Shaping Interference Towards Optimality of Modern Wireless Communication Transceivers

Guido Ferrante

CentraleSupélec

N° d'ordre : 2015-08-TH

CentraleSupélec

ECOLE DOCTORALE STITS
"Sciences et Technologies de l'Information, des Télécommunications et des Systémes"

THÈSE DE DOCTORAT

DOMAINE : Sciences et Technologies de l'Information et de la Communication (STIC)
Spécialité : Télécommunications

Soutenue le 10 avril 2015

par :

Guido Carlo FERRANTE

Façonnement de l'Interférence en vue d'une Optimisation Globale d'un Système Moderne de Communication
(Shaping Interference Towards Optimality of Modern Wireless Communication Transceivers)

Directrice de thèse :	Maria-Gabriella DI BENEDETTO	Professeur, Sapienza
Codirecteur de thèse :	Jocelyn FIORINA	Professeur, CentraleSupélec
Composition du jury :		
Président du jury :	Pierre DUHAMEL	Directeur de recherche, CNRS/CentraleSupélec
Rapporteurs :	Laurent CLAVIER	Professeur, Université de Lille
	Michel TERRÉ	Professeur, CNAM
Examinateur :	Andrea GIORGETTI	Professeur agrégé, Université de Bologne

Sagomatura dell'Interferenza Verso un'Ottimizzazione dei Moderni Sistemi di Ricetrasmissione

Shaping Interference Towards Optimality in Modern Wireless Communication Transceivers

Guido Carlo Ferrante

Dottorato di Ricerca
in Ingegneria dell'Informazione e della Comunicazione
Doctor of Philosophy
in Information and Communication Engineering

Sapienza Università di Roma
Sapienza University of Rome

2015
XXVII Ciclo
27^{th} *Course*

Autore
Author .....

Dipartimento di Ingegneria dell'Informazione, Elettronica e Telecomunicazioni
Department of Information Engineering, Electronics and Telecommunications

Visto
Certified by .....

Prof. Maria-Gabriella Di Benedetto
Tutore di Dottorato *(Thesis Supervisor)*

Visto
Certified by .....

Prof. Jocelyn Fiorina
Co-tutore di Dottorato *(Thesis Co-supervisor)*

SAPIENZA
UNIVERSITÀ DI ROMA

CentraleSupélec

Shaping Interference Towards Optimality of Modern Wireless Communication Transceivers

by

Guido Carlo Ferrante

A Thesis Submitted to the
Department of Information Engineering, Electronics and Telecommunications
at
Sapienza University of Rome
and to the
Department of Telecommunications
at
CentraleSupélec
in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy
in Information and Communication Engineering

April 2015

Supervisors:
Prof. Maria-Gabriella Di Benedetto
and
Prof. Jocelyn Fiorina

This thesis was typeset using the LaTeX typesetting system originally developed by Leslie Lamport, based on TeX created by Donald Knuth.

The body text is set in the 10pt size from Donald Knuth's Concrete Roman font. Other fonts include Sans and Typewriter from Donald Knuth's Computer Modern family.

Most figures were typeset in TikZ.

Acknowledgements

I would like to thank everyone who supported me during the years of my PhD. Support is a somewhat vague word, and this implies that I feel indebted to too many people to remember. My gratitude is doubled for those whom I don't remember at the time of writing.

I am of course indebted to both my advisors, Prof. Di Benedetto and Prof. Fiorina. I spent many hours with them, and I realize how much their time is worth. They encouraged me to pursue the double degree, and I will always be grateful for their help. I hope I have been and continue to be at the level of their expectations.

I want to thank Luca De Nardis: he is a dear friend and in many ways a model for me. Also, not only did I unduly share his office space for a year, but he was always funny and positive. And thoughtful, when needed.

I want to thank my lab mates, in particular Stefano and Giuseppe: we had long discussions about research, career, and life. And Luca Lipa, who has been a close friend of mine since we met, many years ago, in the ACTS lab.

I would like to thank my dear friend Valeria, who was always there for me when I needed her, and never stepped back from saying the truth, which is the way she cared for me; and my dear friend Elira, who is always close despite any distance. I want to thank Nick, who was always keen on coffee breaks, sharing thoughts, and helping me friendly.

I lived one year with my long friend Matt, and I thank him for putting up with me.

Words cannot express the gratitude to my brother: he was a guidance, a friend, a mentor. I know that I can rely on him. I know that I can discuss any matter, and I can grasp some new perspective every time.

Finally, thanks Rome, because you are the only place where I felt at home: Rome sweet Rome.

This thesis is dedicated to the old generations, my grandparents and parents, who valued education above all else, and to the new one, the little Isabel, who is starting her most adventurous trip in this world: the one in the knowledge.

Contents

Synopsis

A communication is impulsive whenever the information-bearing signal is burst-like in time. Examples of the impulsive concept are: impulse-radio signals, that is, wireless signals occurring within short intervals of time; optical signals conveyed by photons; speech signals represented by sound pressure variations; pulse-position modulated electrical signals; a sequence of arrival/departure events in a queue; neural spike trains in the brain. Understanding impulsive communications requires to identify what is peculiar to this transmission paradigm, *i.e.*, different from traditional continuous communications.

In order to address the problem of understanding impulsive *vs.* non-impulsive communications, the framework of investigation must include the following aspects: the different interference statistics directly following from the impulsive signal structure; the different interaction of the impulsive signal with the physical medium; the actual possibility for impulsive communications of coding information into the time structure, relaxing the implicit assumption made in continuous transmissions that time is a mere support.

This thesis partially addresses a few of the above issues, and draws future lines of investigation.

The starting point of our analysis (Chapter 1) is a multiple access scheme where users adopt time-hopping spread-spectrum signals to communicate towards a common receiver. In time-hopping spread-spectrum, a symbol period is divided into N chips, and just a subset of chips is actually used to transmit the radio signal; in other words, the spreading sequence is modeled with a sparse vector, with N_s nonzero entries. Different degrees of sparsity imply different degrees of impulsiveness. In particular, two regimes may be studied as N grows to infinity, corresponding N_s finite or $N_s/N \to \alpha \in (0,1)$. When $N_s = N$, we obtain direct-sequence spread-spectrum. The energy concentration in time-hopping is, therefore, achieved by the uneven use of the degrees of freedom in time. The analysis is conducted in terms of mutual information (with Gaussian inputs) or, whenever feasible, spectral efficiency, in the so-called "large-system limit," where the number of users K is proportional to the number of dimensions N of the spreading sequence, *i.e.*, with fixed load $\beta := K/N$ as both $K \to \infty$ and $N \to \infty$. In order to understand the role of multiuser interference, different receiver structures are considered, namely optimum and linear receivers. The key outcome of the analysis is the following. Spectral efficiency with optimum decoding is higher with direct-sequence than with time-hopping. We show that, in the large-system limit, spectral efficiency increases as N_s increases, even when N_s remains finite, and thus $N_s/N \to 0$. It does not matter that, as N_s increases, interference tends to be Gaussian, and,

therefore, more detrimental—optimum multiuser detection shall cope with that. This is no longer true if we adopt a far simpler receiver, *e.g.* a bank of single-user matched-filters or MMSE filters. In this case, the interference distribution plays a key role, and the case $N_s = 1$ (maximum energy concentration) shows the advantages of impulsive modulation formats: in particular, in a low load, high $\mathcal{E}_b/\mathcal{N}_0$ setting, sparsity allows to achieve a spectral efficiency that is strictly higher than that achievable with direct-sequence.

In the above scheme, transmitted energy is concentrated in bursts, that is the basic idea of impulse-radio communications. Although this characteristic can be attained irrespectively of bandwidth, that defines the "effective duration" of the transmitted pulse, impulse-radio signals have been extensively applied with ultrawide bandwidths. In the opposite case, that is, with narrow bandwidths, the transmission rate would be, indeed, severely affected. The ultra-wideband characteristic of the transmitted signal makes the multipath components of the wireless channel resolvable, whereas the impulse-radio characteristic permits to use simple receiver structures since the signal format avoids intersymbol interference. To further simplify reception, without any performance loss, one must use prefiltering at the transmitter, and specifically a transmit matched filter, also known as time reversal in the ultra-wideband literature. Although prefiltering, and in general precoding, can be used with signals of any bandwidth, transmit matched filter was traditionally used in connection with ultrawide bandwidth. In Chapter 2, the main reason for this is traced back in the statistical behavior of wireless channels. Indeed, the energy of the effective channel, *i.e.*, the channel formed by the cascade of the prefilter and the multipath channel, is monotonically increasing with the bandwidth with usual multipath channels. Since the prefiltering structure requires the knowledge of the multipath channel, Chapter 3 addresses the important issue of the impact of imperfect channel state information on rate and error probability with simple receiver structures, where each user is detected and decoded independently. The investigation links the accuracy needed for the channel estimation with the maximum mutual information achievable with Gaussian inputs. Extension to multiple antennas at the transmitter is considered in Chapter 4, where it is shown that the signal-to-noise ratio achieved with time reversal is not affected by the lack of correlation between channels relative to different antennas, and that multiple antennas increase the energy focusing of time reversal. Chapter 5 compares time reversal with other prefiltering schemes, as a function of the number of fingers of the Rake receiver.

Finally, Chapter 6 discusses interference patterns typically arising with impulsive signals. In particular, two interference distributions seen in Chapter 1 and 3 are shown to belong to a more general family, and a novel interference model arising from binary-valued signals is discussed.

As presented above, the general framework of our analysis permits to address several issues connected with, and implied by, the impulsiveness of transmitted signals. In particular, we are able to markedly separate two characteristics of impulsive signals, namely the time and amplitude statistics, and the bandwidth. On the one hand, the statistics may be measured in several ways; however, the key

feature is the sparsity of the signal into the degrees of freedom occupied in time, *i.e.*, impulsiveness. This peculiar statistics has several implications on fundamental limits of impulsive communications. Under this perspective, we address this issue for a flat-fading multiple access channel with random time-hopping, which implies sparsity of the single-user signal. Simple modifications of the model serve towards the analysis of interesting extensions: fading and multipath channels, non-uniform power constraint over users, frequency-hopping (time-hopping dual), and partial channel knowledge at receiver. On the other hand, the bandwidth has implication on the interaction with the medium where the signal propagates: for example, in the wireless communications setting analyzed, the more the bandwidth, the more the number of resolvable paths—up to the number of multipath components of the channel. We traced back to this interaction the reason for the traditional choice of the time reversal prefilter in connection with ultra-wideband communications. We set the basis for a thorough study of the interplay between the bandwidth of the signal, the multipath channel, the sparsity in time, and the transceiver structure, in both the coded and uncoded regimes, as well as for the robustness analysis of the communication system, that could also be regarded in the bigger picture where the net-ergodic rate of the network and the analog and digital feedback for the channel knowledge acquisition are optimized.

There are several topics that could be envisioned by, but are not specifically addressed in this thesis. In particular, two of the most promising areas of investigation comprises the role of compressive sampling, and in particular super-resolution theory, in the recovery of the sparse signal, and the possibility to encode information directly in time, for example in the interarrival time between pulses, or in their rate. We believe that both investigations may have considerable impact, in particular in the communication and neuroscience communities.

Résumé (français)

Une communication est impulsive chaque fois que le signal portant des informations est intermittent dans le temps et que la transmission se produit à rafales. Des exemples du concept impulsife sont : les signaux radio impulsifs, c'est-à-dire des signaux très courts dans le temps; les signaux optiques utilisé dans les systèmes de télécommunications; certains signaux acoustiques et, en particulier, les impulsions produites par le système glottale; les signaux électriques modulés en position d'impulsions; une séquence d'événements dans une file d'attente; les trains de potentiels neuronaux dans le système neuronal. Ce paradigme de transmission est différent des communications continues traditionnelles et la compréhension des communications impulsives est donc essentielle.

Afin d'affronter le problème des communications impulsives, le cadre de la recherche doit inclure les aspects suivants : la statistique d'interférence qui suit directement la structure des signaux impulsifs; l'interaction du signal impulsif avec le milieu physique; la possibilité pour les communications impulsives de coder l'information dans la structure temporelle. Cette thèse adresse une partie des questions précédentes et trace des lignes indicatives pour de futures recherches.

Chapitre 1

Le point de départ de notre analyse est un système d'accès multiple où les utilisateurs adoptent des signaux avec étalement de spectre par saut temporel *(time-hopping spread spectrum)* pour communiquer vers un récepteur commun. Dans l'étalement de spectre par saut de temps, une période de symbole est divisée en N *chips*, et seulement un sous-ensemble des *chips* est effectivement utilisé pour transmettre le signal radio; en d'autres termes, la séquence d'étalement est modélisée avec un vecteur , avec N_s entrées non nulles. Deux régimes de faible densité sont étudiés quand N tend vers l'infini, correspondant à $N_s/N \to \alpha \in (0,1)$ ou N_s finis. Lorsque $N_s = N$, nous avons un étalement du spectre à séquence directe. La concentration d'énergie dans le temps est donc obtenue par l'utilisation inégale des degrés de liberté disponibles. L'analyse est effectuée en termes d'information mutuelle avec entrées gaussiennes ou, lorsque cela est possible, l'efficacité spectrale, dans la dénommée limite de grand système, où le nombre d'utilisateurs K est proportionnel au nombre de dimensions N de la séquence d'étalement, c'est-à-dire avec $\beta = K/N$ fixée avec $K \to \infty$ et $N \to \infty$. Afin de comprendre le rôle de l'interférence, plusieurs structures de récepteur sont considérées, à savoir la structure optimale et les structures linéaires. Les principaux résultats de l'analyse sont les suivants: l'efficacité spectrale avec décodage optimal est supérieure avec

séquence directe qu'avec saut temporel, en particulier pour $\mathcal{E}_b \gg \mathcal{N}_0$ et $\beta \ll 1$. Peu importe que, quand N_s croît, l'interférence tend à être gaussienne et, par conséquent, plus nuisible : la détection optimale doit faire face à cela. Cela n'est plus vrai si les récepteurs sont linéaires. Dans ce cas, la distribution de l'interférence joue un rôle clé, et le cas $N_s = 1$ (concentration maximale d'énergie) montre pleinement les avantages des formats de modulation impulsive. En particulier, avec $\beta > 1$ et haut $\mathcal{E}_b/\mathcal{N}_0$, la sparsité temporelle permet d'atteindre une efficacité spectrale qui est strictement supérieure à celle obtenue avec séquence directe.

Les résultats de cette section ont été publiés dans les articles suivants :

- G. C. Ferrante, M.-G. Di Benedetto, "Spectral efficiency of Random Time-Hopping CDMA," *IEEE Trans. Inf. Theory*, soumis en Nov. 2013; révisé en Nov. 2014.

Chapitres 2–5

Dans les étalements de spectre par saut temporelle ci-dessus, l'énergie transmise est concentrée en intervalles de courte durée, ce qui forme l'idée de base de communications radio impulsionnelle. Bien que cette propriété peut être vérifiée indépendamment de la bande, qui définit la "durée effective" de l'impulsion transmise, les communications radio d'impulsion ont été largement appliquées dans la bande ultra large. Dans le cas contraire, c'est-à-dire avec des bandes passantes étroites, le taux de transmission serait, en effet, fortement affectée. L'ultra large bande caractéristique du signal transmis permet la résolution de trajets multiples qui caractérise le canal, cette caractéristique impulsive permet l'utilisation des structures de réception simples car le format de signal évite l'interférence entre symboles. Pour simplifier la réception, sans perte de performance, un préfiltre à l'émetteur peut être utilisé, et plus précisément un *transmit matched filter*, également connu comme retournement temporel *(time reversal)* dans la littérature de systèmes à bande ultra large. Bien que le préfiltrage peut être appliqué à des signaux avec largeur de bande quelconque, le *transmit matched filter* a été traditionnellement utilisé dans le cadre de la bande ultra large. Dans le Chapitre 2, la principale raison de l'utilisation de ce filtre pour bande ultra large est reconduit à certaines propriéts statistiques du canal. En effet, l'énergie du canal indiqué comme efficace, c'est-à-dire le canal formé par la cascade du préfiltre et du canal à trajets multiples, est croissante de manière monotone avec la largeur de bande pour les canaux à trajets multiples ordinaires. Puisque le préfiltrage à besoin de connaître le canal à trajets multiples à l'émetteur, le Chapitre 3 aborde la question importante de l'impact de l'information imparfaite du canal sur le taux et la probabilité d'erreur avec des structures de réception simples, où chaque utilisateur est détecté et décodé indépendamment. Nous avons également étudié l'impact de la précision de l'estimation du canal sur l'information mutuelle maximale réalisable avec des entrées gaussiennes. L'extension à plusieurs antennes à l'émetteur est considérée dans le Chapitre 4, où il est montré que le rapport signal-sur-bruit obtenu avec *time reversal* n'est pas affecté par l'absence de corrélation entre les canaux de différentes antennes, et que les antennes multiples augmentent l'énergie concentrée

par *time reversal* au récepteur. Le Chapitre 5 compare le *time reversal* avec d'autres préfiltres, en fonction du nombre de *fingers* du récepteur Rake.

Les résultats de cette section ont été publiés dans les articles suivants :

- G. C. Ferrante, "Is a large bandwidth mandatory to maximally exploit the transmit matched-filter structure?" *IEEE Commun. Lett.*, vol. 18, no. 9, pp. 1555-1558, 2014.
- G. C. Ferrante, J. Fiorina, M.-G. Di Benedetto, "Statistical analysis of the SNR loss due to imperfect time reversal," in *Proc. IEEE Int. Conf. Ultra-Wideband (ICUWB)*, Paris, France, Sep. 1-3, 2014, pp. 36-40.
- G. C. Ferrante, J. Fiorina, M.-G. Di Benedetto, "Time Reversal beamforming in MISO-UWB channels," in *Proc. IEEE Int. Conf. Ultra-Wideband (ICUWB)*, Sydney, Australia, Sep. 15-18, 2013, pp. 261-266.
- G. C. Ferrante, M.-G. Di Benedetto , "Optimum IR-UWB coding under power spectral constraints," in *Proc. ISIVC 2012*, Valenciennes, France, Jul. 4-6, 2012, pp. 192-195.
- G. C. Ferrante, "Time Reversal against optimum precoder over frequency-selective channels," in *Proc. 18th European Wireless Conf.*, Poznan, Poland, Apr. 18-20, 2012, pp. 1-8.
- G. C. Ferrante, J. Fiorina, M.-G. Di Benedetto, "Complexity reduction by combining Time Reversal and IR-UWB," in *Proc. IEEE Wireless Commun. and Networking Conf. (WCNC)*, Paris, France, Apr. 1-4, 2012, pp. 28-31.
- S. Boldrini, G. C. Ferrante, M.-G. Di Benedetto, "UWB network recognition based on impulsiveness of energy profiles," in *Proc. IEEE Int. Conf. Ultra-Wideband (ICUWB)*, Bologna, Italy, Sep. 14-16, 2011, pp. 327-330.

Chapitre 6

Le Chapitre 6 analyse les modèles d'interférence pour des signaux impulsifs. En particulier, nous montrons que deux distributions d'interférence, dejà obtenues aux chapitres 1 et 2, appartiennent à une famille plus générale, et nous trouvons une nouvelle distribution pour l'interférence produit par accès aléatoire à une ressource commune, quand les signaux des utilisateurs sont de forme binaires antipodaux.

Les résultats de cette section ont été publiés dans les articles suivants :

- G. C. Ferrante, M.-G. Di Benedetto, "Closed form asymptotic expression of a random-access interference measure," *IEEE Commun. Lett.*, vol. 18, no. 7, pp. 1107-1110, 2014.

Discussion et travaux futurs

Le cadre général de notre analyse permet de traiter plusieurs problèmes liés à l'impulsivité de signaux. En particulier, nous sommes maintenant en mesure de séparer nettement les deux caractéristiques principales des signaux impulsifs, à savoir faible densité dans le temps et largeur de bande. Nous avons modélisé la faible densité avec le codes des time-hopping, ce qui a plusieurs implications sur les limites fondamentales de communications impulsives. Grâce à de simples modifications, le modèle proposé peut être utilisé pour analyser certaines extensions intéressantes : fading et canaux à trajets multiples, contrainte de puissance non uniforme sur les utilisateurs, frequency-hopping, et connaissance partielle du canal au récepteur. La largeur de bande a des implications sur l'action réciproque avec le milieu dans laquelle le signal se propage : par exemple, dans le cadre analysé des communications sans fil, plus grande sera la bande, plus nombreux seront le chemins résolubles—jusqu'au nombre de chemins multiples composants le canal. Nous avons ainsi expliqué la raison du choix traditionnel du préfiltre time reversal dans le cadre de communications ultra large bande. Nous avons posé la base d'une étude approfondie de l'interaction entre la largeur de bande du signal, le canal à trajets multiples, la faible densité dans le temps, et la structure d'émetteur-récepteur, dans les deux régimes codés et non codés, ainsi que pour l'analyse de la robustesse du système de communication. Il y a plusieurs sujets qui pourraient être envisagées à la suite de notre étude, et ne sont pas abordées dans cette thèse. En particulier, deux des domaines les plus prometteurs sont : le rôle de la théorie de l'acquisition comprimée (compressed sensing) et la possibilité de coder l'information directement dans le temps, par exemple dans l'intervalle temporel entre deux impulsions. Ces deux enquêtes pourraient avoir un impact considérable, en particulier dans la compréhension des problèmes de base et la génération de modèles de phénomènes physiques qui sont typiques des domaines des communications et des neurosciences.

CHAPTER 1

Spectral Efficiency of Random Time-Hopping CDMA

Traditionally paired with impulsive communications, Time-Hopping CDMA (TH-CDMA) is a multiple access technique that separates users in time by coding their transmissions into pulses occupying a subset of N_s chips out of the total N included in a symbol period, in contrast with traditional Direct-Sequence CDMA (DS-CDMA) where $N_s = N$. The object of this work was to analyze TH-CDMA with random coding, by determining whether peculiar theoretical limits were identifiable, with both optimal and sub-optimal receiver structures. Results indicate that TH-CDMA has a fundamentally different behavior than DS-CDMA, where the crucial role played by energy concentration, typical of time-hopping, directly relates with its intrinsic "uneven" use of degrees of freedom.

While Direct-Sequence CDMA (DS-CDMA) is widely adopted and thoroughly analyzed in the literature, Time-Hopping CDMA (TH-CDMA) remains a niche subject, often associated with impulsive ultra-wideband communications; as such, it has been poorly investigated in its information-theoretical limits. This paper attempts to fill the gap, by addressing a reference basic case of synchronous, power-controlled systems, with random hopping.

Time-hopping systems transmit pulses over a subset of chips of cardinality N_s out of the N chips composing a symbol period. In contrast to common DS-CDMA, where each chip carries one pulse, and therefore the number of transmitted pulses per symbol is equal to the number of chips, *i.e.*, $N_s = N$, time-hopping signals may contain much fewer chips in which pulses are effectively used, *i.e.*, $N_s \ll N$. Asymptotically, if the number of used chips is fixed, as the number of chips in a symbol period grows, the fraction of filled-in chips in TH vanishes, *i.e.*, $N_s/N \to 0$, making TH intrinsically different, the performance of which cannot be derived from that of DS. TH *vs.* DS reflect "sparse" *vs.* "dense" spreading, where degrees of freedom, that is, dimensions of the signal space, are "unevenly" *vs.* "evenly" used [1–5]. In our setting, as further explored in the paper, degrees of freedom coincide with chips; while DS "evenly" uses chips, TH adopts the opposite strategy. In this regard, it is evident that DS and TH represent two contrasting approaches,

that will be compared, under the assumption of same bandwidth and same per-symbol energy, in terms of spectral efficiency.

Although we will show that there exist peculiar theoretical limits for TH-CDMA, their derivation can be carried out within the framework developed by Verdú and Shamai [6] and Shamai and Verdú [7], providing a methodology that is valid for investigating general CDMA with random spreading in the so-called *large-system limit* (LSL), where $K \to \infty$, $N \to \infty$, while $K/N \to \beta$ finite; in particular, [6] provides expressions of spectral efficiency for DS power-controlled systems using optimum as well as linear receivers, while [7] removes the power-control assumption and introduces fading. Other seminal contributions towards the understanding of random DS-CDMA, although limited to linear receivers, are those of Tse and Hanly [8], and Tse and Zeitouni [9]. Aside from DS-CDMA, the same framework is aptly used for analyzing other CDMA channels, such as multi-carrier CDMA [10].

The analysis of optimum decoders relies, in general, on the study of the eigenvalue distribution of random matrices describing random spreading. Consolidated results on the statistical distribution of such eigenvalues of DS matrices [11] form the basis for a tractable analysis of theoretical limits in terms of spectral efficiency. In particular, it is shown in [6] that a fixed loss, that depends upon the load, *i.e.*, the ratio β between the number of users K and chips N, is incurred with DS *vs.* orthogonal multiple-access. This loss becomes negligible with optimum decoding when $\beta \gg 1$ while, for $\beta \ll 1$, even a linear receiver such as MMSE is sufficient for achieving this negligible loss; however, this is no longer the case for simpler linear receivers, such as the single-user matched filter (SUMF), that is shown to be limited in spectral efficiency at high SNR. As a matter of fact, the above findings on spectral efficiency of DS-CDMA strongly depend on the statistical properties of the eigenvalue distribution, and as such on the cross-correlation properties of the spreading sequences. By changing the spreading strategy from DS to TH, it can be predicted that different theoretical limits will hold, as will be investigated below. In particular, TH matrices, as rigorously defined in this paper, are a special subset of sparse matrices, where the number of nonzero entries is small compared to the total number of elements. Previous work on sparse CDMA relies on non-rigorous derivations based on replica methods, which are analytical tools borrowed from statistical physics, as pioneered by Tanaka [12], who provides an expression of capacity when inputs are binary. Montanari and Tse [13] propose a rigorous argument for $N_s \to \infty$, proving Tanaka's formula, that is valid up to a maximum load, called *spinodal* ($\beta_s \approx 1.49$). Above the spinodal load, Tanaka's formula remains unproved. Binary sparse CDMA is also analyzed in terms of detection algorithms, in particular in the so-called belief propagation [13–15]. More recently, capacity bounds for binary sparse CDMA are derived in [16, 17]. Still relying on replica methods, [18] and [19] analyze two different regimes, where N_s is either finite or random with fixed mean.

The main contribution of the present work is to provide rigorous information-theoretical limits of time-hopping communications, by inscribing this particular time-domain sparse multiple access scheme into the random matrix framework developed by Verdú and Shamai in [6], for analyzing random spreading. The

present analysis allows comparing TH *vs.* DS with same energy per symbol and same bandwidth constraints, and, therefore, highlights the effect of the energy "concentration," that is typical of TH. A first contribution consists in providing a closed form expression for spectral efficiency of TH with optimum decoding when $N_s = 1$. A second contribution is to prove that the spectral efficiency formula for a bank of single-user matched filter obtained by Verdú and Shamai in [6] for DS systems ($N_s = N$) remains valid if $N_s \to \infty$, $N \to \infty$, and $N_s/N \to \alpha \in (0, 1]$. A third contribution is to provide understanding of when TH performs better than DS.

Based on the above contributions, we are able to present a novel interpretation of TH-CDMA against DS-CDMA, that offers a better understanding of the effect of sparsity in time.

The chapter is organized as follows: in Section 1.1 we describe the model of the synchronous CDMA channel adopted throughout the chapter, and particularized to the special case of time-hopping. Section 1.2 contains the derivation of spectral efficiency of TH-CDMA for different receiver structures, in particular optimum decoding as well as sub-optimal linear receivers, and a comparison with traditional DS-CDMA limits [6]. Conclusions are drawn in Section 4.4.

1.1 Reference Model

We consider the traditional complex-valued multiple access channel model with "no-fading" where the received signal $\mathsf{y}(t)$ is:

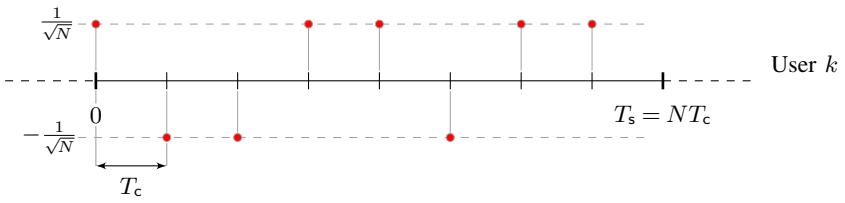
$$\mathsf{y}(t) = \sum_{k=1}^{K} \mathsf{A}\, \mathsf{b}_k \mathsf{s}_k(t) + \mathsf{n}(t), \tag{1.1}$$

where K is the number of users, $\{\mathsf{b}_k \colon k{=}1, \ldots, K\}$ is the set of transmitted symbols, $\mathsf{n}(t)$ the complex Additive White Gaussian Noise process with real and imaginary parts modeled as independent white Gaussian processes both characterized by double-side power spectral density $\mathcal{N}_0/2$, and $\mathsf{s}_k(t)$ is the unit-energy spreading waveform of user k. Based on the "no-fading" hypothesis, coefficient $\mathsf{A} > 0$ in (1.1) is common to all users and for simplicity normalized to one. Under the synchronous hypothesis, the above model that considers only one symbol b_k per user, is sufficient, that is, it can provide a sufficient statistic for optimum detection of $\{\mathsf{b}_k \colon k{=}1, \ldots, K\}$ [20].

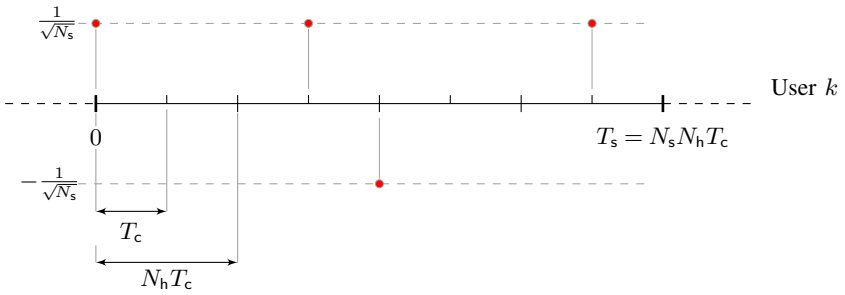
Each spreading waveform, $\mathsf{s}_k(t)$, can be written as the superposition of $N > 0$ orthonormal functions $\{\psi_n(t) \colon n{=}0, \ldots, N-1\}$, that is:

$$\mathsf{s}_k(t) = \sum_{n=0}^{N-1} \mathsf{s}_k[n] \psi_n(t).$$

Typically, a single unit-energy function $\psi(t)$ generates the whole set of orthonormal functions by translation; In this case, denoting with T_s the symbol period, any waveform with autocorrelation function satisfying the Nyquist criterion for a time



(a) DS-CDMA: $N = 8$.



(b) TH-CDMA: $N = 8$, $N_s = 4$, $N_h = 2$.

Figure 1.1: DS-CDMA *vs.* TH-CDMA time-axis structure. The symbol period is divided into $N = 8$ chips in both figures. In DS-CDMA (Fig. 1.1a), each chip is used for transmitting one pulse, hence eight pulses are transmitted per symbol period. The signature sequence shown on figure is $\boldsymbol{s}_k = \frac{1}{\sqrt{8}}[1, -1, -1, 1, 1, -1, 1, 1]^T$. In TH-CDMA (Fig. 1.1b) the symbol period is divided into $N_s = 4$ subgroups of $N_h = 2$ contiguous chips: one pulse only per subgroup is transmitted, that is four pulses in total. The signature sequence shown on figure is $\boldsymbol{s}_k = \frac{1}{2}[1, 0, 0, 1, -1, 0, 0, 1]^T$. Total energy per symbol is identical in both cases, and equal to one.

shift $T_c := T_s/N$, *i.e.* any waveform that is orthogonal to its translated version by multiples of T_c, may generate this set, meanwhile producing ISI-free symbol sequences at the receiver. A typical generating function set is thus:

$$\{\psi_n(t) := \psi(t - nT_c) \colon n = 0, \ldots, N-1\},$$

where T_c is called *chip time*, $\psi(t)$ *chip waveform*, and N is the number of *chips*.

In the present analysis, $\psi(t)$ is, for the sake of simplicity, the minimum bandwidth, that is, zero-excess bandwidth waveform, with unit-energy, and bandlimited to $[-W/2, W/2]$ with $W = 1/T_c$. This choice is common although specific, since there may be infinite possible bandlimited waveforms exceeding the minimum bandwidth and still appropriate, *vs.* infinite possible unlimited bandwidth waveforms, time-limited with duration lower than T_c (see for example Pursley [21] in particular for DS-CDMA with time-limited chip waveforms). The choice of a minimum bandwidth waveform implies in our case that $\psi(t)$ is not time-limited.

By projecting the received signal $\mathrm{y}(t)$ onto the set of orthonormal functions $\{\psi_n(t)\colon n=0,\dots,N-1\}$, a sufficient statistic for optimum detection is obtained:

$$\boldsymbol{y} = \boldsymbol{S}\boldsymbol{b} + \boldsymbol{n}, \tag{1.2}$$

where $\boldsymbol{y} \in \mathbb{C}^N$, and $\boldsymbol{S} = [\boldsymbol{s}_1, \cdots, \boldsymbol{s}_K] \in \mathbb{R}^{N\times K}$ is the spreading matrix,

$$\boldsymbol{S} := \begin{bmatrix} \mathrm{s}_1[0] & \mathrm{s}_2[0] & \cdots & \mathrm{s}_K[0] \\ \mathrm{s}_1[1] & \mathrm{s}_2[1] & \cdots & \mathrm{s}_K[1] \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{s}_1[N-1] & \mathrm{s}_2[N-1] & \cdots & \mathrm{s}_K[N-1] \end{bmatrix}.$$

In addition, $\boldsymbol{b} = [\mathrm{b}_1, \dots, \mathrm{b}_K]^\mathsf{T} \in \mathbb{C}^K$, and $\boldsymbol{n} \in \mathbb{C}^N$ is a circularly symmetric Gaussian vector with zero mean and covariance $\mathcal{N}_0\boldsymbol{I}$. Since the spreading waveforms have unit-energy, the signature sequences have unit norm: $\|\boldsymbol{s}_k\|^2 = 1$.

Matrix $\boldsymbol{S}$ structure is appropriate for describing spread-spectrum systems in general.

In DS-CDMA, the spreading sequences are typically modeled as binary, where:

$$\mathrm{s}_k[i] \in \{-1/\sqrt{N}, +1/\sqrt{N}\}, \quad i = 0, \dots, N-1,$$

is drawn with uniform probability, or spherical, where $\boldsymbol{s}_k$ is a Gaussian random vector with unit norm [6].

In order to cast TH-CDMA in the model described by eq. (1.2), let $N = N_\mathrm{s} \cdot N_\mathrm{h}$, that is the N chips are divided into N_s subgroups, and each of these N_s subgroups is made of N_h contiguous chips. In this case, elements of the signature sequence can take the following values:

$$\mathrm{s}_k[i] \in \{-1/\sqrt{N_\mathrm{s}}, 0, +1/\sqrt{N_\mathrm{s}}\}, \quad i = 0, \dots, N-1,$$

and the structure of the sequence is such that there is one and only one non-zero $\mathrm{s}_k[m]$ within each of the N_s subgroups. Therefore, the number of non-zero elements of each signature sequence is fixed to N_s. Note that, for $N_\mathrm{s} = N$, TH-CDMA reduces to DS-CDMA.

We formally introduce the new structure of spreading sequences by the two following definitions.

Definition 1 (*Sparse vector*). A vector $\boldsymbol{s} = [s_1, \dots, s_N]^\mathsf{T} \in \mathbb{C}_N$ is S-sparse if the subset of its nonzero elements has cardinality S, *i.e.*, $|\{s_i \neq 0\colon i = 1, \dots, N\}| = S$.

Definition 2 ($(N_\mathrm{s}, N_\mathrm{h})$-*sequence, TH and DS sequences and matrices*). A vector $\boldsymbol{s} = [s_1, \dots, s_N]^\mathsf{T} \in \mathbb{C}^{N\times 1}$ is a $(N_\mathrm{s}, N_\mathrm{h})$-sequence when:

1. $N = N_\mathrm{s} \cdot N_\mathrm{h}$, with $N_\mathrm{s} \in \mathbb{N}$ and $N_\mathrm{h} \in \mathbb{N}$;
2. for all $1 \leq m \leq N_\mathrm{s}$, the vector $[s_{1+(m-1)N_\mathrm{h}}, \dots, s_{mN_\mathrm{h}}]^\mathsf{T}$ is 1-sparse, where the nonzero element is either $-1/\sqrt{N_\mathrm{s}}$ or $1/\sqrt{N_\mathrm{s}}$ with equal probability.

A (N_s, N_h)-sequence with $N_s < N$ is a Time-Hopping (TH) sequence; the special case $N_s = N$, *i.e.*, $(N, 1)$-sequences corresponds to binary DS sequences, that will be referred to below simply as DS sequences. A matrix $\boldsymbol{S}$ is called TH *vs.* DS matrix when its columns correspond to TH *vs.* DS sequences. The set of all possible TH *vs.* DS matrices is indicated as TH *vs.* DS ensemble.

Figure 1.1 shows the organization of the time axis for DS-CDMA (Fig. 1.1a) and compares this time pattern against TH-CDMA (Fig. 1.1b).

The unit-norm assumption on spreading sequences implies that comparison of TH-CDMA *vs.* DS-CDMA is drawn under the constraint of same energy per sequence. Note that the $N_s = 1$ case models a strategy of maximum energy concentration in time, while maximum energy spreading in time corresponds to making $N_s = N$, as in DS. Also note that, the two systems operate under same bandwidth constraint given the hypothesis on $\psi(t)$.

1.2 Spectral Efficiency of TH-CDMA

In this section, spectral efficiency of TH-CDMA is derived for different receiver structures, and compared against consolidated results for DS-CDMA [6].

The section is organized as follows: we first analyze the case of optimum decoding (Section 1.2.1), then proceed to linear receivers in sections 1.2.2 and 1.2.3 for single-user matched filters (SUMF), and decorrelator/MMSE receivers, respectively. Finally, Section 1.2.4 contains a synposis.

1.2.1 Optimum decoding

Theoretical framework

In general terms, a key performance measure in the coded regime is spectral efficiency C^{opt} (b/s/Hz) as a function of either signal-to-noise ratio γ or energy per bit $\mathcal{E}_b$-to-noise-$\mathcal{N}_0$, $\mathcal{E}_b/\mathcal{N}_0$.

Referring to model of eq. (1.2), where the dimension of the observed process is N, spectral efficiency is indicated as $C_N^{\text{opt}}(\gamma)$ and is the maximum mutual information between $\boldsymbol{b}$ and $\boldsymbol{y}$ knowing $\boldsymbol{S}$ over distributions of $\boldsymbol{b}$, normalized to N. $C_N^{\text{opt}}(\gamma)$ (b/s/Hz) is achieved with Gaussian distributed $\boldsymbol{b}$, and it is expressed by [6, 22–24]:

$$C_N^{\text{opt}}(\gamma) = \frac{1}{N} \log_2 \det(\boldsymbol{I} + \gamma \boldsymbol{S}\boldsymbol{S}^{\mathsf{T}}), \tag{1.3}$$

where noise has covariance $\boldsymbol{\Sigma_n} = \mathcal{N}_0 \boldsymbol{I}$ and γ is given by [25]:

$$\gamma := \frac{\frac{1}{K}\mathbb{E}\left[\|\boldsymbol{b}\|^2\right]}{\frac{1}{N}\mathbb{E}\left[\|\boldsymbol{n}\|^2\right]} = \frac{\frac{1}{K} \cdot b\mathcal{E}_b}{\frac{1}{N} \cdot N\mathcal{N}_0} = \frac{1}{\beta} \cdot \frac{b}{N} \cdot \frac{\mathcal{E}_b}{\mathcal{N}_0} = \frac{1}{\beta} \cdot C_N^{\text{opt}} \cdot \eta, \tag{1.4}$$

where $\beta := K/N$ is the *load*, $\eta := \mathcal{E}_b/\mathcal{N}_0$, b is the number of bits encoded in $\boldsymbol{b}$ for a capacity-achieving system, and therefore b/N coincides with spectral efficiency

C_N^{opt} of eq. (1.3). Since N is equal to the number of possible complex dimensions, spectral efficiency can, therefore, be interpreted as the maximum number of bits per each complex dimension. Note that the number of complex dimensions coincides in our setting with the degrees of freedom of the system, that is, with the dimension of the observed signal space.

Eq. (1.3) can be equivalently rewritten in terms of the set of eigenvalues $\{\lambda_n(\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}})\colon n=1,\dots,N\}$ of the Gram matrix $\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}}$ as follows:

$$C_N^{\mathrm{opt}}(\gamma)=\frac{1}{N}\sum_{n=1}^{N}\log_2(1+\lambda_n\gamma)=\int_0^{\infty}\log_2(1+\lambda\gamma)\,d\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}}}(\lambda), \tag{1.5}$$

where $\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}}}(x)$ is the so called *empirical spectral distribution* (ESD) defined as [22]:

$$\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}}}(x):=\frac{1}{N}\sum_{n=1}^{N}\mathbb{1}\{\lambda_n(\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}})\le x\}, \tag{1.6}$$

that counts the fraction of eigenvalues of $\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}}$ not larger than x. Being $\boldsymbol{S}$ random, so is the function $\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}}}$, though the limit distribution F of the sequence $\{\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}}}\colon N\ge 1\}$, called *limiting spectral distribution* (LSD), is usually nonrandom [26]. In particular, the regime of interest, referred to as *large-system limit* (LSL), is that of both $N\to\infty$ and $K\to\infty$ while keeping $K/N\to\beta$ finite. Spectral efficiency in the LSL is:

$$C^{\mathrm{opt}}(\gamma):=\int_0^{\infty}\log_2(1+\lambda\gamma)\,d\mathsf{F}(\lambda). \tag{1.7}$$

Therefore, finding the spectral efficiency of CDMA systems with random spreading in the LSL regime reduces to finding the LSD $\mathsf{F}(\lambda)$, that depends on the spreading sequence family only; hence, in the rest of this section, we find the LSD of TH-CDMA with $N_{\mathrm{s}}=1$, which corresponds to a maximum energy concentration in time, as well as asymptotic behaviors of TH-CDMA systems with generic N_{s}.

LSD and spectral efficiency of TH-CDMA systems with $N_{\mathrm{s}}=1$

While for DS-CDMA, spectral efficiency can be computed directly from Marčenko and Pastur result on the ESD of matrices with i.i.d. elements [11], it appears that no analog result is available for neither TH-CDMA matrices nor dual matrices describing frequency-hopping.

We hereby derive the LSD and properties of the ESD of synchronous TH-CDMA when $N_{\mathrm{s}}=1$.

Theorem 1. *Suppose that $\boldsymbol{S}\in\mathbb{R}^{N\times\beta N}$ is a time-hopping matrix, as specified in Definition 2, with $N_{\mathrm{s}}=1$. Then, the ESD of $\{\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}}\colon N\ge 1\}$ converges in*

probability to the distribution function F *of a Poisson law with mean* β*:*

$$\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^\mathsf{T}}(x) \xrightarrow{p} \mathsf{F}(x) = \sum_{k\geq 0} \frac{\beta^k e^{-\beta}}{k!} \mathbb{1}\{k \leq x\}. \tag{1.8}$$

Proof. Denote by $\pi_k \in [N]$ the nonzero element of the kth column $\boldsymbol{s}_k$ of $\boldsymbol{S}$. Then:

$$\boldsymbol{S}\boldsymbol{S}^\mathsf{T} = \sum_{k=1}^{K} \boldsymbol{s}_k \boldsymbol{s}_k^\mathsf{T} = \sum_{k=1}^{K} \boldsymbol{e}_{\pi_k} \boldsymbol{e}_{\pi_k}^\mathsf{T},$$

where $[\boldsymbol{e}_i]_j = \delta_{ij}$, being δ_{ij} the Kronecker symbol. Hence, $\boldsymbol{S}\boldsymbol{S}^\mathsf{T}$ is diagonal, and the nth element on the diagonal, denoted $\nu_n := [\boldsymbol{S}\boldsymbol{S}^\mathsf{T}]_{nn}$, is equal to:

$$\nu_n = |\{k \in [K] \colon \pi_k = n\}| \in [K].$$

The set of eigenvalues $\{\lambda_n \colon n \in [N]\}$ is equal to $\{\nu_n \colon n \in [N]\}$, and, therefore, eigenvalues belong to non-negative integers. The ESD $\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^\mathsf{T}}(x)$ can be written as follows:

$$\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^\mathsf{T}}(x) := \frac{1}{N}\sum_{n=1}^{N} \mathbb{1}\{\lambda_n \leq x\} = \frac{1}{N}\sum_{n=1}^{N} \mathbb{1}\{\nu_n \leq x\} = \frac{1}{N}\sum_{s\leq x} |\{i \in [N] \colon [\boldsymbol{S}\boldsymbol{S}^\mathsf{T}]_{ii} = s\}|, \tag{1.9}$$

where it is intended that the upper bound of the last summation is $\lfloor x \rfloor$. In general, when the ith diagonal element of $\boldsymbol{S}\boldsymbol{S}^\mathsf{T}$ is equal to s, we say that s users are in chip i. Therefore, the last equality indicates that $\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^\mathsf{T}}(x)$ is the fraction of chips with at most x users. We will find the generating function (GF) of:

$$\mathsf{A}(x) := \sum_{s\leq x} |\{i \in [N] \colon [\boldsymbol{S}\boldsymbol{S}^\mathsf{T}]_{ii} = s\}|, \tag{1.10}$$

and we will show that $\mathbb{E}\big[\,\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^\mathsf{T}}(x)\,\big] \to \mathsf{F}(x) := \sum_{0\leq s\leq x} e^{-\beta}\beta^s/s!$ and $\mathrm{Var}\big[\,\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^\mathsf{T}}(x)\,\big] = O(1/N)$, which in particular implies $\mathrm{Var}\big[\,\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^\mathsf{T}}(x)\,\big] \to 0$ in the LSL, therefore proving convergence in probability of $\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^\mathsf{T}}(x)$ to $\mathsf{F}(x)$.

In order to find the GF of $\mathsf{A}(x)$, we use the symbolic method of analytic combinatorics [27, 28]. In our setting, atoms are users, and we define the following combinatorial class:

$$\mathcal{A} = \mathsf{Seq}_N(u\mathsf{Set}_{\leq x}\mathcal{Z} + \mathsf{Set}_{>x}\mathcal{Z}),$$

where $\mathcal{Z}$ is the class containing a single atom, Seq and Set are two basic constructions in analytic combinatorics, and u marks the number of chips with at most x users, *i.e.*, $\mathsf{A}(x)$. $\mathcal{A}$ is mapped to the following bivariate GF:

$$A_x(z,u) = \left(u\sum_{i\leq x}\frac{z^i}{i!} + \sum_{i>x}\frac{z^i}{i!}\right)^N = \left[e^z + (u-1)\phi_x(z)\right]^N, \tag{1.11}$$

where $\phi_x(z) := \sum_{i \le x} z^i/i!$. From $A_x(z,u)$, we can derive $\mathbb{E}[\,\mathsf{A}(x)\,]$ and $\mathbb{E}\big[\,\mathsf{A}(x)^2\,\big]$, as follows [27]:

$$\mathbb{E}[\,\mathsf{A}(x)\,] = \frac{[z^K]\frac{\partial A_x}{\partial u}\big|_{u=1}}{[z^K]A_x(z,1)}, \tag{1.12}$$

$$\mathbb{E}\big[\,\mathsf{A}(x)^2\,\big] = \frac{[z^K]\frac{\partial^2 A_x}{\partial u^2}\big|_{u=1}}{[z^K]A_x(z,1)} + \mathbb{E}[\,\mathsf{A}(x)\,]. \tag{1.13}$$

Hence, from eqs. (1.11) and (1.12), one has:

$$\mathbb{E}\big[\,\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}}}(x)\big] = \sum_{i \le x} \frac{1}{i!} \cdot \frac{K!/(K-i)!}{N^i} \cdot \left(1 - \frac{1}{N}\right)^{K-i} \longrightarrow \sum_{i \le x} \frac{\beta^i e^{-\beta}}{i!}, \tag{1.14}$$

while from eqs. (1.11) and (1.13), one has:

$$\mathbb{E}\Big[\,\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}}}(x)^2\Big] = \left(1 - \frac{1}{N}\right) \sum_{i \le x} \sum_{j \le x} \frac{1}{i!j!} \cdot \frac{K!/(K-i-j)!}{N^{i+j}} \cdot \left(1 - \frac{2}{N}\right)^{K-i-j} + \frac{1}{N}\,\mathbb{E}\big[\,\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}}}(x)\big], \tag{1.15}$$

and, therefore, $\mathrm{Var}\big[\,\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}}}(x)\big]$ is:

$$\mathrm{Var}\big[\,\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}}}(x)\big] = \sum_{i \le x} \sum_{j \le x} \frac{1}{i!j!} \cdot \frac{K!/(K-i-j)!}{N^{i+j}} \cdot \left[\left(1 - \frac{2}{N}\right)^{K-i-j} - \frac{(K-i-j)!K!}{(K-i)!(K-j)!}\left(1 - \frac{1}{N}\right)^{2K-i-j}\right] + O\left(\frac{1}{N}\right).$$

The quantity in brackets is $O(1/N)$ since:

$$\begin{aligned} (1-2/N)^{K-i-j} &= e^{-2\beta} + O(1/N), \\ \{(K-i-j)!K!\}/\{(K-i)!(K-j)!\} &= 1 + O(1/N), \\ (1-1/N)^{2K-i-j} &= e^{-2\beta} + O(1/N). \end{aligned}$$

Hence:

$$\mathrm{Var}\big[\,\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}}}(x)\big] = \sum_{i \le x} \sum_{j \le x} \frac{1}{i!j!}\left[\beta^{i+j} + O\left(\frac{1}{N}\right)\right] O\left(\frac{1}{N}\right) = O\left(\frac{1}{N}\right),$$

uniformly in x. □

In Theorem 2 we derive properties of the Lth moment of the ESD $\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}}}$ with $N_{\mathsf{s}} = 1$, denoted by:

$$m_L := \frac{1}{N}\,\mathsf{Tr}(\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}})^L = \int_0^\infty \lambda^L\, d\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}}}(\lambda),$$

in particular a closed form expression of $\mathbb{E}[\,m_L\,]$ for TH matrices with $N_{\mathsf{s}}=1$ for finite K and N, and we prove convergence in probability to moments of a Poisson distribution with mean β in the LSL.

Theorem 2. *Suppose that $\boldsymbol{S}\in\mathbb{R}^{N\times\beta N}$ is a time-hopping matrix with $N_{\mathsf{s}}=1$. Then, the expectation of the Lth moment m_L is:*

$$\mathbb{E}[\,m_L\,]:=\frac{1}{N}\,\mathbb{E}\,\mathsf{Tr}\Big[(\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}})^L\Big]=\sum_{\ell=1}^{L}\left\{{L\atop\ell}\right\}\frac{K!}{(K-\ell)!}\cdot\frac{1}{N^\ell},\tag{1.16}$$

where $\left\{{L\atop\ell}\right\}$ denotes a Stirling number of the second kind. In the LSL, m_L converges in probability to the Lth moment of a Poisson distribution with mean β, i.e.:

$$m_L\xrightarrow{p}\overline{m}_L:=\sum_{\ell=1}^{L}\left\{{L\atop\ell}\right\}\beta^\ell.$$

that is .

Proof. See Appendix 1.B. □

Lemma 1 (Verifying the Carleman condition). *The sequence of moments $(\overline{m}_L)_{L\geq 0}$ verifies the Carleman condition, i.e.,*

$$\sum_{k\geq 1}\overline{m}_{2k}^{-1/2k}=\infty.$$

Proof. We upper bound $\overline{m}_{2k}$ as follows:

$$\begin{aligned}
\overline{m}_{2k}&\overset{\text{(a)}}{<}\sum_{\ell=1}^{2k}\left\{{2k\atop\ell}\right\}\beta^\ell\\
&\overset{\text{(b)}}{\leq}\sum_{\ell=1}^{2k-1}\frac{1}{2}\binom{2k}{\ell}\ell^{2k-\ell}\beta^\ell+\beta^{2k}\\
&\overset{\text{(c)}}{\leq}\frac{1}{2}(2k-1)^{2k-1}\sum_{\ell=1}^{2k-1}\binom{2k}{\ell}\beta^\ell+\beta^{2k}\\
&\overset{\text{(d)}}{<}\frac{1}{2}(2k-1)^{2k-1}(1+\beta)^{2k}+\beta^{2k}\\
&<(1+\beta)^{2k}(1+(2k)^{2k})
\end{aligned}$$

where: (a) follows from the elementary inequality $(2k)!/(2k-\ell)!=(2k)(2k-1)\cdots(2k-\ell+1)<(2k)^\ell$; (b) from the inequality $\left\{{n\atop\ell}\right\}\leq U(n,\ell):=\frac{1}{2}\binom{n}{\ell}\ell^{n-\ell}$; (c) from upper bounding the term $\ell^{2k-\ell}$ with $(2k-1)^{2k-1}$; (d) from extending the summation over $\ell=0,\ldots,2k$.

From elementary relations between p-norms, one has $(1+(2k)^{2k})^{1/2k} = \|(1,2k)\|_{2k} \leq \|(1,2k)\|_1 = 1+2k$, thus $\overline{m}_{2k}^{1/2k} < (1+\beta)(1+2k)$, and therefore:

$$\sum_{k\geq 1} \overline{m}_{2k}^{-1/2k} > \frac{1}{1+\beta} \sum_{k\geq 1} \frac{1}{1+2k} = \infty,$$

which verifies the Carleman condition. □

In terms of measures, TH-CDMA is thus characterized by the purely atomic measure given by:

$$\mu^{\mathsf{TH}} := \sum_{k\geq 0} \frac{\beta^k e^{-\beta}}{k!}\, \delta_k = \sum_{k\geq 0} \mathsf{f}_k(\beta)\delta_k, \tag{1.17}$$

being $\mathsf{f}_k(\beta) := \beta^k e^{-\beta}/k!$, and δ_k the point mass distribution, *i.e.*, $\delta_k(A) = 1$ if $k \in A$, and $\delta_k(A) = 0$ otherwise. Whence, $\mathsf{F}(x) = \mu^{\mathsf{TH}}((-\infty, x]) = \mu^{\mathsf{TH}}([0,x])$. The above implies peculiar properties of TH-CDMA when compared against DS-CDMA. For convenience, we report here the Marčenko-Pastur law, that is the LSD of eigenvalues of DS-CDMA matrices (see Definition 2), which has measure:

$$\mu^{\mathsf{DS}} = (1-\beta)^+ \delta_0 + \mu_{\mathsf{ac}}^{\mathsf{DS}}, \tag{1.18}$$

where $(x)^+ := \max\{0, x\}$, and $\mu_{\mathsf{ac}}^{\mathsf{DS}}$ is the absolute continuous part of μ^{DS} with density (Radon-Nikodym derivative with respect to the Lebesgue measure m):

$$\frac{d\mu_{\mathsf{ac}}^{\mathsf{DS}}}{dm}(x) = \frac{1}{2\pi\lambda}\sqrt{-(\lambda-\ell^+)(\lambda-\ell^-)}\,\mathbb{1}\{x\in[\ell^-,\ell^+]\} =: \mathsf{f}^{\mathsf{MP}}(x), \tag{1.19}$$

where $\ell^\pm = (1\pm\sqrt{\beta})^2$.

Fig. 1.2 shows Marčenko-Pastur and Poisson laws for $\beta = 1/2$. The Marčenko-Pastur law has, in general, an absolute continuous part with probability density function showed in solid line and an atomic part formed by a point mass at the origin showed with a cross at height $1/2$. The Poisson law has a purely atomic (also known as discrete, or counting) measure with point masses at nonnegative integers showed by dots with heights given by $\mathsf{f}_k(\beta)$ (envelope showed in dashed line).

We use the Poisson LSD to find the spectral efficiency of TH-CDMA with $N_{\mathsf{s}} = 1$ in the LSL, *i.e.* (see eq. (1.5)):

$$\int_0^\infty \log_2(1+\lambda\gamma)\, d\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}}}(\lambda) \xrightarrow{p} \int_0^\infty \log_2(1+\lambda\gamma)\, d\mathsf{F}(\lambda). \tag{1.20}$$

It is important to remark that the above convergence in probability does not follow immediately; in fact, convergence in law does only imply convergence of bounded functionals, but $\log_2(1+\lambda\gamma)$ is not bounded on the support of $\mathsf{F}(\lambda)$. We prove eq. (1.20) in Appendix 1.A, and thus:

$$C_N^{\mathsf{opt}}(\gamma) \xrightarrow{p} C^{\mathsf{opt}}(\gamma) = \sum_{k\geq 0} \frac{\beta^k e^{-\beta}}{k!} \log_2(1+k\gamma). \tag{1.21}$$

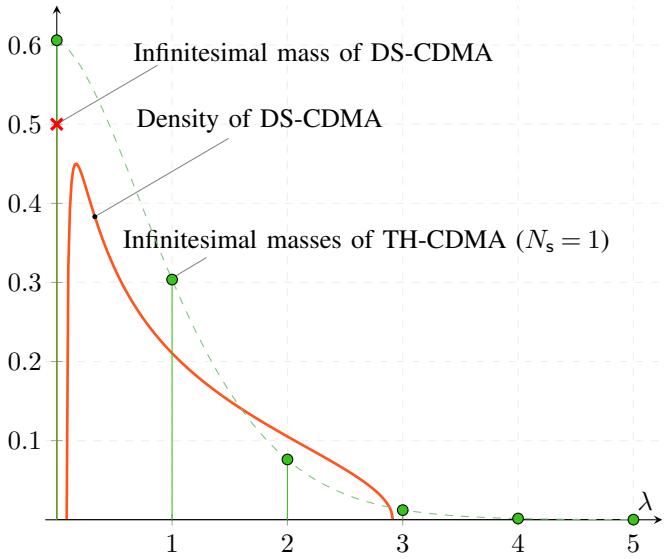


Figure 1.2: Density function of the LSD for DS-CDMA in solide line, and infinitesimal masses of atomic measures of DS-CDMA and TH-CDMA (Poisson law) for $\beta = 1/2$ in cross and dots, respectively. DS-CDMA and TH-CDMA with $N_s = 1$ are governed by Marčenko-Pastur and Poisson laws, respectively.

The capacity of a TH-CDMA system with $N_s = 1$ can be interpreted as follows. Rewrite eq. (1.21) as follows:

$$C^{\text{opt}}(\gamma) = \sum_{k \geq 0} f_k(\beta)\, C_k(\gamma), \tag{1.22}$$

where $C_k(\gamma) := \log_2(1 + k\gamma)$. Hence, $C^{\text{opt}}(\gamma)$ is a sum of channel capacities $C_k(\gamma)$, $k \in \mathbb{N}$, weighted by probabilities $f_k(\beta)$. Since $C_k(\gamma)$ is the capacity of a complex AWGN channel with signal-to-noise ratio $k\gamma$, $k \in \mathbb{N}$, $C^{\text{opt}}(\gamma)$ is equal to the capacity of an infinite set of complex AWGN channels with increasing signal-to-noise ratio $k\gamma$ paired with decreasing probability of being used $f_k(\beta)$. Therefore, TH-CDMA has the same behavior of an access scheme that splits the multiaccess channel into independent channels, each corrupted by noise only, with power gain equal to k, and excited with probability $f_k(\beta)$. Since $f_k(\beta)$ is also the probability that k signatures have their nonzero element in the same dimension, that is for TH-CDMA associated with the event of waveforms having their pulse over the same chip, for small β, that is, $\beta \leq 1$, channels with high capacity (for a fixed γ), that is, with $k \gg 1$, are less frequently used than channels with low capacity; in general, channels with k in a neighborhood of β are used most frequently.

One noticeable difference between DS and TH matrices is that in the former the maximum eigenvalue $\lambda_{\max} \xrightarrow{\text{a.s.}} (1+\sqrt{\beta})^2$ [29], and thus also $\lambda_{\max} \xrightarrow{p} (1+\sqrt{\beta})^2$, while in the latter $\lambda_{\max} \xrightarrow{p} \infty$.

Moreover, there exists a nonzero probability $f_0(\beta) = e^{-\beta}$ such that, also for $\beta > 1$, the zero-capacity channel ($k = 0$) is excited. This probability, that is the

amplitude of the Dirac mass at $\lambda = 0$, is equal to $\mathsf{F}(0)$; it equals the probability that a chip is not chosen by any user or, equivalently, the average fraction of unused chips; and, finally, it equals the high-SNR slope *penalty*, as we will detail below.

It is interesting to analyze the behavior of $r_N := \operatorname{rank} \boldsymbol{S}/N$, that is a random variable for finite N. Figure 1.3 shows with marks $\bar{r}_N := \mathbb{E}[\, r_N \,]$ for TH-CDMA with $N_\mathrm{s} = 1$ and $N_\mathrm{s} = 2$, and for DS-CDMA, when $N = 50$: Monte-Carlo simulations provide point data, represented by marks, with error bars showing one standard deviation of r_N. Solid lines represent the limiting value r of r_N as $N \to \infty$. We will show in the below Theorem 3 that, for $N_\mathrm{s} = 1$, $r_N \xrightarrow{p} 1 - e^{-\beta}$. Almost sure convergence does hold for the Marčenko-Pastur law, hence for DS-CDMA one has $r_N \xrightarrow{\text{a.s.}} \min\{1, \beta\}$. For TH-CDMA with increasing N_s, one might expect r of TH-CDMA to tend to that of DS-CDMA, also suggested by the behavior of the $N_\mathrm{s} = 2$ case shown on figure. In the general $N_\mathrm{s} > 1$ case, we were able to find the upper bound $r \leq 1 - e^{-N_\mathrm{s}\beta}$ only, holding in probability, that is derived in the below Theorem 4, and shown with the dashed line.

Theorem 3. *Suppose that $\boldsymbol{S} \in \mathbb{R}^{N \times \beta N}$ is a time-hopping matrix with $N_\mathrm{s} = 1$. Then: $r_N := \frac{1}{N} \operatorname{rank} \boldsymbol{S} \xrightarrow{p} 1 - e^{-\beta}$.*

Proof. When $N_\mathrm{s} = 1$, each column of $\boldsymbol{S}$ is nonzero in one dimension only, hence, by indicating with $\boldsymbol{T} = [\boldsymbol{t}_1, \dots, \boldsymbol{t}_N] := \boldsymbol{S}^\mathsf{T}$, one has $\operatorname{rank} \boldsymbol{S} = |\{n \in [N] \colon \boldsymbol{t}_n \neq \boldsymbol{0}\}|$. In words, $\operatorname{rank} \boldsymbol{S}$ is equal to the number of nonempty raws of $\boldsymbol{S}$. Therefore, $r_N = 1 - |\{n \in [N] \colon \boldsymbol{t}_n = \boldsymbol{0}\}|/N = 1 - \mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^\mathsf{T}}(0)$, hence, by Theorem 1.8, $r_N \xrightarrow{p} 1 - \mathsf{F}(0)$.

Theorem 4. *Let $N_\mathrm{s} \geq 1$. An upper bound to r is given by:*

$$r \leq \min\{\beta, 1 - e^{-N_\mathrm{s}\beta}\}, \tag{1.23}$$

which holds in probability.

Proof. Rewrite $\boldsymbol{S}$ as follows: $\boldsymbol{S} = [\boldsymbol{S}_1^\mathsf{T}, \cdots, \boldsymbol{S}_{N_\mathrm{s}}^\mathsf{T}]^\mathsf{T}$, where $\{\boldsymbol{S}_i\}_{i=1}^{N_\mathrm{s}}$ are $N_\mathrm{h} \times K$ matrices, $N_\mathrm{h} = N/N_\mathrm{s}$. Using the inequality $\operatorname{rank}(\boldsymbol{A} + \boldsymbol{B}) \leq \operatorname{rank} \boldsymbol{A} + \operatorname{rank} \boldsymbol{B}$, we can upper bound $\operatorname{rank} \boldsymbol{S}$ as follows: $\operatorname{rank} \boldsymbol{S} \leq \sum_{i=1}^{N_\mathrm{s}} \operatorname{rank} \boldsymbol{S}_i$. Since $\{\boldsymbol{S}_i\}_{i=1}^{N_\mathrm{s}}$ are independent, by Theorem 3 one has $r \leq 1 - e^{-N_\mathrm{s}\beta}$ in probability. Moreover, since $\operatorname{rank} \boldsymbol{S}/N \leq \min\{1, \beta\}$ surely, we also have $\bar{r} \leq \beta$, and therefore eq. (1.23). □

Asymptotics

In the following, spectral efficiency, when expressed as a function of $\eta := \mathcal{E}_b/\mathcal{N}_0$, will be indicated by[1] C (b/s/Hz), as suggested in [25], rather than C (b/s/Hz), that denotes spectral efficiency as a function of γ. While an expression of C can be found in terms of the LSD, the same is more difficult for C, given the nonlinear relation between C and C: $\mathsf{C} = C(\eta\mathsf{C}/\beta)$ (c.f. eq. (1.4)).

In order to understand the asymptotic behavior of Cin the low-SNR and high-SNR regimes, *i.e.*, as $\eta \to \eta_{\min} := \inf_{\mathsf{C}>0} \eta(\mathsf{C})$ and $\eta \to \infty$, respectively, Shamai and Verdú [7] and Verdú [25] introduced the following four relevant parameters:

[1] In this subsection, we drop the superscript "opt" for ease of notation.

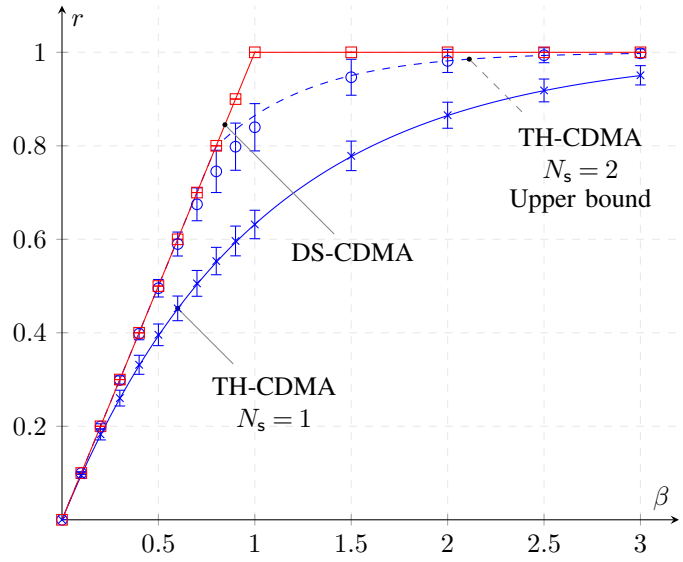


Figure 1.3: Normalized rank r (solid lines) *vs.* load β. The dashed line represents an upper bound of r for TH-CDMA with $N_s = 2$. Crosses, circles, and squares (generally referred to as marks), are obtained by evaluating $\mathbb{E}[\operatorname{rank} \boldsymbol{S}/N]$ by Monte-Carlo simulations of a finite-dimensional system with $N = 50$, for TH-CDMA with $N_s = 1$, TH-CDMA with $N_s = 2$, and DS-CDMA, respectively. Error bars represent one standard deviation of $\operatorname{rank} \boldsymbol{S}/N$.

$\eta_{\min}$: the *minimum energy per bit over noise level* required for reliable communication;

$\mathcal{S}_0$: the *wideband slope* (b/s/Hz/(3 dB));

$\mathcal{S}_\infty$: the *high-SNR slope* (b/s/Hz/(3 dB));

$\mathcal{L}_\infty$: the *high-SNR decibel offset*.

