Study and mitigation techniques of RF impairments for beyond 5G multi-carrier waveforms

Xinying Cheng

▶ **To cite this version:**

Xinying Cheng. Study and mitigation techniques of RF impairments for beyond 5G multi-carrier waveforms. Networking and Internet Architecture [cs.NI]. Sorbonne Université, 2021. English. NNT : 2021SORUS181 . tel-03639457

HAL Id: tel-03639457

https://theses.hal.science/tel-03639457v1

Submitted on 13 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SORBONNE UNIVERSITÉ
CRÉATEURS DE FUTURS DEPUIS 1257

le cnam

École doctorale Informatique, Télécommunications et Électronique
Centre d'Études et de Recherche en Informatique et Communications

THÈSE DE DOCTORAT

présentée par : **Xinying CHENG**
soutenue le : **11 Octobre 2021**

pour obtenir le grade de : **Docteur de Sorbonne Université**

Spécialité : **Informatique, Télécommunications et Électronique**

Study and mitigation techniques of RF impairments for beyond 5G multi-carrier waveforms

THÈSE dirigée par
M. ROVIRAS Daniel — *Professeur, CNAM*

RAPPORTEURS
Mme. LIENARD Martine — *Professeur, Université de Lille*
M. DENEIRE Luc — *Professeur, Université Côte d'Azur*

EXAMINATEURS
M. ALOUINI Mohamed Slim — *Professeur, King Abdullah University of Science and Technology*
M. LOUET Yves — *Professeur, Centrale Supélec*

ENCADRANT
M. SHAIEK Hmaied — *Maître de conférences, CNAM*
M. ZAYANI Rafik — *Maître de conférences, HDR, Innov'COM/Sup'Com*

Acknowledgements

In the long journey of life, this difficult but happy journey, the bit by bit and full growth in my heart, gave me precious life wealth. On the occasion of the completion of this paper, I would like to express my sincerest gratitude to all those who have supported and helped me!

First of all, I want to thank my supervisors : Prof. Daniel ROVIRAS, Dr. Rafik ZAYANI and Dr. Hmaied SHAIEK. All the work of the thesis was completed under their careful guidance and care. From the selection of the topic, conception, research, writing, review and finalization of the paper, they devoted a lot of effort and energy. Not only academically, their positive and optimistic attitude towards life, the approachable personality charm, the modest way of dealing with things, and the calm and broad mind also impressed me a lot. Here, I would like to extend my high respect and heartfelt thanks to them.

During the PhD study, there are too many people who need to be thanked, and my feeling of gratitude is so overwhelming. I would like to sincerely thank all the families, friends and colleagues who have supported and helped me.

In the future journey of life, I will carry your care and encouragement, and return all the care, support, and help to others.

Table des matières

Liste des tableaux

Table des figures

Nomenclature

Abbreviations and Acronyms

ACLR	Adjacent Channel Leakage Ratio
AM/AM	Amplitude to Amplitude
AM/PM	Amplitude to Phase
AP	Autoprecoder
BS	Base Station
BER	Bit Error Rate
CCDF	Cumulative Complementary Distribution Function
CDMA	Code Division Multiple Access
CP	Cyclic Prefix
CSI	Channel State Information
DDR	Dynamic Deviation Reduction
DFT	Discrete Fourier Transform
DL	Deep Learning
DoFs	Degrees of Freedom
DPC	Dirty Paper Coding
DPD	Digital Pre-Distortion
eMBB	enhanced Mobile Broadband
EVM	Error Vector Magnitude
FDMA	Frequency Division Multiple Access
FFT	Fast Fourier Transform
GD	Gradient Descent
HMP	Hybrid Memory Polynomial
IBO	Input Back-Off
ICI	Inter-Carrier-Interference
IDFT	Inverse Discrete Fourier Transform
i.i.d.	independent and identically distributed
IFFT	Inverse Fast Fourier Transform
IMD	Inter-Modulation Distortion
INI	inter-numerology interference
Intra-NI	Intra-Numerology Interference
ISI	Inter-Symbol-Interference

LOS	Line-of-Sight
LS	Least Square
LTE	Long Term Evolution
LTE-A	Long-Term Evolution Advanced
LUT	Look-up Table
mMTC	massive Machine Type Communications
MIMO	Multiple-Input Multiple-Output
MISO	Multiple-Input Single-Output
ML	Machine Learning
MLP	Multi-Layer Perception
MMSE	Minimum Mean Square Error
MP	Matrix Polynomials
MRC	Maximum Ratio Combining
MRT	Maximum Ratio Transmission
MSE	Mean Square Error
MU	Multi-User
MUI	Multi-User Interference
NL	Non Linear
NLD	Nonlinear Distortion
NLOS	Non-Line-of-Sight
NMSE	Normalized Mean Square Error
NN	Neural Networks
NOMA	Non-Orthogonal Multiple Access
OBR	Out-Band Radiation
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
PA	Power Amplifier
PAPR	Peak-to-Average Ratio
PD	Pre-Distortion
PDF	Probability Density Function
PSD	Power Spectral Density
PTS	Partial Transmit Sequence
PU	Primary User
QAM	Quadrature Amplitude Modulation
RF	Radio Frequency
SCS	Subcarrier Spacing
SDMA	Space Division Multiple Access
SER	Symbol Error Rate
SISO	Single-Input Single-Output
SLM	Selective Mapping
SNR	Signal-to-Noise Ratio
SS	Spectrum Sharing
SU	Secondary User
TDD	Time Division Duplexing
TDMA	Time Division Multiple Access

TI	Tone Injection
TR	Tone Reservation
TV	Time Varying
UE	User's Equipment
URLLC	Ultra-Reliable and Low Latency Communications
V-BLAST	Vertical Bell Laboratories Layered Space-Time
VP	Vector Perturbation
WLAN	Wireless Local Area Networks
Wi-Fi	Wireless Fidelity
WiMAX	Worldwide Interoperability for Microwave Access
ZF	Zero Forcing
3GPP	Third Generation Partnership Project
4G	Fourth Generation
5G	Fifth Generation

Notations

x	Scalar.
$\mathbf{x}$	Column vector.
$\mathbf{x}^t$	Row vector.
$\mathbf{X}$	Matrix.
$\bar{\mathbf{X}}$	3-dimension matrix.
$[\cdot]^T$	The transpose operator.
$[\cdot]^H$	The complex conjugate (Hermitian) transpose operator.
$\mathbf{X}^{-1}$	The inverse of matrix $\mathbf{X}$.
$\mathbf{X}^\dagger$	The pseudo-inverse of matrix $\mathbf{X}$.
$\mathbf{I}_N$	The identity matrix of size $N \times N$.
$\mathbf{0}_{M\times N}$	The all zero matrix of size $M \times N$.
$\mathbf{DFT}_N$	DFT matrix of size $N \times N$.
$\mathbf{IDFT}_N$	IDFT matrix of size $N \times N$.
$Diag(\mathbf{x})$	Matrix where the diagonal elements are vector $\mathbf{x}$ and all off-diagonal elements are zero.
$\| \cdot \|_2$	l_2-norm of vector.
$\lceil x \rceil$	The ceil operator, i.e. the smallest integer greater than or equal to x.
$\mathbb{E}\{\cdot\}$	The expectation operator.
$\mid \chi \mid$	The cardinality of set χ.
χ^c	The complement of set χ.

Résumé détaillé de la thèse

La recherche fondamentale vers les réseaux cellulaires au-delà de la 5G est en cours et la vision de 2020 et au-delà comprend un nombre important de cas d'utilisation compte tenu d'un grand nombre d'appareils avec un large éventail de caractéristiques et de demandes. D'une part, la bonne efficacité spectrale donne plus de capacité pour la transmission tandis que d'autre part, la construction d'équipements radio compacts et peu coûteux, flexibles et de haute qualité est une tâche très difficile. Le contexte des travaux de cette thèse est l'étude des techniques de type massive-MIMO en présence d'imperfections RF, notamment celles apportées par les amplificateurs de puissance (PA) non linéaires (NL). Le nombre important de PA dans une station de base de type massive-MIMO crée différentes contraintes : (1) pour des raisons de coût les différents PA peuvent présenter des caractéristiques NL marquées et ainsi dégrader fortement les signaux émis, (2) pour des raisons d'efficacité spectrale, les PA doivent être opérés dans une zone proche de la saturation du PA afin de maximiser leur rendement énergétique. Le résultat de ces deux contraintes entraine ainsi de fortes dégradation des signaux émis tant dans la bande (BER et EVM) que hors bande avec des remontées spectrales. D'autre part, la normalisation 5G a introduit le concept de différentes numérologies conjointement à la technique massive-MIMO. L'influence des différentes numérologies utilisées par les différents utilisateurs sera aussi étudiée dans cette thèse. Le premier chapitre est un chapitre d'état de l'art qui présente les différents éléments utilisés tout au long du travail de thèse : la technique massive-MIMO et les différents précodeurs associés, les différentes numérologies en 5G, les canaux de transmission multitrajets et les caractéristiques des amplificateurs de puissance non-linéaires. Le chapitre 2 est dédié à l'étude de interférences apportées par des utilisateurs utilisant des numérologies différentes en massive-MIMO et à la proposition d'une méthode de réduction de ces interférences. Le troisième chapitre traite de l'effet des PA non linéaires sur les signaux massive-MIMO. Une nouvelle technique de réduction des imperfections apportées par les PA NL est proposée et comparée aux techniques de la littérature.

Enfin le dernier chapitre traite de ce même problème mais avec une approche algorithmique différente. Les algorithmes utilisés dans le chapitre 3 sont de type classique (descente de gradient) alors que l'approche utilisée dans le chapitre 4 est basée sur les nouvelles techniques d'apprentissage profond plus couramment appelées techniques de deep learning. Une nouvelle architecture d'auto-precodeur est ainsi proposée et comparée aux techniques plus classiques en termes de performances et de complexité.

Cette thèse, dans un premier temps, étudie l'utilisation du multiplexage spatial d'utilisateurs, partageant la même bande passante, dont les numérologies associées sont différentes. Une technique de précodage est proposée, qui vise à gérer la transmission et le partage du spectre d'utilisateurs avec des numérologies mixtes (SS). Ensuite, l'interférence inter-numérologie (INI) est analysée et les expressions théoriques de l'INI sont dérivées dans des systèmes MIMO-OFDM massifs.

Ensuite, en termes d'équipement radio à bas coût, un soin particulier est apporté à l'amplificateur de puissance (PA) non linéaire dans la chaîne de communication sans fil. La thèse étudie les caractéristiques du PA, ainsi que ses effet non linéaire sur les transmissions massive-MIMO. Nous proposons une nouvelle technique originale permettant de combattre les effets des PA et comparons celle-ci aux techniques existantes dans la littérature, en termes de performance et de complexité d'implantation. L'objectif est de concevoir un schéma de transmission prenant en considération le précodage multi-utilisateurs (MU), la réduction du rapport de puissance crête-à-moyenne (PAPR) et la pré-distorsion numérique (DPD), en accordant une attention particulière à la complexité de calcul.

Chapter 1 : Etat de l'art et contexte technique

Le chapitre 1 présente le contexte technique des travaux de recherche avec tous les éléments nécessaires pour la suite des travaux : le massive-MIMO et les précodeurs associés, le concept des numérologies en 5G, les canaux multitrajets et les amplificateurs de puissance non-linéaires. Au cours des dix dernières années, l'évolution à long terme (LTE) et l'évolution à long terme avancée (LTE-A) ont été largement utilisées dans le système de communication sans fil de quatrième génération (4G). Dans la norme LTE, le multiplexage à entrées multiples et sorties multiples (MIMO) et le multiplexage par répartition orthogonale de la fréquence (OFDM) sont considérés comme les technologies de base, qui réalisent la transmission de données avec un débit maximal de plus de 100 Mbps. Dans la ver-

sion améliorée du LTE-A, des technologies clés telles que l'agrégation multi-porteuses, la transmission coordonnée à plusieurs points, le relais et le MIMO amélioré sont rajoutées [4] [5].

Afin de répondre aux besoins du développement social futur, le système de communication sans fil de cinquième génération (5G) a vu le jour. Les différentes exigences de service peuvent être classées en trois scénarios principaux : très haut débit avec mobilité des terminaux (eMBB), les communications massives de type machine (mMTC) et les communications ultra-fiables et à faible latence (URLLC) [6]. En particulier, mMTC et URLLC nécessitent des objectifs de capacité critiques tels qu'une densité de connexion de 10^6 appareils $/km^2$, une efficacité énergétique ultra élevée, des terminaux à faible coût, une latence de 1 ms et une mobilité jusqu'à 500 km/h [7], présentant de sérieux défis sur la le déploiement commercial de la 5G. Afin d'atteindre tous les aspects des indicateurs techniques 5G, le système utilise la technique MIMO massif, des communications avec des ondes millimétriques, un réseau hétérogène et d'autres technologies, visant à améliorer l'efficacité du spectre, à augmenter la densité des points d'accès sans fil et, en outre, à étendre les ressources du spectre pour répondre à un débit de transmission plus de mille fois supérieur à celui de la 4G [8]. Le système MIMO massif dans le domaine fréquentiel est illustré par la FIGURE 1. Plus précisément, le MIMO massif déploie un grand nombre d'antennes au niveau de la station de base (BS) pour fournir des services de communication à un nombre relativement restraint d'équipements d'utilisateur (UE) sur la même ressource spectrale. Par rapport au MIMO traditionnel, le MIMO massif apporte non seulement un degré de liberté spatiale plus élevé, mais offre également de nouvelles fonctionnalités et avantages.

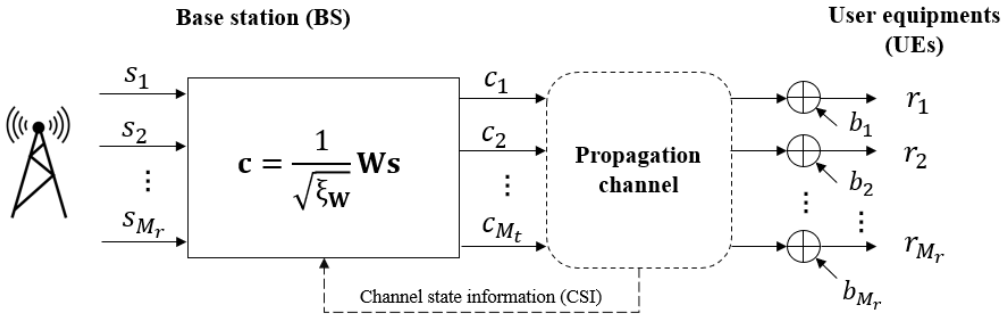


FIGURE 1 – Description du système de liaison descendante MIMO massif dans le domaine fréquentiel avec des antennes d'émission M_t côté BS et M_r UE à antenne unique

Dans le système MIMO massif, le procédé de précodage permet de transmettre simultanément des

données d'une BS avec M_t antennes vers M_r UEs. Du point de vue d'un UE spécifique, l'idée est d'utiliser les propriétés du canal de propagation pour combiner de manière cohérente la partie utile du signal provenant des M_t antennes d'émission tout en combinant de manière destructive la partie interférence reçue du signal.

Les données transmises sont définies par $\mathbf{s}$ de taille M_r et $\mathbf{c}$ de taille M_t. $\mathbf{s}$ est le vecteur de symboles complexes transmis aux différents UEs. Ce vecteur est précodé par la matrice $\mathbf{W}$ dont la taille est $M_t \times M_r$.

ξ_W est un facteur de normalisation conçu pour obtenir une puissance d'émission moyenne ou instantanée égale à σ_s^2. Le canal entre l'antenne d'émission M_t et l'UE M_r peut être défini par un scalaire complexe noté ici $\mathbf{H}$, où $\mathbf{H}(m_r, m_t)$ est le (m_r, m_t)-ième entrée de la matrice de canaux $\mathbf{H}$ de taille $M_r \times M_t$.

Par conséquent, le vecteur de données reçu $\mathbf{r}$ est défini comme suit

$$\mathbf{r} = \frac{1}{\sqrt{\xi_W}}\mathbf{HWc} + \mathbf{b} \tag{1}$$

avec $\mathbf{b}$ le vecteur de bruit.

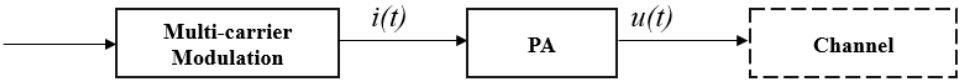


FIGURE 2 – Illustration d'une structure d'émetteur simple avec PA

Dans la chaîne de transmission réelle, le PA non linéaire est largement utilisé. La FIGURE 2 illustre une structure d'émetteur simplifiée avec l'existence d'un PA radiofréquence (RF). Alors, L'enveloppe complexe $i(t)$ du signal à l'entrée du PA peut s'écrire

$$i(t) = \rho(t)e^{j\varphi(t)} \tag{2}$$

, $\rho(t)$ est le module du signal d'entrée et $\varphi(t)$ est la phase du signal d'entrée.

En pratique, le PA est considéré comme un dispositif non linéaire. Le comportement non linéaire

dépendant de la puissance peut être caractérisé par des caractéristiques de conversion d'amplitude en modulation d'amplitude (AM/AM) et d'amplitude en modulation de phase (AM/PM). Sans effet mémoire, les caractéristiques AM/AM et AM/PM sont représentées sous la forme d'une fonction liant le module du signal de sortie en fonction du module du signal d'entrée (AM/AM) et le déphasage du signal de sortie en fonction du module du signal d'entrée (AM/PM). Lorsque le PA est entraîné vers la région non linéaire, la puissance de sortie réelle est inférieure à la sortie amplifiée linéairement. Au fur et à mesure que la puissance du signal d'entrée augmente, la quantité de distorsion non linéaire introduite par le PA augmente progressivement jusqu'àu point de saturation.

Ce phénomène non linéaire entraînera une dégradation des performances dans la bande ainsi que des remontées spectrales hors bande, ce qui entraînera des interférences dans les canaux adjacents. Pour les effets dans la bande, ces distorsions sont quantifiées par le taux d'erreur de symbole (SER) du signal reçu.

Chapter 2 : Interférences inter-numérologies

Comme introduit précédemment, dans les systèmes de communication 5G et au-delà, différentes utilisations et applications nécessitent différents types de services tout en utilisant la même structure de réseau sans fil. Afin de mieux s'adapter à cette diversité, de nombreuses nouvelles méthodes de forme d'onde ont été développées pour répondre à la demande de divers services, cependant, elles n'ont pas suffisamment de flexibilité pour répondre à toutes les exigences du système 5G. Par conséquent, dans ces cas, l'approche classique "taille unique" peut ne pas être possible [9]. Certaines nouvelles recherches tentent de trouver des solutions plus flexibles, telles que les "formes d'onde hybrides" ou les "formes d'onde mixtes" [10]. Actuellement, les systèmes 5G utilisent des «numérologies mixtes» avec une seule forme d'onde au lieu de « formes d'onde mixtes » [11]. La forme d'onde normalisée par la 5G est le multiplexage par division de fréquence orthogonale avec préfixe cyclique (CP-OFDM). Afin de répondre à différents scénarios d'utilisation, différentes numérologies sont permises pour la modulation CP-OFDM. Ces différentes numérologies font référence à différents paramètres de l'OFDM tels que l'espacement des sous-porteuses (SCS), la durée temporelle des symboles et la longueur du préfixe cyclique (CP).

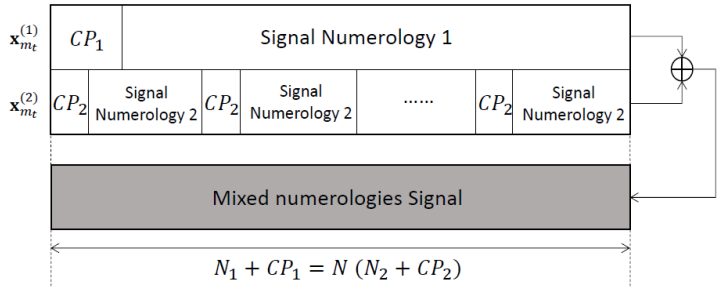
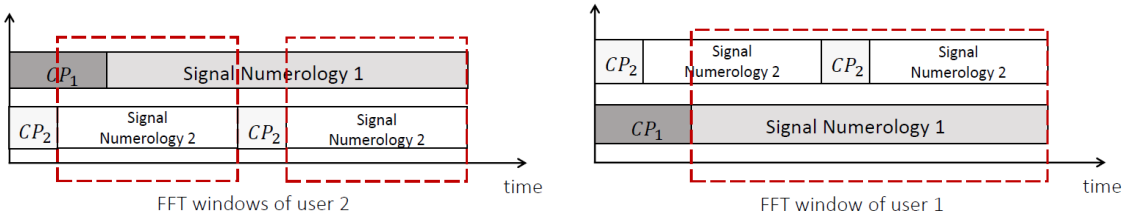


FIGURE 3 – Synchronisation et sommation des signaux. $\mathbf{x}_{m_t}^{(1)}$ est le signal de l'utilisateur 1 sur la mt-ième antenne émettrice, $\mathbf{x}_{m_t}^{(2)}$ est celui de l'utilisateur 2 signal sur la mt-ième antenne émettrice.



(a) INI de l'utilisateur 1 à l'utilisateur 2

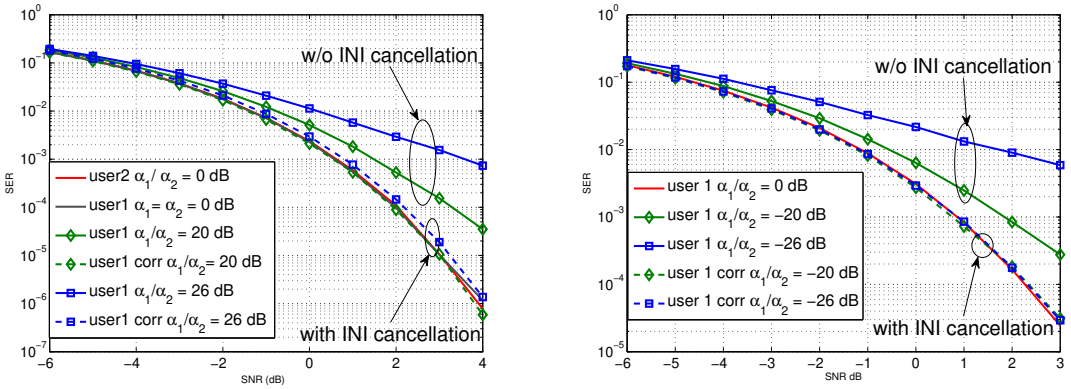
(b) INI de l'utilisateur 2 à l'utilisateur 1

FIGURE 4 – Fenêtres FFT et INI côté récepteur

liaison descendante MIMO massif

Afin de simplifier les schémas, sans perte de généralité,, nous considérons deux utilisateurs ($M_r = 2$) utilisant deux numérologies distinctes. Pour les utilisateurs 1 et 2, $\mathbf{s}^{(1)}, \mathbf{s}^{(2)} \in \mathbb{C}^{1\times N_1}$ sont des signaux produits par M-Modulation d'amplitude en quadrature (M-QAM). Les méthodes du plus petit multiplicateur commun (LCM) [12], comme le montre la FIGURE 3, sont utilisées pour réaliser la synchronisation. On suppose que $N_1 = N \times N_2, CP_1 = N \times CP_2$, où $N = 2^i$ et i est un entier.

