. Mahdi Nejad, and M. H. Madani. "Low-complexity frequency-domain turbo equalisation for doubly-selective HF channel on GPP-based SDR platform". In: *IET Communications* 11.10 (July 2017), pp. 1649–1654 (cit. on pp. 196, 198, 199).

[Rou00] Aline Roumy. "Égalisation et décodage conjoints: méthodes turbo". PhD thesis. Univ. de Cergy-Pontoise, Oct. 2000 (cit. on pp. 3, 11, 104, 109).

[Rou+01] A. Roumy, A. J. Grant, I. Fijalkow, P. D. Alexander, and D. Pirez. "Turbo-equalization: convergence analysis". In: *IEEE Proc. in ICASSP'01*. Vol. 4. May 2001, 2645–2648 vol.4 (cit. on pp. 3, 11, 221).

[RSF14] S. Rangan, P. Schniter, and A. Fletcher. "On the convergence of approximate message passing with arbitrary matrices". In: *IEEE International Symposium on Information Theory*. June 2014, pp. 236–240 (cit. on p. 83).

[RSF17] S. Rangan, P. Schniter, and A. K. Fletcher. "Vector approximate message passing". In: *IEEE International Symposium on Information Theory (ISIT)*. June 2017, pp. 1588–1592 (cit. on pp. 85, 182).

[RSU01] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke. "Design of capacity-approaching irregular low-density parity-check codes". In: *IEEE Transactions on Information Theory* 47.2 (Feb. 2001), pp. 619–637 (cit. on pp. 48, 62, 222).

[RT07] Ram Ramanathan and F. Tchakountio. "Channel Access over Path Segments for Ultra Low Latency MANETs". In: *MILCOM 2007 - IEEE Military Communications Conference*. Oct. 2007, pp. 1–7 (cit. on pp. 21, 23).

[RV05] V. Ramon and L. Vandendorpe. "Predicting the performance and convergence behavior of a turbo-equalization scheme". In: *IEEE Proc. in ICASSP'05*. Vol. 3. Mar. 2005, iii/709–iii/712 Vol. 3 (cit. on pp. 221, 222).

[RW01] Daryl Reynolds and Xiaodong Wang. "Low-complexity Turbo-equalization for diversity channels". In: *Signal Processing* 81.5 (2001), pp. 989 –995 (cit. on p. 110).

[SA11] M. Senst and G. Ascheid. "A combined belief propagation and mean field algorithm for soft carrier phase estimation". In: *2011 8th International Symposium on Wireless Communication Systems*. Nov. 2011, pp. 512–516 (cit. on pp. 66, 92).

[SA12] M. Senst and G. Ascheid. "A message passing approach to iterative Bayesian SNR estimation". In: *2012 International Symposium on Signals, Systems, and Electronics (ISSSE)*. Oct. 2012, pp. 1–6 (cit. on p. 66).

[SBG69] NM Steen, GD Byrne, and EM Gelbard. "Gaussian quadratures for the integrals $\int_0^\infty e^{-x^2} f(x)dx$ and $\int_0^b e^{-x^2} f(x)dx$". In: *Math. of Comput.* 23.107 (July 1969), pp. 661–671 (cit. on p. 228).

[SC05] M. Sikora and D. J. Costello. "A new SISO algorithm with application to turbo equalization". In: *Proceedings of International Symposium on Information Theory*. Sept. 2005, pp. 2031–2035 (cit. on p. 101).

[Sch04] P. Schniter. "Low-complexity equalization of OFDM in doubly selective channels". In: *IEEE Transactions on Signal Processing* 52.4 (Apr. 2004), pp. 1002–1011 (cit. on p. 195).

[SDW17] N. Samuel, T. Diskin, and A. Wiesel. "Deep MIMO detection". In: *Proc. IEEE SPAWC*. July 2017, pp. 1–5 (cit. on p. 190).

[SDW19] N. Samuel, T. Diskin, and A. Wiesel. "Learning to Detect". In: *IEEE Transactions on Signal Processing* 67.10 (May 2019), pp. 2554–2564 (cit. on pp. 3, 11, 87, 90).

[SEA03] A. Sendonaris, E. Erkip, and B. Aazhang. "User cooperation diversity. Part I. System description". In: *IEEE Transactions on Communications* 51.11 (Nov. 2003), pp. 1927–1938 (cit. on p. 42).