In our setting, the low-SNR and high-SNR regimes also correspond to $\mathsf{C} \to 0$ (so called wideband regime [25]) and $\mathsf{C} \to \infty$.

The minimum energy-per-bit $\eta_{\min}$ and the wideband slope $\mathcal{S}_0$ (b/s/Hz/(3 dB)) characterize the affine approximation of C *vs.* $\eta^{\mathrm{dB}} := 10 \log_{10} \eta$ as $\mathsf{C} \to 0$:

$$\eta^{\mathrm{dB}} = \eta_{\min}^{\mathrm{dB}} + \frac{10 \log_{10} 2}{\mathcal{S}_0} \mathsf{C} + o(\mathsf{C}), \quad \mathsf{C} \to 0. \tag{1.24}$$

From eq.s (1.24) and (1.4), one can find $\eta_{\min}$ and $\mathcal{S}_0$ as follows:

$$\eta_{\min} = \lim_{\gamma \downarrow 0} \frac{\beta\gamma}{C(\gamma)} = \frac{\beta}{C'(0)} = \frac{\beta}{\mathbb{E}[\lambda]} \ln 2, \tag{1.25}$$

$$\mathcal{S}_0 = -2 \ln 2 \frac{(C'(0))^2}{C''(0)} = 2 \frac{(\mathbb{E}[\lambda])^2}{\mathbb{E}[\lambda^2]}, \tag{1.26}$$

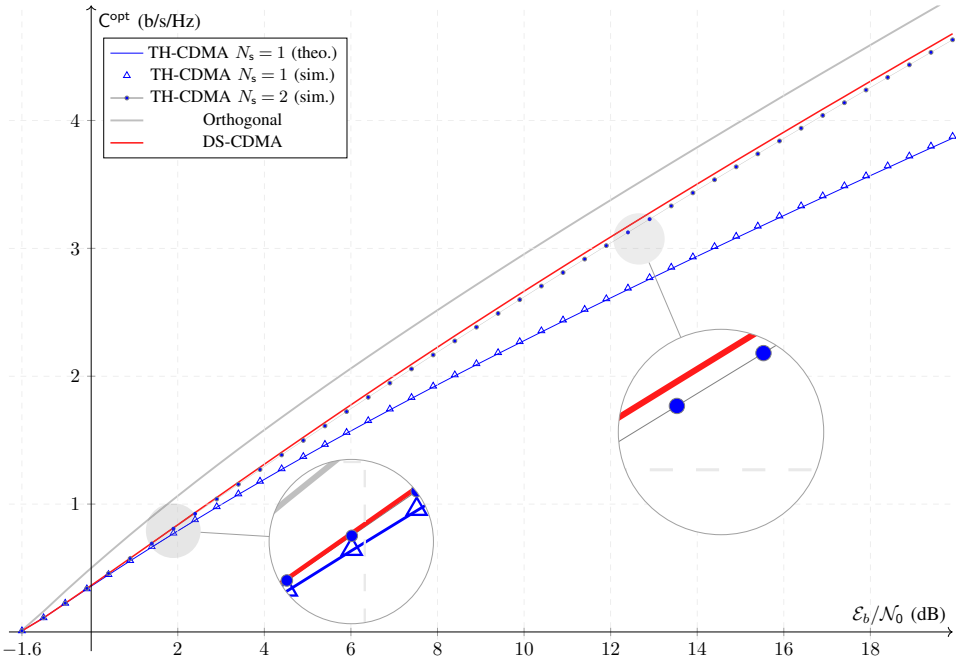


Figure 1.4: Spectral efficiency C^{opt} (b/s/Hz) of TH-CDMA *vs.* DS-CDMA with optimum decoding as a function of $\mathcal{E}_b/\mathcal{N}_0$ (dB) with load $\beta = 1/2$. Orthogonal multiple access is reported for comparison (gray solid line). Analytical expression *vs.* simulation are plotted for the $N_s = 1$ TH-CDMA case in blue solid line and blue triangles, respectively. Blue dots represent the $N_s = 2$ TH-CDMA case obtained by simulation only. DS-CDMA is shown with red solid line. Note on figure that TH-CDMA with $N_s = 1$ and $N_s = 2$ have both similar performance as DS in the wideband regime ($\mathcal{E}_b/\mathcal{N}_0 \to \ln 2$), while departing from it for high SNR when $N_s = 1$. Note on figure that the loss incurred with TH drops to a very small value with as early as $N_s = 2$.

where the expression in the last equality of both eqs. (1.25) and (1.26) is obtained by differentiating $\log_2(1+\lambda\gamma)$ with respect to λ under the integral sign.

The high-SNR slope $\mathcal{S}_\infty$ (b/s/Hz/(3 dB)) and high-SNR decibel offset $\mathcal{L}_\infty$ characterize the affine approximation of C *vs.* η as $\mathsf{C} \to \infty$:

$$\eta^{\text{dB}} = \frac{10\log_{10} 2}{\mathcal{S}_\infty}\,\mathsf{C} + \mathcal{L}_\infty\, 10\log_{10} 2 + O(\log_{10}\mathsf{C}), \quad \mathsf{C} \to \infty.$$

Equivalently, the following relation holds in terms of C *vs.* γ:

$$C(\gamma) = \mathcal{S}_\infty\Big[\,\log_2(\beta\gamma) - \mathcal{L}_\infty\,\Big] + o(1), \quad \gamma \to \infty,$$

from which $\mathcal{S}_\infty$ and $\mathcal{L}_\infty$ are derived as:

$$\mathcal{S}_\infty = \lim_{\eta\uparrow\infty} \frac{\mathsf{C}(\eta)}{\ln\eta}\ln 2 = \lim_{\eta\uparrow\infty} \eta\,\mathsf{C}'(\eta)\ln 2 = \lim_{\gamma\uparrow\infty} \gamma C'(\gamma)\ln 2 = 1 - \mathsf{F}(0) \tag{1.27}$$

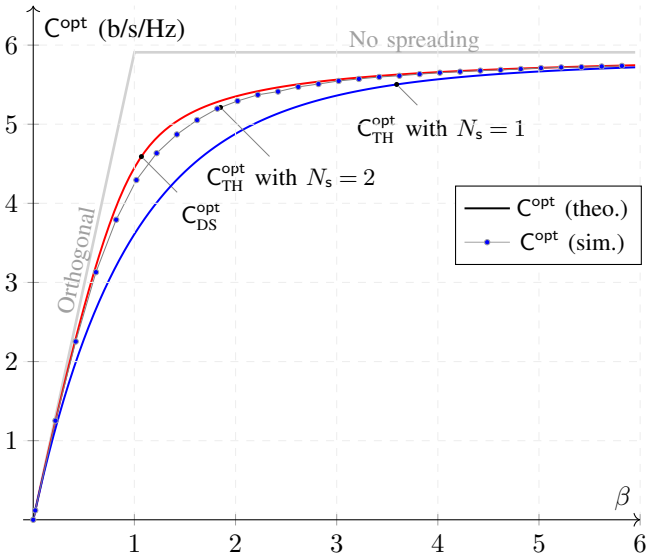


Figure 1.5: Spectral efficiency C^{opt} (b/s/Hz) as a function of β, for $\mathcal{E}_b/\mathcal{N}_0 = 10$ dB. DS and TH with $N_s = 1$ are shown in solid lines, indicating that curves derive from closed form expressions, while values for TH with $N_s = 2$ are shown in dots, indicating that they derive from simulations. Orthogonal access is also reported for reference (gray solid line).

$$\mathcal{L}_\infty = \log_2 \beta + \lim_{\gamma \uparrow \infty} \left[\log_2 \gamma - \frac{C(\gamma)}{\mathcal{S}_\infty} \right], \tag{1.28}$$

where the last equality in eq. (1.27) is obtained by differentiating $\log_2(1+\lambda\gamma)$ with respect to γ and applying the dominated convergence theorem to pass the limit under the integral sign. As a remark, in Appendix 1.C it is shown that $\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^\mathsf{T}}(0) = 1 - \operatorname{rank} \boldsymbol{S}/N$, hence one should be able to prove that $\operatorname{rank} \boldsymbol{S}/N \to \mathcal{S}_\infty$ in some mode of convergence. We can verify this result in the $N_s = 1$ case, where $\mathsf{F}(0) = e^{-\beta}$ and $\frac{1}{N} \operatorname{rank} \boldsymbol{S} \xrightarrow{p} 1 - e^{-\beta}$, as shown by the above Theorem 3.

For TH-CDMA with $N_s = 1$, it can be shown by direct computations that the four above parameters are given by:

$$\eta_{\min} = \ln 2, \tag{1.29}$$

$$\mathcal{S}_0 = 2\frac{\beta}{1+\beta}, \tag{1.30}$$

$$\mathcal{S}_\infty = 1 - e^{-\beta}, \tag{1.31}$$

$$\mathcal{L}_\infty = \log_2 \beta - \frac{1}{1-e^{-\beta}} \sum_{k>1} \frac{\beta^k e^{-\beta}}{k!} \log_2 k. \tag{1.32}$$

For the generic case $N_s > 1$, one can show that asymptotics in the wideband regime are the same as above (see eq. (1.29) and (1.30)). More precisely, we show in Appendix 1.D that $\mathbb{E}[\lambda] = \beta$ surely for any matrix ensemble where columns of $\boldsymbol{S}$

are normalized, and that $\frac{1}{N}\sum_{i=1}^{N}\lambda_i^2 \xrightarrow{p} \beta(\beta+1)$. Therefore, from eqs. (1.25) and (1.26), one has $\eta_{\min} = \ln 2$ surely for any N_s and $\mathcal{S}_0$ in probability as in eq. (1.30), respectively.

Comparison with DS-CDMA results ([6]), eqs. (1.29)-(1.32) show that TH-CDMA has same wideband asymptotic parameters, $\eta_{\min}$ and $\mathcal{S}_0$, as DS-CDMA, while different high-SNR parameters, $\mathcal{S}_\infty$ and $\mathcal{L}_\infty$. In particular, in the high-SNR regime, DS-CDMA achieves $\mathcal{S}_\infty = \min\{1,\beta\}$ [6] while TH-CDMA achieves $\mathcal{S}_\infty = 1 - e^{-\beta}$, that is, TH-CDMA incurs in a slope penalty given by $e^{-\beta}$. At very high loads, $\beta \gg 1$, this penalty becomes negligible, and TH-CDMA high-SNR slope tends to that of DS-CDMA.

Figure 3.6 shows spectral efficiency C (b/s/Hz) of TH-CDMA with $N_s = 1$ (blue solid line) *vs.* DS-CDMA (red solid line) as a function of $\mathcal{E}_b/\mathcal{N}_0$ (dB) with load $\beta = 1/2$; simulation for TH with $N_s = 1$ are also represented on figure (blue triangles) to highlight agreement with theoretical values. Orthogonal multiple access is also reported for comparison (gray solid line) and represents an upper bound on the sum-rate of a multiuser communication scheme. In the wideband regime, where $\mathsf{C} \to 0$, both TH-CDMA and DS-CDMA achieve $\eta_{\min} = \ln 2$ and same wideband slope $\mathcal{S}_0$. At the high-SNR regime, where $\mathcal{E}_b/\mathcal{N}_0 \to \infty$, DS achieves larger high-SNR slope than TH. A simulated case of $N_s = 2$ was also considered in order to understand the effect on C of increased N_s for TH-CDMA (see blue dots on figure). While for any finite N_s the spectral efficiency gap between DS-CDMA and TH-CDMA grows as $\mathcal{E}_b/\mathcal{N}_0$ increases, figure shows that for common values of $\mathcal{E}_b/\mathcal{N}_0$, *e.g.* $\mathcal{E}_b/\mathcal{N}_0 < 20$ dB, $N_s = 2$ pulses only are sufficient to reduce the gap to very small values. Figure 1.5 shows spectral efficiency C^{opt} (b/s/Hz) for TH with $N_s = 1$ (blue solid line) and $N_s = 2$ (dotted line), and for DS (red solid line), for $\mathcal{E}_b/\mathcal{N}_0 = 10$ dB. It is shown that TH achieves lower spectral efficiency with respect to DS. However, the loss is negligible for both $\beta \ll 1$ and $\beta \gg 1$. The gap between the two spectral efficiencies can be almost closed with increased, yet finite, N_s. Simulations suggest that $N_s = 2$ is sufficient to significantly reduce the gap.

1.2.2 Single-User Matched Filter

The output of a bank of SUMF is given by eq. (1.2), that is, $\boldsymbol{y} = \boldsymbol{S}\boldsymbol{b} + \boldsymbol{n}$. Focusing on user 1, one has:

$$\begin{aligned} \mathsf{y}_1 &= \boldsymbol{s}_1^{\mathsf{T}}\boldsymbol{y} \\ &= \mathsf{b}_1 + \sum_{k=2}^{K}\rho_{1k}\mathsf{b}_k + n_1 \\ &= \mathsf{b}_1 + z_1, \end{aligned} \tag{1.33}$$

where $\rho_{1k} := \boldsymbol{s}_1^{\mathsf{T}}\boldsymbol{s}_k$. As shown in [6], spectral efficiency for binary or spherical DS-CDMA when each SUMF is followed by an independent single-user decoder knowing $\boldsymbol{S}$ is [6, 7]:

$$\mathcal{C}_{\text{DS}}^{\text{sumf}}(\beta,\gamma) = \beta\log_2\left(1+\frac{\gamma}{1+\beta\gamma}\right) \quad \text{(b/s/Hz)}. \tag{1.34}$$

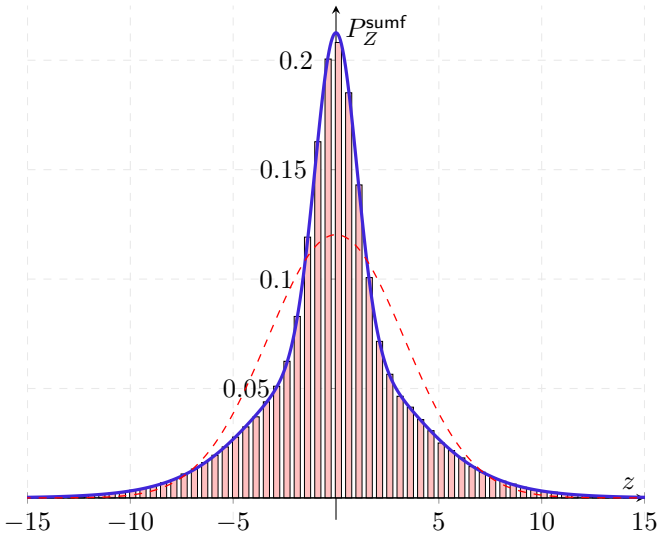


Figure 1.6: Probability density function of the real or imaginary part of the noise-plus-interference term of eq. (1.33) for TH sequences with $\gamma = 10$ dB, $\beta = 1$, and $N_s = 1$ (blue solid line), and comparison against a Gaussian PDF with same mean and variance (red dashed line). This example shows that, contrary to DS-CDMA, P_Z^{sumf} as given in eq. (1.42) may be far from Gaussian.

This result is general, and in particular it does not assume that the PDF of neither inputs nor interference term is Gaussian. Note, however, that, in this case, Gaussian inputs are optimal. In fact, for long spreading sequences, by virtue of the strong laws of large numbers, one has $\sum_{k=2}^{K} \rho_{1k}^2 \xrightarrow{\text{a.s.}} \beta$, and therefore the mutual information per user in bits per channel use is:

$$\begin{aligned} I(\mathsf{y}_1; \mathsf{b}_1 | \boldsymbol{S}) &= I\big(\mathsf{y}_1; \mathsf{b}_1 \big| \rho_{12}, \ldots, \rho_{1K}\big) \\ &= \mathbb{E}\left[\log_2\left(1 + \frac{\gamma}{1 + \gamma \sum_{k=2}^{K} \rho_{1k}^2}\right)\right] \xrightarrow{\text{a.s.}} \log_2\left(1 + \frac{\gamma}{1+\gamma\beta}\right). \end{aligned} \tag{1.35}$$

A similar result does hold for $I(\mathsf{y}_1; \mathsf{b}_1)$ as well.

When interference is not Gaussian, we may expect spectral efficiency to assume a very different form than above. This will prove to be the case for the mutual information of TH-CDMA assuming Gaussian inputs, when N_s remains finite while $N \to \infty$, as investigated below.

Theorem 5. *Suppose that $\boldsymbol{S} \in \mathbb{R}^{N \times \beta N}$ is a time-hopping matrix with generic $N_s < \infty$, and that the receiver is a bank of single-user matched filters followed by independent decoders, each knowing $\boldsymbol{S}$. Assuming Gaussian inputs, mutual information $I_{\text{TH}}^{\text{sumf}}$ (b/s/Hz) is given by:*

$$I_{\text{TH}}^{\text{sumf}}(\beta, \gamma, N_s) \coloneqq \beta I(\mathsf{y}_1; \mathsf{b}_1 | \boldsymbol{S}) = \beta \cdot \sum_{k \geq 0} \frac{(N_s^2 \beta)^k}{k!} e^{-N_s^2 \beta} \log_2\left(1 + \frac{\gamma}{1 + \frac{k}{N_s^2}\gamma}\right), \tag{1.36}$$

Proof. See Appendix 1.E.

In particular, for the $N_s = 1$ case, mutual information is:

$$I_{\mathrm{TH}}^{\mathsf{sumf}}(\beta,\gamma) = \beta \cdot \sum_{k\geq 0} \frac{\beta^k}{k!} e^{-\beta} \log_2\left(1+\frac{\gamma}{1+k\gamma}\right), \tag{1.37}$$

that can be compared to, and interpreted as, eq. (1.21).

Note that eq. (1.37) provides the mutual information of TH-CDMA with $N_s = 1$, and not the spectral efficiency, since Gaussian inputs, rather than optimal ones, are assumed. Hence, we know that spectral efficiency will be larger than or equal to $I_{\mathrm{TH}}^{\mathsf{sumf}}(\beta,\gamma)$. This mutual information expression is, however, sufficient to catch a significant difference between DS-CDMA and TH-CDMA. By comparing eqs. (1.34) and (1.37), we can claim that, while spectral efficiency for DS is bounded at high γ, being:

$$\lim_{\gamma\to\infty} C_{\mathrm{DS}}^{\mathsf{sumf}}(\beta,\gamma) = \beta \log_2\left(1+\frac{1}{\beta}\right), \tag{1.38}$$

spectral efficiency for TH is unbounded. We can indeed derive the below stronger result:

Corollary 1. *Under the hypotheses of Theorem 5, the high-SNR slope of the mutual information* (1.36) *of TH is:*

$$\mathcal{S}_{\infty,\mathrm{TH}}^{\mathsf{sumf}} = \beta e^{-N_s^2\beta}. \tag{1.39}$$

The maximum slope as a function of β is achieved at $\beta = 1/N_s^2$, for which $\mathcal{S}_{\infty,\mathrm{TH}}^{\mathsf{sumf}} = 1/(eN_s^2)$. Since $N_s \geq 1$, the global maximum is $1/e$, and the optimum load is $\beta = 1$. This behavior directly provides an insight from a design standpoint: at high-SNR, the number of chips such that an increase in $\mathcal{E}_b/\mathcal{N}_0$ yields a maximum increase in terms of mutual information is equal to the number of users. As a comparison, for optimum decoding, $\mathcal{S}_\infty$ increases monotonically with β, and its supremum is $\sup \mathcal{S}_\infty = 1$.

Differently from DS, when decoders have no knowledge about cross-correlations of signature sequences of other users, mutual information assumes a very different form, as derived in the following theorem.

Theorem 6. *Suppose that $\boldsymbol{S} \in \mathbb{R}^{N\times\beta N}$ is a time-hopping matrix with generic $N_s < \infty$, and that the receiver is a bank of single-user matched filters followed by independent decoders, each knowing the signature sequence of the user to decode only. Assuming Gaussian inputs, mutual information $I_{\mathrm{TH}^\star}^{\mathsf{sumf}}(\beta,\gamma,N_s)$ (bits/s/Hz) is given by:*

$$I_{\mathrm{TH}^\star}^{\mathsf{sumf}}(\beta,\gamma,N_s) := \beta I(\mathsf{y}_1;\mathsf{b}_1) = \beta \cdot \left[h(P_Y^{\mathsf{sumf}}) - h(P_Z^{\mathsf{sumf}})\right], \tag{1.40}$$

where P_Y^{sumf} and P_Z^{sumf} are the two following Poisson-weighted linear combinations of Gaussian distributions:

$$P_Y^{\mathsf{sumf}} = \sum_{k\geq 0} \frac{(\beta N_{\mathsf{s}}^2)^k}{k!} e^{-\beta N_{\mathsf{s}}^2} \mathcal{CN}\left(0, 1+\gamma+k\gamma/N_{\mathsf{s}}^2\right), \tag{1.41}$$

$$P_Z^{\mathsf{sumf}} = \sum_{k\geq 0} \frac{(\beta N_{\mathsf{s}}^2)^k}{k!} e^{-\beta N_{\mathsf{s}}^2} \mathcal{CN}\left(0, 1+k\gamma/N_{\mathsf{s}}^2\right). \tag{1.42}$$

Proof. See Appendix 1.F.

Despite decoders' lack of knowledge on $\boldsymbol{S}$, a same high-SNR slope as that achieved when decoders have knowledge of $\boldsymbol{S}$ is verified in the $N_{\mathsf{s}}=1$ case, as derived in the following corollary.

Corollary 2. Under the hypotheses of Theorem 6, and when $N_{\mathsf{s}}=1$, the high-SNR slope of the mutual information (1.40) of TH is:

$$\mathcal{S}_{\infty,\mathrm{TH}^\star}^{\mathsf{sumf}} = \beta e^{-\beta}. \tag{1.43}$$

Based on eq. (1.42), it can be checked that the kurtosis of the interference-plus-noise z_1, that we denote Z since it is independent of the user, is:

$$\kappa_Z := \frac{\mathbb{E}\left[|Z|^4\right]}{\mathbb{E}\left[|Z|^2\right]^2} = 2 + \frac{2}{N_{\mathsf{s}}^2} \cdot \frac{\beta\gamma^2}{(1+\beta\gamma)^2}, \tag{1.44}$$

that is always greater than 2, hence showing non-Gaussianity of Z for any β, γ and N_{s}. This non-Gaussian nature is represented on Fig. 1.6, that shows the interference-plus-noise PDF P_Z^{sumf} (solid blue line on figure), as given by eq. (1.42) when $\beta=1$, $\gamma=10$ dB and $N_{\mathsf{s}}=1$, *vs.* a Gaussian distribution with same mean and variance (red dashed line on figure). As shown by figure, P_Z^{sumf}, that is a linear combination, or "mixture," of Gaussian distributions with Poisson weights, cannot be reasonably approximated with a single Gaussian distribution; hence, the Standard Gaussian Approximation does not hold in general. This is the reason for the spectral efficiency gap between DS and TH.

The wideband regime is not affected by decoders' knowledge about crosscorrelations between signature sequences, as summarized by the below corollary, which proof is omitted for brevity.

Corollary 3. The wideband regime parameters derived from either eq. (1.37) or eq. (1.40) are $\eta_{\mathsf{min}} = \ln 2$ and:

$$\mathcal{S}_{0,\mathrm{TH}}^{\mathsf{sumf}} = \frac{2\beta}{1+2\beta}. \tag{1.45}$$

Differently from above, where N_{s} is finite and does not depend on N, we now investigate the case $N_{\mathsf{s}}=\alpha N$ with $\alpha\in(0,1)$, while $N\to\infty$. We show, using an approach similar to that developed in [6], that spectral efficiency of a TH channel with $N_{\mathsf{s}}=\alpha N$, $\alpha\in(0,1)$, is equal to that of a DS system, irrespectively of $\alpha\in(0,1)$.

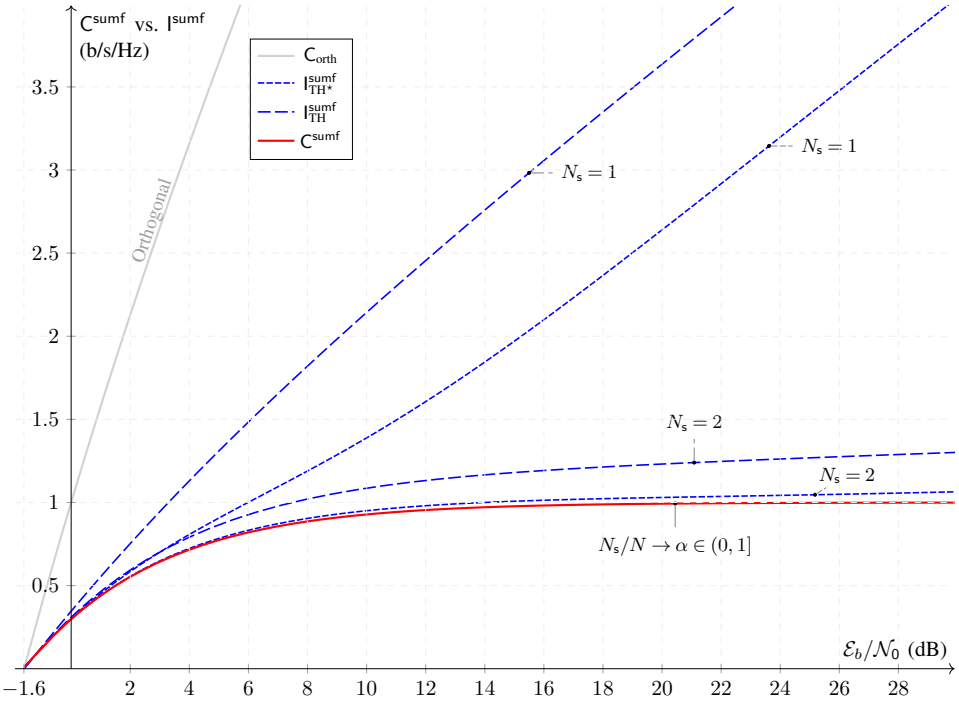


Figure 1.7: Spectral efficiency $\mathsf{C}^{\mathsf{sumf}}$ *vs.* $\mathsf{I}^{\mathsf{sumf}}$ (b/s/Hz) as a function of $\mathcal{E}_b/\mathcal{N}_0$ (dB) with load $\beta = 1$. Closed form expressions of spectral efficiency *vs.* mutual information are plotted in solid *vs.* dashed lines. Simulated mutual information is represented by dotted lines. On figure: SUMF, TH-CDMA, $N_s = 1$, $N_s = 2$ and $N_s = 5$, blue dashed lines; SUMF, TH-CDMA*, $N_s = 1$, blue dashed line; DS-CDMA, red solid line; TH-CDMA with $N_s = \alpha N$ when $N \to \infty$, blue solid line, coinciding with red solid line; TH-CDMA* with $N_s = 2$ and $N_s = 5$, blue dotted lines. Note on figure the crossover of SUMF, TH-CDMA, $N_s = 2$ and SUMF, TH-CDMA*, $N_s = 2$, that shows an example of mutual information becoming greater than conditional mutual information. For reference, orthogonal multiple-access in gray line.

Theorem 7. Suppose that $\boldsymbol{S} \in \mathbb{R}^{N \times \beta N}$ *is a time-hopping matrix with* $N_s = \alpha N$, $\alpha \in (0, 1)$, *and that the receiver is a bank of single-user matched filters followed by independent decoders knowing cross-correlations and input distributions of interfering users. Capacity* $C^{\mathsf{sumf}}(\rho_{12}, \ldots, \rho_{1K}, P_{b_2}, \ldots, P_{b_K})$ *of the single-user channel of eq. (1.33), expressed in bits per user per channel use, converges almost surely to:*

$$C^{\mathsf{sumf}}(\rho_{12}, \ldots, \rho_{1K}, P_{b_2}, \ldots, P_{b_K}) \xrightarrow{\text{a.s.}} \log_2\left(1 + \frac{\gamma}{1 + \beta\gamma}\right), \tag{1.46}$$

irrespective of α.

Proof. See Appendix 1.H.

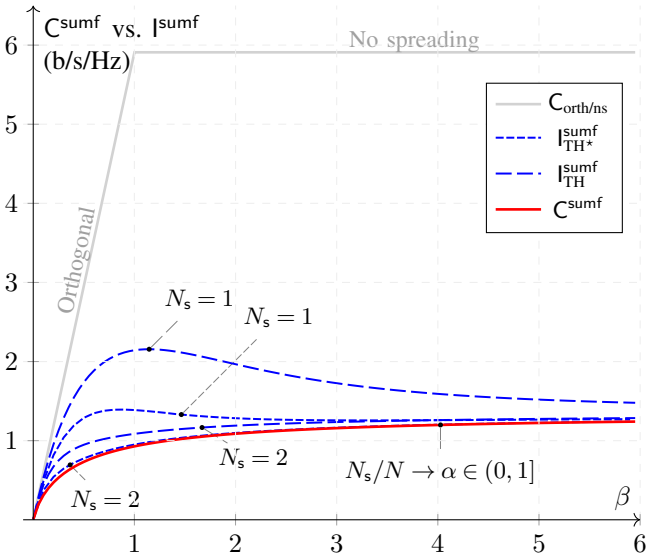


Figure 1.8: Spectral efficiency C^{sumf} *vs.* mutual information I^{sumf} as a function of β for fixed $\mathcal{E}_b/\mathcal{N}_0 = 10$ dB. DS and TH with $N_s/N \to \alpha \in (0, 1)$ are shown in red solid line, and represent the worst performance on figure. Dashed lines correponds to either TH knowing $\boldsymbol{S}$ (large dashing) or TH where decoders know the spreading sequence of the user to decode only (small dashing). Orthogonal access is reported for reference (gray solid line).

Based on eq. (1.46), spectral efficiency coincides with that of DS sequences, as given by eq. (1.34). As a matter of fact, Theorem 7 is a generalization of the result of Verdú and Shamai [6], for TH matrices where the fraction of nonzero entries is α, to which it reduces for $\alpha = 1$.

Figure 1.7 shows spectral efficiency C^{sumf} *vs.* mutual information I^{sumf} (b/s/Hz) as a function of $\mathcal{E}_b/\mathcal{N}_0$ (dB) for DS-CDMA (eq. (1.34), red solid line on figure), TH-CDMA knowning cross-correlations between users (eq. (1.37), blue large-dashed lines) and TH-CDMA without knowing cross-correlations between users, indicated as TH-CDMA* (eq. (1.40), blue small-dashed line), with unit load $\beta = 1$. Spectral efficiency of TH-CDMA when $N_s = \alpha N$, $\alpha \in (0, 1)$, as $N \to \infty$, is equal to that of DS (c.f. eq. (1.46), red solid line). As previously, the orthogonal case (gray solid line) is shown for reference. Note that spectral efficiency is bounded in DS-CDMA and in TH-CDMA when $N_s = \alpha N$, $\alpha \in (0, 1)$, as $N \to \infty$; the value of the limit is 1 on figure (c.f. eq. (1.38)). On the contrary, mutual information is not bounded for both TH-CDMA and TH-CDMA*; in particular, when $N_s = 1$, both TH-CDMA and TH-CDMA* grow with similar slope as $\mathcal{E}_b/\mathcal{N}_0$ increases. Mutual information of systems using multiple pulses per symbol is shown for TH-CDMA* with $N_s = 2$ (small-dashed line) and for TH-CDMA with $N_s = 2$ (eq. (1.36), large-dashed line). These $N_s \neq 1$ cases show that mutual information decreases with respect to the one pulse per symbol case. Figure 1.8 shows spectral efficiency C^{sumf} (b/s/Hz) as a function of β for fixed $\mathcal{E}_b/\mathcal{N}_0 = 10$ dB. Similarly as on fig. 1.7, TH with $N_s = 1$ outperforms other schemes, with and without complete knowledge of $\boldsymbol{S}$. As $\beta \to \infty$,

interference becomes increasingly Gaussian, and mutual information of TH reduces to that of DS, tending to the same limit $1/\ln 2$.

1.2.3 Decorrelator and MMSE

The output of a bank of decorrelators, following the discrete channel $\boldsymbol{y} = \boldsymbol{S}\boldsymbol{b} + \boldsymbol{n}$ (c.f. eq. (1.2)), is given by:

$$\boldsymbol{r} = \boldsymbol{S}^{+}\boldsymbol{y} = \boldsymbol{S}^{+}\boldsymbol{S}\boldsymbol{b} + \boldsymbol{S}^{+}\boldsymbol{n}, \tag{1.47}$$

where $\boldsymbol{S}^{+}$ denotes the Moore-Penrose pseudoinverse; if $\boldsymbol{R} = \boldsymbol{S}^{\mathsf{T}}\boldsymbol{S}$ is invertible, then $\boldsymbol{S}^{+} = (\boldsymbol{S}^{\mathsf{T}}\boldsymbol{S})^{-1}\boldsymbol{S}^{\mathsf{T}}$, otherwise $\boldsymbol{S}^{+}$, according to the Tikhonov regularization, exists and can be computed as the limit $(\boldsymbol{S}^{\mathsf{T}}\boldsymbol{S} + \alpha \boldsymbol{I})^{-1}\boldsymbol{S}^{\mathsf{T}}$ as $\alpha \to 0^{+}$.

In DS-CDMA, for any fixed $\beta \in (0, 1)$, $\boldsymbol{S}$ is almost surely full rank as $N \to \infty$, and therefore, $\boldsymbol{R}$ is almost surely invertible, in which case eq. (1.47) becomes:

$$\boldsymbol{r} = \boldsymbol{b} + \boldsymbol{z}, \tag{1.48}$$

where $\boldsymbol{z} \sim \mathcal{CN}\left(\boldsymbol{0}, \boldsymbol{R}^{-1}\mathcal{N}_0\right)$. Assuming independent single-user decoders, spectral efficiency is [6]:

$$C_{\mathrm{DS}}^{\mathrm{deco}}(\beta, \gamma) = \beta \log_2(1 + \gamma(1 - \beta)). \tag{1.49}$$

The output of a bank of MMSE filters observing $\boldsymbol{y} = \boldsymbol{S}\boldsymbol{b} + \boldsymbol{n}$ (c.f. eq. (1.2)) is:

$$\begin{aligned} \boldsymbol{r} &= \boldsymbol{W}^{\mathsf{T}}\boldsymbol{y} = \boldsymbol{W}^{\mathsf{T}}\boldsymbol{S}\boldsymbol{b} + \boldsymbol{W}^{\mathsf{T}}\boldsymbol{n} \\ &= \boldsymbol{G}\boldsymbol{b} + \boldsymbol{\nu}, \end{aligned} \tag{1.50}$$

where $\boldsymbol{W}^{\mathsf{T}}$ is defined as follows:

$$\boldsymbol{W}^{\mathsf{T}} := \left(\boldsymbol{S}^{\mathsf{T}}\boldsymbol{S} + \frac{1}{\gamma}\boldsymbol{I}\right)^{-1}\boldsymbol{S}^{\mathsf{T}} = \boldsymbol{S}^{\mathsf{T}}\left(\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}} + \frac{1}{\gamma}\boldsymbol{I}\right)^{-1}. \tag{1.51}$$

Note that, as well known, MMSE and decorrelator coincide as $\gamma \to \infty$.

In DS-CDMA, for any fixed $\beta > 0$, it was shown in [6] that:

$$C_{\mathrm{DS}}^{\mathrm{mmse}}(\beta, \gamma) = \beta \log_2\left(1 + \gamma - \frac{1}{4}\mathcal{F}(\beta, \gamma)\right), \tag{1.52}$$

where:

$$\mathcal{F}(\beta, \gamma) = \left[\sqrt{1 + \gamma\ell^{+}} - \sqrt{1 + \gamma\ell^{-}}\right]^2,$$

being $\ell^{\pm} = (1 \pm \sqrt{\beta})^2$ as in eq. (1.19).

We can treat both decorrelator and MMSE as special cases of the linear operator:

$$\boldsymbol{W}^{\mathsf{T}}(\alpha) := (\boldsymbol{S}^{\mathsf{T}}\boldsymbol{S} + \alpha\boldsymbol{I})^{-1}\boldsymbol{S}^{\mathsf{T}} = \boldsymbol{S}^{\mathsf{T}}(\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}} + \alpha\boldsymbol{I})^{-1},$$

for $\alpha \to 0^+$ and $\alpha = 1/\gamma$, respectively. Similarly as eq. (1.50), one has:

$$\boldsymbol{r}(\alpha) = \boldsymbol{W}^{\mathsf{T}}(\alpha)\boldsymbol{y} = \boldsymbol{G}(\alpha)\boldsymbol{b} + \boldsymbol{\nu}(\alpha), \tag{1.53}$$

where dependence on α is now made explicit, and the output for user 1 is:

$$r_1 = G_{11}\mathsf{b}_1 + \sum_{k=2}^{K} G_{1k}\mathsf{b}_k + \nu_1. \tag{1.54}$$

For $N_{\mathrm{s}} = 1$, a closed form expression for the generic element of $\boldsymbol{G}(\alpha)$ is derived in Appendix 1.I, and reads as:

$$G_{ij}(\alpha) = \rho_{ij} \cdot \frac{1}{\alpha + \upsilon_i}, \tag{1.55}$$

where υ_i is:

$$\upsilon_i = \sum_{k=1}^{K} \mathbb{1}\{\rho_{ik} \neq 0\} = \sum_{k=1}^{K} |\rho_{ik}| = \sum_{k=1}^{K} |\rho_{ki}|. \tag{1.56}$$

Denote with $\mathcal{J}_j$ the following set: $\mathcal{J}_j := \{k \in [1\!:\!K]\colon \rho_{jk} \neq 0\}$. Hence, $\upsilon_j = |\mathcal{J}_j|$ is the cardinality of $\mathcal{J}_j$. Denote with $\mathcal{J}'_j := \mathcal{J}_j \backslash \{j\}$. Since $j \in \mathcal{J}_j$, one has $\upsilon'_j := |\mathcal{J}'_j| = \upsilon_j - 1$. We can rewrite eq. (1.119) as follows:

$$r_1 = \frac{1}{\alpha + \upsilon_1}\mathsf{b}_1 + \frac{1}{\alpha + \upsilon_1} \sum_{k \in \mathcal{J}'_1} \rho_{1k}\mathsf{b}_k + \nu_1. \tag{1.57}$$

Note that $\rho_{1k}\mathsf{b}_k$ for $\rho_{1k} \neq 0$ is distributed as b_k, and ν_1 given υ_1 is complex Gaussian with zero mean and conditional variance:

$$\mathrm{Var}\big[\,\nu_1 \big|\, \upsilon_1\,\big] = \mathcal{N}_0 \cdot \frac{1}{(\alpha + \upsilon_1)^2}.$$

Known $\boldsymbol{S}$, r_1 and r_1 given b_1 are both complex Gaussian, hence mutual information expressed in bits per user per channel use is:

$$\begin{aligned} I(\mathsf{b}_1; r_1 | \boldsymbol{S}) = I(\mathsf{b}_1; r_1 | \upsilon'_1) &= \mathbb{E}\left[\log_2\left(1 + \frac{\mathcal{E}/(\alpha + \upsilon'_1)^2}{(\upsilon'_1\mathcal{E} + \mathcal{N}_0)/(\alpha + \upsilon'_1)^2}\right)\right] \\ &= \mathbb{E}\left[\log_2\left(1 + \frac{\gamma}{\upsilon'_1\gamma + 1}\right)\right], \end{aligned}$$

Since $\upsilon'_1 \sim \mathcal{B}(K-1, 1/N)$, in the LSL one has $\upsilon'_1 \xrightarrow{d} \mathcal{P}\,\beta$. Therefore, we proved the following:

Theorem 8. *Suppose that $\boldsymbol{S} \in \mathbb{R}^{N \times \beta N}$ is a time-hopping matrix with $N_{\mathrm{s}} = 1$, and that the receiver is a bank of either decorrelators ($\alpha = 0$) or MMSE filters ($\alpha = 1/\gamma$) followed by independent decoders, each knowing $\boldsymbol{S}$. Assuming Gaussian*

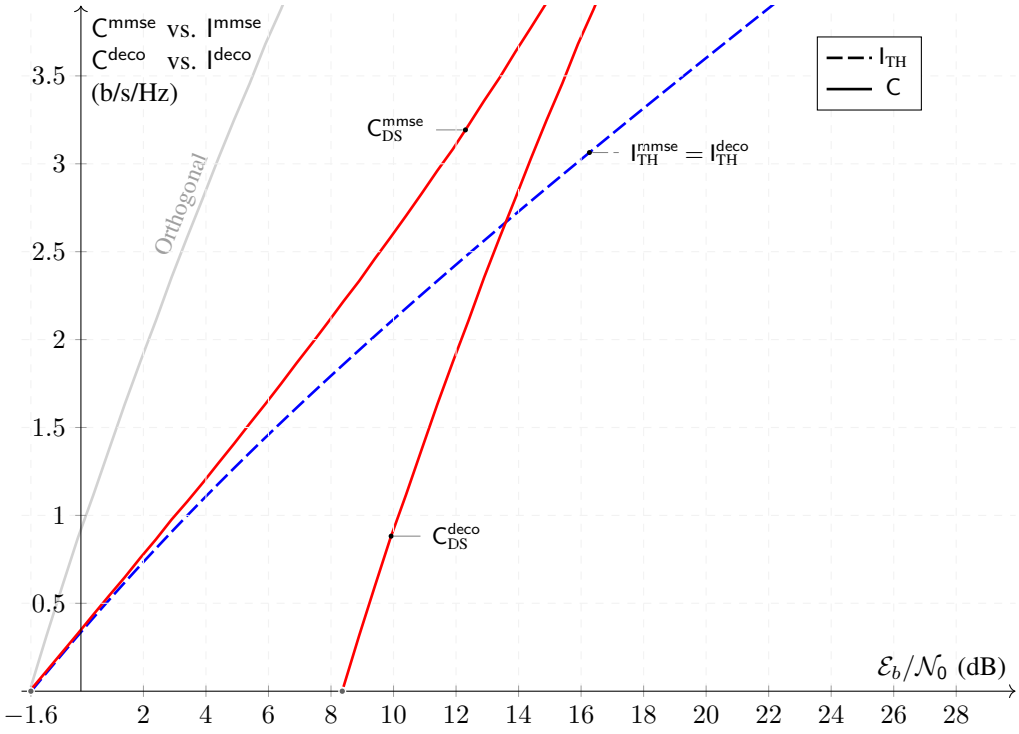


Figure 1.9: Spectral efficiency C^{mmse} *vs.* mutual information I^{mmse} and C^{deco} *vs.* I^{deco} (b/s/Hz) as a function of $\mathcal{E}_b/\mathcal{N}_0$ (dB) with load $\beta = 0.9$. Mutual information of TH-CDMA with $N_s = 1$ (blue dashed line) *vs.* spectral efficiency of DS-CDMA (red solid lines), for decorrelator and MMSE receivers, is shown. It is also shown orthogonal access (gray line) for reference.

inputs, mutual information I_{TH}^{α} *(b/s/Hz) is given by:*

$$I_{TH}^{\alpha}(\beta,\gamma) := \beta I(b_1; r_1|\boldsymbol{S}) = \beta \sum_{k\geq 0} \frac{\beta^k e^{-\beta}}{k!} \log_2\left(1 + \frac{\gamma}{k\gamma + 1}\right). \tag{1.58}$$

Since eq. (1.58) does not depend on α and is equal to eq. (1.37) for SUMF, one explicitly has $I_{TH}^{\alpha} = I_{TH}^{sumf} = I_{TH}^{mmse} = I_{TH}^{deco}$, thus we will write the above quantities interchangeably. With minor modifications of the above argument, it is possible to show that a similar result does hold for any linear receiver $\boldsymbol{W}^{\mathsf{T}}(\alpha)$, $\alpha > 0$, under the assumption $N_s = 1$. Therefore, results for SUMF can be extended verbatim to both decorrelator and MMSE receivers, when $N_s = 1$. This result suggests a striking difference with respect to DS, where spectral efficiency depends on the adopted linear receiver: In TH with $N_s = 1$, SUMF, decorrelator and MMSE all result in the same mutual information.

In order to compare DS and TH for decorrelator and MMSE, we separate the analysis for systems with $\beta < 1$ and $\beta > 1$, to which we refer as underloaded and overloaded, respectively.

Underloaded system ($\beta < 1$).

Decorrelation in DS allows to achieve the maximum high-SNR slope, $\mathcal{S}_{\infty,\mathrm{DS}}^{\mathrm{deco}} = \beta$, that is equal to that of orthogonal multiple access. On the contrary, TH does not fully exploit the capabilities of CDMA in the high-SNR regime, since $\mathcal{S}_{\infty,\mathrm{TH}}^{\mathrm{deco}} = \mathcal{S}_{\infty,\mathrm{TH}}^{\mathrm{sumf}} = \beta e^{-\beta} \leq \beta$. This behavior follows directly from cross-correlation properties of signature sequences of DS *vs.* TH: In DS, the almost sure linear independence of signature sequences, that holds for any $\beta \in (0,1)$, makes $\boldsymbol{R} = \boldsymbol{S}^{\mathsf{T}}\boldsymbol{S}$ almost sure invertible, and thus interference can be mostly removed, which is not the case of TH (c.f. Fig. 1.3 and Theorem 3). However, the optimal high-SNR slope in DS comes at the expense of a minimum $\mathcal{E}_b/\mathcal{N}_0$ equal to $(\ln 2)/(1-\beta)$, that can be much larger than that achieved by TH, namely $\ln 2$; in particular, as $\beta \to 1^-$, the minimum energy-per-bit for DS with decorrelator grows without bound. Therefore, decorrelation with DS should to be considered in a very low load, high-SNR regime only: in this region, it outperforms TH. It can be shown, by comparing eqs. (1.52) and (1.49), that in DS spectral efficiency of MMSE is always larger than that of decorrelator. In particular, it achieves a minimum energy-per-bit equal to $\ln 2$, which is optimal, and also an optimal high-SNR slope.

Overloaded system ($\beta > 1$).

Spectral efficiency of TH and DS with MMSE is similar in the low-SNR regime, with same minimum energy-per-bit and wideband slope. At high-SNR, mutual information of TH is unbounded, while spectral efficiency of DS is bounded, as in the SUMF case. In particular, while the high-SNR slope of TH is equal to $\mathcal{S}_{\infty,\mathrm{TH}}^{\mathrm{sumf}}(\beta) = \beta e^{-\beta}$ for any β, the high-SNR slope of DS with MMSE is:

$$\mathcal{S}_{\infty,\mathrm{DS}}^{\mathrm{mmse}}(\beta) = \beta\, \mathbb{1}\{\beta \in [0,1)\} + \frac{1}{2}\mathbb{1}\{\beta = 1\} + 0 \cdot \mathbb{1}\{\beta > 1\},$$

which implies that, as $\mathcal{E}_b/\mathcal{N}_0 \to \infty$, $\mathsf{C}_{\mathrm{DS}}^{\mathrm{mmse}}$ is infinite for $\beta \leq 1$, while it is finite for $\beta > 1$, and equal to (c.f. eq. (1.52)) [6]:

$$\lim_{\gamma \to \infty} C_{\mathrm{DS}}^{\mathrm{mmse}}(\beta, \gamma) = \beta \log_2 \frac{\beta}{\beta - 1}. \tag{1.59}$$

By comparing this result with eq. (1.38), that refers to SUMF, one also notes that the two limits are different, although as $\beta \to \infty$ both tend to $1/\ln 2$.

Figure 1.9 shows spectral efficiency $\mathsf{C}^{\mathrm{mmse}}$ and $\mathsf{C}^{\mathrm{deco}}$ *vs.* mutual information $\mathsf{I}^{\mathrm{mmse}}$ and $\mathsf{I}^{\mathrm{deco}}$ (b/s/Hz) as a function of $\mathcal{E}_b/\mathcal{N}_0$ (dB) for DS (red solid lines) and TH (blue dashed line), when $\beta = 0.9$. Orthogonal access is also shown for reference (gray solid line). The choice of $\beta = 0.9$ represents a scenario with high interference where eq. (1.49) is still valid, and DS with decorrelation still comparable. MMSE and decorrelator receivers achieve a same mutual information for TH: in the low-SNR regime, $\mathsf{I}_{\mathrm{TH}}^{\mathrm{mmse}} = \mathsf{I}_{\mathrm{TH}}^{\mathrm{deco}}$ and $\mathsf{C}_{\mathrm{DS}}^{\mathrm{mmse}}$ have similar behavior, that departs as $\mathcal{E}_b/\mathcal{N}_0$ increases. Decorrelator with DS achieves the maximum high-SNR slope, which is equal to that of the orthogonal access: note that the two curves on figure are, in fact, translated. This is not the case for TH, for $\boldsymbol{S}$ is not full rank with high probability, and the high-SNR slope is indeed lower. It is shown on figure that

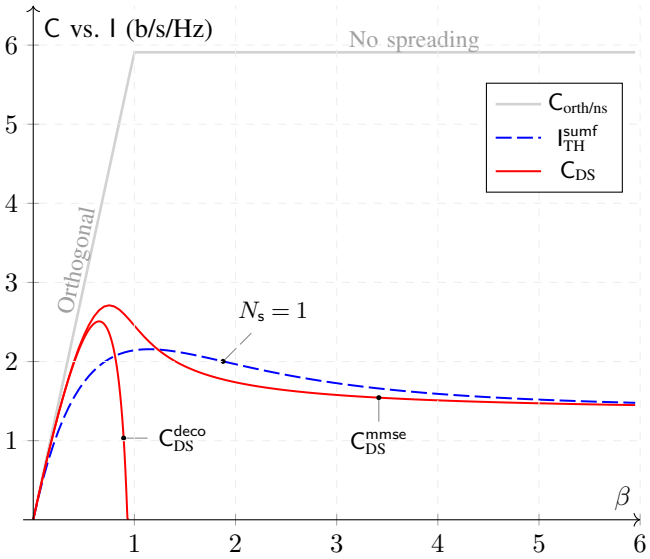


Figure 1.10: Spectral efficiency C_{DS}^{mmse} and C_{DS}^{deco} *vs.* mutual information $I_{TH}^{mmse/deco}$ (b/s/Hz) as a function of β for DS-CDMA (red solid lines) and TH-CDMA (blue dashed line), when $\mathcal{E}_b/\mathcal{N}_0 = 10$ dB. Orthogonal access is reported for reference (gray solid line).

DS with MMSE outperforms linear receivers with TH: this is due to the particular choice of β. Figure 1.10 shows spectral efficiency C_{DS}^{mmse} and C_{DS}^{deco} (red solid lines) *vs.* mutual information $I_{TH}^{sumf} = I_{TH}^{deco} = I_{TH}^{mmse}$ (blue dashed line) as a function of β, when $\mathcal{E}_b/\mathcal{N}_0 = 10$ dB. This figure shows that MMSE with DS is outperformed by TH for large β: in particular, there exists a minimum value of β, say $\bar{\beta}$, in general depending on $\mathcal{E}_b/\mathcal{N}_0$, beyond which the mutual information of TH is higher than the spectral efficiency of DS, although both tending to a same limit as $\beta \to \infty$, that is, $1/\ln 2$. While it is difficult to study $\bar{\beta}$ as a function of $\mathcal{E}_b/\mathcal{N}_0$, the above discussion on the high-SNR slope of DS suggest that $\beta = 1$ marks a transition in DS behavior as $\mathcal{E}_b/\mathcal{N}_0 \to \infty$. Figure 1.11 shows C_{DS}^{mmse} (red solid lines) and $I_{TH}^{sumf} = I_{TH}^{deco} = I_{TH}^{mmse}$ (blue dashed line) as a function of β, for different values of $\eta^{dB} = 10\log_{10}(\mathcal{E}_b/\mathcal{N}_0)$. Figure shows that, as η^{dB} increases, spectral efficiency of DS grows linearly for $\beta \ll 1$, and at about $\beta = 1$ quickly drops towards the limit value given by eq. (1.59), while spectral efficiency of TH remains smooth for any load in the neighborhood of $\beta = 1$ and increases monotonically with η^{dB}.

1.2.4 Synopsis of the TH-CDMA case

Figure 1.12 shows spectral efficiency C or mutual information I (b/s/Hz) *vs.* $\mathcal{E}_b/\mathcal{N}_0$ (dB) for the two extreme cases of optimum decoding and SUMF receivers, when $\beta = 1$. Curves derived from closed form expressions of spectral efficiency are shown for optimum decoding when $N_s = 1$ (top blue solid line), and SUMF when $N_s = \alpha N$ as $N \to \infty$ and $\alpha \in (0, 1)$ (bottom blue solid line). Curves derived

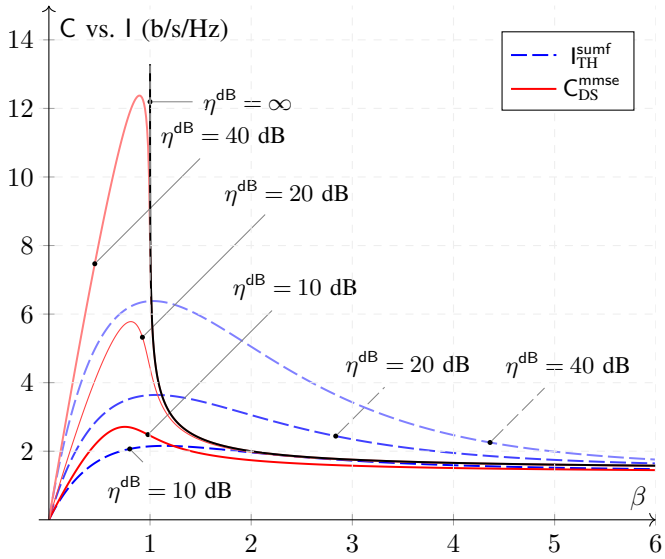


Figure 1.11: Spectral efficiency $\mathsf{C}_{\mathrm{DS}}^{\mathrm{mmse}}$ *vs.* mutual information $\mathsf{I}_{\mathrm{TH}}^{\mathrm{sumf}}$ (b/s/Hz) as a function of β for DS-CDMA (red solid lines) and TH-CDMA (blue dashed lines), for values of $\eta^{\mathrm{dB}} := \mathcal{E}_b/\mathcal{N}_0 \in \{10, 30, 50\}$ dB. Asymptotic value of $\mathsf{C}_{\mathrm{DS}}^{\mathrm{mmse}}$ for $\eta^{\mathrm{dB}} \to \infty$ is also shown for reference (thin solid black line).

from closed form expressions of mutual information assuming Gaussian inputs are shown for SUMF, TH-CDMA (blue dashed line, see label on figure) and SUMF, TH-CDMA* (blue dashed line, see label on figure), when $N_s = 1$ and $N_s = 2$. Finding closed form expressions of spectral efficiency of optimum decoding with generic $N_s > 1$ finite remains an open problem. Simulations provide, however, insights into the behavior of spectral efficiency for this particular case, as shown by $\mathsf{C}_{\mathrm{TH}}^{\mathrm{opt}}$ with $N_s = 2$ (blue dotted line). TH behavior is delimited by DS curves, with optimum decoding *vs.* SUMF (top and bottom red lines). Both upper and lower curves are approached by TH as N_s increases; in particular, we showed that the lower curve describes, in fact, TH when $N_s = \alpha N$, $\alpha \in (0, 1)$, as $N \to \infty$. In between these two extremes lie TH curves with optimum *vs.* linear receivers. In particular, for $N_s = 1$ (maximum energy concentration), mutual information of a receiver as simple as SUMF is not bounded, and also close to optimum decoding with $N_s = 1$. Furthermore, a lack of knowledge in cross-correlations of spreading codes provokes a drop of performance that is, however, not sufficient to degrade mutual information to DS spectral efficiency, with any finite N_s.

Figure 1.13 compares either spectral efficiency C or mutual information I (b/s/Hz), as a function of β, for DS and TH, when $\mathcal{E}_b/\mathcal{N}_0 = 10$ dB. Both DS and TH have similar behaviors when $\beta \ll 1$, for linear and optimum receivers. Irrespective of β, spectral efficiency of DS with optimum decoding is larger than that achieved by TH, the gap being almost closed when $N_s > 1$ finite. Conversely, among linear receivers and access schemes, it is shown that DS with SUMF has the lowest spectral efficiency, which is equal to that of TH when the number of

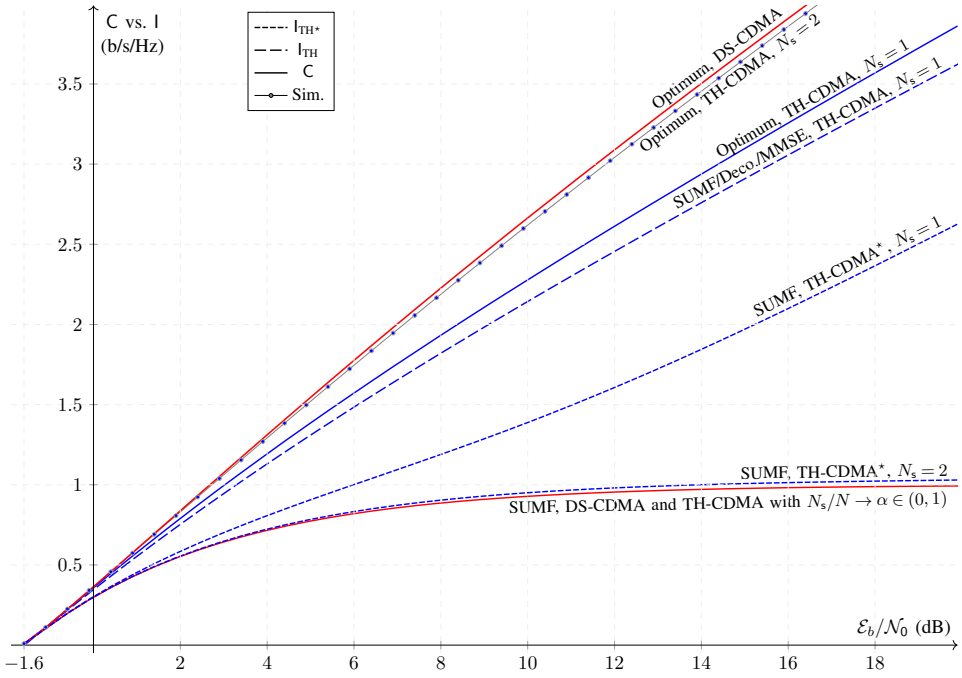


Figure 1.12: Spectral efficiency C *vs.* mutual information I (b/s/Hz) as a function of $\mathcal{E}_b/\mathcal{N}_0$ (dB) with load $\beta = 1$. Optimum *vs.* linear receivers are shown. Top curve shows C_{DS}^{opt} for optimum decoding in DS-CDMA (red solid line). Bottom curve shows C for SUMF, DS-CDMA (red solid line) coinciding with TH-CDMA when N_s goes to infinity proportionally to N, *i.e.*, $\lim_{N\to\infty} N_s/N = \alpha \in (0, 1)$ (red solid line). In between these two extremes: C_{TH}^{opt} curve for optimum decoding, TH-CDMA, $N_s = 2$, simulated values (dotted blue line); C_{TH}^{opt} curve for optimum TH-CDMA, $N_s = 1$ (blue solid line); I_{TH} curve for linear receivers, TH-CDMA, $N_s = 1$ (blue large-dashed line); $I_{TH^\star}^{sumf}$ curve for SUMF, TH-CDMA*, $N_s = 1$ and $N_s = 2$ (blue small-dashed line).

pulses is asymptotically a nonzero fraction of the number of chips. The largest spectral efficiency in DS is obtained with MMSE, which is greater than the mutual information of TH when load is lower than a threshold $\bar{\beta}(\mathcal{E}_b/\mathcal{N}_0)$, depending in general on $\mathcal{E}_b/\mathcal{N}_0$. At higher load, mutual information of TH is larger than spectral efficiency of DS. This analysis is intrinsically conservative, since spectral efficiency of TH will be, in general, larger than or equal to the mutual information obtained assuming Gaussian inputs. Therefore, one should expect that the gap in spectral efficiency between DS and TH with linear receivers is smaller and larger than that showed on figure when $\beta < \bar{\beta}$ and $\beta > \bar{\beta}$, respectively.

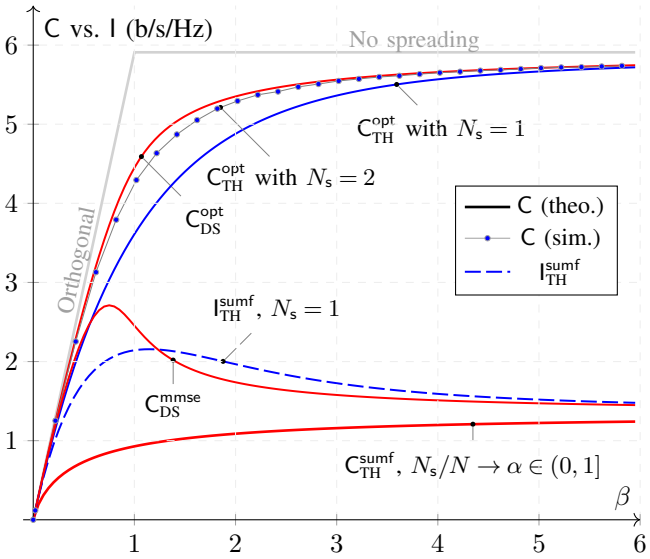


Figure 1.13: Spectral efficiency C or mutual information I (b/s/Hz) as a function of β for DS and TH, when $\mathcal{E}_b/\mathcal{N}_0 = 10$ dB, with optimum and SUMF receivers. Orthogonal access and DS with MMSE receiver are reported for reference.

1.3 Conclusions

Verdú and Shamai showed in [6] that optimum decoding provides a substantial gain over linear decoding in DS-CDMA, with random spreading. In particular, a bank of single-user matched filters followed by independent decoders is bounded in spectral efficiency at high-SNR, and linear multiuser detectors are needed in order to recover a nonzero spectral efficiency high-SNR slope. This behavior is partly due to the "even" use of degrees of freedom—coinciding in our setting with chips—that is intrisic of DS-CDMA [1].

The object of this paper was to analyze TH-CDMA with random hopping, and compare its behavior against DS-CDMA; we interpreted time-hopping in the general framework developed in [6, 7]. The present analysis allowed comparison of TH *vs.* DS with same energy per symbol and same bandwidth constraints, and, therefore, showed the effect of the energy "concentration," that is typical of TH. The degree of "unevenness" in TH-CDMA is directly related to the number of pulses N_s representing each symbol. At one extreme, one has maximum "unevenness," where all energy is concentrated in one pulse ($N_s = 1$), while the other extreme corresponds to maximum "evenness," $N_s = N$, where TH coincides with DS. Particular emphasis has been put on the archetypal case of "unevennes," that is $N_s = 1$, and partial results showing the general behavior when $N_s > 1$ have been derived.

A first result of our analysis was to derive a closed form expression for spectral efficiency of TH-CDMA with optimum decoding when $N_s = 1$, showing that, in this case, DS-CDMA outperforms TH-CDMA, in particular in the high-SNR regime. Same wideband behavior, but lower high-SNR slope, was observed for TH-CDMA

vs. DS-CDMA, that is $\min\{1,\beta\} = \mathcal{S}_{\infty,\mathrm{DS}} > \mathcal{S}_{\infty,\mathrm{TH}} = 1 - e^{-\beta}$. A closed form expression for generic N_s remains an open problem; results based on simulations suggested, however, that the spectral efficiency loss at high-SNR may be considerably reduced while maintaining the number of pulses finite, and we provided evidences that the gap is reduced to a very small value with as low as two pulses per symbol ($N_\mathrm{s} = 2$). This result indicates that the spectral efficiency gap may be substantially reduced while only using a fraction N_s/N of degrees of freedom *per user*, that asymptotically vanishes as N grows.

A different behavior of TH-CDMA with respect to DS-CDMA was observed with linear receivers. Contrarily to DS, spectral efficiency of SUMF for TH with $N_\mathrm{s} = 1$ was unbounded. As suggested, this asymptotic behavior may be traced back to the non-Gaussian distribution of the interference-plus-noise variable observed by each independent single-user decoder, that, in turn, depends on cross-correlation properties of spreading sequences. The same high-SNR slope $\mathcal{S}^{\mathsf{sumf}}_{\infty,\mathrm{TH}} = \beta e^{-\beta}$ was achieved by TH irrespectively of the knowledge that each single-user decoder had about spreading sequences of all other users. It was interesting to note that the maximum slope for TH, providing a hint on greatest energy efficiency, was reached when the number of users K was equal to the number of chips N, *i.e.*, $\beta = 1$, leading to $\mathcal{S}^{\mathsf{sumf}}_{\infty,\mathrm{TH}} = 1/e \approx 0.367879$. On the contrary, for $N_\mathrm{s} = \alpha N$, $\alpha \in (0,1)$, same spectral efficiency as DS-CDMA ($\alpha = 1$) was obtained irrespectively of α for $N \to \infty$.

The bounded nature of spectral efficiency with a SUMF bank in DS-CDMA is overcome, as well known, by using more complex linear receivers, that also account for interference, such as MMSE and decorrelator. Conversely, we showed that, in TH-CDMA, mutual information assuming Gaussian inputs has the same expression, irrespective of the linear receiver used, due to the peculiar structure of TH spreading sequences. TH sequences are indeed "more" likely to be linearly dependent than DS ones, in agreement with the intuition based on the cardinality of binary DS *vs.* TH codes, that is 2^N *vs.* $2N$. This lack of independence led to the impossibility of removing interference, which is instead almost surely feasible for DS, *e.g.* with either decorrelator or MMSE receivers, as long as the load $\beta < 1$. Therefore, in a low load, high-SNR scenario, DS outperforms TH. The opposite is true when $\beta > 1$. In fact, while spectral efficiency in DS with MMSE rapidly drops, in particular with large $\mathcal{E}_b/\mathcal{N}_0$, as soon as β becomes larger than one, mutual information of TH decays softly when one keeps overloading the system, and tends to the same MMSE DS limit. The absence of a spectral efficiency "transition" in the neighborhood of the unit load, that is typical of DS, allows TH to outperform DS with any load larger than $\beta = 1$ for sufficiently high $\mathcal{E}_b/\mathcal{N}_0$.

Beyond the natural extension of the present work to channels with fading, where the effect of an "uneven" use of degrees of freedom typical of TH should be investigated, we do stress that, from the single-user perspective, TH is a particular instance of impulsive signal. As such, the present theoretical setting, if appropriately adapted to asynchronous links, may serve as a basis for refining the understanding of the limits of impulsive communications.

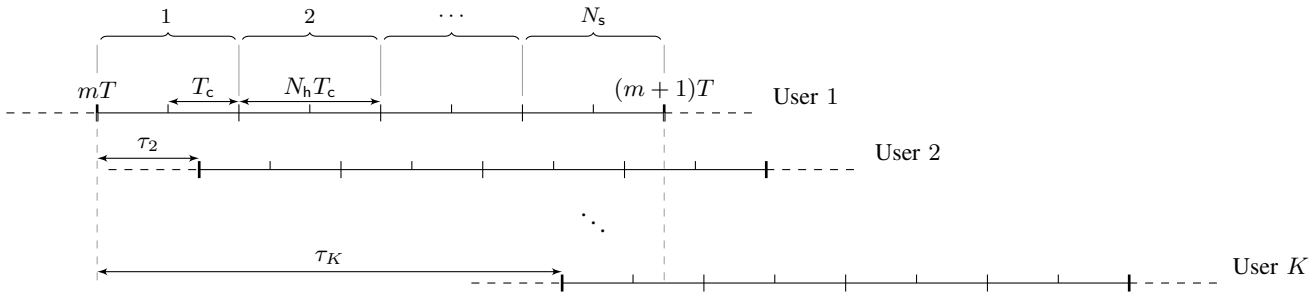


Figure 1.14: Signal model: in picture, $N_s = 4$, $N_h = 2$, therefore $N = N_s \cdot N_h = 8$. Delays are considered with respect to the first User, therefore $\tau_1 = 0$.