Tout d'abord, nous considérons l'utilisateur 2 en utilisant la numérologie 2, qui est avec une taille IFFT/FFT plus petite N_2. L'INI est calculée en utilisant le fait que les tailles des fenêtres de IFFT/FFT sont inégales, comme le montre la FIGURE 4 (a), où la fenêtre FFT de la numérologie 2 est plus petite que celle de la numérologie 1. En raison de l'utilisation du précodeur ZF, l'INI de l'utilisateur 2 vers l'utilisateur 1 est nulle. Par rapport à l'utilisateur 2, la transmission de l'utilisateur 1 n'est pas correctement protégée. L'INI est induite par une fenêtre FFT inégale du côté récepteur, comme le montre la FIGURE 4 (b). La somme des INI de chaque symbole modulé avec la numérologie 2 est l'INI totale.



(a) Performances SER avant et après annulation INI sur l'utilisateur 1 sur la liaison descendante.

(b) Performances SER avant et après annulation INI sur l'utilisateur 1 sur la liaison montante.

FIGURE 5 – Performances SER avant et après annulation INI

liaison montante MIMO massif

Le système de liaison montante est l'inverse du système de liaison descendante présenté dans la section précédente. Nous considérons un système de liaison montante massive MIMO-OFDM mono-cellulaire, où M_t utilisateurs d'une seule antenne qui utilisent différentes numérologies, transmettent des signaux à une station de base équipée de antennes, sur un canal sélectif en fréquence. Notez que dans le système de liaison montante, M_r est significativement plus grand que M_t en raison du fait que maintenant la BS est le récepteur.

Nous utiliserons également deux utilisateurs ($M_t = 2$) avec deux numérologies différentes. En considérant le scénario synchronisé généralisé, similaire au précédent, nous supposons que $N_1 = N \times N_2, CP_1 = N \times CP_2$, où $N = 2^i$ et i est un entier [12] . Du côté de la BS, le signal reçu est séparé en deux branches et le flux de données est ensuite détecté par traitement linéaire par deux détecteurs ZF. Le MUI au sein d'une même numérologie est éliminée par les matrices de détection. Dans un schéma de numérologies mixtes, cependant, L'INI existe entre chaque utilisateur lorsque différentes numérologies sont utilisées dans la même bande.

Dans le chapitre 2, nous donnons l'analyse théorique de l'INI à la fois sur des scénarios de liaison descendante et de liaison montante. Sur cette base, nous pouvons implémenter deux méthodes d'annulation d'INI côté BS pour éliminer l'INI introduite dans le système de numérologie mixte. La FIGURE 5 (a) (b) montre les résultats SER avec ou sans annulation d'INI, respectivement sur la liaison des-

cendante et la liaison montante. Ici, $alpha_1$ et $alpha_2$ indiquent le path loss des deux utilisateurs par rapport à la BS. On voit sur les figures 5 (a) et 5(b) que l'annulation de l'INI fonctionne correctement même avec des utilisateurs ayant des atténuations de path loss très différentes.

Chapter 3 : Précodage MU prenant en compte la distorsion NL pour une liaison descendante MIMO massif avec PA non-linéaires

Afin de répondre aux demandes toujours croissantes des futurs systèmes de communication sans fil, nous avons introduit, dans le chapitre précédent, la gestion flexible de la transmission de numérologies mixtes. Le schéma d'émetteur-récepteur amélioré a été proposé pour les systèmes massifs de liaison descendante et montante MIMO-OFDM. Cependant, quelle que soit la technique de numérologies mixtes utilisée dans le système de communication, les précodeurs MIMO massifs entraînent des signaux en sortie de précodage avec de forts PAPR. Cela est du à la combinaison linéaire d'un nombre élevé de symboles indépendants (théorème de la limite centrale) [13]. La non-linéarité de PA RF, qui est inévitable dans une chaîne de transmission commune, apporte la principale dégradation matérielle du système. La distorsion dans la bande et le rayonnement hors bande (OBR) se produisent avec l'utilisation de PA dans la chaîne de transmission. En conséquence, une distorsion du signal et une rotation de phase apparaissent en raison de la distorsion dans la bande et des interférences dans les canaux adjacents apparaissent avec l'augmentation de l'OBR. L'effet du PA non linéaire sur la conception écoénergétique du MIMO massif est étudié dans [14] et la caractéristique spatiale de la distorsion non linéaire rayonnée par les réseaux d'antennes est calculée dans [15].

Dans le chapitre 3, nous nous sommes concentrés sur le développement d'un schéma de transmission efficace en liaison descendante prenant en considération le précodage MU, la réduction du PAPR et la linéarisation du PA avec une attention particulière accordée à la complexité de calcul. Dans ce chapitre, un problème d'optimisation de précodage MIMO massif considérant les non-linéarités du PA est analysé et formulé en un problème d'optimisation convexe simple, qui est ensuite résolu via l'approche de descente du Gradient (GD). L'algorithme itératif proposé, appelé MU-2P, est capable de pré-compenser la distorsion non linéaire (NLD) causée par les amplificateurs de puissance de la BS. Il exploite les degrés de liberté de grande dimension (DoF), fournis par le grand nombre d'antennes du côté BS. De plus, il conçoit un signal précodé qui, après amplification par le PA et filtrage par le

canal multitrajets, est très similaire à l'information transmise.

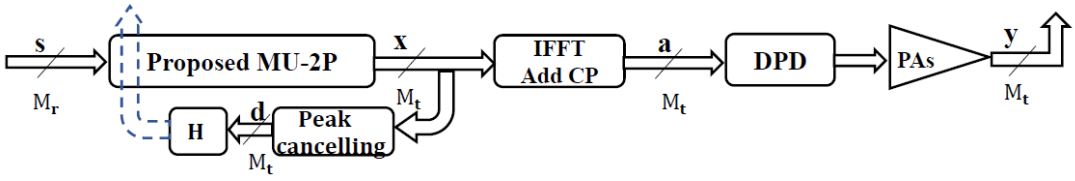


FIGURE 6 – Le schéma proposé, l'algorithme MU-2P, qui combine le précodage MU et la réduction PAPR avec une faible complexité.

De manière similaire au chapitre 2, nous considérons, dans le chapitre 3, un scénario de liaison descendante massive MIMO-OFDM, où le nombre d'antennes de la BS M_t est significativement plus grand que le nombre d'utilisateurs M_r. Le nouveau schéma est représenté sur la FIGURE 6. L'idée clé est de réduire de manière itérative conjointement la MUI et le bruit d'écrêtage $\{\mathbf{d}_{m_t}^t\}$, qui est obtenu en écrêtant les signaux dans le domaine temporel $\{\mathbf{a}_{m_t}^t\}$. Étant donné le seuil d'écrêtage λ, le signal écrêté $\bar{\mathbf{a}}_n$ à la n-ième sous-porteuse peut être obtenu par

$$\bar{\mathbf{a}}_{m_t}(k) = \begin{cases} \mathbf{a}_{m_t}(k), \text{ si } |\mathbf{a}_{m_t}(k)| < \lambda \\ \lambda e^{j\phi(k)}, \text{ si } |\mathbf{a}_{m_t}(k)| > \lambda \end{cases} \tag{3}$$

$\mathbf{a}_{m_t}(k) = |\mathbf{a}_{m_t}(k)|\, e^{j\phi(k)}$ où $\phi(k)$ est la phase de $a_{m_t}(k)$. Afin d'obtenir le meilleur PAPR, le seuil d'écrêtage optimal λ est étroitement lié à la puissance moyenne du signal OFDM σ_a^2 et au rapport entre le nombre de sous-porteuses utilisées et le nombre total de sous-porteuses.

Le bruit d'écrêtage de la m_t-ième antenne d'émission, dans le domaine fréquentiel, est présenté comme $\mathbf{d}_{m_t}^t = FFT(\bar{\mathbf{a}}_{m_t}^t - \mathbf{a}_{m_t}^t)$. Ensuite, dans le domaine fréquentiel, le signal sur la n-ième sous-porteuse du signal d'émission $\tilde{\mathbf{x}}_n$ est représenté par

$$\tilde{\mathbf{x}}_n = \mathbf{x}_n + \mathbf{d}_n, \quad n \in \chi \tag{4}$$

Par conséquent, l'énergie totale de la MUI est déterminée par $\|\mathbf{H}_n\tilde{\mathbf{x}}_n - \mathbf{s}_n\|_2^2$. En minimisant le MSE du symbole attendu $\mathbf{s}_n$ à travers le canal $\mathbf{H}$, le problème peut alors être formulé comme suit

$$\begin{aligned} &\text{minimiser } J(\mathbf{x}_n) = \|\mathbf{H}_n\mathbf{x}_n + \mathbf{H}_n\mathbf{d}_n - \mathbf{s}_n\|_2^2, n \in \chi \\ &\text{sujet à } \mathbf{x}_n = \mathbf{0}_{M_t\times 1}, \quad n \in \chi^c \end{aligned} \tag{5}$$

Afin de résoudre (5), une stratégie de minimisation est utilisée, dans laquelle la fonction objectif est minimisée de manière itérative. Il faut noter que le bruit d'écrêtage utilisé dans la $(l+1)$-ième

itération est calculé à partir de la l-ième itération précédente. Il peut être exprimé comme

$$\mathbf{x}_n^{(l+1)} = \underset{\{\mathbf{x}_n\}}{\operatorname{argmin}} J\left(\mathbf{x}_n, \mathbf{d}_n^{(l)}\right) \tag{6}$$

Ici, nous utilisons la méthode de descente de gradient (GD) la plus raide pour rechercher la direction à chaque itération, où l'itération $l+1$ est déterminée par le gradient négatif de la fonction de coût J à l'itération l, qui est donné par

$$\nabla_x J\left(\mathbf{x}_n^{(l)}, \mathbf{d}_n^{(l)}\right) = 2\mu \mathbf{H}_n^H \left(\mathbf{H}_n \mathbf{x}_n^{(l)} + \mathbf{H}_n \mathbf{d}_n^{(l)} - \mathbf{s}_n\right) \tag{7}$$

,où μ est le taux de mise à jour.

Chapter 4 : Liaison descendante massif MU-MIMO basée sur un apprentissage de bout en bout via un autoprécodeur profond

Dans le chapitre 3, nous avons introduit quelques solutions basées sur des algorithmes classiques basés sur la GD pour compenser la distorsion introduite par les PA et améliorer les performances du système [1] [2]. De plus, nous avons proposé le nouvel algorithme MU-2P, qui effectue conjointement le précodage linéaire et la réduction PAPR avec une complexité de calcul moindre par rapport aux algorithmes existants et tout en maintenant des performances satisfaisantes du système.

Les algorithmes d'apprentissage automatique (Machine Learning ML) et, plus récemment, d'apprentissage en profondeur (Deep Learning DL) se sont avérés efficaces pour effectuer les tâches d'optimisation en temps réel requises pour le MU-MIMO massif. Sohrabi et al. [16] ont proposé un auto-encodeur pour améliorer les performances des systèmes MU-MIMO massifs avec convertisseurs numériques à un bit : les performances de cette stratégie ont été démontrées, mais la complexité de calcul reste un défi en raison de la grande taille des réseaux de neurones déployés. Il convient de noter que le concept d'un auto-encodeur basé sur un réseau de neurones profonds a fait l'objet de recherches approfondies pour les systèmes à entrée unique et sortie unique (SISO) et qu'il a été démontré qu'il peut corriger les imperfections matérielles.

Dans le chapitre 4, nous étudions l'utilisation de DL dans des systèmes de liaison descendante MU-MIMO massifs avec différents canaux d'évanouissement pour traiter les non-linéarités PA. Le concept

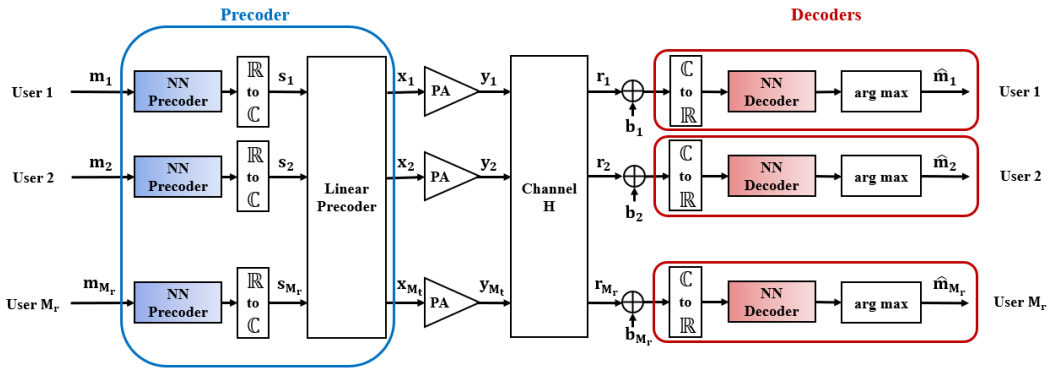


FIGURE 7 – Architecture proposée pour le système auto-précodeur avec correction des imperfections des PA.

important est de modéliser le précodeur d'émission, les PA, le canal et le récepteur comme un seul NN profond, c'est-à-dire comme une tâche de reconstruction de bout en bout avec seulement le précodeur d'émission et le récepteur pouvant être entraînés.

Le système de liaison descendante massive MIMO-OFDM avec autoprécodeur est illustré sur la FIGURE 7. Nous proposons un schéma de précodage en deux étapes : (1) un précodeur réseau neuronal(NN) non linéaire entraîné hors ligne qui s'adapte à n'importe quel canal, et (2) un précodeur linéaire dépendant du canal de transmission qui peut être construit en utilisant les schémas de précodage ZF largement utilisés. Étant donné que la plupart des bibliothèques DL existantes n'offrent que des opérations à valeur réelle, nous devons d'abord transformer les données complexes en données réelles. Puisque le signal transmis est complexe, dans le schéma fonctionnel, il y a un bloc $\mathbb{R}$-to-$\mathbb{C}$ et un $\mathbb{C}$-to-$\mathbb{R}$ bloc, représentant respectivement les mappages réel-complexe et complexe-réel. $\mathbf{m}$ indique la représentation vectorielle unique du symbole émis, c'est-à-dire un vecteur M-dimensionnel avec un élément non nul pour le m-ième élément et zéro pour les éléments restants. Ce vecteur est passé dans le NN non linéaire de transmission, qui comporte de nombreuses couches non linéaires denses et une couche linéaire de sortie. Une couche dense est une multiplication matricielle avec des poids pouvant être appris suivie d'une fonction d'activation non linéaire. Deux neurones linéaires représentant les sorties réelles I et Q du signal complexe en bande de base appartenant à un utilisateur doivent être présents dans la couche linéaire de sortie..

Pendant la phase d'apprentissage, un seul précodeur NN est entraîné, et il sera utilisé pour les M_r utilisateurs au moment du test. Nous générons le signal à valeur réelle $2M_r$ correspondant aux M_r

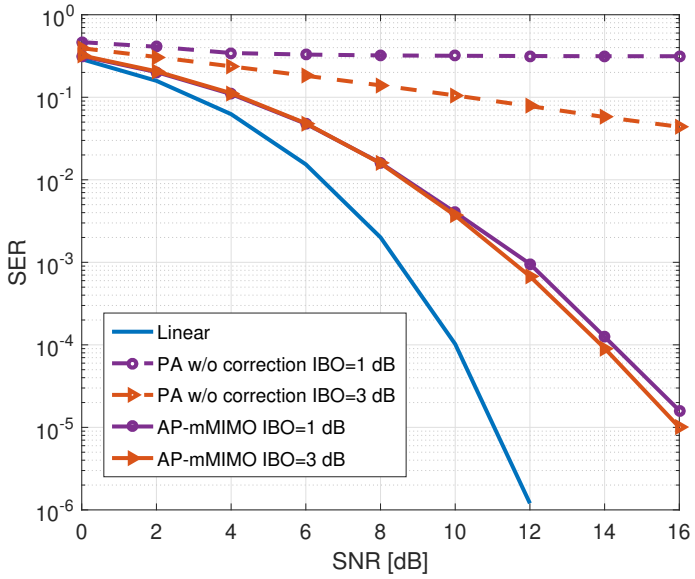


FIGURE 8 – SER en fonction du SNR avec un PA opéré à des IBO de 1 dB et 3 dB

utilisateurs pendant la phase de généralisation, qui sert d'étape initiale de précodage. Un précodeur linéaire traditionnel, tel qu'un ZF, est utilisé dans la deuxième étape. Il convient de noter que le signal d'émission à chaque antenne est ajusté au back-off d'entrée (IBO) requis à la sortie du précodeur linéaire (pour des raisons de simplicité, il n'est pas présenté sur la FIGURE 7).

Du côté du récepteur, un autre NN est utilisé comme décodeur, similaire au précodeur NN émetteur. Toutes les opérations de décodage des utilisateurs dans notre système utilisent le même décodeur NN. Un seul décodeur NN est entraîné pendant la phase d'apprentissage hors ligne, similaire au précodeur, et tous les utilisateurs utiliseront le même décodeur pendant la phase de test en ligne.

La FIGURE 8 illustre le SER en fonction du rapport signal sur bruit (SNR) du schéma proposé lorsque les PA fonctionnent avec des IBO = 1 dB et 3 dB. Comme performances de référence, nous traçons le scénario où il n'y a pas de correction de distorsion PA ("PA sans correction", pire cas) et le cas linéaire où le PA est considéré comme idéal ("Linéaire", cas idéal optimal).

D'après la FIGURE 8, nous pouvons clairement noter que le schéma que nous avons proposé peut apporter une amélioration significative par rapport au cas correction. De plus avec des IBO de 1 et 3dB, les performances de notre système sont relativement proches de celles du système idéal sans PA (4dB de perte à SER= 10^{-4})

General Introduction of the thesis

Context of the PhD study

This work of this PhD study was mainly done in the CEDRIC laboratory of CNAM under supervision of Prof. Daniel ROVIRAS (PhD director, CNAM), Dr. Hmaied SHAIEK (CNAM) and Dr. Rafik ZAYANI (Innov'COM/Sup'Com) and it is funded by Sorbonne Université.

The main objectives of my thesis, "Study and mitigation techniques of RF impairments for beyond 5G multi-carrier waveforms", is the study of massive MIMO techniques with the presence of radio frequency (RF) imperfections, in particular, the non-linear (NL) power amplifiers (PA). On the one hand, it focuses on understanding and developing new massive MIMO systems for beyond 5G communications at the pysical layer. On the other hand, new techniques are established in order to mitigate the PA impairments and increase the system energy efficiency. Besides, the influence of the different numerologies used by the different users will also be studied in this thesis.

Important results and novelties of the thesis

- *Closed form expressions for the inter-numerology interference (INI) in uplink scenario and downlink scenario with massive MIMO-OFDM and different numerologies.*
  To allow flexible management of mixed numerologies transmission, improved transceiver schemes are developed for the massive MIMO-OFDM downlink and uplink systems. For massive MIMO-OFDM systems with mixed numerologies, analytical INI models are created, which could be useful in guiding 5G system design and parameter selection. We derived theoretical INI expressions in closed-form to investigate the effect of INI on users' results theoretically. Theoretical findings match simulation results, proving the validity of our theoretical research.

- *Proposal of a INI cancellation architecture for both uplink and downlink massive MIMO-OFDM scenarios.*

  INI cancellation methods are proposed based on the derived theoretical expressions of INI between different numerologies for massive MIMO-OFDM downlink and uplink. These methods can suppress the INI at the BS side without raising the users' complexity. These findings demonstrated that massive MIMO-OFDM systems can accept mixed numerology transmissions while all users share the same band, indicating that this method can increase the spectrum efficiency as well as meet the needs of future wireless communication systems.

- *Proposal of a low complexity NL distortion-aware multi-user (MU) precoding technique for massive MIMO downlink under PA nonlinearities.*

  We concentrate on the development of an efficient downlink transmission scheme taking into consideration MU precoding, PAPR reduction and PA linearization with a special attention given to the computational complexity. A massive MIMO precoding optimization problem considering PA nonlinearities is analysed and formulated into a simple convex optimization problem, which is then solved via gradient descent (GD) approach. The proposed iterative algorithm is able to pre-compensate nonlinear distortion (NLD) caused by BS power amplifiers. It exploits the high-dimensional degrees of freedom (DoFs), provided by the large number of antennas equipped at the BS side.

- *Proposition of a new neural network (NN) architecture called AP-mMIMO scheme for mitigating the effects of NL PA in massive MU-MIMO downlink transmission.*

  We investigate the use of deep learning (DL) in massive MU-MIMO downlink systems with varying fading channels to deal with PA non-linearities. The important concept is to model the transmit precoder, PAs, channel, and receiver as one deep NN, i.e. as an end-to-end reconstruction task where only the transmit precoder and receive decoders are trainable. The new AP-mMIMO is suitable for varying channel scenarios, requiring less computational complexity for adaptability.

Content of the thesis

The organization of the thesis is presented as follows.

We provide an overview of the evolution of wireless communication systems, as well as their key features, in Chapter 1.

Then, we concentrate on increasing spectral efficiency in massive MIMO-OFDM systems in Chapter 2. The mixed-numerologies technique is used to explore at the usage of spatial multiplexing among users who share the same bandwidth. The INI models were developed to analyze the interference introduced between various users and the downlink and uplink transmission schemes are proposed. The findings indicate that with the proposed cancellation technique, the INI could be reduced with better spectral efficiency. We demonstrat that INI is only generated in frequency selective channels and only when users with larger subcarrier spacing (SCS) interact with users with smaller SCS in the massive MIMO-OFDM downlink scenario. In the uplink situation, however, all users in frequency selective channels suffer from INI.

We switched our attention to energy efficiency in the actual transmission chain in Chapter 3. RF PAs were greatly emphasized in this chapter. The nonlinear features of PA have a negative impact on transmission quality because they cause hardware distortion. We examine various current methods and then propose our novel combined MU precoding and PAPR reduction approach to compensate for this weakness while preserving system efficiency. The method is proposed in a massive MU-MIMO downlink system. The presented combined MU precoding and PAPR reduction approach, which is framed as a simple convex optimization problem, is built using a steepest GD-based methodology. The simulation results show that, as compared to prior approaches, the proposed transmitting scheme offers acceptable performance while needing less computational complexity.

In Chapter 4, we use the deep NN technique to compensate the PA nonlinearities. We attempt to develop an end-to-end solution, as opposed to the traditional optimization approach, which divides the transmission chain into multiple blocks. Two NNs are trained off-line and utilized for all antenna branches and users in our proposed AP-mMIMO system. In comparison to earlier literature, numerical findings clearly demonstrate that the proposed AP-mMIMO scheme offers comparable performance while requiring even less computing complexity than traditional algorithms-based solutions.

Finally, Chapter 5 gives the conclusion and perspectives.

Publications

- **Published papers**

  1. **X. Cheng**, R. Zayani, H. Shaiek and D. Roviras, "Inter-Numerology Interference Analysis and Cancellation for Massive MIMO-OFDM Downlink Systems," in IEEE Access, vol. 7, pp. 177164-177176, 2019, doi : 10.1109/ACCESS.2019.2957194.

  2. **X. Cheng**, R. Zayani, H. Shaiek and D. Roviras, "Analysis and Cancellation of Mixed-Numerologies Interference for Massive MIMO-OFDM UL," in IEEE Wireless Communications Letters, vol. 9, no. 4, pp. 470-474, April 2020, doi : 10.1109/LWC.2019.2959526.

- **Submitted papers**

  1. **X. Cheng**, R. Zayani, H. Shaiek and D. Roviras, "Low Complexity Downlink Transmission scheme for massive MU-MIMO-OFDM under PA nonlinearities," submitted to IEEE Wireless Communications Letters.

  2. **X. Cheng**, R. Zayani, M. Ferecatu and N. Audebert, "Efficient Autoprecoder-based deep learning for massive-MIMO Downlink under PA Non-Linearities," submitted to IEEE Wireless Communications Letters.