[Sen+11] M. Senst et al. “How the Framework of Expectation Propagation Yields an Iterative IC-LMMSE MIMO Receiver”. In: *GLOBECOM’11*. Dec. 2011 (cit. on pp. 3, 11, 54, 66, 77, 92, 124, 137, 138, 140, 170).

[SF09] M. Sabbaghian and D. Falconer. “An Analytical Approach for Finite Block Length Performance Analysis of Turbo Frequency-Domain Equalization”. In: *IEEE Journal on Vehicular Technology* 58.3 (Mar. 2009), pp. 1292–1301 (cit. on pp. 147, 223).

[SF+11] C. Studer, S. Fateh, et al. “ASIC Implementation of Soft-Input Soft-Output MIMO Detection Using MMSE Parallel Interference Cancellation”. In: *IEEE Journal of Solid-State Circuits* 46.7 (July 2011), pp. 1754–1765 (cit. on p. 127).

[SGL06] A. Scaglione, D. L. Goeckel, and J. N. Laneman. “Cooperative communications in mobile ad hoc networks”. In: *IEEE Signal Processing Magazine* 23.5 (Sept. 2006), pp. 18–29 (cit. on p. 21).

[SH00] Mathini Sellathurai and Simon Haykin. “TURBO-BLAST for high-speed wireless communications”. In: *Proc. of the IEEE WCNC*. Vol. 1. 2000, pp. 315–320 (cit. on p. 202).

[SH03] A. Scaglione and Y.-W. Hong. “Opportunistic large arrays: cooperative transmission in wireless multihop ad hoc networks to reach far distances”. In: *IEEE Transactions on Signal Processing* 51.8 (Aug. 2003), pp. 2082–2092 (cit. on pp. 21, 44).

[Sha48] C. E. Shannon. “A mathematical theory of communication”. In: *The Bell System Technical Journal* 27.3 (July 1948), pp. 379–423 (cit. on pp. 38, 48).

[SKJ94] H. Sarı, G. Karam, and I. Jeanclaud. “Frequency-domain equalization of mobile radio and terrestrial broadcast channels”. In: *Proc. IEEE GLOBECOM’04*. Nov. 1994, 1–5 vol.1 (cit. on pp. 4, 12, 31, 160, 161, 164).

[SMF+17a] I. Santos, J. J. Murillo-Fuentes, et al. “Expectation Propagation as Turbo Equalizer in ISI Channels”. In: *IEEE Transactions on Communications* 65.1 (Jan. 2017), pp. 360–370 (cit. on pp. 137, 140, 141, 176, 215).

[SMF+17b] I. Santos, J. J. Murillo-Fuentes, et al. “Probabilistic Equalization With a Smoothing Expectation Propagation Approach”. In: *IEEE Transactions on Wireless Communications* 16.5 (May 2017), pp. 2950–2962 (cit. on pp. 137, 141, 253, 258).

[SMF+18] I. Santos, J. J. Murillo-Fuentes, et al. “Turbo EP-based Equalization: a Filter-Type Implementation”. In: *IEEE Transactions on Communications* (2018) (cit. on pp. 137–139, 141, 142, 176, 195).

[Sou+19] Eric Soubigou, Serdar Şahin, Antonio Maria Cipriano, Charly Poulliat, and Romain Chayot. “Analyse Multicritères des Performances et de la Complexité des Turbo-égaliseurs à Complexité Réduite à base de Treillis et de Filtres”. In: *27ème Colloque GRETSI sur le Traitement du Signal et des Images (GRETSI 2019)*. Lille, FR: GRETSI CNRS, 2019, pp. 1–4 (cit. on p. 18).

[SRF06] N. Sellami, A. Roumy, and I. Fijalkow. "The impact of both a priori information and channel estimation errors on the MAP equalizer performance". In: *IEEE Transactions on Signal Processing* 54.7 (July 2006), pp. 2716–2724 (cit. on p. 100).

[SRF08] N. Sellami, A. Roumy, and I. Fijalkow. "A Proof of Convergence of the MAP Turbo-Detector to the AWGN Case". In: *IEEE Transactions on Signal Processing* 56.4 (Apr. 2008), pp. 1548–1561 (cit. on p. 100).

[SSM06] B. Sirkeci-Mergen, A. Scaglione, and G. Mergen. "Asymptotic analysis of multistage cooperative broadcast in wireless networks". In: *IEEE Transactions on Information Theory* 52.6 (June 2006), pp. 2531–2550 (cit. on pp. 2, 10, 239).