1.4 Asynchronous channel

1.4.1 Model

As for the asynchronous channel model, we adopt the model proposed in [20] for CDMA systems, where the received signal is:

$$y(t) = \sum_{k=1}^{K} \sum_{m=-M}^{M} A\, b_k[m]\, s_k(t - mT - \tau_k) + n(t), \tag{1.60}$$

in which we introduce the possibility for the signature waveform to change from one symbol to the next, and propose the following straightforward generalization:

$$y(t) = \sum_{k=1}^{K} \sum_{m=-M}^{M} A\, b_k[m]\, s_{k,m}(t - mT - \tau_k) + n(t), \tag{1.61}$$

where the codeword of length $n := 2M+1$ is considered for each user, the spreading waveform $s_{k,m}(t)$ of user k in the m^{th} epoch $mT + \tau_k$ may depend on m, that is:

$$s_{k,m}(t) = \sum_{i=0}^{N-1} s_{k,m}[i]\, \psi_i(t);$$

$b_k[m]$ is the m^{th} symbol of user k, and $\tau_k \in [0, T]$ indicates its delay. Users are ordered according to increasing τ, that is, $\tau_1 \leq \tau_2 \leq \ldots \leq \tau_K$, and τ_1 is set to $\tau_1 = 0$. As previously, $A = 1$.

Note that in this case, the anaylis focuses on set of M symbols per user *vs.* 1 of the synchronous case since this has proved to be necessary in order to derive a sufficient statistic for $\{b_k[m] : k = 1, \ldots, K; m = 1, \ldots, M\}$ [20].

A procedure similar to the one adopted in [30] allows to derive a discrete channel model equivalent to the continuous-time channel of eq. (1.61); the equivalence is defined in terms of channel capacity, and also in terms of capability of producing sufficient statistics leading to optimum inference of $\boldsymbol{b}$.

To this end, define the m^{th} symbol vector $\boldsymbol{b}[m] \in \mathbb{C}^K$ as:

$$\boldsymbol{b}[m] := (\mathsf{b}_1[m], \ldots, \mathsf{b}_K[m])^{\mathsf{T}}, \quad m = -M, \ldots, M,$$

and form vector $\boldsymbol{b}$ by stacking the above vectors:

$$\boldsymbol{b} = (\boldsymbol{b}[-M]^{\mathsf{T}}, \ldots, \boldsymbol{b}[M]^{\mathsf{T}})^{\mathsf{T}}. \tag{1.62}$$

This vector can be compactly written using the Kronecker product as follows:

$$\boldsymbol{b} = \sum_{m=-M}^{M} \tilde{\boldsymbol{e}}_m \otimes \boldsymbol{b}[m], \tag{1.63}$$

where $\tilde{\boldsymbol{e}}_m := \boldsymbol{e}_{m+M+1}$. Explicitly, the $(k+mK)$-th component of $\boldsymbol{b}$ is $\mathsf{b}_k[m]$. Note that, according to the above notation, the asynchronous channel with K users and codewords length n is equivalent to a synchronous channel with nK users and codeword length one.

Now consider projecting the continuous-time process $\mathsf{y}(t)$ onto the set of all spreading waveforms:

$$\begin{aligned}\{v_{mK+k}(t) := \mathsf{s}_{k,m}(t - mT - \tau_k) \colon k = 1, \ldots, K,\\ m = -M, \ldots, M\},\end{aligned} \tag{1.64}$$

that is, consider the inner product between $\mathsf{y}(t)$ and the generic $v_{k+mK}(t)$:

$$\gamma_{mK+k} = \langle \mathsf{y}(t), v_{k+mK}(t) \rangle .$$

This is the output of a filter that is matched to the signature sequence of user k at symbol m sampled *at the corresponding epoch*, that is, at time $mT + \tau_k$.

Stacking these projections as for $\boldsymbol{b}$, one has:

$$\begin{aligned}\boldsymbol{\gamma} &= \sum_{m=-M}^{M} \tilde{\boldsymbol{e}}_m \otimes \boldsymbol{\gamma}[m] \\ &= \sum_{m=-M}^{M} \tilde{\boldsymbol{e}}_m \otimes \sum_{k=1}^{K} \boldsymbol{e}_k \gamma_{mK+k},\end{aligned} \tag{1.65}$$

that is, a vector $\boldsymbol{\gamma} \in \mathbb{C}^{nK}$ providing a sufficient statistic for the optimum detection of $\boldsymbol{b}$.

Based on eqs. (1.62) and (1.65), the equivalent discrete channel model can therefore be expressed by:

$$\boldsymbol{\gamma} = \boldsymbol{R}\boldsymbol{b} + \boldsymbol{\nu}, \tag{1.66}$$

where components of matrix $\boldsymbol{R} \in \mathbb{R}^{nK \times nK}$ are given by:

$$(\boldsymbol{R})_{ij} := \langle v_j(t), v_i(t) \rangle = \langle v_i(t), v_j(t) \rangle = \int_{\mathbb{R}} v_j(t) v_i(t) dt,$$

$$
\begin{aligned}
(\boldsymbol{R})_{mK+k,m'K+k'} &= \langle \mathsf{s}_{k',m'}(t-m'T-\tau_{k'}),\, \mathsf{s}_{k,m}(t-mT-\tau_k)\rangle \\
&= \sum_{j=0}^{N-1}\sum_{j'=0}^{N-1} \mathsf{s}_{k,m}[j]\,\mathsf{s}_{k',m'}[j'] \\
&\qquad\qquad \langle \psi(t-m'T-\tau_{k'}-j'T_{\mathrm{c}}),\, \psi(t-mT-\tau_k-jT_{\mathrm{c}})\rangle \\
&\equiv \boldsymbol{s}_{k,m}^{\mathsf{T}}\boldsymbol{\Phi}[m,m';k,k']\boldsymbol{s}_{k',m'} \qquad (1.67)
\end{aligned}
$$

$$
\text{with:}\quad \big(\boldsymbol{\Phi}[m,m';k,k']\big)_{j,j'} := \langle\, \psi(t-m'T-\tau_{k'}-j'T_{\mathrm{c}}),\, \psi(t-mT-\tau_k-jT_{\mathrm{c}})\,\rangle.
$$

---

showing that $\boldsymbol{R}$ is symmetric.

Since n symbols are transmitted for each user and given eqs. (1.62), (1.65) and (1.66), $\boldsymbol{R}$ has the following block-matrix structure:

$$
\boldsymbol{R} \;=\; \sum_{m=-M}^{M}\sum_{m'=-M}^{M} \tilde{\boldsymbol{e}}_m\tilde{\boldsymbol{e}}_{m'}^{\mathsf{T}} \otimes \boldsymbol{R}[m,m'], \qquad (1.68)
$$

where each block $\boldsymbol{R}[m,m']$ is $K\times K$ with components:

$$
\begin{aligned}
(\boldsymbol{R}[m,m'])_{k,k'} &= (\boldsymbol{R})_{mK+k,m'K+k'} \\
&= \langle v_{m'K+k'}(t)\,,\, v_{mK+k}(t)\rangle\,.
\end{aligned}
$$

Since $\boldsymbol{R}$ is symmetric, one also has $\boldsymbol{R}[m',m]=\boldsymbol{R}[m,m']^{\mathsf{T}}$. Given (1.68), the following relation holds for $m=-M,\ldots,M$:

$$
\boldsymbol{\gamma}[m] \;=\; \sum_{m'=-M}^{M} \boldsymbol{R}[m,m']\boldsymbol{b}[m'] + \boldsymbol{\nu}[m]. \qquad (1.69)
$$

Compared to CDMA systems where signatures do not change across symbol periods, note that, as previously indicated, the proposed model releases this constraint, as reflected in the general block structure of $\boldsymbol{R}$. Note that the proposed model reduces to its traditional form [20] when signatures do not vary across symbol periods, for which one has:

$$
\boldsymbol{R} \;=\; \sum_{m=-M}^{M}\sum_{m'=-M}^{M} \tilde{\boldsymbol{e}}_m\tilde{\boldsymbol{e}}_{m'}^{\mathsf{T}} \otimes \boldsymbol{R}[m-m'],
$$

having defined $\boldsymbol{R}[m-m'] := \boldsymbol{R}[m,m']$. In this case, $\boldsymbol{R}$ is also block-Toeplitz.

Note that if the channel were synchronous, one would have $\boldsymbol{R}=\boldsymbol{S}^{\mathsf{T}}\boldsymbol{S}$, with $\boldsymbol{S}$ is a block-diagonal matrix with the generic m^{th} diagonal block has the following

block-diagonal structure:

$$\boldsymbol{S}[m,m] \coloneqq \begin{bmatrix} \boldsymbol{s}_{1,m} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{s}_{2,m} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{s}_{K,m} \end{bmatrix}. \tag{1.70}$$

Therefore, $\boldsymbol{S}$ is:

$$\begin{aligned} \boldsymbol{S} &= \sum_{m=-M}^{M} \tilde{\boldsymbol{e}}_m \tilde{\boldsymbol{e}}_m^{\mathsf{T}} \otimes \boldsymbol{S}[m,m] \\ &= \sum_{m=-M}^{M} \tilde{\boldsymbol{e}}_m \tilde{\boldsymbol{e}}_m^{\mathsf{T}} \otimes \sum_{k=1}^{K} \boldsymbol{e}_k \boldsymbol{e}_k^{\mathsf{T}} \otimes \boldsymbol{s}_{k,m}. \end{aligned}$$

We may express $\boldsymbol{R}$ in terms of *chip-matched filter* output sampled at epoch $mT_{\mathrm{s}} + \tau_k$ for user k and symbol m. To this end, note that the generic element of $\boldsymbol{R}$ can be written in terms of matrices:

$$\{\boldsymbol{\Phi}[m,m';k,k'] \in \mathbb{R}^{N\times N} : m,m' = -M,\dots,M, \\ k,k' = 1,\dots,K\},$$

as defined in eq. (1.67). Note that $\boldsymbol{\Phi}[m,m';k,k']$ is Toeplitz and only depends on difference $m-m'$ rather than on m and m'. Therefore, denoting by $\boldsymbol{\Phi}[m,m']$ the block matrix obtained by grouping the $K \times K$ blocks $\{\boldsymbol{\Phi}[m',m;k',k] : k,k' = 1,\dots,K\}$, that is:

$$\boldsymbol{\Phi}[m,m'] = \sum_{k=1}^{K}\sum_{k'=1}^{K} \boldsymbol{e}_k \boldsymbol{e}_{k'}^{\mathsf{T}} \otimes \boldsymbol{\Phi}[m,m';k,k'],$$

it results that $\boldsymbol{\Phi}[m,m']$ is block-Toeplitz. Finally, consider the matrix obtained by grouping the $n \times n$ blocks $\boldsymbol{\Phi}[m,m']$ as follows:

$$\boldsymbol{\Phi} = \sum_{m=-M}^{M}\sum_{m'=-M}^{M} \tilde{\boldsymbol{e}}_m \tilde{\boldsymbol{e}}_{m'}^{\mathsf{T}} \otimes \boldsymbol{\Phi}[m,m'].$$

By construction, $\boldsymbol{\Phi}$ is block-Toeplitz with respect to blocks $\{\boldsymbol{\Phi}[m,m'] : m,m' = -M,\dots,M\}$ and also with respect to smaller blocks $\{\boldsymbol{\Phi}[m,m';k,k'] : k,k' = 1,\dots,K\}$.The generic (i,j) element of $\boldsymbol{\Phi}$ can be expressed by:

$$(\boldsymbol{\Phi})_{ij} = \langle u_j(t), u_i(t) \rangle = \int_{\mathbb{R}} u_j(t) u_i(t) dt,$$

having defined:

$$u_{mNK+kK+j}(t) \coloneqq \psi(t - mT - \tau_k - jT_c),$$

which also shows that $\boldsymbol{\Phi}$ is symmetric.

We can express $\boldsymbol{R}$ in terms of spreading sequences and matrix $\boldsymbol{\Phi}$ as follows:

$$\boldsymbol{R} = \boldsymbol{S}^{\mathsf{T}} \boldsymbol{\Phi} \boldsymbol{S}.$$

This model duly reduces to the synchronous model when delays are null, as can be derived in a straightforward way.

Note that a linear time-invariant, frequency-selective channel with impulse response $c(t)$ might be taken into account by substituting $\psi(t)$ with $\psi * c(t)$. In this case, $\boldsymbol{\Phi}$ would account for both asynchronicity and frequency-selectivity.

Note that the assumption of zero-excess bandwidth pulses was, as initially indicated, introduced for mere simplification, and any orthonormal family of pulses, as in particular time-limited waveforms, fit the model; UWB impulse-radio communications are appropriately modeled.

1.4.2 Impulsiveness

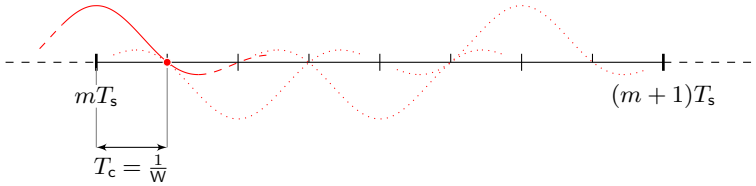
A non-impulsive system occupying frequency band $[-\mathsf{W}/2, \mathsf{W}/2]$ (Hertz) and using a channel for T (seconds) may transmit at most $\mathsf{W}T$ orthogonal waveforms by transmitting a waveform that occupies a bandwidth W every $1/\mathsf{W}$ seconds for T seconds. Symbol period is therefore $1/\mathsf{W}$, and effective duration is $1/\mathsf{W}$ as well, that is in non-impulsive systems these two quantities coincide.

Conversely, in impulsive systems, that is, systems using pulses with bandwidth $\mathsf{W}' \geq \mathsf{W}$, *while keeping fixed the symbol period*, the channel may still used every $1/\mathsf{W}$ seconds but transmission occurs over bandwidth W'; effective duration is thus $1/\mathsf{W}'$. Impulsiveness can, therefore, be measured by ratio W'/W, that is, the ratio of transmission bandwidth over minimum Nyquist bandwidth. We call this ratio *impulsiveness index* $\imath$.

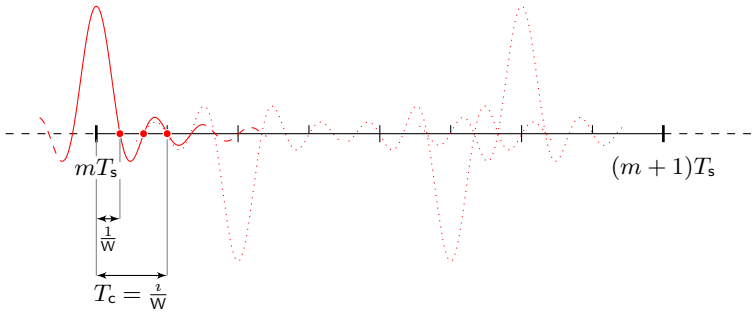
While spread-spectrum systems use a bandwidth larger than what is the minimum required to transmit at a given *symbol* rate, impulsive systems use a larger bandwidth than the minimum required to transmit at a given *chip* rate. Therefore, impulsiveness can be viewed as another way for spreading bandwidth. In non-impulsive spread-spectrum systems, bandwidth spreading is described by the number of chips in one symbol period, named *spreading factor* $N = T/T_{\mathsf{c}} \in \mathbb{N}$. Similarly, impulsiveness index $\imath$ specifies spreading as system bandwidth *vs.* chip rate; impulsive systems may use bandwidth $\mathsf{W} \gg 1/T_{\mathsf{c}}$, *i.e.*, $T_{\mathsf{eff}} \ll T_{\mathsf{c}}$, therefore transmitting pulses that are much shorter than the chip interval. The condition $\imath \in \mathbb{N}$ guarantees the Nyquist criterion to be satisfied.

In general, we propose to define impulsiveness resorting to the concept of *effective duration* akin to that of *effective bandwidth* as described in [31]:

$$\imath := \frac{T_{\mathsf{c}}}{T_{\mathsf{eff}}}\,, \tag{1.71}$$



(a) Non-impulsive



(b) Impulsive

Figure 1.15: Illustration of the impulsiveness index concept. A non-impulsive system, where $\imath = 1$, and therefore $T_c = 1/W$, is shown in Fig. (a), while an impulsive system, where $\imath = 3$, and therefore $T_c = 3/W$, is shown in Fig. (b).

where *effective duration* is defined as to satisfy:

$$T_{\text{eff}} \cdot \max_{t \in \mathbb{R}} |\psi(t)|^2 = \int_{-\infty}^{\infty} |\psi(t)|^2 \, dt = 1. \tag{1.72}$$

Note that for zero-excess bandwidth pulses occupying frequency band $[-W/2, W/2]$, as used throughout this chapter for the sake of simplicity, one has $T_{\text{eff}} = 1/W$, and therefore $\imath = T_c W$. However, the impulsiveness index is well defined also for the opposite case of strictly time-limited pulses; for example, for rectangular pulses with duration Δ (seconds), eq. (1.72) yields $T_{\text{eff}} = \Delta$, as intuitive.

Figure 1.15 shows an example of a spreading waveform for non-impulsive communications (Fig. 1.15a), where pulse bandwidth is directly related to chip rate, *vs.* impulsive communications (Fig. 1.15b), where pulse effective duration is shorter than the chip interval, and thus bandwidth is larger than chip rate, specifically $\imath$ times $1/T_c$.

In the literature of the last fifteen years, the concept of peaky signaling arose (*e.g.* [1, 5, 25, 32]). Although related, the concepts of impulsiveness and peakdeness are complementary, as we explain in the next section.

1.4.3 Impulsiveness *vs.* Peakedness

The notion of peakedness of a random variable is strictly related to the limitedness of its maximum absolute value or, in a relaxed sense, to the limitedness of its fourth moment and kurtosis [1, 5, 32]. In general, a distribution with large kurtosis is indicated as "peaky" [1].

As example of non-peaky distribution, consider a discrete binary r.v. X uniformly distributed, $P_X = (1/2)\delta_{-\mathsf{A}} + (1/2)\delta_{+\mathsf{A}}$; X has kurtosis $\kappa_X = 1$. On the contrary, as example of "peaky" distribution, consider $P_X = 1/(2p)\delta_{-\mathsf{A}} + (1-p)\delta_0 + 1/(2p)\delta_{+\mathsf{A}}$; X has kurtosis $\kappa_X = 1/p$, that is increasingly "peaky" as $p \to 0^+$. In general, a "peaky" distribution has the zero value, or nearly zero values compared to the standard deviation, with high probability with respect to the probability under the tails, that is, for values greater than the standard deviation; therefore, leptokurtic distributions are "peaky". A r.v. drawn from a "peaky" distribution is characterized by rare large and frequent mild deviations.

A communication system that transmits over bandwidth $[-\mathsf{W}/2, \mathsf{W}/2]$ for T seconds is characterized by $\mathsf{W}T$ degrees of freedom indicating that the channel may carry $\mathsf{W}T$ input symbols towards output; the capacity of the channel is the maximum number of bits, that can be encoded in these $\mathsf{W}T$ symbols, with a vanishing error probability. Under fixed power constraint, the distribution of inputs over these degrees of freedom determines peakedness. In particular, if the kurtosis of inputs is "relatively" high, then the input distribution is "peaky".

Increasingly "peaky" inputs lead to reaching the wideband capacity, *i.e.*capacity for spectral efficiency tending to zero, for transmissions over fading channels when the receiver has no a priori knowledge on fading coefficients [33–35]. Conversely, in the same context, Gallager and Médard [1, 2] show that non-peaky inputs lead to zero mutual information. In particular, these authors highlight a "bandwidth scaling" property for multipath fading channels, that can be described as follows: assuming an infinite number of path scattering model, mutual information of spread-spectrum signals tends to zero as bandwidth grows if energy and fourth moment of inputs *scale* with $1/\mathsf{W}$ and $1/\mathsf{W}^2$, respectively, that is, if degrees of freedom are used "evenly", in a typical non-peaky fashion. Finally, Verdú [25] defines the family of input distributions, called *flash*, achieving capacity at the wideband regime. This family encompasses inputs that are distributed as the mixture of two probability distributions, the first tending to a Dirac mass in the origin, and the second that vanishes as the signal-to-noise ratio per degree of freedom tends to zero.

Degrees of freedom can be exploited in time only, in frequency only, or in both time and frequency.

As example of a system that is non-peaky neither in time nor in frequency, consider DS-CDMA as in [1, 2]. As example of a system that is "peaky" in frequency, but not in time, consider M-FSK with symbol duration T and bandwidth W divided into M equal slices of width $1/T$, hence $M = \mathsf{W}T$. For transmitting a message m with $1 \leq m \leq M$, one frequency slice only is used. Inputs X present thus 1 out of M nonzero value occurences, and therefore are distributed as $P_X = 1/M\delta_1 + (1 - 1/M)\delta_0$, with kurtosis $\kappa_X = M$. As example of a system that is "peaky" in both

time and frequency, consider FSK with duty cycle, as in [3, 5]. As example of a system that is "peaky" in time, but not in frequency, consider M-PPM with symbol effective-duration $1/\mathrm{W}$, hence $M = \mathrm{W}T$. Inputs are distributed as above as in the case of M-FSK.

The last example is particularly significant to draw an analogy between peaky inputs in time, that are those that present "many" occurrences of the zero value, and impulsiveness.

Impulsive systems are intrinsically "peaky" since a subset only of the available degrees of freedom is used; this is particularly evident in the time domain since many occurrences of the zero value can be observed. Impulsiveness reshapes indeed the input distribution $\overline{X}$ of a non-impulsive signal with $\imath = 1$ by adding a Dirac mass at the origin as follows: $P_X = (1 - 1/\imath)\delta_0 + (1/\imath)P_{\overline{X}}$. Therefore, the intuition may be that impulsiveness provides a mean that is complementary to duty-cycle in order to achieve capacity at the wideband regime; as impulsiveness grows, inputs become increasingly "peaky", and may *flash* at the limit $\imath \to \infty$.

Moreover, TH-CDMA signals, as opposed to DS-CDMA, are "peaky"; also, for any fixed N_s, as N grows, these signals are increasingly "peaky", while tending to *flash* at the limit $N \to \infty$.

Impulsive CDMA signals in general have increased peakedness with respect to non-impulsive CDMA signals; therefore, for DS-CDMA, impulsiveness is a mean towards peakedness, while for TH-CDMA signals it is a mean towards increased peakedness.

1.4.4 First results

Spectral efficiency of the asynchronous channel can be derived, similarly to the synchronous case, from the equivalent discrete channel model $\boldsymbol{\gamma} = \boldsymbol{R}\boldsymbol{b} + \boldsymbol{\nu}$ (c.f. eq. (1.66)) as:

$$C^{N,n} = \frac{1}{nN} \log\det(\boldsymbol{\Sigma}_{\boldsymbol{\gamma}|\boldsymbol{R}} \boldsymbol{\Sigma}_{\boldsymbol{\nu}}^{-1}) \quad (\mathrm{b/s/Hz}).$$

Since the covariance of $\boldsymbol{\nu}$ and the conditional covariance of $\boldsymbol{\gamma}$ given $\boldsymbol{R}$ are, respectively:

$$\begin{aligned} \boldsymbol{\Sigma}_{\boldsymbol{\nu}} &= \mathcal{N}_0 \boldsymbol{R}, \\ \boldsymbol{\Sigma}_{\boldsymbol{\gamma}|\boldsymbol{R}} &= \mathbb{E}\left[(\boldsymbol{R}\boldsymbol{b} + \boldsymbol{\nu})(\boldsymbol{R}\boldsymbol{b} + \boldsymbol{\nu})^\dagger \,\middle|\, \boldsymbol{R} \right] \\ &= \mathcal{E}\boldsymbol{R}\boldsymbol{R}^\mathsf{T} + \mathcal{N}_0 \boldsymbol{R}, \end{aligned}$$

being $\mathbb{E}[\,\boldsymbol{b}\boldsymbol{b}^\dagger\,] = \mathcal{E}\boldsymbol{I}$, then:

$$C^{N,n}(\gamma) = \frac{1}{nN} \log\det(\boldsymbol{I} + \gamma \boldsymbol{R}) \quad (\mathrm{b/s/Hz}), \tag{1.73}$$

having denoted with $\gamma = \mathcal{E}/\mathcal{N}_0$ as previous.

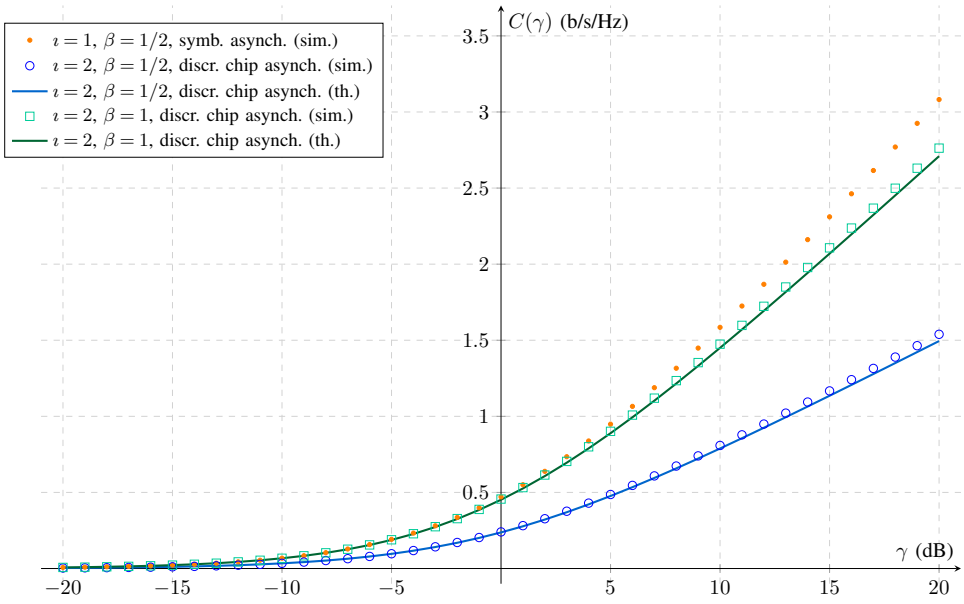


Figure 1.16: Spectral efficiency C as function of γ (dB) for the asynchronous channel with $N_s = 1$.

When an impulsive communication is employed, then the bandwidth used is $\imath$ times larger than the minimum required, therefore eq. (1.73) generalizes to:

$$C^{N,n}(\gamma) = \frac{1}{\imath} \cdot \frac{1}{nN} \log\det(\boldsymbol{I} + \gamma \boldsymbol{R}) \quad \text{(b/s/Hz)}. \tag{1.74}$$

Investigation of $C^{N,n}(\gamma)$ when either $n \to \infty$ or $N \to \infty$ for special families of random spreading signatures in an asynchronous setting is the major challange.

For non-varying signatures across symbol periods, a prior work of Verdú [30] reports the expression of capacity for asynchronous DS-CDMA channel with $n \to \infty$ and N finite, once signature sequences are fixed and nonrandom. For random DS-CDMA, Cottatellucci et al. [36] find the expression of the SINR achieved by an MMSE receiver using the REFORM method proposed by Girko [37] and exploit the celebrated relation between mutual information and mmse [38] in order to find capacity with the optimum receiver.

Cottatellucci's analysis of the DS-CDMA asynchronous channel in [36] proceeds in two steps. In the first step, it is assumed that the maximum delay between users is limited to a chip interval, that is $0 \leq \tau_K \leq T_c$: the so obtained channel is called *chip asynchronous, symbol quasi-synchronous* channel. In the second step, it is shown that results obtained in the previous step continue to be valid also when $0 \leq \tau_K \leq T_s$: this is called *symbol asynchonous* channel.

In what follows, we limit our attention to the chip-asynchronous, symbol quasi-synchronous TH-CDMA channel.

As in [36], the assumption of limitedness of τ_K to T_c implies that, in the large system limit, ISI becomes negligible.

Therefore, we can study a channel where one symbol only is transmitted.

We further assume discrete delays, *i.e.*, delays that are multiples of $1/\mathsf{W}$: $\tau_k = m/\mathsf{W}$ with $0 \le m \le \imath - 1$. This kind of "discrete-asynchronism" effectively turns the channel into a set of $\imath$ parallel channels, each populated by a number of users equal to the number of users with same delay: therefore, capacity of the discrete-asynchronous channel is the sum of capacities of these parallel channels.

Intuition behind this result lies on orthogonality of users with different delays. When two users have indeed different delays, their waveforms are orthogonal irrespective of the signature sequences: this follows from the property of bandlimited pulses to be orthogonal if delayed by multiples of $1/\mathsf{W}$. Therefore, dividing the set of users into equivalent classes where the equivalence is defined by having the same delay and calling k_ℓ the number of users with delay ℓ/W, the original channel is decomposed into at most $\imath$ orthogonal channels, since there are at most $\imath$ delays, each channel having symbols of N chips, using a bandwidth W, and being populated by k_ℓ users, hence having load equal to k_ℓ/N. This construction implies that $k_1 + \cdots + k_\imath = K$.

Now let prove these intuitions in light of the asynchronous channel model developed in Section 1.4.1 and the general spectral efficiency expression given by eq. (1.74), both specialized to the case $n = 1$.

Consider a composition of K into $\imath$ parts, that is, $K = k_1 + \cdots + k_\imath$, where $0 \le k_\ell \le K$. If two users j and j' have different delays, *i.e.*, $\tau_j \neq \tau_{j'}$, then $\forall \nu, \nu' = 0, \dots, N-1$:

$$\langle \psi(t - \tau_j - \nu T_\mathsf{c}), \psi(t - \tau_{j'} - \nu' T_\mathsf{c}) \rangle = 0.$$

Therefore, $\boldsymbol{\Phi}[0,0;i,j] = 0$ when $\tau_i \neq \tau_j$, and $\boldsymbol{\Phi}[0,0]$ is a $KN \times KN$ block-diagonal matrix with at most $\imath$ diagonal blocks of dimensions $k_1 N \times k_1 N$, $\dots$, $k_\imath N \times k_\imath N$.

As defined in eq. (1.66), $\boldsymbol{R}$ inherits this block-diagonal structure, with diagonal blocks of dimensions $k_1 \times k_1$, $\dots$, $k_\imath \times k_\imath$. Call these diagonal blocks $\boldsymbol{R}_{k_1}$, $\dots$, $\boldsymbol{R}_{k_\imath}$. From eq. (1.74), we have:

$$\begin{aligned} C^N(\gamma) &= \frac{1}{N\imath} \log\det(\boldsymbol{I}_K + \gamma \boldsymbol{R}) \\ &= \frac{1}{\imath} \sum_{\ell=1}^{\imath} \frac{1}{N} \log\det(\boldsymbol{I}_{k_\ell} + \gamma \boldsymbol{R}_{k_\ell}) \\ &= \frac{1}{\imath} \sum_{\ell=1}^{\imath} C^N_{k_\ell}(\gamma) \qquad \text{(b/s/Hz)}, \end{aligned}$$

where each $\boldsymbol{R}_{k_\ell}$ describes a multiple access channel with k_ℓ synchronous users and N chips having capacity $C^N_{k_\ell}(\gamma)$ (b/s/chip) as found in Section 1.2. Ensemble average

of C^N is therefore:

$$\overline{C}^N(\gamma) = \mathbb{E}\left[\frac{1}{\imath}\sum_{\ell=1}^{\imath} C_{k\ell}^N(\gamma)\right] = \frac{1}{\imath}\sum_{\ell=1}^{\imath}\mathbb{E}\big[\,C_{k\ell}^N(\gamma)\big]\,,$$

where the expectation is on k_ℓ, that is the number of users with delay equal to $(\ell-1)/\mathsf{W}$. Assuming uniform i.i.d. delays, $k_\ell \sim \mathrm{Binom}(K,\epsilon)$ with $\epsilon = 1/\imath$, irrespective of ℓ, hence:

$$\overline{C}^N(\gamma) = \sum_{k=1}^{K}\left[\binom{K}{k}\epsilon^k(1-\epsilon)^{K-k}\right]\overline{C}_k^N(\gamma) \qquad \text{(b/s/Hz)}.$$

Assuming as previous spectral efficiency to be self-averaging, the large-system limit yields spectral efficiency $C(\gamma)$:

$$C(\gamma) = \int_0^\infty \log(1+\lambda\gamma)\, d\overline{\mathsf{F}}(\lambda) \qquad \text{(b/s/Hz)}, \tag{1.75}$$

where:

$$\overline{\mathsf{F}}(\lambda) = \lim_{K\to\infty}\sum_{k=1}^{K}\binom{K}{k}\epsilon^k(1-\epsilon)^{K-k}\;\overline{\mathsf{F}}_k^N(\lambda), \tag{1.76}$$

having denoted with $\overline{\mathsf{F}}_k^N(\lambda)$ the e.s.d. of a synchronous CDMA channel with load k/N averaged over the spreading sequences, and with $\overline{\mathsf{F}}(\lambda)$ the large-system limit e.s.d. that is found by averaging $\overline{\mathsf{F}}_k^N(\lambda)$ with respect to k and let $K\to\infty$.

In general $\epsilon = 1/\imath$ is function of N. There are, therefore, different ways we can approach the large-system limit depending on the behavior of $\alpha := K\epsilon$ as $N\to\infty$.

Case 1) $\boxed{\imath = O(1)}$

This is case when $N\to\infty$, $K/N\to\delta\in(0,\infty)$ and $\epsilon = 1/\imath$ is constant, and thus also the bandwidth is constant. The load of the system is:

$$\beta \triangleq \frac{\#\,\text{users}}{\#\,\mathsf{DoF}} = \frac{K}{\imath N} = \epsilon\delta. \tag{1.77}$$

Case 2) $\boxed{\imath = \Theta(N)}$

When $\imath$ grows as N, the number of DoFs grows as N^2; therefore, also the number of users K has to grow as N^2 in order to keep spectral efficiency nonzero.

This is the case when $K/\imath = \alpha = \Theta(N)$. For the sake of simplicity, consider:

$$\frac{K}{\imath} = \alpha = \delta N.$$

Note that in this case, as the impulsiveness index, also the bandwidth tends to infinity in the large-system limit. The load of the system is:

$$\beta \triangleq \frac{\#\,\text{users}}{\#\,\text{DoF}} = \frac{K}{\imath N} = \frac{\alpha}{N} = \delta. \tag{1.78}$$

Therefore we can actually call β the constant factor $\beta = \alpha/N$.

TH-CDMA with $N_s = 1$.

We find closed form expression of (1.75) for TH-CDMA with $N_s = 1$ in the two cases specified above.

In general, spectral efficiency of TH-CDMA with $N_s = 1$ is given by:

$$\overline{C}^N(\gamma) = \mathbb{E}_k\big[\,\mathbb{E}_{\mathsf{x}|k}[\log(1+\mathsf{x}\gamma)\,]\big] \qquad \text{(b/s/Hz)},$$

where the inner expectation is over $\mathsf{x}|k \sim \text{Pois}(k/N)$ and the outer expectation is over $k \sim \text{Binom}(K, 1/\imath)$.

Case 1) Quite surprisingly, although in this case the Binomial distribution does not tend to a Poisson distribution, we can anyway regard the (unconditioned) distribution of x as Poissonian:

$$\begin{aligned}
\mathsf{f}_{\mathsf{x}}(x) &= \sum_{k=1}^{K} \mathsf{f}_k \mathsf{f}_{\mathsf{x}|k}(x) \\
&= \sum_{k=1}^{K} \binom{K}{k} \epsilon^k (1-\epsilon)^{K-k} \sum_{i\geq 0} \frac{(k/N)^i}{i!} e^{-\frac{k}{N}} \delta(x-i) \\
&= \sum_{i\geq 0} \left[\sum_{k=1}^{K} \binom{K}{k} \epsilon^k (1-\epsilon)^{K-k} \frac{(k/N)^i}{i!} e^{-\frac{k}{N}} \right] \delta(x-i) \\
&\asymp \sum_{i\geq 0} \frac{(\epsilon\delta)^i}{i!} e^{-\epsilon\delta} \delta(x-i),
\end{aligned}$$

hence, asymptotically, $\mathsf{x} \sim \text{Pois}(\epsilon\delta)$.

Case 2) In this case, the Binomial distribution tends to a Poisson distribution as follows: $k \sim \text{Pois}(\alpha)$, $\mathsf{x}|k \sim \text{Pois}(k/N)$ and $\mathsf{x} \sim \text{Pois}(\beta)$.

Therefore, in both cases, spectral efficiency is given by:

$$C(\gamma) = \mathbb{E}[\log(1+\mathsf{x}\gamma)] \qquad \text{(b/s/Hz)}, \tag{1.79}$$

as in a synchronous channel with load given respectively by eqs. (1.77) and (1.78).

1.5 Future work

In this section, several proposal for extending the analysis of previous sections are presented in order to account:

(E1) non-uniform power constraints over users;

(E2) fading channels;

(E3) multipath channels;

(E4) frequency-hopping;

(E5) asynchronism.

E1. We call *non-uniform* power constraint a constraint $\mathbb{E}\left[\,|\mathsf{b}_k|^2\,\right]$ that depends on k, *i.e.*,

$$\mathbb{E}\left[\,|\mathsf{b}_k|^2\,\right] \leq \mathcal{E}_k, \quad k = 1, \dots, K,$$

that, when symbols of different users are independent, leads to:

$$\boldsymbol{\Sigma_b} \triangleq \mathbb{E}[\boldsymbol{b}\boldsymbol{b}^\dagger] = \mathcal{E}, \tag{1.80}$$

where $\mathcal{E}$ is diagonal. Spectral efficiency is, therefore, given by:

$$\begin{aligned} C &= \lim_{N\to\infty} \sup_{P_{\boldsymbol{b}}} \frac{1}{N} I(\boldsymbol{y}; \boldsymbol{b} | \boldsymbol{S}) \\ &= \lim_{N\to\infty} \frac{1}{N} \log\det(\boldsymbol{I} + \frac{1}{\mathcal{N}_0} \boldsymbol{S}\mathcal{E}\boldsymbol{S}^{\mathsf{T}}). \end{aligned} \tag{1.81}$$

E2. The extension of model (1.2) to account channels with fading is straightforward. Denote with $\boldsymbol{A}$ the $K \times K$ diagonal matrix of the fading coefficients $\{\mathsf{A}_1, \dots, \mathsf{A}_K\}$. The flat-fading channel model reads as:

$$\boldsymbol{y} = \boldsymbol{S}\boldsymbol{A}\boldsymbol{b} + \boldsymbol{n} = \boldsymbol{H}\,\boldsymbol{b} + \boldsymbol{n}. \tag{1.82}$$

For the sake of simplicity, it is common to assume (e.g. [39]) that $\{\mathsf{A}_k \colon k = 1, \dots, K\}$ are i.i.d. with unit variance, $\mathbb{E}\left[\,|\mathsf{A}_k|^2\,\right] = 1$. This corresponds to channels having, on average, unit gain, that may be assumed w.l.o.g. (in the opposite case, the different gains may be included in a non-uniform power constraint, such as in ***E1***). Spectral efficiency, when receiver has knowledge of the fading coefficients, is:

$$C = \lim_{N\to\infty} \frac{1}{N} \log\det(\boldsymbol{I} + \gamma \boldsymbol{H}\boldsymbol{H}^\dagger), \tag{1.83}$$

and, therefore, the analysis reduces to that of finding the eigenvalue distribution of $\boldsymbol{H}\boldsymbol{H}^\dagger = \boldsymbol{S}\boldsymbol{A}\boldsymbol{A}^\dagger\boldsymbol{S}^{\mathsf{T}}$, or equivalently $\boldsymbol{H}^\dagger\boldsymbol{H} = \boldsymbol{A}^\dagger\boldsymbol{S}^{\mathsf{T}}\boldsymbol{S}\boldsymbol{A}$. In case a non-uniform power constraint is set, one has, similarly to eq. (1.85):

$$C = \lim_{N\to\infty} \frac{1}{N} \log\det(\boldsymbol{I} + \frac{1}{\mathcal{N}_0} \boldsymbol{S}\boldsymbol{A}\mathcal{E}\boldsymbol{A}^\dagger\boldsymbol{S}^{\mathsf{T}}). \tag{1.84}$$

Suppose to set the maximum power consumption of the network, *i.e.*, $\mathrm{Tr}\left[\,\boldsymbol{\Sigma_b}\,\right] \leq K\mathcal{E}$. If each transmitter has knowledge of $\boldsymbol{A}$, one can find the spectral efficiency as

follows:

$$C = \lim_{N\to\infty} \sup_{\substack{P_{\boldsymbol{b}} \\ \boldsymbol{\Sigma_b}\succeq 0:\ \mathsf{Tr}[\,\boldsymbol{\Sigma_b}\,]\le K\mathcal{E}}} \frac{1}{N} I(\boldsymbol{y};\boldsymbol{b}|\boldsymbol{S},\boldsymbol{A})$$
$$= \lim_{N\to\infty} \sup_{\boldsymbol{\Sigma_b}\succeq 0:\ \mathsf{Tr}[\,\boldsymbol{\Sigma_b}\,]\le K\mathcal{E}} \frac{1}{N} \log\det(\boldsymbol{I} + \frac{1}{\mathcal{N}_0}\boldsymbol{S}\boldsymbol{A}\boldsymbol{\Sigma_b}\boldsymbol{A}^\dagger\boldsymbol{S}^\mathsf{T}), \tag{1.85}$$

where the optimum $\boldsymbol{\Sigma_b}$ is expected to be non-diagonal, thus requiring cooperating users, and follow a water-filling argument. It would be interesting to find a closed form expression of eq. (1.85) and compare the resulting spectral efficiency *vs.* a suboptimum uniform power constraint, as given by eq. (1.84) with diagonal $\mathcal{E}$.

E3. The multipath channel between user k and the receiver may be modeled with a convolution matrix $\boldsymbol{C}_k$ of dimensions $(N+L)\times N$, where L is the length of the channel expressed in number of taps, corresponding to the delay spread $T_d \approx L/\mathsf{W}$ seconds. The channel model becomes:

$$\boldsymbol{y} = \sum_{k=1}^{K} \boldsymbol{C}_k \boldsymbol{s}_k \mathsf{b}_k + \boldsymbol{n} = \sum_{k=1}^{K} \boldsymbol{h}_k \mathsf{b}_k + \boldsymbol{n} = \boldsymbol{H}\boldsymbol{b} + \boldsymbol{n}, \tag{1.86}$$

where:

$$\boldsymbol{H} = [\boldsymbol{h}_1, \ldots, \boldsymbol{h}_K] = \sum_{k=1}^{K} \boldsymbol{e}_k^\mathsf{T} \otimes \boldsymbol{h}_k.$$

Spectral efficiency is given again by eqs. (1.84) and (1.85) by formally replacing $\boldsymbol{S}\boldsymbol{A}$ with $\boldsymbol{H}$. In the present case, however, we may exploit the convolutional structure of $\boldsymbol{C}_k$ in order to attempt finding a closed form solution.

E4. Frequency-hopping may be regarded as the time-hopping dual modulation. We present in the following the basic time-hopping signaling scheme, and analogies and differences with frequency-hopping. Investigation of a further generalization is proposed. Strictly speaking, several expressions below are valid asymptotically only, as degrees of freedom in time and/or frequency tend to infinity; we do not account this specific issue here.

In time-hopping, a waveform $\psi_W(t)$ strictly bandlimited to W is transmitted with period $1/W$ for $t \in [0, T]$. By projecting onto the set $\{\psi_W(t - k/W)\colon k \ge 0\}$, one has:

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{b} + \boldsymbol{n}, \tag{1.87}$$

where $\boldsymbol{b}$ is the vector symbol transmitted (or the set of transmitted symbols with time-hopping), $\boldsymbol{n} \sim \mathcal{CN}(0, \mathcal{N}_0)$, and $\boldsymbol{H}$ assumes a different form depending on the channel that is experienced. In particular, when the channel is time-selective, then $\boldsymbol{H} = \boldsymbol{A} = \mathrm{diag}\,\mathsf{A}_0, \ldots, \mathsf{A}_{N-1}$, with typically i.i.d. diagonal elements [39]; when the channel is frequency-selective, $\boldsymbol{H} = \boldsymbol{C}$ is a convolution matrix, such that $H_{ij} = \left\langle c(t), (1/\sqrt{W})\psi_W(t - (i-j)/W) \right\rangle$, which is asymptotically circulant, and therefore diagonalizable with a Fourier basis, *i.e.*, $\boldsymbol{H} = \boldsymbol{C} = \boldsymbol{F}^\dagger \boldsymbol{G} \boldsymbol{F}$, where $\boldsymbol{G}$ is diagonal

with diagonal elements equal to the DFT of the first column of $\boldsymbol{C}$. In case where both time- and frequency-selective fading are present, one has: $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{F}^{\dagger}\boldsymbol{G}\boldsymbol{F}\boldsymbol{b} + \boldsymbol{n}$.

On the contrary, in frequency-hopping (*e.g.* [3]), transmitted signals belong to the orthonormal family

$$\left\{ \exp(j2\pi f_m t) \exp(-j2\pi \frac{W}{2} t) \frac{1}{\sqrt{T}} \mathbb{1}\{t \in [0, T_s)\} \colon f_m = m/T_s,\, m = 0, \dots, WT_s - 1 \right\}.$$

The presence of the term $\exp(-j2\pi \frac{W}{2} t)$ implies that $f \in [-W/2, W/2]$. In presence of a frequency-selective channel, the signal projected onto the above family is: $\tilde{\boldsymbol{y}} = \tilde{\boldsymbol{H}}\tilde{\boldsymbol{b}} + \tilde{\boldsymbol{n}}$, where $\tilde{\boldsymbol{H}}$ is diagonal. On the other hand, in case of time-selective fading, $\tilde{\boldsymbol{H}}$ is circulant, in a dual fashion with respect to frequency-selective fading in time. The relationship between signaling expressed in time or frequency domain is given by the Fourier matrix $\boldsymbol{F}$ and its inverse $\boldsymbol{F}^{-1} = \boldsymbol{F}^{\dagger}$.

A further generalization is not to divide the $N = WT_s$ degrees of freedom available in "slices" that occupies the whole frequency resource for a short time (time-hopping) or the whole time resource in a narrow frequency band (frequency-hopping). For the sake of clarity, consider time and frequency as two coordinate axis on a plane: the time-frequency (TF) plane. Time *vs.* frequency resources are not independent—there exist a coupling expressed by an uncertainty principle. Indeed, one can divide the TF plane in rectangles with dimensions Δt and Δf, such that $\Delta t \cdot \Delta f \geq 1$. This partition can be obtained, for example, by the Weyl-Heisenberg (WH) family. A suitable choice of Δt and Δf alongside WH signaling is able to diagonalize a given channel—under the rather technical assumption of being underspread, which is largely verified by most indoor and outdoor channels. Several investigations of TF signals in connection with wireless communications can be found in [40]. Future analyses may find closed-form expressions of spectral efficiency, by starting from relations similar to eq. (1.85), under the general perspective of a TF discretization, by further consider the effect of pratical impairments due to imperfect channel knowledge and correlations in channels coefficients.

E5. In symbol-asynchronous systems, where users may experience delay $\tau \in [0, T_s)$, the general formula for the spectral efficiency is given by eq. (1.73). Find a closed form as both $N \to \infty$ and $n \to \infty$ seems to be a formidable task in general; on the other hand, a success in this direction, that could maybe rely on Girko's methods [37], would probably uncover a way to address several issues related to asynchronism in spread-spectrum systems, as well as systems which can be formally reconducted to a spread-spectrum description.

1.6 Conclusion

Verdú and Shamai showed in [6] that optimum decoding provides a substantial gain over linear decoding in DS-CDMA, with random spreading. In particular, a bank of single-user matched-filters followed by independent decoders is limited in spectral efficiency at high SNR, and linear multiuser detectors are needed in order to recover

a nonzero spectral efficiency high-SNR slope. This behavior is partly due to the "even" use of degrees of freedom — coinciding in our setting with chips — that is intrisic of DS-CDMA [1].

The object of this chapter was to analyze TH-CDMA with random hopping, and compare its behavior against DS-CDMA; we interpreted time-hopping in the general framework developed in [6, 7]. The present analysis allowed comparison of TH *vs.* DS with same energy per symbol and same bandwidth constraints, and, therefore, showed the effect of the energy "concentration", that is typical of TH. The degree of "unevenness" in TH-CDMA is directly related to the number of pulses N_s representing each symbol. At one extreme, one has maximum "unevenness", where all energy is concentrated in one pulse ($N_s = 1$); the other extreme corresponds to maximum "evenness", $N_s = N$, where TH coincides with DS.

A first result of our analysis was to derive a closed form expression for spectral efficiency of TH-CDMA with optimum decoding in the archetypal case of "unevennes", that is $N_s = 1$, showing that, in this case, DS-CDMA outperforms TH-CDMA; same wideband behavior, but lower high-SNR slope, was observed for TH-CDMA *vs.* DS-CDMA, that is $\min\{1, \beta\} = \mathcal{S}_{\infty,\mathrm{DS}} > \mathcal{S}_{\infty,\mathrm{TH}} = 1 - e^{-\beta}$. A closed form expression for generic N_s remains an open problem; results based on simulations suggested, however, that the spectral efficiency loss may become negligible with as low as two pulses per symbol ($N_s = 2$). This result indicates that the spectral efficiency gap between TH-CDMA and DS-CDMA may be substantially reduced to a very small value, although only using a fraction N_s/N of degrees of freedom *per user*, that asymptotically vanishes as N grows. With respect to the non-asymptotic $N_s > 1$ generic case with N finite, we developed a systematic method for finding approximations of spectral efficiency at the wideband regime ($\mathsf{C} \to 0$). In particular, for small values of N (example $N = 6$), $\mathcal{E}_b/\mathcal{N}_0$ (dB) *vs.* C was independent of N_s, up to a second order approximation.

A different behavior of TH-CDMA with respect to DS-CDMA was observed with a SUMF linear receiver. Contrarily to DS, TH-CDMA spectral efficiency with $N_s = 1$ was unbounded for increasing SNR. As suggested, this asymptotic behavior may be traced back to the non-Gaussian distribution nature of the interference-plus-noise variable observed by each independent single-user decoder, that, in turn, depends on cross-correlation properties of spreading sequences. The same high-SNR slope $\mathcal{S}^{\mathsf{sumf}}_{\infty,\mathrm{TH}} = \beta e^{-\beta}$ was achieved by TH irrespectively of the knowledge that each single-user decoder had about spreading sequences of all other users. For generic N_s, we were able to show that, when N_s obeys the rule $N_s = \alpha N$, $\alpha \in (0, 1)$, same spectral efficiency as DS-CDMA ($\alpha = 1$) was obtained irrespectively of α for $N \to \infty$. The unstructured generic N_s case, where each single-user decoder had no knowledge on spreading sequences of other users, was analyzed by simulation, and results showed that, for increased N_s, the high-SNR slope decreased although never reaching zero (spectral efficiency tending but not reaching a horizontal asymptote). It was interesting to note that the maximum slope for TH, providing a hint on greatest energy efficiency, was reached when the number of users K was equal to the number of chips N, *i.e.*, $\beta = 1$, leading to $\mathcal{S}^{\mathsf{sumf}}_{\infty,\mathrm{TH}} = 1/e \approx 0.367879$.

The bounded nature of spectral efficiency with a SUMF bank in DS-CDMA is overcome, as well known, by using more complex linear receivers, that also account for interference, such as MMSE and decorrelator. Conversely, we showed that, in TH-CDMA, the introduction of complexity in the linear receiver led to only small gains over SUMF, due to the structure of spreading sequences; TH spreading sequences are "more" likely to be linearly dependent than DS ones, in agreement with the intuition based on the cardinality of binary DS *vs.* TH codes, that is 2^N *vs.* $2N$. Reinforcing this argument, note that, for TH with generic finite N_s, we proved that same spectral efficiency was achieved with SUMF, decorrelator and MMSE, when decoders had knowledge on cross-correlations of spreading sequences.

Beyond natural extensions of the present work to channels with fading, multipath, and asynchronism, where the effect of an "uneven" use of degrees of freedom typical of TH should be investigated, we do stress that, from the single-user perspective, TH is a particular instance of impulsive signals. As such, the present theoretical setting, if appropriately adapted to asynchronous links, may serve as a basis for refining the understanding of impulsive communications bound laws.

Appendix

1.A Proof of $C_N^{\mathsf{opt}}(\gamma) \xrightarrow{p} C^{\mathsf{opt}}(\gamma)$

In order to prove eq. (1.20), we have to show that $\mathbb{P}\big(|C_N^{\mathsf{opt}}(\gamma) - C^{\mathsf{opt}}(\gamma)| \geq \epsilon\big) \to 0$ for all $\epsilon > 0$, that is:

$$\mathbb{P}\left(\left|\int_0^\infty \log_2(1+\lambda\gamma)\, d\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^\mathsf{T}}(\lambda) - \int_0^\infty \log_2(1+\lambda\gamma)\, d\mathsf{F}(\lambda)\right| \geq \epsilon\right) \to 0, \quad \forall \epsilon > 0. \tag{1.88}$$

It is useful to rewrite integrals with respect to measures. In particular, let μ_N^{TH} be a measure such that $\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^\mathsf{T}}(x) = \mu_N^{\mathsf{TH}}((-\infty, x])$, and denote by $\bar{\mu}_N^{\mathsf{TH}}$ its expectation. Also, recall that μ^{TH} is given by eq. (1.17). As already pointed out in the proof of Theorem 1, eigenvalues of $\boldsymbol{S}\boldsymbol{S}^\mathsf{T}$ belong to non-negative integers, therefore both measures are discrete (*i.e.*, purely atomic), and one has:

$$\begin{aligned}
&\left|\int_0^\infty \log_2(1+\lambda\gamma)\mu_N^{\mathsf{TH}}(d\lambda) - \int_0^\infty \log_2(1+\lambda\gamma)\mu^{\mathsf{TH}}(d\lambda)\right| \\
&\qquad = \left|\sum_{k=0}^{K} \log_2(1+k\gamma)\mu_N^{\mathsf{TH}}(k) - \sum_{k=0}^{\infty} \log_2(1+k\gamma)\mu^{\mathsf{TH}}(k)\right| \\
&\qquad\qquad \leq \left|\sum_{k=0}^{K} \log_2(1+k\gamma)[\mu_N^{\mathsf{TH}}(k) - \bar{\mu}^{\mathsf{TH}}(k)]\right| \\
&\qquad\qquad\qquad + \left|\sum_{k=0}^{K} \log_2(1+k\gamma)[\bar{\mu}_N^{\mathsf{TH}}(k) - \mu^{\mathsf{TH}}(k)]\right| \\
&\qquad\qquad\qquad\qquad + \left|\sum_{k>K} \log_2(1+k\gamma)\mu^{\mathsf{TH}}(k)\right|,
\end{aligned} \tag{1.89}$$

where (c.f. eq. (1.14)):

$$\bar{\mu}_N^{\mathsf{TH}}(k) = \frac{1}{k!} \cdot \frac{K!/(K-k)!}{N^k} \cdot \left(1 - \frac{1}{N}\right)^{K-k}, \quad \mu^{\mathsf{TH}}(k) = \frac{\beta^k e^{-\beta}}{k!},$$

and, for ease of notation, we write $\mu_N^{\mathsf{TH}}(k)$ and $\mu^{\mathsf{TH}}(k)$ for $\mu_N^{\mathsf{TH}}(k)$ and $\mu^{\mathsf{TH}}(\{k\})$, respectively. It can be shown by elementary methods that, in the LSL, the last two terms in eq. (1.89) both tend to zero. Since:

$$\operatorname{Var}\left[\sum_{k=0}^{K}\log_2(1+k\gamma)\mu_N^{\mathsf{TH}}(k)\right] \le \log_2(1+K\gamma)^2\operatorname{Var}\left[\,\mu_N^{\mathsf{TH}}([0,K])\,\right]$$
$$= \log_2(1+\beta N\gamma)^2\,\operatorname{Var}\left[\,\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}}}(x)\,\right] = O\!\left(\frac{(\log_2 N)^2}{N}\right),$$

eq. (1.88) holds.

1.B Proof of Theorem 2

This appendix is split in two parts. In the first part, we will find average moments $\mathbb{E}[\,m_L\,]$ for finite dimensional systems, where both K and N are finite. In the second part, we will prove convergence in probability of m_L to the Lth moment of a Poisson distribution in the LSL, by showing that $\operatorname{Var}[\,m_L\,]\to 0$.

Part 1: Average Moments of TH-CDMA matrices with $N_{\mathsf{s}}=1$.

The Lth moment of the ESD can be written as follows:

$$m_L = \frac{1}{N}\operatorname{Tr}(\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}})^L = \frac{1}{N}\sum_{i_1,\dots,i_L}\operatorname{Tr}\left[s_{i_1}s_{i_1}^{\mathsf{T}}s_{i_2}s_{i_2}^{\mathsf{T}}\cdots s_{i_{L-1}}s_{i_{L-1}}^{\mathsf{T}}s_{i_L}s_{i_L}^{\mathsf{T}}\right]$$
$$= \frac{1}{N}\sum_{i_1,\dots,i_L} s_{i_L}^{\mathsf{T}}s_{i_1}\cdot s_{i_1}^{\mathsf{T}}s_{i_2}\cdot\cdots\cdot s_{i_{L-1}}^{\mathsf{T}}s_{i_L} \tag{1.90}$$

where summations span over $i_l = 1,\dots,K$. By taking the expectation of eq. (1.90), the following expression for the Lth average moment is derived:

$$\mathbb{E}[\,m_L\,] = \frac{1}{N}\sum_{i_1,\dots,i_L}\mathbb{E}\left[\,s_{i_L}^{\mathsf{T}}s_{i_1}\cdot s_{i_1}^{\mathsf{T}}s_{i_2}\cdot\cdots\cdot s_{i_{L-1}}^{\mathsf{T}}s_{i_L}\,\right]$$
$$= \frac{1}{N}\sum_{i_1,\dots,i_L}\mathbb{P}\big(\pi_{i_L}=\pi_{i_1}\wedge\pi_{i_1}=\pi_{i_2}\wedge\cdots\wedge\pi_{i_{L-1}}=\pi_{i_L}\big)$$
$$= \frac{1}{N}\sum_{i_1,\dots,i_L}\mathbb{P}(\,\pi_{i_1}=\pi_{i_2}=\cdots=\pi_{i_L}\,)$$
$$= \frac{1}{N}\sum_{i_1,\dots,i_L}\sum_{p_1,\dots,p_K}\frac{1}{N^K}\mathbb{1}\{p_{i_1}=\cdots=p_{i_L}\}. \tag{1.91}$$

Sums in eq. (1.91) span over all possible L-uple $(i_1,\dots,i_L)\in[K]^L$ and $(p_1,\dots,p_K)\in[N]^K$. In order to derive a closed form expression, we partition the set $[K]^L$ as follows. Consider ℓ nonempty subsets of indeces $(i_1,\dots,i_L)$; hence, $1\le\ell\le L$. Stirling numbers of the second kind enumerates the number of partitions of a set of L elements into ℓ nonempty subsets, therefore there are $\left\{{L\atop\ell}\right\}$ possible partitions. Assign a (different) value in $[K]^L$ to each subset in the partition; the number of possible assignments is $K(K-1)\cdots(K-\ell+1)$. For any fixed ℓ, there are $K-\ell$ free summations, hence each partition with ℓ parts is counted $N^{K-\ell+1}$ times.

Therefore, eq. (1.91) can be rewritten as:

$$\begin{aligned}\mathbb{E}[m_L] &= \frac{1}{N^{K+1}} \sum_{p_1,\dots,p_K} \sum_{i_1,\dots,i_L} \mathbb{1}\{p_{i_1} = \cdots = p_{i_L}\} \\ &= \frac{1}{N^{K+1}} \sum_{\ell=1}^{L} \left\{ {L \atop \ell} \right\} \frac{K!}{(K-\ell)!} N^{K-\ell+1} \\ &= \sum_{\ell=1}^{L} \left\{ {L \atop \ell} \right\} \frac{K!}{(K-\ell)!} \frac{1}{N^{\ell}}. \end{aligned} \tag{1.92}$$

In the LSL, one has:

$$\mathbb{E}[m_L] = \sum_{\ell=1}^{L} \left\{ {L \atop \ell} \right\} \beta^{\ell}, \tag{1.93}$$

that is exactly the Bell polynomial $\mathrm{B}_L(\beta)$ providing the Lth moment of a Poisson distribution with mean β.

Remark 1. Interestingly, the Lth moment of the Marčenko-Pastur law (c.f. eq. (1.19)) can be expressed as follows (see *e.g.* [22, 29]):

$$m_L^{\mathsf{MP}} = \sum_{\ell=1}^{L} \mathcal{N}_{\ell} \beta^{\ell}, \quad \mathcal{N}_{\ell} = \frac{1}{L} \binom{L}{\ell} \binom{L}{\ell - 1}, \tag{1.94}$$

where $\mathcal{N}_{\ell}$ is the number of *non-crossing partitions* of the set $[L]$ into ℓ blocks, also known as *Narayana number*. As a remark, the sum of Narayana numbers over $[L]$ is the Lth Catalan number, that has many combinatorial interpretations (see *e.g.* [41, 42]).

On the other hand, eq. (1.93) is formally similar to eq. (1.94), with Stirling number of the second kind in place of Narayana numbers. While the latter enumerate non-crossing partitions only, the former enumerate all partitions, both crossing and non-crossing ones. As a remark, the sum of Stirling numbers over $[L]$, or, equivalently, the value of $\mathbb{E}[m_L]$ with unit load, is equal to the Lth Bell number.

Part 2: $\mathrm{Var}[m_L] \to 0$.

We will find $\mathbb{E}\big[m_L^2\big]$, from which $\mathrm{Var}[m_L] = \mathbb{E}\big[m_L^2\big] - \mathbb{E}[m_L]^2$. Proceeding as for eq. (1.90) and (1.91), we have:

$$\begin{aligned}\mathbb{E}\big[m_L^2\big] &= \frac{1}{N^2} \sum_{\substack{i_1,\dots,i_L \\ j_1,\dots,j_L}} \mathbb{E}\Big[s_{i_L}^{\mathsf{T}} s_{i_1} \cdot s_{i_1}^{\mathsf{T}} s_{i_2} \cdots s_{i_{L-1}}^{\mathsf{T}} s_{i_L} \cdot s_{j_L}^{\mathsf{T}} s_{j_1} \cdot s_{j_1}^{\mathsf{T}} s_{j_2} \cdots s_{j_{L-1}}^{\mathsf{T}} s_{j_L} \Big] \\ &= \frac{1}{N^2} \sum_{\substack{i_1,\dots,i_L \\ j_1,\dots,j_L}} \mathbb{P}\big(\pi_{i_1} = \pi_{i_2} = \cdots = \pi_{i_L} \wedge \pi_{j_1} = \pi_{j_2} = \cdots = \pi_{j_L} \big) \qquad (1.95) \\ &= \frac{1}{N^2} \sum_{\substack{i_1,\dots,i_L \\ j_1,\dots,j_L}}^{*} \mathbb{P}\big(\pi_{i_1} = \pi_{i_2} = \cdots = \pi_{i_L} \wedge \pi_{j_1} = \pi_{j_2} = \cdots = \pi_{j_L} \big) \qquad (1.96) \end{aligned}$$

$$+\frac{1}{N^2}\sum_{\substack{i_1,\dots,i_L\\ j_1,\dots,j_L}}^{**}\mathbb{P}(\pi_{i_1}=\pi_{i_2}=\cdots=\pi_{i_L}\wedge\pi_{j_1}=\pi_{j_2}=\cdots=\pi_{j_L}), \tag{1.97}$$

where sums in eq. (1.96) and in eq. (1.97) are over indeces such that $\{\pi_{i_1},\pi_{i_2},\dots,\pi_{i_L}\}\cap\{\pi_{j_1},\pi_{j_2},\dots,\pi_{j_L}\}=\emptyset$ and $\{\pi_{i_1},\pi_{i_2},\dots,\pi_{i_L}\}\cap\{\pi_{j_1},\pi_{j_2},\dots,\pi_{j_L}\}\neq\emptyset$, respectively. Suppose that $\{\pi_{i_1},\pi_{i_2},\dots,\pi_{i_L}\}\cap\{\pi_{j_1},\pi_{j_2},\dots,\pi_{j_L}\}=\emptyset$. Hence:

$$\begin{aligned}&\mathbb{P}(\pi_{i_1}=\pi_{i_2}=\cdots=\pi_{i_L}\wedge\pi_{j_1}=\pi_{j_2}=\cdots=\pi_{j_L})\\&\quad=\mathbb{P}(\pi_{i_1}=\pi_{i_2}=\cdots=\pi_{i_L})\cdot\mathbb{P}(\pi_{j_1}=\pi_{j_2}=\cdots=\pi_{j_L})=\frac{1}{N^{\ell_1-1}}\cdot\frac{1}{N^{\ell_2-1}},\end{aligned} \tag{1.98}$$

where, akin to the previous section, ℓ_1 and ℓ_2 are the number of parts of a partition of $\{i_1,\dots,i_L\}$ and $\{j_1,\dots,j_L\}$, respectively. Therefore, for disjoint subsets of indeces, eq. (1.96) reduces to:

$$\begin{aligned}&\frac{1}{N^2}\sum_{\substack{i_1,\dots,i_L\\ j_1,\dots,j_L}}^{*}\mathbb{P}(\pi_{i_1}=\pi_{i_2}=\cdots=\pi_{i_L}\wedge\pi_{j_1}=\pi_{j_2}=\cdots=\pi_{j_L})\\&\qquad=\sum_{1\le\ell_1+\ell_2\le L}\left\{{L\atop\ell_1}\right\}\left\{{L\atop\ell_2}\right\}\frac{K!}{(K-\ell_1-\ell_2)!}\frac{1}{N^{\ell_1+\ell_2}},\end{aligned} \tag{1.99}$$

which remains finite in the LSL. Suppose now that:

$$\{\pi_{i_1},\pi_{i_2},\dots,\pi_{i_L}\}\cap\{\pi_{j_1},\pi_{j_2},\dots,\pi_{j_L}\}\neq\emptyset.$$

In this case, at least for one k it happens that $j_k\in\{i_1,\dots,i_L\}$, and therefore there is only one long chain of equalities in eq. (1.95), *i.e.*,

$$\mathbb{P}(\pi_{i_1}=\pi_{i_2}=\cdots=\pi_{i_L}=\pi_{j_1}=\pi_{j_2}=\cdots=\pi_{j_L}),$$

where there are at most $2L-1$ different RVs. Hence, for any fixed partition of $\{i_1,\dots,i_L,j_1,\dots,j_L\}$ into $\ell\le 2L-1$ parts, the above probability is $1/N^{\ell-1}$, and we can assign to each part a (different) value in $[K]$ in $K(K-1)\cdots(K-\ell+1)$ ways, similarly as in the previous section. Therefore, each part accounts for $O(1/N)$, and since the number of partitions remains finite, the overall sum in eq. (1.97) is $O(1/N)$. Since eq. (1.99) is equal to $\mathbb{E}[m_L]^2$ up to an $O(1/N)$ term, we have $\mathrm{Var}[m_L]\to 0$ in the LSL.

1.C Relationship between Rank and high-SNR slope

From the definition of ESD, one has:

$$\mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^\intercal}(0)=\frac{1}{N}\sum_{i=1}^{N}\mathbb{1}\{\lambda_i=0\},$$

and the number of zero eigenvalues is the geometric multiplicity of $\lambda = 0$, that is equal to the dimension of the associated eigenspace, which is the nullity subspace of $\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}}$. In general, from the Rank-Nullity Theorem, it results:

$$\begin{aligned}\dim \operatorname{Ker} \boldsymbol{S}\boldsymbol{S}^{\mathsf{T}} &= N - \dim \operatorname{Im} \boldsymbol{S}\boldsymbol{S}^{\mathsf{T}} \\ &= N - \operatorname{rank} \boldsymbol{S}\boldsymbol{S}^{\mathsf{T}} \\ &= N - \operatorname{rank} \boldsymbol{S},\end{aligned}$$

and, therefore:

$$1 - \frac{1}{N} \operatorname{rank} \boldsymbol{S} = \mathsf{F}_N^{\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}}}(0) \to \mathsf{F}(0),$$

from which $\mathcal{S}_\infty$ can be obtained as $1 - \mathsf{F}(0)$.