Chapitre 1

Technical background

Contents

In this chapter, we present the sate-of-the-art presenting all of the necessary aspects for the rest of the thesis studies. In the first part, I will give an overview of wireless systems, with a focus on cellular radiocommunication systems from 1G to 5G. Following that, MIMO and massive MIMO technologies will presented. The technique of mixed-numerologies will then introduced, as well as the wireless propagation channels. Finally, I will evoke the problem of the power amplifier (PA), presenting the major hardware imperfection in massive MIMO based wireless systems.

1.1 Wireless communication systems : 5G and beyond

Since the 1980s, wireless communication systems have undergone renewal and evolution at a rate of approximately every 10 years. In the last forty years, the wireless communication industry has developed rapidly and has been deeply integrated into human daily life.

The first generation (1G) of wireless communication system was using Frequency Division Multiple Access (FDMA) technology to transmits analog modulated voice communication signals. The second generation (2G) was transformed from analog communication to digital communication, with Time Division Multiple Access (TDMA) and FDMA technology to provide low-rate data services with a peak rate of up to 64 kbps. Third Generation (3G) system has adopted Code Division Multiple Access (CDMA) which can reach the peak rate calculated in Mbps level. Due to the significant improvement of system data transmission capacity, 3G started to support mobile multimedia services such as broadband Internet access, video calls and online games. In the past ten years, long-term evolution (LTE) and Long-Term Evolution Advanced (LTE-A) have been widely used in the fourth generation (4G) wireless communication system. Inside LTE, multiple-input multiple-output (MIMO) and Orthogonal Frequency Division Multiplexing (OFDM) are considered the core technologies, which realize data transmission with a peak rate of more than 100 Mbps. In the upgraded version of LTE-A, key technologies such as multi-carrier aggregation, Coordinated Multiple-Point transmission, relay, and enhanced MIMO are included [4] [5]. The evolution of wireless communication systems, industrial communication and social development complement each other. On the one hand, emerging demands in the industry as well as in the society are stimulating the progress of wireless communication systems ; on the other hand, the improvement of data transmission capabilities in wireless communication systems provides more possibilities for social and industrial communication development. In recent years, the popularization of mobile networks and smart terminals has made people's lifestyles restructured around

various new applications. New applications such as mobile medical care, autonomous driving, smart homes, and virtual reality have been continuously proposed. Then, as a result, low latency, high speed and massive connections have become the new requirements.

In order to meet the needs of future social and industry development, the Fifth Generation (5G) wireless communication system came into our sight. The different service requirements can be categorized into three main scenarios : enhanced mobile broadband (eMBB), massive machine type communications (mMTC), and ultra-reliable and low latency communications (URLLC) [6]. In particular, mMTC and URLLC require critical capability objectives such as 10^6 devices/km^2 connection density, ultra high energy efficiency, low cost terminals, 1 ms latency and mobility up to 500 kmph [7], presenting serious challenges on 5G commercial deployments. In order to achieve all aspects of 5G and beyond technical indicators, the system uses massive MIMO, millimeter wave communications, heterogeneous network and other technologies, aiming to improve spectrum efficiency, increase the density of wireless access points and furthermore, expand spectrum resources to meet a transmission rate more than a thousand times higher than that of 4G [8]. Unlike previous systems that distinguish different users or data streams mainly in the frequency domain, time domain, and code domain, one of the key points of 5G system research is how to focus more on the space domain using Space Division Multiple Access (SDMA) technology or on the power domain using Non -Orthogonal Multiple Access (NOMA) technology. Therefore, massive MIMO technology that can dig a large amount of space resources is one of the mainstream and research hot-spots of 5G and beyond research.

1.2 Massive MIMO

In modern communication systems, MIMO is a technology that relies on multiple antennas at the transmitter and receiver to either transmit data simultaneously on parallel streams or to increase the link reliability. MIMO can usually be traced back to research papers on multi-channel digital transmission systems in the 1970s [17][18][19]. The mathematical techniques for dealing with mutual interference in these papers promoted the research and development of MIMO. By the mid-1980s, Bell Laboratories further developed the research of MIMO, studying time division multiplexing in multi-user systems [20]. In the 1990s, a MIMO prototype system was created and field tested [21]. Bell Laboratories established a laboratory prototype in 1998, demonstrating its vertical Bell Laboratories Layered Space-Time (V-BLAST) technology [22]. In wireless communication systems, MIMO techno-

logy can fully use the space resources [23] and improve system channel capacity and other performance without requiring additional transmission power and spectrum bandwidth. In particular, for a point-to-point MIMO system, when the channels are not correlated with each other, the system capacity has a linear growth relationship with the minimum of the transmitting antenna and the receiving antenna [24]. Multi-user MIMO distinguishes users by spatial location and shares multiplexing gains with all users, which can increase the number of users served while increasing the system capacity [25]. After decades of development, MIMO technology has become a basic element of modern wireless communication standards. Related standards include but are not limited to WLAN/WiFi [26], WiMAX and LTE-A [27]. In the previous communication systems, the number of antennas was still quite small. For example, the LTE-A standard supports up to 8 antennas. In the latest 5G standard, base stations can already support 64 antenna elements [28], but there is still a certain gap between the concept of large-scale (tens of thousands). Therefore, the performance gain brought by MIMO technology with small number of antennas is unable to meet the demand for high-speed data transmission in the future. In order to further increase the transmission rate of the system and break the limit of the existing cellular system, literature [29] proposed an evolved version of MIMO technology in the spatial dimension, namely massive MIMO. Specifically, massive MIMO deploys a large number of antennas at the base station (BS) to provide communication services for a relatively small number of user's equipment (UE) on the same spectrum resource. Compared with traditional MIMO, massive MIMO not only brings a higher degree of spatial freedom, but also provides new features and advantages.

1.2.1 Precoding

In a massive MIMO system, the precoding method allows to provide SDMA and simultaneously transmit data from a BS with M_t antennas to M_r UEs. In the rest of the thesis, we always consider BS with M_t antennas and M_r UEs with only one antenna. From the perspective of a specific UE, the idea is to use the propagation channel properties to coherently add the useful part of the signal from the M_t transmit antennas while destructively adding the received interference part of the signal. Both nonlinear and linear precoding techniques can be used in typical MIMO systems. Compared with linear precoding methods, nonlinear methods such as dirty paper coding (DPC) [30], vector perturbation (VP) [31] and lattice assisted methods [32] increase the complexity of implementation, but the performance is also improved. However, as the number of BS antennas increases, linear precoding

has proven to be almost optimal [29], [33]. Therefore, it is more practical to use less complex linear precoding techniques in massive MIMO systems. The frequency domain system model is shown by FIGURE 1.1

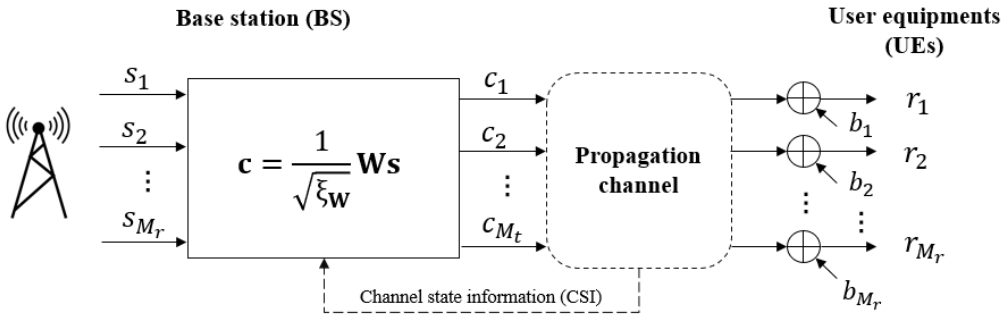


FIGURE 1.1 – Description of the frequency domain massive MIMO downlink system model with M_t transmit antennas on the BS side and M_r single-antenna UE

The transmitted data are defined by the following vector :

$$\mathbf{s} = \begin{bmatrix} s_1 & s_2 & \cdots & s_{M_r} \end{bmatrix}^T \tag{1.1}$$

where s_{m_r} is the complex data transmitted to the UE m_r with a variance of σ_s^2. This vector $\mathbf{s}$ is precoded by the following $(M_t \times M_r)$ matrix :

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_1 & \mathbf{W}_2 & \cdots & \mathbf{W}_{M_r} \end{bmatrix} \tag{1.2}$$

The transmitted precoded data are defined by the following vector :

$$\mathbf{c} = \begin{bmatrix} c_1 & c_2 & \cdots & c_{M_r} \end{bmatrix}^T = \frac{1}{\sqrt{\xi_W}}\mathbf{W}\mathbf{c}, \tag{1.3}$$

where ξ_W is a normalization factor designed to obtain an average or instantaneous transmit power equal to σ_s^2.

The channel in frequency domain between the transmit antenna m_t and the UE m_r can be defined by a complex scalar denoted here by $\mathbf{H}$, where $\mathbf{H}(m_r, m_t)$ is the (m_r, m_t)-th entry of the channel matrix $\mathbf{H}$ of size $M_r \times M_t$.

Therefore, the received data vector $\mathbf{y}$ is defined as follows :

$$\mathbf{y} = \begin{bmatrix} y_1 & y_2 & \cdots & y_{M_r} \end{bmatrix}^T = \frac{1}{\sqrt{\xi_W}}\mathbf{HWc} + \mathbf{b}, \tag{1.4}$$

with $\mathbf{b}$ the noise vector :

$$\mathbf{b} = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_{M_r} \end{bmatrix}^T \tag{1.5}$$

whose entries are i.i.d circularly-symmetric complex Gaussian distribution with zero-mean and variance σ_b^2.

There are several precoding technqiues, which are listed below.

Maximization of the Received Power (MRT)

The first strategy is to maximize the received power of each UE without considering its multi-user interference (MUI). This is done using the matched filter of the channel, implemented by :

$$\mathbf{W} = \mathbf{H}^H \tag{1.6}$$

If ξ_w is defined to have the instantaneous power equal to σ_s then :

$$\xi_{\mathbf{W}} = \sum_{m_t=1}^{M_r} \sum_{m_r=1}^{M_r} |H(m_t, m_r)|^2 \tag{1.7}$$

In the literature, this precoding method is called Maximum Ratio Transmission (MRT) and was first introduced in [34] for single-user MIMO transmission. Therefore, if $M_r = 1$, this solution is the best because there is no MUI. In addition, when the number of transmitting antennas tends to infinity, MUI disappears[29], this solution is still the best. $\xi_{\mathbf{W}}$ can also be set to have an average transmit power equal to σ_s, which is [29] :

$$\xi_{\mathbf{W}} = M_t \times M_r \tag{1.8}$$

Minimization of the MUI (ZF) The second strategy is to minimize MUI. This can be done by the well-known zero-forcing (ZF) recorder, also known as channel reversal [35], where :

$$\mathbf{W} = \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H\right)^{-1} \tag{1.9}$$

In [35], the author pointed out that the average normalization factor is difficult to use because it can be infinite, for example when $M_r = M_t$. This is why the instantaneous normalization factor is a better choice. In this case, it can be obtained by :

$$\xi_{\mathbf{W}} = \sum_{m_t=1}^{M_t} \sum_{m_r=1}^{M_r} |\mathbf{W}(m_t, m_r)|^2 \tag{1.10}$$

For ZF combining technology, this precoder completely removes MUI. Under low noise power, the performance of the ZF precording is better than that of the MRT precoding, due to the fact that the performance of MRT is limited by MUI [36]. However, under high noise power conditions, the performance of MRT precoding is better than that of ZF precoding. In fact, for a ZF precoding, the normalization coefficient $\xi_{\mathbf{W}}$ may be very high, and the received power will then decrease. In addition, ZF is more complicated than MRT technology because it needs to perform matrix inversion. The ZF precoding will be used in Chapter 2 to Chapter 4.

Trade-off Between Received Power and MUI (MMSE)

One way to improve the performance of the ZF precoding under high noise power is to "regularize" the inverse matrix $\left(\mathbf{H}\mathbf{H}^H\right)^{-1}$ [35]. The precoding matrix is as follows :

$$\mathbf{W} = \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H + \alpha \mathbf{I}_{M_r}\right)^{-1}, \alpha \in \mathbb{R}^+ \tag{1.11}$$

This precoding method is called transmit minimum mean square error (MMSE) or regularized zero forcing (RZF). In fact, if the value of α is high, $\mathbf{W}$ tends to become an MRT precoder, but if the value of α is low, $\mathbf{W}$ tends to become a ZF precoder.

1.2.2 Null space

In mathematics, the kernel of a linear map, also known as the null space or nullspace, is the linear subspace of the domain of the map which is mapped to the zero vector [37]. Null Space and Nullity are concepts in linear algebra which are used to identify the linear relationship among attributes. The null space of any matrix $\mathbf{A}$ is the set of matrix $\mathbf{B}$ such that $\mathbf{AB} = \mathbf{0}$ and $\mathbf{B}$ is not zero. It can also be thought as the solution obtained from $\mathbf{AB} = \mathbf{0}$ where $\mathbf{A}$ is the known channel matrix $\mathbf{H}$ of size $M_r \times M_t$. The size of the null space of the matrix provides us with the number of linear relations among attributes. There exists several ways to get the null space of the channel matrix $\mathbf{H}$.

The first way is to use the singular value decomposition[38]. There exists a SVD of $\mathbf{H}$:

$$\mathbf{H} = \mathbf{U}\Sigma\mathbf{V}^H \tag{1.12}$$

where $\mathbf{U}$ is an $M_r \times M_r$ complex unitary matrix, Σ is an $M_r \times M_t$ rectangular diagonal matrix with non-negative real numbers on the diagonal, and $\mathbf{V} = \left[\mathbf{v}^1, \mathbf{v}^2, \ldots, \mathbf{v}^{M_t}\right]$ is an $M_t \times M_t$ complex unitary matrix. The diagonal entries σ_i of Σ are known as the singular values of $\mathbf{H}$. In MIMO transmission, $M_t \times M_r$-dimensional $\mathbf{V}^d = \left[\mathbf{v}^1, \mathbf{v}^2, \ldots, \mathbf{v}^{M_r}\right]$ is used as the Beam-forming matrix of the M_r data stream. Then, the rest part of the matrix $\mathbf{V}^0 = \left[\mathbf{v}^{M_r+1}, \ldots, \mathbf{v}^{M_t}\right]$ spans the null space of the channel. Therefore :

$$\mathbf{H}\mathbf{V}^0 = 0 \tag{1.13}$$

The null space of the channel can also be obtained by the following method :

$$\mathbf{V}^0 = \mathbf{I} - \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H\right)^{-1} \mathbf{H} \tag{1.14}$$

to obtain the zero result after the multiplication.

The key idea of the channel null space is that, when extra information transmitted in the null space of the propagation channel with the current user, no interference is seen at the current user location. In the literature, many researchers use the channel null space to transmit the signal of secondary user (SU) without interfering the primary user (PU). SU and PU are operating at the same frequency band, therefore the reciprocity of wireless channel $\mathbf{H}$ can be assumed. In [39], the coexistence of PU and multiple SUs based on adaptive null space is proposed for MIMO Orthogonal Frequency Division Multiple Access (OFDMA) uplink. Xiong *el al.* [40] consider a downlink scenario with a SU base station with multiple SUs and PUs. The SU BS notifies all SUs of the null space of the channel matrix between itself and the PU. However, the performance of null space application is relevant to the knowledge of the channel state information (CSI). When CSI is not fully known, there are effects in terms of capacity and bit error rate (BER). The concept of null space will be used in Chapter 3 when a signal that lies in the null space is added for all M_r users.

1.2.3 Degrees of freedom

A fundamental problem for massive MIMO is to determine the system capacity. From [41], the capacity of a point-to-point MIMO system with M input and N output increases linearly with $min(M, N)$

in the high signal-to-noise ratio (SNR) region.. For wireless systems with limited power and bandwidth, another dimension "space" is proposed that it can be used like time and frequency. Similar to time division and frequency division multiplexing, massive MIMO systems have the possibility of multiplexing signals in space. For example, using the SVD of the channel, it is possible to generate orthogonal slots in space by dividing time or frequency into multiple channels.

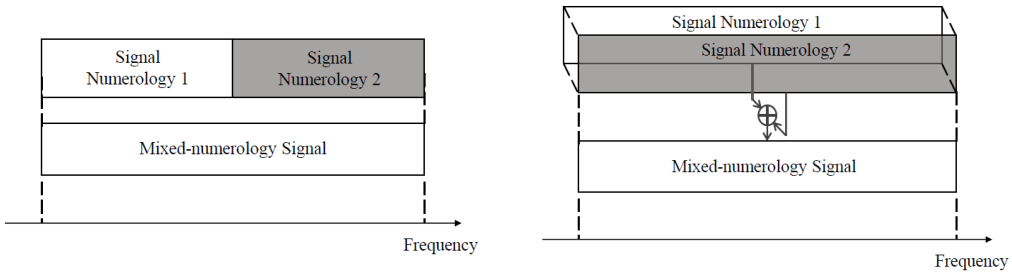
1.3 Numerologies

As introduced before, in the 5G and beyond communication systems, different uses and applications require different kinds of services while using the same wireless network structure. In order to better adapt to this diversity, many new waveform methods have been developed to meet the demand of various services, however, they do not have enough flexibility to meet all 5G system requirements. Therefore, in these cases, the classical "one size fits all" approach may not be possible [9]. Some new research attempts to find more flexible solutions, such as the "hybrid waveforms" or "mixed waveforms" [10]. Currently, 5G systems use "mixed numerologies" with a single waveform instead of "mixed waveforms" [11], which is cyclic prefix orthogonal frequency division multiplexing (CP-OFDM). This method is designed to meet system requirements based on frame design. According to [3], seven numerologies structures are defined for the new 5G radio (NR), as shown on TABLE 1.1. Different numerologies refer to different parameters setting in OFDM such as subcarrier spacing (SCS), symbol duration and cyclic prefix (CP) length. One of the first studies which introduced the multi-numerology or mixed-numerology system is [42].

TABLE 1.1 – Numerology structures in 5G [3]

Type	Spectrum efficiency	Δf (kHz)	# of symbols in one subframe	T_{cp} (μs)
Type-A1	93.3 %	15	14	4.76
Type-A2	93.3 %	30	28	2.38
Type-A3	93.3 %	60	56	1.19
Type-A4	80.0 %	60	48	4.17
Type-A5	93.3 %	120	112	0.60
Type-A6	93.3 %	240	224	0.30
Type-A7	93.3 %	480	448	0.15

In the literature, there exist mainly two ways of mixed numerologies implementation, as shown on



(a) Mixed numerologies with divided subbands (b) Mixed numerologies with sharing full bandwidth

FIGURE 1.2 – Mixed numerologies implementations

FIGURE 1.2. The most accepted way proposed to support diverse services is to divide the bandwidth into several subband, then flexible allocation of different numerologies are implemented in each subband [43] [44]. This implementation in frequency domain is shown on FIGURE 1.2 (a). Although, this usage of numerology multiplexing significantly improves the system flexibility, each user cannot access the full bandwidth. Furthermore, the interference between users belonging to different numerologies appears and the inter-numerology interference (INI) affects the system performance.

Recently, some works [45] [46] have dealt with mixed numerologies spectrum sharing (SS), which can be also described as overlapping mixed numerologies. This method, as shown on FIGURE 1.2 (b), allocates the full bandwidth to all users using different numerologies (i.e., users are sharing the same time/frequency resources). However, under this solution, it is impossible to avoid interference using windowed/filtered waveforms and the INI models are more complex. INI will be studied in Chapter 2 for uplink and downlink in a massive MIMO context.

1.4 Propagation channels

The characteristics of the propagation channel depend on the type of the communication. Regarding cellular communications, in order to send signal from the transmitter to the receiver, the electromagnetic wave eventually passes through the propagation channel. When it passes through the channel, some phenomena may occur, leading to signal distortion or fading. These phenomena can be categorized mainly in three parts : reflection, diffraction and scattering [47].

Reflection happens when the electromagnetic wave propagates over large obstacles (such as the

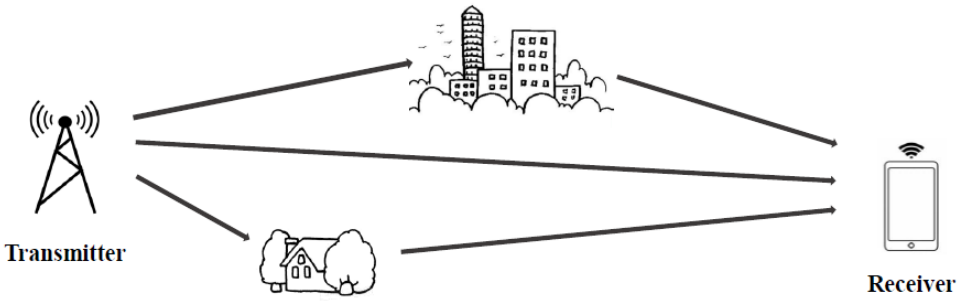


FIGURE 1.3 – Illustration of a multipath propagation channel for cellular communications

ground or buildings) whose size is very large compared to the wavelength at which it propagates. Diffraction occurs when the path between the transmitter and receiver is blocked by a surface with obvious irregularities (edges) [48]. This phenomenon allows the electromagnetic wave to travel around the curved surface and propagate behind obstacles. Therefore, even if the receiver does not have a line of sight to the transmitter, the receiver will still receive information from the transmitter. When electromagnetic waves propagate on rough surfaces with irregularities smaller than the wavelength, scattering occurs. Due to these phenomena, the received signal is a combination of signals from different paths. Each of these signals is an attenuated and delayed version of the original signal. Attenuation and delay depend on the nature of the obstacle encountered and the distance the signal travels. Such a propagation channel is considered a linear propagation channel, and is called a multi-path propagation channel. This phenomena is illustrated by FIGURE 1.3.

The various received signals are obtained from a direct path called line-of-sight (LOS) and another path called non-line-of-sight (NLOS). In practice, the LOS path does not always exist. In fact, if the obstacle is higher than the transmitter or receiver antenna, the LOS direction will be lost. However, if it exists, the signal received from the LOS path usually represents the most dominant signal [48]. In addition, these phenomena, also called fading, are either large-scale fading or small-scale fading. The large-scale fading is the mean attenuation of the signal and it relies on the distance between the transmitter and the receiver. The small-scale fading are caused by the multi-path components introduced before. The received signal is thus the sum of all paths.

The reflection, diffraction and scattering result in the reception of the transmitted signal with different amplitudes, phases and delays from the multi-path, then, the wireless multi-path channel can

be modeled as a linear time varying (TV) filter with impulse response $h(t,\tau)$ given by [48] :

$$h(t,\tau) = \sum_{k=0}^{L_h-1} a_k(t,\tau) e^{j\{2\pi f_k \tau_k(t) + \phi_k(t,\tau)\}} \delta\left(\tau - \tau_k(t)\right) \tag{1.15}$$

, where L_h stands for the different paths. a_k, f_k and τ_k represent the real amplitude, the transmitting frequency and the delay of the k-th path, respectively. The term $\{2\pi f_k \tau_k(t) + \phi_k(t,\tau)\}$ is the phase shift and $\delta(.)$ is the Dirac delta function.

1.4.1 Channel characterization

Delay spread

The delay spread is the difference in time between two pathways that arrive at somewhat different times. There are three different forms of delay spread.

Mean excess delay It reflects the weighted average delay and can be presented as :

$$\bar{\tau} = \frac{\sum_{k=0}^{L_h-1} a_k^2 \tau_k}{\sum_{k=0}^{L_h-1} a_k^2} \tag{1.16}$$

Maximum excess delay : It denotes the delay of the most recent electromagnetic wave, whose magnitude is not ignored in comparison to the first one. It is referred to as τ_{max}.