[SSS02] S. Song, A. C. Singer, and K. Sung. "Turbo equalization with an unknown channel". In: *IEEE International Conference on Acoustics, Speech, and Signal Processing*. Vol. 3. May 2002, pp. III–2805–III–2808 (cit. on p. 60).

[SV05] H. H. Sneessens and L. Vandendorpe. "Soft decode and forward improves cooperative communications". In: *1st IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, 2005.* Dec. 2005, pp. 157–160 (cit. on p. 42).

[SZW+15] P. Sun, C. Zhang, Z. Wang, et al. "Iterative Receiver Design for ISI Channels Using Combined Belief- and Expectation-Propagation". In: *IEEE Signal Processing Letters* 22.10 (Oct. 2015), pp. 1733–1737 (cit. on pp. 66, 77, 137, 253, 258).

[Tak17] K. Takeuchi. "Rigorous dynamics of expectation-propagation-based signal recovery from unitarily invariant measurements". In: *IEEE Proc. in Int. Symp. on Information Theory (ISIT)*. June 2017, pp. 501–505 (cit. on pp. 224, 225).

[Tan+18] Xiaosi Tan, Weihong Xu, Yair Be'ery, Zaichen Zhang, Xiaohu You, and Chuan Zhang. "Improving Massive MIMO Belief Propagation Detector with Deep Neural Network". In: *arXiv preprint arXiv:1804.01002* (2018) (cit. on p. 91).

[Tan81] R. Tanner. "A recursive approach to low complexity codes". In: *IEEE Transactions on Information Theory* 27.5 (Sept. 1981), pp. 533–547 (cit. on p. 54).

[Tao15] J. Tao. "Single-Carrier Frequency-Domain Turbo Equalization With Various Soft Interference Cancellation Schemes for MIMO Systems". In: *IEEE Transactions on Communications* 63.9 (Sept. 2015), pp. 3206–3217 (cit. on pp. 4, 12, 164, 165, 175, 179, 202, 208).

[Tao16] J. Tao. "On Low-Complexity Soft-Input Soft-Output Decision-Feedback Equalizers". In: *IEEE Communications Letters* 20.9 (Sept. 2016), pp. 1737–1740 (cit. on pp. 3, 12, 114, 116, 119, 120, 125, 144, 155).

[TB00] Stephan Ten Brink. "Designing iterative decoding schemes with the extrinsic information transfer chart". In: *AEU Int. J. Electron. Commun.* Vol. 54. 6. Jan. 2000 (cit. on pp. 49, 52, 63, 65, 92, 131, 173, 187, 222).

[TB+16] S. Tedik Başaran, G. Karabulut Kurt, M. Uysal, and İ. Altunbaş. "A tutorial on network coded cooperation". In: *IEEE Communications Surveys Tutorials* 18.4 (2016), pp. 2970–2990 (cit. on p. 42).

[TBH02] M. Tüchler, S. Ten Brink, and J. Hagenauer. "Measures for Tracing Convergence of Iterative Decoding Algorithms". In: *in Proc. 4th IEEE/ITG Conf. on Source and Channel Coding*. 2002, pp. 53–60 (cit. on p. 225).

[TH00] Michael Tüchler and Joachim Hagenauer. "Turbo equalization using frequency domain equalizers". In: *Proc. of the Allerton Conference*. Oct. 2000 (cit. on pp. 4, 12).

[TH01] M. Tüchler and J. Hagenauer. "Linear time and frequency domain turbo equalization". In: *Proc. IEEE 54rd Veh. Technology Conf.* Vol. 4. Nov. 2001, pp. 2773–2777 (cit. on pp. 4, 12, 109, 110, 163, 164, 174, 175).

[Tho+17] R. R. Thomas, M. Cardone, R. Knopp, D. Tuninetti, and B. T. Maharaj. "A Practical Feasibility Study of a Novel Strategy for the Gaussian Half-Duplex Relay Channel". In: *IEEE Transactions on Wireless Communications* 16.1 (Jan. 2017), pp. 101–116 (cit. on pp. 21, 42).

[TK+02] M. Tüchler, R. Koetter, et al. "Turbo equalization: principles and new results". In: *IEEE Transactions on Communications* 50.5 (May 2002), pp. 754–767 (cit. on pp. 3, 11, 66, 109–114, 116, 118, 119, 121, 125, 132, 137, 144, 145).