1.D Asymptotics in the wideband regime for $N_{\mathsf{s}} > 1$

1) Minimum energy-per-bit.

We will show that:

$$\bar{\lambda}_N := \frac{1}{N} \sum_{i=1}^{N} \lambda_i = \beta$$

surely. In fact, while $\{\lambda_i \colon i \in [N]\}$ are RVs, $\bar{\lambda}_N$ is not, for any N. This follows from:

$$\frac{1}{N} \sum_{i=1}^{N} \lambda_i = \frac{1}{N} \operatorname{Tr}(\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}}) = \frac{1}{N} \operatorname{Tr}(\boldsymbol{S}^{\mathsf{T}}\boldsymbol{S}) = \frac{1}{N} \sum_{i=1}^{K} \boldsymbol{s}_i^{\mathsf{T}} \boldsymbol{s}_i = \frac{1}{N} \cdot K = \beta,$$

and therefore $\mathbb{E}[\lambda] = \beta$ since $\bar{\lambda}_N \to \mathbb{E}[\lambda]$. Hence, the minimum energy-per-bit is:

$$\eta_{\mathsf{min}} = \frac{\beta}{\mathbb{E}[\lambda]} \ln 2 = \ln 2.$$

2) Wideband slope $\mathcal{S}_0$.

We will show that:

$$S_N := \frac{1}{N} \sum_{i=1}^{N} \lambda_i^2 \xrightarrow{p} \beta(1+\beta).$$

Denoting with $\rho_{ij} = (\boldsymbol{S}^{\mathsf{T}}\boldsymbol{S})_{ij}$, since:

$$\begin{aligned}\frac{1}{N} \sum_{i=1}^{N} \lambda_i^2 = \frac{1}{N} \operatorname{Tr}(\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}})^2 = \frac{1}{N} \operatorname{Tr}(\boldsymbol{S}^{\mathsf{T}}\boldsymbol{S})^2 = \frac{1}{N} \sum_{i=1}^{K} \sum_{j=1}^{K} \rho_{ij}^2 \\ = \beta + \frac{1}{N} \sum_{i=1}^{K} \sum_{\substack{j=1 \\ j \neq i}}^{K} \rho_{ij}^2, \quad (1.100)\end{aligned}$$

one has:

$$\mathbb{E}[\,S_N\,] = \beta + \frac{1}{N}K(K-1)\,\mathbb{E}\big[\,\rho_{12}^2\,\big]\,.$$

Denote with:

$$\rho := \rho_{12} = \boldsymbol{s}_1^{\mathsf{T}}\boldsymbol{s}_2 = \sum_{m=1}^{N_{\mathsf{s}}} \boldsymbol{s}_{1m}^{\mathsf{T}}\boldsymbol{s}_{2m},$$

where $\boldsymbol{s}_k = [\boldsymbol{s}_{k1}^{\mathsf{T}}, \ldots, \boldsymbol{s}_{kN_{\mathsf{s}}}^{\mathsf{T}}]^{\mathsf{T}}$ is an $(N_{\mathsf{s}}, N_{\mathsf{h}})$-sequence (see Definition 2). Indicate with $\rho_m := \boldsymbol{s}_{1m}^{\mathsf{T}}\boldsymbol{s}_{2m}$. The moment generating function (MGF) of ρ is:

$$M_\rho(t) := \mathbb{E}\big[\,e^{t\rho}\,\big] = \mathbb{E}\Big[\,e^{t(\rho_1+\cdots+\rho_{N_{\mathsf{s}}})}\,\Big] = \big(\,\mathbb{E}\big[\,e^{t\rho_1}\,\big]\,\big)^{N_{\mathsf{s}}} = \left(\int e^{t\xi}\mu(d\xi)\right)^{N_{\mathsf{s}}},$$

where:

$$\mu = \frac{1}{2N_{\mathsf{h}}}\delta_{-\frac{1}{N_{\mathsf{s}}}} + \left(1 - \frac{1}{N_{\mathsf{h}}}\right)\delta_0 + \frac{1}{2N_{\mathsf{h}}}\delta_{\frac{1}{N_{\mathsf{s}}}}.$$

Hence:

$$M_\rho(t) = \left[\left(1 - \frac{1}{N_{\mathsf{h}}}\right) + \frac{1}{N_{\mathsf{h}}}\cosh\left(\frac{t}{N_{\mathsf{s}}}\right)\right]^{N_{\mathsf{s}}}, \tag{1.101}$$

and $\mathbb{E}\big[\,\rho^2\,\big] = M_\rho''(0) = 1/N$, and also:

$$\mathbb{E}[\,S_N\,] = \beta + \beta\frac{K-1}{N} \to \beta(1+\beta).$$

We will now show that $\mathrm{Var}[\,S_N\,] = O(1/N)$, proving convergence in probability of S_N to $\beta + \beta^2$. From eq. (1.100), one has:

$$\begin{aligned}
\mathbb{E}\big[\,S_N^2\,\big] &= \beta^2 + \frac{2\beta}{N}K\cdot\frac{K-1}{N} + \frac{1}{N^2}\,\mathbb{E}\left(\sum_{i=1}^{K}\sum_{q=1}^{K}\sum_{\substack{j=1\\ j\neq i}}^{K}\sum_{\substack{r=1\\ r\neq q}}^{K}\rho_{ij}^2\rho_{qr}^2\right)\\
&= \beta^2 + \frac{2\beta}{N}K\cdot\frac{K-1}{N} + \frac{1}{N^2}K(K-1)\big\{\,\mathbb{E}\big[\,\rho^4\,\big] + (K(K-1)-1)\,\mathbb{E}\big[\,\rho^2\,\big]^2\big\}\\
&= \beta^2 + 2\beta^3 + \beta^4 + O\left(\frac{1}{N}\right)
\end{aligned}$$

hence $\mathrm{Var}\big[\,S_N^2\,\big] = O(1/N)$.

1.E Mutual Information of SUMF when single-user decoders have knowledge on cross-correlations.

The SUMF channel for user 1, as given by eq. (1.33), is:

$$\mathsf{y}_1 = \mathsf{b}_1 + \sum_{k=2}^{K}\rho_{1k}\mathsf{b}_k + n_1.$$

Assuming Gaussian inputs, $\mathsf{b}_i \sim \mathcal{CN}(0, \mathcal{E})$, the conditional mutual information on $\{\rho_{12}, \ldots, \rho_{1K}\}$ expressed in bits per channel use per user is:

$$I(\mathsf{y}_1; \mathsf{b}_1 | \rho_{12}, \ldots, \rho_{1K}) = \mathbb{E}\left[\log_2\left(1 + \frac{\gamma}{1+\varsigma\gamma}\right)\right], \tag{1.102}$$

where expectation is over $\{\rho_{12}, \ldots, \rho_{1K}\}$, and $\varsigma := \sum_{k=2}^{K} \rho_{1k}^2$. We find below the PDF of ς in the LSL.

From eq. (1.101), the characteristic function (CF) of the generic RV $\rho := \rho_{1k}$ is:

$$\begin{aligned}
\varphi_\rho(t) := \mathbb{E}\left[e^{it\rho}\right] &= \left[1 - \frac{N_s}{N}\left(1 - \cos\left(\frac{t}{N_s}\right)\right)\right]^{N_s} \\
&= \sum_{m=0}^{N_s} \binom{N_s}{m} (-1)^m \frac{N_s^m}{N^m}\left[1 - \cos\left(\frac{t}{N_s}\right)\right]^m \\
&= 1 - \frac{N_s^2}{N}\left[1 - \cos\left(\frac{t}{N_s}\right)\right] + O\left(\frac{1}{N^2}\right),
\end{aligned} \tag{1.103}$$

and the CF of ς is:

$$\varphi_\varsigma(t) = \mathbb{E}\left[e^{it\varsigma}\right] = \mathbb{E}\left[e^{it\rho^2}\right]^{K-1}. \tag{1.104}$$

The last expectation can be computed as:

$$\begin{aligned}
\mathbb{E}\left[e^{it\rho^2}\right] &= \frac{1}{2\pi}\int_{\mathbb{R}} d\omega\, \varphi_\rho(\omega) \int_{\mathbb{R}} dx\, e^{itx^2} e^{-i\omega x} \\
&= \frac{1}{2\pi}\int_{\mathbb{R}} d\omega\, \varphi_\rho(\omega) \sqrt{i\frac{\pi}{t}} e^{-i\frac{\omega^2}{4t}} \\
&= 1 + \frac{N_s^2}{N}\left[e^{it/N_s^2} - 1\right] + O\left(\frac{1}{N^2}\right).
\end{aligned}$$

In the LSL, eq. (1.104) becomes:

$$\varphi_\varsigma(t) \to e^{\beta N_s^2 (e^{it/N_s^2} - 1)},$$

which is the CF of a Poisson RV with measure:

$$\mu_\varsigma = \sum_{k \geq 0} \frac{(\beta N_s^2)^k}{k!} e^{-\beta N_s^2} \delta_{k/N_s^2}.$$

Therefore, from eq. (1.102), mutual information converges to:

$$I(\mathsf{y}_1; \mathsf{b}_1 | \rho_{12}, \ldots, \rho_{1K}) \to \sum_{k \geq 0} \frac{(\beta N_s^2)^k}{k!} e^{-\beta N_s^2} \log_2\left(1 + \frac{\gamma}{1 + \frac{k}{N_s^2}\gamma}\right).$$

1.F Proof of Theorem 6

Consider the output of the SUMF of user 1, that is given by eq. (1.33), divided by $\sqrt{N_0}$:

$$Y_1 = b_1 + \sum_{k=2}^{K} \rho_{1k} b_k + N_1 = b_1 + J_1 + N_1 = b_1 + Z_1,$$

where $N_1 \sim \mathcal{CN}(0,1)$ and we assume $b_k \sim \mathcal{CN}(0,\gamma)$. The goal is to find mutual information $I(Y_1;b_1)$, that reduces to find P_{Y_1} and $P_{Y_1|b_1}$, both of which easily follow from P_{J_1}.

From eq. (1.103), one can write the CF of each term $\rho_{1k}b_k$ as:

$$\varphi_{\rho b}(t) = \mathbb{E}[\varphi_\rho(bt)] = 1 - \frac{N_s^2}{N}\left(1 - e^{-\frac{\gamma}{2N_s^2}t^2}\right) + O\left(\frac{1}{N^2}\right),$$

and, therefore, the CF of J in the LSL is:

$$\varphi_J(t) = \varphi_{\rho b}(t)^{K-1} \to \exp\left(\beta N_s^2\left(e^{-\frac{\gamma}{2N_s^2}t^2} - 1\right)\right) = \sum_{k\geq 0} \frac{(\beta N_s^2)^k}{k!} e^{-\beta N_s^2} e^{-\frac{k\gamma}{2N_s^2}t^2},$$

which is the CF of:

$$P_J = \sum_{k\geq 0} \frac{(\beta N_s^2)^k}{k!} e^{-\beta N_s^2} \mathcal{CN}\left(0, k\gamma/N_s^2\right).$$

Therefore, Z_1 and Y_1 are distributed as:

$$P_Z = \sum_{k\geq 0} \frac{(\beta N_s^2)^k}{k!} e^{-\beta N_s^2} \mathcal{CN}\left(0, 1 + k\gamma/N_s^2\right). \tag{1.105}$$

and:

$$P_Y = \sum_{k\geq 0} \frac{(\beta N_s^2)^k}{k!} e^{-\beta N_s^2} \mathcal{CN}\left(0, 1 + \gamma + k\gamma/N_s^2\right). \tag{1.106}$$

Mutual information can be obtained as $I(Y_1;b_1) = h(P_Y) - h(P_Z)$.

1.G Proof of Eq. (1.43)

The goal is to find the quantity $\mathcal{S}^{\text{sumf}}_{\infty,\text{TH}^\star} := \lim_{\gamma\uparrow\infty} \gamma dI/d\gamma$, where I (nats/s/Hz) is given in Theorem 6. Assuming that a limit does exist, we will upper and lower bound $h(P)$ with bounds having the same first derivative as $\gamma \to \infty$. Here P is a generic linear combination of Gaussian distributions with possibly different moments and weighted by a real sequence $\mathsf{w}_k \geq 0$, that is,

$$P := \sum_{k\geq 0} \mathsf{w}_k \mathcal{CN}\left(\mu_k, \sigma_k^2\right).$$

Upper bound. Apply the elementary inequality $(x_1+x_2)\ln(x_1+x_2) \geq x_1 \ln x_1 + x_2 \ln x_2$ with $x_1 \geq 0$ and $x_2 \geq 0$, aptly generalized to a denumerable number of summands, to the entropy of P, that is:

$$\begin{aligned} h(P) &= -\int_{\mathbb{C}} dz \left(\sum_{k\geq 0} w_k \, \mathcal{CN}\left(z; \mu_k, \sigma_k^2\right) \right) \ln \left(\sum_{i\geq 0} w_i \, \mathcal{CN}\left(z; \mu_i, \sigma_i^2\right) \right) \\ &\leq -\int_{\mathbb{C}} dz \sum_{k\geq 0} w_k \, \mathcal{CN}\left(z; \mu_k, \sigma_k^2\right) \ln\left[w_k \, \mathcal{CN}\left(z; \mu_k, \sigma_k^2\right) \right] \\ &= \sum_{k\geq 0} -w_k \ln w_k + \sum_{k\geq 0} w_k \, h(\mathcal{CN}\left(\mu_k, \sigma_k^2\right)) =: h_{\mathsf{P}}(P) + h_{\mathsf{G}}(P), \end{aligned}$$

where $h(\mathcal{CN}\left(\mu_k, \sigma_k^2\right)) = h(\mathcal{CN}\left(0, \sigma_k^2\right)) = \ln(\pi e \sigma_k^2)$, irrespective of μ_k; $h_{\mathsf{P}}(P)$ is constant in γ, while $h_{\mathsf{G}}(P)$, that is a weighted sum of entropies $\{h(\mathcal{CN}\left(0, \sigma_k^2\right))\}_{k\geq 0}$, may depend on γ via $\{\sigma_k^2\}_{k\geq 0}$.

Lower bound. It is easy to show via variational calculus that the solution of the following problem:

$$\min_g \int_{\mathbb{R}} -dx \, f(x) \ln g(x) \quad \text{s.t.} \quad \int_{\mathbb{R}} dx \, g(x) = 1$$

for sufficiently well-behaved functions $f\colon \mathbb{R} \to [0, +\infty)$ is $g(x) = f(x)$, this result being a generalization of the Gibbs' inequality to continuous functions. We can apply the above result for $f = \mathcal{CN}\left(\mu_k, \sigma_k^2\right)$, from which $h(P)$ is bounded from below as follows:

$$\begin{aligned} h(P) = \sum_{k\geq 0} w_k \int_{\mathbb{C}} -dz \, \mathcal{CN}\left(z; \mu_k, \sigma_k^2\right) \ln\left(\sum_{i\geq 0} w_i \, \mathcal{CN}\left(z; \mu_i, \sigma_i^2\right) \right) \\ \geq \sum_{k\geq 0} w_k \, h(\mathcal{CN}\left(0, \sigma_k^2\right)) =: h_{\mathsf{G}}(P). \end{aligned}$$

Bounds. From above bounds, it follows that $h_{\mathsf{G}}(P) \leq h(P) \leq h_{\mathsf{G}}(P) + h_{\mathsf{P}}(P)$, and since $h_{\mathsf{P}}(P)$ is constant in γ, one has:

$$\lim_{\gamma\to\infty} \gamma \frac{dh(P)}{d\gamma} = \lim_{\gamma\to\infty} \gamma \frac{dh_{\mathsf{G}}(P)}{d\gamma}, \tag{1.107}$$

provided that the limit on the LHS does exist. By specifying upper and lower bounds to Poisson weights, $w_k := e^{-\beta}\beta^k/k!$, and P equal to either P_Y or P_Z as given by eqs. (1.41) and (1.42), respectively, it results:

$$\begin{aligned} S_{\infty,\mathrm{TH}^\star}^{\mathsf{sumf}} &= \lim_{\gamma\to\infty} \gamma \frac{d}{d\gamma} \{\beta \cdot \left[h(P_Y) - h(P_Z) \right]\} \\ &= \beta \cdot \lim_{\gamma\to\infty} \gamma \frac{d}{d\gamma} \{h_{\mathsf{G}}(P_Y) - h_{\mathsf{G}}(P_Z)\}, \end{aligned} \tag{1.108}$$

and, by direct computations, it follows that:

$$\frac{dh_{\mathsf{G}}(P_Z)}{d\gamma} = \frac{1-e^{-\beta}}{\gamma} + o\left(\frac{1}{\gamma^2}\right), \tag{1.109}$$

$$\frac{dh_{\mathsf{G}}(P_Y)}{d\gamma} = \frac{1}{\gamma} + o\left(\frac{1}{\gamma^2}\right). \tag{1.110}$$

hence $\mathcal{S}_{\infty,\mathrm{TH}^\star}^{\mathsf{sumf}} = \beta e^{-\beta}$.

1.H Spectral Efficiency of SUMF for $N_{\mathsf{s}} = \alpha N$, $\alpha \in (0,1]$, As $N \to \infty$.

In this appendix we will show that spectral efficiency of the single-user SUMF channel:

$$\begin{aligned} \mathsf{y}_1 &= \mathsf{b}_1 + \sum_{k=2}^{K} \rho_{1k}\mathsf{b}_k + n_1, \\ &= \mathsf{b}_1 + z_1 \end{aligned} \tag{1.33}$$

for TH-CDMA with $N_{\mathsf{s}} = \alpha N$, $\alpha \in (0,1]$, is same as that of DS-CDMA, therefore generalizing a previous result of Verdú and Shamai [6] to which we reduce when $\alpha = 1$. The argument below follows closely that developed in [6].

Part 1: Non-Gaussian capacity bounds.

Spectral efficiency C of channel of eq. (1.33) is obtained by knowing cross-correlations $\mathcal{R}_1 = \{\rho_{1k}\}_{k=2}^{K}$ and input distributions $\mathcal{F}_1 = \{\mathsf{f}_{\mathsf{b}_k}\}_{k=2}^{K}$. In general, z_1 is non-Gaussian, and C depends on both $\mathcal{R}_1$ and $\mathcal{F}_1$, *i.e.*, $C(\mathcal{R}_1, \mathcal{F}_1)$. We denote with $\tilde{z}_1$ a Gaussian r.v. with same mean and variance as z_1, and, for simplicity, $C(\mathcal{R}_1, \mathcal{F}_1)$ by C. A classic result [43] allows to bound C as follows:

$$\tilde{C} \le C \le \tilde{C} + \mathcal{D}(\, z_1 \| \tilde{z}_1 \,), \tag{1.111}$$

where $\tilde{C}$ is the spectral efficiency of a Gaussian channel with noise variance given by:

$$\mathrm{Var}[\, z_1 \,|\, \mathcal{R}_1 \,] = \mathcal{N}_0 + \mathcal{E} \sum_{k=2}^{K} \rho_{1k}^2,$$

that is,

$$\tilde{C} = \log\left(1 + \frac{\mathcal{E}}{\mathrm{Var}[\, z_1 \,|\, \mathcal{R}_1 \,]}\right), \tag{1.112}$$

and where $\mathcal{D}(\, z_1 \| \tilde{z}_1 \,)$ is the Kullback-Leibler divergence between the distribution of the interference-plus-noise term z_1 and a Gaussian distribution with same variance; we recall that the divergence $\mathcal{D}(\, P \| Q \,)$ between two distributions P and Q with

densities p and q, respectively, is defined as follows:

$$\mathcal{D}(\,P||Q\,)=\int p(x)\ln\frac{p(x)}{q(x)}dx,$$

and $\mathcal{D}(\,P||Q\,)\geq 0$ with equality when $P=Q$.

We will prove below that $\mathcal{D}(\,z_1||\tilde{z}_1\,)\xrightarrow{\text{a.s.}}0$ when $N_\text{s}=\alpha N$ with $\alpha\in(0,1]$ by showing that z_1 is asymptotically Gaussian distributed.

We proceed by verifying the Lindeberg-Feller condition as proposed in [6]:

$$\lim_{K\to\infty}\sum_{k=2}^{K}\mathbb{E}\Big[\,\rho_{1k}^2\mathrm{b}_k^2\,\mathbb{1}_{\{\rho_{1k}^2\mathrm{b}_k^2>\xi\}}\,\Big|\,\rho_{1k}\,\Big]=0,\quad\forall\xi>0;$$

for our purposes, it is sufficient to show that:

$$\sum_{k=2}^{K}\mathbb{E}\Big[\,\rho_{1k}^2\mathrm{b}_k^2\,\mathbb{1}_{\{\rho_{1k}^2\mathrm{b}_k^2>\xi\}}\,\Big|\,\rho_{1k}\,\Big]\leq$$

$$(K-1)\cdot\frac{h}{N}\,\mathbb{E}\Big[\,\mathrm{b}_2^2\,\mathbb{1}_{\{\mathrm{b}_2^2>N\xi/h\}}\,\Big]+\sum_{k=2}^{K}\rho_{1k}^2\,\mathbb{1}_{\{N\rho_{1k}^2>h\}}\xrightarrow{K\to\infty}0,\quad(1.113)$$

where the inequality is true $\forall h\geq 0$.

Part 2: Verification of the Lindeberg-Feller condition. The first term in the right-hand side of eq. (1.113) tends to zero, irrespective of N_s, since $\mathrm{Var}[\,\mathrm{b}_2\,]$ is finite while $N\xi/h\to\infty$ for any ξ and h.

The second term, thanks to the law of large numbers, tends a.s. to:

$$\sum_{k=2}^{K}\rho_{1k}^2\,\mathbb{1}_{\{N\rho_{1k}^2>h\}}\xrightarrow{\text{a.s.}}(K-1)\cdot\frac{\beta}{K}\,\mathbb{E}\Big[\,N\rho_{12}^2\,\mathbb{1}_{\{N\rho_{12}^2>h\}}\,\Big],$$

therefore it is sufficient to show that:

$$\mathbb{E}\Big[\,N\rho_{12}^2\,\mathbb{1}_{\{N\rho_{12}^2>h\}}\,\Big]\xrightarrow{N\to\infty}0.\quad(1.114)$$

Call $\chi=\sqrt{N}\rho_{12}$; the c.f. of χ is:

$$\mathcal{C}_\chi(u)=\left[\left(1-\frac{1}{N_\text{h}}\right)+\frac{1}{2N_\text{h}}\left(e^{j2\pi u\frac{\sqrt{N}}{N_\text{s}}}+e^{-j2\pi u\frac{\sqrt{N}}{N_\text{s}}}\right)\right]^{N_\text{s}}.\quad(1.115)$$

When $N_\text{s}=\alpha N$ with fixed $\alpha\in(0,1]$, we have $N=N_\text{s}N_\text{h}=\alpha N\cdot N_\text{h}$, hence $N_\text{h}=1/\alpha$; as $N\to\infty$, eq. (1.115) becomes:

$$\mathcal{C}_\chi(u)=e^{-2\pi^2\cdot\frac{1}{\alpha N_\text{h}}\cdot u^2}=e^{-2\pi^2u^2},\quad(1.116)$$

therefore χ is asymptotically Gaussian distributed with zero mean and unit variance. Applying the Markov inequality:

$$\mathbb{E}\left[\chi^2 \mathbb{1}_{\{\chi^2>h\}}\right] \leq \frac{1}{h},$$

eq. (1.114) is satisfied as $h \to \infty$.

1.I Closed form expression of eq. (1.53) for a General Class of Linear Receivers.

The discrete synchronous multiple-access channel considered is (c.f. eq. (1.2)):

$$\boldsymbol{y} = \boldsymbol{S}\boldsymbol{b} + \boldsymbol{n}.$$

The output of a generic linear receiver $\boldsymbol{W}^{\mathsf{T}}$ is as follows:

$$\begin{aligned}\boldsymbol{r} &:= \boldsymbol{W}^{\mathsf{T}}\boldsymbol{y} = \boldsymbol{W}^{\mathsf{T}}\boldsymbol{S}\boldsymbol{b} + \boldsymbol{W}^{\mathsf{T}}\boldsymbol{n} \\ &= \boldsymbol{G}\boldsymbol{b} + \boldsymbol{\nu},\end{aligned} \tag{1.117}$$

where $\boldsymbol{G} = \boldsymbol{W}^{\mathsf{T}}\boldsymbol{S}$ and $\boldsymbol{\nu} \sim \mathcal{CN}\left(\mathbf{0}, \mathcal{N}_0 \boldsymbol{W}^{\mathsf{T}}\boldsymbol{W}\right)$. We consider the following linear receiver structure parametrized by α and ϵ:

$$\boldsymbol{W}^{\mathsf{T}} = \boldsymbol{S}^{\mathsf{T}}(\epsilon \boldsymbol{S}\boldsymbol{S}^{\mathsf{T}} + \alpha \boldsymbol{I})^{-1}; \tag{1.118}$$

by setting $\epsilon = 1$, decorrelator and MMSE receivers are obtained as special cases for $\alpha \to 0$ and $\alpha = 1/\gamma$, respectively; by setting $\epsilon = 0$ and $\alpha \neq 0$, one obtains SUMF.

By focusing on user 1, the output of channel of (1.117) can be written as eq. (1.119), which is reported here for reference:

$$r_1 = G_{11}\mathsf{b}_1 + \sum_{k=2}^{K} G_{1k}\mathsf{b}_k + \nu_1. \tag{1.119}$$

As in the proof of Theorem 1, we say that s users are in chip i when the ith diagonal element of $\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}}$ is equal to s. Since $\boldsymbol{S}\boldsymbol{S}^{\mathsf{T}}$ is diagonal, one can write:

$$\left[(\epsilon \boldsymbol{S}\boldsymbol{S}^{\mathsf{T}} + \alpha \boldsymbol{I})^{-1}\right]_{ii} = \frac{1}{\alpha + \epsilon u_i},$$

where u_i is the number of users in chip i; since $\boldsymbol{s}_i \in \{\pm \boldsymbol{e}_n\}_{n=1}^{N}$, one has $\boldsymbol{s}_i = (-1)^{\mathsf{a}_i}\boldsymbol{e}_{\pi_i}$ for some $\mathsf{a}_i \in \{0,1\}$ and $\pi_i \in [N]$, and therefore u_i can be formally expressed as $u_i = |\{k \in [K] \colon \pi_k = i\}|$. The generic element G_{ij} in eq. (1.119) is explicitly given by:

$$\begin{aligned}G_{ij} &= \boldsymbol{s}_i^{\mathsf{T}}(\epsilon \boldsymbol{S}\boldsymbol{S}^{\mathsf{T}} + \alpha \boldsymbol{I})^{-1}\boldsymbol{s}_j \\ &= (-1)^{\mathsf{a}_i}\boldsymbol{e}_{\pi_i}^{\mathsf{T}}(\epsilon \boldsymbol{S}\boldsymbol{S}^{\mathsf{T}} + \alpha \boldsymbol{I})^{-1}\boldsymbol{e}_{\pi_j}(-1)^{\mathsf{a}_j}\end{aligned}$$

$$= (-1)^{\mathrm{a}_i+\mathrm{a}_j}\delta_{\pi_i,\pi_j}\cdot\frac{1}{\alpha+\epsilon u_{\pi_i}} = \rho_{ij}\cdot\frac{1}{\alpha+\epsilon v_i},$$

where we denoted by $v_i := u_{\pi_i}$, that is also equal to the number of spreading sequences equal to either $\boldsymbol{s}_i$ or $-\boldsymbol{s}_i$, *i.e.*:

$$v_i = \sum_{k=1}^{K}|\rho_{ik}| = \sum_{k=1}^{K}|\rho_{ki}|.$$

With similar computations, the generic element of the conditional covariance matrix of the noise vector in eq. (1.117) given $\{v_i\}$, and, therefore, given $\boldsymbol{S}$, is:

$$\left[\boldsymbol{\Sigma}_{\boldsymbol{\nu}|\boldsymbol{S}}\right]_{ij} = \mathbb{E}\Big[\boldsymbol{W}^{\mathsf{T}}\boldsymbol{n}\boldsymbol{n}^{\dagger}\boldsymbol{W}\,\Big|\,\boldsymbol{S}\Big]_{ij} = \mathcal{N}_0[\boldsymbol{W}^{\mathsf{T}}\boldsymbol{W}]_{ij} = \mathcal{N}_0\rho_{ij}\cdot\frac{1}{(\alpha+\epsilon v_i)^2}.$$

CHAPTER 2

Is A Large Bandwidth Mandatory to Maximally Exploit the Transmit Matched-Filter Structure?

Transmitted signals can be focused in time or space, *i.e.*, at a particular receiver location, by introducing a prefilter in the transmitter. Prefiltering also allows to contain inter-symbol and multiuser interference. Performance bounds in terms of signal-to-noise ratio gain, achievable by prefiltering, depend on both transmitted signal bandwidth and channel characteristics. This chapter analyzes the above dependence for transmit matched-filters, also known as time-reversal prefilters, and channels with multipath.

Theoretical results show that single-cluster channels verify a condition by which the gain is monotonic non-decreasing with bandwidth. Multi-cluster channels, such as those described by the Saleh-Valenzuela model, seem to follow a similar behavior, as suggested by simulation of the IEEE 802.15.3a channel model. As such, the transmit matched-filter is particularly suitable for signals with large bandwidths, as in acoustics and ultra-wideband communications.

The interplay between transmitted signals, that are under designer's control, and channel, that is set by Nature, has an impact on system performance bounds. For systems with prefiltering, the dependence of performance bounds in terms of signal-to-noise ratio (SNR) on both bandwidth and channel has not been specifically addressed. For particular contexts, such as ultra-wideband (UWB) communications, the general claim is that the larger the bandwidth the higher the SNR. The intuition is that, while bandwidth grows, the prefilter can take advantage of an increased fraction of the channel; this intuition, however, has not been thoroughly investigated.

The goal of this paper is to address the above issue by investigating whether increased bandwidth implies increased SNR, when the channel is affected by multipath. To this end, two different transceiver structures are considered in the analysis. In the first structure, a simple pulse, *i.e.*, a zero-excess bandwidth pulse as will be further defined, is transmitted. In the second structure, the same zero-excess

bandwidth pulse is filtered by a transmit matched-filter [44–46], that has an impulse response proportional to the time-reversed version of the channel impulse response. In acoustics [47,48], and UWB [49], this prefilter is known as time-reversal, although its first appearence may go back to [50, 51], where it was introduced as pre-Rake. These two structures are compared based on the SNR gain G, that is achievable by introducing prefiltering.

The paper is organized as follows: Section 4.2 contains the two reference system models and defines the system performance measure. Results are presented and discussed in Section 2.2. Section 6.3 contains the conclusion.

2.1 System Model and Performance Measure

The two transceiver structures under analysis are shown in Fig. 2.1. No Inter-Symbol Interference (ISI) is assumed, *e.g.* either one symbol only, or symbol sequences with symbols modulating waveforms with vanishing crosscorrelations at the receiver, are transmitted.

In Structure 1, the transmitted signal is $x_1(t) = b\psi(t)$, where $\psi(t)$ has unit energy and b is the symbol to be transmitted, with $\mathbb{E}\left[\,|b|^2\,\right] = \mathcal{E}$. The noiseless received signal $y_1(t)$ is, as a function of frequency, $\hat{y}_1(f) = b\hat{\psi}(f)\hat{c}(f)$, where $\hat{c}(f)$ is the channel transfer function.

In Structure 2, the transmitted signal is $x_2(t)$. As a function of frequency, one has $\hat{x}_2(f) = \hat{x}_1(f) \cdot \sqrt{\xi}\hat{c}^*(f) = b\hat{\psi}(f)\sqrt{\xi}\hat{c}^*(f)$, where ξ is a constant such that $x_1(t)$ and $x_2(t)$ have same energy $\mathcal{E}$. Structure 2 requires a perfect knowledge of the channel at the transmitter. The noiseless received signal $y_2(t)$ is as a function of frequency: $\hat{y}_2(f) = \sqrt{\xi} \cdot b\hat{\psi}(f)|\hat{c}(f)|^2$.

In the two structures, $n(t)$ is a White Gaussian Noise process with spectrum height $\sigma^2 = \mathcal{N}_0/2$, and the receiver is a matched-filter followed by a sampler. Based on the no-ISI hypothesis, the channel can be modeled as Additive White Gaussian Noise (AWGN). In the first structure, the received signal $r_1(t) = y_1(t) + n(t)$ is projected onto a waveform proportional to $y_1(t)$. In the second structure, $r_2(t) = y_2(t) + n(t)$ is projected onto a waveform proportional to $y_2(t)$. The projection of the generic received signal $r_k(t)$ onto $y_k(t)$, $k \in \{1, 2\}$, is:

$$\langle r_k, y_k\rangle = \langle \hat{r}_k, \hat{y}_k\rangle \triangleq \int_{\mathbb{R}} \hat{r}_k(f)\,\hat{y}_k^*(f)\,df = \langle \hat{y}_k, \hat{y}_k\rangle + \nu_k, \tag{2.1}$$

where ν_k is a real Gaussian random variable (r.v.) with $\mathrm{Var}[\nu_k] = \sigma^2\langle \hat{y}_k, \hat{y}_k\rangle = \sigma^2\|\hat{y}_k\|^2$. The SNR at the receiver, after the sampler, is:

$$\mathrm{SNR}_i \triangleq \frac{\mathbb{E}[\,|\langle \hat{y}_i, \hat{y}_i\rangle|^2\,]}{\mathrm{Var}[\nu_i]}. \tag{2.2}$$

With reference to Fig. 2.1, one has $\mathrm{SNR}_1 = (\mathcal{E}/\sigma^2) \cdot \|\hat{\psi}\hat{c}\|^2$ and $\mathrm{SNR}_2 = (\mathcal{E}/\sigma^2) \cdot \|\hat{\psi}|\hat{c}|^2\|^2/\|\hat{\psi}\hat{c}\|^2$.

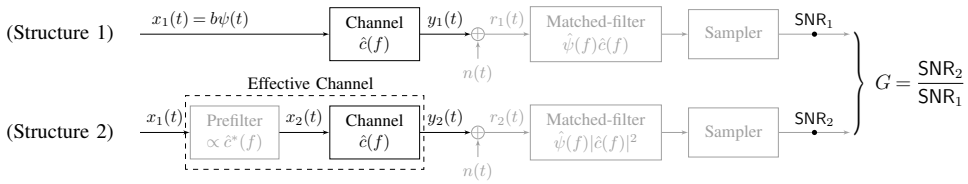


Figure 2.1: The two reference system models. Structure 1 corresponds to a traditional transmission with no prefiltering at the transmitter and matched-filtering at the receiver. Structure 2 corresponds to prefiltering at the transmitter with transmit matched-filter, and matched-filtering at the receiver. In both structures, $\psi(t)$ is a zero-excess bandwidth waveform. It is assumed that prefiltering does not alter energy, that is, $x_1(t)$ and $x_2(t)$ have same energy $\mathcal{E}$.

Gain G allows comparison of the two structures and is defined as follows:

$$G \triangleq \frac{\mathsf{SNR}_2}{\mathsf{SNR}_1} = \frac{||\hat{\psi}|\hat{c}|^2||^2}{||\hat{\psi}\hat{c}||^4}. \tag{2.3}$$

For the sake of simplicity, $\psi(t)$ is assumed as a zero-excess bandwidth waveform with band $[-W/2, W/2]$, *i.e.*, $\psi(t) = \sqrt{W}\,\mathrm{sinc}(\pi W t)$, where $\mathrm{sinc}(x) \triangleq (\sin x)/x$, for which SNR_1 and SNR_2 become:

$$\mathsf{SNR}_1 = \frac{\mathcal{E}}{\sigma^2} \cdot \frac{1}{W} \int_{-\frac{W}{2}}^{\frac{W}{2}} |\hat{c}(f)|^2 df, \tag{2.4}$$

$$\mathsf{SNR}_2 = \frac{\mathcal{E}}{\sigma^2} \cdot \int_{-\frac{W}{2}}^{\frac{W}{2}} |\hat{c}(f)|^4 df \Big/ \int_{-\frac{W}{2}}^{\frac{W}{2}} |\hat{c}(f)|^2 df, \tag{2.5}$$

and gain G is:

$$G = W \cdot \frac{\int_{-\frac{W}{2}}^{\frac{W}{2}} |\hat{c}(f)|^4 df}{\left[\int_{-\frac{W}{2}}^{\frac{W}{2}} |\hat{c}(f)|^2 df\right]^2} = \frac{W}{2} \cdot \frac{\int_{0}^{\frac{W}{2}} |\hat{c}(f)|^4 df}{\left[\int_{0}^{\frac{W}{2}} |\hat{c}(f)|^2 df\right]^2}. \tag{2.6}$$

Gain G also provides a hint on performance in the coded regime, as measured by maximal mutual information $I_k = \frac{1}{2}\ln(1 + \mathsf{SNR}_k)$, $k \in \{1, 2\}$, nats/channel use. Since the system is baseband, the no-ISI hypothesis holds, for example, by making the symbol period T_s greater than the channel delay spread T_d, by which $I_k = \frac{1}{T_s} \cdot \frac{1}{2}\ln(1 + \mathsf{SNR}_k)$ nats/s. Similarly to the SNR gain G, an *information gain* G_I can be defined as follows:

$$G_I \triangleq \frac{I_2}{I_1} = \frac{\ln(1 + \mathsf{SNR}_2)}{\ln(1 + \mathsf{SNR}_1)} = \frac{\ln(1 + G \cdot \mathsf{SNR}_1)}{\ln(1 + \mathsf{SNR}_1)}.$$

In the low-SNR regime ($\mathsf{SNR}_k \ll 1$, $k \in \{1, 2\}$) one has $G_I \approx \mathsf{SNR}_2/\mathsf{SNR}_1 = G$, hence G_I reduces to G. In the high-SNR regime ($\mathsf{SNR}_k \gg 1$) one has $G_I \approx$

$\ln(\mathsf{SNR}_2)/\ln(\mathsf{SNR}_1) = 1 + \ln(G)/\ln(\mathsf{SNR}_1)$, and for channels with bounded G, *e.g.* multipath channels, G_I reduces to unity as $\mathsf{SNR}_1 \to \infty$. G is, therefore, the most important parameter to be analyzed to determine performance of systems in both uncoded and coded regimes, under the no-ISI hypothesis.

2.2 System Analysis Based on Gain G

In this section, we find the lower bound of G and the asymptotic value of G as $W \to \infty$ for multipath channels (section 2.2.1), and derive a necessary and sufficient condition for G to be a monotonic non-decreasing function of W, also showing examples and counterexamples of channels verifying the condition (section 2.2.2).

2.2.1 Gain G Limit Values

1) Lower bound of G.

By applying the Cauchy-Schwarz inequality to the denominator of eq. (2.3), one has $\|\hat{\psi}\hat{c}\|^2 = \langle \hat{\psi}, \hat{\psi}\hat{c}^*\hat{c}\rangle \leq \|\hat{\psi}\| \cdot \|\hat{\psi}\hat{c}^*\hat{c}\|$. Therefore, $G \geq 1$, implying that Structure 2 in Fig. 2.1 cannot underperform Structure 1. Equality is achieved iff $\hat{\psi} \propto \hat{\psi}\hat{c}^*\hat{c}$, *i.e.* iff $|\hat{c}(f)|$ is constant for $f \in \operatorname{supp}\hat{\psi} \triangleq \{\nu \in \mathbb{R} : \hat{\psi}(\nu) \neq 0\}$, as can also be directly shown by methods of variational calculus. The minimum gain $G = 1$ is thus obtained when the magnitude of the channel transfer function is nonzero and flat, for frequency intervals where the amplitude of the spectrum of transmitted signals is nonzero. Typically, this is true when the channel is perfect (channel transfer function with constant amplitude) at least within the transmitted signal bandwidth. This condition easily holds for narrowband communications.

2) Asymptotic value of G as $W \to \infty$.

Consider the following multipath channel:

$$c(t) = \sum_{k \geq 0} \alpha(\tau_k)\delta(t - \tau_k), \tag{2.7}$$

where τ_k and $\alpha(\tau_k) \in \mathbb{R}$ are delay and amplitude of ray k. Eq. (2.7) also models realizations of channels with clusters. Supposing that a same interarrival time between two consecutive rays is not possible,[1] it can be shown that [52], as $W \to \infty$, the gain is:

$$G = 2 - \sum_{k \geq 0} \alpha(\tau_k)^4 \Big/ \left[\sum_{k \geq 0} \alpha(\tau_k)^2\right]^2, \tag{2.8}$$

and, therefore, $G < 2$ irrespective of channel amplitudes.

[1] This hypothesis often holds with probability one for random channels, since interarrival times are usually independent.

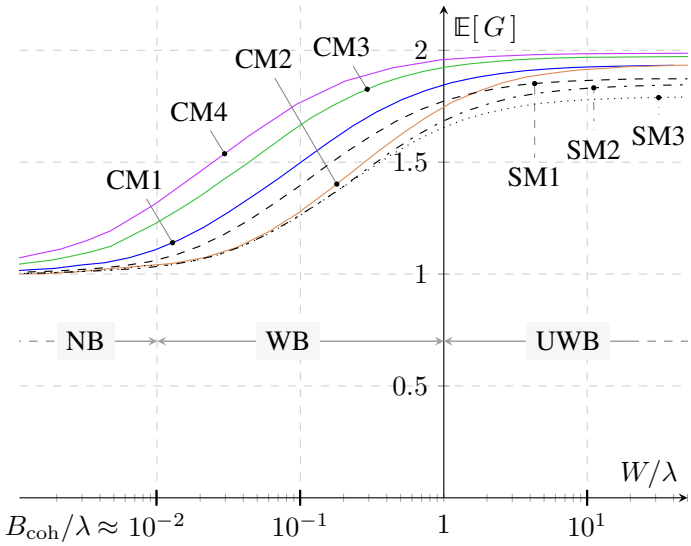


Figure 2.2: Average gain $\mathbb{E}[G]$ vs. W/λ for channel models CM1-CM4 of the IEEE 802.15.3a standard [53], and single-cluster models SM1-SM3, that refers to channel of eq. (2.7), with parameters $(\lambda, \gamma, \sigma_r)$ equal to $(2, 4, \frac{1}{2})$, $(2, 2, \frac{1}{4})$ and $(2, 2, \frac{1}{2})$, respectively, where $1/\lambda$ and γ are expressed in nanoseconds.

2.2.2 Slope of G vs. W

From above, G is bounded as follows: $1 = \lim_{W\to 0} G(W) \leq \lim_{W\to\infty} G(W) < 2$. In this section, the condition by which $G(W)$ is a monotonic non-decreasing function is derived. For the sake of generality, the condition is derived for random channels, but it can be applied to nonrandom channels as well. Suppose that $G(W)$ is random, with average $\mathbb{E}[G]$. The slope of $\mathbb{E}[G]$ is nonnegative iff:

$$\frac{d}{dw}\,\mathbb{E}[G] \geq 0, \quad \forall w > 0, \tag{2.9}$$

where $w \triangleq W/2$.

Figure 2.2 shows $\mathbb{E}[G]$ vs. W/λ obtained by means of Monte-Carlo simulations for UWB channels following the IEEE 802.15.3a standard model [53], where $1/\lambda$ is the average intra-cluster interarrival time between two rays, with values in the range $[0.4, 2]$ nanoseconds. Figure 2.2 suggests that channel models valid for very large bandwidths, up to several gigahertz, verify eq. (2.9).

Also shown in Fig. 2.2 is when a signal can be considered as narrowband, wideband or UWB, as a function of the statistical parameters of the channel it experiences. In a multipath channel, most of the power lies in $[0, T_d]$, *i.e.*, $[0, 1/B_{\text{coh}}]$, where B_{coh} is the *coherence bandwidth*. When $W < B_{\text{coh}}$, there is about only one resolved path. B_{coh} may serve, therefore, to define narrowband signals, the condition being $W < B_{\text{coh}}$. Similarly, λ may define UWB signals, the condition

being $W > \lambda$, for which a significant fraction of the multipath components can be resolved. For $W \in [B_{\text{coh}}, \lambda]$, the signal may be considered as wideband.

In general, a direct computation of $\mathbb{E}[\,G\,]$ is a formidable task, even for simple models that do not account for clustering. Therefore, a condition, that does not require the explicit knowledge of $\mathbb{E}[\,G\,]$, for eq. (2.9) to hold, is derived below.

Proposition 1. *Let $\hat{c}(f)$ be the frequency response of a realization of a random multipath channel, and denote by $S(f) = |\hat{c}(f)|^2$ its squared magnitude. Then $\mathbb{E}[\,G\,]$ is a monotonic non-decreasing function, i.e., condition of eq. (2.9) holds, iff:*

$$F(w) \triangleq \frac{1}{w^2} \iint_{[0,w]^2} \mathbb{E}\big[\,S(f_1)S(f_2)^2\,\big]\,df_1 df_2 + \frac{1}{w} \int_{[0,w]} \mathbb{E}\big[\,S(f)S(w)^2 - 2S(f)^2 S(w)\,\big]\,df \geq 0. \quad (2.10)$$

Proof. Interchange differentiation and expectation operators in eq. (2.9), and compute dG/dw; then discard the denominator, since it is always positive, and interchange the order of expectations and integrations. □

Note on eq. (2.10) that $F(w)$ is the derivative of $\mathbb{E}[\,G\,]$ up to a positive factor depending on $|\hat{c}(f)|^2$.

An example of a channel model that verifies eq. (2.10) is the "Single-cluster Model" (SM). This model is similar to the IEEE 802.15.3a-CM1 [53] when restricted to the first cluster. In particular, the SM channel impulse response is given by eq. (2.7), where $\{\tau_k\}_{k\geq 0}$ and $\{\alpha(\tau_k)\}_{k\geq 0}$ are random variables. Interarrival times $\{\tau_{k+1} - \tau_k\}_{k\geq 0}$ are independent and exponentially distributed with average $\mathbb{E}[\,\tau_{k+1} - \tau_k\,] = 1/\lambda$. Conditioned on τ_k, the distribution of the path amplitude $\alpha(\tau_k)$ is an even function, and $|\alpha(\tau_k)|$ follows a log-normal distribution. Even conditional moments are $\mathbb{E}\big[\,\alpha(\tau)^{2n}\,\big|\,\tau\,\big] = (\Omega_0\, e^{-\tau/\gamma})^n e^{2n(n-1)\sigma_r^2}$ for $n \geq 1$, where $\gamma > 0$ is the ray decay factor, σ_r is the standard deviation of the Gaussian r.v. generating the log-normal r.v., and Ω_0 is the variance of $\alpha(0)$. Odd conditional moments are nil. In order to compute $\mathbb{E}\big[\,S(f_1)S(f_2)^2\,\big]$, the approach proposed in [54] for $\mathbb{E}[\,S(f_1)S(f_2)\,]$ is generalized here:

$$\mathbb{E}\left[\prod_{i=1}^{M} S(f_i)\right] = \mathbb{E}\left\{ \sum_{\mu=1}^{2M} \sum_{k_\mu \geq 0} \mathbb{E}\left[\prod_{m=1}^{2M} \alpha(\tau_{k_m}) \,\middle|\, \{\tau_{k_n}\}_{n=1}^{2M}\right] \times \exp\left[-j2\pi \sum_{\ell=1}^{M} f_\ell (\tau_{k_{2\ell-1}} - \tau_{k_{2\ell}})\right] \right\}, \quad (2.11)$$

that, for $M = 3$ and $f_3 \to f_2$, provides $\mathbb{E}\big[\,S(f_1)S(f_2)^2\,\big]$. In the right-hand side of eq. (2.11), the inner expectation is over amplitudes conditioned on delays $\{\tau_{k_m}\}_{m=1}^{2M}$, and the outer expectation is over delays only. Inner expectations are nonzero iff the set $\{\alpha(\tau_{k_1}), \ldots, \alpha(\tau_{k_{2M}})\}$ can be partitioned into subsets with even cardinality. With probability one, the set of indices $\{k_1, \ldots, k_{2M}\}$ can be equivalently partitioned. Logical conditions ensuring the partition can be derived. In the

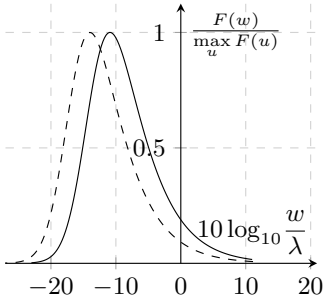


Figure 2.3: $F(w)/\max_u F(u)$ vs. w/λ (dB) for the single-cluster channel model and different $(\lambda, \gamma, \sigma_r)$ configurations. Solid curve: $(2, 2, \frac{1}{2})$, $(2, 2, \frac{1}{4})$. Dashed curve: $(4, 2, \frac{1}{2})$, $(2, 4, \frac{1}{2})$. $1/\lambda$ and γ are expressed in nanoseconds. $F(w)$ is positive, showing that eq. (2.10) is verified.

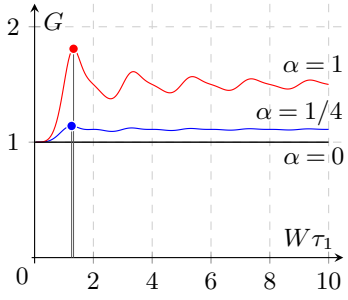


Figure 2.4: Gain G vs. $W\tau_1$ for the two-paths channel $c(t) = \delta(t) + \alpha\delta(t - \tau_1)$ with $\alpha \in \{0, \frac{1}{4}, 1\}$. For $\alpha \neq 0$, G is not monotonic non-decreasing: there exist local maxima and minima, and a global maximum (see dots) with respect to $W\tau_1$ (indicated by vertical lines) depending on α.

simple case of $M = 2$, there are four logical conditions, *i.e.*, $k_1 = k_2 = k_3 = k_4$; $k_1 = k_2 \neq k_3 = k_4$; $k_1 = k_3 \neq k_2 = k_4$; $k_1 = k_4 \neq k_2 = k_3$ (see Appendix 2.A for a detailed derivation). For $M = 3$, there are 31 logical conditions, that are not reported here for brevity. Once the conditions are set, the closed form expression of $\mathbb{E}\left[S(f_1)S(f_2)^2\right]$ can be derived, since each inner expectation is equal to an even conditional moment with appropriate order. Figure 2.3 represents $F(w)$, normalized to its maximum, for different γ and σ_r values, and shows that $F(w)$ is positive for any relevant w.

A channel that does not verify eq. (2.10) is as simple as a two-paths nonrandom channel with impulse response $c(t) = \delta(t) + \alpha\delta(t - \tau_1)$, $\alpha \in \mathbb{R}$, $\tau_1 > 0$. Since the channel is nonrandom, $\mathbb{E}[G] = G$. The channel frequency response squared magnitude is:

$$S(f) = |\hat{c}(f)|^2 = 1 + \alpha^2 + 2\alpha\cos(2\pi f\tau_1). \tag{2.12}$$

The limit case $\alpha = 0$ reduces to the channel with constant spectrum (c.f. Sec. 2.2.1). The value of G as $W\tau_1 \to \infty$ is $1 + 2\alpha^2/(1 + \alpha^2)^2$, that is maximum for $|\alpha| = 1$, for which $G = 3/2$. Figure 2.4 shows G vs. $W\tau_1$ for $\alpha \in \{0, \frac{1}{4}, 1\}$. Note that, for any fixed α, there exists an optimum $W\tau_1$ that maximizes G (see dots on figure for $\alpha = 1/4$ and $\alpha = 1$). For example, for $\alpha = 1$, the maximum gain is reached for $W\tau_1 \approx 1.32465$, and is equal to $G \approx 1.80905$. Figure 2.4 shows that G is not monotonic non-decreasing with W for any α.

2.3 Future work

Three assumptions of the analysis thus far were: 1) no-ISI; 2) transmit matched filter precoding; 3) multipath channel. In this section, we set the basis for a future,

general analysis, where all three assumptions are removed, and thus accounting for generic FIR channels, prefiltering schemes, and symbol rate $1/T_s$. In particular, the latter is not restricted to be lower than, or equal to, the coherence bandwidth, given by $B_c = 1/T_d$.

Capacity of a deterministic channel with the following discrete representation:

$$r_i = \sum_{\ell} h_\ell b_{i-\ell} + n_i, \quad i \in \mathbb{Z}, \tag{2.13}$$

is given by [55, 56]:

$$C = \int_{-\pi}^{\pi} \frac{d\theta}{2\pi} \cdot \frac{1}{2} \ln\left(1 + \frac{S_b(e^{j\theta})|H(e^{j\theta})|^2}{S_n(e^{j\theta})}\right) \quad \text{nats/channel use},$$

where $S_b(e^{j\theta})$ and $S_n(e^{j\theta})$ are the spectra of the transmitted symbol sequence and noise, respectively, and $|H(e^{j\theta})|^2$ is the squared magnitude of the DTFT of the sequence (h_ℓ). When the channel is random and unknown at the encoder, capacity is averaged over the channel (see *e.g.* [39]):

$$C = \int_{-\pi}^{\pi} \frac{d\theta}{2\pi} \cdot \mathbb{E}\left[\frac{1}{2} \ln\left(1 + \frac{S_b(e^{j\theta})|H(e^{j\theta})|^2}{S_n(e^{j\theta})}\right)\right] \quad \text{nats/channel use}.$$

The above discrete representation can be derived from a continuous-time description as follows. Let the received signal be:

$$r(t) = \sum_m b_m x(t - m/W) * c(t) + n(t),$$

where $n(t)$ is an AWGN with spectral height $\mathcal{N}_0/2$, $c(t)$ is the channel impulse response, and $x(t)$ is the unit energy waveform associated with the m-th symbol. In the present paper, $x(t)$ can be either a zero-excess bandwidth waveform $\psi(t)$ with band $[-W/2, W/2]$, or the prefiltered waveform $\alpha\psi * c^{\leftarrow}(t)$, where $c^{\leftarrow}(t) = c(-t)$ and $\alpha = 1/\|\psi * c^{\leftarrow}(t)\| = 1/\|\psi * c(t)\|$. In both cases, the bandwidth of $x(t)$ is W, and the projection onto $\{\psi(t - i/W) \colon i \in \mathbb{Z}\}$ provides a sufficient statistic for optimum inference of (b_m):

$$\begin{aligned}
r_i &\triangleq \langle r(t), \psi(t - i/W)\rangle \\
&= \left\langle \sum_m b_m x(t - m/W) * c(t) + n(t), \psi(t - i/W) \right\rangle \\
&= \sum_m b_m \langle x(t - m/W) * c(t), \psi(t - i/W)\rangle + \langle n(t), \psi(t - i/W)\rangle \\
&= \sum_m b_m \left\langle \sum_j x_j \psi(t - (m + j)/W) * c(t), \psi(t - i/W) \right\rangle + n_i \\
&= \sum_m b_m \sum_j x_j \left\langle \sum_\ell c_\ell \psi(t - (m + j + \ell)/W), \psi(t - i/W) \right\rangle + n_i
\end{aligned}$$

$$= \sum_m b_m \sum_j x_j \sum_\ell c_\ell \delta_{m+j+\ell,i} + n_i$$
$$= \sum_m b_m \sum_j x_j c_{i-(m+j)} + n_i,$$

that has the form (2.13) as $h_i = \sum_j x_j c_{i-j}$, hence:

$$C = \int_{-\pi}^{\pi} \frac{d\theta}{2\pi} \cdot \mathbb{E}\left[\frac{1}{2} \ln\left(1 + \frac{S_b(e^{j\theta})|X(e^{j\theta})|^2 |C(e^{j\theta})|^2}{\mathcal{N}_0/2}\right)\right] \quad \text{nats/channel use,}$$

where the expectation is taken over the channel spectrum, and $S_n(e^{j\theta}) = \mathcal{N}_0/2$ since $n_i \sim \mathcal{N}(0, \mathcal{N}_0/2)$. The channel is used every $1/W$ seconds, therefore the capacity in nats/s is:

$$C = W \int_{-\pi}^{\pi} \frac{d\theta}{2\pi} \cdot \mathbb{E}\left[\frac{1}{2} \ln\left(1 + \frac{S_b(e^{j\theta})|X(e^{j\theta})|^2 |C(e^{j\theta})|^2}{\mathcal{N}_0/2}\right)\right] \quad \text{nats/s.} \qquad (2.14)$$

For Structure 1, one has $x_j = \delta_j$, $\mathbb{E}\left[b_m^2\right] = \mathcal{E}$, and the capacity is given by:

$$C_1 = \int_{-\pi}^{\pi} \frac{d\theta}{2\pi} \cdot \mathbb{E}\left[\frac{1}{2} \ln\left(1 + \frac{S_b(e^{j\theta})|C(e^{j\theta})|^2}{\mathcal{N}_0/2}\right)\right] \quad \text{nats/channel use.}$$

For Structure 2, one has $X(e^{j\theta}) = \alpha C(e^{j\theta})^*$, and the capacity is given by:

$$C_2 = \int_{-\pi}^{\pi} \frac{d\theta}{2\pi} \cdot \mathbb{E}\left[\frac{1}{2} \ln\left(1 + \frac{S_b(e^{j\theta})\alpha^2|C(e^{j\theta})|^4}{\mathcal{N}_0/2}\right)\right] \quad \text{nats/channel use.}$$

It can be shown that:

$$\alpha^2 = \frac{1}{\|\psi * c(t)\|^2} = \left[\int_{-\pi}^{\pi} \frac{d\theta}{2\pi} |C(e^{j\theta})|^2\right]^{-1}.$$

A *capacity gain* can be defined as follows:

$$G_C = C_2/C_1.$$

A careful analysis of the capacity gain in terms of $X(e^{j\theta})$ (prefiltering) and $C(e^{j\theta})$ (channel), which also depends on the two parameters W and T_d, represents a future line of investigation.

Let see how this reduces to the previous analysis in absence of ISI. Under the hypothesis of vanishing crosscorrelation of waveforms modulating symbols, the channel becomes an AWGN channel with multipath. This assumption holds, for example, by transmitting symbols farther apart than the channel delay spread T_d, *i.e.*, the symbol period T_s is such that $T_s > T_d$, *i.e.*, such that:

$$r(t) = \sum_m b_m x(t - mT_s) * c(t) + n(t).$$

There is, indeed, no ISI, and the channel can be represented (for $m=0$) by:

$$r_i = b_0 \sum_j x_j c_{i-j} + n_i.$$

Hereinafter, we will write b instead of b_0 for notation simplicity.

Equivalently, optimum inference of b is based on:

$$z = \langle bx * c(t) + n(t), x * c(t) \rangle = b\|x * c(t)\|^2 + \nu, \quad \nu \sim \mathcal{N}(0, \|x * c(t)\|^2 \mathcal{N}_0/2).$$

For Structure 1, one has:

$$z = b\|\psi * c(t)\|^2 + \nu, \quad \nu \sim \mathcal{N}(0, \|\psi * c(t)\|^2 \mathcal{N}_0/2),$$

and the maximum mutual information $\max_{P_b} I(b;z)$ is achieved with Gaussian inputs and assumes the well-known AWGN capacity expression:

$$I_1 = \frac{1}{2}\ln(1+\mathsf{SNR}_1) \quad \text{nats/channel use}, \tag{2.15}$$

where here the channel is used every T_s seconds, and:

$$\mathsf{SNR}_1 = \|\psi * c(t)\|^2 \cdot \frac{\mathcal{E}}{\mathcal{N}_0/2}$$

is that analyzed in the Letter, being $\mathcal{E} = \mathrm{Var}[b]$.

Similarly, for Structure 2, one has:

$$I_2 = \frac{1}{2}\ln(1+\mathsf{SNR}_2) \quad \text{nats/channel use}, \tag{2.16}$$

where:

$$\mathsf{SNR}_2 = \frac{\|\psi * c * c^{\leftarrow}(t)\|^2}{\|\psi * c(t)\|^2} \cdot \frac{\mathcal{E}}{\mathcal{N}_0/2} = G \cdot \mathsf{SNR}_1,$$

being $G = \mathsf{SNR}_2/\mathsf{SNR}_1$ the SNR gain analyzed in the Letter.

Expressed in terms of nats/s, we have:

$$I_k = \frac{1}{T_s} \cdot \frac{1}{2}\ln(1+\mathsf{SNR}_k) \ \text{ nats/s}, \qquad k = 1, 2.$$

We adopted the notation I_1 and I_2 in place of C_1 and C_2, respectively, to stress that the above are mutual informations rather than capacities, that are derived under the constraint of a symbol period T_s larger than the minimum, *i.e.*, $T_s > 1/W$.

Similarly to the capacity gain defined above, we can define the following *(mutual) information gain*:

$$G_I = \frac{I_2}{I_1} = \frac{\ln(1+\mathsf{SNR}_2)}{\ln(1+\mathsf{SNR}_1)} = \frac{\ln(1+G\cdot\mathsf{SNR}_1)}{\ln(1+\mathsf{SNR}_1)}.$$

In the low-SNR regime, that is, when $\mathsf{SNR}_k \ll 1$, $k = 1,2$, one has:

$$G_I \approx \frac{G \cdot \mathsf{SNR}_1}{\mathsf{SNR}_1} = G,$$

and therefore the information gain is equal to the SNR gain G studied in the paper.

In the high-SNR regime, that is, when $\mathsf{SNR}_k \gg 1$, one has:

$$G_I \approx \frac{\ln(\mathsf{SNR}_2)}{\ln(\mathsf{SNR}_1)} = 1 + \frac{\ln(G)}{\ln(\mathsf{SNR}_1)}.$$

In particular, for those channels (*e.g.* multipath channels) having bounded G, $G_I \longrightarrow 1$ for large SNR.

2.4 Conclusion

This chapter investigated whether a large bandwidth is mandatory to maximally exploit the potential benefits of the transmit matched-filter, in the absence of ISI. To this end, two transceiver structures, with and without prefiltering, were compared based on a parameter indicating the SNR gain G achieved by introducing the prefilter. Performance depended on both transmitted signal bandwidth W and channel frequency response squared magnitude. Limit values of G when the channel is affected by multipath were derived, for $W \to 0$ and $W \to \infty$, and it was proven that G for $W \to 0$ reaches the minimum, and is equal to one. A condition for G to be a monotonic non-decreasing function of W was also derived for generic random channels. This condition was then specified for single-cluster multipath channel models, and verified in the particular case of exponentially distributed interarrival times and absolute path amplitudes following a log-normal distribution. Simulation experiments of channels following the IEEE 802.15.3a standard channel model showed a monotonic non-decreasing gain in W for the four models of the standard, suggesting that G may also be monotonic non-decreasing for channels with clusters. However, the analysis of a channel with two paths only was shown to be a very simple counterexample, for which the gain was not a monotonic function of W.

In conclusion, a large bandwidth is not mandatory to achieve a maximum gain with transmit matched filter, although this seems to be the case for models describing realistic channels, such as the IEEE 802.15.3a model.

Results were obtained in terms of SNR gain G. As discussed in Section II, an information gain G_I can be defined, similarly to G, as the ratio between the mutual information achieved with and without prefiltering. It was proven that, under the no-ISI hypothesis, G_I essentially reduces to G. This would be no longer true if the no-ISI hypothesis were removed, in which case an achievable "capacity gain" should be analyzed. This will be the goal of future work together with two generalization: 1) channel models that expand beyond multipath; 2) prefilter structures that expand beyond transmit matched-filters. The former can be analyzed by specifying fairly general channel models and studying when eq. (2.10) holds with respect to

the parameters specified in the channel model. The latter can be investigated by specifying a prefilter that is proportional to $\hat{c}^*(f)/(|\hat{c}(f)|^2+\lambda)$, that encompasses both transmit zero-forcing and transmit MMSE, for particular values of λ.

The setting used in this chapter, therefore, may serve towards a deeper understanding of prefiltering effects beyond the transmit matched-filter case in multipath channels.

Appendix

2.A Derivation of the $M = 2$ case of eq. (2.11).

The goal is to evaluate $\mathbb{E}[\,S(f_1)S(f_2)\,]$ when the channel is:

$$c(t) = \sum_{k\geq 0} \alpha(\tau_k)\delta(t-\tau_k).$$

Therefore:

$$S(f) = \left| \sum_{k\geq 0} \alpha(\tau_k) e^{-j2\pi f \tau_k} \right|^2 = \sum_{k_1\geq 0}\sum_{k_2\geq 0} \alpha(\tau_{k_1})\alpha(\tau_{k_2}) e^{-j2\pi f(\tau_{k_1}-\tau_{k_2})}.$$

Note that, in the previous expression, both $\{\tau_k\}$ and $\{\alpha(\tau_k)\}$ are random. Then:

$$S(f_1)S(f_2) = \sum_{k_1\geq 0}\sum_{k_2\geq 0}\sum_{k_3\geq 0}\sum_{k_4\geq 0} \alpha(\tau_{k_1})\alpha(\tau_{k_2})\alpha(\tau_{k_3})\alpha(\tau_{k_4}) \\ e^{-j2\pi f_1(\tau_{k_1}-\tau_{k_2})} e^{-j2\pi f_2(\tau_{k_3}-\tau_{k_4})}.$$

Since the expectation $\mathbb{E}[\,S(f_1)S(f_2)\,]$ is on both $\{\tau_k\}$ and $\{\alpha(\tau_k)\}$, but conditional moments of amplitudes are provided for fixed delays, we may split the expectation in the previous expression by conditioning first on delays $\{\tau_k\}$:

$$\mathbb{E}[\,S(f_1)S(f_2)\,] = \mathbb{E}\Big\{ \sum_{k_1\geq 0}\sum_{k_2\geq 0}\sum_{k_3\geq 0}\sum_{k_4\geq 0} \mathbb{E}\Big[\alpha(\tau_{k_1})\alpha(\tau_{k_2})\alpha(\tau_{k_3})\alpha(\tau_{k_4}) \,\Big|\, \{\tau_i\}_{i\geq 0} \Big] \\ e^{-j2\pi f_1(\tau_{k_1}-\tau_{k_2})} e^{-j2\pi f_2(\tau_{k_3}-\tau_{k_4})} \Big\}. \quad (2.17)$$

Since amplitudes for different delays are uncorrelated, the inner expectation is nonzero when $\alpha(\tau_{k_1})\alpha(\tau_{k_2})\alpha(\tau_{k_3})\alpha(\tau_{k_4}) = \alpha(\tau_{k'_1})^{n_1}\alpha(\tau_{k'_2})^{n_2}$ with both n_1 and n_2 even numbers, and $k'_i \in \{k_1, k_2, k_3, k_4\}$. For the case under analysis, the conditions are:

- $k_1 = k_2 = k_3 = k_4$;
- $k_1 = k_2$, $k_3 = k_4$, and $k_1 \neq k_3$;
- $k_1 = k_3$, $k_2 = k_4$, and $k_1 \neq k_2$;
- $k_1 = k_4$, $k_2 = k_3$, and $k_1 \neq k_2$.