Root mean square (RMS) delay : It's written as σ_τ or τ_{RMS}, and it is delivered by

$$\sigma_\tau = \sqrt{\overline{\tau^2} - \bar{\tau}^2} \tag{1.17}$$

where

$$\overline{\tau^2} = \frac{\sum_{k=0}^{L_h-1} a_k^2 \tau_k^2}{\sum_{k=0}^{L_h-1} a_k^2} \tag{1.18}$$

Coherence bandwidth : B_c

The coherence bandwidth, presented as B_c, denotes the frequency range in which the channel is considered flat. With the following expression, it is related to the RMS delay spread given by equation (1.17) :

$$B_c \simeq \frac{1}{\alpha_\tau \sigma_\tau} \tag{1.19}$$

where $\alpha_\tau \in \mathbb{R}^{*+}$. It's worth noting that the coherence bandwidth can be used to determine if a channel is frequency selective or not. If B_c exceeds the signal bandwidth B_S, the channel is considered flat. When $B_c < B_s$, however, the channel is considered to be frequency selective.

Doppler spread B_D

The Doppler spread, commonly known as the Doppler spectrum, is defined as the frequency range across which the received signal spectrum from a single frequency signal is essentially non-zero. The Doppler spread, presented by f_d, is defined as the frequency change of an electromagnetic wave and is expressed as :

$$f_d = \frac{v}{\lambda}\cos(\theta) \tag{1.20}$$

where the velocity, the wavelength signal and the angle between the direction of mobile motion and the direction of wave arrival are represented by v, λ and θ, respectively.

The channel is described as slow fading if the doppler spread B_D is less than the signal bandwidth. If B_D is significantly higher than the signal bandwidth, however, the channel is classified as a fast fading channel.

Coherence time : T_c

The coherence time,T_c, refers to the period of time during which the channel is regarded invariant. With the following expression, it is inversely proportional to the Doppler shift :

$$T_c = \frac{1}{\alpha_f f_m} \tag{1.21}$$

where f_m denotes the maximum Doppler shift and α_f denotes a positive real coefficient that varies with the temporal correlation function.

1.4.2 Channel taps distribution

The wireless multi-path channel has been modeled by a linear TV impulse response in equation (1.15). Therefore, for the frequency selective channel, it is the sum of many delta functions, representing signals arriving at different moments with different amplitudes on each path. Depending on whether the LOS path exist or not, all signal amplitudes of all multi-path signals follow the Rayleigh distribution (no LOS) or the Rician distribution (LOS path).

Rayleigh distribution The Rayleigh fading channel model is a statistical channel model that can be used to simulate multi-path components when it is not a direct LOS path between the transmitter and the receiver. If there is no direct path between the transmitter and receiver, *i.e.* the propagation channel is not in line of sight (NLOS), the Rayleigh fading channel model can be used to simulate all multi-path components. The probability density function (PDF) of a Rayleigh random variable x is given by [49] :

$$p(x) = \begin{cases} \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{1.22}$$

where σ is the standard deviation of the real part of x.

Rician distribution When the multi-path component is a direct LOS path between the transmitter and the receiver, the Rician fading channel model is used. The PDF of the Rician distribution of a random variable x is given by [49] :

$$p(x) = \begin{cases} \frac{x}{\sigma^2} e^{-\frac{x^2+D^2}{2\sigma^2}} I_0\left(\frac{Dx}{\sigma^2}\right), & x \geq 0, D \geq 0 \\ 0, & x < 0 \end{cases} \tag{1.23}$$

where σ^2, D and $I_0(.)$ represent the variance of the real or the imaginary part, the magnitude of the LOS component and the modified first-kind Bessel function with zero-order, respectively.

1.5 PA

1.5.1 PA characteristics

AM/AM and AM/PM

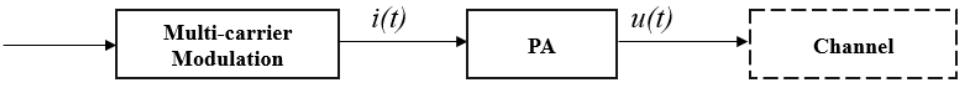


FIGURE 1.4 – Illustration of a simple transmitter structure with PA

FIGURE 1.4 illustrates a simplified transmitter structure with the existence of radio frequency(RF)

power amplifier (PA). Then, The complex envelope $i(t)$ of the signal at the input of the PA can be written as :

$$i(t) = \rho(t)e^{j\varphi(t)} \tag{1.24}$$

where $\rho(t)$ is the input signal modulus and $\varphi(t)$ is the input signal phase.

In practice, PA is considered as a nonlinear device. The power dependent nonlinear behavior can be characterized by AM/AM (amplitude to amplitude) and AM/PM (amplitude to phase) conversion characteristics. Then, the amplified complex envelope signal $u(t)$ can be written as :

$$u(t) = F_a(\rho(t))e^{j(F_p(\rho(t))+\varphi(t))} \tag{1.25}$$

where $F_a(\rho(t))$ is the AM/AM characteristic of the HPA and $F_p(\rho(t))$ is the AM/PM characteristic of the PA. When the PA is operated in its non-linear region, the actual output power is lower than the linearly amplified output. As the input signal power increases, the amount of non-linear distortion introduced by the PA gradually increases until it reaches saturation point.

This non-linear phenomenon will lead to in-band performance degradation and out-of-band spectrum regrowth, which leads to adjacent channel interference. For in-band signals, these distortions are quantified by the symbol error rate (SER) of the received signal and by the error vector magnitude (EVM) on the received constellation. In the following chapters, we will use the SER and the normalized mean square error (NMSE) in order to quantify in-band effects. It has to be pointed out that NMSE is generally expressed in dB while EVM is in percentage, but these two notions are very similar. PA nonlinearity is present for both BS PA (downlink, DL) and UE PA (uplink, UL). For UL and DL it is necessary to increase energy efficiency in order to increase the life duration of batteries in the UE and for decreasing the overall BS consumption at the BS side. The dynamic long-term memory effect of the HPA provides asymmetry in the process of sideband adjacent channel leakage ratio (ACLR) or intermodulation distortion (IMD) varying with envelope frequency [50] [51].

IBO and 1dB compression point

As we can see on FIGURE 1.5, in order to avoid or at least to reduce the nonlinear effects, the HPA is usually operated at a given Input Back-Off (IBO) from its 1dB compression point [52]. The 1dB

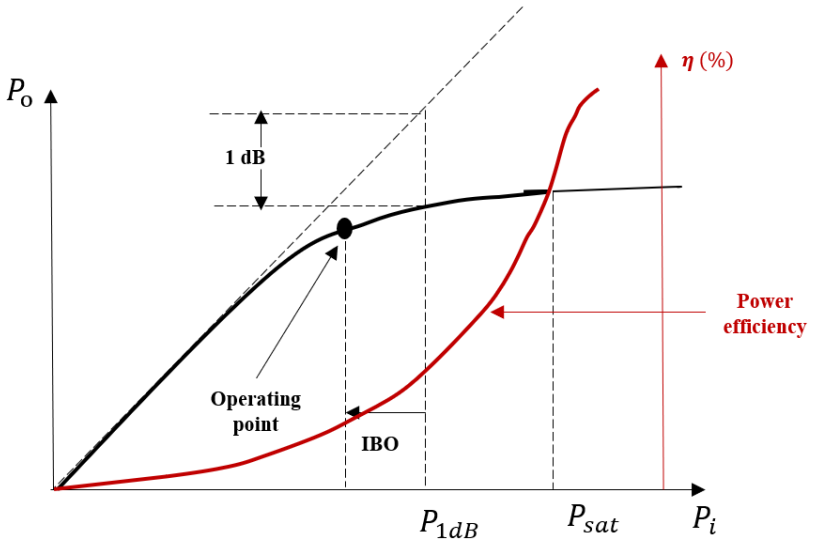


FIGURE 1.5 – Illustration of a nonlinear PA AM/AM characteristic with definition of 1 dB compression point

compression point refers to the input power level where the transfer characteristics of the amplifier have dropped by 1 dB from the ideal linear characteristics. The IBO is defined, in log scale, as follows

$$\text{IBO} = 10\log_{10}\left(\frac{P_{1\text{ dB}}}{P_i}\right) \tag{1.26}$$

where $P_{1\text{ dB}}$ is the input power at the 1dB compression point and P_i is the mean power of the input signal. Generally, the the linear region size of a PA mainly depends on the distance between the saturation point input power level P_{sat} and the $P_{1\text{ dB}}$. In addition, IBO adjustment aims to reduce the nonlinear distortion, that is, sacrificing efficiency in exchange for the linear operation, which is also the main reason for the contradiction between linearity and efficiency of the PA.

1.5.2 PA model

In 2G and previous communication systems, RF PA often exhibits weak memory effects due to the narrow signal bandwidth and low peak-to-average ratio (PAPR). The corresponding early power amplifier behavior models are mostly memory-less models or quasi-memory-less models, such as Polar Saleh model[53], Berman model [54] and look-up table (LUT) model [55]. However, with wireless communication systems development and frequency bandwidth increase PA must be considered as memory

devices. At the same time, the RF PA shows more stronger dynamic nonlinear characteristics. Under this circumstance, the quasi-memory-less based nonlinear model is no longer enough to accurately describe the non-linear behavior of PA.

Volterra series are well-known method to model nonlinear systems with memory accurately. By trimming and designing the Volterra series, a variety of polynomial models have been proposed one after another. The memory polynomial (MP) model only intercepts several items on the main diagonal of the full Volterra series [56] and it is the polynomial model with the lowest computational complexity. Based on the MP model, D.R.Morgan *el.al.* added the multi-diagonal basis functions of the Volterra series, and proposed a generalized memory polynomial (GMP) model [57]. There are also other models based on the Volterra series such as Dynamic Deviation Reduction (DDR) model [58] and hybrid memory polynomial (HMP) model [59] , *etc.*

In order to further enhance the flexibility of nonlinear modeling, neural networks (NN) are considered to be very useful with their flexible and high-precision nonlinear fitting capabilities. The early PA NN model used two complex multi-layer perception networks to fit the characteristics of AM-AM and AM-PM respectively [60], but the disadvantage of this is that the non-linearity of PA is divided into two parts and are fed to two independent networks for fitting. It is difficult to ensure that these two independent networks converge to the optimal point at the same time. In [61], the author uses directly complex NNs for nonlinear modeling, but the complex coefficient updating in the NN requires the use of complex gradient descent operations, and this stage will bring high computational complexity [62] [63].

In LTE and 5G NR reports, several PA models are provided. We will introduce two of them, which will be used in the following chapters.

Polynomial model

This model is based on the measured commercial 4GHz LTE PA [64]. It simultaneous exhibit both AM/AM and AM/PM distortions. The measured AM/AM and AM/PM curves of the PA model are approximated using a full-rank polynomial model with degree $P = 9$. The signal $u(t)$ at the output of the PA device can be written as :

$$u(t) = \sum_{l=1}^{P} a_l \rho(t) |\rho(t)|^{l-1} \tag{1.27}$$

where a_l are the complex coefficients of the polynomial approximation, obtained by using the classical

least square (LS) method. Then, the complex soft envelope of the signal $u(t)$ is :

$$S(\rho(t)) = F_a(\rho(t))e^{jF_p(\rho(t))} = \sum_{l=1}^{P} a_l \rho(t)^n \tag{1.28}$$

Modified Rapp model

Rapp model is proposed by the 3GPP for the NR evaluation [65]. In modified Rapp, which resembles closely to realistic PAs, AM/AM and AM/PM conversions can be given by :

$$F_a(\rho(t)) = \frac{G\rho(t)}{\left(1 + \left|\frac{G\rho(t)}{V_{sat}}\right|^{2p}\right)^{\frac{1}{2p}}}, \quad F_p(\rho(t)) = \frac{A\rho(t)^q}{\left(1 + \left(\frac{\rho(t)}{B}\right)^{2p}\right)} \tag{1.29}$$

where G is the linear gain, V_{sat} is the saturation level, p is the smoothness factor and A, B and q are fitting parameters.

1.5.3 PA nonlinear compensation

As described above, the wireless system suffers from PA nonlinearities, when the power efficiency of this latter is high, i.e., it is operated with low IBO. Therefore, in order to increase the PA efficiency without damaging the system performance, several solutions are proposed. The techniques can be divided into two classess : 1) the pre-distortion (PD) module which is implemented before PA and 2) PAPR reduction module which is the signal adjustment. When pre-distorsion is applied on the base-band sampled digital signal we call it digtal pre-distorsion (DPD). This kind of predistosion will be used in the following of the presented work. These two complementary technique are always implemented together.

DPD

DPD technology was proposed by James K. Cavers *et. al* [66] in the 1990s. After more than 20 years of development, DPD has become one of the most basic blocks in current wireless communication systems. The core idea of DPD is to first extract an inverse model of RF PA, and then cascade the extracted inverse model before the input of PA, so that the cascaded system of DPD and PA is linear [67]. As shown before in FIGURE 1.5, When the input power of PA is small, its input-output gain is approximately constant. When the input signal power gradually increases, PA gradually enters the nonlinear region and exhibits strong nonlinear characteristics. In order to compensate for this negative

effect, we insert a pre-distortion module before the input of PA. The gain of this pre-distortion module changes with the input signal power, and the trend is opposite to that of the PA, so that the final RF PA output signal is relative linear amplification to the original transmitted signal. The method is shown on FIGURE 1.6. Through the application of DPD technology, the linear working area of PA can be effectively extended, so that it can produce higher output power without obvious nonlinear distortion effects, and greatly improve the efficiency.

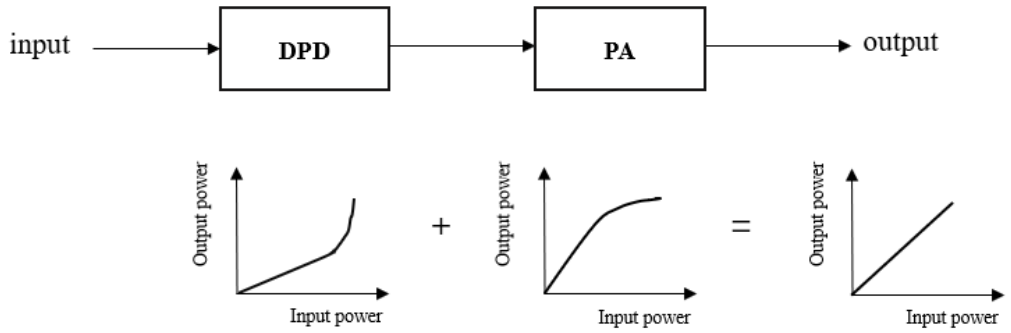


FIGURE 1.6 – Illustration of DPD method

PAPR reduction

Unlike single-carrier transmission where the variations of the signal are small and depend only on the shaping filter and the digital modulation used, multi-carrier transmission are characterized by a high dynamic range of the transmitted signal. Indeed, the time-domain signal of a multi-carrier modulation is the result of the sum of a large number of subcarriers, where the envelope of the signal has a huge fluctuation. This is often referred to high PAPR. This characteristic makes multi-carrier transmission very sensitive to the HPA nonlinearities.

In general, the PAPR of an OFDM signal is the ratio of the maximum power (peak power) of the current OFDM symbol to its average power [68] [69]. It can be written as :

$$PAPR[x(t)] = \frac{\max_{0\leq t\leq NT}\left[|x(t)|^2\right]}{P_{\text{average}}} \tag{1.30}$$

where the $P_{average}$ is given by :

$$P_{\text{average}} = \frac{1}{T}\int_0^T |x(t)|^2 dt \tag{1.31}$$

with N samples of the OFDM signal and sample duration T.

For a better approximation of the PAPR of OFDM signals at continuous time domain, the sampled signal is L-times over-sampled. Therefore, the output sample $x_k, 0 \leq k \leq NL-1$, after $NL-IFFT$ can be expressed as :

$$x_k = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} X_n e^{\frac{j2\pi nk}{NL}} \tag{1.32}$$

Then, the PAPR of x_k is defined as :

$$PAPR\left[x_k\right] = \frac{\max_{0\leq k\leq NL-1}\left[|x_k|^2\right]}{E\left[|x_k|^2\right]} \tag{1.33}$$

where $\mathbf{E}[.]$ is the mathematical expectation operator.

The PAPR of the multi-carrier signal is a random variable. In the literature, we are generally interested in the cumulative complementary distribution function (CCDF) of PAPR. The distribution is a key factor in multi-carrier system as it evaluates the PAPR level. It denotes the probability that the PAPR of the signal exceeds a given threshold $PAPR_0$, as shown on the following equation.

$$CCDF[x_k]\left(PAPR_0\right) = \text{Prob}\left(PAPR[x_k] > PAPR_0\right) \tag{1.34}$$

In the literature, there are several methods of reducing PAPR which are classified into three categories : coding methods, probabilistic methods and signal addition methods.

Coding methods [70] are early technology used for PAPR reduction. The basic idea of this technology is to encode the input binary data and to select only the code words characterized by a low PAPR, in order to avoid transmitting multi-carrier symbols at a high PAPR. However, it is applicable only for systems with small number of subcarriers[70].

There are different types of probabilistic methods for PAPR reduction such as Partial Transmit Sequence (PTS) and Selective Mapping (SLM) methods. In the PTS method [71] [72], each input data block is divided into sub-blocks. The subcarriers in each subblock are weighted by a phase factor, such that the PAPR of the signal is minimized. The principle of the SLM method [71] [73] is to make phase rotations for each symbol and to choose the symbol having the smallest PAPR. The drawback of these two methods is the necessity to send side information to the receiver.

The signal addition methods consist in adding the PAPR reduction signal x_{papr} to the useful signal

x_k, aiming to reduce the peaks and reach the result $PAPR\,[x_k + x_{papr}] < PAPR[x_k]$. There are several methods in literature like Tone Reservation (TR) [74], Clipping and Filtering [75], Tone Injection (TI) [76], *etc.* TR and TI are PAPR reduction techniques without degradation of the OFDM signal (i.e., no in-band effects due to the PAPR reduction technique) while clipping degrades the OFDM signal integrity. The advantage of these methods is that no side information is necessary.

Chapitre 2

Massive MIMO for Inter-numerology interference cancellation

Contents

2.1 Introduction

The new 5G wireless system is expected to enable a number of services with varying specifications. MIMO technology was developed to meet the demand for increased spectral efficiency and improved connection reliability through spacial multiplexing gain and antenna diversity gain. This chapter will cover the use of spatial multiplexing to link users who share the same bandwidth but have different numerologies. Precoding designs are presented with the aim of managing mixed numerologies spectrum sharing (SS) transmission. Then, in massive MIMO-OFDM systems, we investigate internumerology interference (INI) and derive theoretical expressions for its radiation pattern. Mixed numerology schemes are being proposed, in which the numerology varies depending on the service characteristics required. Subcarrier spacing (SCS), symbol size, and cyclic prefix (CP) length are all parameters that can be set in OFDM using different numerologies. In [77], authors presented the first studies to present the multi-numerology or mixed numerology method. Mixed numerologies schemes are also included in the Third Generation Partnership Project (3GPP) NR standardization and have been the subject of numerous studies [78][79][80][81][82]. According to the above studies, the most commonly accepted approach for supporting diverse services is to divide the bandwidth into many subbands, then introduce flexible numerology allocation in each subband. Although the use of numerology multiplexing increases system stability, interference between users of different numerologies exists, and inter-numerology interference (INI) has an effect on system efficiency. In [83], an INI model is constructed as a function of numerology spacing, overlapping windows, and channel frequency response. While theoretical expressions are developed, this model is limited to a SISO Windowed-OFDM system and users occupying adjacent subbands. In [84], which investigated INI problems and explained their underlying causes such as SCS, number of activated subcarriers, control, and so on, factors that affect INI in a mixed numerologies OFDM system are discussed. This work, however, is still limited to SISO systems with adjacent bands, and only simulation results have been presented. Also, where the INI occurs near the junction of the two subbands, [85] produces similar findings as the two papers above. Some recent researches [86] [12] have focused on mixed numerologies SS, also known as overlapping mixed numerologies. Unlike non-overlapping mixed numerologies systems, it is difficult to prevent interference using windowed/filtered waveforms in mixed numerologies SS systems (i.e., users share the same time/frequency resources). In [12], the authors proposed a new transceiver design for classical MIMO-OFDM using a mixed numerologies SS scheme, where the users have different SCS.

The interference pattern is deduced, but no theoretical model is suggested, and only simulation results are shown.

It is beneficial to allocate the maximum bandwidth to all users in massive MIMO-OFDM systems in order to maintain high spectral performance when offering scalable services. It is worth noting that this SS scheme differs from the SISO/classical MIMO scenario, in which adjacent bandwidths divided by a guard band are allocated to different users and INI occurs at the subband edges [81][82]. In addition, unlike in a SISO/classical MIMO system with different numerologies, MUI in a massive MIMO system has two parts : Intra-Numerology Interference (Intra-NI) and INI, where the former refers to interference between users of the same numerology and the latter refers to interference between users of different numerologies. Due to the various antennas equipped at the BS, the Intra-NI is cancelled by precoders or detection method using the high dimensional degrees of freedom (DoFs). In comparison to classical MIMO, some hypothesis can be applied to massive MIMO to simplify the computation of an analytical model for the INI with high accuracy. Furthermore, how MU massive MIMO-OFDM systems work in mixed numerologies SS transmission is still unknown. In addition to the studied schemes and analysis, a good INI cancellation algorithm is needed for good results.

In this chapter, we will bring the following contents :

- To allow flexible management of mixed numerologies SS transmission, improved transceiver schemes are developed for the massive MIMO-OFDM downlink and uplink systems. Many patterns, such as SCS, channel selectivity, and power allocation, are analyzed in the INI.
- For massive MIMO-OFDM systems with mixed numerologies, analytical INI models are created, which could be useful in guiding 5G system design and parameter selection. We derive theoretical INI expressions in closed-form to investigate the effect of INI on users' results theoretically. Theoretical findings match simulation results, proving the validity of our theoretical research.
- We propose an INI cancellation methods based on the derived theoretical expressions of INI between different numerologies for massive MIMO-OFDM downlink and uplink. The results show that the proposed cancellation method can reduce the INI and improve the transmission quality.

Special Notations : In this chapter, the superscript represent the numerology index (e.g. $\mathbf{X}^{(1)}$). The k-th element in the vector $\mathbf{x}$ is denoted as $\mathbf{x}[k]$. For a $M \times N$ matrix $\mathbf{X}$, $\mathbf{X}(m,:)$ denote the m-th

line and $\mathbf{X}(:, n)$ denote the n-th column.

2.2 Massive MIMO-OFDM downlink : INI theoretical analysis

2.2.1 Massive MIMO-OFDM downlink system model

We consider a massive MIMO-OFDM downlink scheme with various numerologies. The BS has M_t antennas that serve M_r single-antenna users with different numerologies over a frequency-selective channel, where M_t is far larger than M_r. NUM numerologies, represented by index num, where $num = 1, ..., NUM$, can be used to divide the M_r users into NUM groups. N_{num} and CP_{num} denote the IFFT/FFT size and CP size of group num, respectively. The proposed transceiver architecture differs from that of [86], in which we added zeros to each user's signal structure to allow flexible INI management. Our transceiver can suppress Intra-NI by using several precoders in this way, and the added zeros enable INI cancellation, which will be addressed later. The BS sends data, over the m_t-th antenna, to the m_r-th user via channel $\sqrt{\alpha_{m_r}}\mathbf{ht}_{m_r,m_t}$, where α_{m_r} is the large-scale fading, $\mathbf{ht}_{m_r,m_t} \in \mathbb{C}^{1\times D}$ is the channel impulse response between transmitting antenna m_t and user m_r, $m_t = 1...M_t$, $m_r = 1...M_r$ and D is the number of taps.