[Tob87] F. A. Tobagi. "Modeling and performance analysis of multihop packet radio networks". In: *Proceedings of the IEEE* 75.1 (Jan. 1987), pp. 135–155 (cit. on pp. 1, 9).

[Tom05] S. Tomasin. "Overlap and save frequency domain DFE for throughput efficient single carrier transmission". In: *Proceedings of the IEEE PIMRC'05*. Vol. 2. Sept. 2005, pp. 1199–1203 (cit. on p. 198).

[Ton00] Andrea M Tonello. "Space-time bit-interleaved coded modulation with an iterative decoding strategy". In: *Proc. of the IEEE VTC-Fall'00*. Vol. 1. IEEE. 2000, pp. 473–478 (cit. on pp. 201, 209).

[TPB16] R. Tajan, C. Poulliat, and M. Boucheret. "Circular Faster Than Nyquist: Transmitter and iterative receiver design". In: *2016 9th International Symposium on Turbo Codes and Iterative Information Processing (ISTC)*. Sept. 2016, pp. 241–245 (cit. on p. 193).

[TRK05] V. D. Trajković, P. B. Rapajic, and R. A. Kennedy. "Turbo DFE algorithm with imperfect decision feedback". In: *IEEE Signal Processing Letters* 12.12 (Dec. 2005), pp. 820–823 (cit. on pp. 116, 125).

[TS11] M. Tüchler and A. C. Singer. "Turbo Equalization: An Overview". In: *IEEE Transactions on Information Theory* 57.2 (Feb. 2011), pp. 920–952 (cit. on pp. 3, 11, 48, 101, 109–111, 122, 141, 163, 222).

[Tse+02] Yu-Chee Tseng, Sze-Yao Ni, Yuh-Shyan Chen, and Jang-Ping Sheu. "The Broadcast Storm Problem in a Mobile Ad Hoc Network". In: *Wireless Networks* 8.2 (Mar. 2002), pp. 153–167 (cit. on pp. 21, 43).

[TSK02] M. Tüchler, A. C. Singer, and R. Koetter. "Minimum mean squared error equalization using a priori information". In: *IEEE Transactions on Signal Processing* 50.3 (Mar. 2002), pp. 673–683 (cit. on pp. 3, 11, 12, 48, 60, 91, 109, 110, 118, 119, 126, 127, 129, 139, 156, 163).

[TTA06] Hiromichi Tomeba, Kazuaki Takeda, and Fumiyuki Adachi. "Overlap MMSE-frequency-domain equalization for multicarrier signal transmissions". In: *Proceedings of the 9th Symp. on Wireless Personal Mulimedia Communications*. 2006 (cit. on p. 198).

[TUY14] M. N. Tehrani, M. Uysal, and H. Yanikomeroglu. "Device-to-device communication in 5G cellular networks: challenges, solutions, and future directions". In: *IEEE Communications Magazine* 52.5 (May 2014), pp. 86–92 (cit. on p. 20).

[TV05] David Tse and Pramod Viswanath. *Fundamentals of wireless communication*. Cambridge university press, 2005 (cit. on pp. 21, 29, 39, 40, 75, 201, 202, 207).

[TW+11] J. Tao, J. Wu, et al. "Enhanced MIMO LMMSE Turbo Equalization: Algorithm, Simulations, and Undersea Experimental Results". In: *IEEE Transactions on Signal Processing* 59.8 (Aug. 2011), pp. 3813–3823 (cit. on p. 116).

[Van+15] J. Van Hecke, P. Del Fiorentino, R. Andreotti, V. Lottici, F. Giannetti, L. Vandendorpe, and M. Moeneclaey. "Accurate modeling of the predicted κESM-based link performance metric for BIC-OFDM systems". In: *2015 IEEE Symposium on Communications and Vehicular Technology in the Benelux (SCVT)*. Nov. 2015, pp. 1–6 (cit. on p. 222).

[VAP06] C. M. Vithanage, C. Andrieu, and R. J. Piechocki. "Approximate inference in hidden Markov models using iterative active state selection". In: *IEEE Signal Processing Letters* 13.2 (Feb. 2006), pp. 65–68 (cit. on p. 101).

[VB+10] R. Visoz, A. O. Berthet, et al. "Semi-Analytical Performance Prediction Methods for Iterative MMSE-IC Multiuser MIMO Joint Decoding". In: *IEEE Transactions on Communications* 58.9 (Sept. 2010), pp. 2576–2589 (cit. on pp. 7, 15, 147, 148, 150, 218, 221–223, 226, 230, 231, 250).