Therefore, explicitly, eq. (2.17) becomes:

$$\begin{aligned}\mathbb{E}[S(f_1)S(f_2)] = &\mathbb{E}\left[\sum_{k_1\geq 0}\mathbb{E}\Big[\alpha(\tau_{k_1})^4\,\Big|\,\{\tau_i\}_{i\geq 0}\Big]\right]\\ &+\mathbb{E}\left[\sum_{k_1\geq 0}\sum_{k_3\geq 0}\mathbb{E}\Big[\alpha(\tau_{k_1})^2\alpha(\tau_{k_3})^2\,\Big|\,\{\tau_i\}_{i\geq 0}\Big]\right]\\ &+\mathbb{E}\left[\sum_{k_1\geq 0}\sum_{k_2\geq 0}\mathbb{E}\Big[\alpha(\tau_{k_1})^2\alpha(\tau_{k_2})^2\,\Big|\,\{\tau_i\}_{i\geq 0}\Big]\,e^{-j2\pi f_1(\tau_{k_1}-\tau_{k_2})}e^{-j2\pi f_2(\tau_{k_1}-\tau_{k_2})}\right]\\ &+\mathbb{E}\left[\sum_{k_1\geq 0}\sum_{k_2\geq 0}\mathbb{E}\Big[\alpha(\tau_{k_1})^2\alpha(\tau_{k_2})^2\,\Big|\,\{\tau_i\}_{i\geq 0}\Big]\,e^{-j2\pi f_1(\tau_{k_1}-\tau_{k_2})}e^{-j2\pi f_2(\tau_{k_2}-\tau_{k_1})}\right].\end{aligned} \tag{2.18}$$

Since, in the Letter, we assumed that:

$$\mathbb{E}\big[\alpha(t)^{2n}\,\big|\,t\big] = (\Omega_0\,e^{-t/\gamma})^n\cdot e^{2n(n-1)\sigma_r^2},\quad n\geq 1,$$

inner expectations are:

$$\mathbb{E}\Big[\alpha(\tau_{k_1})^4\,\Big|\,\tau_{k_1}\Big] = (\Omega_0\,e^{-\tau_{k_1}/\gamma})^2\cdot e^{4\sigma_r^2},$$

and:

$$\begin{aligned}\mathbb{E}\Big[\alpha(\tau_{k_1})^2\alpha(\tau_{k_3})^2\,\Big|\,\{\tau_{k_1},\tau_{k_3}\},\tau_{k_1}\neq\tau_{k_3}\Big] &= \mathbb{E}\Big[\alpha(\tau_{k_1})^2\,\Big|\,\tau_{k_1}\Big]\,\mathbb{E}\Big[\alpha(\tau_{k_3})^2\,\Big|\,\tau_{k_3}\Big]\\ &= (\Omega_0\,e^{-\tau_{k_1}/\gamma})(\Omega_0\,e^{-\tau_{k_3}/\gamma}).\end{aligned}$$

Therefore, eq. (2.18) becomes:

$$\mathbb{E}[S(f_1)S(f_2)] \tag{2.19}$$

$$\begin{aligned}&= \mathbb{E}\left[\sum_{k_1\geq 0}(\Omega_0\,e^{-\tau_{k_1}/\gamma})^2\cdot e^{4\sigma_r^2}\right]\\ &+\mathbb{E}\left[\sum_{k_1\geq 0}\sum_{k_3\geq 0}(\Omega_0\,e^{-\tau_{k_1}/\gamma})(\Omega_0\,e^{-\tau_{k_3}/\gamma})\right]\\ &+\mathbb{E}\left[\sum_{k_1\geq 0}\sum_{k_2\geq 0}(\Omega_0\,e^{-\tau_{k_1}/\gamma})(\Omega_0\,e^{-\tau_{k_2}/\gamma})e^{-j2\pi f_1(\tau_{k_1}-\tau_{k_2})}e^{-j2\pi f_2(\tau_{k_1}-\tau_{k_2})}\right]\\ &+\mathbb{E}\left[\sum_{k_1\geq 0}\sum_{k_2\geq 0}(\Omega_0\,e^{-\tau_{k_1}/\gamma})(\Omega_0\,e^{-\tau_{k_2}/\gamma})e^{-j2\pi f_1(\tau_{k_1}-\tau_{k_2})}e^{-j2\pi f_2(\tau_{k_2}-\tau_{k_1})}\right].\end{aligned} \tag{2.20}$$

Now expectations are only over delays. Assuming, as in the Letter, Poisson arrivals, *i.e.*, $\mathbb{E}[\tau_{k+1}-\tau_k]=1/\lambda$, $k\geq 0$, and the possible presence of a line-of-sight (LOS) component, it results:

$$\begin{aligned}
&\mathbb{E}[S(f_1)S(f_2)] \qquad (2.21)\\
&= \int_0^\infty (\Omega_0\, e^{-\tau/\gamma})^2 \cdot e^{4\sigma_r^2}(\lambda+\delta(\tau)I_{\mathsf{LOS}})d\tau\\
&+ \int_0^\infty \int_0^\infty (\Omega_0\, e^{-\tau_1/\gamma})(\Omega_0\, e^{-\tau_3/\gamma})(\lambda+\delta(\tau)I_{\mathsf{LOS}})d\tau_1(\lambda+\delta(\tau)I_{\mathsf{LOS}})d\tau_3\\
&+ \int_0^\infty \int_0^\infty (\Omega_0\, e^{-\tau_1/\gamma})(\Omega_0\, e^{-\tau_2/\gamma})e^{-j2\pi f_1(\tau_1-\tau_2)}e^{-j2\pi f_2(\tau_1-\tau_2)}\\
&\qquad\qquad (\lambda+\delta(\tau)I_{\mathsf{LOS}})d\tau_1(\lambda+\delta(\tau)I_{\mathsf{LOS}})d\tau_2\\
&+ \int_0^\infty \int_0^\infty (\Omega_0\, e^{-\tau_1/\gamma})(\Omega_0\, e^{-\tau_2/\gamma})e^{-j2\pi f_1(\tau_1-\tau_2)}e^{-j2\pi f_2(\tau_2-\tau_1)}\\
&\qquad\qquad (\lambda+\delta(\tau)I_{\mathsf{LOS}})d\tau_1(\lambda+\delta(\tau)I_{\mathsf{LOS}})d\tau_2, \qquad (2.22)
\end{aligned}$$

when $I_{\mathsf{LOS}}=1$ when the line-of-sight component is present, and $I_{\mathsf{LOS}}=0$ otherwise.

These integrals can be solved in closed form, providing the following result:

$$\begin{aligned}
\mathbb{E}[\,S(f_1)S(f_2)\,] = {}& \Omega^2\frac{1}{2}e^{4\sigma_r^2}(\gamma\lambda+2I_{\mathsf{LOS}})\\
&+\Omega^2(\gamma\lambda+I_{\mathsf{LOS}})^2\\
&+\frac{\Omega^2\left\{(\gamma\lambda+I_{\mathsf{LOS}})^2+4\pi^2\gamma^2\,(f_1+f_2)^{\,2}I_{\mathsf{LOS}}\right\}}{1+4\pi^2\gamma^2\,(f_1+f_2)^{\,2}}\\
&+\frac{\Omega^2\left\{(\gamma\lambda+I_{\mathsf{LOS}})^2+4\pi^2\gamma^2\,(f_1-f_2)^{\,2}I_{\mathsf{LOS}}\right\}}{1+4\pi^2\gamma^2\,(f_1-f_2)^{\,2}}. \qquad (2.23)
\end{aligned}$$

CHAPTER 3

Some results on Time Reversal *vs.* All-Rake Transceivers in Multiple Access Channels

Time reversal, that is prefiltering of transmitted signals with time reversed channel impulse responses, may be used in single user communications in order to move complexity from the receiver to the transmitter, and in multiuser communications to also modify statistical properties of multiuser interference. Imperfect channel estimation may, however, affect pre- *vs.* post- filtering schemes in a different way. This issue is the object of this chapter; Robustness of time reversal (TR) *vs.* All-Rake (AR) transceivers, in multiple access communications, with respect to channel estimation errors, is investigated. Results of performance analyses in terms of symbol error probability and spectral efficiency when the receiver is structured either by a bank of matched filters or by 1Rake, followed by independent decoders, indicate that AR is slightly more robust than time reversal but requires in practice more complex channel estimation procedures since all channels must be simultaneously inferred in the multiuser communication setting.

Field equivalence principles [57–60] state that the radiated field within a volume V, with boundary ∂V, enclosing a source, can be computed by considering, in place of actual source, an infinity of equivalent virtual sources placed on ∂V.

Suppose the source S_O is pointwise, impulsive, and located in a point $O \in V$. The electromagnetic problem of finding radiated field in V can be solved based on the Green function.

From a communication perspective, knowing the channel at all points of ∂V would allow, in principle, to understand the nature of a source S_O, that is, the location of a pointwise source within volume V, that radiated the field observed on ∂V. Sensing the channel on ∂V would require a multiantenna system and a perfect knowledge of impulse responses of channels between O and all points on ∂V.

Time reversal is a technique that takes advantage of the above physical phenomenon and that was also proposed in acoustics [47, 48, 61]. By prefiltering transmissions with a scaled version of the channel impulse response, reversed in

time, allows simplification of receiver design, since the channel is compensated by precoding. Time reversal also focuses signals in space, given that there is only one "correct" location of the receiver that experiences the specific "time reversed" channel impulse response.

Pioneering work on single-antenna time reversal spread-spectrum communications dates back to the nineties, where the time reversal pre-filter was named *pre-Rake* [50,51]. The basic idea was to pre-filter the transmitted pulse with the channel impulse response reversed in time, therefore matching the transmitted signal with the subsequent channel.

Precoding techniques for multiuser spread-spectrum systems were developed along similar lines of receive filters: transmit Zero-Forcing (ZF) [62], that attempts to pre-equalize the channel by flattening the effective channel formed by the cascade of the pre-filter and the actual channel, is optimum in the high-SNR regime; transmit matched-filter (MF), that has been recognized to be equivalent to the pre-Rake filter in [44], that conversely is optimum in the low-SNR regime; and finally, transmit MMSE (Wiener) filter minimizing the SINR was derived in [46] following previous attempts [63, 64].

In recent years, along with the fast developing of narrowband MIMO systems, pre-coding techniques using multiple antennas at the transmitter were thoroughly studied (for a complete overview on MIMO precoding see [65]). Since the mathematical formulation of multiuser spread-spectrum is very close to that of MIMO communications (see [66] for an overview of this analogy), MIMO linear precoders can be derived along similar techniques.

Time reversal was proposed in connection to UWB communications in [49], that also addressed equalization through an MMSE receiver. In [67], early experimental data, showing the feasibility of time reversal, were collected. Following, experimental investigations on multiple-antenna systems with time reversal [68–70], and performance analyses [71], were also pursued. In [72, 73], compensation for pulse distortion in connection to time reversal was investigated. In [74], the trade-off between the complexity of transmitter *vs.* receiver in terms of number of paths was analyzed. In [75], the insensitivity of time reversal to the lack of correlation between channels in a MISO system was investigated. Finally, in [76], the effect of time reversal on statistical properties of multiuser interference in communication *vs.* positioning was explored.

The above investigations were all carried out based on the hypothesis of perfect channel estimation. This hypothesis, however, is strong, since it is unrealistic, irrespectively of whether channel estimation is performed at the transmitter or at the receiver. While previous works addressed the comparison of pre- and post-channel filtering, the problem remains of realistic imperfect channel estimation and of how this affects performance for time reversed *vs.* receiver-based channel estimation schemes.

This chapter addresses the above problem, by comparing single-antenna systems using time reversal, against receiver-based equalization schemes such as the exemplary case of an AR receiver.

The adopted network model considers multiple access by K user terminals (UTs) communicating to one basestation (BS), where both UTs and the BS have one antenna only, and communication between each UT and BS adopts ultra-wideband, impulse-radio signaling.

Performance comparison of TR *vs.* AR transceivers will be carried out in terms of effect of imperfect channel state information (CSI) on symbol error probability of a generic information-bearing symbol for a given UT (see [77] for a work on a close topic regarding CDMA systems). The analysis will further explore robustness of TR *vs.* AR, by finding the maximum achievable rate for the uplink channel. Finally, the maximum information rate, that takes into account channel estimation overhead, will be explored.

The chapter is organized as follows: Section 3.1 contains the system model; Section 3.2 is devoted to the performance analysis in terms of symbol error probability; Section 3.3 contains results of comparison in terms of uplink rate of the network; Section 3.4 suggests future works, and conclusions are drawn in Section 4.4.

3.1 Reference Model

3.1.1 Network Model

A multiple access channel where K independent sources transmit information-bearing symbols to a common sink is considered (uplink communication channel). Borrowing the terminology from the cellular network field, sources of information are called user terminals (UTs) and the sink is called basestation (BS). However, UTs and BS are intendend to designate more than what the name implies. For example, in a typical WLAN, a BS is a fixed (*e.g.* desktop) or mobile receiver (*e.g.* tablet, laptop, mobile phone) and UTs are peripherals or other fixed *vs.* mobile devices. Figure 3.1 shows the adopted network model.

A generic UT_k transmits data, that is encoded into a sequence of information-bearing symbols $\{b_k[m]\colon m\in\mathbb{Z}\}$. This set of symbols is partitioned into blocks of n symbols each, $\{\boldsymbol{b}_k[i]\colon i\in\mathbb{Z}\}$ where $\boldsymbol{b}_k[i]=(b_k[in],\ldots,b_k[(i+1)n-1])^{\mathsf{T}}$. The

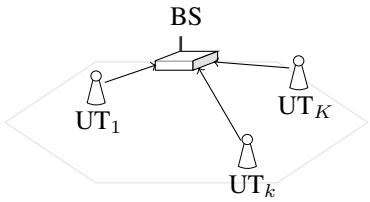


Figure 3.1: Network model: several user terminals (UTs) transmit information-bearing symbols to a common sink, *i.e.*, basestation (BS).

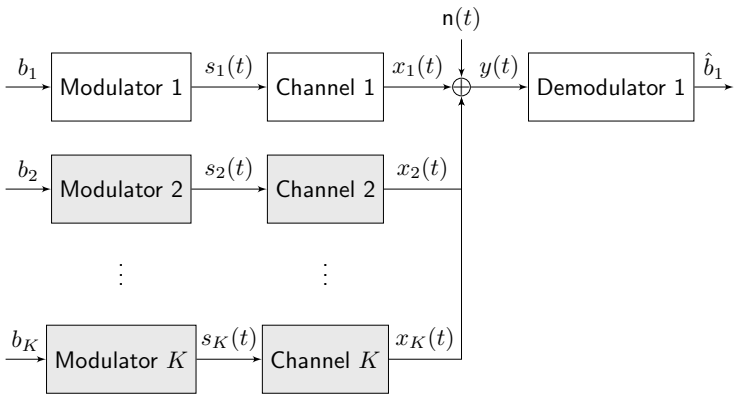


Figure 3.2: System model. The transmitter is formed by K modulators. Radiated signals are a eted by K di erent channels and white gaussian noise $n(t)$ at the receiver. Receveir consists in one demodulator shown on figure for the example case of user 1.

length n of the block in terms of symbols will be linked below with the coherence time T_{coh} of the channel. Each block is transmitted by the following signal:

$$s_{k,i}(t; \boldsymbol{b}_k[i]) \;=\; \sum_{m=in}^{(i+1)n-1} b_k[m] \mathrm{g}_{k,m}(t - mT_{\mathrm{s}}), \tag{3.1}$$

where T_{s} (sec) is the symbol period and $\mathrm{g}_{k,m}(t)$ is the unit energy waveform associated with the m-th symbol of user k. In general, $\mathrm{g}_{k,m}(t)$ is a spread-spectrum signal at user k prefilter output, and has band $[-\mathrm{W}/2, \mathrm{W}/2]$, that is, its spectrum is nonzero for $|f| \leq \mathrm{W}/2$. Assuming that $\{\mathrm{g}_{k,m}(t - mT_{\mathrm{s}})\colon m = in, \ldots, (i+1)n-1\}$ are orthonormal, or very mildly crosscorrelated, the energy of $s_{k,i}(t; \boldsymbol{b}_k[i])$ in eq. (3.1) is $n\,\mathbb{E}\left[\,|b_k[m]|^2\,\right]$; Since the block has duration nT_{s}, the average power is $\mathcal{P}_k = \mathbb{E}\left[\,|b_k[m]|^2\,\right]/T_{\mathrm{s}}$.

In the adopted model, demodulation at BS is performed on a block-by-block basis. Index i, that specifies the block number, is thus dropped. Consider for the sake of simplicity $i = 0$ in eq. (3.1):

$$s_k(t; \boldsymbol{b}_k) = \sum_{m=0}^{n-1} b_k[m] \mathrm{g}_{k,m}(t - mT_s). \tag{3.2}$$

Figure 4.1 shows the system model, including K modulators producing K transmitted signals, $s_k(t)$, $k = 1, \ldots, K$, affected by propagation within K different channels and corrupted at the receiver by white gaussian noise $n(t)$. The receiver consists in one demodulator.

Transmitted signal $s_k(t;\boldsymbol{b}_k)$ of each user propagates over a multipath channel with impulse response $c_k(t)$ and is distorted into $x_k(t)$:

$$x_k(t;\boldsymbol{b}_k) = s_k(t;\boldsymbol{b}_k) * c_k(t) = \sum_{\ell=0}^{\infty} c_{k,\ell} s_k(t-\tau_{k,\ell};\boldsymbol{b}_k), \tag{3.3}$$

where $\{c_{k,\ell}\colon \ell \geq 0\}$ and $\{\tau_{k,\ell}\colon \ell \geq 0\}$ are amplitudes and delays of the paths of $c_k(t)$, respectively.

The received signal is:

$$y(t;\boldsymbol{b}) = \sum_{k=1}^{K} x_k(t;\boldsymbol{b}_k) + n(t), \tag{3.4}$$

where $n(t)$ is a white Gaussian noise with flat power spectral density $\mathcal{N}_0/2$ (W/Hz). Throughout the chapter, the receiver estimates transmitted symbols of user k, $\{b_k[m]\colon k=1,\dots,K; m=0,\dots,n-1\}$, on a symbol-by-symbol basis, by considering users $j \neq k$ as unknown interference over user k; for example, Fig. 4.1 shows the demodulation of user 1. As detailed below, transmissions are symbol-synchronous but not necessarily chip-synchronous, therefore the symbol-by-symbol demodulation does not imply any performance loss. In the adopted model, the receiver is a single user detector, and as such suboptimal, since it does not take into account the possibility of joint multiuser detection. How channel is estimated and how error affected estimated channels play a role in the model will be explained further down in this section in association with the different modulation and demodulation structures. Expliciting signals for the symbol at time epoch $m=0$, and denoting by $b_k = b_k[0]$, eqs. (3.2), (3.3) and (3.4) become:

$$s_k(t;b_k) = b_k g_{k,0}(t), \tag{3.5}$$

$$x_k(t;b_k) = s_k(t;b_k) * c_k(t) = \sum_{\ell=0}^{\infty} c_{k,\ell} s_k(t-\tau_{k,\ell};b_k), \tag{3.6}$$

$$y(t;b_1,\dots,b_K) = \sum_{k=1}^{K} x_k(t;b_k) + n(t). \tag{3.7}$$

3.1.2 Single User Channel

Since the system symbol-synchronous, analysis may refer to transmission of one generic symbol, that is chosen as symbol $m=0$, $b[0]$, denoted by b. If transmission does not foresee prefiltering, that is, a zero-excess bandwidth pulse $\psi(t)$ with bandwidth W and unit energy is transmitted to modulate b, the received signal is:

$$\begin{aligned} y(t;b) = x(t;b) + n(t) &\overset{(a)}{=} b \sum_{\nu=0}^{N-1} s[\nu]\psi(t-\nu T_c) * c(t) + n(t) \\ &\overset{(b)}{=} b \sum_{\nu=0}^{N-1} s[\nu] \sum_{\ell=0}^{L_i} c[\ell]\psi(t-\ell/W-\nu T_c) + n(t) \end{aligned}$$

$$\overset{(c)}{=} b \sum_{\nu=0}^{N-1} \mathsf{s}[\nu] \sum_{\ell=0}^{L\imath} c[\ell]\psi(t-(\ell+\nu\imath)/\mathsf{W}) + n(t), \tag{3.8}$$

where in (a) the spreading sequence $\boldsymbol{s} = (\mathsf{s}[0], \dots, \mathsf{s}[N-1])^{\mathsf{T}}$ and the chip period T_{c} are made explicit; (b) follows from $\psi(t)$, and therefore also $\psi(t) * c(t)$, being bandlimited to $\mathsf{W}/2$; and (c) follows from assuming $T_{\mathsf{c}} = \imath/\mathsf{W}$, being $\imath$ a positive integer called *impulsiveness index*, that introduces a model to account for a pulse duration shorter than chip duration, as common in UWB communications. In the following, time-hopping is considered, for which all $\mathsf{s}[\nu]$ are zero, but one.

By projecting eq. (3.8) onto $\{\psi(t-k/\mathsf{W})\colon k = 0, \dots, (N+L+1)\imath - 1 - 1\}$, the following discrete model is obtained:

$$\boldsymbol{y} = \boldsymbol{C}\boldsymbol{x}\, b + \boldsymbol{n}, \quad \boldsymbol{x} = \boldsymbol{s} \otimes \boldsymbol{e}_1^{\imath}, \tag{3.9}$$

where $\boldsymbol{e}_1^{\imath} = [1, \boldsymbol{0}_{(\imath-1)\times 1}^{\mathsf{T}}]^{\mathsf{T}}$ is the first vector of the canonical basis of $\mathbb{R}^{\imath}$, $\boldsymbol{x}$ is $N\imath \times 1$, and $\boldsymbol{C}$ is Toeplitz with dimensions $(N+L+1)\imath - 1 \times N\imath$ and elements $C_{ij} = c[\imath - j]$. Since $c(t)$ is causal, it is assumed that $c[\ell] = 0$ for $\ell < 0$, and since $c(t)$ has finite delay spread T_d, it is assumed that $c[\ell] = 0$ for $\ell > L\imath$, being $T_d = LT_{\mathsf{c}} = L\imath/\mathsf{W}$; therefore, $\boldsymbol{C}$ is banded Toeplitz.

In general, for a system with prefiltering, with prefiltering impulse response $p(t)$, eq. (3.9) generalizes to (see *e.g.* [66]):

$$\boldsymbol{y} = \boldsymbol{C}\boldsymbol{P}\boldsymbol{x}\, b + \boldsymbol{n}, \tag{3.10}$$

where $\boldsymbol{P}$ is a Toeplitz matrix with dimensions $(N+2L)\imath \times (N+L+1)\imath - 1$ and elements $P_{ij} = p[\imath - j] = p * \psi((\imath - j)/\mathsf{W})$.

Prefiltering is introduced in order to compensate channel effects; in particular, prefiltering is based on an estimated version of the channel impulse response. In other words, imperfect prefiltering may be matched to channel estimation error patterns. If prefiltering is imperfect, as will be justified in Subsection 3.1.4, the error due to the estimation process can be modeled as a white Gaussian process $\xi(t)$, that is added to $\boldsymbol{P}$ as follows:

$$\hat{\boldsymbol{P}} = \alpha(\boldsymbol{P} + \boldsymbol{\Xi}), \tag{3.11}$$

where $\Xi_{ij} = \xi_{\imath - j} \sim \mathcal{N}(0, \sigma_\xi^2)$, where σ_ξ^2 accounts for estimation accuracy, and $\alpha > 0$ is such that $||\hat{\boldsymbol{P}}\boldsymbol{x}||^2 = ||\boldsymbol{P}\boldsymbol{x}||^2$.

No prefiltering, All-Rake receiver.

The traditional (or conventional) receiver is a matched-filter, *i.e.*, an AR receiver in the case of a multipath channel. Knowing the time-hopping spreading sequence $\boldsymbol{x}$ and the resolved channel $\boldsymbol{c}$, a sufficient statistic for b is obtained by projecting the received signal $\boldsymbol{y}$ onto $\boldsymbol{h} = \boldsymbol{C}\boldsymbol{x}$, or, equivalently, onto $\boldsymbol{h}/||\boldsymbol{h}||$:

$$z^{\mathrm{AR}} = \frac{\boldsymbol{h}^{\mathsf{T}}}{||\boldsymbol{h}||}\boldsymbol{y} = ||\boldsymbol{h}||b + \nu,$$

where $\nu \sim \mathcal{N}(0, \mathcal{N}_0/2)$.

As occurs in the prefiltering, also the AR receiver is affected by possible channel estimation errors. If the AR is provided with imperfect channel state information (CSI), that is, operates using an estimation $\hat{\boldsymbol{c}}$ of channel $\boldsymbol{c}$ that is impaired by an error $\boldsymbol{\xi} \sim \mathcal{N}(\boldsymbol{0}, \sigma_\xi^2 \boldsymbol{I}_{(Li+1)})$, then the AR combines paths through $\hat{\boldsymbol{h}} := \hat{\boldsymbol{C}}\boldsymbol{x}$ instead of $\boldsymbol{h} = \boldsymbol{C}\boldsymbol{x}$, and inference of b is based on:

$$\hat{z}^{\mathrm{AR}} = \frac{\hat{\boldsymbol{h}}^{\mathsf{T}}}{\|\hat{\boldsymbol{h}}\|}\boldsymbol{y} = \frac{(\boldsymbol{h}+\boldsymbol{\chi})^{\mathsf{T}}}{\|\boldsymbol{h}+\boldsymbol{\chi}\|}(\boldsymbol{h}b + \boldsymbol{n}) = \frac{(\boldsymbol{h}+\boldsymbol{\chi})^{\mathsf{T}}}{\|\boldsymbol{h}+\boldsymbol{\chi}\|}\boldsymbol{h}b + \frac{(\boldsymbol{h}+\boldsymbol{\chi})^{\mathsf{T}}}{\|\boldsymbol{h}+\boldsymbol{\chi}\|}\boldsymbol{n}, \tag{3.12}$$

where $\boldsymbol{\chi} = [\boldsymbol{0}_{j_{\boldsymbol{x}}}^{\mathsf{T}}, \boldsymbol{\xi}^{\mathsf{T}}, \boldsymbol{0}_{Ni-j_{\boldsymbol{x}}}^{\mathsf{T}}]^{\mathsf{T}}$, being $j_{\boldsymbol{x}}$ the nonzero dimension of $\boldsymbol{x}$.

Time Reversal prefiltering, 1Rake receiver.

The time reversal prefilter is represented by $p[j] = \alpha c[Li - j]$, where $\alpha > 0$ guarantees that prefiltered and non prefiltered transmitted waveforms have same energy. Time-hopping implies $\boldsymbol{s} \in \{\boldsymbol{e}_\nu\}_{\nu=1}^{N}$, and $\boldsymbol{x} = \boldsymbol{s} \otimes \boldsymbol{e}_1^i \in \{\boldsymbol{e}_{(\nu-1)i+1}\}_{\nu=1}^{N}$. A 1Rake receiver is given by $\boldsymbol{e}_{Li+j_{\boldsymbol{x}}}$. Denoting by $\boldsymbol{T}$ the time-reversal prefilter matrix, one has:

$$z^{\mathrm{TR}} = \boldsymbol{e}_{Li+j_{\boldsymbol{x}}}^{\mathsf{T}}\boldsymbol{y} = \boldsymbol{e}_{Li+j_{\boldsymbol{x}}}^{\mathsf{T}}\boldsymbol{C}\boldsymbol{T}\boldsymbol{x}\,b + \boldsymbol{e}_{Li+j_{\boldsymbol{x}}}^{\mathsf{T}}\boldsymbol{n} = \boldsymbol{e}_{Li+j_{\boldsymbol{x}}}^{\mathsf{T}}\boldsymbol{H}\boldsymbol{x}\,b + n_{Li+j_{\boldsymbol{x}}}, \tag{3.13}$$

being $n_{Li+j_{\boldsymbol{x}}} \sim \mathcal{N}(0, \mathcal{N}_0/2)$.

If the transmitter is provided with imperfect CSI, then model of eq. (3.11) holds, and eq. (3.13) becomes:

$$\hat{z}^{\mathrm{TR}} = \boldsymbol{e}_{Li+j_{\boldsymbol{x}}}^{\mathsf{T}}\boldsymbol{C}\hat{\boldsymbol{T}}\boldsymbol{x}\,b + \boldsymbol{e}_{Li+j_{\boldsymbol{x}}}^{\mathsf{T}}\boldsymbol{n} = \boldsymbol{e}_{Li+j_{\boldsymbol{x}}}^{\mathsf{T}}\boldsymbol{C}[\alpha(\boldsymbol{T}+\boldsymbol{\Xi})]\boldsymbol{x}\,b + n_{Li+j_{\boldsymbol{x}}}. \tag{3.14}$$

AR vs. TR.

As well-known [76], TR is equivalent to a system without prefiltering and AR in terms of the signal-to-noise ratio. From a single user perspective, there is no performance difference in both uncoded (symbol error probability) and coded (channel capacity) regimes between the two transceiver structures. Moreover, previous work [74] suggested that sets of equivalent systems can be obtained with partial Rakes compensating for partial time reversal transmitter structures. In the case of imperfect CSI, the comparison of the different structures is the object of this chapter.

3.1.3 Multiuser Channel

A straightforward extension of eq. (3.9) to K users is as follows:

$$\boldsymbol{y} = \sum_{k=1}^{K} \boldsymbol{C}_k \boldsymbol{x}_k b_k + \boldsymbol{n}, \tag{3.15}$$

where $\boldsymbol{x}_k = \boldsymbol{s}_k \otimes \boldsymbol{e}_{l_k}^{\iota}$ and $1 \leq l_k \leq \iota$ models the chip-asynchronism by making l_k i.i.d. according to a uniform distribution. This extension holds based on the hypothesis that all UTs are symbol-synchronous. This hypothesis is reasonable since, as further discussed in Subsection 3.1.4, the BS broadcasts in a link setup phase a known sequence to the UTs. Denoting by $\boldsymbol{h}_k = \boldsymbol{C}_k \boldsymbol{x}_k$ the spreading sequence $\boldsymbol{x}_k$ after transition in the multipath channel, and by:

$$\boldsymbol{H} = \sum_{k=1}^{K} \boldsymbol{e}_k^{\mathsf{T}} \otimes \boldsymbol{h}_k = [\boldsymbol{h}_1, \ldots, \boldsymbol{h}_K]$$

the spreading matrix, eq. (3.9) can also be rewritten as follows:

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{b} + \boldsymbol{n}, \tag{3.16}$$

where $\boldsymbol{b} = (b_1, \ldots, b_K)^{\mathsf{T}}$. For systems with prefiltering, eq. (3.15) generalizes to:

$$\boldsymbol{y} = \sum_{k=1}^{K} \boldsymbol{C}_k \boldsymbol{P}_k \boldsymbol{x}_k \, b_k + \boldsymbol{n}, \tag{3.17}$$

where matrices $\boldsymbol{P}_k$ and $\boldsymbol{C}_k$ have same dimensions as $\boldsymbol{P}$ and $\boldsymbol{C}$ of eq. (3.10), respectively, and eq. (3.16) holds with $\boldsymbol{h}_k = \boldsymbol{C}_k \boldsymbol{P}_k \boldsymbol{x}_k$. In the presence of imperfect CSI, $\boldsymbol{P}_k$ in eq. (3.17) is substituted by $\hat{\boldsymbol{P}}_k$, as defined in eq. (3.11), where estimation errors are independent with respect to k.

No prefiltering, All-Rake receiver.

The decision variable following the matched filter of user k is:

$$z_k^{\mathrm{AR}} = \|\boldsymbol{h}_k\| b_k + \sum_{\substack{j=1 \\ j \neq k}}^{K} \frac{\boldsymbol{h}_k^{\mathsf{T}}}{\|\boldsymbol{h}_k\|} \boldsymbol{h}_j \, b_j + \frac{\boldsymbol{h}_k^{\mathsf{T}}}{\|\boldsymbol{h}_k\|} \boldsymbol{n} = \|\boldsymbol{h}_k\| b_k + I_k + \nu_k, \tag{3.18}$$

where $\boldsymbol{h}_i = \boldsymbol{C}_i \boldsymbol{x}_i$, I_k represents the MUI, and $\nu_k = \frac{\boldsymbol{h}_k^{\mathsf{T}}}{\|\boldsymbol{h}_k\|} \boldsymbol{n} \sim \mathcal{N}(0, \mathcal{N}_0/2)$.

If the AR is provided with imperfect CSI, then signal $\boldsymbol{y}$ is projected onto $\hat{\boldsymbol{h}}_k$ instead of $\boldsymbol{h}_k$, hence:

$$\hat{z}_k^{\mathrm{AR}} = \frac{\hat{\boldsymbol{h}}_k^{\mathsf{T}}}{\|\hat{\boldsymbol{h}}_k\|} \boldsymbol{h}_k b_k + \sum_{\substack{j=1 \\ j \neq k}}^{K} \frac{\hat{\boldsymbol{h}}_k^{\mathsf{T}}}{\|\hat{\boldsymbol{h}}_k\|} \boldsymbol{h}_j \, b_j + \frac{\hat{\boldsymbol{h}}_k^{\mathsf{T}}}{\|\hat{\boldsymbol{h}}_k\|} \boldsymbol{n} = \hat{a}_{kk}^{\mathrm{AR}} b_k + \hat{I}_k^{\mathrm{AR}} + \hat{\nu}_k^{\mathrm{AR}}, \tag{3.19}$$

where $\hat{\nu}_k^{\mathrm{AR}} = \frac{\hat{\boldsymbol{h}}_k^{\mathsf{T}}}{\|\hat{\boldsymbol{h}}_k\|} \boldsymbol{n} \sim \mathcal{N}(0, \mathcal{N}_0/2)$.

Time Reversal prefiltering, 1Rake receiver.

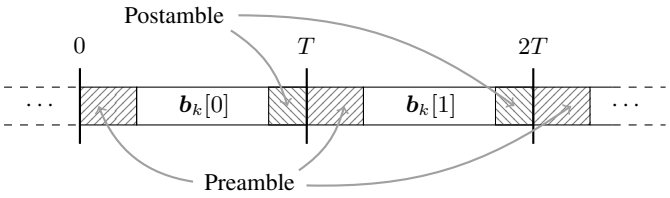


Figure 3.3: Data transmission structured into blocks for UT_k.

With time reversal, the decision variable for user k becomes:

$$z_k^{\mathrm{TR}} = \boldsymbol{e}_{q_k}^{\mathsf{T}} \boldsymbol{C}_k \boldsymbol{T}_k \boldsymbol{x}_k \, b_k + \sum_{\substack{j=1 \\ j \neq k}}^{K} \boldsymbol{e}_{q_k}^{\mathsf{T}} \boldsymbol{C}_j \boldsymbol{T}_j \boldsymbol{x}_j \, b_j + \boldsymbol{e}_{q_k}^{\mathsf{T}} \boldsymbol{n}, \tag{3.20}$$

where $q_k = L\imath + j_{\boldsymbol{x}_k}$ is the delay (in samples) to which the 1Rake is synchronized. In the presence of imperfect CSI, the decision variable is:

$$\hat{z}_k^{\mathrm{TR}} = \boldsymbol{e}_{q_k}^{\mathsf{T}} \boldsymbol{C}_k \alpha_k (\boldsymbol{T}_k + \boldsymbol{\Xi}_k) \boldsymbol{x}_k \, b_k + + \sum_{\substack{j=1 \\ j \neq k}}^{K} \alpha_j \boldsymbol{e}_{q_k}^{\mathsf{T}} \boldsymbol{C}_j (\boldsymbol{T}_j + \boldsymbol{\Xi}_j) \boldsymbol{x}_j \, b_j + \boldsymbol{e}_{q_k}^{\mathsf{T}} \boldsymbol{n}. \tag{3.21}$$

AR vs. TR.

As well-known [76, 78], time reversal usually increases the kurtosis of the interference at the output of Rake receivers. This follows from the fact that the effective channel impulse response formed by the combination of prefilter and multipath channel has a peaked behavior, whereas without time reversal the behavior is non-peaked. While in the single user case the two schemes are equivalent, this equivalence does not hold in the multiuser case. The impact of estimation errors will be investigated below.

3.1.4 Channel Estimation and Data Transmission

For both AR and TR, the channel impulse response estimation takes place, at least partially, both in the transmitter and in the receiver.

Actual transmission of the set of information-bearing symbols requires, therefore, additional symbols to be sent either in a preamble or in a postamble of the block [79], as shown in Fig. 3.3.

Training is the simplest estimation process to evaluate the necessary channel state information. Time-Division Duplexing (TDD) is assumed as commonly witnessed in impulse-radio as well as common WLAN transmissions. Since precoding of UT_k does not depend on channels experienced by users $j \neq k$, feedback is not a necessary feature, given that channel is reciprocal. Note that there is no dedicated training since precoding is supposed to be disjoint, that is, the precoding vector of each UT does uniquely depend on the channel between its transmitter and the BS

and is in particular independent of channels and precoding vectors of other UTs (see [80] for a thorough discussion).

Transmission follows a scheme that is shown at a glance in Fig. 3.4a and in more detail in Fig. 3.4b. Figure 3.4 summarizes the organization of the different links (downlink, *i.e.*, broadcast, *vs.* uplinks) over time, where durations of data, preamble and postamble are indicated in terms of number of chips (N_t^{DL}, N_t^{UL}, N_{d}, N_g^{UL}). In particular, Fig. 3.4b shows an information exchange between the BS and each UT consisting in four phases:

1. **Downlink Channel Training**: the BS broadcasts a training sequence of length $N_t^{\mathrm{DL}} i$ samples, known by the set of UTs, followed by a zero-padding sequence of Li samples (idle period), that allows each UT to receive the training sequence smeared by the channel; each UT estimates the channel based on the received samples and the knowledge of the transmitted training sequence, as further detailed in the remaining part of this section. This training sequence may be also used for network synchronization at symbol level.

2. **Uplink Channel Training**: each UT transmits a training sequence of length $N_t^{\mathrm{UL}} i$ samples, known by the BS, followed by a zero-padding sequence of length Li samples; these training sequences are assumed pseudo-noise (PN) sequences rather than orthogonal given that each UT chooses its sequence independently from the others. The BS estimates channels through the observation of the superposition of training sequences, that have been distorted by respective channels.

3. **Data Transmission**: each UT sends n information-bearing symbols of a block, corresponding to $N_{\mathrm{d}} i = nNi$ samples;

4. **Idle**: each UT sends a zero-padded postamble of duration Li samples.

During the downlink training, the BS broadcasts its training sequence to the UTs. With reference to model of Section 3.1, and in particular to eq. (3.9) and impulsiveness index i, the received signal at UT_k is:

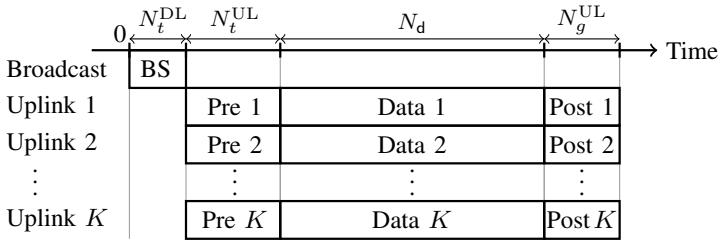
$$\boldsymbol{y}_k^{\mathrm{DL}} = \boldsymbol{C}_k \boldsymbol{v}^{\mathrm{DL}} + \boldsymbol{n}_k, \quad \boldsymbol{v}^{\mathrm{DL}} = \boldsymbol{\phi}_k^{\mathrm{DL}} \otimes \boldsymbol{e}_1^i, \tag{3.22}$$

where $\boldsymbol{y}_k^{\mathrm{DL}}$ is the $(N_t^{\mathrm{DL}} i + Li) \times 1$ vector of received samples, $\boldsymbol{C}_k$ is the $(N_t^{\mathrm{DL}} i + Li) \times N_t^{\mathrm{DL}} i$ Toeplitz channel matrix, $\boldsymbol{\phi}_k^{\mathrm{DL}}$ is the $N_t^{\mathrm{DL}} \times 1$ training sequence, $\boldsymbol{v}^{\mathrm{DL}}$ is the $N_t^{\mathrm{DL}} i \times 1$ training sequence accounting for impulsiveness, and $\boldsymbol{n}_k$ is the $(N_t^{\mathrm{DL}} i + Li) \times 1$ white Gaussian noise vector.

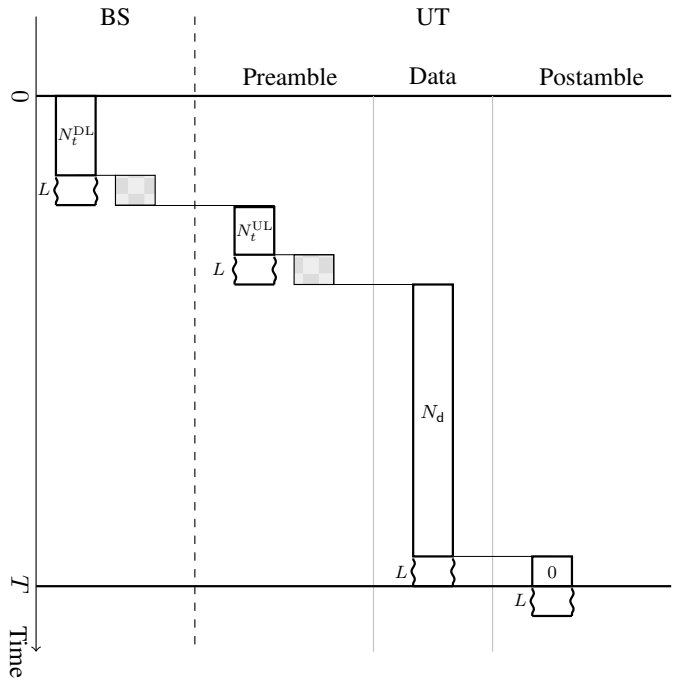
Eq. (3.22) can be rewritten as follows:

$$\boldsymbol{y}_k^{\mathrm{DL}} = \boldsymbol{\Upsilon}^{\mathrm{DL}} \boldsymbol{c}_k + \boldsymbol{n}_k, \tag{3.23}$$

where now $\boldsymbol{\Upsilon}^{\mathrm{DL}}$ is a $(N_t^{\mathrm{DL}} i + Li) \times Li$ Toeplitz matrix and $\boldsymbol{c}_k$ is the $Li \times 1$ channel vector.



(a) Transmission scheme: at a glance.



(b) Detailed transmission scheme.

Figure 3.4: Transmission schemes. In *(a)*: transmission scheme at a glance. In *(b)*: detailed transmission scheme. *Phase 1*: the BS broadcasts a training sequence of length $N_t^{\mathrm{DL}}\imath$ samples (corresponding to $N_t^{\mathrm{DL}}\imath/\mathsf{W}$ seconds) that is received by each UT starting at time 0. Each multipath channel spreads the sequence for $L\imath$ samples, hence each UT listens from time 0 to time $N_t^{\mathrm{DL}}\imath + L\imath$ samples. This training sequence may be also used for network synchronization at symbol level. *Phase 2*: Once the training sequence is received, each UT transmits its own training sequence of length $N_t^{\mathrm{UL}}\imath$ samples to the BS (preamble). By reciprocity, channels spread these sequences for $L\imath$ samples, therefore each UT remains idle for $L\imath$ samples. *Phase 3*: Each UT transmits a sequence of information-bearing symbols for $N_{\mathrm{d}}\imath = nN\imath$ samples. *Phase 4*: Each UT transmits a sequence of null symbols denoted with 0 (postamble).

In order to minimize the signal-plus-interference-to-noise ratio, UT_k may use an MMSE estimation of $\boldsymbol{c}_k$, where the cause of interference is due to multipath. However, the use of PN sequences as training sequences is very common, due to their good autocorrelation properties. In fact, PN sequences have periodic ACF of the following form [81, 82]:

$$\rho_{\phi}[i] = \|\boldsymbol{\phi}\|^2 \left(-\frac{1}{N_t^{\mathrm{DL}}} + \delta_{i,0} \right), \quad 0 \leq i \leq N_t^{\mathrm{DL}} - 1,$$

that is asymptotically impulse-like, as $N_t^{\mathrm{DL}} \gg 1$.

Asymptotically then, and dropping the superscript DL to unclutter notation, $\boldsymbol{\Upsilon}^{\mathsf{T}}\boldsymbol{\Upsilon} \approx \|\boldsymbol{\phi}\|^2 \boldsymbol{I} = \|\boldsymbol{\upsilon}\|^2 \boldsymbol{I}$, and MMSE reduces to a matched-filter, and estimation is as follows:

$$\boldsymbol{z}_k \approx \boldsymbol{\Upsilon}^{\mathsf{T}}\boldsymbol{\Upsilon}\,\boldsymbol{c}_k + \boldsymbol{\Upsilon}^{\mathsf{T}}\boldsymbol{n}_k = \|\boldsymbol{\upsilon}\|^2 \boldsymbol{c}_k + \boldsymbol{\nu}_k, \qquad \boldsymbol{\nu}_k \sim \mathcal{N}(\boldsymbol{0}, \sigma_N^2 \|\boldsymbol{\upsilon}\|^2 \boldsymbol{I}). \tag{3.24}$$

Dividing by $\|\boldsymbol{\upsilon}\|^2$ the previous expression yields:

$$\hat{\boldsymbol{c}}_k := \frac{1}{\|\boldsymbol{\upsilon}\|^2} \boldsymbol{z}_k \approx \boldsymbol{c}_k + \frac{1}{\|\boldsymbol{\upsilon}\|} \boldsymbol{\nu}_k = \boldsymbol{c}_k + \boldsymbol{\nu}'_k, \tag{3.25}$$

where $\boldsymbol{\nu}'_k \sim \mathcal{N}(\boldsymbol{0}, (\sigma_N^2/\|\boldsymbol{\upsilon}\|^2)\boldsymbol{I})$. Note that, as well-known (*e.g.* [83]), the estimation can be made as accurate as desired by increasing $\|\boldsymbol{\upsilon}\|^2$. For antipodal sequences, say $\phi[i] \in \{-\mathsf{A}_t, \mathsf{A}_t\}$ with $\mathsf{A}_t > 0$, the energy of the training sequence is $\mathsf{A}_t^2 N_t^{\mathrm{DL}}$; therefore, $\|\boldsymbol{\upsilon}\|^2$ can be increased either by increasing power spent on training, that is, by increasing A_t, or by increasing time spent for training, that is, by increasing N_t^{DL}, or both.

In the uplink training, the BS receives the superposition of the sequences of users each filtered by the corresponding channel, that is:

$$\boldsymbol{y}^{\mathrm{UL}} = \sum_{k=1}^{K} \boldsymbol{\Upsilon}_k^{\mathrm{UL}} \boldsymbol{c}_k + \boldsymbol{n} \equiv \boldsymbol{\Upsilon}^{\mathrm{UL}} \boldsymbol{c} + \boldsymbol{n}, \tag{3.26}$$

having defined:

$$\boldsymbol{\Upsilon}^{\mathrm{UL}} := \sum_{k=1}^{K} \boldsymbol{e}_k^{\mathsf{T}} \otimes \boldsymbol{\Upsilon}_k^{\mathrm{UL}}, \qquad \boldsymbol{c} := \sum_{k=1}^{K} \boldsymbol{e}_k \otimes \boldsymbol{c}_k. \tag{3.27}$$

As previously, the superscript UL is dropped to unclutter notation.

The goal of the BS is to linearly estimate $\boldsymbol{c}$ by observing $\boldsymbol{y}$, knowing $\boldsymbol{\Upsilon}$:

$$\boldsymbol{z} = \boldsymbol{W}^{\mathsf{T}} \boldsymbol{y} \equiv \sum_{k=1}^{K} \boldsymbol{e}_k \otimes \boldsymbol{z}_k = \boldsymbol{W}^{\mathsf{T}} \boldsymbol{\Upsilon} \boldsymbol{c} + \boldsymbol{W}^{\mathsf{T}} \boldsymbol{n}, \tag{3.28}$$

where $\boldsymbol{z}$ is the $KLi \times 1$ vector of channel estimations, being $\boldsymbol{z}_k$ the $Li \times 1$ vector representing $\boldsymbol{c}_k$ estimate, and $\boldsymbol{W}^{\mathsf{T}}$ is the $KLi \times N_t^{\mathrm{UL}} i$ matrix representing the estimator. All common linear estimators, that is ZF (Zero-Forcing), RZF (Regularized Zero-Forcing), MMSE (Minimum Mean Square Error) and MF (Matched Filter),

can be described by the following expression, parametrized by ξ and ζ:

$$\boldsymbol{W}^{\mathsf{T}} = (\xi \boldsymbol{\Upsilon}^{\mathsf{T}} \boldsymbol{\Upsilon} + \zeta \boldsymbol{I})^{-1} \boldsymbol{\Upsilon}^{\mathsf{T}}. \tag{3.29}$$

Indeed, MMSE is obtained with $(\xi, \zeta) = (1, \sigma_N^2)$; ZF with $(\xi, \zeta) = (1, 0)$; MF with $(\xi, \zeta) = (0, 1)$; RZF with $(\xi, \zeta) = (1, z)$.

In the simple case of ZF, the form assumed by eq. (3.28) is as follows:

$$\boldsymbol{z} = \boldsymbol{c} + (\boldsymbol{\Upsilon}^{\mathsf{T}} \boldsymbol{\Upsilon})^{-1} \boldsymbol{\Upsilon}^{\mathsf{T}} \boldsymbol{n} \equiv \boldsymbol{c} + \boldsymbol{\nu}, \quad \boldsymbol{\nu} \sim \mathcal{N}(\boldsymbol{0}, \sigma_N^2 (\boldsymbol{\Upsilon}^{\mathsf{T}} \boldsymbol{\Upsilon})^{-1}) \tag{3.30}$$

and, therefore, the ℓ-th tap of the channel of generic user k is:

$$z_k[\ell] = c_k[\ell] + \nu_k[\ell].$$

Here, $\nu_k[\ell]$ is a correlated Gaussian random variable with variance coinciding with the $((k-1)L\imath + \ell + 1)$-th diagonal element of $\sigma_N^2 (\boldsymbol{\Upsilon}^{\mathsf{T}} \boldsymbol{\Upsilon})^{-1}$.

Assuming all UTs are transmitting the same power, *i.e.*, $\|\boldsymbol{v}_k\|^2$ is the same for each $1 \leq k \leq K$, the approximation $\boldsymbol{\Upsilon}_j^{\mathsf{T}} \boldsymbol{\Upsilon}_i = \|\boldsymbol{v}\|^2 \boldsymbol{I} \delta_{ji}$ allows to assume uncorrelated estimation errors, since $\boldsymbol{\Upsilon}^{\mathsf{T}} \boldsymbol{\Upsilon} = \|\boldsymbol{v}\|^2 \boldsymbol{I}$, and thus:

$$\boldsymbol{z} = \boldsymbol{c} + \boldsymbol{\nu}, \quad \boldsymbol{\nu} \sim \mathcal{N}(\boldsymbol{0}, (\sigma_N^2 / \|\boldsymbol{v}\|^2) \boldsymbol{I}), \quad N_t^{\mathrm{UL}} \gg L. \tag{3.31}$$

3.1.5 Performance measures

In both system structures, the statistic for inferring the transmitted symbol b_k of user k can be written in the following form:

$$z_k = a_{kk} b_k + \sum_{\substack{j=1 \\ j \neq k}}^{K} a_{kj} b_j + \nu_k = a_{kk} b_k + I_k + \nu_k,$$

where ν_k is a r.v. representing noise, and $\{a_{kj} : j = 1, \ldots, K\}$ are r.*vs.* depending on multipath channels, random time-hopping codes, random delays, and estimation errors.

Two performance measures are considered.

In the *uncoded* regime, the probability of error as defined by:

$$P_e = \frac{1}{2} \mathbb{P}\Big(z_k < 0 \,\Big|\, b = \sqrt{\mathcal{E}} \Big) + \frac{1}{2} \mathbb{P}\Big(z_k > 0 \,\Big|\, b = -\sqrt{\mathcal{E}} \Big),$$

is considered.

In the *coded* regime, mutual information with Gaussian inputs and a bank of matched-filters followed by independent decoders is considered; for the generic user k, this is given by:

$$I(b_k; z_k) \quad \text{nats/channel use}, \tag{3.32}$$

where $I(b_k; z_k)$ is the mutual information between the transmitted symbol b_k and the decision variable z_k. Since a channel use corresponds to $NT_{\mathrm{c}} = N\imath/W$ seconds,

the sum-rate achieved by the set of K users is:

$$R \triangleq W \frac{\beta}{\lambda} I(b;z) \quad \text{nats/s}, \tag{3.33}$$

having indicated with $I(b;z)$ the mutual information (3.32) for a generic user. Finally, a spectral efficiency equal to

$$\mathcal{R} \triangleq \frac{\beta}{\lambda} I(b;z) \quad \text{(nats/s)/Hz} \tag{3.34}$$

is obtained.

3.2 Probability of Error

3.2.1 Single User

The main contribution of this subsection is to show that imperfect TR and AR achieves the same probability of error, and, therefore, that the same accuracy is needed for channel estimation at transmitter and receiver in order to achieve a given error probability.

With reference to decision variables $\hat{z}^{\mathrm{TR}}$ of eq. (3.14) and $\hat{z}^{\mathrm{AR}}$ of eq. (3.12), the probability of error, in both cases, is:

$$P_e = \frac{1}{2}\mathbb{P}(z<0 \,|\, b=\mathsf{A}) + \frac{1}{2}\mathbb{P}(z>0 \,|\, b=-\mathsf{A}) = \mathbb{P}(z<0 \,|\, b=\mathsf{A}), \tag{3.35}$$

where the first equality follows from b belonging to $\{-\mathsf{A},\mathsf{A}\}$ with equal probability, and the second equality follows from the distribution of z being an even function. For the power constraint, it results $\mathsf{A}=\sqrt{\mathcal{E}}$. Equivalence of P_e for the two cases is derived by showing that $\hat{z}^{\mathrm{TR}}$ and $\hat{z}^{\mathrm{AR}}$ have the same distribution.

To this end, rewrite the decision variable $\hat{z}^{\mathrm{AR}}$ conditioned on $b=\mathsf{A}$. Without loss of generality, and for the sake of simplicity, consider $\boldsymbol{x}=\boldsymbol{e}_1$. Then:

$$\hat{z}^{\mathrm{AR}} = ||\boldsymbol{c}||^2\mathsf{A} + \boldsymbol{c}^{\mathsf{T}}(\boldsymbol{n}+\boldsymbol{\xi}\mathsf{A}) + \boldsymbol{\xi}^{\mathsf{T}}\boldsymbol{n} = ||\boldsymbol{c}||^2\mathsf{A} + \boldsymbol{c}^{\mathsf{T}}\boldsymbol{\xi}\mathsf{A} + \boldsymbol{n}^{\mathsf{T}}(\boldsymbol{\xi}+\boldsymbol{c}).$$

Similarly, the decision variable $\hat{z}^{\mathrm{TR}}$ conditioned on $b=\mathsf{A}$ is:

$$\hat{z}^{\mathrm{TR}} = \mathsf{A}\,\boldsymbol{c}^{\mathsf{T}}(\boldsymbol{c}+\boldsymbol{\xi}) + n||\boldsymbol{c}+\boldsymbol{\xi}||,$$

where $n \sim \mathcal{N}(0,\sigma_N^2)$. By comparing the two expressions, $\hat{z}^{\mathrm{AR}}$ is equivalent to $\hat{z}^{\mathrm{TR}}$, the equivalence being defined as producing the same P_e, iff term $\boldsymbol{n}^{\mathsf{T}}(\boldsymbol{\xi}+\boldsymbol{c})$ in $\hat{z}^{\mathrm{AR}}$ is distributed as $n||\boldsymbol{c}+\boldsymbol{\xi}||$ in $\hat{z}^{\mathrm{TR}}$. This is, indeed, the case; by choosing an orthogonal matrix $\boldsymbol{Q}$ such that $\boldsymbol{Q}(\boldsymbol{c}+\boldsymbol{\xi}) = ||\boldsymbol{c}+\boldsymbol{\xi}||\boldsymbol{e}_1$, one has:

$$\boldsymbol{n}^{\mathsf{T}}(\boldsymbol{c}+\boldsymbol{\xi}) = \boldsymbol{n}^{\mathsf{T}}\boldsymbol{Q}^{\mathsf{T}}\boldsymbol{Q}(\boldsymbol{c}+\boldsymbol{\xi}) = \boldsymbol{n}_1'^{\mathsf{T}}||\boldsymbol{c}+\boldsymbol{\xi}||\boldsymbol{e}_1 = n_1'||\boldsymbol{c}+\boldsymbol{\xi}||,$$

where $n_1' \sim \mathcal{N}(0,\sigma_N^2)$, hence the equivalence in terms of distributions, and, therefore, probability of error is verified.

Based on the described training algorithm, samples of the channel estimated by A are:

$$\hat{\boldsymbol{c}} = \boldsymbol{c} + \boldsymbol{\xi} \in \mathbb{R}^{L+1}, \tag{3.36}$$

where L is the channel length in terms of samples ($L = T_d W$, being T_d the delay spread of the channel); $\boldsymbol{c} = [c_0, c_1, \ldots, c_L]^{\mathsf{T}}$ is the vector of samples of the channel impulse response, where:

$$c[i] = \langle c(t), \psi(t - i/W) \rangle \triangleq \int_{-\infty}^{\infty} c(t)\psi(t - i/W)dt \; ;$$

$\boldsymbol{\xi} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}\sigma_\xi^2)$ is a Gaussian random vector modeling the uncertainty of estimation, being $\sigma_\xi^2 = \sigma_N^2 / \mathcal{E}_{\text{tr}}^{\text{r}}$ where $\mathcal{E}_{\text{tr}}^{\text{r}}$ is the energy of the training sequence received by A during the Training Phase I [83].

Focus on transmission of one symbol only, that for the sake of simplicity is assumed binary, $b \in \{-1, 1\}$. Assume that the symbol period is formed of N chips of duration $1/W$; therefore, the symbol period has duration $T_s = N/W$. The sampled received signal is:

$$\boldsymbol{y} = \sqrt{\mathcal{E}} \boldsymbol{C} \hat{\boldsymbol{p}} b + \boldsymbol{n},$$

where $\mathcal{E}$ is the trasmitted energy per symbol, $\boldsymbol{C}$ is the $(N+2L) \times (N+L)$ convolution matrix describing the channel, $\hat{\boldsymbol{p}}$ is the $(N + L) \times 1$ transmitted waveform, and $\boldsymbol{n} \sim \mathcal{N}(\boldsymbol{0}, \sigma_N^2 \boldsymbol{I})$. In particular, $\hat{p}[i] = \hat{c}[L - i]$ for $0 \leq i \leq L$, and $\hat{p}[i] = 0$ otherwise, where $\hat{c}[\ell]$ are the elements of $\hat{\boldsymbol{c}}$ as in eq. (3.36). A 1Rake yields to the following decision variable, upon which depends the decision on b:

$$\alpha = \sqrt{\mathcal{E}}\, \boldsymbol{c}^{\mathsf{T}} \frac{\boldsymbol{c} + \boldsymbol{\xi}}{\|\boldsymbol{c} + \boldsymbol{\xi}\|} b + n, \tag{3.37}$$

and $n \sim \mathcal{N}(0, \sigma_N^2)$. Define:

$$\zeta \triangleq \frac{\boldsymbol{c}^{\mathsf{T}}(\boldsymbol{c} + \boldsymbol{\xi})}{\|\boldsymbol{c}\| \cdot \|\boldsymbol{c} + \boldsymbol{\xi}\|}, \tag{3.38}$$

then eq. (3.37) becomes:

$$\alpha = \sqrt{\mathcal{E}}\, \|\boldsymbol{c}\|\, \zeta\, b + n. \tag{3.39}$$

The two performance measures analyzed are the signal-to-noise ratio SNR:

$$\mathsf{SNR} \triangleq \frac{\mathcal{E}\|\boldsymbol{c}\|^2 \zeta^2}{\sigma_N^2}, \tag{3.40}$$

where ζ^2 is the SNR loss, that is a r.v. since $\boldsymbol{\xi}$ is a random vector, and the symbol error probability:

$$P_e = \mathbb{P}(\, \alpha < 0 \,|\, b = 1 \,). \tag{3.41}$$

As derived in Appendix 3.A, the PDF of ζ is:

$$\mathsf{f}_\zeta(z) = \mathsf{f}_{\mathcal{T}'_L\left(\frac{|\boldsymbol{c}|}{\sigma_\xi}\right)}\left(\frac{\sqrt{L}}{\sqrt{1 - z^2}} z\right) \cdot \frac{\sqrt{L}}{\sqrt{(1 - z^2)^3}} \;, \; |z| < 1 \;, \tag{3.42}$$

where $\mathrm{f}_{\mathcal{T}'_L(\|\boldsymbol{c}\|/\sigma_\xi)}$ is the Student $\mathcal{T}$-distribution with L degrees of freedom and non-central parameter $\|\boldsymbol{c}\|/\sigma_\xi$. Figure 3.5 shows simulated histograms *vs.* $\mathrm{f}_\zeta(z)$ for different values of L and $\sigma_\xi/\|\boldsymbol{c}\|$. Values of L corresponds, for example, to a channel with delay spread of $T_d \approx 50$ ns, and a signal with a bandwidth of $W \approx 0.5$ GHz. Values of $\sigma_\xi/\|\boldsymbol{c}\|$ depends on $\mathcal{E}_{\mathrm{tr}}^{\mathrm{r}}$, and, therefore, on the training sequence power and the duration of the training. By comparing Fig. 3.5(a) *vs.* (b), it is shown, as can be expected, that ζ approaches 1 when the estimation improves. It can be shown that, as $\sigma_\xi^2 \to 0$, *i.e.*, for vanishing estimation errors, the PDF tends to $\delta(z-1)$, and no loss in eq. (3.39) occurs. By comparing Fig. 3.5(a) *vs.* (c), it is shown that ζ departs from 1 as L increases. This can be intuitively justified as follows: for a given accuracy, the total uncertainty on the channel increases with the number of resolved paths $L = T_d W$, and, therefore, as L increases, the variance of the estimation error on each path must decrease to avoid reduced, or even worse, performance. Moreover, physical multipath channels become more and more sparse as bandwidth increases: the number of multipath components, that is, the number of resolvable paths, is indeed limited, and so is the energy gain that can be carried by the whole channel. Hence, as bandwidth increases, there are no multipath components in an increasing fraction of the L taps composing the resolved channel, that, therefore, become a mere source of nuisance.

Note that, in eq. (3.42), $\boldsymbol{c}$ is regarded as nonrandom since during each coherence time the channel remains constant. This allows to derive, for example, the symbol error probability that affects the system during a particular coherence time, as studied below. However, if average performance over multiple coherence time is of interest, then the PDF of ζ must be regarded as the conditional pdf given $\boldsymbol{c}$. Nonetheless, although ζ depends upon $\boldsymbol{c}$, f_ζ depends upon $\boldsymbol{c}$ only via $\|\boldsymbol{c}\|$. Not the entire channel realization affects ζ, but just its energy $\|\boldsymbol{c}\|^2$. Therefore, although in the following $\boldsymbol{c}$ is regarded as nonrandom, it is just $\|\boldsymbol{c}\|$ to be nonrandom, that is a fading coefficient.

There are two main detrimental effects that imperfect estimation implies on SNR and P_e.

The first effect is a reduction in the SNR. The variance of the useful term $\sqrt{\mathcal{E}}\,\|\boldsymbol{c}\|\,\zeta\, b$ in eq. (3.39) is, indeed, equal to $\mathcal{E}\|\boldsymbol{c}\|^2\zeta^2$: with perfect CSIT, $\zeta = 1$, while with imperfect CSIT, $\zeta \in [-1, 1)$, and therefore $\zeta^2 \in [0, 1)$ measures the loss of variance in the useful term, and thus in SNR (see eq. (3.40)), due to the imperfect knowledge of the channel.

The second effect is the presence of a symbol error probability floor depending on σ_ξ^2, irrespective of the amount of power spent in transmission (Data Transmission Phase), and only depending on the energy spent during the Training Phase I. Accuracy of estimation can bound, therefore, the achievability of low (uncoded) symbol error probability. The symbol error probability floor appears as $\mathcal{E}/\sigma_N^2 \to \infty$; in this asymptotic case, the decision variable tends to:

$$\bar{\alpha} = \sqrt{\mathcal{E}}\,\frac{\boldsymbol{c}^{\mathsf{T}}(\boldsymbol{c}+\boldsymbol{\xi})}{\|\boldsymbol{c}+\boldsymbol{\xi}\|}\, b,$$

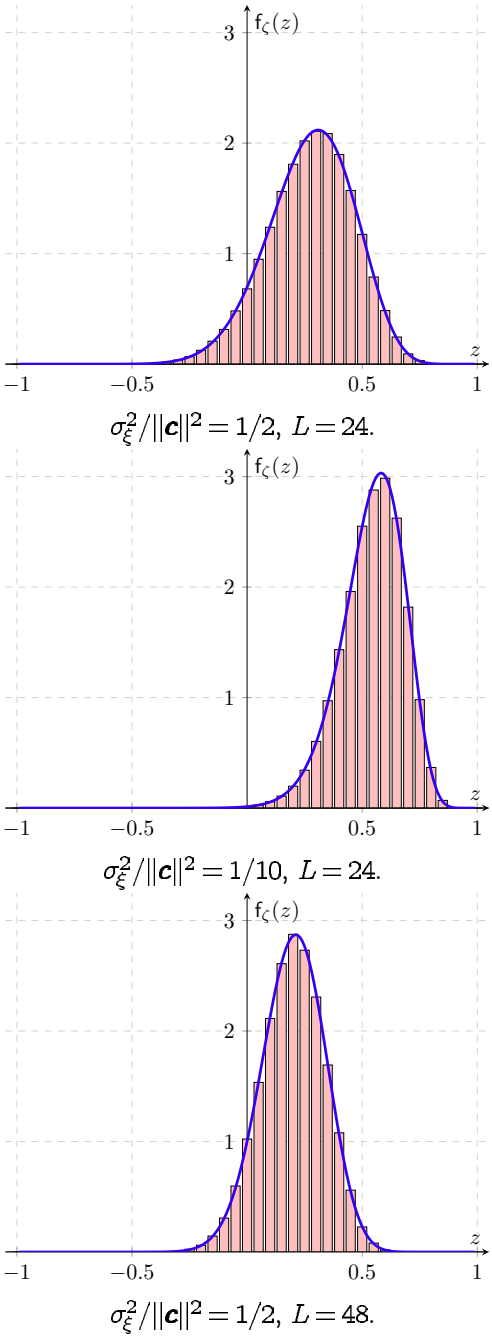


Figure 3.5: PDF of ζ: simulated histograms *vs.* theoretical expression of eq. (3.42). Parameters: L is the number of resolvable paths of the channel; σ_ξ^2 is the variance of the estimation error.

and decision on b is made based on the sign of $\bar{\alpha}$, hence:

$$\begin{aligned} P_e^{\text{floor}} &= \mathbb{P}(\,\bar{\alpha} < 0 \,|\, b = 1\,) \\ &= \mathbb{P}(\,\boldsymbol{c}^{\mathsf{T}}(\boldsymbol{c}+\boldsymbol{\xi}) < 0\,) = Q\!\left(\frac{\|\boldsymbol{c}\|}{\sigma_\xi}\right). \end{aligned}$$

The exact symbol error probability is (see eqs. (3.41) and (3.39)):

$$\begin{aligned} P_e &= \mathbb{P}(\,\alpha < 0 \,|\, b = 1\,) \\ &= \mathbb{P}(\sqrt{\mathcal{E}}\|\boldsymbol{c}\|\zeta + n < 0\,) \\ &= \int_{-1}^{1} dz\, Q\!\left(\frac{\sqrt{\mathcal{E}}\,\|\boldsymbol{c}\|\,z}{\sigma_N}\right) \mathrm{f}_\zeta(z; L, \|\boldsymbol{c}\|/\sigma_\xi), \end{aligned} \tag{3.43}$$

where the dependence on L and $\|\boldsymbol{c}\|/\sigma_\xi$ (see eq. (3.42)) is explicited in $\mathrm{f}_\zeta(z; L, \|\boldsymbol{c}\|/\sigma_\xi)$.