Then,

$$\mathbf{hf}_{m_r,m_t}^{(num)} = FFT(\sqrt{\alpha_{m_r}}\mathbf{ht}_{m_r,m_t}, N_{num}). \tag{2.1}$$

represents the channel frequency response with FFT size N_{num}. For a massive MIMO-OFDM system, the channel frequency response is denoted by $\bar{\mathbf{H}}^{(num)} \in \mathbb{C}^{M_r\times M_t\times N_{num}}$, where

$$\bar{\mathbf{H}}^{(num)}(m_r, m_t, :) = \mathbf{hf}_{m_r,m_t}^{(num)}. \tag{2.2}$$

The construction of $\bar{\mathbf{H}}^{num}$ is illustrated on FIGURE. 2.1.

In this chapter, in order to keep the scheme simple without losing its generality, we consider two users ($M_r = 2$) using two separate numerologies, as shown in FIGURE 2.2. For users 1 and 2, $\mathbf{s}^{(1)}, \mathbf{s}^{(2)} \in \mathbb{C}^{1\times N_1}$ are signals produced by M-quadrature amplitude modulation (QAM). The least common multiplier (LCM) methods [12], as shown in FIGURE 2.3, is used to achieve synchronization. We assume that $N_1 = N \times N_2, CP_1 = N \times CP_2$, where $N = 2^i$ and i is an integer. This generalized synchronized scenario has also been considered in previous works [80] [83]. $\mathbf{x}_{m_t}^{(1)}$ is the modulated OFDM symbol on the m_t-th transmitting antenna for user 1 and $\mathbf{x}_{m_t}^{(2,n)}$ is user 2's n-th modulated

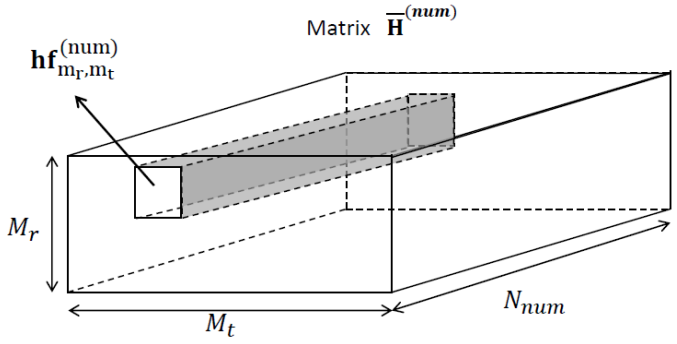


FIGURE 2.1 – Structure of matrix $\bar{\mathbf{H}}^{num}$

OFDM symbol on the m_t-th transmitting antenna. $\mathbf{x}^{(1)}$ has a total symbol length of N times the symbol length of $\mathbf{x}^{(2,n)}$, and all of the symbols are aligned.

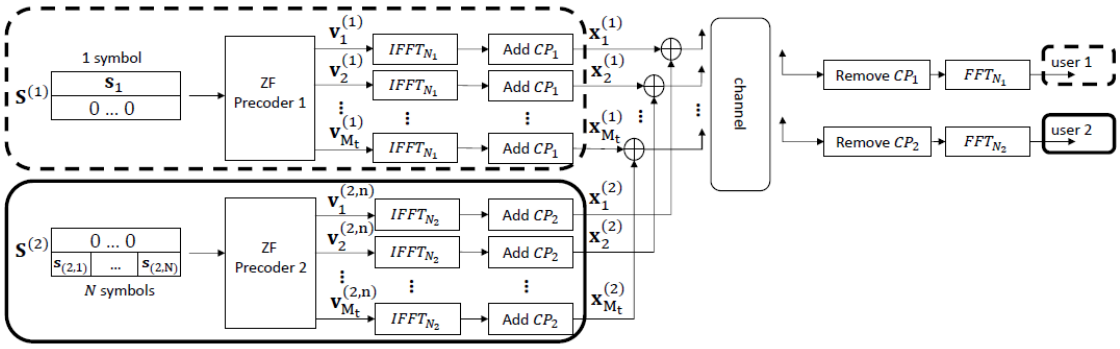


FIGURE 2.2 – System model of the massive MIMO OFDM downlink with two different numerologies : M_t transmit antennas at the BS, two sigle-antenna terminals, two blocks illustrate two different numerologies with OFDM of N_1 and N_2 subcarriers.

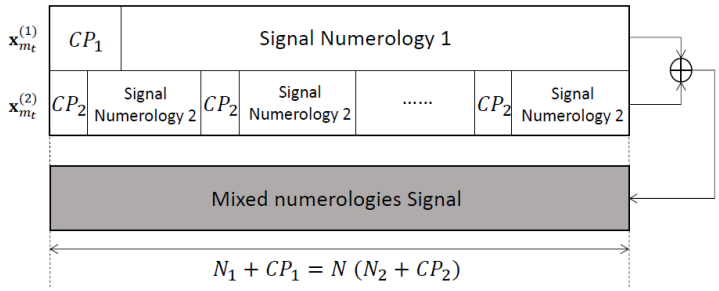


FIGURE 2.3 – Signal synchronization and summation. $\mathbf{x}_{m_t}^{(1)}$ is user 1's signal on mt-th transmitting antenna, $\mathbf{x}_{m_t}^{(2)}$ is user 2's signal on m_t-th transmitting antenna.

At the BS, a linear precoding scheme is used to eliminate the Intra-NI at the receivers. We consider

two linear precoding divisions, each for one numerology, where two zero-forcing (ZF) precoders are used to allow flexible management [87]. On FIGURE 2.2, the first line of matrix $\mathbf{S}^{(1)}$ is the data knowledge vector $\mathbf{s}^{(1)}$ in numerology 1 for user 1, while the second line is set to zero to protect user 2. It is also worth noting that matrix $\mathbf{S}^{(2)}$ contains data vector $\mathbf{s}^{(2)}$ for user 2 in the second branch (solid-line block), while null-vector is prepared for user 1.

Then, at each branch, ZF precoding is performed, which is designed to fully eliminate Intra-NI between users within the same serving cell as introduced in the previous chapter. As a result, the signal matrices $\mathbf{S}^{(1)}, \mathbf{S}^{(2)} \in \mathbb{C}^{2\times N_1}$ are linearly coded as follows

$$\mathbf{u}_{n_1}^{(1)} = \frac{1}{\sqrt{\zeta_{n_1}^{(1)}}} \mathbf{W}_{n_1}^{(1)} \mathbf{s}_{n_1}^{(1)}, n_1 = 1, ..., N_1, \tag{2.3}$$

$$\mathbf{u}_{n_2}^{(2,n)} = \frac{1}{\sqrt{\zeta_{n_2}^{(2,n)}}} \mathbf{W}_{n_2}^{(2)} \mathbf{s}_{n_2}^{(2,n)}, n_2 = 1, ..., N_2 \,\&\, n = 1, ...N. \tag{2.4}$$

where $\zeta_{n_1}^{(1)}$, $\zeta_{n_2}^{(2,n)}$ are power normalization factors, $\mathbf{W}_{n_{num}}^{(num)} \in \mathbb{C}^{M_t \times M_r}$ denotes the ZF precoding matrices for n_{num}-th OFDM subcarrier of numerology num. $\mathbf{s}_{n1}^{(1)} = \mathbf{S}^{(1)}(:, n_1) \in \mathbb{C}^{2\times 1}, \mathbf{u}_{n1}^{(1)} \in \mathbb{C}^{M_t \times 1}$ represent the original signal on the n_1-th subcarrier and the precoded vectors of the n_1-th subcarrier with numerology 1, $\mathbf{s}_{n_2}^{(2,n)} = \mathbf{S}^{(2)}(:, n_2 + (n-1)\cdot N_2) \in \mathbb{C}^{2\times 1}, \mathbf{u}_{n_2}^{(2,n)} \in \mathbb{C}^{M_t \times 1}$ are the original signal of the n_2-th subcarrier for n-th symbol and the precoded vector of the n_2-th subcarrier for n-th symbol with numerology 2. The ZF precoding matrices are constructed as

$$\mathbf{W}_{n_{num}}^{(num)} = \mathbf{H}_{n_{num}}^{(num)H} \left(\mathbf{H}_{n_{num}}^{(num)} \mathbf{H}_{n_{num}}^{(num)H} \right)^{-1} \tag{2.5}$$

where $\mathbf{H}_{n_{num}}^{(num)} = \bar{\mathbf{H}}^{(num)}(:, :, n_{num}) \in \mathbb{C}^{2\times M_t}$ is the MIMO channel frequency response of the n_{num}-th OFDM subcarrier with numerology num. We assume that channel matrices are perfectly estimated at the BS, which can be calculated using time division multiplexing (TDD) systems' channel reciprocity. [88][89].

After precoding, the M_t-dimension vector $\mathbf{u}_{n_1}^{(1)}$ and $\mathbf{u}_{n_2}^{(2,n)}$ are reordered in order to construct matrices $\mathbf{V}^{(1)}$ and $\mathbf{V}^{(2,n)}$, according to the following mapping

$$\mathbf{v}_{m_t}^{(1)} = \left[\mathbf{u}_1^{(1)}[m_t] \dots \mathbf{u}_{N_1}^{(1)}[m_t] \right], \tag{2.6}$$

$$\mathbf{v}_{m_t}^{(2,n)} = \left[\mathbf{u}_1^{(2,n)}[m_t] \dots \mathbf{u}_{N_2}^{(2,n)}[m_t]\right]. \tag{2.7}$$

where $\mathbf{v}_{m_t}^{(1)} \in \mathbb{C}^{1\times N_1}$ and $\mathbf{v}_{m_t}^{(2,n)} \in \mathbb{C}^{1\times N_2}$ denote each line of the precoded signals.

The structure of matrix $\mathbf{V}^{(1)}$ is shown in FIGURE 2.4, and the structure of matrix $\mathbf{V}^{(2,n)}$ is similar. The time-domain signals are then obtained by using IFFTs and adding CPs. For each transmitting antenna, signals of two different numerologies are added point by point before transmission over wireless channel. FIGURE 2.3 also illustrates the process of adding at each transmit antenna m_t. It is worth noting that users in the proposed system share the same bandwidth, resulting in high spectrum performance. Inter-symbol interference (ISI) is avoided at the receiver side if the length of CP_{num} is greater than the length of the multi-path channel.

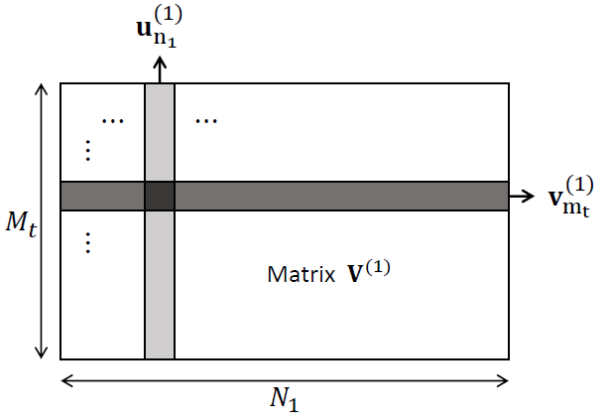


FIGURE 2.4 – Structure of matrix $\mathbf{V}^{(1)}$

Since $\mathbf{H}_{n_{num}}^{(num)} \left(\mathbf{H}_{n_{num}}^{(num)}\right)^{\dagger} = \mathbf{I}_{Mr}$, transmitting $\mathbf{u}_{n_{num}}^{(num)} = \left(\mathbf{H}_{n_{num}}^{(num)}\right)^{\dagger} \mathbf{s}_{n_{num}}^{(num)}$ perfectly eliminate all the Intra-NI, which is part of the total MUI. The suppression of Intra-NI with ZF precoders can be expressed as follows for, respectively, user 1 and user 2.

$$\sum_{m_t=1}^{M_t} \mathbf{hf}_{2,m_t}^{(1)}[n_1]\mathbf{v}_{m_t}^{(1)}[n_1] = 0, n_1 = 1, ..., N_1, \tag{2.8}$$

$$\sum_{m_t=1}^{M_t} \mathbf{hf}_{1,m_t}^{(2)}[n_2]\mathbf{v}_{m_t}^{(2,n)}[n_2] = 0, n = 1, ..., N \& \; n_2 = 1, ..., N_2. \tag{2.9}$$

It is worth noting that equations (2.8) and (2.9) are only fulfilled when $M_t \gg M_r$, indicating that the ZF precoders work well and can fully suppress Intra-numerology MUI. This obviously does not

apply in SISO systems, and even less so in classical MIMO, where the number of BS transmitting antennas is insufficient. As a result, the analysis of INI between users using different numerologies that we will perform in the next few paragraphs will only be true for massive MIMO-OFDM systems. It will also be different from previous studies for SISO and classical MIMO systems. The obtained signal can be written as follows, in the case of INI between users who use different numerologies.

$$\mathbf{r}_1[n_1] = \mathbf{s}_1[n_1] + \mathbf{ini}^{(2,1)}[n_1] + \mathbf{b}_1[n_1], n_1 = 1, ..., N_1, \tag{2.10}$$

$$\mathbf{r}_{2,n}[n_2] = \mathbf{s}_{2,n}[n_2] + \mathbf{ini}_n^{(1,2)}[n_2] + \mathbf{b}_{2,n}[n_2], n_2 = 1, ..., N_2. \tag{2.11}$$

where $\mathbf{y}_1$ is the received signal by user 1, $\mathbf{y}_{2,n}$ is the n-th received signal obtained by user 2. INI from numerology 2 to numerology 1 and INI from numerology 1 to n-th numerology 2 symbol are $\mathbf{ini}^{(2,1)}$ and $\mathbf{ini}_n^{(1,2)}$, respectively. The received noises $\mathbf{b}_1$ and $\mathbf{b}_{2,n}$ are i.i.d circularly-symmetric complex Gaussian distributions with zero-mean and ${\sigma_b}^2$ variance.

2.2.2 Inter-numerology interference analysis

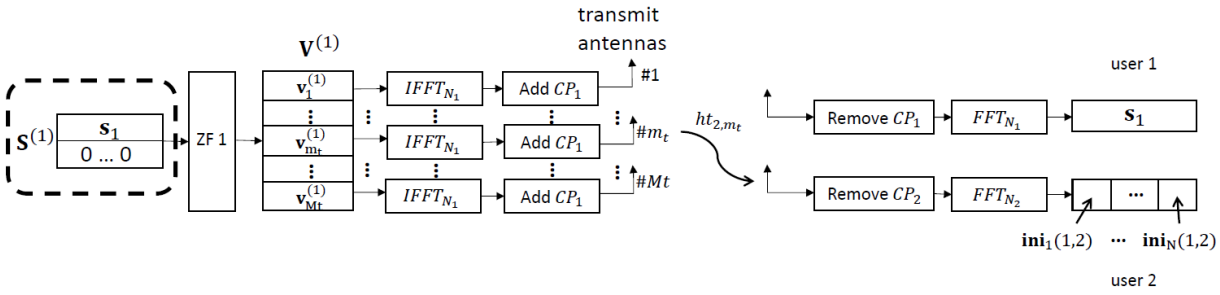
INI from numerology 1 with N_1 to numerology 2 with $N_2 = N_1/N$

First, we consider the user 2 using numerology 2, which is with smaller IFFT/FFT size N_2.

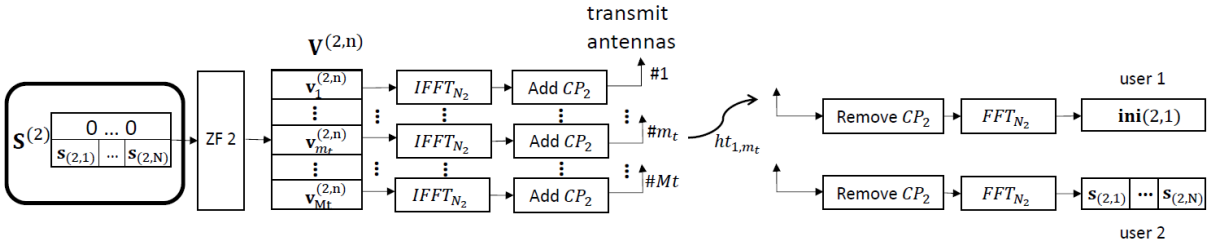
User 2 receives an all-zero signal if only the dash-line block in FIGURE 2.2 is enabled. Since the ZF precoder removes all MUI, we only get INI from user 1 which is using a different numerology. FIGURE 2.5 (a) demonstrates the transmission chain. For user 2, a single symbol modulated with numerology 1 corresponds to N INI symbols. INI is derived from the IFFT/FFT window sizes being unequal, as seen in FIGURE 2.6 (a), where the FFT window of numerology 2 is smaller than that of numerology 1. In the case of a transmission between transmit antenna m_t and user 2, the n-th received symbol can be expressed , in frequency domain, as follows

$$\mathbf{r}_{2,n,m_t} = \mathbf{G}^{(1)}\mathbf{B}_n^{(1)} Diag\left(\mathbf{hf}_{2,m_t}^{(1)}\right)\left(\mathbf{v}_{m_t}^{(1)}\right)^T. \tag{2.12}$$

where $\mathbf{r}_{2,n,m_t} \in \mathbb{C}^{1\times N_2}$ is the received n-th symbol on user 2 with numerology 2, transmitting from the m_t-th transmit antenna. $\mathbf{G}^{(1)} = \mathbf{DFT}_{N_2} \in \mathbb{C}^{N_2\times N_2}$, $\mathbf{hf}_{2,m_t}^{(1)} \in \mathbb{C}^{1\times N_1}$ is the MIMO channel frequency response between m_t-th transmitting antenna and user 2. $\mathbf{B}_n^{(1)} \in \mathbb{C}^{N_2\times N_1}$ is the n-th $N_2 \times N_1$ part of



(a) Transmission where zero-signal is prepared for user 2. At the receiver side, user 2 will only receives INI

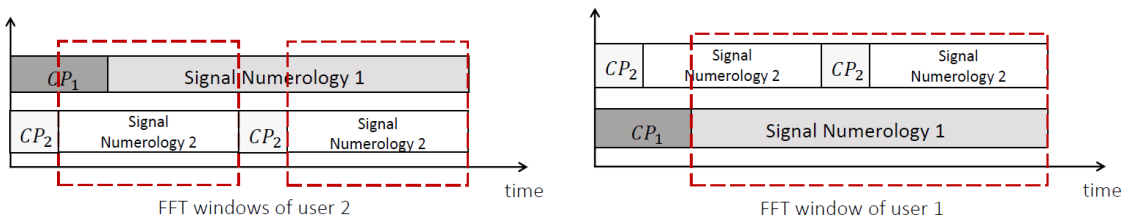


(b) Transmission where zero-signal is prepared for user 1. At the receiver side, user 1 will only receives INI

FIGURE 2.5 – Systems used when analysing different INIs

matrix $\mathbf{B}^{(1)}$, where $\mathbf{B}^{(1)} \in \mathbb{C}^{N_1 \times N_1}$ is a rotated version of matrix IDFT. The structure of matrix $\mathbf{B}^{(1)}$ is represented on FIGURE 2.7.

From numerology 1 to numerology 2, the INI can then be seen as the summation of $\mathbf{y}_{2,n,m_t}$ in



(a) INI from user 1 to user 2

(b) INI from user 2 to user 1

FIGURE 2.6 – FFT windows and INI on the receiver side

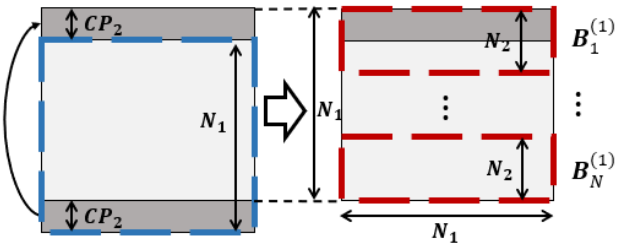


FIGURE 2.7 – Structure of matrix $\mathbf{B}^{(1)}$ and matrices $\mathbf{B}_n^{(1)}$. The blue dash block is $\mathbf{IDFT}_{N_1}$, we rotate the last $CP_2 \times N_1$ part above to get $\mathbf{B}^{(1)}$ as shown on the left side. Then, $\mathbf{B}_n^{(1)}$ is the n-th $N_2 \times N_1$ part of $\mathbf{B}^{(1)}$ as the red dash blocks shown on the right side.

(2.12) over all transmitting antennas. This INI is given by :

$$\begin{aligned}\mathbf{ini}_n^{(1,2)} &= \sum_{mt=1}^{M_t} \mathbf{r}_{2,n,m_t} \\ &= \sum_{m_t=1}^{M_t} \mathbf{G}^{(1)}\mathbf{B}_n^{(1)} Diag\left(\mathbf{hf}_{2,m_t}^{(1)}\right)\left(\mathbf{v}_{m_t}^{(1)}\right)^T = \mathbf{G}^{(1)}\mathbf{B}_n^{(1)} \sum_{n_1=1}^{N_1}\sum_{m_t=1}^{M_t} \mathbf{hf}_{2,m_t}^{(1)}[n_1]\mathbf{v}_{m_t}^{(1)}[n_1].\end{aligned} \tag{2.13}$$

According to (2.8), we have $\sum_{m_t=1}^{M_t} \mathbf{hf}_{2,m_t}^{(1)}[n_1]\mathbf{v}_{m_t}^{(1)}[n_1] = 0$, then, we get

$$\mathbf{ini}_n^{(1,2)} = 0 \qquad n = 1, ..., N. \tag{2.14}$$

.

As a conclusion, we can say that there is no INI from numerology 1 to numerology 2 since the proposed precoding scheme protects all subcarriers occupied by user 2.

Remark 1 : Performance of users using small numerology are well protected against interference coming from higher numerology, i.e., their performance is not degradated.

As a result, unlike SISO/classical MIMO, where INI exists on small numerology (IFFT/FFT size) [83][86], the INI of users with small IFFT/FFT size is completely suppressed when using the proposed transceiver design for massive MIMO-OFDM. This result shows another interest of Massive MIMO technology for OFDM transmissions using different numerologies.

INI from numerology 2 with N_2 to numerology 1 with $N_1 = N \times N_2$

In comparison to user 2, user 1's transmission is not properly protected by the ZF precoder, the INI analysis will be presented in the following.

When only the solid-line block on FIGURE 2.1 is enabled, user 1 receives zeros. In this case, the received signal on user 1 is the INI induced by signal from user 2, as illustrated by FIGURE 2.5 (b). In this scenario, N symbols of numerology 2 refer to one symbol with user 1 since $N_1 = N \times N_2$. The INI is induced by an uneven FFT window on the receiver side, as seen in FIGURE 2.6 (b). The sum of INI from each symbol modulated with numerology 2 is the total INI, which can be expressed as

$$\mathbf{ini}^{(2,1)} = \sum_{n=1}^{N} \mathbf{ini}_{n}^{(2,1)} = \sum_{n=1}^{N} \sum_{n_2=1}^{N_2} \mathbf{ini}_{n,n_2}^{(2,1)}. \tag{2.15}$$

where $n = 1, .., N$, $n_2 = 1, ..., N_2$, $\mathbf{ini}_{n}^{(2,1)}$ represents the interference from the n-th symbol with numerology 2 and $\mathbf{ini}_{n,n2}^{(2,1)}$ represents the interference from the n_2-th subcarrier in n-th symbol with numerology 2.