[VBC06] R. Visoz, A. O. Berthet, and S. Chtourou. "Frequency-Domain Block Turbo-Equalization for Single-Carrier Transmission Over MIMO Broadband Wireless Channel". In: *IEEE Transactions on Communications* 54.12 (Dec. 2006), pp. 2144–2149 (cit. on pp. 164, 165, 175, 202, 208).

[VBL08] R. Visoz, A. O. Berthet, and M. Lalam. "A Novel Fast Semi-Analytical Performance Prediction Method for Iterative MMSE-IC Multiuser MIMO Joint Decoding". In: *Proc. IEEE GLOBECOM'08*. Nov. 2008, pp. 1–5 (cit. on p. 75).

[Ver98] S. Verdù. *Multiuser detection*. Cambridge Univ. Press, 1998 (cit. on p. 104).

[VG06] D. Veronesi and D. L. Goeckel. "CTH15-4: Multiple Frequency Offset Compensation in Cooperative Wireless Systems". In: *IEEE Globecom 2006*. Nov. 2006, pp. 1–5 (cit. on pp. 239, 246).

[VHG01] M. Vollmer, M. Haardt, and J. Gotze. "Comparative study of joint-detection techniques for TD-CDMA based mobile radio systems". In: *IEEE Journal on Selected Areas in Communications* 19.8 (Aug. 2001), pp. 1461–1475 (cit. on pp. 196, 197).

[Vis+05] R. Visoz et al. "A new class of iterative equalizers for space-time BICM over MIMO block fading multipath AWGN channel". In: *IEEE Transactions on Communications* 53.12 (Dec. 2005), pp. 2076–2091 (cit. on pp. 109, 110).

[Vit67] Andrew Viterbi. "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm". In: *IEEE Transactions on Information Theory* 13.2 (1967), pp. 260–269 (cit. on p. 51).

[VLC96] P. A. Voois, I. Lee, and J. M. Cioffi. "The effect of decision delay in finite-length decision feedback equalization". In: *IEEE Transactions on Information Theory* 42.2 (Mar. 1996), pp. 618–621 (cit. on p. 107).

[Wal06] J. MacLaren Walsh. "Distributed iterative decoding and estimation via expectation propagation: performance and convergence". Ph.D. dissertation. Ithaca: Cornell University, May 2006 (cit. on p. 54).

[Wan+15] B. Wang, K. Wang, Z. Lu, T. Xie, and J. Quan. "Comparison study of non-orthogonal multiple access schemes for 5G". In: *2015 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*. June 2015, pp. 1–5 (cit. on p. 40).

[Wan+19] Xiaojie Wang, Chulong Liang, Li Ping, and Stephan ten Brink. "Achievable Rate Region for Iterative Multi-User Detection via Low-cost Gaussian Approximation". In: *ISIT 2019 - International Symposium on Information Theory*. IEEE. 2019, pp. 1–4 (cit. on pp. 201, 203).

[WC02] Brad Williams and Tracy Camp. "Comparison of Broadcasting Techniques for Mobile Ad Hoc Networks". In: *Proceedings of the 3rd ACM International Symposium on Mobile Ad Hoc Networking &Amp; Computing*. MobiHoc '02. Lausanne, Switzerland: ACM, 2002, pp. 194–205 (cit. on p. 43).

[WGV06] S. Wei, D. L. Goeckel, and M. C. Valenti. "Asynchronous cooperative diversity". In: *IEEE Transactions on Wireless Communications* 5.6 (June 2006), pp. 1547–1557 (cit. on pp. 21, 42, 44, 239, 246).

[WH03] Wim Wiegerinck and Tom Heskes. "Fractional Belief Propagation". In: *Advances in Neural Information Processing Systems 15*. Ed. by S. Becker, S. Thrun, and K. Obermayer. MIT Press, 2003, pp. 438–445 (cit. on p. 74).

[WH77] Won Lee and F. Hill. "A Maximum-Likelihood Sequence Estimator with Decision-Feedback Equalization". In: *IEEE Transactions on Communications* 25.9 (Sept. 1977), pp. 971–979 (cit. on p. 103).

[Wib96] N. Wiberg. "Codes and decoding on general graphs". Ph.D. dissertation. Linköping, Sweden: Linköping University, 1996 (cit. on p. 99).