Figure 3.6 shows P_e *vs.* $\mathcal{E}/\sigma_N^2$ for different values of $\sigma_\xi^2/\|\boldsymbol{c}\|^2$ and L, and compares the symbol error rate obtained through Monte-Carlo simulations (points with error bars on figure) with the analytical expression of P_e (solid lines) given by eq. (3.43). Observe that, as expected, increased accuracy of the estimation yields to decreased P_e^{floor}. Furthermore, for fixed σ_ξ^2 and $\mathcal{E}/\sigma_N^2$, increased L yields to increased P_e: increasing L by a factor of two implies a loss of approximately 3 dB for $P_e^{\text{floor}} \ll P_e \ll 1$ (see $5 \leq \mathcal{E}/\sigma_N^2 \leq 25$ dB on figure), while the floor does not depend on L, hence same performance is reached for high SNR irrespective of the bandwidth.

3.2.2 Multiuser

In the multiuser setting, although the expression for the probability of error remains as in eq. (3.35), there are three sources of errors: thermal noise, imperfect CSI, and multiuser interference (MUI). In particular, as $\mathcal{E}/\sigma_N^2$ increases, the last two factors both lead to a probability error floor, *i.e.*, $P_e \to P_e^{\text{floor}}(\beta, \sigma_\xi^2) > 0$ as $\mathcal{E}/\sigma_N^2 \to \infty$, being $\beta = K/N$ the load of the system.

Figures 3.7 and 3.8 show the probability of error P_e *vs.* $\mathcal{E}/\sigma_N^2$ (dB) for systems with $\imath = 1$ *vs.* $\imath = 5$, respectively, and for different values of β and σ_ξ^2, $(\beta_1, \beta_2, \beta_3) = (\sigma_1^2, \sigma_2^2, \sigma_3^2) = (0, \frac{1}{20}, \frac{1}{10})$. In particular, in both figures, the left-hand side plot (Fig. 3.7a and 3.8a) refers to TR, while the r.h.s. plot (Fig. 3.7b and 3.8b) refers to AR. Figures show that floors depend on a combination of both imperfect CSI and MUI. For low-SNR, *i.e.*, $\mathcal{E} \ll \sigma_N^2$, P_e is not very sensitive to estimation errors: on figures, systems with same load have similar P_e at low-SNR, and leads to different P_e^{floor} at high-SNR, $\mathcal{E} \gg \sigma_N^2$. On the contrary, for fixed estimation error variance σ_ξ^2, increased β implies increased P_e for any SNR, and contributes to a higher error floor. All figures are obtained by Monte-Carlo simulations of finite-dimensional systems with $N = 200$ chips. By comparing Figs. 3.7 (a) and (b), that both refer to systems with $\imath = 1$, AR is shown to outperform TR when imperfect CSI is the main cause of error, and vice versa when the load is the main cause of error: compare, for example, the two cases (β_3, σ_1) and (β_1, σ_2). A similar behavior can be observed with impulsive systems (Fig. 3.8) with even more emphasis. Figure 3.8 (a) and

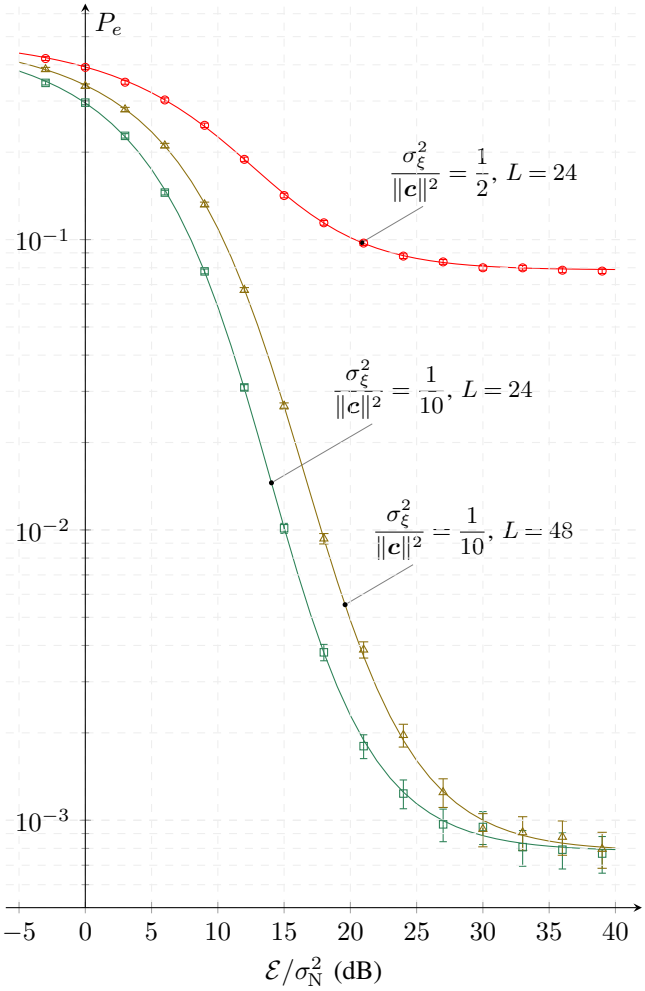
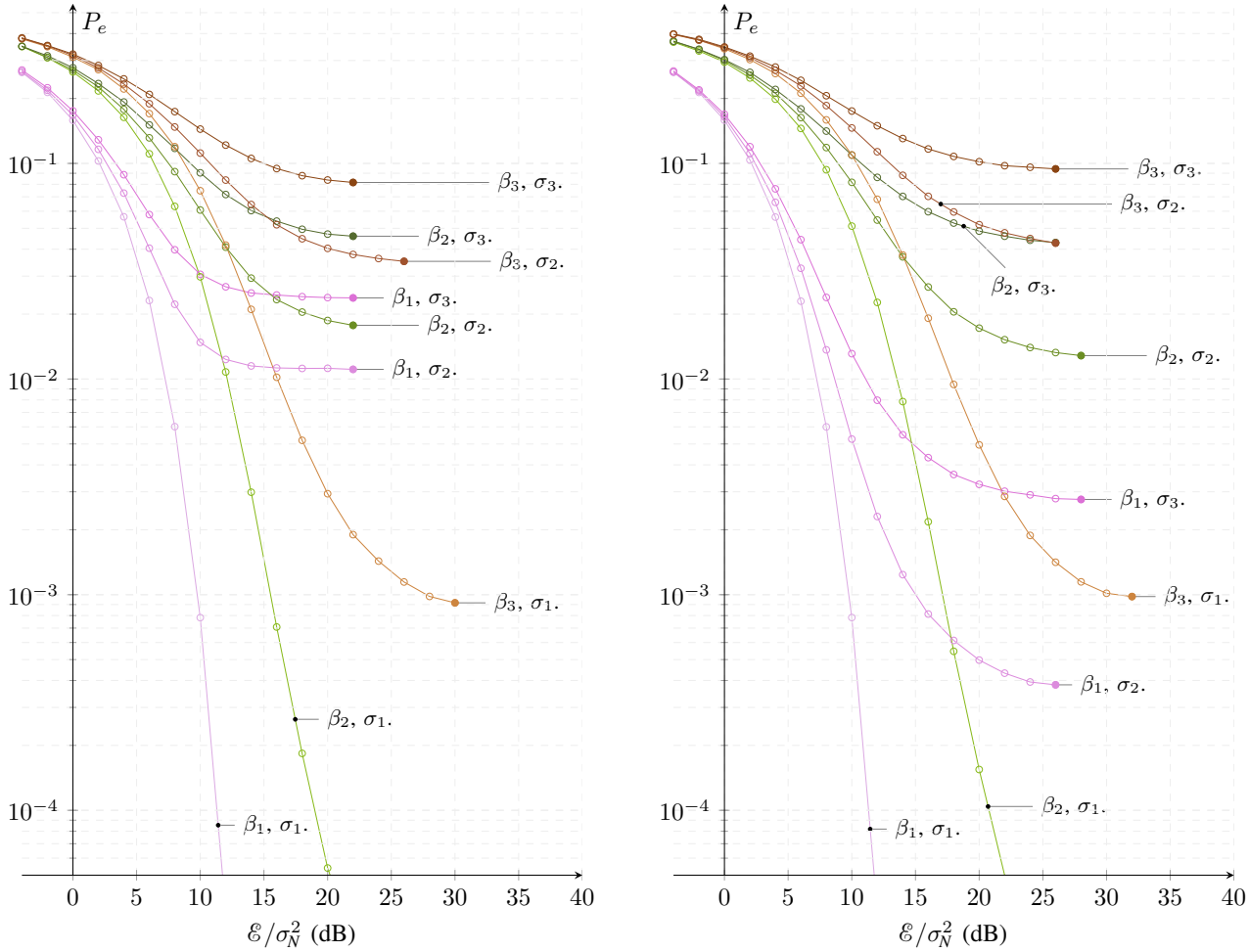


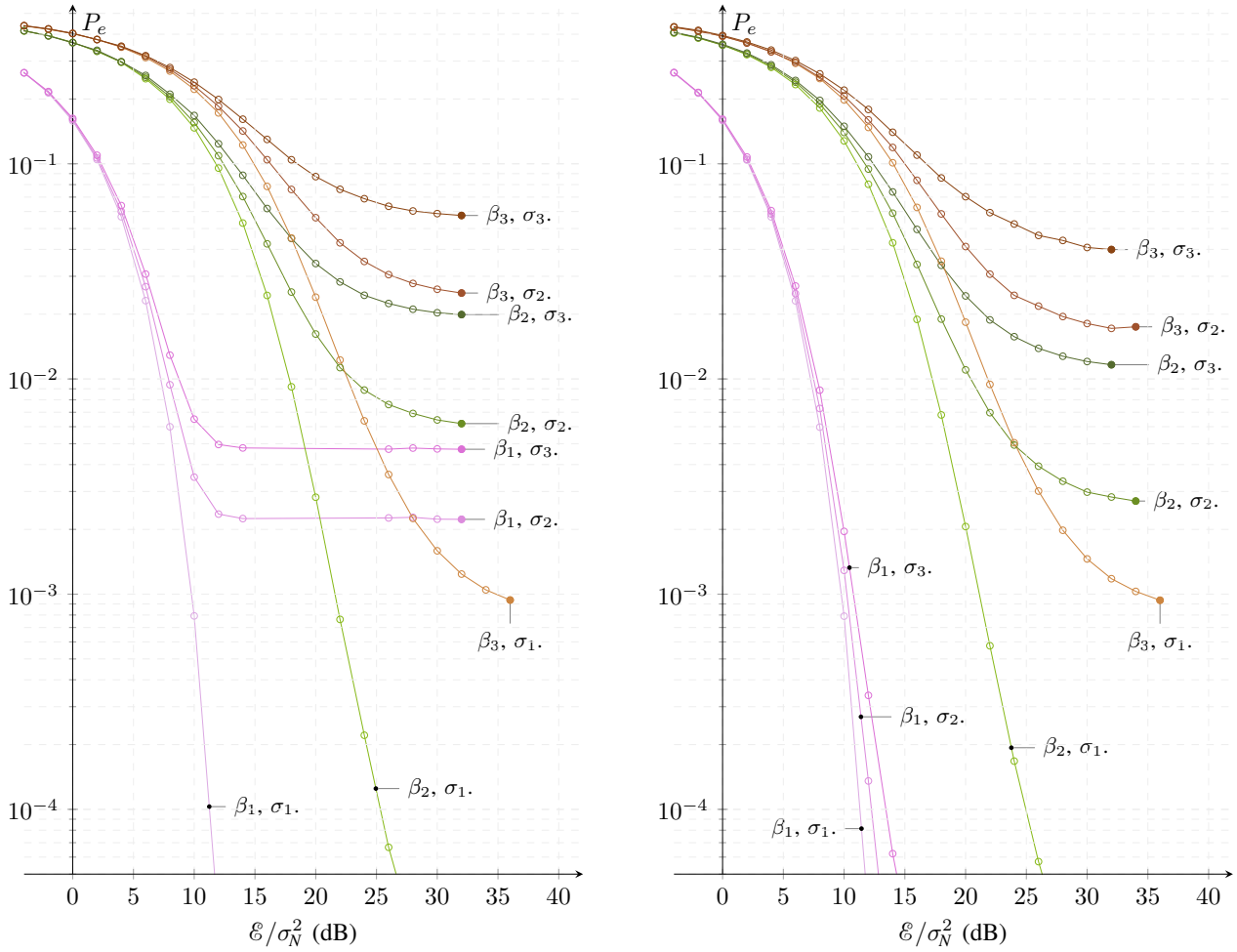
Figure 3.6: Symbol error probability P_e *vs.* $\mathcal{E}/\sigma_N^2$ for di erent values of the number of resolved paths L and the estimation error variance σ_ξ^2.

(b) shows performance of systems with $i = 5$, and indicates that performance of TR is strongly limited by the presence of estimation errors, while when MUI is the limiting cause TR and AR have similar performance (see for example curves with (β_3, σ_1) and (β_1, σ_2)).



(a) TR (b) AR

Figure 3.7: Probability of error P_e *vs.* $\mathcal{E}/\sigma_N^2$ (dB) for systems with $i = 1$ and $(\beta_1, \beta_2, \beta_3) = (\sigma_1^2, \sigma_2^2, \sigma_3^2) = (0, \frac{1}{20}, \frac{1}{10})$. Figure (a) refers to TR while (b) to AR.



(a) TR

(b) AR

Figure 3.8: Probability of error P_e *vs.* $\mathcal{E}/\sigma_N^2$ (dB) for impulsive ($\iota = 5$) systems with $(\beta_1, \beta_2, \beta_3) = (\sigma_1^2, \sigma_2^2, \sigma_3^2) = (0, \frac{1}{20}, \frac{1}{10})$. Figure (a) refers to TR while (b) to AR.

3.3 Mutual information, Sum-Rate, and Spectral Efficiency

In this section, mutual information (3.32) is derived for AR and TR. The other merit figures (3.33) and (3.34) follows directly, although all the elements for a comparison are already included in (3.32).

3.3.1 Derivation of Mutual Information

The decision variable for both the imperfect TR (c.f. eq. (3.21)) and AR (c.f. eq. (3.19)) can be cast in the following form:

$$\hat{z}_k = \hat{a}_{kk} b_k + \sum_{\substack{j=1 \\ j \neq k}}^{K} \hat{a}_{kj} b_j + \hat{\nu}_k = \hat{a}_{kk} b_k + \hat{S}_k + \hat{\nu}_k, \tag{3.44}$$

where $\hat{\nu}_k = n_{q_k} \sim \mathcal{N}(0, \mathcal{N}_0/2)$.

Let specify and give an interpretation of the terms $\hat{a}_{ki}$, $i = 1, \ldots, K$, for both TR and AR.

TR coupling coefficients.

For TR, the term $\hat{a}_{kk}$ is given by:

$$\hat{a}_{kk}^{\mathrm{TR}} = \boldsymbol{e}_{q_k}^{\mathsf{T}} \boldsymbol{C}_k \alpha_k (\boldsymbol{T}_k + \boldsymbol{\Xi}_k) \boldsymbol{x}_k = \boldsymbol{c}_k^{\overleftarrow{\mathsf{T}}} \frac{\boldsymbol{t}_k + \boldsymbol{\xi}_k}{\|\boldsymbol{t}_k + \boldsymbol{\xi}_k\|} = \boldsymbol{c}_k^{\mathsf{T}} \frac{\boldsymbol{c}_k + \boldsymbol{\xi}_k^{\overleftarrow{}}}{\|\boldsymbol{c}_k + \boldsymbol{\xi}_k^{\overleftarrow{}}\|},$$

where $\boldsymbol{v}^{\overleftarrow{}} \in \mathbb{R}^n$ denotes a vector with same components of vector $\boldsymbol{v}$ in reversed order, *i.e.*, $[\boldsymbol{v}^{\overleftarrow{}}]_i = \boldsymbol{v}_{n-i+1}$, $1 \leq i \leq n$. The term $\hat{a}_{kj}$, $j \neq k$, is:

$$\hat{a}_{kj}^{\mathrm{TR}} = \boldsymbol{e}_{q_k}^{\mathsf{T}} \alpha_j \boldsymbol{C}_j (\boldsymbol{T}_j + \boldsymbol{\Xi}_j) \boldsymbol{x}_j = \boldsymbol{e}_{q_k}^{\mathsf{T}} \frac{1}{\|\boldsymbol{t}_j + \boldsymbol{\xi}_j\|} \boldsymbol{C}_j (\boldsymbol{t}'_j + \boldsymbol{\xi}'_j) = \boldsymbol{e}_{q_k}^{\mathsf{T}} \frac{1}{\|\boldsymbol{t}_j + \boldsymbol{\xi}_j\|} (\boldsymbol{\gamma}'_j + \boldsymbol{\zeta}'_j), \tag{3.45}$$

where $\boldsymbol{\gamma}'_i = [\boldsymbol{0}^{\mathsf{T}}_{j_{\boldsymbol{x}_i} \times 1}, \boldsymbol{\gamma}_i^{\mathsf{T}}, \boldsymbol{0}^{\mathsf{T}}_{(N i - j_{\boldsymbol{x}_i}) \times 1}]^{\mathsf{T}}$ being $\boldsymbol{\gamma}_j$ the $(2(L+1)i - 1) \times 1$ autocorrelation sequence of $(c_j[\ell])_{\ell=0}^{Li}$, and, similarly, $\boldsymbol{\zeta}'_i = [\boldsymbol{0}^{\mathsf{T}}_{j_{\boldsymbol{x}_i} \times 1}, \boldsymbol{\zeta}_i^{\mathsf{T}}, \boldsymbol{0}^{\mathsf{T}}_{(N i - j_{\boldsymbol{x}_i}) \times 1}]^{\mathsf{T}}$ with $\boldsymbol{\zeta}_i$ a Gaussian random vector with non-identity correlation.

In order to provide an interpretation of the above expressions, it is useful to start with the case of no estimation error. In general, the decision variable at the output of the matched filter of user k is given by the q_k-th sample of the sum of both intended and interference signals, plus noise. In the special case of no estimation errors, $a_{kk} = \|\boldsymbol{c}_k\|$ is the square root energy of channel k, *i.e.*, the maximum tap of the effective channel, while a_{kj} is either equal to zero if the effective channel of user j, that occupies $2(L+1)i - 1$ out of Ni degrees of freedom in a symbol period, is not present at delay q_k, or to a random resolved path of the effective channel of user j, the randomness owing to random hopping and asynchronism. In presence of estimation errors, $\hat{a}_{kk}$ is smaller than, although in general in the neighbourhood of, the square root energy of channel k due to the mismatch between $\hat{\boldsymbol{h}}_k$ and $\boldsymbol{h}_k$, and

$\hat{a}_{kj}$ is either equal to zero if the *perturbed* effective channel of user j, that occupies $2(L+1)i-1$ out of Ni degrees of freedom in a symbol period, is not present at delay q_k, or equal to a random path of the *perturbed* effective channel of user j, the perturbation owing to the imperfect channel estimation of user j.

AR coupling coefficients.

For AR, the set of coupling coefficients $\{\hat{a}_{ki}\}_{i=1}^{K}$ are:

$$\hat{a}_{kk}^{\mathrm{AR}} = \frac{\hat{\boldsymbol{h}}_k^{\mathsf{T}}}{\|\hat{\boldsymbol{h}}_k\|}\boldsymbol{h}_k = \|\boldsymbol{c}_k\|^2 + \boldsymbol{\xi}_k^{\mathsf{T}}\boldsymbol{c}_k, \tag{3.46}$$

$$\hat{a}_{kj}^{\mathrm{AR}} = \frac{\hat{\boldsymbol{h}}_k^{\mathsf{T}}}{\|\hat{\boldsymbol{h}}_k\|}\boldsymbol{h}_j, \quad j \neq k \tag{3.47}$$

where $\hat{\boldsymbol{h}}_k = \boldsymbol{h}_k + \boldsymbol{\chi}_k$, $\boldsymbol{h}_k = \boldsymbol{C}_k\boldsymbol{x}_k$, and $\boldsymbol{\chi}_k = \boldsymbol{\Xi}_k\boldsymbol{x}_k$. We can think of $\hat{\boldsymbol{h}}_k/\|\hat{\boldsymbol{h}}_k\|$ as the perturbed direction along which the received signal is projected in order to decode user k; $\hat{a}_{kk}^{\mathrm{AR}}$ represents, therefore, the "mismatch" between the perturbed and unperturbed channels of user k; $\hat{a}_{kj}^{\mathrm{AR}}$ represents the coupling between user k, that is perturbed, and another user j. As in the TR case, a channel impulse response occupies a fraction, that is approximately equal to $(L+1)/N$, of the available degrees of freedom in a symbol period; Opposite of the TR case, where the perturbation affects user j in $\hat{a}_{kj}^{\mathrm{TR}}$, user k is perturbed in the AR case (through $\hat{\boldsymbol{h}}_k$), user j appearing with the true channel impulse response h_j.

Derivation.

Being each term in the r.h.s. of eq. (3.44) independent from the other terms, mutual information $I(z_k;b_k)$ can be derived once the distributions of $\hat{a}_{kk}$ and $\hat{S}_k$ are known. The former depends on both the random channel impulse response and estimation errors of user k, and the latter on the random channel impulse responses and estimation errors, and the random delays with respect to user k. Hence, the final form assumed by $I(z_k;b_k)$ strongly depends on the channel model; however, in the following, the effect of the time-hopping and random asynchronism will be enucleated, without enter in the computation of a mutual information when a particular channel model is adopted; this last task is addressed by simulations, where the IEEE 802.15.3a model [53], that is valid for bandwidths up to several gigahertz, is selected.

As for TR, since the effective channel of user j occupies a fraction $f = (2(L+1)i-1)/(Ni)$, and user k, due to the assumptions on independence and uniformity of hopping codes and asynchronism, selects uniformly at random one of the Ni samples available per symbol period, then a_{kj}, $j \neq k$, is equal to zero with probability $1-f$, and is distributed as the generic path of the effective channel h_j with probability f, that is:

$$P_{a_{kj}^{\mathrm{TR}}} = (1-f)\delta_0 + fP^{\mathrm{TR}}, \tag{3.48}$$

where P^{TR} indicates the distribution of the generic path of the effective channel of user j (that is independent of j). In presence of estimation errors, the above argument holds, that is, $\hat{a}_{kj}$, $j \neq k$, is equal to zero with probability $1-f$, and is

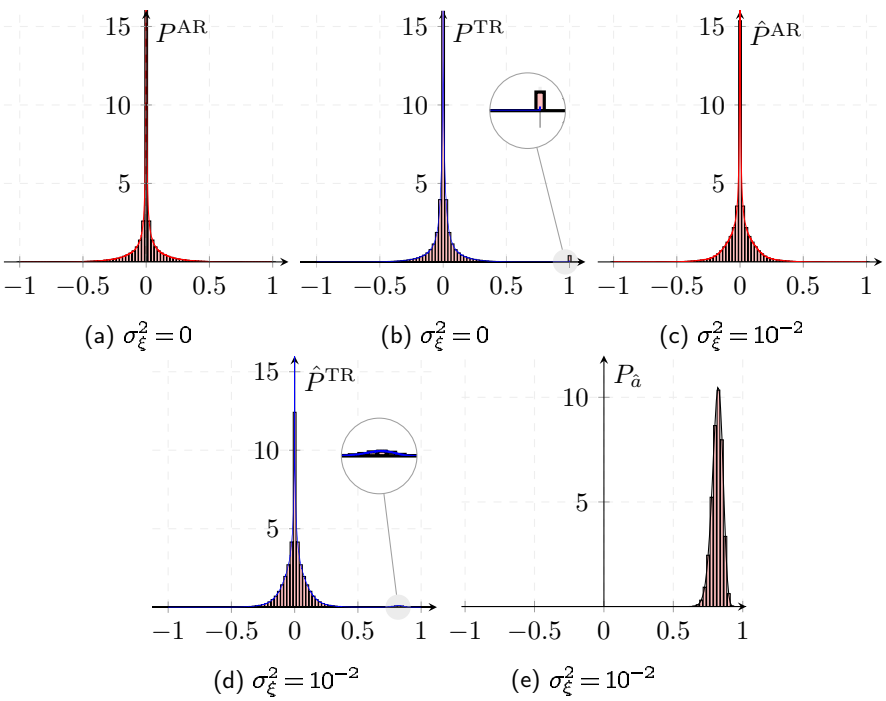


Figure 3.9: PDFs $\hat{P}$ and $P_{\hat{a}}$ of the variables $\hat{a}_{kj}$, $j \neq k$, and $\hat{a}_{kk}$, respectively. Distributions do not depend on the particular user k. Figs. *(a)* and *(b)* correspond to a variable $\hat{a}_{kj}$, $j \neq k$, in systems without estimation errors; Figs. *(c)* and *(d)* correspond to a variable $\hat{a}_{kj}$, $j \neq k$, in systems with estimation errors with variance $\sigma_\xi^2 = 0.01$; Fig. *(e)* corresponds to a variable $\hat{a}_{kk}$ in systems with a same error variance.

distributed as the generic path of the *perturbed* effective channel $\hat{\boldsymbol{h}}_j$ with probability f:

$$P_{\hat{a}_{kj}^{\mathrm{TR}}} = (1-f)\delta_0 + f\hat{P}^{\mathrm{TR}}, \tag{3.49}$$

where $\hat{P}^{\mathrm{TR}}$ indicates the distribution of the generic path of the *perturbed* effective channel of user j (that is independent of j).

As for AR, let start by finding the distribution of a_{kj}, $j \neq k$, *i.e.*, the coupling coefficient between two users in absence of estimation error. Each channel spans a subspace of dimension $(L+1)\imath$ in a space with $(N+L+1)\imath - 1$ dimensions[1]; in other words, just $(L+1)\imath$ entries of $\boldsymbol{h}_i$ are nonzero. From the hypotheses of independence and uniformity of delays due to asynchronism between users and time-hopping codes, there exists a probability f such that the inner product $\boldsymbol{h}_k^{\mathsf{T}}\boldsymbol{h}_j$ is nonzero, and the remaining probability $1-f$ that the inner product is zero. We may think of the "nonzero event" as the partial overlapping between two channels. As $N \gg L$, it results $f \approx (2(L+1)\imath - 1)/(N\imath)$, where the assumption $N \gg L$ allows to neglect border effects. Indicating with P^{AR} the distribution of a_{kj}^{AR} conditioned on the nonzero event, one has:

$$P_{a_{kj}^{\mathrm{AR}}} = (1-f)\delta_0 + fP^{\mathrm{AR}}. \tag{3.50}$$

In presence of estimation errors, the above discussion remains valid, since an error $\boldsymbol{\chi}_k$ changes, in general, the direction of vector $\hat{\boldsymbol{h}}_k$ with respect to $\boldsymbol{h}_k$, *i.e.*, $\hat{\boldsymbol{h}}_k$ and h_k are, in general, not collinear, but it does not change the subspace spanned by the two channels, *i.e.*, the subspace spanned by the true channel is equal to the subspace spanned by the perturbed channel. Indicating with $\hat{P}^{\mathrm{AR}}$ the distribution of $\hat{a}_{kj}^{\mathrm{AR}}$ conditioned on the nonzero event, one has:

$$P_{\hat{a}_{kj}^{\mathrm{AR}}} = (1-f)\delta_0 + f\hat{P}^{\mathrm{AR}}. \tag{3.51}$$

$\hat{P}^{\mathrm{AR}}$ reduces to P^{AR} when the estimation error is nil.

Figure 3.9 shows the distributions P^{TR}, $\hat{P}^{\mathrm{TR}}$, P^{AR} and $\hat{P}^{\mathrm{AR}}$ in eqs. (3.48), (3.49), (3.50) and (3.51), respectively, and the distribution $P_{\hat{a}}$ of the term $\hat{a}_{kk}$, assuming the Channel Model 1 (CM1) specified in the IEEE 802.15.3a standard. All simulations refer to a system with fixed chip duration $T_{\mathrm{c}} = 1$ ns and bandwidth $\mathrm{W} = 1/T_{\mathrm{c}}$. Power control is assumed; in particular, $\|\boldsymbol{c}_k\|^2 = 1$ for all users, $k = 1, \ldots, K$. The delay spread of each channel impulse response is fixed at a value $T_d = 50$ ns that includes most of the energy of typical CM1 channels. For a given bandwidth W, the length of the channel expressed in number of samples per channel is $T_d\mathrm{W}$, *i.e.*, $\boldsymbol{c}_k$ is a $(\lfloor T_d\mathrm{W} \rfloor + 1) \times 1$ vector.

Figures 3.9 *(a) vs. (b)* show the distributions of the coupling coefficient a_{kj}, $j \neq k$, in case of no estimation errors, for AR and TR, respectively. As may be expected, the variance of the latter is larger than the variance of the former, as follows from the property of time reversal to increase the total energy of the effective channel; In the specific case, the $\mathrm{Var}[a_{kj}^{\mathrm{AR}}] \approx 0.0099$ while $\mathrm{Var}[a_{kj}^{\mathrm{TR}}] \approx 0.0173$. In Fig. 3.9 *(b)*

[1] The number of degrees of freedom in a symbol period is $N\imath$; in the large system limit, as $N \gg L$, the difference between $(N+L+1)\imath - 1$ and $N\imath$ due to the convolution is negligible.

it is highlighted the presence of a strong interference ($a_{kj}^{\mathrm{TR}}=1$) that is not present in the AR case; in the TR case, there is, indeed, among the $2(L+1)\imath-1$ paths of the effective channel, one path with amplitude equal to the square-root energy of the channel, $\|\boldsymbol{c}_k\|=1$, that is, therefore, selected with probability $1/(2(L+1)\imath-1)$. P^{TR} is also more leptokurtic than P^{AR}, showing a kurtosis approximately equal to 34 *vs.* 12 reached by the AR variable. Figures 3.9 *(c) vs. (d)* show the distributions of the coupling coefficient $\hat{a}_{kj}$, $j\neq k$, in case of estimation error with per sample variance $\sigma_\xi^2=0.01$, for AR and TR, respectively. Both variance and kurtosis of TR are still larger than those of AR; in particular, it results $\mathrm{Var}[a_{kj}^{\mathrm{AR}}]\approx 0.0099$ *vs.* $\mathrm{Var}[a_{kj}^{\mathrm{TR}}]\approx 0.0149$, and $\mathrm{kurt}[a_{kj}^{\mathrm{AR}}]\approx 6.9$ *vs.* $\mathrm{kurt}[a_{kj}^{\mathrm{TR}}]\approx 21.7$. Figure 3.9 *(e)* shows the distribution of the term $\hat{a}_{kk}$, that is the same for both TR and AR, and is represented for $\sigma_\xi^2=0.01$. Note that all variables $\hat{a}_{ki}$ can be regarded as normalized cross-correlations, hence $-1\leq\hat{a}_{ki}\leq 1$.

In terms of c.fs., eqs. (3.51) and (3.49) becomes:

$$\varphi_{\hat{a}_{kj}}(u)=\mathbb{E}\left[\,e^{ju\alpha}\,\right]=1-f(1-\hat{\varphi}(u)),\qquad \hat{\varphi}(u)=\int_{\mathbb{R}}d\alpha\,e^{j\alpha u}\hat{P}(\alpha),$$

being $\hat{P}$ equal to either $\hat{P}^{\mathrm{AR}}$ or $\hat{P}^{\mathrm{TR}}$ in eqs. (3.51) and (3.49), respectively. In general, given two independent r.*vs.* X and Y and their product $Z=XY$, it results $\varphi_Z(u)=\mathbb{E}[\,\varphi_Y(uX)\,]=\mathbb{E}[\,\varphi_X(uY)\,]$; therefore, the r.v. $\hat{a}_{kj}b_j$ has c.f.:

$$\varphi_{\hat{a}_{kj}b_j}(u)=\mathbb{E}\left[\,\varphi_{\hat{a}_{kj}}(b_ju)\,\right]=1-f(1-\mathbb{E}[\,\hat{\varphi}(b_ju)\,])=1-f(1-\overline{\varphi}(u)),$$

where the expectation is with respect $b_j\sim\mathcal{N}(0,\mathcal{E})$, and $\overline{\varphi}(u)$ is independent of j. Since $\{\hat{a}_{kj}b_j\}$ with $j\in\{1,\ldots,K\}\setminus k$ are independent, then $\hat{S}_k$ has c.f.:

$$\varphi_{\hat{S}_k}(u)=\left\{1-\frac{2(L+1)\imath-1}{N\imath}(1-\overline{\varphi}(u))\right\}^{K-1},$$

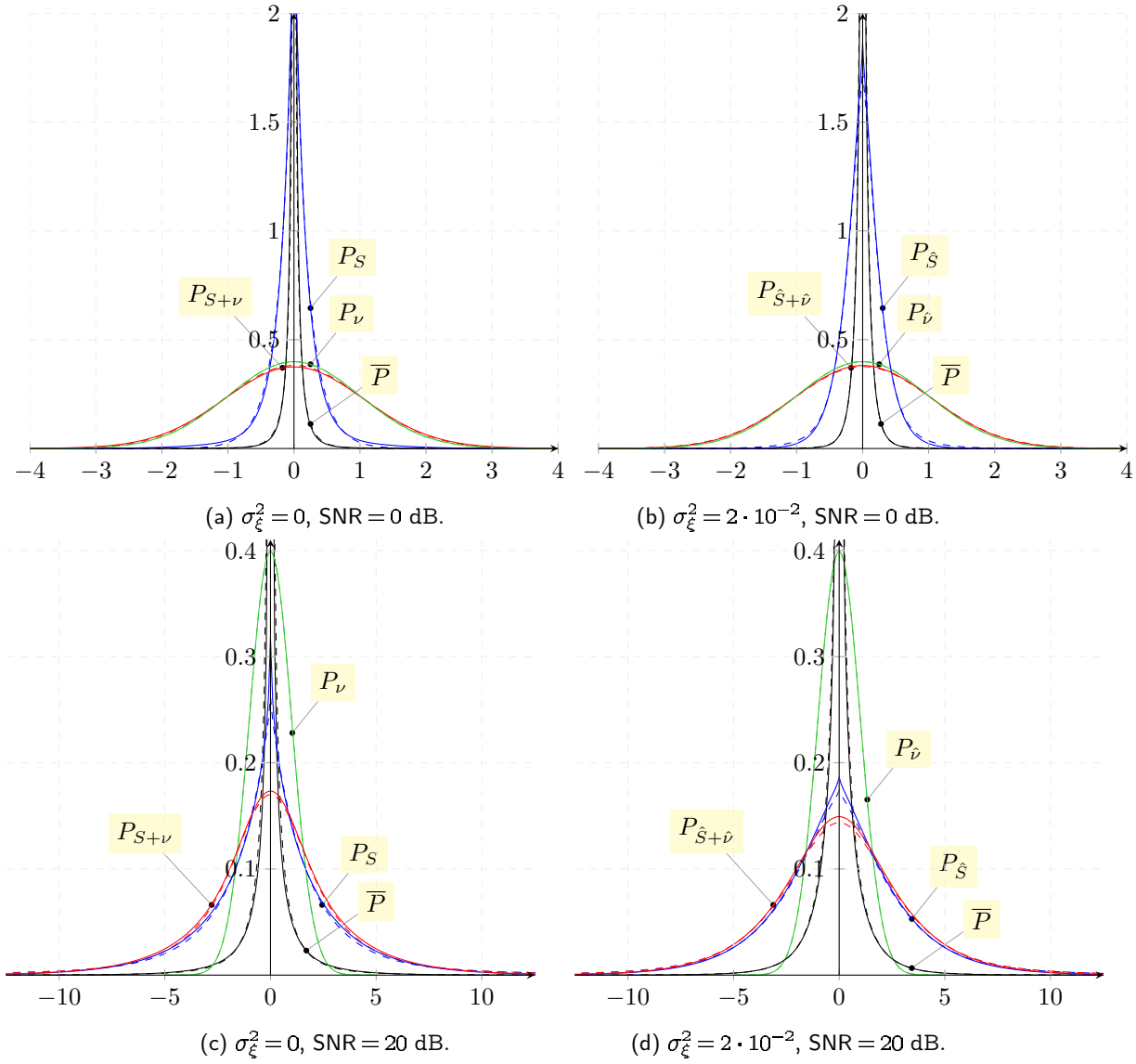
that, in the large system limit, where $K\to\infty$, $N\to\infty$, $K/N\to\beta$, converges to:

$$\varphi_{\hat{S}_k}(u)=e^{-\beta_{\mathrm{eff}}[1-\overline{\varphi}(u)]}=\sum_{r\geq 0}\frac{\beta_{\mathrm{eff}}^r}{r!}e^{-\beta_{\mathrm{eff}}}[\overline{\varphi}(u)]^r,\tag{3.52}$$

where $\beta_{\mathrm{eff}}=\beta(2(L+1)\imath-1)/\imath$ is the effective load; without multipath ($L=1$) and one pulse per chip ($\imath=1$), β_{eff} reduces to the usual load β as given by K/N. The interference-plus-noise variable has thus c.f. given by:

$$\varphi_{\hat{S}_k+\hat{\nu}_k}(u)=\varphi_{\hat{S}_k}(u)\varphi_{\hat{\nu}_k}(u)=e^{-\beta_{\mathrm{eff}}[1-\overline{\varphi}(u)]-\frac{\sigma_N^2}{2}u^2}.$$

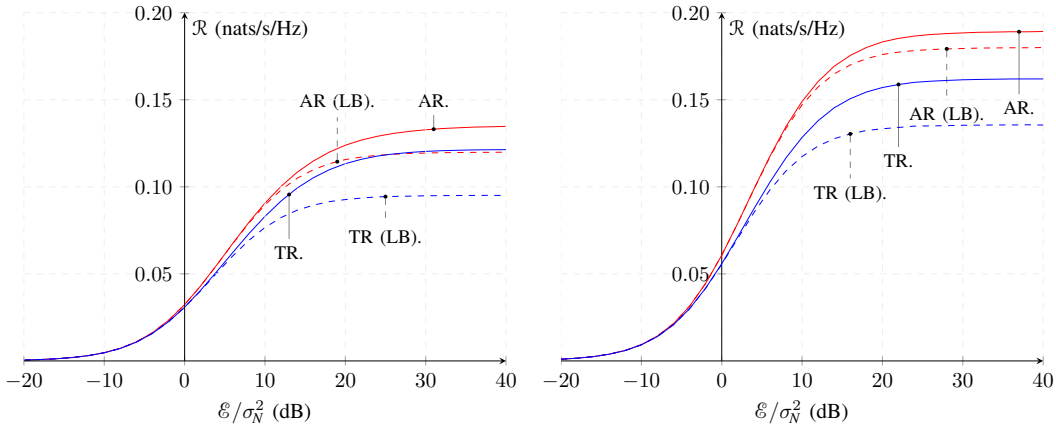
Figure 3.10 shows distributions $\overline{P}$, $P_{\hat{S}}$, $P_{\hat{\nu}}$, and $P_{\hat{S}+\hat{\nu}}$, that correspond to c.fs. $\overline{\varphi}$, $\varphi_{\hat{S}}$, $\varphi_{\hat{\nu}}$, and $\varphi_{\hat{S}+\hat{\nu}}$ defined above, for different values of σ_ξ^2 and $\mathrm{SNR}=\mathcal{E}/\sigma_N^2$, and fixed value of load $\beta=0.1$. Figures 3.10 *(a)* and *(b)* show a noise-limited scenario, where $\mathrm{SNR}=0$ dB, without and with estimation errors, respectively: simulations show that the interference-plus-noise variable is not significantly affected by estimation errors (c.f. $P_{S+\nu}$ and $P_{\hat{S}+\hat{\nu}}$ curves), although P_S, as well as $\overline{P}$,



(a) $\sigma_\xi^2 = 0$, SNR $= 0$ dB. (b) $\sigma_\xi^2 = 2 \cdot 10^{-2}$, SNR $= 0$ dB.

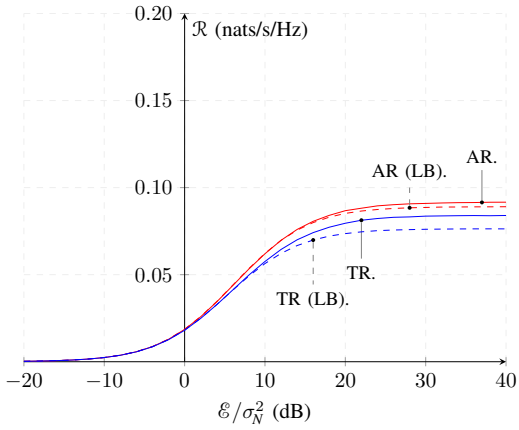
(c) $\sigma_\xi^2 = 0$, SNR $= 20$ dB. (d) $\sigma_\xi^2 = 2 \cdot 10^{-2}$, SNR $= 20$ dB.

Figure 3.10: PDF $\overline{P}$ of interfering terms $\hat{a}_{kj} b_j$, $j \neq k$, given $\hat{a}_{kj} \neq 0$, P_ν of the noise variable $\hat{\nu}_k$, $P_{\hat{S}}$ of interference $\hat{S}_k$, and $P_{\hat{S}+\hat{\nu}}$ of interference-plus-noise $\hat{S}_k + \hat{\nu}_k$. Error variance σ_ξ^2 and SNR $= \mathcal{E}/\sigma_N^2$ are specified below each subfigure. The noise variance σ_N^2 is normalized to 1. The load β of the network is fixed and equal to $\beta = 0.1$. Solid *vs.* dashed curves refer to TR *vs.* AR, respectively.

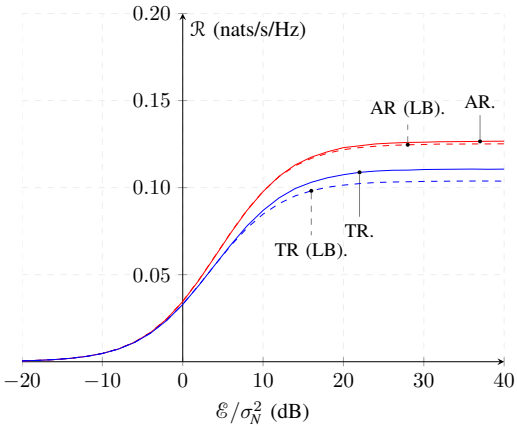


(a) $\imath = 1$, $\beta = 0.1$, $\sigma_\xi^2 = 0$

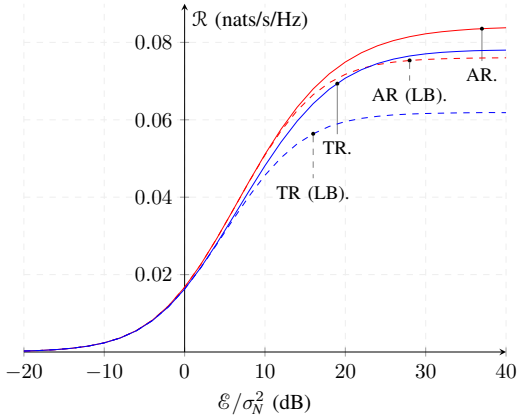
(b) $\imath = 1$, $\beta = 0.2$, $\sigma_\xi^2 = 0$



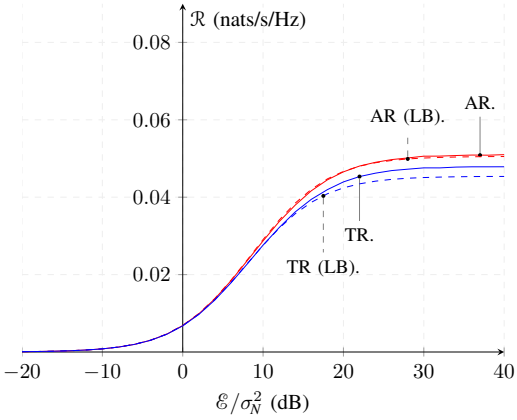
(c) $\imath = 1$, $\beta = 0.1$, $\sigma_\xi^2 = 2 \cdot 10^{-2}$



(d) $\imath = 1$, $\beta = 0.2$, $\sigma_\xi^2 = 2 \cdot 10^{-2}$



(e) $\imath = 2$, $\beta = 0.1$, $\sigma_\xi^2 = 0$



(f) $\imath = 2$, $\beta = 0.1$, $\sigma_\xi^2 = 2 \cdot 10^{-2}$

Figure 3.11: Spectral efficiency $\mathcal{R}$ (solid curve) and lower bound assuming Gaussian interference (dashed curve) *vs.* SNR $= \mathcal{E}/\sigma_N^2$ (dB) for di erent values of load β and error variance σ_ξ^2. Impulsiveness index is fixed to $\imath = 1$ in Figs. from (a) to (d), and to $\imath = 2$ in Figs. (e) and (f).

becomes less leptokurtic in presence of estimation errors. Figures 3.10 *(a)* and *(b)* show an interference-limited scenario, where SNR $= 20$ dB: in this case P_S is far from Gaussian, and so is the interference-plus-noise PDF $P_{S+\nu}$; the effect of the estimation error is to decrease the kurtosis of P_S, and so that of $P_{S+\nu}$.

Knowing the distribution of z_k conditioned on b_k, or equivalently its c.f., mutual information $I(z_k;b_k) = h(z_k) - h(z_k|b_k)$ follows directly. Hence, the c.f. of z_k given b_k is:

$$\varphi_{z_k|b_k}(u) = \varphi_{\hat{S}_k+\hat{\nu}_k}(u)\varphi_{\hat{a}_{kk}}(b_k u);$$

hence, the c.f. of $\hat{z}_k$ is:

$$\varphi_{\hat{z}_k}(u) = \mathbb{E}\left[\,\varphi_{\hat{z}_k|b_k}(u)\,\right] = \varphi_{\hat{S}_k+\hat{\nu}_k}(u)\,\mathbb{E}\left[\,\varphi_{\hat{a}_{kk}}(b_k u)\,\right],$$

where the expectation is over $b_k \sim \mathcal{N}(0,\mathcal{E})$. Explicitly, one has:

$$h(\hat{z}_k) = \int_{\mathbb{R}} -P_{\hat{z}_k}(z)\ln P_{\hat{z}_k}(z)dz, \quad P_{\hat{z}_k}(z) = \frac{1}{2\pi}\int_{\mathbb{R}} du\, e^{-juz}\varphi_{\hat{z}_k}(u),$$

and:

$$h(\hat{z}_k|b_k) = \int_{\mathbb{R}} \Phi_{0,\mathcal{E}}(b) h(\hat{a}_{kk}b + \hat{S}_k + \hat{\nu}_k) = \int_{\mathbb{R}} \Phi_{0,\mathcal{E}}(b)\int_{\mathbb{R}} -P_{\hat{z}_k|b_k=b}(z)\ln P_{\hat{z}_k|b_k=b}(z)dz,$$

where $\Phi_{0,\mathcal{E}}$ denotes a Gaussian distribution with zero mean and variance $\mathcal{E}$, and:

$$P_{\hat{z}_k|b_k=b}(z) = \frac{1}{2\pi}\int_{\mathbb{R}} du\, e^{-juz}\varphi_{\hat{z}_k|b_k=b}(u).$$

The above derivation allows to find $I(\hat{z}_k;\hat{b}_k) = h(\hat{z}_k) - h(\hat{z}_k|b_k)$ as a function of the distribution $\hat{P}$ of $\hat{a}_{kj}$, $j \neq k$, and the distribution $P_{\hat{a}_{kk}}$ accounting for the loss of correlation incurred by the user to be decoded because of the estimation error. Both $\hat{P}$ and $P_{\hat{a}_{kk}}$ accounts for the channel model and the estimation error, in particular its variance.

As baseline comparison, we also provide the following lower bound $\underline{I}(\hat{z}_k;b_k)$ for $I(\hat{z}_k;b_k)$, that is achieved when the interference is Gaussian:

$$I(\hat{z}_k;b_k) \geq \underline{I}(\hat{z}_k;b_k) = \mathbb{E}\left[\frac{1}{2}\ln\left(1 + \frac{\mathcal{E}\,\hat{a}_{kk}^2}{\mathrm{Var}[\,\hat{S}_k\,] + \mathrm{Var}[\,\hat{\nu}_k\,]}\right)\right],$$

where the expectation is over $\hat{a}_{kk}$. The corresponding spectral efficiency lower bound is $\underline{\mathcal{R}} = (\beta/\imath)\underline{I}(\hat{z}_k;b_k)$ (c.f. eq. (3.34)).

Results are shown on Fig. 3.11, where $\mathcal{R}$ (solid curve) and $\underline{\mathcal{R}}$ (dashed curve) are presented as a function of $\mathcal{E}/\sigma_N^2$, for different values of β and σ_ξ^2. The receiver structure shows a mutual information floor at high SNR. By comparing Figs. 3.11 (a) and (b), one observes that $\mathcal{R}$ increases sublinearly as β increases, while by comparing Figs. 3.11 (a) and (c), or Figs. 3.11 (b) and (d), a reduction in spectral efficiency due to the presence of an estimation error is observed. $\mathcal{R}$ scales with $\imath$ as shown on Figs. 3.11 (e) and (f).

In each of these simulations, AR outperforms TR. However, note that the gap $\mathcal{R} - \underline{\mathcal{R}}$ may be viewed as a measure of the non-Gaussianity of $\hat{z}_k$ and $\hat{z}_k|b_k$, and is indeed higher in the TR case with respect to the AR case, because of the different distribution of the interference term, that is more leptokurtic in the TR case. There could be, therefore, a room for TR to outperform AR.

3.4 Future work

In order to generalize the analysis developed in this chapter, the following two assumptions may be relaxed:

1. single user detection: different receiver structures that use multiuser detection may be considered, in first place linear multiuser detection (RZF, MMSE), that can be obtained by leaving eq. (3.29) in the general form;

2. unit gain of channel impulse responses: a complete characterization of the wireless channel should include fading/shadowing.

The first generalization can proceed by projecting eq. (3.17) onto a filter as specified by eq. (3.29) in the general form. Results are envisioned to be very different with linear receivers that account for the interference, such as RZF and MMSE, in constrast to the simple MF bank that is used above.

The second generalization can be obtained by removing the constraint $||\boldsymbol{c}_k|| = 1$ for each and every k, that also implies that power transmitted by all users is identical. In this case, performance is not envisioned to differ qualitatively from the present analysis.

For both generalizations, a comparison of transceiver structures and fading statistics with different degrees of sparsity of the signal transmitted by each user, *e.g.* selecting DS-CDMA *vs.* TH-CDMA, will complete the above analysis.

Furthermore, the mutual information analysis can be extended by considering the maximum net ergodic rate,

$$R_{\text{net}} = \max_{N_t^{\text{DL}}} \left(1 - \frac{N_t^{\text{DL}}}{nN}\right) R,$$

where the factor in paretheses, namely $\rho \triangleq 1 - \frac{N_t^{\text{DL}}}{nN}$, accounts for the data transmission fraction of time, and R is the rate achievable during that period. Since the latter depends on the former, *i.e.*, $R = R(\rho)$, a maximum with respect to ρ is expected, as also suggested in [80] and confirmed by preliminary investigations. The analysis of R_{net} is important from a practical viewpoint since it suggests the optimal duration of training phases that maximize the data rate.

3.5 Conclusions

In this chapter, the problem of characterizing system performance for single antenna systems using time reversal in the case of imperfect channel estimation was addressed. The analyzed setting included one BS and several UTs, and the uplink communication channel was considered in the investigation. Each UT adopted impulse-radio ultra-wideband communication with prefiltering, and the receiving BS adopted a 1Rake; degrees of impulsiveness were reflected by an impulsiveness index that ranges from $\imath = 1$ to $\imath \to \infty$ for ideal impulsiveness. In order to evaluate time reversal behavior, this communication scheme was compared against a reference configuration with no prefiltering and AR at the receiver. The effect of imperfect channel estimation on both transceiver configurations was analyzed. Channel estimation error was modeled as an additive Gaussian noise based on the output of a training phase that was used to tune transmitter and receiver structures. The comparison was performed for both the single user channel and the multiuser channel with power control. Modeling of the channels was obtained based on the 802.15.3a CM1 model. The two communication schemes, TR and AR, were compared based on two different performance parameters: probability of error and mutual information as a function of signal-to-noise ratio.

Results highlighted that, for the single user channel, probability of error for TR and AR coincided, while for the multiuser channel, AR outperformed TR when imperfect CSI was the main cause of error, and the two schemes had similar performance when the load, as measured by the ratio between the number of terminals K and the number of chips N in a symbol period, $\beta = K/N$, was the main cause of error, irrespectively of the degree of impulsiveness.

In terms of spectral efficiency, we provided lower and upper bound expressions, and analyzed the two structures with different impulsiveness index $\imath$ and load β. Results expressed by spectral efficiency $\mathcal{R}$ (nats/s/Hz) as a function of signal-to-noise ratio indicated that, for low-SNR, $\mathcal{R}$ was similar for the two systems, while for higher SNR values, AR outperforms TR. However, remind that, in practical scenarios, it would be simpler for a TR system to acquire a better estimation of the channel with respect to an AR system, since the estimation error variance depends on the energy of the training sequence only, and in the TR case the training sequence is transmitted by a basestation rather than a device, and may require weaker energy consumption constraints. Furthermore, in the presence of estimation errors, a reduction in $\mathcal{R}$ was observed, due to both a mismatch with the user to decode, and a reduced kurtosis of the interference term.

Both generalizations and extensions of the analysis developed in this chapter are presented.

Appendix

3.A Derivation of the PDF of ζ and ζ^2

1) PDF of ζ.

An orthogonal transformation on $\boldsymbol{\xi}$ greatly simplifies the expression of ζ. We can think of $\boldsymbol{c}$ as the $(L+1)$-tuple of coordinates of a vector in $\mathbb{R}^{L+1}$ with respect to the canonical basis $\mathcal{B}$. A different orthonormal basis $\mathcal{B}'$ such that only the first coordinate of the vector is non-zero can be found, for example via the Gram-Schmidt orthogonalization. By denoting with $\boldsymbol{c}'$ the coordinates of the vector with respect to $\mathcal{B}'$, it results $\boldsymbol{c}' = (c'_1, 0, \dots, 0)^\mathsf{T}$. For convenience, we choose the first vector of $\mathcal{B}'$ as $\boldsymbol{c}/\|\boldsymbol{c}\|$, hence $c'_1 = \|\boldsymbol{c}\|$.

Call $\boldsymbol{Q}$ the matrix that changes coordinates from $\mathcal{B}$ to $\mathcal{B}'$; then $\boldsymbol{c}' = \boldsymbol{Q}\boldsymbol{c}$. It is a well-known result that $\boldsymbol{Q}$, that is the matrix that changes coordinates of vectors between two orthogonal bases, is an *orthogonal* matrix, *i.e.*, $\boldsymbol{Q}^{-1} = \boldsymbol{Q}^\mathsf{T}$. As a consequence, $\boldsymbol{Q}$ is also an *isometry*, that is, vectors transformed under the action of $\boldsymbol{Q}$ do not change their norm: $\|\boldsymbol{c}'\| = \|\boldsymbol{c}\|$.

We can rewrite ζ as follows:

$$\zeta = \frac{\boldsymbol{c}^\mathsf{T}(\boldsymbol{c}+\boldsymbol{\xi})}{\|\boldsymbol{c}\| \cdot \|\boldsymbol{c}+\boldsymbol{\xi}\|} = \frac{\boldsymbol{c}^\mathsf{T}\boldsymbol{Q}^\mathsf{T}\boldsymbol{Q}(\boldsymbol{c}+\boldsymbol{\xi})}{\|\boldsymbol{Q}\boldsymbol{c}\| \cdot \|\boldsymbol{Q}(\boldsymbol{c}+\boldsymbol{\xi})\|} = \frac{\boldsymbol{c}'^\mathsf{T}}{\|\boldsymbol{c}'\|} \cdot \frac{\boldsymbol{c}'+\boldsymbol{\xi}'}{\|\boldsymbol{c}'+\boldsymbol{\xi}'\|} = \frac{c'_1+\xi'_1}{\|\boldsymbol{c}'+\boldsymbol{\xi}'\|} \,.$$

Since $\boldsymbol{c}'/\|\boldsymbol{c}'\| = [1, 0, \dots, 0]^\mathsf{T}$, one has:

$$\|\boldsymbol{c}'+\boldsymbol{\xi}'\| = \sqrt{(c'_1+\xi'_1)^2 + \xi'^2_2 + \cdots + \xi'^2_{L+1}} = \sqrt{(c'_1+\xi'_1)^2 + \|\boldsymbol{\xi}'_{-1}\|^2} \,,$$

where $\boldsymbol{\xi}'_{-1} := (\xi'_2, \dots, \xi'_{L+1})^\mathsf{T}$, being $\{\xi'_k \colon k = 2, \dots, L+1\}$ a set of i.i.d. Gaussian r.*vs.* Define:

$$x := \frac{c'_1+\xi'_1}{\sigma_\xi} \sim \mathcal{N}(c'_1/\sigma_\xi, 1)$$

and:

$$y := \frac{\|\boldsymbol{\xi}'_{-1}\|}{\sigma_\xi} \sim \chi_L \,;$$

hence ζ is as follows:

$$\zeta = \frac{x}{\sqrt{x^2+y^2}} = \frac{x/y}{\sqrt{1+(x/y)^2}}, \tag{3.53}$$

therefore ζ depends only on the ratio x/y. It is useful to define the following variable:

$$t := \frac{x}{\frac{1}{\sqrt{L}}y} \; ; \tag{3.54}$$

in fact, t is distributed according to a known distribution, that is the non-central Student $\mathcal{T}$-distribution with L degrees of freedom and non-central parameter $\sqrt{\lambda} := c_1'/\sigma_\xi = \|\boldsymbol{c}\|/\sigma_\xi$ [84], that we denote by $\mathcal{T}'_L(\sqrt{\lambda})$ and has the following canonical form:

$$\mathrm{f}_{\mathcal{T}'_L(\sqrt{\lambda})}(t) = \frac{2^L e^{-\frac{\lambda}{2}} L^{1+\frac{L}{2}}}{\pi (t^2+L)^{\frac{L+1}{2}}} \Gamma\left(\frac{L+1}{2}\right) H_{-1-L}\left(-\sqrt{\frac{\lambda}{2}}\frac{t}{\sqrt{t^2+L}}\right) \; ,$$

where $H_n(x)$ is the Hermite polynomial [84]. Then, since from eqs. (3.53) and (3.54), it results:

$$\zeta = \frac{t}{\sqrt{L+t^2}},$$

the pdf of ζ is obtained by a change of variables[2] as follows:

$$\mathrm{f}_\zeta(z) = \mathrm{f}_{\mathcal{T}'_L(\sqrt{\lambda})}\left(\frac{\sqrt{L}}{\sqrt{1-z^2}}z\right)\frac{\sqrt{L}}{\sqrt{(1-z^2)^3}} \; , \quad |z| \leq 1 \; .$$

2) PDF of ζ^2.

Write ζ^2 as follows:

$$\begin{aligned}
\zeta^2 &= 1 - \frac{\|\boldsymbol{\xi}'_{-1}\|^2}{(c_1'+\xi_1')^2 + \|\boldsymbol{\xi}'_{-1}\|^2} \\
&= 1 - \frac{1}{1+\dfrac{(c_1'+\xi_1')^2}{\|\boldsymbol{\xi}'_{-1}\|^2}} \\
&= 1 - \frac{1}{1+\dfrac{1}{L}\dfrac{\left(\frac{c_1'}{\sigma_\xi}+\frac{\xi_1'}{\sigma_\xi}\right)^2}{\frac{1}{L}\left\|\frac{\boldsymbol{\xi}'_{-1}}{\sigma_\xi}\right\|^2}} = 1 - \frac{1}{1+\dfrac{1}{L}\dfrac{x^2}{y^2/L}} \\
&= 1 - \frac{1}{1+\frac{1}{L}\Psi},
\end{aligned}$$

having defined $\Psi = \frac{x^2}{y^2/L}$. The pdf of ζ^2 can be traced back to a known distribution. In fact, since $x^2 \sim \chi'^2_1(c_1'^2/\sigma_\xi^2)$ and $y^2 \sim \chi^2_L$, ζ^2 depends on the ratio of two independent chi-square distributions. It is known as (non-central) $\mathcal{F}$ (ratio) *distribution* the pdf that describes the ratio of two independent chi-square distributions [84]. Precisely, if $X \sim \phi'^2_n(\lambda)$, $Y \sim \phi'^2_m(\eta)$, then $Z = \frac{X/n}{Y/m}$ has a *doubly non-central* $\mathcal{F}$ *ratio distribution* of orders (n,m) and non-centrality parameters (λ,η),

$$Z \sim \mathcal{F}'_{n,m}(\lambda,\eta).$$

[2] The pdf of ζ can be obtained from that of t by guaranteeing that $\left|\mathrm{f}_{\mathcal{T}'_L(\sqrt{\lambda})}(t)dt\right| = |\mathrm{f}_\zeta(z)dz|$.

In the present case, $n = 1$, $m = L$, $\lambda = c_1'^2/\sigma_\xi^2 = \|\boldsymbol{c}\|^2/\sigma_\xi^2$ and $\eta = 0$, hence:

$$\Psi \sim \mathcal{F}'_{1,L}(c_1'^2/\sigma_\xi^2, 0).$$

Since:

$$\zeta^2 = 1 - \frac{1}{1 + \frac{1}{L}\Psi},$$

the pdf ζ^2 can be derived from that of Ψ by a change of variables from Ψ to ζ^2, and assumes the following form:

$$\mathrm{f}_{\zeta^2}(x; L, \lambda) = \frac{e^{-\frac{\lambda}{2}}(1-x)^{\frac{L}{2}-1}}{\sqrt{x}B\left(\frac{1}{2}, \frac{L}{2}\right)} {}_1F_1\left(\frac{L+1}{2}; \frac{1}{2}; \frac{\lambda}{2}x\right) \mathbb{1}_{[0,1]}(x),$$

where $B(\cdot,\cdot)$ is the Beta function and ${}_1F_1(\cdot;\cdot;\cdot)$ is the Kummer confluent hypergeometric function [84].

CHAPTER 4

Some results on MISO Time Reversal

In this chapter, we investigate Multiple-Input Single-Output (MISO) UWB communications over a multipath channel. After introducing a tractable discrete-time multipath MISO channel model accounting for correlation between multiple channels, analytical results for single-user transmissions are drawn in the following two cases: no precoding at transmitter with All-Rake receiver vs. time reversal (TR) at transmitter with 1-Rake receiver. Channel state information at both transmitter (CSIT) and receiver (CSIR) is assumed. Optimality of TR conditioned on 1-Rake receiver structure is proved. Robustness brought by TR combined with 1-Rake with respect to lack of correlation between channels vs. no precoding is shown. Simulations with realistic reference channels show that combining time reversal with multiple transmit antennas amplifies the performance improvement due to each technique when used alone.

4.1 Introduction

Time reversal is a signaling scheme that focuses a signal in both time and space. When transmitted signals have impulsive nature, as in Impulse Radio Ultra Wide Band (IR-UWB), time reversal further increases impulsiveness thanks to the many resolvable paths of UWB channels, even in the single-antenna setting [78,85]. When multiple antennas at transmitter side are used, time reversal is a beamformer focusing the wideband signal at receiver. Early experimental results for MISO and MIMO time reversal are collected, respectively, in [67] and in [68–70]. MISO beamforming with N_t antennas in flat-fading channels provides SNR gain of N_t [79]. On the other hand, in Single-Input Single-Output (SISO) multipath channels, a prefilter, *i.e.*, a filter at transmitter side, matched to the realization of the channel impulse response (CIR), which corresponds to time reversal, can focus in time the signal received: a transceiver using time reversal and 1-Rake achieves same performance in terms of SNR as that obtained without time reversal and with All-Rake [76]. This is useful when transmitter cost and complexity are not crucial as opposed to receiver simplicity. However, time reversal can also provide SNR gain when more complex receiver is used, that is when more than one finger is employed in Rake.

In this chapter we aim at providing theoretical results on MISO time reversal and extending known analyses proposed for SISO time reversal [74, 76, 78, 86] to MISO time reversal; in particular, we study the SNR gain due to time reversal beamforming and investigate the impact of the presence of more fingers in the Rake receiver when multiple antennas at transmitter are used.

The chapter is organized as follows. In Section 4.2 it is described the continuous-time system with realistic channel model and it is proposed a simplified tractable discrete-time channel model. In Section 4.3, analytical and simulation results are derived. Finally, conclusions and future work are drawn in Section 4.4.

4.2 System model

The basic equation of a binary PAM-UWB signal is the following:

$$s(t) = \sum_{m \in \mathbb{Z}} \mathsf{b}[m] p(t - mT_s - c_m T_c) \tag{4.1}$$

where $s(t)$ is the signal sent by the transmitter, c_m is the pseudorandom user-specific time hopping code, $\mathsf{b}[m] \in \{\pm \mathsf{A}\}$, where $\mathsf{A} > 0$, is the m-th symbol sent at time interval T_s, and T_c is called chip period [87]. We have used here Pulse Amplitude Modulation (PAM) but also Pulse Position Modulation [87] can be adopted. When no time reversal is used, $p(t)$ is a transmitted pulse of very short duration denoted in this chapter by $\psi(t)$, typically a Scholtz pulse or a bandlimited pulse. Conversely, when time reversal is used, $p(t)$ is the output of a linear time-invariant filter called *time reversal precoder* or *prefilter* with input waveform $\psi(t)$.

Single-user MISO communication with N_t antennas at transmitter side is considered. We focus on the reception of one symbol only, where each symbol is transmitted by amplitude modulating $p(t)$. Under this assumption, ISI is negligible since symbol period T_s is typically longer than channel delay spread and time-hopping code can ensure a time guard.

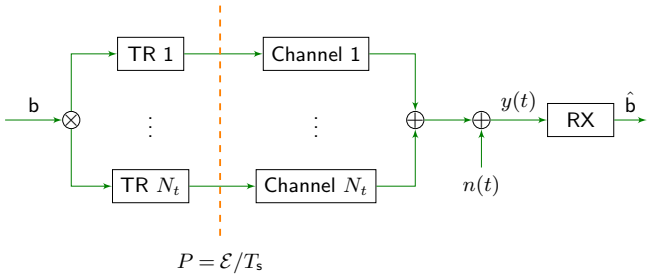


Figure 4.1: System model: MISO TR. N_t antennas are fed by N_t time reversal prefilters, each of which is matched to the corresponding multipath channel. A correlation receiver is adopted in order to estimate the transmitted symbol. Partial TR and Rake can be used in order to reduce complexity of the transceiver.

Received signal $y(t)$ can be written as follows, where we omitted the time-hopping code because of the one-symbol assumption:

$$y(t) = \mathsf{b} \int_{\mathbb{R}} \boldsymbol{h}^{\mathsf{T}}(t-\tau)\, \boldsymbol{p}(\tau) d\tau + n(t), \tag{4.2}$$

where $\boldsymbol{h}(t)$ and $\boldsymbol{p}(t)$ are the $N_t \times 1$ channel impulse response (CIR) and transmitted signal real-valued vectors, $n(t)$ is an Additive White Gaussian Noise (AWGN) with power spectral density σ_N^2, and b is the transmitted symbol. Note that all antennas transmit the same symbol during a symbol period. The i-th component of $\boldsymbol{h}(t)$ and $\boldsymbol{p}(t)$ is denoted by $h_i(t)$ and $p_i(t)$. Power constraint is intended as follows:

$$\frac{1}{T_{\mathsf{s}}} \int_{\mathbb{R}} \mathsf{A}^2 \|\boldsymbol{p}(t)\|^2 dt = \frac{1}{T_{\mathsf{s}}} \sum_{i=1}^{N_t} \int_{\mathbb{R}} \mathsf{A}^2 p_i^2(t) dt \leq P,$$

where T_{s} is the symbol period and P is the power spent by the N_t antennas, and, therefore, by the transmitter.