If we focus on the m_t-th transmit antenna, we can derive the theoretical expression of the obtained signal on user 1 as

$$\mathbf{r}_{1,n,m_t} = \begin{cases} \mathbf{G}_1^{(2)}\mathbf{B}_1^{(2)} Diag\left(\mathbf{hf}_{1,m_t}^{(1)}\right) \mathbf{G}_2^{(2)}\mathbf{B}_2^{(2)} \left(\mathbf{v}_{m_t}^{(2,n)}\right)^T \\ for \quad n = 1, \\ \\ \mathbf{G}_2^{(2)}\mathbf{ID}_n^{(2)}\mathbf{B}_3^{(2)} Diag\left(\mathbf{hf}_{1,m_t}^{(1)}\right) \mathbf{G}_2^{(2)}\mathbf{B}_2^{(2)} \left(\mathbf{v}_{m_t}^{(2,n)}\right)^T \\ for \quad n = 2, ..., N. \end{cases} \tag{2.16}$$

where $\mathbf{y}_{1,n,m_t}$ is the received INI caused by the n-th symbol with numerology 2. $\mathbf{hf}_{1,m_t}^{(1)} \in \mathbb{C}^{1\times N_1}$ is the MIMO channel frequency response between the m_t-th transmitting antenna and user 1. All the other matrices in the expression are

$$\mathbf{G}_1^{(2)} = \begin{bmatrix} \mathbf{0}_{N_1\times CP_1} & \mathbf{DFT}_{N_1} \end{bmatrix} \in \mathbb{C}^{N_1\times(N_1+CP_1)}, \mathbf{G}_2^{(2)} = \mathbf{DFT}_{N_1} \in \mathbb{C}^{N_1\times N_1},$$

$$\mathbf{B}_1^{(2)} = \begin{bmatrix} \mathbf{IDFT}_{N_1} \\ \mathbf{0}_{CP_1\times N_1} \end{bmatrix} \in \mathbb{C}^{(N_1+CP_1)\times N_1}, \mathbf{B}_3^{(2)} = \mathbf{IDFT}_{N_1} \in \mathbb{C}^{N_1\times N_1},$$

$$\mathbf{ID}_n^{(2)} = \begin{bmatrix} \mathbf{0}_{T_1\times(N_1/2)} & \mathbf{0}_{T_1\times(N_1/2)} \\ \mathbf{I}_{T_2} & \mathbf{0}_{T_2\times(N_1-T_2)} \\ \mathbf{0}_{T_3\times(N_1/2)} & \mathbf{0}_{T_3\times(N_1/2)} \end{bmatrix} \in \mathbb{C}^{N_1\times N_1} \ .$$

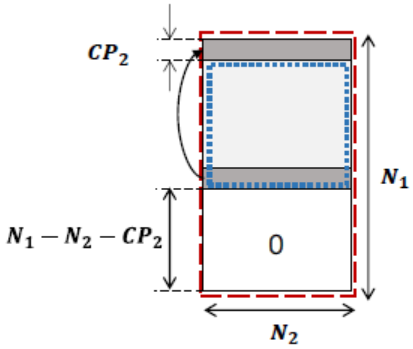


FIGURE 2.8 – Structure of matrix $\mathbf{B}_2^{(2)}$ with red dash box. The blue dot block is $\mathbf{IDFT}_{N_2}$, we copy the last $CP_2 \times N_2$ part above and add zeros below to extend the dimension to $N_1 \times N_2$.

for matrix $\mathbf{ID}_n^{(2)}$, we have

$$\begin{cases} T_1 = N_1 - (N - n + 1)(N_2 + CP_2), \\ T_2 = \begin{cases} N_2 + CP_2 + D \quad for \quad n = 2, ..., N-1 \\ N_2 + CP_2 \quad for \quad n = N, \end{cases} \\ T_3 = \begin{cases} (N-n)(N_2 + CP_2) - D \quad for \quad n = 2, ..., N-1 \\ (N-n)(N_2 + CP_2) \quad for \quad n = N. \end{cases} \end{cases} \tag{2.17}$$

The structure of matrix $\mathbf{B}_2^{(2)} \in \mathbb{C}^{N_1 \times N_2}$ is illustrated on FIGURE 2.8.

The signal obtained by user 1 is the summation from all transmit antennas, thus :

$$\mathbf{ini}_n^{(2,1)} = \sum_{m_t=1}^{M_t} \mathbf{r}_{1,n,m_t}. \tag{2.18}$$

From (2.16), we can rewrite (2.18) into

$$\mathbf{ini}_n^{(2,1)} = \sum_{m_t=1}^{M_t} \mathbf{E}_n Diag\left(\mathbf{hf}_{1,m_t}^{(1)}\right) \mathbf{Z} \left(\mathbf{v}_{m_t}^{(2,n)}\right)^T. \tag{2.19}$$

where

$$\mathbf{E}_n \in \mathbb{C}^{N_1 \times N_1} = \begin{cases} \mathbf{G}_1^{(2)} \mathbf{B}_1^{(2)} & for \quad n = 1, \\ \mathbf{G}_2^{(2)} \mathbf{ID}_n^{(2)} \mathbf{B}_3^{(2)} & for \quad n = 2, ..., N. \end{cases} \tag{2.20}$$

$\mathbf{Z} = \mathbf{G}_2^{(2)} \mathbf{B}_2^{(2)} \in \mathbb{C}^{N_1 \times N_2}$ is identical for all symbols. The INI on the n_1-th subcarrier with nume-

rology 1 from the n-th symbol with numerology 2 is then

$$\begin{aligned}\mathbf{ini}_n^{(2,1)}[n_1] &= \sum_{n_1'=1}^{N_1}\sum_{n_2=1}^{N_2}\sum_{m_t=1}^{M_t}\mathbf{E}_n(n_1,n_1')\mathbf{hf}_{1,m_t}^{(1)}[n_1']\mathbf{Z}(n_1',n_2)\mathbf{v}_{m_t}^{(2,n)}[n_2] \\ &= \sum_{n_1'=1}^{N_1}\sum_{n_2=1}^{N_2}\mathbf{E}_n(n_1,n_1')\mathbf{Z}(n_1',n_2)\sum_{m_t=1}^{M_t}\mathbf{hf}_{1,m_t}^{(1)}[n_1']\mathbf{v}_{m_t}^{(2,n)}[n_2].\end{aligned} \tag{2.21}$$

where $n_1 = 1, ..., N_1, n_1' = 1, ..., N_1$, $n_2 = 1, ..., N_2$ and $m_t = 1, ..., M_t$.

Due to the frequency selectivity of the channel response, $\mathbf{ini}_n^{(2,1)}$ in (2.21) is not equal to zero. (2.9) is not satisfied since $\mathbf{hf}_{1,m_t}^{(1)}[n_1'] \neq \mathbf{hf}_{1,m_t}^{(2)}[n_2]$, where $n_1' = 1, ..., N_1, n_2 = 1, ..., N_2$. In a selective channel, we can deduce that user 1 would be affected by INI induced by user 2.

Remark 2 : In a constant channel, $\mathbf{ini}^{(2,1)}$ *is zero due to the channel frequency response characteristic.*

The two MIMO channels $\mathbf{hf}_{1,m_t}^{(1)} = FFT(\mathbf{ht}_{1,m_t}, N_1) \in \mathbb{C}^{1\times N_1}$ and $\mathbf{hf}_{1,m_t}^{(2)} = FFT(\mathbf{ht}_{1,m_t}, N_2) \in \mathbb{C}^{1\times N_2}$ are two frequency responses of the same impulse response with different length. When the channel remains constant, we get the following equation :

$$\mathbf{hf}_{1,m_t}^{(1)}[n_1] = \mathbf{hf}_{1,m_t}^{(2)}[n_2]. \tag{2.22}$$

where $n_1 = 1, ...N_1, n_2 = 1, ...N_2$.

From (2.9) and (2.22), all of the elements inside the two frequency responses are identical. Then, for all of user 1's subcarriers, we have

$$\mathbf{ini}_n^{(2,1)}[n_1] = \sum_{n_1'=1}^{N_1}\sum_{n_2=1}^{N_2}\mathbf{E}_n(n_1,n_1')\mathbf{Z}(n_1',n_2)\sum_{m_t=1}^{M_t}\mathbf{hf}_{1,m_t}^{(2)}[n_2]\mathbf{v}_{m_t}^{(2,n)}[n_2] = 0. \tag{2.23}$$

Unlike SISO and classical MIMO systems, Massive MIMO, ensures no INI when mixed numerologies are used over flat-fading channels.

Remark 3 : $\mathbf{ini}^{(2,1)}$ *is increased by channel selectivity and difference between* N_1 *and* N_2*.*

Vector $\mathbf{hf}_{1,m_t}^{(1)}$ is the interpolation of vector $\mathbf{hf}_{1,m_t}^{(2)}$ when the channel is frequency-selective. In a frequency-selective channel, in order to build vector $\mathbf{hf}_{1,m_t}^{(1)}$, $N-1$ values are added between every two

points in $\mathbf{hf}_{1,mt}^{(2)}$. From this characteristic, we can obtain

$$\mathbf{hf}_{1,m_t}^{(1)}[(n_2-1)N+1] = \mathbf{hf}_{1,m_t}^{(2)}[n_2]. \tag{2.24}$$

where $n_2 = 1, ...N_2$.

We can begin by considering the n_2-th subcarrier of the n-th symbol with numerology 2. Its impact on the n_1-th subcarrier with numerology 1 is derived in (2.21). According to (2.38) and (2.9), the influence of n_2-th subcarrier is zero only when $n_1' = (n_2-1)N+1$, because under this situation, we could obtain $\mathbf{hf}_{1,m_r}^{(1)}n_1'] = \mathbf{hf}_{1,m_r}^{(2)}[n_2]$. For other values of n_1' ($n_1' \neq (n_2-1)N+1$), the difference between $\mathbf{hf}_{1,m_r}^{(1)}[n_1']$ and $\mathbf{hf}_{1,m_r}^{(2)}[n_2]$ is closely related to the channel selectivity and the difference between N_1 and N_2. Greater selectivity and a larger difference between N_1 and N_2 result in a larger difference in the frequency responses of the two channels, resulting in high INI.

Remark 4 : Interference is directly influenced by power distribution for different users.

We use a basic power distribution scheme to keep it simple and concentrate on the INI analysis. The power delegated to the m_r-th user (p_{m_r}) is proportional to the inverse of its path-loss $\sqrt{\alpha_{m_r}}$. The global power allocation is given by

$$\mathbf{P} = \frac{\rho}{M_t}\mathbf{A}^{-1}. \tag{2.25}$$

where ρ is a normalization factor that guarantees the power constraint $tr(\mathbf{P}) = p_{Tx}$ is satisfied, where p_{Tx} is the global power at the transmitter side. $\mathbf{P} = Diag(p_1, ..., p_{M_r})$ and $\mathbf{A} = Diag(\alpha_1, ..., \alpha_{M_r})$. Then, as user m_r's large-scale fading increases, the user's transmission power increases.

In our system, having two users with different path-loss, the generated INI power from user 1 on user 2 ($p^{(1,2)}$) and that from user 2 on user 1 ($p^{(2,1)}$) satisfy

$$\frac{p^{(1,2)}}{p^{(2,1)}} = \frac{\alpha_2}{\alpha_1}. \tag{2.26}$$

For example, if $\alpha_2 < \alpha_1$, which means that user 2 is farther away from the BS than user 1, user 1 will receive more INI as a result of user 2's improved transmitting power.

2.3 Massive MIMO-OFDM uplink : INI theoretical analysis

2.3.1 Massive MIMO-OFDM uplink system model

The uplink system is the reverse of the downlink system introduced in the previous section. We consider a single-cell massive MIMO-OFDM uplink system, where M_t single-antenna users that are using different numerologies, transmit signals to a BS equipped with M_r antennas, over a frequency-selective channel. Note that in the uplink system, M_r is significantly larger than M_t due to the fact that now the BS is the receiver. The users can be also divided into NUM groups using NUM numerologies. Other definitions such as channel response remain the same with the previous section.

We also use two users ($M_t = 2$) with two different numerologies, as shown on FIGURE 2.9. Considering the generalized synchronized scenario, similar as before, we assume that $N_1 = N \times N_2, CP_1 = N \times CP_2$, where $N = 2^i$ and i is an integer [12]. At the BS side, the received signal is separated into two branches and the data stream are then detected through linear processing by two ZF detectors $\bar{\mathbf{P}}^{(1)} \in \mathbb{C}^{2\times M_r \times N1}$ and $\bar{\mathbf{P}}^{(2)} \in \mathbb{C}^{2\times M_r \times N_2}$.

$\mathbf{P}_{n_{num}}^{(num)} = \bar{\mathbf{P}}^{(num)}(:,:,n_i) \in \mathbb{C}^{2\times M_r}, num = 1,2$ represent the ZF detection matrix for the n_{num}-th OFDM subcarrier of numerology num, which is given as follows

$$\mathbf{P}_{n_{num}}^{(num)} = \left(\mathbf{H}_{n_{num}}^{(num)H}\mathbf{H}_{n_{num}}^{(num)}\right)^{-1}\mathbf{H}_{n_{num}}^{(num)H} \tag{2.27}$$

where $\mathbf{H}_{n_{num}}^{(num)} = \bar{\mathbf{H}}^{(num)}(:,:,n_{num}) \in \mathbb{C}^{M_r\times 2}, i = 1,2$ is the frequency-domain MIMO channel of the n_{num}-th OFDM subcarrier with numerology num. Due to the ZF detection characteristics, the following equations are valid

$$\sum_{m_r=1}^{M_r} \bar{\mathbf{P}}^{(1)}(2,m_r,n_1)\mathbf{hf}_{m_r,1}^{(1)}[n_1]\mathbf{s}^{(1)}[n_1] = 0 \tag{2.28}$$

$$\sum_{m_r=1}^{M_r} \bar{\mathbf{P}}^{(2)}(1,m_r,n_2)\mathbf{hf}_{m_r,2}^{(2)}[n_2]\mathbf{s}^{(2,n)}[n_2] = 0 \tag{2.29}$$

MUI within the same numerology is eliminated by the detection matrices. In a mixed numerologies scheme, INI impacts users sharing the same band. In the next section, we will carry analysis of the INI and its corresponding radiation pattern.

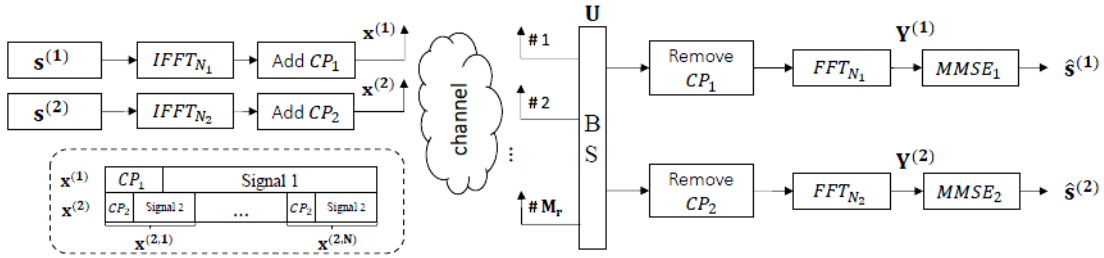


FIGURE 2.9 – System model of the massive MIMO OFDM uplink with two different numerologies : M_r receiving antennas at the BS, two sigle-antenna terminals, two blocks illustrate two different numerologies with OFDM of N_1 and N_2 subcarriers.

2.3.2 Inter-numerology interference analysis

INI from numerology 1 with N_1 to numerology 2 with $N_2 = N_1/N$

First, we consider the INI for user 2 with numerology 2, which employs a small IFFT/FFT size. We assume a noise-free transmission channel and $\mathbf{s}^{(2)} = \mathbf{0}$ to simply focus on the INI caused by user 1. $\mathbf{r}_{2,n,m_r} \in \mathbb{C}^{1\times N_2}$ represents the received n-th symbol on user 2 with numerology 2, received from the m_r-th receiving antenna, can be written as follow

$$\mathbf{y}_{2,n,m_r} = \mathbf{G}^{(1)}\mathbf{Q}_n^{(1)} Diag\left(\mathbf{hf}_{mr,1}^{(1)}\right)\left(\mathbf{s}^{(1)}\right)^T \tag{2.30}$$

where the matrix definition is the same as equation (2.12). Then, the BS uses $\bar{\mathbf{P}}^{(2)}$ for the detection of user 2's signal and the interference of each subcarrier in n-th symbol can be written as

$$\mathbf{ini}_n^{(1,2)}[n_2] = \sum_{m_r=1}^{M_r} \bar{\mathbf{P}}^{(2)}(2, m_r, n_2)\mathbf{r}_{2,n,m_r}[n_2] \tag{2.31}$$

From equation (2.31), the INI on user 2 is caused by all subcarriers from user 1.

INI from numerology 2 with N_2 to numerology 1 with $N_1 = N \times N_2$

In this part, we will analyze at the INI received by user 1 with numerology 1, which has a large IFFT/FFT size. Similarly, with $\mathbf{s}^{(1)} = \mathbf{0}$, we consider a noise-free transmission channel.

If we focus on the m_r-th receiving antenna, we can derive the theoretical expression of the obtained

signal on user 1 as

$$\mathbf{r}_{1,m_r} = \sum_{n=1}^{N} \mathbf{E}_n Diag\left(\mathbf{hf}_{m_r,2}^{(1)}\right) \mathbf{Z}\left(\mathbf{s}^{(2,n)}\right)^T \tag{2.32}$$

where the definition of matrices $\mathbf{E}_n$ and $\mathbf{Z}$ are the same as them in equation (2.19)

Then, the BS uses $\bar{\mathbf{P}}^{(1)}$ for the detection of user1's signal and the interference of each subcarrier can be written as

$$\mathbf{ini}^{(2,1)}[n_1] = \sum_{m_r=1}^{M_r} \bar{\mathbf{P}}^{(1)}(1, m_r, n_1)\mathbf{r}_{1,mr}[n_1] \tag{2.33}$$

From (2.33), we can see that $\mathbf{ini}^{(2,1)}[n_1]$ is caused by all subcarriers of user 2.

Remark 5 : Due to the channel frequency response characteristics, in a constant channel, the purposed transceiver scheme for massive MIMO-OFDM based mixed numerologies SS uplink transmissions does not produce any INI for both user 1 and user 2 (i.e. $\boldsymbol{ini}_n^{(1,2)} = \boldsymbol{0}, \quad \boldsymbol{ini}^{(2,1)} = \boldsymbol{0}$ *).*

Different sizes of FFTs produce the same result for a constant channel (i.e. the number of channel taps $D = 1$). For any $m_t = 1...M_t$ and $m_r = 1...M_r$, We can get

$$\mathbf{hf}_{m_r,m_t}^{(1)}[n_1] = \mathbf{hf}_{m_r,m_t}^{(2)}[n_2] \tag{2.34}$$

for all $n_1 = 1, ..., N_1, n_2 = 1...N_2$.

From (2.27), The ZF detection matrices are derived from the frequency response of the channel, which is constant in a constant channel, shown as

$$\bar{\mathbf{P}}^{(1)}(m_t, m_r, n_1) = \bar{\mathbf{P}}^{(2)}(m_t, m_r, n_2) \tag{2.35}$$

for all $n_1 = 1...N_1, n_2 = 1...N_2$.

Then, to show the INI on each subcarrier, we should rewrite equations (2.31) and (2.33) in detail. The INI on each subcarrier of user 1 and user 2 is expressed in equations (2.36) and (2.37), respectively. From (2.34) and (2.35), in a constant channel, elements in vector $\mathbf{hf}_{m_r,2}^{(1)}$ and vector $\mathbf{hf}_{m_r,2}^{(2)}$ are the same, which will leads to $\bar{\mathbf{P}}^{(1)}(1, m_r, n_1) = \bar{\mathbf{P}}^{(2)}(1, m_r, n_2)$, $\bar{\mathbf{P}}^{(1)}(2, m_r, n_1) = \bar{\mathbf{P}}^{(2)}(2, m_r, n_2)$ for all $n_1 = 1...N_1, n_2 = 1...N_2$. With these substitutions and equations (2.28), (2.29), it is clear that $\mathbf{ini}_n^{(1,2)} = 0$

and $\mathbf{ini}^{(2,1)} = 0$, for $n = 1...N$.

$$\mathbf{ini}_n^{(1,2)}[n_2] = \sum_{k_2=1}^{N_2} \sum_{n_1=1}^{N_1} \mathbf{G}^{(1)}(n_2, k_2)\mathbf{QQ}_n^{(1)}(k_2, n_1) \underbrace{\sum_{m_r=1}^{M_r} \bar{\mathbf{P}}^{(2)}(2, m_r, n_2)\mathbf{hf}_{m_r,1}^{(1)}[n_1]\mathbf{s}^{(1)}[n_1]}_{equals\ to\ 0\ in\ constant\ channel\ (2.28)} \tag{2.36}$$

$$\mathbf{ini}^{(2,1)}[n_1] = \sum_{n=1}^{N} \sum_{k_1=1}^{N_1} \sum_{n_2=1}^{N_2} \mathbf{E}_n(n_1, k_1)\mathbf{Z}(k_1, n_2) \underbrace{\sum_{m_r=1}^{M_r} \bar{\mathbf{P}}^{(1)}(1, m_r, n_1)\mathbf{hf}_{m_r,2}^{(1)}[k_1]\mathbf{s}^{(2,n)}[n_2]}_{equals\ to\ 0\ in\ constant\ channel\ (2.29)} \tag{2.37}$$

Remark 6 : The INIs for both user 1 and user 2 increases with channel selectivity and the difference between N_1 and N_2.

When the channel is frequency-selective, vector $\mathbf{hf}_{m_r,m_t}^{(1)} \in \mathbb{C}^{1\times N_1}$ is the interpolation of vector $\mathbf{hf}_{m_r,m_t}^{(2)} \in \mathbb{C}^{1\times N_2}$ for all $m_t = 1...M_t$ and $m_r = 1...M_r$, where $N-1$ values are added between every two points in $\mathbf{hf}_{m_r,m_t}^{(2)}$ to construct vector $\mathbf{hf}_{m_t,m_t}^{(1)}$. We have derived this characteristic as follow

$$\mathbf{hf}_{m_r,m_t}^{(1)}[(n_2-1)N+1] = \mathbf{hf}_{m_r,m_t}^{(2)}[n_2] \tag{2.38}$$

where $n_1 = 1,...N_1, n_2 = 1,...N_2$.

Equation (2.38) shows that there are only a few same-values within the two channel frequency responses for the same time-domain channel between user m_t and receiving antenna m_r. For other values (i.e. $n_1 \neq (n_2-1)N+1$), the difference between $\mathbf{hf}_{m_r,m_t}^{(1)}[n_1]$ and $\mathbf{hf}_{1,m_r}^{(2)}[n_2]$ is closely related to the channel selectivity and the difference between N_1 and N_2. Greater selectivity and a larger difference between N_1 and N_2 result in a larger difference in the frequency responses of the two channels, resulting in more interference.

Remark 7 : The INIs at user 1 and user 2 are directly affected by the power allocation for different users.

In the previous analysis, we use a simple power allocation schemes where the transmission power of the m_t-th user is equal for the sake of simplicity and to concentrate on the INI analysis. As noted before, the received signal is the summation of all the transmitted signals. As a consequence, the power of INIs produced is influenced by the unequal receiving power of each user. Greater large-scale fading in user m_t (represented by α_{m_t}) results in lower receiving power and INI power on other users. The cell's most distant user suffers the most INI.

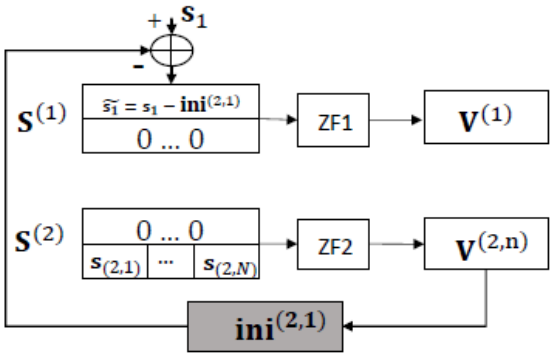


FIGURE 2.10 – INI cancellation for downlink.