[Wit+02] M. Witzke, S. Baro, F. Schreckenbach, and J. Hagenauer. "Iterative detection of MIMO signals with linear detectors". In: *Conference Record of the Thirty-Sixth Asilomar Conference on Signals, Systems and Computers, 2002.* Vol. 1. Nov. 2002, pp. 289–293 (cit. on pp. 48, 101, 165, 175, 202, 208).

[WLL08] Y. Wang, Y. C. Liang, and W. S. Leon. "Frequency Domain Equalization and Interference Cancellation for TD-SCDMA Downlink in Fast Time-Varying Environments". In: *IEEE Transactions on Vehicular Technology* 57.1 (Jan. 2008), pp. 648–653 (cit. on p. 198).

[WP99] Xiaodong Wang and H. V. Poor. "Iterative (turbo) soft interference cancellation and decoding for coded CDMA". In: *IEEE Transactions on Communications* 47.7 (July 1999), pp. 1046–1061 (cit. on pp. 3, 11, 48, 54, 59, 91, 101, 109, 110, 117, 163, 206, 207).

[WRM+17] K. Wu, G. Ren, X. Meng, et al. "Spectral-Efficient Band Allocation Scheme for Frequency-Domain Pulse-Shaping-Based SC-FDMA Systems". In: *IEEE Transactions on Vehicular Technology* 66.9 (Sept. 2017), pp. 8249–8262 (cit. on pp. 166, 176).

[WS73] T. Walzman and M. Schwartz. "Automatic equalization using the discrete frequency domain". In: *IEEE Transactions on Information Theory* 19.1 (Jan. 1973), pp. 59–68 (cit. on pp. 161, 164).

[WSG08] Z. Wang, H. R. Sadjadpour, and J. J. Garcia-Luna-Aceves. "A Unifying Perspective on the Capacity of Wireless Ad Hoc Networks". In: *IEEE INFOCOM 2008 - The 27th Conference on Computer Communications.* Apr. 2008, pp. 211–215 (cit. on pp. 21, 22).

[Wu+14] S. Wu, L. Kuang, Z. Ni, J. Lu, D. Huang, and Q. Guo. "Low-Complexity Iterative Detection for Large-Scale Multiuser MIMO-OFDM Systems Using Approximate Message Passing". In: *IEEE Journal of Selected Topics in Signal Processing* 8.5 (Oct. 2014), pp. 902–915 (cit. on pp. 3, 11, 80).

[WXY09] H. Wang, X. Xia, and Q. Yin. "Computationally efficient equalization for asynchronous cooperative communications with multiple frequency offsets". In: *IEEE Transactions on Wireless Communications* 8.2 (Feb. 2009), pp. 648–655 (cit. on pp. 21, 239).

[WZ07] J. Wu and Y. R. Zheng. "Low Complexity Soft-Input Soft-Output Block Decision Feedback Equalization". In: *IEEE GLOBECOM 2007 - IEEE Global Telecommunications Conference.* Nov. 2007, pp. 3379–3383 (cit. on p. 114).

[WZ08] J. Wu and Y. R. Zheng. "Low complexity soft-input soft-output block decision feedback equalization". In: *IEEE Journal on Selected Areas in Communications* 26.2 (Feb. 2008), pp. 281–289 (cit. on p. 114).

[XC11] Y. Xiao and L. J. Cimini. "Spectral efficiency of distributed cooperative relaying". In: *2011 45th Annual Conference on Information Sciences and Systems.* Mar. 2011, pp. 1–6 (cit. on p. 43).

[XC12] Y. Xiao and L. J. Cimini. "Spectral efficiency of cooperative relaying with imperfect channel estimation". In: *2012 IEEE International Conference on Communications (ICC)*. June 2012, pp. 4616–4620 (cit. on p. 43).

[Yen+11] G. Yenihayat, F. A. Onat, E. Kolagasioglu, and A. O. Yilmaz. "Cooperative Relaying Schemes for Narrow-Band Frequency Hopping Wireless Ad Hoc Networks". In: *2011 IEEE 73rd Vehicular Technology Conference (VTC Spring)*. May 2011, pp. 1–5 (cit. on pp. 21, 44).

[YFW01] Jonathan S Yedidia, William T Freeman, and Yair Weiss. "Bethe free energy, Kikuchi approximations, and belief propagation algorithms". In: *Advances in neural information processing systems 13* (2001) (cit. on p. 68).