Given the large bandwidth of UWB systems, the propagation channel is resolved into multiple paths. In general, knowledge of CIR realizations at transmitter side can be exploited for prefilter design; MISO TR design adopts a bank of N_t prefilters, where each antenna element is fed by the output of a prefilter modulated by b and matched to the CIR between that antenna element and the receiver, hence:

$$p_n(t) = \int_{\mathbb{R}} \psi(\tau) h_n(\tau - t) d\tau = (\psi * h_n^{\mathsf{r}})(t), \tag{4.3}$$

where $h_n^{\mathsf{r}}(t) = h_n(-t)$; denoting by $\mathcal{E}$ the maximum energy spent by the transmitter during a symbol period, power constraint reads as:

$$\mathsf{A}^2 \sum_{n=1}^{N_t} \int_{\mathbb{R}} \left| \psi * h_n^{\mathsf{r}}(t) \right|^2 dt \leq P T_{\mathsf{s}} \triangleq \mathcal{E}.$$

Since the multipath channel $h_i(t)$ can be written as:

$$h_i(t) = \sum_{k \geq 0} a_k^i \delta(t - \tau_k^i),$$

partial time reversal prefilters can be also designed, where just a subset of paths is considered [74].

Finally, receiver is assumed to be a matched filter that projects $y(t)$ onto the following template:

$$\int_{\mathbb{R}} \boldsymbol{h}^{\mathsf{T}}(t-\tau)\, \boldsymbol{p}(\tau) d\tau \equiv \sum_{k \geq 0} \tilde{a}_k \psi(t - t_k).$$

where $\{\tilde{a}_k, t_k \colon k \geq 0\}$ represents the set of amplitudes and delays of fingers in the Rake. However, we can also consider a partial Rake receiver where only a subset of $\{\tilde{a}_k, t_k \colon k \geq 0\}$ is known and used: when the strongest paths only are considered,

the receiver is known as *Selective-Rake*, whereas when all paths are considered, it is known as *All-Rake*.

4.2.1 Channel Model: Continuous-Time

We consider for simulation results the channel proposed in [88] that combines the standard SISO channel IEEE 802.15.3a [53] with the Kronecker channel correlation model for MIMO. We briefly summarize below this model in the special case of MISO and uniform linear array.

Channel impulse response between the i-th transmit antenna and the receiver is as follows:

$$h_i(t) = \sum_{\ell,k\geq 0} a^i_{k,\ell}\delta(t - T_\ell - \tau_{k,\ell} - \Delta^i)\delta(\phi^{\mathrm{t}} - \phi^{\mathrm{t}}_{k,\ell})\delta(\phi^{\mathrm{r}} - \phi^{\mathrm{r}}_{k,\ell}),$$

where $a^i_{k,\ell}$ is the signed amplitude of ray k in cluster ℓ, T_ℓ is cluster ℓ delay, $\tau_{k,\ell}$ is the delay of ray k in cluster ℓ with respect to T_ℓ, $\Delta^i = (i-1)d/c$ is the incremental delay due to antenna i position where d is the distance between two antennas and c is the speed of light, and $\phi^{\mathrm{t}}_{k,\ell}$ and $\phi^{\mathrm{r}}_{k,\ell}$ are the angle-of-departure and angle-of-arrival of ray k in cluster ℓ, respectively. Amplitude of ray k in cluster ℓ for CIR i has the form: $a^i_{k,\ell} = z_{k,\ell}\xi_\ell\beta^i_{k,\ell}$, where $z_{k,\ell}$ is the equiprobable ± 1 amplitude sign due to random reflections, and ξ_ℓ and $\beta^i_{k,\ell}$ are lognormal distributed cluster and ray fading, respectively. Amplitudes and delays statistics are in accordance with the standard IEEE 802.15.3a SISO channel. Multiple transmit antennas correlation is taken into account by assigning a correlation structure to $\boldsymbol{a}^{\mathsf{T}}_{k,\ell} = (a^1_{k,\ell}, \ldots, a^{N_t}_{k,\ell})$, that is by requiring that $\mathbb{E}[\boldsymbol{a}_{k,\ell}\boldsymbol{a}^{\mathsf{T}}_{k,\ell}] = \boldsymbol{\Sigma}^{\boldsymbol{a}}_{k,\ell}$. Since only $\beta^i_{k,\ell} = 10^{(\mu_{k,\ell}+\nu^i_{k,\ell})/20}$ in $a^i_{k,\ell}$ depends on the antenna, we just need to assign a correlation structure to $(\beta^i_{k,\ell})_{i=1}^{N_t}$, and, in turn, to $\boldsymbol{\nu}_{k,\ell} = (\nu^1_{k,\ell}, \ldots, \nu^{N_t}_{k,\ell})^{\mathsf{T}}$. Since $\nu^i_{k,\ell}$ is Gaussian, then $\boldsymbol{\nu}_{k,\ell}$ is a zero mean multivariate normal distribution with covariance matrix $\boldsymbol{\Sigma}^{\boldsymbol{\nu}}_{k,\ell}$ that can be derived once fixed $\boldsymbol{\Sigma}^{\boldsymbol{a}}_{k,\ell}$. For further details refer to [88].

4.2.2 Simplified Discrete-Time Channel Model

A more tractable channel model for analytical derivations is described below under the assumption of unit-energy, ideal baseband W-bandlimited pulse $\psi(t)$ with band $[-\mathsf{W}/2, \mathsf{W}/2]$.

SISO model

Single-Input Single-Output frequency-selective continuous-time AWGN channels can be described by the relation:

$$y(t) = \sum_{n=0}^{N-1} h(t) * \mathsf{b}[n]\psi(t - nT_{\mathsf{s}}) + n(t),$$

where $\{\mathsf{b}[n]\colon 0 \le n \le N-1\}$ is the set of transmitted symbols, $\psi(t)$ is the bandlimited waveform, T_{s} is the symbol period, $h(t)$ is the channel impulse response, and $n(t)$ is a realization of the continuous-time AWGN process with power spectral density σ_N^2.

In impulse-radio ultra-wideband communications, adjacent symbols are separated in time by an interval much longer than $1/\mathsf{W}$: in fact, $1/\mathsf{W}$ is shorter than the chip period and each symbol period counts usually tens of chips. Therefore, called $M = T_{\mathsf{s}}\mathsf{W}$, previous relation can be written as follows:

$$y(t) = \sum_{m=0}^{NM-1} h(t) * \tilde{\mathsf{b}}[m]\psi(t - m/\mathsf{W}) + n(t),$$

where $\tilde{\mathsf{b}}[nM] = \mathsf{b}[n]$ for $0 \le n \le N-1$, and $\tilde{\mathsf{b}}[i] = 0$ when i is not a multiple of M. Since one symbol only is considered, the channel model is as follows:

$$y(t) = \sum_{m=0}^{M-1} h(t) * \tilde{\mathsf{b}}[m]\psi(t - m/\mathsf{W}) + n(t).$$

At receiver, projecting this signal onto $\{\psi(t - m/\mathsf{W})\colon m \in \mathbb{Z}\}$ yields the following discrete-time channel:

$$y_k = \sum_{m\in\mathbb{Z}} h_{k-m}x_m + n_k,$$

where $x_m := \tilde{\mathsf{b}}[m]$, $\{n_k\colon k \in \mathbb{Z}\}$ is a set of i.i.d. Gaussian random variables with variance σ_N^2, and

$$h_\ell := \frac{1}{\sqrt{\mathsf{W}}} \int_{\mathbb{R}} h(t)\psi(t - \ell/\mathsf{W})\,dt.$$

Typically, most of the energy is included in the first paths, say $0 \le \ell \le L$. Without loss of generality, $L \ll M$ since M can be made as large as needed. Since $x_m = 0$ for $m < 0$ and $m > M-1$, then

$$y_k \approx \sum_{m=0}^{M-1} h_{k-m}x_m + n_k,$$

and this relation can be written in vector form as follows:

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n},$$

where $\boldsymbol{H}$ is a $(M+L)\times M$ Toeplitz matrix with first column equal to $(h_0, h_1, \ldots, h_L, 0, \ldots, 0)^{\mathsf{T}}$, $\boldsymbol{x} = (x_i)_{i=0}^{M-1}$ and $\boldsymbol{n} = (n_i)_{i=0}^{M+L}$. One-shot communication implies that $x_m = \mathsf{b}[0]\delta_{m,0}$, therefore $\boldsymbol{x} = \mathsf{b}[0]\boldsymbol{e}_1$, where $\boldsymbol{e}_i$ is the $M \times 1$ vector with all zero entries but i-th element equal to 1. Note that $\mathsf{b}[0]$ is equal to b of eq. (4.2), therefore hereinafter we will denote it with b. Power constraint is expressed as $\mathbb{E}\big[\,\|\boldsymbol{x}\|^2\,\big] = \mathsf{A}^2 \le \mathcal{E}$.

Note that we can consider $\boldsymbol{y}$ as the concatenation of two vectors, $\overline{\boldsymbol{y}}$ and $\underline{\boldsymbol{y}}$, such that $\boldsymbol{y} = [\overline{\boldsymbol{y}}^{\mathsf{T}}\underline{\boldsymbol{y}}^{\mathsf{T}}]^{\mathsf{T}}$, where $\overline{\boldsymbol{y}}$ is $M \times 1$ and $\underline{\boldsymbol{y}}$ is $L \times 1$. Similarly, $\boldsymbol{H} = [\overline{\boldsymbol{H}}^{\mathsf{T}}\underline{\boldsymbol{H}}^{\mathsf{T}}]^{\mathsf{T}}$ and

$\boldsymbol{n} = [\overline{\boldsymbol{n}}^{\mathsf{T}} \underline{\boldsymbol{n}}^{\mathsf{T}}]^{\mathsf{T}}$, where $\overline{\boldsymbol{H}}$ is $M \times M$ and $\underline{\boldsymbol{H}}$ is $L \times M$. Since $M \gg L$, we can neglect $\underline{\boldsymbol{y}}$. Prefiltering can be accounted by a $M \times M$ matrix $\boldsymbol{P}$ as follows: $\overline{\boldsymbol{y}} = \overline{\boldsymbol{H}}\boldsymbol{P}\boldsymbol{x} + \overline{\boldsymbol{n}}$. With prefiltering, transmitted vector is $\boldsymbol{P}\boldsymbol{x}$, therefore power constraint is $\mathbb{E}\big[\,\|\boldsymbol{P}\boldsymbol{x}\|^2\,\big] = \mathsf{A}^2\|\boldsymbol{p}\|^2 \leq \mathcal{E}$, being $\boldsymbol{p}$ the first column of $\boldsymbol{P}$. Since the prefilter is a linear time-invariant system, $\boldsymbol{P}$ is Toeplitz and $\boldsymbol{p}$ is the vector containing projections of the prefilter impulse response $p(t)$ onto $\{\psi(t - m/\mathsf{W})\colon 0 \leq m \leq M-1\}$.

In order to provide a tractable model for the MISO channel, we propose first a simplified channel model for the SISO channel, which helps in illustrating the behavior of the system with and without TR. For SISO NLOS channels, it seems reasonable to consider $\{h_\ell\colon 0 \leq \ell \leq L\}$ independent, zero-mean Gaussian random variables with variance profile $\{v_\ell\colon 0 \leq \ell \leq L\}$, *i.e.*, $\mathrm{Var}[\,h_\ell\,] = v_\ell$. Therefore, the simplest model we propose is: $h_\ell = \sqrt{v_\ell} h_w(\ell)$, where $\{h_w(\ell)\colon 0 \leq \ell \leq L\}$ are i.i.d. unit variance Gaussian random variables. Furthermore, when the channel has also a LOS component, we propose a generalization that reads as follows:

$$h_\ell = \sqrt{v_\ell}\left\{\sqrt{\frac{\kappa}{1+\kappa\rho}}\delta_{\ell,0} + \sqrt{\frac{1}{1+\kappa\rho}}h_w(\ell)\right\},$$

where κ, similarly to the "Ricean factor" in flat-fading channels, accounts for the fraction of energy in the LOS component with respect to the NLOS component *in the first path*, and ρ accounts for the fraction of energy of the first path with respect to the sum of the energy of all paths, namely $\rho = v_0/\mathsf{G}$, where $\mathsf{G} = \sum_{\ell \geq 0} v_\ell$ is the total channel gain.

MISO model

The discrete-time model is as follows:

$$y(k) = \sum_{m=0}^{M-1} \boldsymbol{h}^{\mathsf{T}}(k-m)\boldsymbol{x}(m) + n(k),$$

where $\boldsymbol{h}(m)$ is a $N_t \times 1$ vector where the i-th element is the channel impulse response between the i-th transmit antenna and the receiver projected onto $\psi(t - m/\mathsf{W})/\sqrt{\mathsf{W}}$, and $\boldsymbol{x}(m) = \mathsf{b}\mathbf{1}\delta_{m,0}$, where $\mathbf{1}$ is a $N_t \times 1$ vector of '1'. Stacking $\{y(k)\colon 0 \leq k \leq M-1\}$ in a vector, $\overline{\boldsymbol{y}}$, the previous relation can be written as:

$$\overline{\boldsymbol{y}} = \overline{\boldsymbol{H}}\boldsymbol{x} + \overline{\boldsymbol{n}}, \quad \text{with} \quad \boldsymbol{x} = \boldsymbol{e}_1 \otimes \mathbf{1}\mathsf{b}, \tag{4.4}$$

where $\boldsymbol{x}$ is a $MN_t \times 1$ vector of symbols and $\overline{\boldsymbol{H}}$ is a $M \times MN_t$ block-Toeplitz channel matrix structured as follows:

$$\overline{\boldsymbol{H}} = \sum_{m=0}^{M-1} \boldsymbol{J}_m \otimes \boldsymbol{h}^{\mathsf{T}}(m) = \sum_{\ell=0}^{L} \boldsymbol{J}_\ell \otimes \boldsymbol{h}^{\mathsf{T}}(\ell), \tag{4.5}$$

where $\boldsymbol{J}_m$ is a $M \times M$ matrix with ones on the m-th subdiagonal. Power constraint is: $\mathbb{E}\big[\,\|\boldsymbol{x}\|^2\,\big] = N_t\mathsf{A}^2 \leq \mathcal{E}$. Prefiltering can be accounted by a $MN_t \times MN_t$ block-

Toeplitz matrix $\boldsymbol{P}$ structured as follows:

$$\boldsymbol{P} = \sum_{m=0}^{M-1} \boldsymbol{J}_m \otimes \boldsymbol{P}(m). \tag{4.6}$$

Thus, the channel assumes the following form: $\overline{\boldsymbol{y}} = \overline{\boldsymbol{H}}\boldsymbol{P}\boldsymbol{x} + \overline{\boldsymbol{n}}$; from eq. (4.4), we have:

$$\overline{\boldsymbol{y}} = \overline{\boldsymbol{H}}\boldsymbol{p}\mathsf{b} + \overline{\boldsymbol{n}}. \tag{4.7}$$

where:

$$\boldsymbol{p} := \boldsymbol{P}(\boldsymbol{e}_1 \otimes \mathbf{1}) = \sum_{m=0}^{M-1} \boldsymbol{J}_m \boldsymbol{e}_1 \otimes \boldsymbol{P}(m)\mathbf{1} \equiv \sum_{m=1}^{M} \boldsymbol{e}_m \otimes \boldsymbol{p}(m),$$

having denoted $\boldsymbol{p}(m+1) := \boldsymbol{P}(m)\mathbf{1}$. In general, $\boldsymbol{P}(m)$ jointly precodes symbols for all antennas; in MISO TR prefilters are decoupled, that is, each antenna is fed by the output of a prefilter knowing the channel impulse response between that antenna and the receiver only. This structure implies that $\boldsymbol{P}(m)$ is diagonal, therefore also uniquely defined by a vector that we call $\boldsymbol{p}(m+1)$, that is consistent with previous notation:

$$\boldsymbol{P}(m) = \mathrm{diag}(\boldsymbol{p}(m+1)).$$

Power constraint reads:

$$\mathbb{E}\left[\, \|\boldsymbol{P}\boldsymbol{x}\|^2 \,\right] = \mathsf{A}^2 \|\boldsymbol{p}\|^2 \leq \mathcal{E}. \tag{4.8}$$

In order to derive analytical results, we propose the following tractable simplified channel model for MISO channel, that similarly to SISO channels is defined as:

$$\boldsymbol{h}^{\mathsf{T}}(\ell) = \sqrt{v_\ell} \left\{ \sqrt{\frac{\kappa}{1+\kappa\rho}} \mathbf{h}_0^{\mathsf{T}} \delta_{\ell,0} + \sqrt{\frac{1}{1+\kappa\rho}} \boldsymbol{h}_w^{\mathsf{T}}(\ell) \boldsymbol{\Theta}_t^{1/2} \right\},$$

where matrix $\boldsymbol{\Theta}_t$ is introduced as in the Kronecker model in order to account correlation between channels of different antennas, $\mathbf{h}_0$ is a deterministic component that account for the line-of-sight path, and $\boldsymbol{h}_w(\ell) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$. We want that $\mathbb{E}\left[\, \|\boldsymbol{h}(\ell)\|^2 \,\right] = N_t v_\ell$, therefore we set $\mathrm{Tr}\,\boldsymbol{\Theta}_t = N_t$ and $\|\mathbf{h}_0\|^2 = N_t$. In this way, correlation matrix of $\boldsymbol{h}_w(\ell)$ is as follows:

$$\mathbb{E}\Big[\, \boldsymbol{h}(\ell)\boldsymbol{h}^{\mathsf{T}}(\ell) \,\Big] = v_\ell \cdot \left\{ \frac{\kappa}{1+\kappa\rho} \mathbf{h}_0 \mathbf{h}_0^{\mathsf{T}} \delta_{\ell,0} + \frac{1}{1+\kappa\rho} \boldsymbol{\Theta}_t \right\}. \tag{4.9}$$

A generalization to MIMO would be straightforward, taking into account also correlations at receiver antennas with a matrix $\boldsymbol{\Theta}_r$. However, since in this chapter we investigate MISO channels, we denote hereinafter $\boldsymbol{\Theta}_t$ by $\boldsymbol{\Theta}$, being the subscript redundant.

4.2.3 Performance measure

The key performance measure that we are interested in is SNR. The problem is to estimate b from eq. (4.7) through a linear receiver, that is a $M \times 1$ vector $\boldsymbol{w}$ onto which $\overline{\boldsymbol{y}}$ is projected. We want to compute the SNR of the decision variable $z = \boldsymbol{w}^{\mathsf{T}}\overline{\boldsymbol{y}}$, that is:

$$z = \boldsymbol{w}^{\mathsf{T}}\overline{\boldsymbol{H}}\boldsymbol{p}\mathsf{b} + \boldsymbol{w}^{\mathsf{T}}\overline{\boldsymbol{n}} =: \zeta + \nu,$$

where $\zeta := \boldsymbol{w}^{\mathsf{T}}\overline{\boldsymbol{H}}\boldsymbol{p}\mathsf{b}$ and $\nu := \boldsymbol{w}^{\mathsf{T}}\overline{\boldsymbol{n}}$. Therefore, SNR of z is as follows:

$$\mathsf{SNR} := \frac{\mathbb{E}\left[\,|\zeta|^2\,\right]}{\mathbb{E}\left[\,|\nu|^2\,\right]} = \frac{\mathsf{A}^2}{\sigma_N^2} \cdot \frac{(\boldsymbol{w}^{\mathsf{T}}\overline{\boldsymbol{H}}\boldsymbol{p})^2}{\|\boldsymbol{w}\|^2},$$

where expectations are taken conditioned on the channel realization. This quantity is, therefore, still a random variable, since the channel is a random process. We are primary interested in its mean value, $\mathbb{E}[\,\mathsf{SNR}\,]$.

4.3 Results

4.3.1 Analytical Derivations

We derive expressions of SNR in terms of channel correlation, receiver number of fingers and transmitter number of antennas. In particular, we compute the two opposed cases of strongly correlated and independent channels, characterized by $\boldsymbol{\Theta} = \mathbf{1}\mathbf{1}^{\mathsf{T}}$ and $\boldsymbol{\Theta} = \boldsymbol{I}$, respectively.

All-Rake ($\boldsymbol{w} = \overline{\boldsymbol{H}}\boldsymbol{p}$), General Precoder

$$\begin{aligned}\mathsf{SNR} &= \frac{\mathsf{A}^2}{\sigma_N^2} \cdot \|\overline{\boldsymbol{H}}\boldsymbol{p}\|^2 \\ &= \frac{\mathsf{A}^2}{\sigma_N^2} \cdot \left\| \sum_{\ell=0}^{M-1} \sum_{m=1}^{M} \boldsymbol{e}_{m+\ell} \otimes \boldsymbol{h}^{\mathsf{T}}(\ell)\boldsymbol{p}(m) \right\|^2,\end{aligned}$$

intending that the sum is extended over $1 \leq m + \ell \leq M$.

All-Rake ($\boldsymbol{w} = \overline{\boldsymbol{H}}\boldsymbol{p}$), No Precoding

With no precoding, $\boldsymbol{P} \propto \boldsymbol{I}$, therefore $\boldsymbol{P}(m) \propto \boldsymbol{I}_{N_t}\delta_{m,0}$. According to the power constraint, it is sufficient to specialize the general relation for SNR to:

$$\boldsymbol{p}(m) = \mathbf{1}\sqrt{\frac{\mathcal{E}}{N_t\mathsf{A}^2}}\delta_{m,1},$$

that yields:

$$\mathsf{SNR} = \frac{\mathcal{E}}{N_t \sigma_N^2} \sum_{\ell=0}^{L} \left(\boldsymbol{h}^{\mathsf{T}}(\ell)\mathbf{1}\right)^2. \tag{4.10}$$

Note that, in the SISO case, this reduces to the well-known result $\mathsf{SNR} = (\mathcal{E}/\sigma_N^2)\sum_\ell h(\ell)^2$.

Using the tractable proposed channel model, the average SNR is:

$$\mathbb{E}[\mathsf{SNR}] = \frac{\mathcal{E}}{N_t \sigma_N^2} \sum_{\ell=0}^{L} \mathbf{1}^{\mathsf{T}} \mathbb{E}\left[\boldsymbol{h}(\ell)\boldsymbol{h}(\ell)^{\mathsf{T}}\right] \mathbf{1},$$

that, using eq. (4.9), is:

$$\begin{aligned}\mathbb{E}[\mathsf{SNR}] &= \frac{\mathcal{E}}{N_t \sigma_N^2} \sum_{\ell=0}^{L} v_\ell \left\{ \frac{\kappa}{1+\kappa\rho} (\mathbf{h}_0^{\mathsf{T}}\mathbf{1})^2 \delta_{\ell 0} + \frac{1}{1+\kappa\rho} \mathbf{1}^{\mathsf{T}} \boldsymbol{\Theta} \mathbf{1} \right\} \\ &= \frac{\mathcal{E}\,\mathsf{G}}{N_t \sigma_N^2} \left\{ \frac{\kappa\rho}{1+\kappa\rho} \cdot \mathsf{H} + \frac{1}{1+\kappa\rho} \cdot \theta \right\}. \end{aligned} \tag{4.11}$$

having set $\mathsf{H} = (\mathbf{h}_0^{\mathsf{T}}\mathbf{1})^2$, $\theta = \mathbf{1}^{\mathsf{T}}\boldsymbol{\Theta}\mathbf{1}$ and $\mathsf{G} = \sum_{\ell=0}^{L} v_\ell$.

Let analyze particular cases. Assume deterministic LOS component equal for all antennas, that is $\mathbf{h}_0 = \mathbf{1}$, then $\mathsf{H} = N_t^2$. For strongly correlated channels, namely for $\boldsymbol{\Theta} = \mathbf{1}\mathbf{1}^{\mathsf{T}}$, we have $\theta = N_t^2$. In this case:

$$\mathbb{E}[\mathsf{SNR}] = \frac{\mathcal{E}\mathsf{G}}{\sigma_N^2} N_t.$$

For independent channels, that is $\boldsymbol{\Theta} = \boldsymbol{I}$, we have $\theta = N_t$, and therefore:

$$\mathbb{E}[\mathsf{SNR}] = \frac{\mathcal{E}\,\mathsf{G}}{\sigma_N^2} \left\{ \frac{\kappa\rho}{1+\kappa\rho} \cdot N_t + \frac{1}{1+\kappa\rho} \right\}.$$

1-Rake, time reversal

A 1-Rake is $\boldsymbol{w} = \boldsymbol{e}_i$, where in the case of time reversal $i = L+1$. Time reversal is the use of the precoder:

$$\boldsymbol{p}^\star = \frac{\sqrt{\mathcal{E}}}{\mathsf{A}} \cdot \frac{1}{\|\boldsymbol{h}^{\mathsf{r}}\|} \boldsymbol{h}^{\mathsf{r}},$$

where $\boldsymbol{h}^{\mathsf{r}} = \boldsymbol{c}^{L+1}$ and $\boldsymbol{c}^{L+1}$ is the $(L+1)$-th row of $\overline{\boldsymbol{H}}$. We can prove that this precoder maximizes the SNR when 1-Rake is used.

In fact, consider 1-Rake $\boldsymbol{w} = \boldsymbol{e}_i$, $1 \le i \le M$. Since $\boldsymbol{e}_i^{\mathsf{T}} \overline{\boldsymbol{H}} \boldsymbol{p} = \boldsymbol{c}^{i\mathsf{T}} \boldsymbol{p}$, where $\boldsymbol{c}^i$ is the i-th row of $\overline{\boldsymbol{H}}$, the problem is that of finding the precoding vector:

$$\boldsymbol{p}^\star = \arg\max_{\boldsymbol{p}} \mathsf{SNR} = \arg\max_{\boldsymbol{p}} (\boldsymbol{c}^{i\mathsf{T}} \boldsymbol{p})^2,$$

subject to power constraint $\mathsf{A}^2 \|\boldsymbol{p}\|^2 \le \mathcal{E}$. This problem is solved, using for instance the Cauchy-Schwarz inequality, by $\boldsymbol{p}_i^\star \propto \boldsymbol{c}^i$, that is time reversal. Each $i \ge L+1$

Table 4.1: Parameters used in simulations.

Channel Model	IEEE 802.15.3a-CH1
Pulse	Scholtz, $\tau = 200$ ps
Antenna distance (d)	0.1 m
Angle spread (Ω)	38°

is sufficient for taking into account the whole delay spread of channels, therefore exploiting the whole gain offered. For the sake of simplicity, we choose $i = L + 1$.

With the time reversal precoder, achieved SNR is:

$$\mathsf{SNR} = \frac{\mathcal{E}}{\sigma_N^2} \|\boldsymbol{h}^{\mathrm{r}}\|^2 = \frac{\mathcal{E}}{\sigma_N^2} \sum_{\ell=0}^{L} \|\boldsymbol{h}(\ell)\|^2. \tag{4.12}$$

A well-known result in SISO channels is that this SNR is equal to that obtained with All-Rake receiver and no precoding. This is no longer true with MISO, since eq. (4.12) is not equivalent to eq. (4.10).

Using the tractable proposed channel model, the average value of (4.12) is:

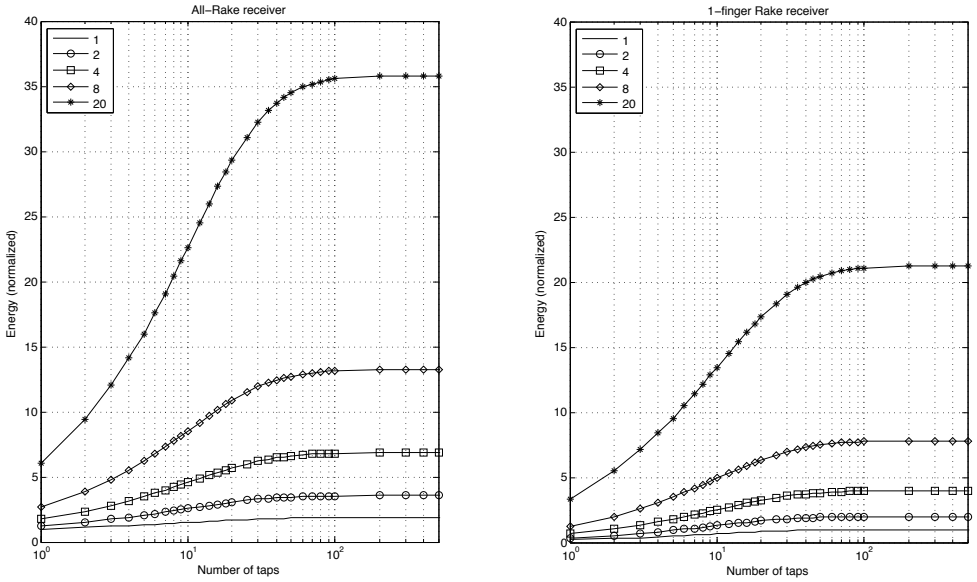
$$\mathbb{E}[\mathsf{SNR}] = \frac{\mathcal{E}\mathsf{G}}{\sigma_N^2} N_t. \tag{4.13}$$

The meaning of this relation is that time reversal is insensitive, on average and in terms of SNR, to many parameters characterizing the channel as in (4.11). In particular, time reversal is robust to the lack of correlation as opposed to All-Rake without precoding: in fact, 1-Rake, time reversal systems average SNR is larger than that achieved by systems with All-Rake receiver and without precoding, beiung equal only when channels between different transmit antennas and receiver are identical.

4.3.2 Simulation Results

In addition to analytical derivations and comparison of performance (in terms of SNR) of systems using time reversal with 1-Rake vs. All-Rake without prefiltering, we investigate through simulations performance dependence on number of antennas, type of receiver (1-Rake vs. All-Rake) and number of taps in the time reversal prefilter. In fact, transmitter may also select a subset of channel paths to form TR: when one tap only is employed, we have no prefiltering; when the whole channel impulse response is taken into account, we have full TR; between this two extrema, when the number of taps in the precoder is limited, we have in general partial TR, and paths are selected with decreasing amplitude. Simulations are performed with the realistic channel described in Subsection 4.2.1. Parameters used in simulations are reported in Table 4.1.

In Figs. 4.2a and 4.2b it is shown the energy collected by an All-Rake and by a 1-Rake, respectively, normalized to the average energy collected without time reversal



(a) Energy collected by an All-Rake

(b) Energy collected by a 1-Rake.

Figure 4.2: Energy collected with two di erent receivers normalized to the average energy collected without time reversal, with single antenna, as a function of the number of taps of precoder.

with single antenna, as function of the number of paths considered in time reversal. In both cases, time reversal can be used for reducing the number of antennas while maintaining fixed the energy collected by the receiver. For instance, in Fig. 4.2a, a system with 20 antennas without time reversal collects the same energy of a system with 4 antennas employing time reversal with 25 taps. Note that, moreover, energy gain due to time reversal remains roughly constant as function of the number of antennas, therefore the total gain due to multiple antennas and time reversal is roughly the product of each gain. A relative increase of focusing, although of small entity, is observed on Figs. 4.3 showing the percentage of collected energy as function of the number of fingers in the Rake receiver with time reversal. We confirmed by simulations the better performance of time reversal with 1-Rake with respect to no time reversal with All-Rake, as expected from the analyical derivation.

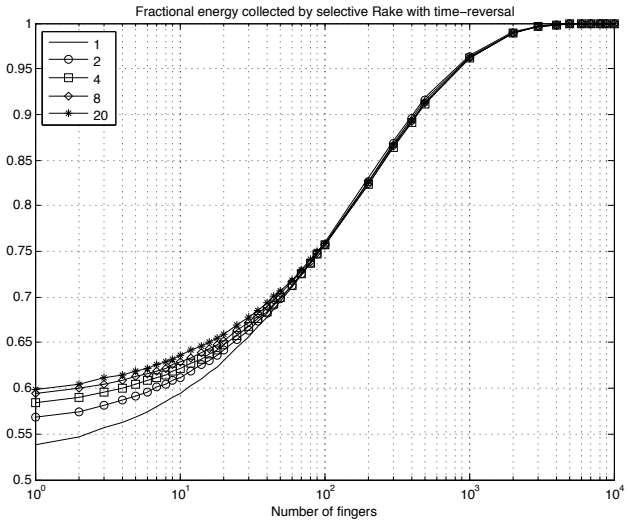


Figure 4.3: Fractional energy: time reversal

4.4 Conclusion and Future work

In this chapter we studied combination of time reversal and multiple transmit antennas providing theoretical arguments for robustness of communications using TR and simulation results emphasizing focusing properties of TR and possible trade-offs in system design. In general, thanks to its focusing properties, time reversal allows to use a 1-Rake receiver in place of receivers requiring multiple fingers in systems without prefiltering. We proved that MISO TR is optimum when a 1-Rake receiver is used and that systems with no time reversal and All-Rake receiver are less robust to channels correlation with respect to systems with time reversal and 1-Rake receiver, which performance in terms of average SNR is better. Simulation results with realistic channel model and analytical derivations with a more tractable channel model we proposed were presented. We observed that combination of MISO and TR amplify performance of both techniques used independently as well as increased focusing of MISO TR with respect to SISO TR; we might expect higher gains with a channel model considering a more realistic propagation model.

A future generalization may consider multiuser beamforming, in particular relaxing the assumption on the bank of separate time reversal prefilters, one per antenna, and by studying a joint TR design. Another development may be towards understanding the broadcast time reversal channel: in its simplest form, two users have to be served simultaneously by a transmitter with multiple antennas using a time reversal prefilter. It may be studied how to optimally combine the two channel impulse responses in the joint time reversal prefilter, and compare these results against different prefiltering schemes, such as RZF, by taking into account the correlation between the two channels. Finally, we suggest to study how the antenna

configuration, *i.e.*, geometry, impacts the performance of TR-based transmitters, which may be relevant in distributed antenna setups.

CHAPTER 5

Some results on SISO Time Reversal

In this chapter, a comparison between two optimum prefilters for frequency-selective SISO channels under different assumptions on the receiver is drawn. The first design foresees a jointly optimized transceiver (optimum prefilter and all-Rake receiver) while the second constrains the receiver to be minimum complex (1-Rake) and consequently uses the optimized prefilter, that is, time reversal. Assuming an 802.15.3a channel model, analytical results based on the description of the channel as a doubly stochastic point process are proposed for the latter case. Simulations provide performance comparison with the former design showing the maximum prefiltering gain gap and the minimum Rake complexity required by the optimum prefilter for a performance improvement.

Notation

$\lvert x\rvert$	absolute value (element-wise if $\boldsymbol{x}$ is a vector)
$\lVert\boldsymbol{x}\rVert_p$	p-norm where $\boldsymbol{x}$ is a vector (p=2 if not specified)
$\mathbb{N}$	set of natural numbers (1, 2, ...)
$\overline{\mathbb{N}}$	set of non-negative integers ($\overline{\mathbb{N}} = \mathbb{N} \cup \{0\}$)
$\vartheta(x)$	Heaviside theta (or unit-step) function
$X \overset{\mathcal{D}}{\sim} p_X(x)$	X is distributed according to the probability density function $p_X(x)$
$\mathbb{E}_X\, f(X)$	expectation of $f(X)$ where $f(X)$ is a function of random variable (r.v.) X

For notation simplicity, the expectation takes no precedence over any other operator. For example, the following holds:

$$\mathbb{E}_X\, f(X) + g(X) = \mathbb{E}_X\{f(X) + g(X)\}$$
$$\mathbb{E}_Y\, \mathbb{E}_{X|Y}\, f(X)g(Y) = \mathbb{E}_Y\{\mathbb{E}_{X|Y}\{f(X)g(Y)\}\}$$

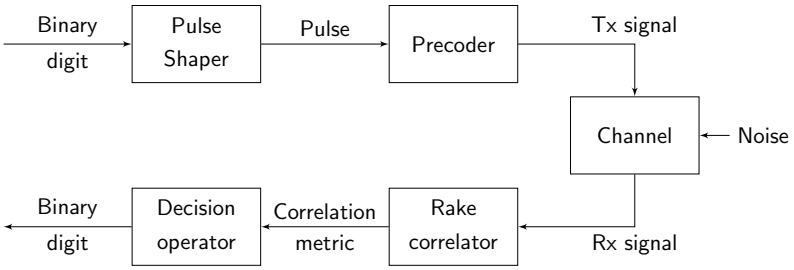


Figure 5.1: System model: block diagram.

5.1 Introduction

Communications over frequency-selective channels benefit from pre-equalization of transmitted waveforms. The optimum receiver in absence of interference is the Rake receiver, but its complexity grows with the required number of fingers – that is ideally equal to the number of paths– and the performance gain provided by adding a new finger is always below a prediction based on a linear improvement. Furthermore, prefiltering optimality is conditioned on the structure of the receiver. Therefore, much work [66] has focused on finding sub-optimal receiver and transceiver joint optimum design.

When the receiver is constrained to one finger only, time reversal is the optimum prefilter (see Appendix 5.A for a simple proof). We compare the design composed of time reversal and 1-Rake with the joint optimum transceiver, assuming an 802.15.3a channel model. The channel is described as a doubly stochastic one-dimensional point process, that is a simplification of previous works on the topic [89–91]. We find the average prefiltering gain brought by time reversal, and compare it with the maximum gain reached with the optimum prefilter. Driven by the underlying motivation of simplifying the receiver, we compute the minimum number of fingers required by the Rake receiver to guarantee an average performance improvement with respect to time reversal.

The chapter is organized as follows: Section 5.2 is devoted to the description of the system model; Sections 5.3 and 5.4 present time reversal vs. Optimum Prefilter theoretical performance; Section 5.5 introduces the channel model and analyzes the performance of time reversal for this specific case. Finally, simulation results and conclusions are presented in Section 5.6.

For clarity of presentation, complex analytical derivations and descriptions are included in Appendices. In particular, Appendix 5.A provides a proof of the optimality of time reversal when the receiver is constrained to one finger only. Appendix 5.B contains the derivation of an expression for the time reversal prefiltering gain. In Appendix 5.C and 5.D we describe the point process model of the channel and in Appendix 5.E we specialize the time reversal prefiltering gain to this channel family.

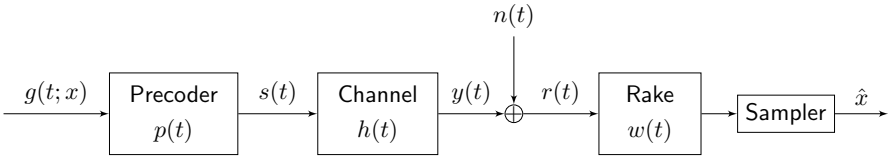


Figure 5.2: System model (continuous-time): $y(t)$ is the useful signal and $r(t) = y(t) + n(t)$ the received signal corrupted by the AWGN $n(t)$ with variance σ_N^2. The receiver is implemented by a filter with impulse response $w(t) := y(-t)$ followed by a sampler at $t = 0$.

5.2 System model

We consider a point-to-point communication through a frequency-selective channel. The system is represented in block diagram form in Fig. 5.1. We model each block as a linear time-invariant system with impulse response shown in Fig. 5.2. We assume:

(a1) a signaling scheme based on the two normalized waveforms $\{g(t;x)\}$, being x the information-bearing symbol;

(a2) a power constraint at the transmitter, that is, independent of the prefilter, the power of the transmitted signal is constant and equal to P_0 (w.l.o.g. we can set $P_0 = 1$);

(a3) independent inter-arrival times between paths in the multipath channel.

For the sake of simplicity we consider PAM as the modulation technique, that is $g(t;a) = ag(t)$, where $g(t)$ is the unit-energy pulse and $a = 2b-1$ is the information-bearing symbol, being b the bit of information.

The optimum demodulator is the filter matched to the whole useful received signal $y(t)$. In the case of a multipath channel, it is known as Rake correlator [92,93]. This is depicted in Fig. 5.2 that shows the continuous-time model of the system. We can derive an equivalent discrete-time model (see Fig. 5.3): the optimal decoder $\boldsymbol{w}$ is the one that maximizes, having fixed $\boldsymbol{p}$, the SNR of the correlation metric, and it is the matched filter to the effective channel $\boldsymbol{Hp}$:

$$\mathsf{SNR} = \frac{\mathbb{E}|\boldsymbol{w}^{\mathsf{H}}\boldsymbol{H}\boldsymbol{p}x|^2}{\mathbb{E}(\boldsymbol{w}^{\mathsf{H}}\boldsymbol{n}\boldsymbol{n}^{\mathsf{H}}\boldsymbol{w})} = \frac{|\boldsymbol{w}^{\mathsf{H}}\boldsymbol{H}\boldsymbol{p}|^2}{\sigma_N^2\|\boldsymbol{w}\|^2} \implies \boldsymbol{w} \propto \boldsymbol{H}\boldsymbol{p}\ .$$

The resulting $\mathsf{SNR}^\star = \boldsymbol{p}^{\mathsf{H}}\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H}\boldsymbol{p}/\sigma_N^2 = \mathbb{E}(\|\boldsymbol{y}\|^2)/\sigma_N^2$ depends only on the energy of the useful received signal.

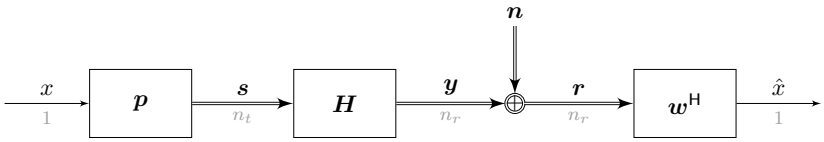


Figure 5.3: System model (discrete-time): x represents the symbol to be transmitted, $\boldsymbol{p} \in \mathbb{R}^{n_t}$ the prefilter sampled impulse response, $\boldsymbol{s}$ the transmitted signal, $\boldsymbol{H} \in \mathbb{R}^{n_t \times n_r}$ the linear convolution channel matrix, $\boldsymbol{r} = \boldsymbol{y} + \boldsymbol{n}$ the received signal with useful component $\boldsymbol{y}$ and AWGN $\boldsymbol{n} \in \mathbb{R}^{n_r}$, and $\boldsymbol{w} \in \mathbb{R}^{n_r}$ the correlator.

5.3 Time Reversal SNR in SISO Frequency-Selective Channels

We describe the SISO frequency-selective channel as

$$h(t) = \sum_{k \geq 0} h_k \delta(t - \tau_k),$$

where h_k and τ_k are amplitude and delay of the k^{th} path, respectively.

Note 1. Delays are taken in ascending order, $\tau_k < \tau_{k+1}$.

Note 2. In practice, we would take only a finite number of paths, say $L+1$.

Given a channel realization, we can design the prefilter according to different criteria. One criterion is the maximization of the SNR at the 1-Rake receiver. Under this assumption, the optimum filter is time reversal, represented by a special choice of $p(t)$, that is

$$p(t) = \frac{\sqrt{P_0}}{||\boldsymbol{h}||} \sum_{k \geq 0} h_k \delta(t + \tau_k).$$

Proof. See Appendix 5.A.

Note 3. We do not address the issue of the causality of this filter; however, it is evident that in order to model a real system we need to introduce a delay of at least τ_L.

Remark 2. Normalization is required by the power constraint assumption (**a2**).

Definition 3. We call *prefiltering gain* the quantity $\rho := \mathcal{E}_y / \mathcal{E}_h$.

Remark 3. The prefiltering gain is absent without prefiltering, that is $\rho = 1$ when $p(t) = \delta(t)$.

Proposition 2 (**Prefiltering gain of time reversal**). *Assume that* (**a3**) *holds. Then, having fixed the continuous-time channel impulse response $h(t)$, it results almost surely (that is, with probability 1) that*

$$\rho = 2 - \frac{||\boldsymbol{h}||_4^4}{||\boldsymbol{h}||_2^4}. \tag{5.1}$$

where $\boldsymbol{h} := (h_k)_{k \geq 0}$.

Proof. See Appendix 5.B. □

Remark 4. This result is no longer true in the discrete-time model. However, it is an excellent approximation if the discrete-time channel impulse response is represented by a sparse vector. In the case of multipath channel, the higher the bandwidth of the receiver, the better is the approximation.

Remark 5. This result is a property of the continuous-time channel impulse response (CIR) and it remains valid if and only if the temporal duration of the pulse is lower than the minimum distance between two paths, that is, $\operatorname{supp} g(t) \leq \min\{|\tau_i - \tau_j| : i \neq j\}$. Conversely, due to the presence of inter-pulse interference (IPI), the wider the pulse, the worse is the approximation.

Now, the aim is to bound ρ. Since $||\boldsymbol{h}||_4/||\boldsymbol{h}||_2 > 0$, a trivial upper bound of ρ is given by 2. A lower bound follows from the Cauchy-Schwarz inequality once we accept a rather technical assumption verified numerically by simulations (see Fig. 5.3 for more details):

Lemma 2. *Assume that*

$$\mathbb{E}\frac{||\boldsymbol{h}||_4^4}{||\boldsymbol{h}||_2^4} \leq \frac{\mathbb{E}||\boldsymbol{h}||_4^4}{\mathbb{E}||\boldsymbol{h}||_2^4},$$

then the following inequality holds:

$$\mathbb{E}\rho = 2 - \mathbb{E}\frac{||\boldsymbol{h}||_4^4}{||\boldsymbol{h}||_2^4} \geq 2 - \frac{\mathbb{E}||\boldsymbol{h}||_4^4}{\mathbb{E}||\boldsymbol{h}||_2^4} \geq 2 - \frac{\mathbb{E}||\boldsymbol{h}||_4^4}{\left(\mathbb{E}||\boldsymbol{h}||_2^2\right)^2}.$$

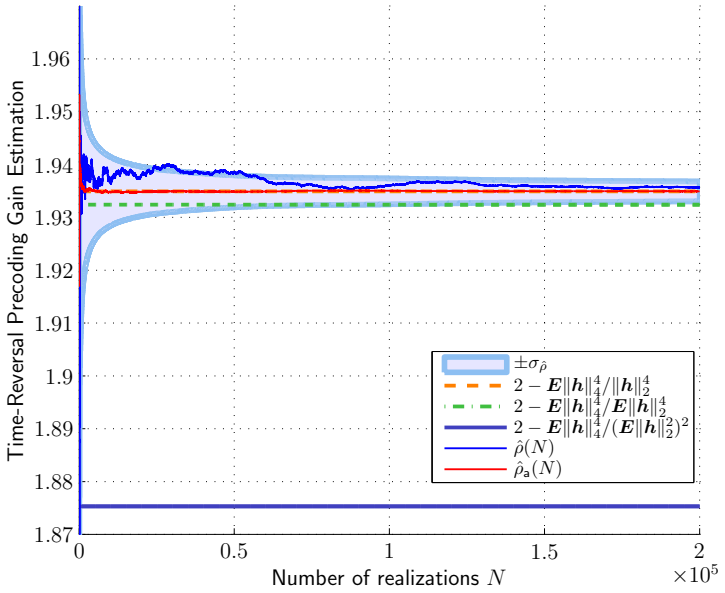


Figure 5.3: This plot suggests that assumption of *Lemma* 2 is verified with IEEE 802.15.3a (CM1). Here we denote with $\boldsymbol{h}$ a discrete-time CIR sampled at 10 GHz. Furthermore, $\hat{\rho}(N)$ is the ensemble average of ρ (see Def. 3) over a set of N channel realizations, $\hat{\rho}(N) := \sum_i \rho_i / N$, and the accuracy of the mean estimator within $\pm\sigma_{\hat{\rho}}$, that is one standard deviation, is represented by the filled region. As $N \to \infty$, $\hat{\rho}(N)$ tends to the expectation of ρ (computed over discrete-time channels). Since Prop. 2 is valid only in continuous-time, we show a di erent curve for the average of the RHS of eq. (5.1), that is, $\hat{\rho}_{\mathsf{a}}(N)$. The dashed curve shows the asymptotic value of $\hat{\rho}_{\mathsf{a}}(N)$. This value is below the asymptotic value of $\hat{\rho}(N)$, that means that, in discrete-time, the estimation of $\hat{\rho}$ with $\hat{\rho}_{\mathsf{a}}$ is conservative, at least asymptotically. The dot-dashed curve shows that it is conservative to replace the expectation of the ratio with the ratio of expectations, verifying the assumption of *Lemma* 2. Finally, the thick solid curve show the lower bound that derives from the Cauchy-Schwarz inequality.

5.4 Optimum SNR in SISO Frequency-Selective Channels

The optimum prefilter is designed to maximize SNR under the power constraint:

$$\begin{cases} \underset{\boldsymbol{p} \in \mathbb{R}^{n_t}}{\text{maximize}} & \boldsymbol{p}^{\mathsf{H}} \boldsymbol{H}^{\mathsf{H}} \boldsymbol{H} \boldsymbol{p} \\ \text{subject to} & \mathsf{Tr}(\boldsymbol{p}\boldsymbol{p}^{\mathsf{H}}) = 1 \end{cases}$$

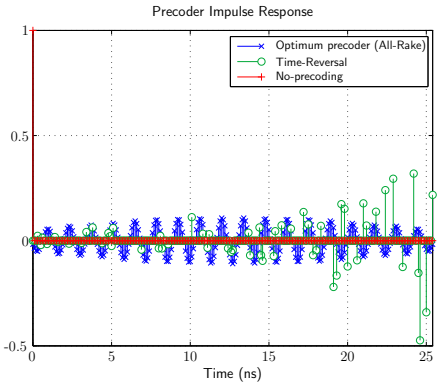
having set $P_0 = 1$. This is also optimum from an information theoretic point of view as it maximizes the capacity of the link.

The solution of this optimization problem is given by the eigenvector of matrix $\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H}$ corresponding to the maximum eigenvalue $\lambda_{\max}$, or equivalently by the right singular vector $\boldsymbol{u}_{\max}$ of matrix $\boldsymbol{H}$ corresponding to the maximum singular value $\sigma_{\max}$. In turn, it results that $\lambda_{\max} = \sigma_{\max}^2$.

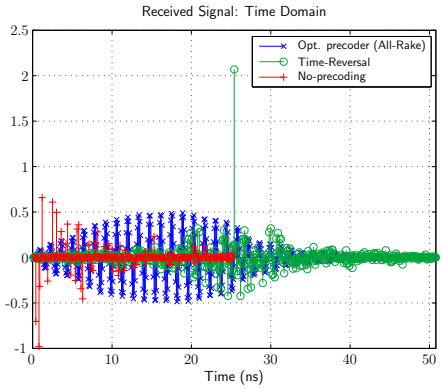
Remark 6. The optimum prefilter tries to focus the power at the neighbourhood of the channel spectrum peak frequency. We can justify informally this behavior noting that $\boldsymbol{H}$ is a convolutional matrix, thus it is at least asymptotically diagonalized by a Fourier matrix $[\boldsymbol{F}]_{mn} := e^{j2\pi mn/n_t}$ (like a Toeplitz matrix):

$$\boldsymbol{H} \doteq \boldsymbol{F}^{\mathsf{H}} D_{\mathbf{H}} \boldsymbol{F} \implies \boldsymbol{H}^{\mathsf{H}}\boldsymbol{H} \doteq \boldsymbol{F}^{\mathsf{H}} D_{\mathbf{H}}^{\mathsf{H}} D_{\mathbf{H}} \boldsymbol{F} = \boldsymbol{F}^{\mathsf{H}} D_{|\mathbf{H}|^2} \boldsymbol{F}$$

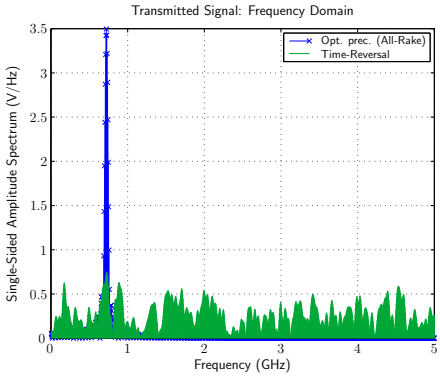
where $D_{\mathbf{H}} = \operatorname{diag}(H[0], \ldots, H[n_t - 1])$, being $H[k]$ the k^{th} sample of the n_t-points DFT of $\boldsymbol{h}$, say $\mathbf{H} = \boldsymbol{F}^{\mathsf{H}}\boldsymbol{h}$, and $\doteq$ emphasizing the asymptotic nature of the equality. The last diagonal matrix is intended to be generated from $|\mathbf{H}|^2 = \operatorname{diag}(|H[0]|^2, \ldots, |H[n_t - 1]|^2)$ with the absolute value on the matrix acting element-wise. With this diagonalization, we can simply read the eigenvalue spectrum of $\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H}$ as the power spectrum of $\boldsymbol{h}$, $\lambda_k = |H[k]|^2$. Therefore, the maximum eigenvalue occurs at the discrete-frequency $\bar{k}$ that maximizes the discrete power spectrum of the channel. Conversely, the optimum prefilter tries to focus the energy where the channel attenuation is the least, that is in the neighbourhood of $\bar{k}$. This behavior is shown in Fig. 5.4*(d)*.



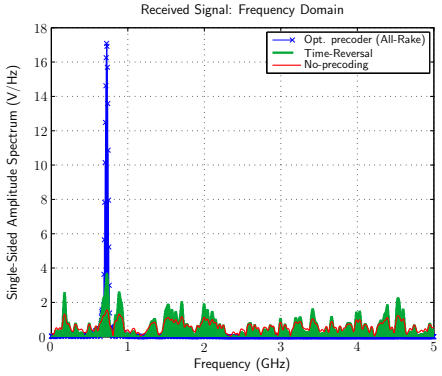
(a) Prefilter impulse response (sampling time: $\Delta t = 0.1$ [ns]).



(b) Received signal in time domain (sampling time: $\Delta t = 0.1$ [ns]).



(c) Transmitted signal: amplitude spectrum.



(d) Received signal: amplitude spectrum.

Figure 5.4: Example of (a) prefilter impulse response, and transmit and received signals *(b)–(d)* with a fixed channel realization drawn from the IEEE 802.15.3a standard.

5.5 The Case of IR-UWB Channels

In line with previous works on the topic [89–91], we use point process theory to model the channel impulse response. Hence, the channel is described as a doubly stochastic point process, that is formally derived from a Cox process (doubly stochastic Poisson point process).

The starting point is the channel model proposed by Saleh and Valenzuela in [94] and further adapted to UWB by Foerster et al. in [53]. The expression for the channel impulse response is

$$h(t) = \sum_{\ell \geq 0} \sum_{k \geq 0} h_{k\ell} \delta(t - T_\ell - \tau_{k\ell})$$

where T_ℓ is the ℓ^{th} cluster start time and $\tau_{k\ell}$ is the delay of the k^{th} path in the ℓ^{th} cluster relative to T_ℓ. According to the model in [53]

- $\mathcal{C} := (T_\ell)_{\ell \in \mathbb{N}}$ is a Poisson process of intensity Λ, and
- $(\tau_{k\ell})_{k,\ell \in \mathbb{N}}$ is another independent Poisson process of intensity λ.

Therefore, defining $\mathcal{R}_\ell := (T_\ell + \tau_{k\ell})_{k \in \mathbb{N}}$, we have $\forall \ell$ that $\mathcal{R}_\ell | T_\ell$ is a Poisson process of intensity $\varrho_{\mathcal{R}_\ell | T_\ell}(t) = \lambda \vartheta(t - T_\ell)$. The idea is now to model the (deterministic) presence of the first path of each cluster at the cluster start time using generalized functions in the intensity function. Thus, in order to model that $T_0 = 0$ (in LOS environment) and $\tau_{0\ell} = 0$ almost surely, we define:

- $\overline{\mathcal{C}} := (T_\ell)_{\ell \in \overline{\mathbb{N}}}$ that is a point process of intensity $\varrho_{\overline{\mathcal{C}}}(t) = \Lambda \vartheta(t) + \delta(t)$, and
- $\overline{\mathcal{R}}_\ell | T_\ell$ that is a point process of intensity $\varrho_{\overline{\mathcal{R}}_\ell | T_\ell}(t) = \lambda \vartheta(t - T_\ell) + \delta(t - T_\ell)$.

By the independence of the latter processes, also their superimposition

$$\overline{\mathcal{R}} | \overline{\mathcal{C}} := \bigcup_{\ell=0}^{\infty} \overline{\mathcal{R}}_\ell | T_\ell$$

is a point process with intensity function given by

$$\varrho_{\overline{\mathcal{R}} | \overline{\mathcal{C}}}(t) = \sum_{\ell=0}^{\infty} \varrho_{\overline{\mathcal{R}}_\ell | T_\ell}(t) = \sum_{\ell=0}^{\infty} \lambda \vartheta(t - T_\ell) + \delta(t - T_\ell)$$

Remark 7. In NLOS environments it results that $\varrho_{\overline{\mathcal{C}}}(t) = \Lambda \vartheta(t)$.

Thus, regarding $(T_\ell)_{\ell \in \mathbb{N}}$ as random, this intensity function becomes a random field

$$Z(t) := \sum_{\ell=0}^{\infty} \lambda \vartheta(t - T_\ell) + \delta(t - T_\ell) = \sum_{\bar{c} \in \overline{\mathcal{C}}} \lambda \vartheta(t - \bar{c}) + \delta(t - \bar{c}).$$

This means that $\overline{\mathcal{R}}$ conditioned on Z is a point process with intensity function Z, therefore it is a doubly stochastic process with random intensity measure Z and

intensity function:

$$\varrho_{\overline{\mathcal{R}}}(t) = \mathbb{E}Z(t).$$

Lemma 3. *The intensity function of the point process representing the channel is*

$$\varrho_{\overline{\mathcal{R}}}(t) = \lambda(1+\Lambda t) + \Lambda + \delta(t).$$

Proof. See Appendix 5.C. □

The statistical properties of the channel amplitudes $(h_{k\ell})_{k,\ell\in\bar{\mathbb{N}}}$ are described in [53]. We recall them in Appendix 5.D and use them in Appendix 5.E in order to find the following result:

Proposition 3 (**Bounds to prefiltering gain of TR**). *Assume that the channel is described as previously, then*

$$0 \leq 2 - \mathbb{E}\rho \leq \frac{r_4(1+\lambda\gamma/2)(1+\Lambda\Gamma/2)}{(1+\lambda\gamma)^2(1+\Lambda\Gamma)^2} \tag{5.2}$$

Proof. See Appendix 5.E. □

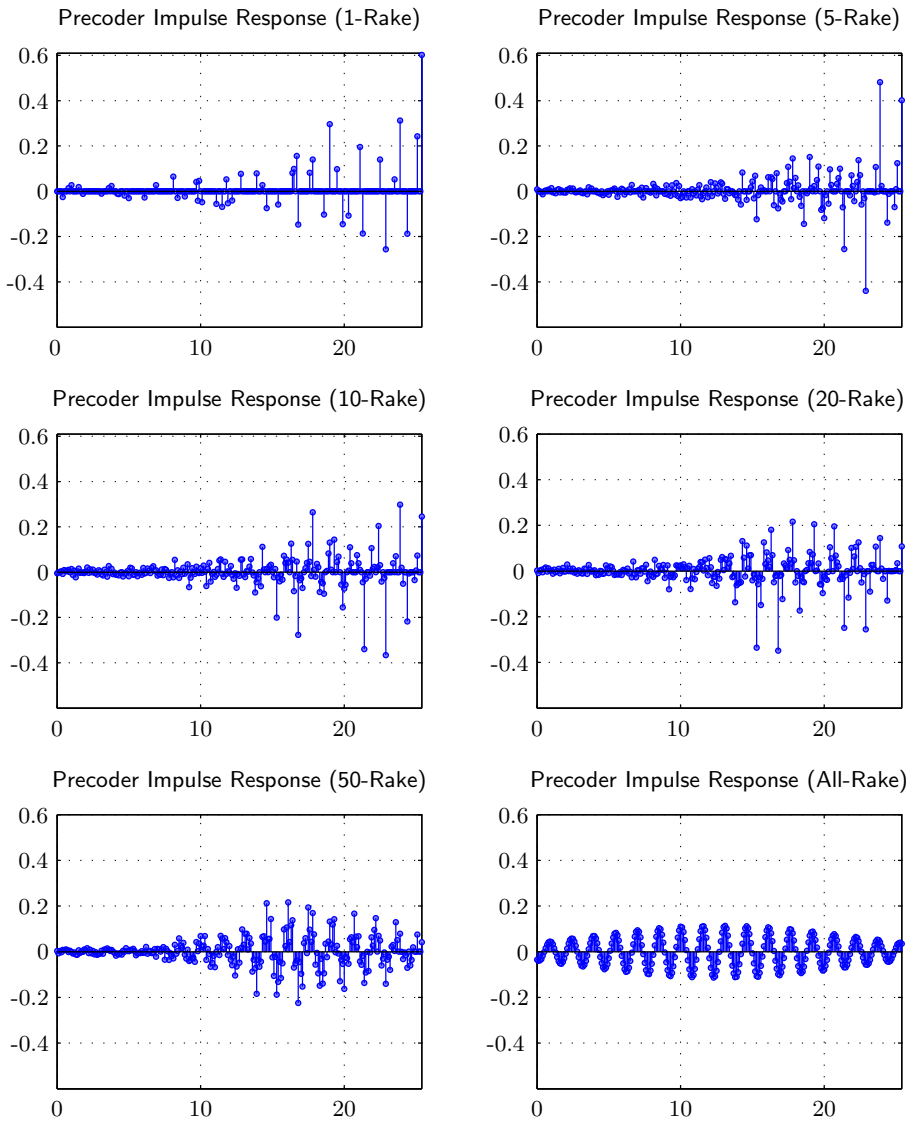


Figure 5.5: Prefilter impulse responses optimized with respect to the number of fingers of Rake receivers. The impulse responses are obtained by maximizing numerically the SNR at the receiver. This is an empirical study, since there is no proof that each of these impulse responses actually maximize the SNR, but the prefilter optimized for the All-Rake receiver matches the theory.

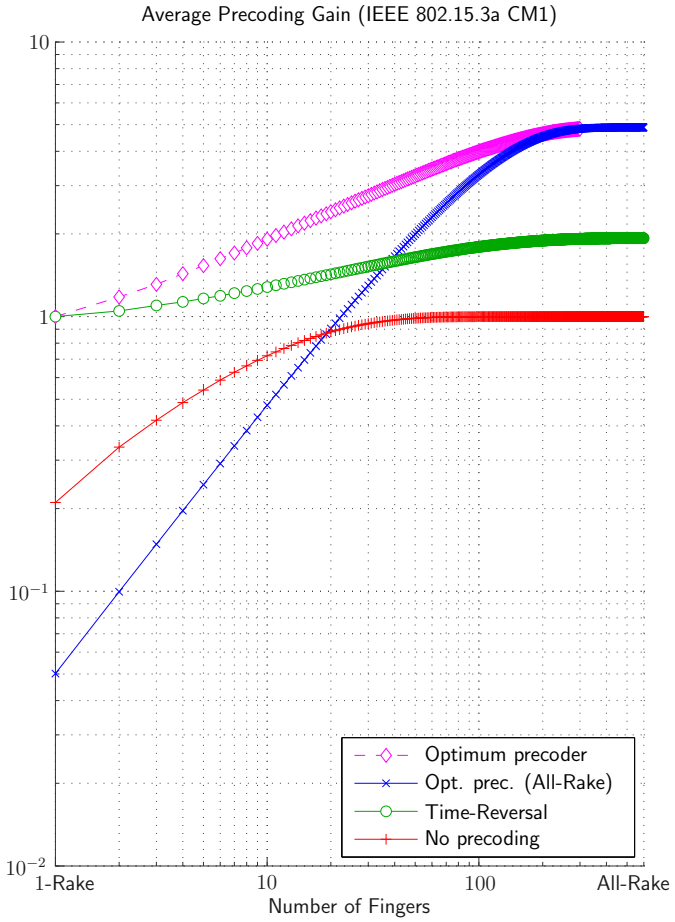


Figure 5.6: Prefiltering gain as a function of the number of fingers of the Rake receiver. The asymptotically optimum precoder (crosses) is optimum under the condition that the receiver is an All-Rake. The optimum precoder (diamonds) is obtained empirically by using prefilters as shown in Fig. 5.5. Performance of systems with no precoding (pluses) and time reversal (circles) at the transmitter side is shown for comparison. The receiver bandwidth is 5 GHz; however, the qualitative behavior of the curves is loosely dependent from the bandwidth.

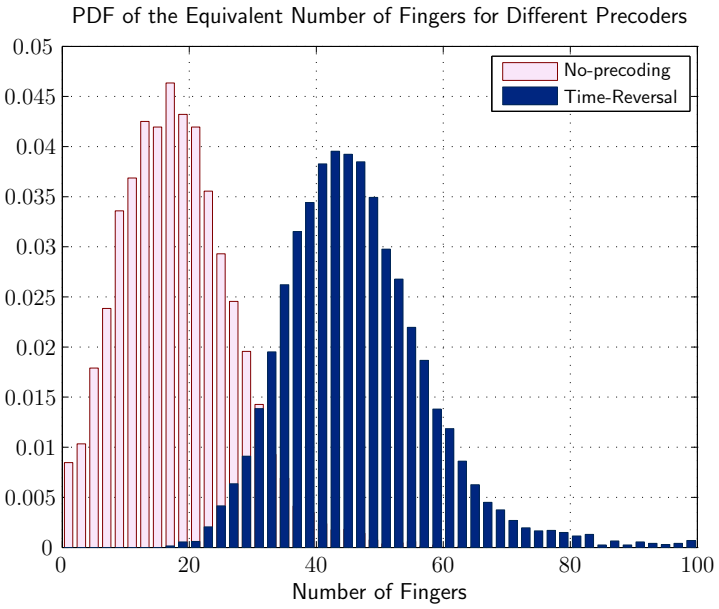


Figure 5.7: Estimation of the PDF of that number of fingers such that di erent prefilters have same performance (see also Fig. 5.6). The histogram on the left (with the lighter background) refers to Optimum Prefilter (All-Rake) and No-prefiltering; the histogram on the right (with the darker background) refers to Optimum Prefilter (All-Rake) and time reversal.

5.6 Conclusion

In this chapter we compared optimum prefilters operating in SISO frequency-selective channels. The SISO channels are characterized by a complexity feature that we define in association with the number of fingers of the Rake receiver. The analysis is carried out in the absence of MUI. Figures 5.4*(a)–(d)* showed the behavior in time and frequency of different prefilters. It is shown that the optimum prefilter focuses the energy in a neighbourhood of the frequency corresponding to the maximum channel gain (see Fig. 5.4*(d)*). A drawback is the spreading in time of both the impulse response of this filter (see Fig. 5.4(a)) and the received signal (see Fig. 5.4(b)), implying that a larger number of fingers is required in order to collect the same energy that would be collected with other prefilters. Eventually, further analysis is required in order to verify compliance with UWB masks.

A comparison of the gain of different prefiltering schemes is shown on Fig. 5.6. The prefiltering gain is the value that the curves tend to asymptotically, as the number of fingers grows to infinity. Other points on the curves show the energy normalized to the channel gain $\mathcal{E}_h$ as a function of the number of fingers of the Rake receiver. For the simplest Rake receiver (one finger only), time reversal is optimum: it allows to collect the maximum energy with a 1-Rake, that is equal to $\mathcal{E}_h$. At the other extreme, the optimum prefilter, that is joint-optimized with the optimum receiver (All-Rake), allows to collect the maximum energy. Figure 5.6 further shows that the asymptotically optimum prefilter, *i.e.*, optimum only when the receiver is an All-Rake, outperforms time reversal only if the receiver has a number of fingers larger than a threshold, that is a random variable whose pdf is estimated in Fig. 5.7. Conversely, this analysis suggests that it is mandatory to require receiver complexity if a prefilter that is optimum when the receiver is an All-Rake is employed.

Future work may investigate analytically optimum prefilters as a function of the number of fingers of the Rake receiver. In general, a deeper investigation of prefiltering techniques with multipath channels is of interest. Moreover, extension to multiuser communications is possible and may reveal peculiar properties of prefilters operating in time over multipath channels.