In our system with two users having different path-loss (α_1 and α_2), the generated INI power from user 1 on user 2 ($p^{(1,2)}$) and that from user 2 on user 1 ($p^{(2,1)}$) satisfy

$$\frac{p^{(1,2)}}{p^{(2,1)}} = \frac{\alpha_1}{\alpha_2} \tag{2.39}$$

If $\alpha_2 < \alpha_1$, for example, it means that user 2 is further away from the BS than user 1. Then, at the BS, the obtained INI power on user 2 is higher than the INI power on user 1.

2.4 INI cancellation

In this section, we present new precoding schemes to remove the INI, thanks to the developed INI analytical expressions. The main idea is to reduce the complexity of the single-antenna users by doing all the processing at the BS. In this regard, we proposed two INI cancellation methods for both downlink and uplink massive MIMO-OFDM systems.

2.4.1 Downlink INI cancellation

Only numerologies with large IFFT/FFT sizes suffer from INI as compared to numerologies with small IFFT/FFT sizes, as seen in the previous sections. In massive MIMO systems, only the uplink transmission estimates the maximum channel state. The uplink channel estimation is used for precoding in the downlink transmission, but discrepancies between the uplink and downlink channels are adjusted by a reciprocity adjustment method [90]. In this regard, the proposed INI cancellation approach is implemented at the BS side, so the receivers are not incredibly complex.

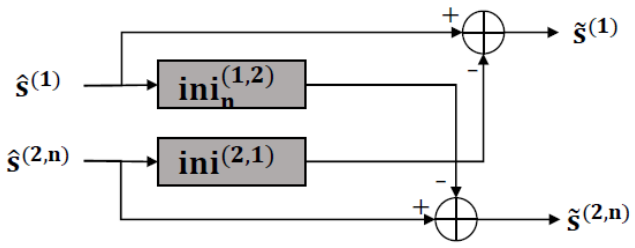


FIGURE 2.11 – INI cancellation for uplink.

At the BS side, FIGURE 2.10 illustrates the proposed INI cancellation method's scheme. The key concept is to measure the INI from numerology 2 to numerology 1 ahead of time using knowledge of the MIMO channel response and signals following user 2's precoding. Then, instead of transmitting $\mathbf{s}_1$, we transmit $\widetilde{\mathbf{s}}_1 = \mathbf{s}_1 - \mathbf{ini}^{(2,1)}$ in the transmission part. As we demonstrated in section 2.2.2, this calibration does not introduce any INI at user 2, and our proposed transmission scheme perfectly protects the transmission of user 2. The transmit power with INI cancellation is $p_{Tx} = p_u + p_{ini}$, where p_u is the power allocated to users and p_{ini} is the power allocated to INI cancellation. It is worth noting that in equation (2.25), ρ is adjusted to fulfill the constraint imposed by the total transmit power p_{Tx} when taking the p_{ini} into account. According to some simulations, the p_{ini} is insignificant compared to the power allocated to users (p_u), and it is worth noting that this would result in a minor signal-to-noise ratio (SNR) loss.

2.4.2 Uplink INI cancellation

As analyzed before, on uplink, INIs are created for each user in a frequency-selective channel. With the support of the INI analytical results presented previously, we implement an INI cancellation scheme. By performing all processing at the BS [91], one of the key ideas in massive MIMO uplink transmission is to reduce the complexity of the transmitters (users) while maintaining good transmission efficiency. The full channel state is also calculated on the uplink in massive MIMO, which can be used in ZF detection and INI cancellation. The proposed INI cancellation approach is implemented at the BS side, as shown by FIGURE 2.11, and it does not add any complexity to the users in this regard. INI is estimated after the symbol estimation, then $\tilde{\mathbf{s}}^{(1)} = \hat{\mathbf{s}}^{(1)} - \mathbf{ini}^{(2,1)}$ and $\tilde{\mathbf{s}}^{(2,n)} = \hat{\mathbf{s}}^{(2,n)} - \mathbf{ini}_n^{(1,2)}$ is done in order to suppress the INI.

Furthermore, the proposed INI cancellation scheme can be extended to any method that employs

more than two numerologies. The INI cancellation can be applied effectively for a given numerology by measuring and extracting the INIs from all other users that use different numerologies. It is worth noting, however, that the proposed INI cancellation approach is just as sensitive to imperfect CSI as ZF detections.

2.5 Simulation results

2.5.1 Parameters

The accuracy of the derived analytical expressions for INI in a mixed numerologies SS large MIMO-OFDM downlink scheme is evaluated in this section. There are two single-antenna users who use two different numerologies but share the same bandwidth in the single cell and the BS is equipped with one hundred antennas. For downlink transmission, this scenario means $M_t = 100$ and $M_r = 2$ while for uplink transmission, it means $M_t = 2$ and $M_r = 100$. A 16-QAM with Gray mapping is considered. Two pairs of numerologies are taken into consideration, with the first pair $N_1 = 1024, N_2 = 512, CP_1 = 72, CP_2 = 36$ and the second pair $N_1 = 1024, N_2 = 256, CP_1 = 72, CP_2 = 18$. The time-domain channel responses $\mathbf{ht}_{mr,mt}[d] = c_d$ have circularly symmetric Gaussian distributed values of zero mean and unit variance, where $d = 1, ..., D$. Four wireless channels are represented by a tap delay line with $D = 1, 2, 8, 18$ taps. It is worth noticing that the CP duration is often adequate to suppress the ISI induced by the channel.

In our simulations, we consider the transmission without noise first when testing the INI. The normalized mean-square error (NMSE) is known as

$$\mathrm{NMSE}_{num} = \frac{\| \hat{\mathbf{s}}_{num} - \mathbf{s}_{num} \|_2}{\| \mathbf{s}_{num} \|_2}, num = 1, 2. \tag{2.40}$$

Note that $\hat{\mathbf{s}}_{num}$ represents the estimated symbols and $\mathbf{s}_{num}$ represents original transmitted symbols.

2.5.2 Downlink simulation results

TABLE 2.1 shows the NMSE values in dB for the user of interest for the transmission method mentioned in the previous section under different conditions. It is worth mentioning that all of the simulation results in TABLE 2.1 are with equal transmitting power. TABLE 2.1 shows that the INI is of -300dB when the interfering user utilizes higher IFFT/FFT size compared to the user of

TABLE 2.1 – NMSE (dB) of user of interest with interfering numerology and different channels for massive MIMO-OFDM downlink

User of Interest	user1 ($N_1 = 1024$)		user2 ($N_2 = 512$)	user2 ($N_2 = 256$)
Interfering User	user2($N_2 = 512$)	user2($N_2 = 256$)	user1($N_1 = 1024$)	
$D = 1$	-300	-300	-300	-300
$D = 2$	-52	-48	-300	-300
$D = 8$	-42	-37	-300	-300
$D = 18$	-38	-33	-300	-300

interest. This verdict remains valid regardless the channel selectivity. It should be noted that the value of -300 dB is related to the following : (a) We used Matlab's floating point operation. (b) There are only a few users who use the service (2 users in the simulations). The degrees of freedom will decrease as the number of users increases, and the NMSE values will be affected. This NMSE can also be harmed by fixed point operation and incorrect channel estimation. These findings are in full accordance with the study conducted in section 2.2.2. When the difference between the IFFT/FFT size increases in a frequency-selective channel, however, user 1 suffers more INI from numerology 2. User 1's reliability is worse when interference is from $N_2 = 256$ than when interference is from $N_2 = 512$ for the same frequency-selective channel with $D = 2$. (4 dB difference). This discrepancy may result in poor transmission quality. Meanwhile, when comparing the performance of the same pair of N_1 and N_2 under various channel selectivity, we can see the performance decrease with the frequency selectivity of the propagation channel. Indeed, for $N_1 = 1024$ and $N_2 = 512$, There is a 10 dB degradation under channel $D = 8$ compared to $D = 2$ and a 14 dB degradation under channel $D = 18$. The degradation is 11 dB and 15 dB for $N_1 = 1024, N_2 = 256$. These findings support what we found in section 2.2.2 .

The NMSE of INI is often very small (-33 dB in the worst case) and should have a minimal effect on SNRs, as shown in TABLE 2.1. It should be noted, however, that these figures are for users with the same path-loss : $\alpha_1 = \alpha_2$ ($\alpha_1/\alpha_2 = 0$ dB).If we consider a more realistic scenario, we can use the equation (2.41) [92] for the LTE path-loss model.

$$\text{Pathloss(dB)} = 128.1 + 37.6\log_{10}(d_k) \tag{2.41}$$

where d_k is the distance from the BS in km.

TABLE 2.2 – Scenario parameters and simulation results on user 1 for massive MIMO-OFDM downlink

(a) Different scenarios

	distance user 2/BS (km)	path-loss of user 2 (dB)	distance user 1/BS (km)	path-loss of user 1 (dB)	α_1/α_2 (dB)
case 1	1	128.1	0.3	108.44	19.66
case 2	1	128.1	0.25	105.46	22.64
case 3	1	128.1	0.2	101.82	26.28

(b) NMSE values before and after INI cancellation

User of Interest	user1 ($N_1 = 1024$)			
Interfering User	user2($N_2 = 512$)		user2($N_2 = 256$)	
α_1/α_2 (dB)	original (dB)	corrected (dB)	original (dB)	corrected (dB)
19.66	-22	-300	-18	-300
22.64	-19	-300	-14	-300
26.28	-16	-300	-11	-300

Since the analysis in the previous sections and the simulation results shown in TABLE 2.1 suggest that user 2 performs the best, we will concentrate on the transmission of user 1 in the following simulations. For example, in a scenario with a cell radius of 1 km, the path-loss attenuation is shown in TABLE 2.2 (a), for various user/BS distances.

TABLE 2.2 (b) gives the performance on user 1 for various path-loss and with channel tap number $D = 8$. This performance is given in terms of NMSE when applying or/not INI cancellation at the BS side.

It can be seen that as the path-loss on user 2 grows, user 1 experiences more INI. When $N_1 = 1024, N_2 = 512$ and $D = 8$, there is a 20 dB degradation when $\alpha_1/\alpha_2 = 20$ dB compared to $\alpha_1/\alpha_2 = 0$ dB in TABLE 2.1. When $\alpha_1/\alpha_2 = 26$ dB, this degradation increases to 26 dB. These results are in agreement to what we found in section 2.2.2. In the lower sub-table, it can also be shown that, after the INI cancellation is applied at the BS, user 1's transmission performs as well as user 2, reaching -300 dB for all path-loss values.

FIGURE 2.12 shows the symbol error rate (SER) vs. E_b/N_0 using $N_1 = 1024, N_2 = 512$ and $D = 18$ to validate the INI analysis on user 1.

We can see in FIGURE 2.12 that when the path-loss of user 2 increases, user 1's output suffers

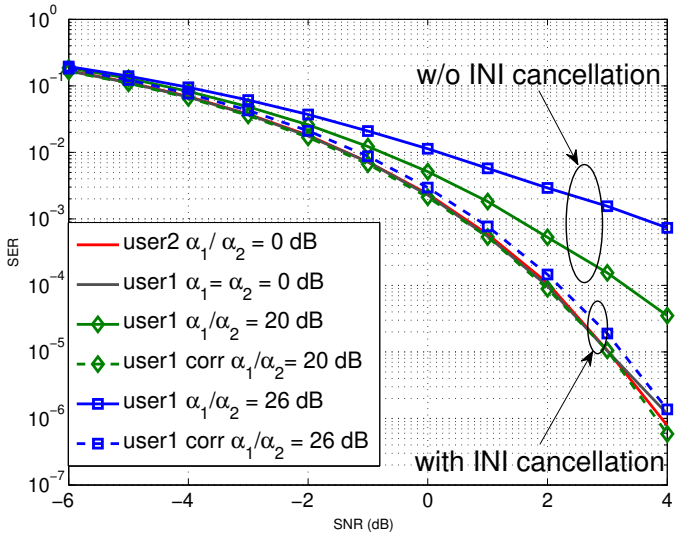


FIGURE 2.12 – SER performance with and without INI cancellation on user 1. $\alpha_1/\alpha_2 = 0, 20, 23$ and 26 dB.

significantly. We could compare $\alpha_1/\alpha_2 = 0$ dB to $\alpha_1/\alpha_2 = 26$ dB, for example. When $SNR = 4$ dB, user 1 will achieve SER $\leq 10^{-6}$ when $\alpha_1/\alpha_2 = 0$ dB, but when $\alpha_1/\alpha_2 = 26$ dB, user 1's SER is greater than 10^{-5}.

2.5.3 Uplink simulation results

TABLE 2.3 – NMSE (dB) of user of interest with interfering numerology and different channels for massive MIMO-OFDM uplink

User of Interest	user1 ($N_1 = 1024$)		user2($N_2 = 512$)	user2($N_2 = 256$)
Interfering User	user2($N_2 = 512$)	user2($N_2 = 256$)	user1($N_1 = 1024$)	
$D = 1$	-300	-300	-300	-300
$D = 2$	-56	-52	-52	-49
$D = 8$	-47	-45	-44	-40
$D = 18$	-43	-40	-39	-36

TABLE 2.3 shows the NMSE values in dB for the user of interest for the uplink system described in the previous sections under various conditions. Notice that $\alpha_1 = \alpha_2$ is used in all simulation results in TABLE 2.3. It shows that INIs are found on all users except for the flat fading channel. Furthermore, these findings are consistent with the analysis presented in section 2.3.2, demonstrating that channel

selectivity and increased N_1/N_2 ratio raise INIs. For example, under channel with $D = 8$, there is 9 dB of degradation on $NMSE_1$ compared to channel $D = 2$ and 13 dB under channel with 18 taps for the same pair of numerologies, $N_1 = 1024, N_2 = 512$. Meanwhile, when comparing performance under the same channel selectivity, different pairs of numerologies behave differently, with users suffering more INI as the IFFT/FFT size gap grows. User 1's performance is worse when the interference is from $N_2 = 256$ than when the interference is from $N_2 = 512$ over the same frequency-selective channel with $D = 2$. (4 dB difference). From table 2.3, we can see also that for a given channel selectivity (with $D > 1$), the degradation is higher when the interferer's numerology is higher compared to that of the user of interest.

Similar to our simulation of massive MIMO-OFDM downlink system, we consider some more realistic scenarios with the LTE path-loss model given by equation (2.41) [92].With a cell radius of 1 km, the different path-loss attenuations are shown in TABLE 2.4 (a), for various user/BS distances.

The impact of path-loss on system performance is shown in TABLE 2.4, with D = 8 as the number of channel taps. It can be seen that as path-loss increases, users experience more INI. As an example, when $N_1 = 1024, N_2 = 512$, there is a 21 dB degradation on $NMSE_2$ when $\alpha_1 = \alpha_2 = 20dB$ compared to $\alpha_1/\alpha_2 = 0dB$, and this degradation increases to 28 dB when $\alpha_1/\alpha_2 = 26dB$. These findings match our study in section 2.3.2. The lower sub-table also shows that, after the INI cancellation was introduced at the BS, transmission of user 1 and user 2 reached about -70 dB for all path-loss values.

Using $N1 = 1024, N2 = 512, D = 8$ and three separate pairs of locations, FIGURE 2.13 shows the symbol error rate (SER) vs. signal-to-noise ratio (SNR) of different users with/without INI cancellation.

In comparison to user 2, we can see that when the path-loss increases, user 1's performance suffers from INI significantly. When $SNR = 3$ dB, for example, user 1 could achieve $SER \leq 10^{-5}$ when $\alpha1/\alpha2 = 0$ dB, while this SER is reduced to 10^{-3} when $\alpha1/\alpha2 = -26$ dB.

It can also be seen in the figure that, regardless of path-loss, user 1's performance increases significantly after the INI is cancelled. In certain cases (i.e. $\alpha1 > \alpha2$), these results can also be obtained for user 2.

TABLE 2.4 – Scenario parameters and simulation results on user 1 for massive MIMO-OFDM uplink

(a) Different scenarios

	distance user 2/BS (km)	path-loss of user 2 (dB)	distance user 1/BS (km)	path-loss of user 1 (dB)	α_1/α_2 (dB)
case 1	1	128.1	1	128.1	0
case 2	0.3	108.44	1	128.1	19.66
case 3	0.2	101.82	1	128.1	26.28
case 4	1	128.1	0.3	108.44	-19.66
case 5	1	128.1	0.2	101.82	-26.28

(b) NMSE values before and after INI cancellation with $N_1 = 1024$, $N_2 = 512$ and $D = 8$

α_1/α_2 (dB)	$NMSE_1$ original (dB)	$NMSE_1$ corrected (dB)	$NMSE_2$ original (dB)	$NMSE_2$ corrected (dB)
0	-47	-70	-44	-69
20	-66	-71	-23	-69
26	-72	-72	-16	-70
-20	-28	-69	-65	-69
-26	-20	-69	-70	-72

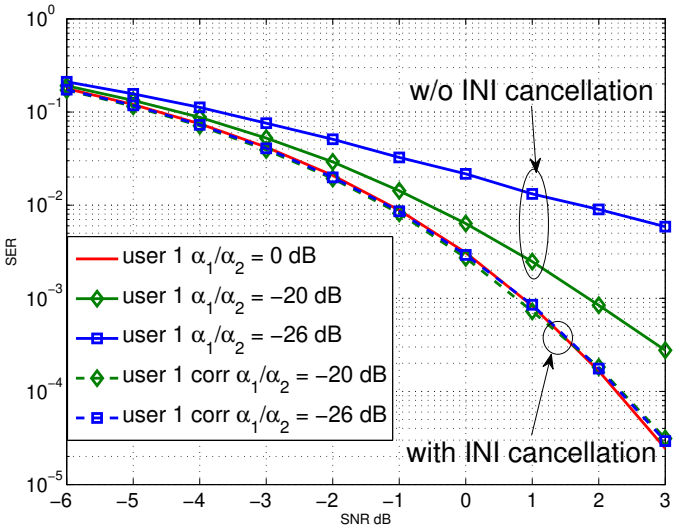


FIGURE 2.13 – SER performance before and after INI cancellation on user 1. $\alpha_1/\alpha_2 = 0, -20$ and -26 dB.

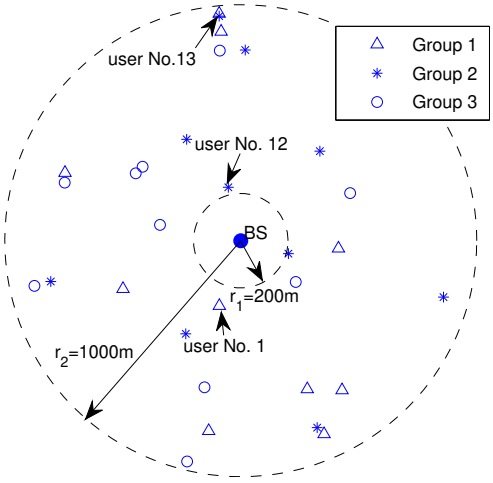


FIGURE 2.14 – User distribution in a ring where BS is in the center. The inner radius $r_1 = 200$ m and the outer radius $r_2 = 1000$ m. $M_t = 100$, $M_r = 30$.

2.5.4 Multi-user cellular system

The results of the simulations above are for a simple cell with only two different numerologies taken into account and each category using the same numerology has just one user. In reality, each category can have multiple users, and a single cell can contain several groups. We will, in this seciton, use a downlink scenario as an example. Users are evenly distributed in the ring around the BS, with inner radius $r1 = 200$ m and outer radius $r2 = 1000$ m, as shown in FIGURE 2.14 with $M_t = 100$ and $M_r = 30$. The thirty users are divided into three groups : Group 1 includes users 1 to 10, Group 2 includes users 11 to 20, and Group 3 includes users 21 to 30, using three different numerologies of IFFT/FFT sizes of 1024, 512, and 256, respectively.

The INI between different users in the cell using our proposed transceiver design is shown on FIGURE 2.15, without considering path-loss under a frequency-selective channel with $D = 18$, where the x axis represents the interfering user index, the y axis represents the user of interest (UoI), and the color represents the UoI's NMSE values. We can see that there is no INI between users in the majority of cases (the dark-blue part). The other two classes do not produce any INI on the users from group 3 (UoI No. 21 to No. 30) who use IFFT/FFT size equal to 256. Furthermore, when INI

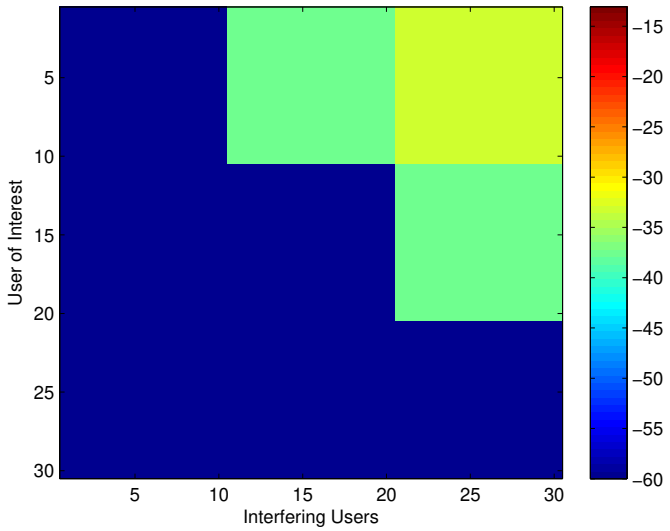


FIGURE 2.15 – NMSE values for different users without considering the path-loss. Group 1 : user No.1 to No.10 with $N_1 = 1024$. Group 2 : user No.11 to No.20 with $N_2 = 512$. Group 3 : user No.21 to No.30 with $N_3 = 256$. Channel taps $D = 18$.

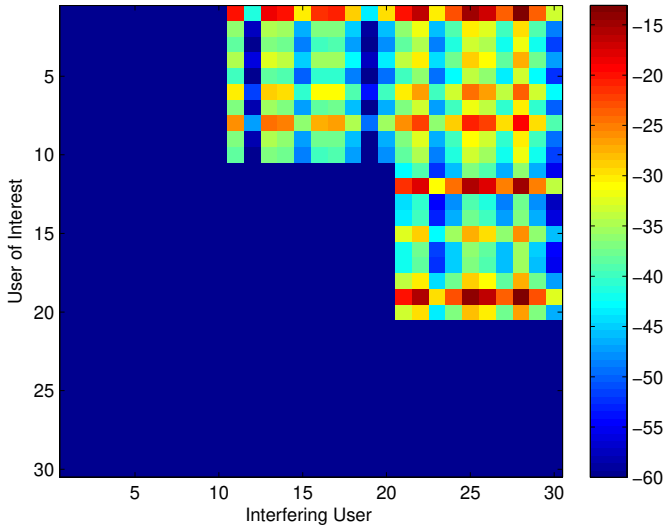


FIGURE 2.16 – NMSE values for different users considering the path-loss presented on Fig.12. Group 1 : user No.1 to No.10 with $N_1 = 1024$. Group 2 : user No.11 to No.20 with $N_2 = 512$. Group 3 : user No.21 to No.30 with $N_3 = 256$. Channel taps $D = 18$.

is created, the NMSE values are around -35 dB (as shown in TABLE 2.1), implying that there is no need to implement INI cancellation when the users have similar path-loss.