[YFW03] J.S. Yedidia, W.T. Freeman, and Y. Weiss. "Understanding Belief Propagation and Its Generalizations". In: *Exploring Artificial Intelligence in the New Millennium*. Ed. by G. Lakemeyer and B. Nebel. Morgan Kaufmann Publishers, Jan. 2003. Chap. 8, pp. 239–236 (cit. on pp. 49, 53, 54, 67, 70).

[YFW05] J. S. Yedidia, W. T. Freeman, and Y. Weiss. "Constructing free-energy approximations and generalized belief propagation algorithms". In: *IEEE Transactions on Information Theory* 51.7 (July 2005), pp. 2282–2312 (cit. on pp. 67, 71).

[YGP07] X. Yuan, Q. Guo, and Li Ping. "Evolution Analysis of Iterative LMMSE-APP Detection for Coded Linear System with Cyclic Prefixes". In: *IEEE Proc. in Int. Symp. on Information Theory (ISIT)*. June 2007, pp. 71–75 (cit. on pp. 7, 15, 222, 223, 226, 228).

[Yua+08] Xiaojun Yuan et al. "Low-complexity iterative detection in multi-user MIMO ISI channels". In: *IEEE S.P. Lett.* 15 (2008), pp. 25–28 (cit. on pp. 202, 206, 210).

[ZDRD19] Alessio Zappone, Marco Di Renzo, and Mérouane Debbah. "Wireless Networks Design in the Era of Deep Learning: Model-Based, AI-Based, or Both?" In: *arXiv preprint arXiv:1902.02647* (2019) (cit. on p. 87).

[Zha+15a] C. Zhang, C. N. Manchón, Z. Wang, and B. H. Fleury. "Message-Passing Receivers for Single Carrier Systems with Frequency-Domain Equalization". In: *IEEE Signal Processing Letters* 22.4 (Apr. 2015), pp. 404–407 (cit. on pp. 3, 11, 80, 166).

[Zha+15b] J. Zhang, L. Yang, L. Hanzo, and H. Gharavi. "Advances in Cooperative Single-Carrier FDMA Communications: Beyond LTE-Advanced". In: *IEEE Communications Surveys Tutorials* 17.2 (2015), pp. 730–756 (cit. on pp. 21, 31, 194).

[Zha+16] D. Zhang, L. L. Mendes, M. Matthé, I. S. Gaspar, N. Michailow, and G. P. Fettweis. "Expectation Propagation for Near-Optimum Detection of MIMO-GFDM Signals". In: *IEEE Transactions on Wireless Communications* 15.2 (Feb. 2016), pp. 1045–1062 (cit. on pp. 166, 176).

[Zha+18] C. Zhang, J. Butepage, H. Kjellstrom, and S. Mandt. "Advances in Variational Inference". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2018), pp. 1–1 (cit. on p. 88).

[ZLL06] Shengli Zhang, Soung Chang Liew, and Patrick P. Lam. "Hot Topic: Physical-layer Network Coding". In: *Proceedings of the 12th Annual International Conference on Mobile Computing and Networking*. MobiCom '06. Los Angeles, CA, USA: ACM, 2006, pp. 358–365 (cit. on pp. 2, 10).

[ZPL90] K. Zhou, J. G. Proakis, and F. Ling. "Decision-feedback equalization of time-dispersive channels with coded modulation". In: *IEEE Transactions on Communications* 38.1 (Jan. 1990), pp. 18–24 (cit. on p. 108).

[ZV05] B. Zhao and M. C. Valenti. "Practical relay networks: a generalization of hybrid-ARQ". In: *IEEE Journal on Selected Areas in Communications* 23.1 (Jan. 2005), pp. 7–18 (cit. on pp. 42, 43).

Résumé — Les réseaux ad hoc mobiles (MANETs) sont des systèmes de communication sans fil rapidement déployables et qui fonctionnent avec une coordination minimale, ceci afin d'éviter les pertes d'efficacité spectrale induites par la signalisation. Les stratégies de transmissions coopératives présentent un intérêt pour les MANETs, mais la nature distribuée de tels protocoles peut augmenter le niveau d'interférence avec un impact autant plus sévère que l'on cherche à pousser les limites des efficacités énergétique et spectrale. L'impact de l'interférence doit alors être réduit par l'utilisation d'algorithmes de traitement du signal au niveau de la couche PHY, avec une complexité calculatoire raisonnable.