Appendix

5.A Time Reversal is Optimum if the Receiver is 1-Rake

This appendix aims at providing a quick proof of the optimality of time reversal when the receiver is constrained to one finger only. The problem can be stated in continuous-time as follows: Given the channel impulse response (CIR) $h(t)$, find the function $p(t)$ solution of the problem:

$$\max_{p} \max_{t} |(h * p)(t)|^2.$$

We can constrain w.l.o.g. the solution p to provide the maximum at $t = 0$; the problem becomes

$$\max_{p} \left| \int_{\mathbb{R}} h(\tau) p(-\tau) d\tau \right|^2.$$

Defining $f(t) := p(-t)$, the last integral is the inner product $\langle h, f \rangle$ and the solution derives from the Cauchy-Schwarz inequality: $f \propto h$, that implies $p(t) \propto h(-t)$, that is, time reversal.

5.B Proof of Time Reversal Gain.

The received useful signal $y(t)$ is as follows:

$$\begin{aligned} y(t) &= \frac{1}{\sqrt{\mathcal{E}_h}} \sum_{k \geq 0} \sum_{i \geq 0} h_k h_i \delta(t + \tau_k - \tau_i) \\ &= \frac{1}{\sqrt{\mathcal{E}_h}} \left\{ \left[\sum_{k \geq 0} h_k^2 \right] \delta(t) + \sum_{k \geq 0} \sum_{\substack{i \geq 0 \\ i \neq k}} h_k h_i \delta(t + \tau_k - \tau_i) \right\} \\ &= \sqrt{\mathcal{E}_h} \delta(t) + \frac{1}{\sqrt{\mathcal{E}_h}} \sum_{k \geq 0} \sum_{\substack{i \geq 0 \\ i \neq k}} h_k h_i \delta(t + \tau_k - \tau_i) \end{aligned}$$

The assumption $(\mathbf{a3})$ implies that $\forall k, k', i, i'$ with $k \neq k'$ and $i \neq i'$, one has $P(\tau_k - \tau_{k'} = \tau_i - \tau_{i'}) = 0$; therefore, almost surely, all previous $\delta(\cdot)$ are located at different times. The energy $\mathcal{E}_{y_N}$ of $y(t)$ considering only the firsts N paths tends

asymptotically (as $N \to \infty$) to:

$$\begin{aligned}\mathcal{E}_{y_N} &\xrightarrow{\text{a.s.}} \mathcal{E}_h + \frac{1}{\mathcal{E}_h}\sum_{k=0}^{\infty}\sum_{\substack{i=0\\ i\neq k}}^{\infty} h_k^2 h_i^2 \\ &= \mathcal{E}_h + \frac{1}{\mathcal{E}_h}\left[\sum_{k=0}^{\infty}\sum_{i=0}^{\infty} h_k^2 h_i^2 - \sum_{k=0}^{\infty} h_k^4\right] \\ &= \mathcal{E}_h + \frac{1}{\mathcal{E}_h}\left(\mathcal{E}_h^2 - ||\boldsymbol{h}||_4^4\right) \\ &= \mathcal{E}_h + \mathcal{E}_h - \frac{||\boldsymbol{h}||_4^4}{\mathcal{E}_h} \\ &= \mathcal{E}_h\left[2 - \frac{||\boldsymbol{h}||_4^4}{||\boldsymbol{h}||_2^4}\right]\end{aligned}$$

where $\boldsymbol{h} := (h_k)_{k\geq 0}$. It follows that

$$\rho_N := \frac{\mathcal{E}_{y_N}}{\mathcal{E}_h} \xrightarrow{\text{a.s.}} \rho := 2 - \frac{||\boldsymbol{h}||_4^4}{||\boldsymbol{h}||_2^4}.$$

5.C Proof of Lemma 3: Channel Process Intensity Function.

The aim is to compute $\mathbb{E}Z(t)$. This can be accomplished by applying the Campbell's formula that allows to determine expectations of functions evaluated on the point process. Precisely, given a point process $\mathcal{X} \subseteq \mathbb{R}^m$ with measure N, mean measure $\mu := \mathbb{E}N$ and intensity $\varrho_{\mathcal{X}}(x)$, and a function $f\colon \mathcal{X} \to \mathbb{R}$, it holds that:

$$\begin{aligned}\mathbb{E}\sum_{x\in\mathcal{X}} f(x) &= \mathbb{E}\int f(x)N(dx) \\ &= \int f(x)\mu(dx) = \int f(x)\varrho_{\mathcal{X}}(x)dx,\end{aligned}$$

where the integral is extended over the domain of $\mathcal{X}$.

We recall that

$$Z(t) = \sum_{\overline{c}\in\overline{\mathbb{C}}} \delta(t-\overline{c}) + \lambda\vartheta(t-\overline{c}) = \sum_{\ell\geq 0} \delta(t-T_\ell) + \lambda\vartheta(t-T_\ell)$$

and that $\varrho_{\overline{\mathbb{C}}}(t) = \Lambda + \delta(t)$. Therefore

$$\begin{aligned}&\mathbb{E}\sum_{\ell\geq 0}\delta(t-T_\ell) + \lambda\,\mathbb{E}\sum_{\ell\geq 0}\vartheta(t-T_\ell) \\ &= \int_0^\infty \delta(t-x)[\Lambda + \delta(x)]dx + \lambda\int_0^\infty \vartheta(t-x)[\Lambda+\delta(x)]dx \\ &= \Lambda + \delta(t) + \lambda(\Lambda t + 1).\end{aligned}$$

5.D Statistical Description of IEEE 802.15.3a Channel Path Amplitudes

Let us summarize the statistical properties of the channel amplitudes proposed in [53] as the following three constraints:

(c1) *exponential decay*: $\mathbb{E}(h_{k\ell}^2) = \Omega_0 e^{-T_\ell/\Gamma} e^{-\tau_{k\ell}/\gamma}$

being γ and Γ the intra-cluster and cluster decay factors, respectively;

(c2) *independent cluster fading*: each cluster is affected by a Log-Normal fading (cluster fading) independently from other clusters, that is, the paths within the ℓ^{th} cluster, say $(h_{k\ell})_{k\in\overline{\mathbb{N}}}$, are scaled by (the same) log-normal r.v. ξ_ℓ statistically independent from $\{\xi_{\ell'}\}_{\ell'\neq\ell}$;

(c3) *independent ray fading*: each path is affected by a Log-Normal fading (ray fading) independently from other paths, that is, the generic path $h_{k\ell}$ is scaled by a r.v. $\beta_{k\ell}$ statistically independent from $\{\beta_{k'\ell'}\}_{\substack{k'\neq k\\ \ell'\neq\ell}}$.

Remark 8. The generic $h_{k\ell} =: h(T_\ell, T_\ell + \tau_{k\ell}) = h(c, r)$, $c \in \overline{\mathcal{C}}$, $r \in \overline{\mathcal{R}}_\ell | T_\ell$ can be viewed as the mark of the point located in r.

In line with [53], the generic channel amplitude can be decomposed as follows:

$$h(c, r) = p(c, r)\xi(c)\beta(c, r)$$

where $p(c, r) \overset{\mathcal{D}}{\sim} \mathcal{U}\{-1, +1\}$ takes into account the path sign due to reflections and $\xi(\cdot)$ and $\beta(\cdot, \cdot)$ follows from the natural extension of Remark 8 to ξ_ℓ and $\beta_{k\ell}$, respectively. We further decompose the term $\beta(c, r)$ in a deterministic part, say $\eta(c, r)$, and a random part, say $\zeta(r)$:

$$\beta(c, r) = \eta(c, r)\zeta(r)$$

and call $\alpha(c, r) := \xi(c)\beta(c, r)$.

Remark 9. In this way $\xi(c)$ and $\zeta(r)$ are i.i.d. while dealing with $\beta(c, r)$ leads just to independence. Furthermore, if we were dealing with $\alpha(c, r)$, we would only have conditional independence of $\alpha(c, \cdot)$ with respect to c, that is, the amplitudes are indepedent only once conditioned to the cluster.

Constraints (**c1**)–(**c3**) uniquely define the channel process. Nevertheless, it remains a freedom in the formal description and we use it to slightly simplify the analytical formulation found in [53]. We recap the latter in (A) and propose the former in (B). Hereinafter, we use the following notation:

- n_i stands for a Normal process such that $\forall t\ n_i(t) \overset{\mathcal{D}}{\sim} \mathcal{N}(0, 1)$,
- $\sigma_1 > 0$ is the standard variation of the Normal r.v. associated with the cluster fading,

$\sigma_2 > 0$ is the standard variation of the Normal r.v. associated with the ray fading,

$\kappa = (\ln 10)/20$ is the conversion factor from dB to Np, thus 1 [dB] $= \kappa$ [Np].

(A) In [53] it is assumed that $\forall c > 0, r > 0$

$$\xi(c) := \mathrm{e}^{\kappa\sigma_1 n_1(c)} \overset{\mathcal{D}}{\sim} \ln\mathcal{N}(0, \kappa^2\sigma_1^2)$$
$$\zeta(r) := \mathrm{e}^{\kappa\sigma_2 n_2(r)} \overset{\mathcal{D}}{\sim} \ln\mathcal{N}(0, \kappa^2\sigma_2^2)$$

and $\eta(c, r)$ is such that it is satisfied $(\mathbf{c1})$. Therefore, having defined $\eta(c, r) := \mathrm{e}^{\kappa\mu(c,r)}$, they impose that

$$\begin{aligned}\mathbb{E}\{\xi(c)^2\zeta(r)^2\eta(c,r)^2\} &= \eta(c,r)^2\mathrm{e}^{\kappa^2\sigma_1^2}\mathrm{e}^{2\kappa^2\sigma_2^2} \\ &\equiv \Omega_0\mathrm{e}^{-c/\Gamma}\mathrm{e}^{-(r-c)/\gamma}\end{aligned}$$

$$\iff \mu(c,r) = \frac{1}{2\kappa}\left[\ln\Omega_0 - \frac{c}{\Gamma} - \frac{r-c}{\gamma}\right] - \kappa(\sigma_1^2 + \sigma_2^2)$$

thus

$$\eta(c,r) = \left[\Omega_0 e^{-c/\Gamma} e^{-(r-c)/\gamma}\right]^{1/2} e^{-\kappa^2(\sigma_1^2+\sigma_2^2)}$$

In this formulation, we have

$$\begin{aligned}\mathbb{E}_{\beta(c,r)}\,\beta(c,r)^2 &= \eta(c,r)^2\,\mathbb{E}_{\zeta(r)}\,\zeta(r)^2 \\ &= \Omega_0\mathrm{e}^{-c/\Gamma}\mathrm{e}^{-(r-c)/\gamma}e^{-2\kappa^2\sigma_1^2} \\ \mathbb{E}_{\xi(c)}\,\xi(c)^2 &= e^{2\kappa^2\sigma_1^2}.\end{aligned}$$

From a pragmatic point of view, the paths within a cluster, prior to be affected by cluster fading, are thought with a (conditional) second moment lower than the second moment that they have to exhibit to be compliant with $(\mathbf{c1})$, that is, after the cluster fading. Therefore, the cluster fading has the effect to rise the conditional second moment.

In general, we can write the moments as follows:

$$\begin{aligned}\mathbb{E}_{\xi(c)}\,\xi(c)^{2n} &= e^{2n^2\kappa^2\sigma_1^2} \\ \mathbb{E}_{\zeta(r)}\,\zeta(r)^{2n} &= e^{2n^2\kappa^2\sigma_2^2} \\ \mathbb{E}_{\beta(c,r)}\,\beta(c,r)^{2n} &= \eta(c,r)^{2n}\,\mathbb{E}_{\zeta(r)}\,\zeta(r)^{2n}\end{aligned}$$

(B) We define differently $\xi(\cdot)$ and $\zeta(\cdot)$ introducing a gauge parameter such that $\mathbb{E}\xi(c)^2 = 1$ and $\mathbb{E}\zeta(r)^2 = 1$:

$$\xi(c) := \mathrm{e}^{\kappa\sigma_1 n_1(c)}\mathrm{e}^{-\kappa^2\sigma_1^2} \overset{\mathcal{D}}{\sim} \ln\mathcal{N}(-\kappa^2\sigma_1^2, \kappa^2\sigma_1^2)$$
$$\zeta(r) := \mathrm{e}^{\kappa\sigma_2 n_2(r)}\mathrm{e}^{-\kappa^2\sigma_2^2} \overset{\mathcal{D}}{\sim} \ln\mathcal{N}(-\kappa^2\sigma_2^2, \kappa^2\sigma_2^2).$$

With the same of $\eta(c,r)$, we find that

$$\mu(c,r) = \frac{1}{2\kappa}\left[\ln\Omega_0 - \frac{c}{\Gamma} - \frac{r-c}{\gamma}\right],$$

thus:

$$\eta(c,r) = \left[\Omega_0 e^{-c/\Gamma} e^{-(r-c)/\gamma}\right]^{1/2}.$$

In this formulation, we have

$$\begin{aligned}\mathbb{E}_{\beta(c,r)}\,\beta(c,r)^2 &= \eta(c,r)^2\,\mathbb{E}_{\zeta(r)}\,\zeta(r)^2\\ &= \Omega_0 \mathrm{e}^{-c/\Gamma}\mathrm{e}^{-(r-c)/\gamma}\\ \mathbb{E}_{\xi(c)}\,\xi(c)^2 &= 1.\end{aligned}$$

From a pragmatic point of view, the paths within a cluster, prior to be affected by cluster fading, are thought with a (conditional) second moment equal to the second moment that they have to exhibit to be compliant with (**c1**). Therefore, the cluster fading has no effect on the second moment.

In general, we can write the moments as follows:

$$\begin{aligned}\mathbb{E}_{\xi(c)}\,\xi(c)^{2n} &= e^{2(n^2-n)\kappa^2\sigma_1^2}\\ \mathbb{E}_{\zeta(r)}\,\zeta(r)^{2n} &= e^{2(n^2-n)\kappa^2\sigma_2^2}\\ \mathbb{E}_{\beta(c,r)}\,\beta(c,r)^{2n} &= \eta(c,r)^{2n}\,\mathbb{E}_{\zeta(r)}\,\zeta(r)^{2n}\end{aligned}$$

Note that the last equality holds either in (A) and (B).

5.E Proof of Proposition 3: TR Gain Lower Bound.

Notation. For a given realization C of $\overline{\mathcal{C}}$, we write $\xi(C) := \{\xi(c)\}_{c\in C}$ and similarly for $\overline{\mathcal{R}}$. All summations over c or r are understood to be over $c\in\overline{\mathcal{C}}$ or $r\in\overline{\mathcal{R}}(c)|c$, respectively.

We want to compute

$$\mathbb{E}\sum_c\sum_r h(c,r)^{2n} = \mathbb{E}\sum_c\sum_r \xi(c)^{2n}\beta(c,r)^{2n}$$

with $n=1$ and $n=2$, where $\alpha(c,r)$ is the amplitude of the path occurred at the time r belonging to the cluster begun in c. Let us start reducing the expectations only over the point process:

$$
\begin{aligned}
&\mathbb{E}\sum_c \sum_r \xi(c)^{2n}\beta(c,r)^{2n} \\
&= \mathbb{E}_{\overline{\mathcal{C}}}\, \mathbb{E}_{\overline{\mathcal{R}}(\overline{c})|\overline{c}}\, \mathbb{E}_{\xi(\overline{c}),\beta(\overline{c},\overline{\mathcal{R}}(\overline{c}))|\overline{\mathcal{R}}(\overline{c}),\overline{c}} \sum_c \sum_r \xi(c)^{2n}\beta(c,r)^{2n} \\
&= \mathbb{E}_{\overline{\mathcal{C}}}\, \mathbb{E}_{\overline{\mathcal{R}}(\overline{c})|\overline{c}} \sum_c \mathbb{E}_{\xi(c)}\, \xi(c)^{2n}\, \mathbb{E}_{\beta(c,\overline{\mathcal{R}}(c))|\overline{\mathcal{R}}(c),c} \sum_r \beta(c,r)^{2n} \\
&= \mathbb{E}_{\xi}\, \xi^{2n}\, \mathbb{E}_{\overline{\mathcal{C}}}\, \mathbb{E}_{\overline{\mathcal{R}}(\overline{c})|\overline{c}} \sum_c \sum_r \mathbb{E}_{\beta(c,r)|r,c}\, \beta(c,r)^{2n}
\end{aligned}
$$

and further decomposing β:

$$
\begin{aligned}
&= \mathbb{E}_{\xi}\, \xi^{2n}\, \mathbb{E}_{\overline{\mathcal{C}}}\, \mathbb{E}_{\overline{\mathcal{R}}(\overline{c})|\overline{c}} \sum_c \sum_r \eta(c,r)^{2n}\, \mathbb{E}_{\zeta(r)}\, \zeta(r)^{2n} \\
&= \mathbb{E}_{\xi}\, \xi^{2n}\, \mathbb{E}_{\zeta}\, \zeta^{2n}\, \mathbb{E}_{\overline{\mathcal{C}}}\, \mathbb{E}_{\overline{\mathcal{R}}(\overline{c})|\overline{c}} \sum_c \sum_r \eta(c,r)^{2n}
\end{aligned}
$$

Now we can rely on the independence of each cluster process to obtain:

$$
= \mathbb{E}_{\xi}\, \xi^{2n}\, \mathbb{E}_{\zeta}\, \zeta^{2n}\, \mathbb{E}_{\overline{\mathcal{C}}} \sum_c \mathbb{E}_{\overline{\mathcal{R}}(c)|c} \sum_r \eta(c,r)^{2n}
$$

and the last expectation can be handled with the Campbell's formula. From Appendix 5.D, using the formulation (B), we can explicitly write

$$
\begin{aligned}
&= e^{2(n^2-n)\kappa^2(\sigma_1^2+\sigma_2^2)}\, \mathbb{E}_{\overline{\mathcal{C}}} \sum_c \mathbb{E}_{\overline{\mathcal{R}}(c)|c} \sum_r \left[\Omega_0 e^{-c/\Gamma} e^{-(r-c)/\gamma}\right]^n \\
&= \Omega_0^n e^{2(n^2-n)\kappa^2(\sigma_1^2+\sigma_2^2)}\, \mathbb{E}_{\overline{\mathcal{C}}} \sum_c e^{-nc/\Gamma}\, \mathbb{E}_{\overline{\mathcal{R}}(c)|c} \sum_r e^{-n(r-c)/\gamma}
\end{aligned}
$$

The process $\overline{\mathcal{R}}(c)|c$ has an intensity given by $\varrho_{\overline{\mathcal{R}}(c)|c}(t) = \lambda\vartheta(t-c) + \delta(t-c)$, see Section 5.5, thus

$$
\mathbb{E}_{\overline{\mathcal{R}}(c)|c} \sum_r e^{-n(r-c)/\gamma} = \int_{\mathbb{R}} e^{-n(r-c)/\gamma} \varrho_{\overline{\mathcal{R}}(c)|c}(r)dr = 1 + \lambda\gamma/n
$$

Now, the previous expression becomes

$$
= \Omega_0^n e^{2(n^2-n)\kappa^2(\sigma_1^2+\sigma_2^2)}\, (1+\lambda\gamma/n)\, \mathbb{E}_{\overline{\mathcal{C}}} \sum_c e^{-nc/\Gamma}
$$

whose expectation can be solved also recurring to the Campbell's formula, this time with $\varrho_{\overline{\mathcal{C}}}(t) = \Lambda\vartheta(t) + \delta(t)$, see Section 5.5:

$$
\mathbb{E}_{\overline{\mathcal{C}}} \sum_c e^{-nc/\Gamma} = \int_{\mathbb{R}} e^{-nc/\Gamma} \varrho_{\overline{\mathcal{C}}}(c)dc = 1 + \Lambda\Gamma/n
$$

We then arrive at the final expression of the expectation:

$$
\mathbb{E}\sum_c \sum_r h(c,r)^{2n} = \Omega_0^n r_{2n}\, (1+\lambda\gamma/n)\, (1+\Lambda\Gamma/n)
$$

having set $r_{2n} := e^{2(n^2-n)\kappa^2(\sigma_1^2+\sigma_2^2)}$.

Specializing this formula to $n=1$ and $n=2$ respectively we obtain

$$\mathbb{E}\sum_c \sum_r h(c,r)^2 = \Omega_0 \left(1+\lambda\gamma\right)\left(1+\Lambda\Gamma\right)$$
$$\mathbb{E}\sum_c \sum_r h(c,r)^4 = \Omega_0^2 e^{4\kappa^2(\sigma_1^2+\sigma_2^2)} \left(1+\lambda\gamma/2\right)\left(1+\Lambda\Gamma/2\right)$$

that are the required formulas to compute the ratio.

CHAPTER 6

Closed Form Asymptotic Expression of a Random-Access Interference Measure

A model describing the cumulative effect of the independent access of K users to a shared resource, formed by N elements, is proposed, based on which an integer interference measure ζ is defined. While traditional cases can be reconducted to reference well-known results, for which ζ is either Gaussian or Poissonian (see, for example, Appendix 1.H and eq. (1.37), respectively, and also eq. (3.52) for a Poisson-mixture distribution), the proposed model provides a framework that offers the tool for understanding the different nature of ζ. In particular, an asymptotic closed form expression ($K\to\infty$, $N\to\infty$, $K/N\to\beta\in(0,\infty)$) for ζ distribution is provided for systems presenting constructive *vs.* destructive interference, and as such is applicable to characterizing statistical properties of interference in a wide range of random multiple access channels.

6.1 System Model

When many users access a common resource independently, they may interfere with each other. A resource can be viewed, in general, as a set of elements that are used to transmit information. For example, at the physical layer, the resource is the set of degrees of freedom that carry the information-bearing signal: a system using bandwidth W for time T with N_t antennas can access WTN_t degrees of freedom belonging to time, frequency and space domains; at the medium access layer, the resource is usually time supporting either continuous or slotted packet transmission.

In the proposed model, the resource is a discrete set of N elements[1] $[1{:}N]$, *i.e.*, N *slots*. This description holds when the resource is discrete, or can be aptly discretized. Resource is shared independently by K users: user k chooses a subset

[1] In this chapter, two common notations in combinatorics are used: the set $\{m, m+1, \dots, m+n-1\}$ is denoted by $[m{:}m+n-1]$, and, when $m=1$, it is simply written as $[n]$.

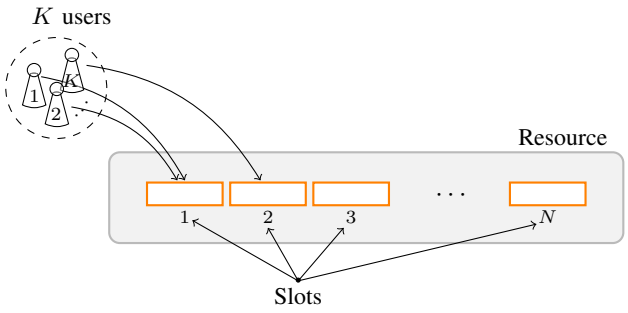


Figure 6.1: Abstract setting: a resource composed on N parts (*slots*) is randomly accessed by K users. Each user assigns to the accessed slot a numerical value (*label*), randomly and independently from other users.

$\mathcal{L}_k$ of L (irrespective of k) slots, ignoring the other users choice, and assigns a label $S^\star_{nk}$ to each accessed slot $n \in \mathcal{L}_k$.

Figure 6.1 illustrates the abstract setting, where a resource is made up of N slots and K users access randomly to a subset of slots, which is shown for $L = 1$.

Let $S_{nk} = S^\star_{nk}$ for $n \in \mathcal{L}_k$, and $S_{nk} = 0$ otherwise, that is, S_{nk} is zero for the non-accessed slots of user k, while it is equal to the assigned labels for the accessed slots. Therefore, $S^\star_{nk} \in \mathcal{S}^\star$, where $\mathcal{S}^\star$ is the set of possible values that the label may assume, and $S_{nk} \in \mathcal{S}$, where $\mathcal{S}$ and $\mathcal{S}^\star$ may or may not be equal.

Let Z_{nk} be the sum of labels assigned to slot n by all users but user k, that is:

$$Z_{nk} = \sum_{\substack{i=1 \\ i \neq k}}^{K} S_{ni}. \tag{6.1}$$

By specifying $\mathcal{S}^\star$ and L, the proposed model encompasses the problem of statistically describing multiple access interference for communication systems in which interference has a quantal nature, or can be reconducted to the model of eq. (6.1). Let present four examples: the first three address well-known problems, and are intended to clarify concepts and notations defined above, while the last addresses a novel problem, that is solved thanks to results proved in this chapter.

Example 1. This example may refer to throughput of Slotted Aloha at the medium access layer [95]. Assume that resource is time, that is slotted in N equal parts, and that K users, each willing to transmit a single packet, independently choose one of the N slots. In this setting, $L = 1$ and $S^\star_{nk}$ can be interpreted as the binary variable indicating the presence $(S^\star_{nk} = 1)$ or absence $(S^\star_{nk} = 0)$ of packet of user k within slot n. The goal is to find the number of colliding packets with user k packet, *i.e.*, the number of packets in the slot n_k selected by user k. Given n_k, it results $S^\star_{nk} = \delta_{n,n_k}$. Moreover, $\mathcal{S}^\star = \{1\}$ and $\mathcal{S} = \{0, 1\}$. Z_{nk} in eq. (6.1) indicates the number of packets in slot n that are transmitted by all users but user k, and, therefore, the goal is to find $Z_{n_k k}$. In the large-system limit, that is, as $N \to \infty$,

$K \to \infty$, and $(K/N) \to \beta \in (0, \infty)$, both Z_{nk} and $Z_{n_k k}$ are distributed according to a Poisson distribution with mean β.

Example 2. This example may describe demodulation of direct-sequence spread-spectrum (DSSS) signals with a single-user matched filter bank. Assume that resource is time, that is slotted in N chips of duration Δ, and supports transmission of synchronous random DSSS signals [6]. Let the received signal be

$$y(t) = \sum_{k=1}^{K} b_k s_k(t) + n(t), \tag{6.2}$$

where K is the number of users, $\{s_k(t)\}_{k=1}^{K}$ is the set of unit energy transmitted spreading waveforms, $\{b_k\}_{k=1}^{K}$ is the set of transmitted antipodal symbols, and $n(t)$ is a white Gaussian noise process with power spectrum σ_N^2. Direct-sequence implies $s_k(t) = \sum_{n=0}^{N-1} s_k[n]\psi(t - n\Delta)$, where $\{s_k[n] \colon k = 1, \ldots, K; n = 0, \ldots, N-1\}$ are i.i.d. r.v.s assuming values in $\{-(1/\sqrt{N}), (1\sqrt{N})\}$ with equal probability (see *e.g.* [6]). In this scenario, $L = N$, $\mathcal{S}^\star = \{-(1/\sqrt{N}), (1/\sqrt{N})\}$, $\mathcal{S}^\star = \mathcal{S}$, $S_{nk} = b_k s_k[n]$, and Z_{nk} is the signal interfering with that of user k, within chip n, at the output of a chip-matched filter. The goal is to find the distribution of Z_{nk}, that can be used, for example, in order to find the capacity of the system. In the large-system limit, Z_{nk} is Gaussian distributed with zero mean and unit variance. Generally, a Lindeberg condition suggests that if a fixed fraction, however small, of degrees of freedom is "uniformly" used, then Z_{nk} is Gaussian distributed; otherwise, if few degrees of freedom are used in the large-system limit, *e.g.* a finite number, then the asymptotic distribution may be not Gaussian, as is usually the case.

Example 3. This example may describe demodulation of binary PPM time-hopping spread-spectrum (THSS) signals with a single-user matched filter bank, where interference is at physical rather than medium access layer compared to Example 1, and can only be constructive. Assume that resource is time, that is divided in $N/2$ chips of duration Δ, and supports transmission of synchronous binary PPM THSS signals (see *e.g.* [96]). Let the received signal be

$$y(t) = \sum_{k=1}^{K} \psi(t - c_k\Delta - \epsilon b_k) + n(t), \tag{6.3}$$

where K is the number of users, c_k is uniformly distributed over $[0 \colon (N/2) - 1]$ assuming $N/2$ an integer, $\{b_k\}_{k=1}^{K}$ is the set of binary transmitted symbols, and, for the sake of simplicity, $\psi(t)$ is a zero-excess bandwidth waveform with band $[-W/2, W/2]$, and $1/W = \Delta/2 = \epsilon$. In this model, there are N slots of duration $\Delta/2$, $\mathcal{S}^\star = \{1\}$, $\mathcal{S} = \{0, 1\}$, and Z_{nk} may be regarded as the interference of the output of a filter slot-matched to slot n. In the large-system limit, Z_{nk} is distributed according to a Poisson distribution with mean β, as in the first example.

Example 4. Example 4 is similar to Example 2, except for the random spreading sequences that now belong to the time-hopping family (see *e.g.* [97,98]), *i.e.*, for any fixed k, $s_k[n] \in \{-1, 1\}$, with equal probability, for only one chip $n_k \in [0 \colon N-1]$. In this case, Z_{nk} is the interference, to which contribute both constructive and

distructive terms $b_i s_i[n]$ for $i \neq k$, at the output of a filter chip-matched to chip n, $\mathcal{S}^\star = \{-1, 1\}$, and $\mathcal{S} = \{-1, 0, 1\}$. Moreover, $Z_{n_k k}$ is the interference at the output of the single-user matched filter of user k. The distribution of $Z_{n_k k}$ is unknown, and can be found thanks to the result presented in this chapter.

As hinted by Example 4, this chapter finds, in the large-system limit, the closed form distribution of:

$$\zeta_k \triangleq Z_{n_k k} = \sum_{\substack{i=1 \\ i \neq k}}^{K} S_{n_k i}. \tag{6.4}$$

The chapter is organized as follows: in Section 6.2 the main result is presented and proved; essential analytic combinatorics are recapped in Appendix 6.A. Conclusions are drawn in Section 6.3.

6.2 Main Result

Theorem 9. *Let N be the number of slots of a resource, that is shared by K users. The generic user $k \in [1{:}K]$ selects one slot only $n_k \in [1{:}N]$, and assigns to this slot a label $S^\star_{n_k k}$ that is a r.v. taking value in $\{-1, 1\}$ with equal probability. Then, ζ_k, as defined in eq. (6.4), is distributed in the large-system limit, that is, for $N \to \infty$, $K \to \infty$, and $K/N \to \beta \in (0, \infty)$, as:*

$$P_\zeta = \sum_{m \in \mathbb{Z}} e^{-\beta} I_m(\beta)\, \delta_m, \tag{6.5}$$

irrespective of k, where $I_m(\beta)$ is the modified Bessel function of the first kind.

Proof. We provide two proofs. The first proof is probabilistic: S_{nk} is regarded as a r.v. and the pdf of ζ is derived straightforwardly via algebraic manipulations. The second proof is based on results of analytic combinatorics: the probability $\mathbb{P}(\zeta = z)$ is derived by considering all the ways, and the associated probability, a particular value z of ζ can be obtained; the law of large numbers guarantees that the so obtained result holds with probability one in the large-system limit.

First Proof.

For fixed k, ζ_k is the sum of $(K-1)$ i.i.d. random variables $\{S_{n_k i}\}_{i=1, i\neq k}^{K}$, each of which is distributed according to:

$$P_S = \frac{1}{2N}\delta_{-1} + \left(1 - \frac{1}{N}\right)\delta_0 + \frac{1}{2N}\delta_1. \tag{6.6}$$

Denoting by $b(n, p; r) = \binom{n}{r} p^r (1-p)^{n-r}$, one has:

$$P_{\zeta_k} = \sum_{i=0}^{K-1} b(K-1, 1/N; i) \frac{1}{2^i} \sum_{\ell=0}^{i} \binom{i}{\ell} \delta_{2\ell - i}.$$

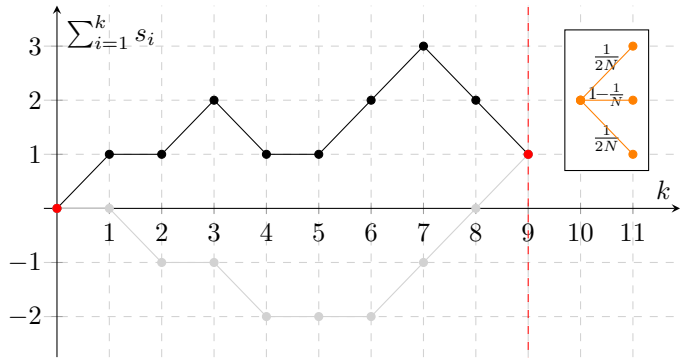


Figure 6.2: Two simple lattice walks are shown. In the two cases, $K = 10$, and therefore the walk length is 9, and steps are $\mathcal{S} = \{-1, 0, 1\}$, as shown in the box at the north-east corner. Both walks start from $(0, 0)$ and end at $(9, 1)$: the darker corresponds to the sequence $(1, 0, 1, -1, 0, 1, 1, -1, -1)$, while the lighter corresponds to $(0, -1, 0, -1, 0, 0, 1, 1, 1)$.

In order to obtain the asymptotic pdf, rewrite the previous relation as follows:

$$P_{\zeta_k} = \sum_{|m|<K} \left\{ \sum_{j=0}^{J_m} c_{|m|+2j,j} \right\} \delta_m, \quad J_m = \frac{K-1-|m|}{2},$$

where this time the contribution to the amplitude of each Dirac mass is isolated in the term in parentheses, that is defined as $c_{i\ell} = b(K-1, 1/N; i)\frac{1}{2^i}\binom{i}{\ell}$. In the large-system limit, the Binomial distribution tends to a Poisson distribution with mean β, $b(\beta N - 1, 1/N; i) \to e^{-\beta}(\beta^i/i!)$, and the term in parentheses reduces to $e^{-\beta} I_{|m|}(\beta)$, being:

$$I_m(\beta) = \sum_{i=0}^{\infty} \frac{1}{i!} \cdot \frac{1}{(m+i)!} \left(\frac{\beta}{2}\right)^{2i+m}.$$

Finally, P_{ζ_k} becomes:

$$P_\zeta = \sum_{m \in \mathbb{Z}} e^{-\beta} I_{|m|}(\beta)\, \delta_m,$$

irrespective of k. The theorem follows since $I_{|m|}(\beta) = I_m(\beta)$ when $m \in \mathbb{Z}$.

Second Proof.

As expressed by eq. (6.4), ζ_k is a sum of the form:

$$\zeta = \sum_{i=1}^{K-1} s_i, \quad s_i \in \{-1, 0, 1\},$$

where subscript k is discarded. In order to find $\mathbb{P}(\zeta = h)$, the number of ways h can be obtained as sum of elements of a sequence $(s_1, \ldots, s_{K-1})$ is counted, and let

the probability each that sequence appears be:

$$\Pr(\zeta = h) \quad = \sum_{\substack{s_k \in \mathcal{S},\, k \in [K-1] \\ \sum_{k=1}^{K-1} s_k = h}} P(s_1, \ldots, s_{K-1}). \tag{6.7}$$

Sequence $(s_1, \ldots, s_{K-1})$ can be regarded as an unconstrained, simple walk of length $K-1$ in the lattice $\mathbb{Z} \times \mathbb{Z}$ (refer to Appendix 6.A for definitions and theorems of analytic combinatorics that are used in this chapter). Figure 6.2 shows two such walks with $|\mathcal{S}| = 3$ possible steps, $\mathcal{S} = \{-1, 0, 1\}$, where notation for simple walks is adopted. Associated with these steps are weights $1/(2N)$ and $(1 - 1/N)$ (see box at north-east corner of Fig. 6.2) such that the characteristic polynomial of $\mathcal{S}$ is:

$$P(y) = \frac{1}{2N}\frac{1}{y} + \left(1 - \frac{1}{N}\right) + \frac{1}{2N} y = 1 - \frac{1}{N}\left(1 - \frac{y}{2} - \frac{1}{2y}\right).$$

Thanks to weights, the probability a particular point in $\mathbb{Z} \times \mathbb{Z}$ is reached can be computed. In order to find $\mathbb{P}(\zeta = h)$, summation in eq. (6.7) is over walks starting from $(0, 0)$ and ending at $(K-1, h)$, and the probability within the sum is that associated to each walk, that is the product of probabilities associated to steps composing the walk. The generating function of these walks is:

$$W(x, y) = \frac{1}{1 - xP(y)} = \sum_{k \geq 0} P(y)^k x^k,$$

the coefficient $[x^{K-1} y^h] W(x, y)$ giving the probability to reach $(K-1, h)$:

$$\begin{aligned} [x^{K-1} y^h] W(x, y) &= [y^h] P(y)^{K-1} \\ &= [y^h] \left[1 - \frac{1}{N}\left(1 - \frac{y}{2} - \frac{1}{2y}\right)\right]^{K-1}. \end{aligned}$$

In the large-system limit, the quantity in brackets converges to:

$$[x^{K-1}] W(x, y) \longrightarrow e^{\beta \alpha(y)} \triangleq \bar{P}(y),$$

with $\alpha(y) = 1 - y/2 - 1/(2y) = -(y-1)^2/(2y)$. In order to find $[y^h] \bar{P}(y)$ and therefore $[x^{K-1} y^h] W(x, y)$, Cauchy's integral formula can be used as follows:

$$\begin{aligned} [y^h] \bar{P}(y) &= \frac{1}{2\pi j} \oint_\gamma \frac{\bar{P}(y)}{y^h} \cdot \frac{dy}{y} = \frac{1}{2\pi j} \oint_{|y|=1} \frac{\bar{P}(y)}{y^h} \cdot \frac{dy}{y} \\ &= \frac{1}{2\pi} \int_0^{2\pi} \bar{P}(e^{j\theta}) e^{-jh\theta} d\theta \\ &= e^{-\beta} \int_0^{2\pi} e^{\beta \cos\theta} e^{-jh\theta} d\theta = e^{-\beta} I_h(\beta). \end{aligned}$$

Therefore, ζ assumes the integer value h with probability $e^{-\beta} I_h(\beta)$, hence the theorem. □

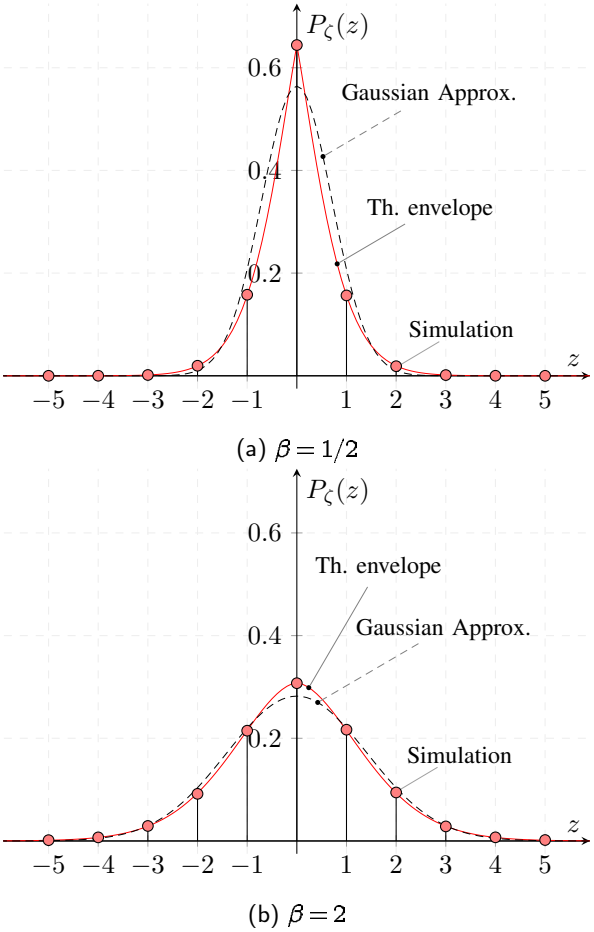


Figure 6.3: Theoretical envelope (red dashed line) *vs.* simulated histogram (filled circles) of ζ for $\beta = 1/2$ (subfig. (a)) and $\beta = 2$ (subfig. (b)). A Gaussian PDF with same mean and variance is shown for reference.

Figure 6.3 shows simulations (filled circles at integer values) *vs.* theoretical envelope $e^{-\beta} I_{|z|}(\beta)$, $z \in \mathbb{R}$ (red solid line) of P_ζ, for $\beta = 1/2$ (Fig. 6.3(a)) and $\beta = 2$ (Fig. 6.3(b)). Simulated values are drawn from Monte-Carlo simulations of 10^5 finite dimensional systems with $N = 100$. A Gaussian r.v. with same mean and variance (black dashed line) is reported for reference. As hinted by figures, the envelope of the distribution is increasingly Gaussian as β increases. In particular, odd moments of ζ are null, while the two first even moments are $\mathbb{E}\left[\zeta^2\right] = \beta$ and $\mathbb{E}\left[\zeta^4\right] = \beta(1 + 3\beta)$, hence the kurtosis is $\kappa = 3 + 1/\beta$. Since $\beta > 0$, ζ is always leptokurtic, and $\kappa \to 3$ as $\beta \to \infty$, suggesting that a Gaussian approximation may hold for $\beta \gg 1$.

6.3 Future work and Conclusion

In this chapter, a model describing systems where users access a resource independently was proposed. Each user assigns labels to accessed slots: the label is a numerical value with sign, *i.e.* accounting for polarity. Based on this model, an interference measure called ζ_k for the generic user k that considers the cumulative value of other users labels in terms of their sum was considered. In particular, the case where each user accesses one slot only and assigns a label -1 or 1 with equal probability to the accessed slot was addressed. A closed form expression of the distribution of this cumulative value was found in the large-system limit: it was shown that, if the cardinality of the population of users is a fraction β of the number of available slots, then the distribution converges to a novel expression that is in general far from Gaussian, and may be approximated by a Gaussian distribution for $\beta \gg 1$. Two proofs, one probabilistic and the other based on analytic combinatorics, were provided.

The second proof presents a potentially fruitful framework that can be used to derive several generalizations. Firstly, the labels are restricted to be binary in this work: by removing this constraint, conditions under which the interference does not follow neither a Gaussian nor a Poisson distribution can be studied. Secondly, the number of slots accessed by each user is here constrained to one, as both the total number of slots N and the population K grows at the same rate $\beta = K/N$. This constraint also can be removed, and the maximum number of slots, as a function of N, that each user may access without reducing the interference to behave as a Gaussian random variable can be studied. Lastly, as a third generalization, the relation with so-called "stable distributions," that represent distributions describing the sum of an infinite number of properly scaled i.i.d. random variables, can be investigated.

Appendix

6.A Basics on Analytic Combinatorics of Lattice Paths

We refer mostly to the seminal work of Banderier and Flajolet [99].

Definition 4 (*Lattice Path or Walk*). A *lattice path* (or walk) is a sequence $(v_1, \dots, v_n) \in \mathcal{S}^n$ where n is the *length* of the path and

$$\mathcal{S} = \{(a_1, b_1), \dots, (a_m, b_m) \colon (a_i, b_i) \in \mathbb{Z} \times \mathbb{Z}\}$$

is the set of *steps*. A path is:

- *directed* if $a_i > 0$;
- *simple* if $a_i = 1$ (in this case the set of steps is written as $\mathcal{S} = \{b_1, \dots, b_m\}$);
- *unconstrained* (resp. *constrained*) if $v \in \mathbb{Z} \times \mathbb{Z}$ (resp. $v \in \mathbb{Z} \times \mathbb{Z}_{\geq 0}$).

We can assign a *weight* to each allowed step, that is $\mathcal{S} \ni s \mapsto w(s) \in \mathbb{R}$. The following definition is the starting point of the analytic approach:

Definition 5 (*Characteristic Polynomial*). Let $\mathcal{S}$ be the set of steps of a simple walk and w_i the weight associated to b_i. The characteristic polynomial of $\mathcal{S}$ is:

$$P(y) = \sum_{i=1}^{m} w_i \, y^{b_i}.$$

The ending point of a walk is $\sum_{i=1}^{n} v_i$ that, for simple walks, assumes the form $(K - 1, h)$, where h is called *final altitude*. Denote by $\mathcal{W}_{nk}$ the class of walks with length n and final altitude k, and let $W_{nk} = |\mathcal{W}_{nk}|$.

Definition 6 (*Generating Function*). The *generating function* of $\mathcal{W}_{nk}$ is defined as:

$$W(x, y) = \sum_{n,k} W_{nk} \, x^n y^k,$$

where $x \in \mathbb{C}$ is a mark for the length and $y \in \mathbb{C}$ is a mark for the final altitude.

The following theorem links $W(x, y)$ with $P(y)$:

Theorem 10. *The generating function of a simple walk is:*

$$W(x, y) = \frac{1}{1 - xP(y)}. \tag{6.8}$$

Proof (sketch). Rewrite $W(x,y)$ as follows:

$$W(x,y) = \sum_n \left[\sum_k W_{nk}\, y^k \right] x^n = \sum_n w_n(y) x^n,$$

where $w_n(y) = [x^n]W(x,y)$ is a Laurent polynomial in y where $[y^h]w_n(y)$ is the (possibly weighted) number of ways to reach the final altitude h in n steps. Since the only altitude reachable in 0 steps is 0, then $w_0(y) = 1$; at step 1, the reachable altitudes are described by $P(y)$. In general $w_k(y) = w_{k-1}(y)P(y)$, and therefore a summation over $k \geq 0$ yields to eq. (6.8).

CHAPTER 7

Conclusion and Future works

Impulsive communications belong to the subset of communications that are sparse in time. Impulsiveness is, therefore, strictly related to both sparsity in time of the transmitted signal and properties of the physical medium that the signal passes through. Impulsive communications are those where the received signal occupies a fraction of the received signal space. The typical example in the wireless communication field is represented by impulse-radio ultra-wideband, where transmissions occur at bursts and the transmitted signal occupies a small fraction of the symbol period.

In this thesis we addressed several issues concerning impulsive communications and the effect of interference on their performance. In particular, we formalized essential characteristics of multiuser impulsive communications, namely sparsity and random hopping, and we identified their information-theoretical limits in multiple access channels with power-control. We then investigated the interplay between the transmitted signal bandwidth and the statistical properties of a multipath channel, when transmitter uses a prefiltering scheme called transmit matched-filter (also known as time reversal in ultra-wideband and speech communities). Several works on the analysis of performance of systems using time reversal is then presented, for both SISO and MISO systems, and in presence of imperfect knowledge of the channel. Finally, the interference of a general multiple access model accounting for constructive *vs.* destructive events is investigated: we showed that several systems can be described with this model, and a new asymptotic distribution formula for the interference is derived.

7.1 Conclusion

We summarize below the main conclusions that can be drawn from each chapter of this thesis.

In Chapter 1, information-theoretical limits are derived for impulsive *vs.* non-impulsive multiuser communications, represented by time-hopping *vs.* direct-sequence CDMA, that are prototypes of sparse *vs.* dense multiple access formats. We compared spectral efficiency of random time-hopping with consolidated results

regarding direct-sequence. We found that there is a regime where time-hopping allows to achieve spectral efficiency higher than that achievable by direct-sequence, namely when the system is overloaded and the receiver is linear. This result provides a hint for and a partial justification to designing networks with nodes that transmit bursty signals in time: this signaling scheme is appropriate whenever the network load is high, $\beta \gg 1$, while keeping the receiver design simple, *e.g.* owing to very large bandwidths: the performance gain with respect to traditional direct-sequence networks is increasingly relevant as $\mathcal{E}_b/\mathcal{N}_0$ increases.

In Chapter 2, we investigated the interplay between transmitted signal bandwidth and statistical properties of a multipath channel, when transmitter uses a prefiltering scheme called transmit matched-filter (also known as time reversal in ultra-wideband and speech communities). We derived a necessary and sufficient condition for the performance of the system to increase as bandwidth increases. We showed that multipath channels that are described by or strictly related to the Saleh-Valenzuela model, satisfy the condition, therefore suggesting a theoretical justification for the wide adoption of time reversal in ultra-wideband communications.

Chapters 3–5 are dedicated to detailed study of time reversal in various settings. In Chapter 3, we analyzed performance of two different transceiver structures, namely transceivers with time reversal prefiltering and 1-Rake receiver *vs.* no prefiltering and All-Rake receiver. Performance is analyzed in terms of error probability in both single-user and multiuser settings; preliminary mutual information analyses has been also conducted. It was shown that in the single-user setting the two system designs are equivalent in terms of error probability and sensitivity to perturbations, while behaving differently in the multiuser setting, where the system using time reversal shows higher error probability. Partial results on mutual information do not change qualitatively the conclusion drawn from the comparison based on the error probability metric. In Chapter 4, we analyzed MISO ultra-wideband with time reversal, in particular showing that performance is not affected by the the lack of correlation between channels from each transmitting antenna to the receiver. In Chapter 5, we compared SISO time reversal with other prefilters, as a function of the number of fingers of a Rake receiver, and we showed the gain that can be obtained by the different transceiver structures.

Finally, in Chapter 6 we analyzed a general multiple access model where users can interfere with each other antipodally in either a constructive or destructive fashion. We detailed few settings of practical relevance that are described by this model, but several others, including queues in particular regimes, may be aptly described as well. We found the asymptotic distribution of the interference in closed form, that cannot be reduced to any known probability distribution.

7.2 Future works

In several chapters of this thesis, we already detailed future works that are related to the content of the chapter. The kind of detailed works aims at extending results obtained under certain sets of assumptions, and can form the basis for investigations

that either complete or extend the scope of the works that are presented in the thesis. We hope also that several other works, different from the ones that we explicitly proposed, will be inspired by this thesis.

In this section we briefly discuss future works that necessitate models that are either more general or intrinsically different with respect to those presented in this thesis. Firstly, although this thesis is focused on impulsive communications, that are sparse in time, it is interesting to investigate more general cases of sparsity across dimensions of the signal space. Secondly, it is interesting to investigate the sum-rate of a random ad hoc network where each node uses sparse spreading, and compare results with those achieved with dense spreading. Different scaling laws are envisioned in the sparse case. Finally, we believe that the most ambitious future work lies in embracing a different perspective, namely the possibility to formalize impulsive communications as those that regard time not as a mere support of the communications, but as the resource within which coding information through inter-arrival times between events, such as signal receptions. In this regard, [100–102] are seminal contributions that may form the basis for this ambitious project.

Bibliography

[1] M. Medard and R. G. Gallager, "Bandwidth scaling for fading multipath channels," *IEEE Trans. Inf. Theory*, vol. 48, no. 4, pp. 840–852, 2002.

[2] R. Gallager and M. Medard, "Bandwidth scaling for fading channels," in *Proc. IEEE Int. Symp. Inf. Theory*, 1997, p. 471.

[3] E. Telatar and D. Tse, "Capacity and mutual information of wideband multipath fading channels," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1384–1400, 2000.

[4] D. Porrat, D. Tse, and S. Nacu, "Channel uncertainty in ultra-wideband communication systems," *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 194–208, 2007.

[5] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: information-theoretic and communications aspects," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2619–2692, 1998.

[6] S. Verdù and S. Shamai, "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 622–640, 1999.

[7] S. Shamai and S. Verdù, "The impact of frequency-flat fading on the spectral efficiency of CDMA," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1302–1327, 2001.

[8] D. Tse and S. Hanly, "Linear multiuser receivers: effective interference, effective bandwidth and user capacity," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 641–657, 1999.

[9] D. Tse and O. Zeitouni, "Linear multiuser receivers in random environments," *IEEE Trans. Inf. Theory*, vol. 46, no. 1, pp. 171–188, 2000.

[10] A. M. Tulino, L. Li, and S. Verdu, "Spectral efficiency of multicarrier CDMA," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 479–505, 2005.

[11] V. A. Marcenko and L. A. Pastur, "Distribution of eigenvalues for some sets of random matrices," *Mathematics of the USSR-Sbornik*, vol. 1, no. 4, pp. 457–483, 1967.

[12] T. Tanaka, "A statistical-mechanics approach to large-system analysis of CDMA multiuser detectors," *IEEE Trans. Inf. Theory*, vol. 48, no. 11, pp. 2888–2910, 2002.

[13] A. Montanari and D. Tse, "Analysis of belief propagation for non-linear problems: The example of CDMA (or: How to prove tanaka's formula)," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Punta del Este, 2006, pp. 160–164.

[14] D. Guo and C.-C. Wang, "Multiuser detection of sparsely spread CDMA," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 3, pp. 421–431, 2008.

[15] T. Tanaka and M. Okada, "Approximate belief propagation, density evolution, and statistical neurodynamics for CDMA multiuser detection," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 700–706, 2005.

[16] K. Alishahi, F. Marvasti, V. Aref, and P. Pad, "Bounds on the sum capacity of synchronous binary CDMA channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3577–3593, 2009.

[17] S. Korada and N. Macris, "Tight bounds on the capacity of binary input random CDMA systems," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5590–5613, 2010.

[18] J. Raymond and D. Saad, "Sparsely spread CDMA—a statistical mechanics-based analysis," *J. Phys. A: Math. Theor.*, vol. 40, no. 41, pp. 12 315–12 333, 2007.

[19] M. Yoshida and T. Tanaka, "Analysis of sparsely-spread CDMA via statistical mechanics," in *Proc. IEEE Int. Symp. Inf. Theory*, 2006, pp. 2378–2382.

[20] S. Verdù, *Multiuser Detection*, 1st ed. New York: Cambridge University Press, 1998.

[21] M. B. Pursley, "Performance evaluation for phase-coded spread-spectrum multiple-access communication–Part I: System analysis," *Communications, IEEE Transactions on*, vol. 25, no. 8, pp. 795–799, 1977.

[22] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 1, pp. 1–182, 2004.

[23] A. M. Tulino, *The Eta and Shannon transforms: a bridge between random matrices and wireless communications*, ser. Lecture notes. World Scientific Publishing, 2009, vol. 18, ch. 4, pp. 95–138.

[24] R. Couillet and M. Debbah, *Random matrix methods for wireless communications*. Cambridge University Press, 2011.

[25] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319–1343, 2002.

[26] Z. Bai and J. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices*, 2nd ed., ser. Springer Series in Statistics. New York: Springer, 2010.

[27] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*. Cambridge University Press, 2009.

[28] R. Sedgewick and P. Flajolet, *An Introduction to the Analysis of Algorithms*. Boston, MA, USA: Pearson Education, 2013.

[29] Z. D. Bai and J. W. Silverstein, "No eigenvalues outside the support of the limiting spectral distribution of large dimensional sample covariance matrices," *Ann. Probab.*, vol. 26, pp. 316–345, 1998.

[30] S. Verdu, "The capacity region of the symbol-asynchronous gaussian multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 35, no. 4, pp. 733–751, 1989.

[31] S. Benedetto and E. Biglieri, *Principles of Digital Transmission: With Wireless Applications*. Norwell, MA, USA: Kluwer Academic Publishers, 1999.

[32] A. Lozano and D. Porrat, "Non-peaky signals in wideband fading channels: Achievable bit rates and optimal bandwidth," *IEEE Trans. Wireless Commun.*, vol. 11, no. 1, pp. 246–257, 2012.

[33] J. Richters, *Communication over fading dispersive channels*, ser. Technical report. MIT, RLE, 1967.

[34] I. Abou-Faycal, M. Trott, and S. Shamai, "The capacity of discrete-time memoryless rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1290–1301, 2001.

[35] R. S. Kennedy, *Fading dispersive communication channels*. Wiley-Interscience, 1969.

[36] L. Cottatellucci, R. Muller, and M. Debbah, "Asynchronous CDMA systems with random spreading: Part i: Fundamental limits," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1477–1497, 2010.

[37] V. Girko, *Theory of Stochastic Canonical Equations: Volumes I and II*, ser. Mathematics and Its Applications. Springer Netherlands, 2001.

[38] D. Guo, S. Shamai, and S. Verdu, "Mutual information and minimum mean-square error in gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1261–1282, 2005.

[39] A. Tulino, G. Caire, S. Shamai, and S. Verdú, "Capacity of channels with frequency-selective and time-selective fading," *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp. 1187–1215, 2010.

[40] F. Hlawatsch and G. Matz, *Wireless communications over rapidly time-varying channels*. Academic Press, 2011.

[41] R. P. Stanley, *Enumerative combinatorics*. Cambridge University Press, 2011, vol. 1.

[42] ——, *Enumerative combinatorics.* Cambridge University Press, 2011, vol. 2.

[43] I. S., "On the capacity of channels with additive non-gaussian noise," *Information and Control*, vol. 37, no. 1, pp. 34–39, 1978.

[44] M. Joham, W. Utschick, and J. Nossek, "On the equivalence of prerake and transmit matched filter," in *Proc. 10th Aachen Symp. Signal Theory*, Aachen, Germany, 2001, pp. 313–318.

[45] ——, "Linear transmit processing in MIMO communications systems," *IEEE Trans. on Signal Process.*, vol. 53, no. 8, pp. 2700–2712, 2005.

[46] M. Joham, K. Kusume, M. H. Gzara, W. Utschick, and J. A. Nossek, "Transmit Wiener filter for the downlink of TDD DS-CDMA systems," in *Proc. IEEE 7th Int. Symp. Spread Spectr. Techn. and Appl. (ISSSTA)*, Prague, Czech Republic, 2002, pp. 9–13.

[47] A. Derode, P. Roux, and M. Fink, "Robust acoustic time reversal with high-order multiple scattering," *Phys. Rev. Lett.*, vol. 75, pp. 4206–4209, 1995.

[48] M. Fink, G. Montaldo, and M. Tanter, "Time reversal acoustics," in *Proc. IEEE Ultrason. Symp.*, vol. 2, Montreal, Canada, 2004, pp. 850–859.

[49] T. Strohmer, M. Emami, J. Hansen, G. Papanicolaou, and A. Paulraj, "Application of time-reversal with MMSE equalizer to UWB communications," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, vol. 5, Dallas, TX, USA, 2004, pp. 3123–3127.

[50] R. Esmailzadeh and M. Nakagawa, "Pre-rake diversity combination for direct sequence spread spectrum communications systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, vol. 1, Geneva, Switzerland, 1993, pp. 463–467.

[51] ——, "Pre-rake diversity combination for direct sequence spread spectrum mobile communications systems," *IEICE Trans. Commun.*, vol. 76, no. 8, pp. 1008–1015, 1993.

[52] G. C. Ferrante, "Time-reversal against optimum precoder over frequency-selective channels," in *Proc. 18th European Wireless Conf.*, Poznan, Poland, 2012, pp. 1–8.

[53] J. Foerster, "Channel modeling sub-committee report final," IEEE, Tech. Rep. Doc. P802.15-02/490r1-SG3a, 2003.

[54] W.-D. Wu, C.-C. Lee, C.-H. Wang, and C.-C. Chao, "Signal-to-interference-plus-noise ratio analysis for direct-sequence ultra-wideband systems in generalized saleh-valenzuela channels," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 3, pp. 483–497, 2007.

[55] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Wiley, 2006.

[56] Y.-H. Kim, "Feedback capacity of stationary gaussian channels," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 57–85, 2010.

[57] S. A. Schelkunoff, "Some equivalence theorems of electromagnetics and their application to radiation problems," *Bell System Technical Journal*, vol. 15, no. 1, pp. 92–112, 1936.

[58] ——, "On diffraction and radiation of electromagnetic waves," *Phys. Rev.*, vol. 56, pp. 308–316, 1939.

[59] J. A. Stratton and L. J. Chu, "Diffraction theory of electromagnetic waves," *Phys. Rev.*, vol. 56, pp. 99–107, 1939.

[60] Y. F. Chen, "A mathematical formulation of the equivalence principle," *IEEE Tran. Microw. Theory Techn.*, vol. 37, no. 10, pp. 1576–1581, 1989.

[61] M. Fink, "Time reversal of ultrasonic fields. i. basic principles," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 39, no. 5, pp. 555–566, 1992.

[62] Z. Tang and S. Cheng, "Interference cancellation for DS-CDMA systems over flat fading channels through pre-decorrelating," in *Proc. IEEE Int. Symp. Personal, Indoor and Mobile Radio Commun.*, vol. 2, 1994, pp. 435–438.

[63] B. R. Vojcic and W. M. Jang, "Transmitter precoding in synchronous multiuser communications," *IEEE Trans. Commun.*, vol. 46, no. 10, pp. 1346–1355, 1998.

[64] A. N. Barreto and G. Fettweis, "Capacity increase in the downlink of spread spectrum systems through joint signal precoding," in *Proc. IEEE Int. Conf. Commun. (ICC)*, vol. 4, 2001, pp. 1142–1146.

[65] M. Vu and A. Paulraj, "Mimo wireless linear precoding," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 86–105, 2007.

[66] D. P. Palomar and Y. Jiang, "MIMO transceiver design via majorization theory," *Found. Trends Commun. Inf. Theory*, vol. 3, no. 4, pp. 331–551, 2006.

[67] A. E. Akogun, R. C. Qiu, and N. Guo, "Demonstrating time reversal in ultra-wideband communications using time domain measurements," in *Int. Instrumentation Symp.*, Knoxville, TN, USA, 2005, pp. 8–12.

[68] N. H. T., J. B. Andersen, G. F. Pedersen, P. Kyritsi, and P. C. F. Eggers, "Time reversal in wireless communications: a measurement-based investigation," *IEEE Trans. Wireless Commun.*, vol. 5, no. 8, pp. 2242–2252, 2006.

[69] R. C. Qiu, C. Zhou, N. Guo, and J. Q. Zhang, "Time reversal with MISO for ultrawideband communications: Experimental results," *IEEE Antennas Wireless Propag. Lett.*, vol. 5, no. 1, pp. 269–273, 2006.

[70] C. Zhou, N. Guo, and R. C. Qiu, "Time-reversed ultra-wideband (UWB) multiple input multiple output (MIMO) based on measured spatial channels," *IEEE Trans. Veh. Technol.*, vol. 58, no. 6, pp. 2884–2898, 2009.

[71] N. Guo, B. M. Sadler, and R. C. Qiu, "Reduced-complexity uwb time-reversal techniques and experimental results," *IEEE Trans. Wireless Commun.*, vol. 6, no. 12, pp. 4221–4226, 2007.

[72] R. Qiu, J. Zhang, and N. Guo, "Detection of physics-based ultra-wideband signals using generalized rake with multiuser detection (MUD) and time-reversal mirror," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 4, pp. 724–730, 2006.

[73] R. C. Qiu, "Generalized time domain multipath channel and its application in ultra-wideband (UWB) wireless optimal receiver—Part III: system performance analysis," *IEEE Trans. Wireless Commun.*, vol. 5, no. 10, pp. 2685–2695, 2006.

[74] G. C. Ferrante, J. Fiorina, and M.-G. Di Benedetto, "Complexity reduction by combining time reversal and IR-UWB," in *Proc. IEEE Wireless Commun. and Networking Conf. (WCNC)*, Paris, France, 2012, pp. 28–31.

[75] ——, "Time reversal beamforming in MISO-UWB channels," in *Proc. IEEE Int. Conf. Ultra-Wideband (ICUWB)*, Sydney, Australia, 2013, pp. 261–266.

[76] L. De Nardis, J. Fiorina, D. Panaitopol, and M.-G. Di Benedetto, "Combining UWB with time reversal for improved communication and positioning," *Telecommunication Systems*, vol. 52, no. 2, pp. 1145–1158, 2013.

[77] X. Wang and J. Wang, "Effect of imperfect channel estimation on transmit diversity in CDMA systems," *IEEE Trans. Veh. Technol.*, vol. 53, no. 5, pp. 1400–1412, 2004.

[78] J. Fiorina, G. Capodanno, and M.-G. Di Benedetto, "Impact of time reversal on multi-user interference in IR-UWB," in *Proc. IEEE Int. Conf. Ultra-Wideband (ICUWB)*, 2011, pp. 415–419.

[79] E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications*. New York: Cambridge University Press, 2008.

[80] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2845–2866, 2010.

[81] D. Sarwate and M. Pursley, "Crosscorrelation properties of pseudorandom and related sequences," *Proc. IEEE*, vol. 68, no. 5, pp. 593–619, 1980.

[82] M. A., "On pseudo-random and orthogonal binary spreading sequences," *Int. J. Information Techn.*, vol. 2, no. 12, pp. 543–550, 2008.

[83] E. Biglieri, *Coding for Wireless Channels.* New York: Springer-Verlag, 2005.

[84] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions: with formulas, graphs, and mathematical tables.* New York: Dover Publications, 1964.

[85] H. Saghir, M. Heddebaut, F. Elbahhar, A. Rivenq, and J. Rouvaen, "Time-reversal UWB wireless communication-based train control in tunnel," *Journal of Communications*, vol. 4, no. 4, pp. 248–256, 2009.

[86] D. Panaitopol, J. Fiorina, and M. Di Benedetto, "Trade-off between the number of fingers in the prefilter and in the rake receiver in time reversal ir-uwb," in *Proc. IEEE Int. Conf. Ultra-Wideband (ICUWB)*, 2009, pp. 819–823.

[87] M.-G. Di Benedetto and G. Giancola, *Understanding Ultra Wide Band radio fundamentals.* Prentice Hall, 2004.

[88] X. Hong, C.-X. Wang, B. Allen, and W. Q. Malik, "Correlation-based double-directional stochastic channel model for multiple-antenna ultra-wideband systems," *IET microw., antennas and propag.*, vol. 1, no. 6, pp. 1182–1191, 2007.

[89] M. L. Jakobsen, T. Pedersen, and B. H. Fleury, "Analysis of the stochastic channel model by saleh & valenzuela via the theory of point processes," in *International Zurich Seminar on Communications*, 2012, pp. 115–118.

[90] J. A. Gubner and K. Hao, "A computable formula for the average bit error probability as a function of window size for the ieee 802.15.3a uwb channel model," *IEEE Tran. Microw. Theory Techn.*, vol. 54, no. 4, pp. 1762–1768, 2006.

[91] K. Hao and J. A. Gubner, "The distribution of sums of path gains in the ieee 802.15.3a uwb channel model," *IEEE Trans. Wireless Commun.*, vol. 6, no. 3, pp. 811–816, 2007.

[92] R. Price and P. Green, "A communication technique for multipath channels," *Proc. IRE*, vol. 46, no. 3, pp. 555–570, 1958.

[93] J. G. Proakis and M. Salehi, *Digital Communications.* McGraw-Hill Higher Education, 2008.

[94] A. A. M. Saleh and R. A. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE J. Sel. Areas Commun.*, vol. 5, no. 2, pp. 128–137, 1987.

[95] N. Abramson, "The throughput of packet broadcasting channels," *IEEE Trans. Commun.*, vol. 25, no. 1, pp. 117–128, 1977.

[96] G. Giancola and M.-G. Di Benedetto, "A novel approach for estimating multi-user interference in impulse radio UWB networks: The pulse collision model," *Signal Processing*, vol. 86, no. 9, pp. 2185 – 2197, 2006.

[97] G. C. Ferrante and M.-G. Di Benedetto, "Spectral efficiency of random time-hopping CDMA," *submitted to the IEEE Trans. Inf. Theory*, 2013. Revised on Nov. 2014, [Online]. Available: http://arxiv.org/abs/1311.4830.

[98] J. Fiorina and W. Hachem, "On the asymptotic distribution of the correlation receiver output for time-hopped uwb signals," *IEEE Trans. on Signal Process.*, vol. 54, no. 7, pp. 2529–2545, 2006.

[99] B. C. and F. P., "Basic analytic combinatorics of directed lattice paths," *Theoretical Computer Science*, vol. 281, no. 1–2, pp. 37–80, 2002.

[100] V. Anantharam and S. Verdú, "Bits through queues," *IEEE Trans. Inf. Theory*, vol. 42, no. 1, pp. 4–18, 1996.

[101] R. Sundaresan and S. Verdu, "Robust decoding for timing channels," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 405–419, 2000.

[102] R. Sundaresan and S. Verdú, "Capacity of queues via point-process channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2697–2709, 2006.