Des avancées récentes sur les techniques de conception de récepteurs numériques itératifs proposent d'exploiter l'inférence bayésienne approximée et des techniques de passage de message associés afin d'améliorer le potentiel des turbo-détecteurs plus classiques. Entre autres, la propagation d'espérance (EP) est une technique flexible, qui offre des compromis attractifs de complexité et de performance dans des situations où la propagation de croyance conventionnel est limité par sa complexité calculatoire. Par ailleurs, grâce à des techniques émergentes de l'apprentissage profond, de telles structures itératives peuvent être projetés vers des réseaux de détection profonds, où l'apprentissage des hyper-paramètres algorithmiques améliore davantage les performances.

Dans cette thèse nous proposons des égaliseurs à retour de décision à réponse impulsionnelle finie basée sur la propagation d'espérance (EP) qui apportent des améliorations significatives, en particulier pour des applications à haute efficacité spectrale vis à vis des turbo-détecteurs conventionnels, tout en ayant l'avantage d'être asymptotiquement prédictibles. Nous proposons un cadre générique pour la conception de récepteurs dans le domaine fréquentiel, afin d'obtenir des architectures de détection avec une faible complexité calculatoire. Cette approche est analysée théoriquement et numériquement, avec un accent mis sur l'égalisation des canaux sélectifs en fréquence, et avec des extensions pour de la détection dans des canaux qui varient dans le temps ou pour des systèmes multi-antennes. Nous explorons aussi la conception de détecteurs multi-utilisateurs, ainsi que l'impact de l'estimation du canal, afin de comprendre le potentiel et le limite de cette approche.

Pour finir, nous proposons une méthode de prédiction performance à taille finie, afin de réaliser une abstraction de lien pour l'égaliseur domaine fréquentiel à base d'EP. L'impact d'un modélisation plus fine de la couche PHY est évalué dans le contexte de la diffusion coopérative pour des MANETs tactiques, grâce à un simulateur flexible de couche MAC.

Mots clés : Réseau ad hoc mobile (MANET), communication coopérative, diffusion coopérative, détection itérative, égalisation, estimation de channel, inférence bayésienne, propagation de l'espérance, algorithmes de passage de message, dépliement profond, prédiction semi-analytique.

Abstract — Mobile ad hoc networks (MANETs) are rapidly deployable wireless communications systems, operating with minimal coordination in order to avoid spectral efficiency losses caused by overhead. Cooperative transmission schemes are attractive for MANETs, but the distributed nature of such protocols comes with an increased level of interference, whose impact is further amplified by the need to push the limits of energy and spectral efficiency. Hence, the impact of interference has to be mitigated through with the use PHY layer signal processing algorithms with reasonable computational complexity.

Recent advances in iterative digital receiver design techniques exploit approximate Bayesian inference and derivative message passing techniques to improve the capabilities of well-established turbo detectors. In particular, expectation propagation (EP) is a flexible technique which offers attractive complexity-performance trade-offs in situations where conventional belief propagation is limited by computational complexity. Moreover, thanks to emerging techniques in deep learning, such iterative structures are cast into deep detection networks, where learning the algorithmic hyper-parameters further improves receiver performance.

In this thesis, EP-based finite-impulse response decision feedback equalizers are designed, and they achieve significant improvements, especially in high spectral efficiency applications, over more conventional turbo-equalization techniques, while having the advantage of being asymptotically predictable. A framework for designing frequency-domain EP-based receivers is proposed, in order to obtain detection architectures with low computational complexity. This framework is theoretically and numerically analysed with a focus on channel equalization, and then it is also extended to handle detection for time-varying channels and multiple-antenna systems. The design of multiple-user detectors and the impact of channel estimation are also explored to understand the capabilities and limits of this framework.

Finally, a finite-length performance prediction method is presented for carrying out link abstraction for the EP-based frequency domain equalizer. The impact of accurate physical layer modelling is evaluated in the context of cooperative broadcasting in tactical MANETs, thanks to a flexible MAC-level simulator.

Keywords: Mobile ad-hoc network (MANET), cooperative communications, cooperative broadcast, iterative detection, equalization, channel estimation, Bayesian inference, expectation propagation, message passing algorithms, deep unfolding, semi-analytic prediction.

Institut de Recherche en Informatique de Toulouse / Toulouse INP, 2 Rue Charles Camichel 31000 Toulouse