When we consider the path-loss of various users, as seen in FIGURE 2.14, the INI for each user varies. FIGURE 2.16 describes various NMSE values on various UoI belonging to various groups. As seen in this figure, some users have a major effect on others (red and orange points). We can compare the different INI impacts on UoI No. 1 from users No. 12 and No. 13, for example, where user No. 1 belongs to group 1 ($N1 = 1024$) and users No. 12 and 13 belong to group 2 ($N2 = 512$). The INI generated from user No. 12 to user No. 1 induces the NMSE value to be around -40 dB (blue point), while the INI generated from user No. 13 to user No. 1 induces the NMSE value to be around -20 dB. (red point). In this case, it's more important, for user No.1, to cancel the INI generated by user No.13. The difference in path-loss between users results in different INI impacts, as seen in FIGURE 2.14, where user No.1 and user No.12 have equal path-loss while user No.1 and user No.13 have a significant difference in path-loss. To summarize, INI cancellation is interesting, or even important, when users from two different groups have a significant difference in terms of path-loss.

2.6 Conclusion

In comparison to mixed numerologies SISO and classical MIMO systems, where users use adjacent frequency bands, in this chapter, we first introduced new transmission schemes designed for MU massive MIMO-OFDM based 5G that supports different services using different numerologies while sharing the same band. The efficiency of these mixed numerologies in massive MIMO-OFDM downlink / uplink systems was then investigated. We derived theoretical expressions of the INI for the occurring INI in the mixed numerologies scheme, which were checked by simulations. We showed that, in the massive MIMO-OFDM downlink scenario, INI is only created in frequency selective channels and only when users with large SCS communicate with users with small SCS. However, in uplink scenario, all users suffer from INI in frequency selective channels.

Furthermore, we have verified that the theoretical INI analysis agrees with simulation findings, suggesting that INI is influenced by SCS, channel selectivity, and power allocation. Last but not least, we proposed INI cancellation schemes based on the developed close-form INI expressions that can suppress the INI at the BS side without raising the users' complexity. These findings demonstrated

that massive MIMO-OFDM systems can accept mixed numerology transmissions while all users share the same band, indicating that this method can increase the spectrum efficiency as well as meet the needs of future wireless communication systems.

Chapitre 3

NL Distortion-aware MU Precoding for massive MIMO Downlink under PA nonlinearities

Contents

3.1 Introduction

The upcoming 5G and beyond wireless communication systems are expected to support a variety of services with different needs while their energy and spectral efficiencies are of crucial importance. Massive multi-user (MU) MIMO system, which was first introduced in [93], with a BS equipped with a large number of antennas, providing services for multiple single-antenna users [94] simultaneously. The number of single-antenna users is much smaller than the number of antennas in the BS, and the created multiplexing gain is shared by everyone, thereby, improving the spectral and energy efficiencies [95] [96] [97]. In addition, simple linear signal processing methods, such as MF and ZF are considered as widely used precoding techniques for massive MIMO downlinks, which have the potential to reduce transmitter power consumption, realize signal transmission and eliminate MUI [98].

However, massive MIMO precoders bring transmit signals with high PAPR, which suffer from the hardware impairments [13]. The non-linearity of the RF PA, which is unavoidable in a common transmission chain, especially when the global system energy-efficiency is high (i.e., the PA is operated with low IBO where its power efficiecy is high). brings the main hardware impairment in the system. In-band distortion and Out-Band Radiation (OBR) occur with the existence of PA in the transmission chain. As a result, signal distortion and phase rotation appear due to the in-band distortion and adjacent channel interference appears with the rise of OBR. The effect of non-linear PA on the energy-efficient design of massive MIMO is studied in [14] and the spatial characteristic of the non-linear distortion radiated from antenna arrays is shown in [15]. In order to avoid severe distortion, one method is to ensure that each PA operates in its linear region (i.e., a region where its transmission characteristics are sufficiently linear), corresonding to operate at high IBOs. However, this is not considered as a practical solution for its lack of power efficiency. It will result in huge operating expenses for large-scale BSs with a great amount of antennas.

Recently, many research works have studied new precoding techniques to reduce the PAPR of the transmitted signals in massive MU-MIMO. Low PAPR Tomlinson-Harashima precoding schemes and iterative precoding method were proposed in [99] and [100], under the system of MU-MISO downlinks, respectively. However, specific signal processing is required in these solutions on the receiver side (i.e., in the single-antenna terminals), thereby increasing the whole system complexity and reducing their attractiveness. In [101] and [102], authors proposed the optimisation algorithms to reduce the PAPR

in the MU-MIMO systems. In stead of solving the convex problem, another algorithm-based solution is proposed in [103], in which the authors use relaxation of the nonconvex constraints to obtain a convex problem that approximates the original problem. In [104], authors introduced a method, in which certain tones are preserved, aiming to reduce PAPR in large-scale MU-MIMO-OFDM systems. Another peak-clipping scheme is proposed by Prabhu et al. [105], for PAPR reduction in an OFDM-based massive MIMO system, where certain antennas at the BS side are preserved in order to compensate the added peak-clipping signals. However, these methods lack efficiency because of their preservation on antennas or tones, which will lead to spectral efficiency loss. Therefore, they are not realistic solutions for future wireless networks.

There are also some researches aiming to jointly perform MU precoding and PAPR reduction for massive MIMO-OFDM, like the papers [106] [107] [1], where algorithms EM-TGM-GAMP, PROXIINF-ADMM and MU-PP-GDm were proposed, respectively. These recent algorithms showed great performance both in MUI and OBR while satisfying PAPR conditions. However, all of them, despite they do not consider PA nonlinearities, still have a quite high computational complexity. More recently, Zayani et al. have proposed the MU-PNL-GDm algorithm performing jointly MU precoding, PAPR reduction and digital-predistortion [2]. The proposed algorithm provides good performance but its complexity is still challenging.

Motivated by these previous researches, in this chapter, we are encouraged to concentrate on the development of an efficient downlink transmission scheme taking into consideration MU precoding, PAPR reduction and PA linearization with a special attention given to the computational complexity. In this chapter, a massive MIMO precoding optimization problem considering PA nonlinearities is analysed and formulated into a simple convex optimization problem, which is then solved via gradient descent (GD) approach. The proposed iterative algorithm, referred to as MU-2P, is able to pre-compensate nonlinear distortion (NLD) caused by BS power amplifiers. It exploits the high-dimensional degrees of freedom (DoFs), provided by the large number of antennas equipped at the BS side. Moreover, it designs a precoded signal that, when passed through the PAs and the channel, is very similar to the transmitted information.

Our proposed method presents the following improvements compared with the existing methods, e.g. [107]-[108] :

- First of all, the most valuable improvement is that our method reaches the same performance

with lower computational complexity over all existing methods [107][106][108]. Indeed, the introduced precoding optimization method enables low-complexity iterative algorithm where only matrix-vector multiplications are performed avoiding large-scale pseudo-inverse matrix computation needed for ZF precoding.

- Contrary to the method proposed in [1], in which MU precoding and PAPR reduction are jointly performed through exploiting large-scale null-space channel matrix, our method exploits DoFs avoiding the high computational complexity needed for computing the null-space channel matrix as well as for performing peak-cancelling signal precoding.
- In [2], authors introduce a transmission scheme with lower complexity compared to the method proposed in [108]. Nevertheless, this scheme needs to use the PA models inside the iterative algorithm leading to a high complexity, because we have a large number of PAs. This issue is avoided in our proposed approach. Furthermore, the scheme in [2] is tested with single carrier based massive MIMO system while our proposed approach treats multi-carrier transmission scheme, as OFDM.
- The proposed method can guarantee perfect MUI cancellation without introducing any additional OBR. On the contrary, other proposed methods like [107] cannot guarantee complete MUI and OBR cancellation. Therefore, our precoder has the leading importance for future wireless networks.

Special Notations : In this chapter, we use $\mathbf{x}_m$ to present the m-th column, and $\mathbf{x}_n^t$ to present the n-th row. The k-th element in the vector $\mathbf{x}$ is denoted as $\mathbf{x}(k)$.

3.2 Massive MIMO Dwonlink system model

We consider, a massive MU MIMO-OFDM downlink scenario, where the number of antennas equipped at BS M_t is significantly larger than the number of single-antenna user M_r. Here, each antenna is equipped with a PA, as illustrated in FIGURE 3.1. The signal vector $\mathbf{s}_n \in \mathbb{C}^{M_r \times 1}$ contains the symbols associated with the n-th OFDM subcarrier for M_r users, where $n = 1, ..., N$ and it is chosen from a complex-valued constellations. In fact, OFDM systems usually specify certain preserved sub-carriers, which are used for guard-bands which are usually located at both ends of the spectrum). Therefore, the available sub-carriers are divided into two sets χ and χ^c, where the sub-carriers in the set χ are used for data transmission, and the complementary set χ^c is prepared as guard-bands. In

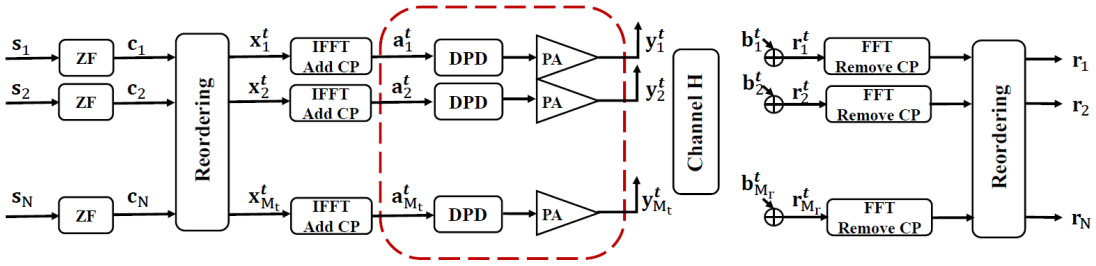


FIGURE 3.1 – MU MIMO-OFDM downlink system. The dash box represents the PA and its DPD compensation stage at each transmission chain

addition, we set $\mathbf{s}_n = \mathbf{0}_{M_r \times 1}$ for $n \in \chi^c$ so that there is no signal transmission on the guard band.

In order to suppress the MUI at the receivers, linear precoding scheme is implemented at the BS, designed to cancel the MUI completely. Thus, the transmitted signal, corresponding to the n-th subcarrier, through the M_t-th antennas, $\mathbf{c}_n \in \mathbb{C}^{M_t \times 1}$ is linearly coded as

$$\mathbf{c}_n = \frac{1}{\sqrt{\zeta_w}} \mathbf{W}_n \mathbf{s}_n, n = 1, ..., N \tag{3.1}$$

where ζ_w is the power normalization factor designed to obtain an average or instantaneous transmit power equal to P_t. $\mathbf{W}_n \in \mathbb{C}^{M_t \times M_r}$ denotes the ZF precoding matrix for n-th OFDM subcarrier.

ZF precoding scheme is widely considered [87] to completely cancel the MUI. Note that in massive MIMO, due to $M_r \ll M_t$, many precoding matrices are proposed to achieve perfect MUI elimination. Among them, the most widely used form is

$$\mathbf{W}_n = \mathbf{H}_n^H \left(\mathbf{H}_n \mathbf{H}_n^H \right)^{-1}, \quad \forall n \tag{3.2}$$

where $\mathbf{H}_n \in \mathbb{C}^{M_r \times M_t}$ is the MIMO channel matrix, in frequency domain, of the n-th OFDM subcarrier. We assume that the channel matrices are transparent and fully known by the BS side. In time division duplexing (TDD) system, channel information can be determined by using the channel reciprocity [89] [88].

After precoding, the M_t-dimension vector $\mathbf{c}_n$ is reordered to construct $\mathbf{x}_n$, according to the following mapping :

$$\left[\mathbf{x}_1^t ; \ldots ; \mathbf{x}_{M_t}^t \right] = \left[\mathbf{c}_1 \ldots \mathbf{c}_N \right]^T \tag{3.3}$$

Here, N-dimensional vector $\mathbf{x}_{m_t}$ represents the frequency domain signal to be transmitted via the m_t-th antenna. Then, IFFT is applied on $\mathbf{x}_{m_t}$ in order to obtain the time-domain signal $\mathbf{a}_{m_t}^t$, CP is then

added after to the time-domain signals at each branch. Assuming that the length of CP is larger than that of the channel multipath, then, both inter-symbol interference (ISI) and inter-carrier interference (ICI) are eliminated.

For the case of the massive MU-MIMO with non-linear PAs, the precoded symbols in (1) are fed, towards the BS antennas, through M_t parallel transmit chains with PAs. The resulting amplified symbols are :

$$\left[\mathbf{y}_1^t; \ldots; \mathbf{y}_{M_t}^t\right] = \left[p_1(\mathbf{a}_1^t); p_2(\mathbf{a}_2^t); \ldots; p_{M_t}(\mathbf{a}_{M_t}^t)\right] \tag{3.4}$$

where $p_{m_t}(.)$ denotes the nonlinear amplification operation of the m_t-th PA. In order to simplify the system and without any loss of generality, we specify the input-output relationship of the wireless channel in the frequency-domain. Then, the signals received by M_r users can be written as

$$\mathbf{r}_n = \mathbf{H}_n \mathbf{c}_n + \mathbf{b}_n \tag{3.5}$$

Where $\mathbf{b}_n$ is the added Gaussian noise, and its entry is an i.i.d circularly- symmetric distribution with zero mean and σ_b^2 variance.

Consider a typical OFDM system with N subcarriers, the instantaneous amplitude level of the time domain signal $\mathbf{a}_{m_t}^t$ shows very high peak values compared to its average value, leading to higher PAPR. The PAPR of $\mathbf{a}_{m_t}^t$ is the ratio of the highest peak power to its average power value. It is defined as :

$$\text{PAPR}\left(\mathbf{a}_{m_t}^t\right) = \frac{\max_{0\leq n\leq N-1}\left[\left|\mathbf{a}_{m_t}(n)\right|^2\right]}{\mathbb{E}\left[\left|\mathbf{a}_{m_t}(k)\right|^2\right]} = \frac{N\left\|\mathbf{a}_{m_t}^t\right\|_\infty^2}{\left\|\mathbf{a}_{m_t}^t\right\|_2^2} \tag{3.6}$$

For our MU-MIMO system, the major sources of distortions in the transmitter is the nonlinear PAs, especially when they are operated close to their saturation regions to increase the power efficiency (i.e., with low IBO). Now, we let $\mathbf{a}_{m_t}(n) = \rho(n)e^{j\phi(n)}$ be the n-th sampling point that is to be transmitted and amplified via the antenna m_t, where $\rho(n)$ and $\phi(n)$ denote, respectively, the magnitude and phase of that sample. Then, the relation between the baseband equivalent input and output signals of the PA in the m_t-th antenna branch can be written as

$$\mathbf{y}_{m_t}^t(n) = g(\alpha\rho(k))e^{j(\phi+\Psi(\alpha\rho(k)))} \tag{3.7}$$

where $g(.)$ is the amplitude-to-amplitude (AM-AM) conversion and $\Psi(.)$ is the amplitude-to-phase (AM-PM) conversion of the PA. The factor α is a multiplicative coefficient applied at the input of the PA to obtain the PA working point according to the given IBO. Here, all IBOs are given in dB related to the PA saturation point. The factor α needed to ensure a signal $\mathbf{a}_{m_t}(n)$ with a given IBO value, which is obtained through $\sqrt{P_{sat}/(10^{\frac{IBO}{10}}P_t)}$, where P_{sat} is the saturation inout power and P_t is the emitted power. The functions $g(\rho)$ and $\Psi(\rho)$ are modelled by the modified Rapp model proposed by the 3GPP for the New Radio (NR) evaluation [109]. In this modified Rapp model, which reflects closely to realistic PAs, AM/AM and AM/PM conversions can be described by :

$$g(\rho) = \frac{Gu}{\left(1 + \left|\frac{Gu}{V_{sat}}\right|^{2p}\right)^{\frac{1}{2p}}}, \quad \Psi(\rho) = \frac{Au^q}{(1 + (\frac{u}{B})^q)} \tag{3.8}$$

where G is the small signal gain, V_{sat} is the saturation level, p is the smoothness factor and A, B and q are fitting parameters. Neglecting the memory can be justified as we are primarily looking for how the basic PA nonlinear distortion are processed and tackled towards high energy-efficient massive MU-MIMO downlink. As future works, we intend to make equivalent analysis when considering memory PA effects.

3.3 Existing transmission schemes

In order to simplify the system presentation, we present the different schemes with simple block diagram as shown on FIGURE 3.2. The branches are shown with only one big arrow with a slash, representing the size.

3.3.1 Reference 1 and 2 :

As shown in FIGURE 3.2, there are two reference schemes Ref. 1 and Ref. 2. Ref. 1 is the classical ideal MU-MIMO-OFDM scheme with ZF precoding, so ideal PAs are considered without any PA nonlinearities (i.e., no PA distortion), which represents the best quality of transmission. This is the scenario shown on FIGURE 3.1 but without the dash box. On the contrary, Ref. 2 represents the worst scenario among all the schemes, with only DPD processing before each NL PA at each branch, which is the entire transmission chain shown on FIGURE 3.1.

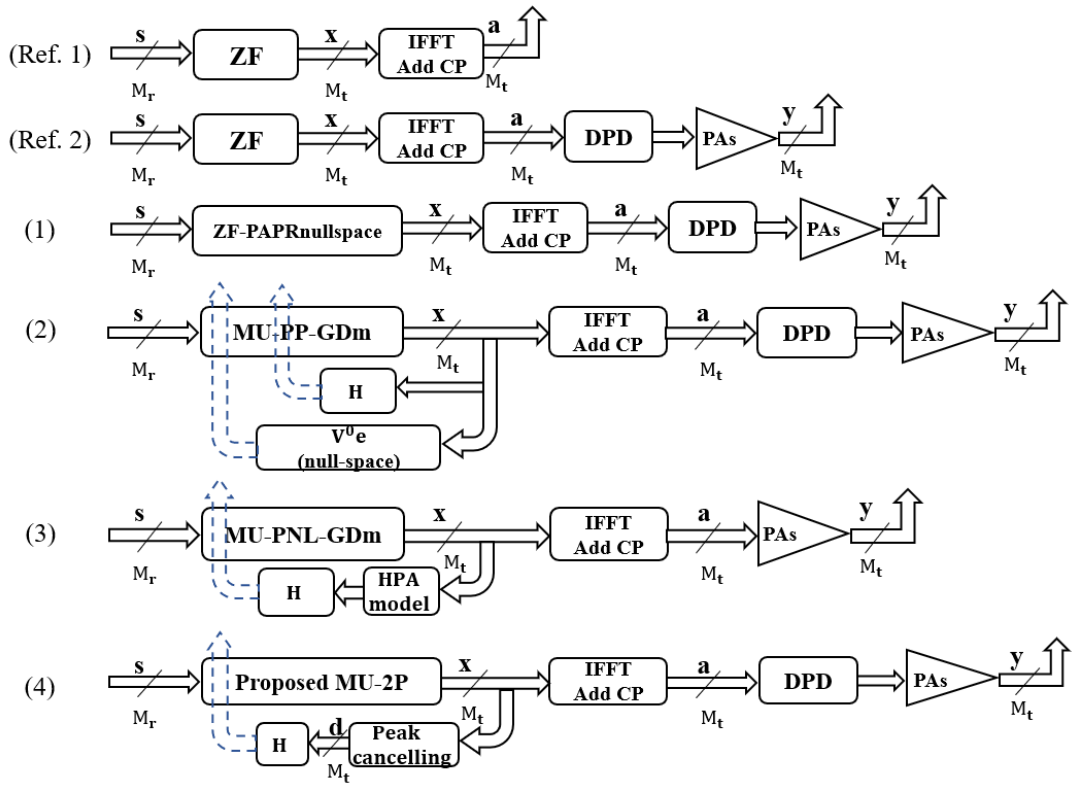


FIGURE 3.2 – Ref. 1 Classical ideal Massive MU-MIMO-OFDM system : ZF precoder is used to remove MUI and there is no PA in the transmission chain. Ref. 2 Classical Massive MU-MIMO-OFDM system under PA nonlinearities : ZF precoder is used to remove MUI and DPD is used for each PA. (1) Massive MU-MIMO-OFDM downlink scenario, where ZF-PAPRnullspace is applied. (2) The joint MU precoding and PAPR reduction solution MU-PP-GDm [1] is used to simultaneously remove the MUI and reduce the PAPR. (3) Joint MU precoding, PAPR reduction and DPD solution,so the MU-PNL-GDm [2] is applied. There is no DPD outside of the algorithm block. (4) The proposed scheme, algorithm MU-2P, which combines MU precoding and PAPR reduction with low complexity, is applied.

3.3.2 Scheme 1 : ZF precoding + PAPR reduction using null-space + DPD

In this scheme, we consider to perform separately the MU precoding, the PAPR reduction and PA linearisation. Here, we consider the ZF for MU precoding, the peak-cancelling signals projected onto the null-space of the MIMO channel for PAPR reduction and neural network (NN) based DPD for PA linearisation. Specifically, PAPR reduction is achieved by adding peak-cancelling signals $\{\mathbf{p}_{m_t}^t\}$ to precoded signals $\mathbf{x}_{m_t}^t$. The N-dimensional vector $\mathbf{p}_{m_t}^t$ is the peak-cancelling signal added to the precoded signal of the m_t-th transmit antenna, while $\mathbf{p}_n \in \mathbb{C}^{M_t \times 1}$ is the frequency-domain peak-cancelling signal collected from the M_t peak-cancelling signals at the n-th subcarrier. In order to do not damage the transmission quality, the peak-cancelling signals must be projected onto the null-space of the MIMO channel, satisfying :

$$\mathbf{H}_n \mathbf{p}_n = \mathbf{0}_{M_r \times 1}, \quad n \in \chi \tag{3.9}$$

There are two methods to compute the null-space $\mathbf{V}_n$ using the singular value decomposition (SVD) [38] or $\mathbf{V}_n = \mathbf{I}_{M_t} - \mathbf{H}_n^H \left(\mathbf{H}_n \mathbf{H}_n^H\right)^{-1} \mathbf{H}$.

Therefore :

$$\mathbf{H}_n \mathbf{V}_n = \mathbf{0}_{M_r \times M_t}, \quad \forall n \tag{3.10}$$

The key point is to design a peak cancelling signal in the frequency-domain such that $\mathbf{p}_n$ is given by

$$\mathbf{p}_n = \mathbf{V}_n \mathbf{e}_n \tag{3.11}$$

After the clipping method applied to reduce the PAPR, the original frequency-domain clipping noise related to the n-th subcarrier is $\mathbf{d}_n \in \mathbb{C}^{M_t \times 1}$. Then, the aim of this scheme is to find the solution of the following convex optimization problem :

$$\begin{array}{ll} \underset{\dot{\mathbf{e}}_n}{\text{minimize}} & G\left(\mathbf{e}_n\right) = \left\|\mathbf{V}_n \mathbf{e}_n - \mathbf{d}_n\right\|_2^2, n \in \chi \\ \text{subject to} & \dot{\mathbf{e}}_n = \mathbf{0}_{M_t \times 1}, n \in \chi^c \end{array} \tag{3.12}$$

The above formulation leads to an iterative algorithm in which the peak-cancelling signal is constrained in the null space of the related MIMO channel matrix. This algorithm, which referred to as ZF-PAPRnullspace, aims at performing separately the MUI elimination and PAPR reduction into two separate parts, where MU precoding is done by a predetermined precoder.

3.3.3 Scheme 2 : joint precoding and PAPR reduction using null-space [1]+ DPD

While ZF-PAPRnullspace + DPD proposed one solution of PAPR reduction using null-space while applying ZF precoding before. However, scheme 2 performs jointly the MU precoding and PAPR reduction. It consists in an iterative algorithm which can performs the PAPR reduction process and the MU precoding one by one. Compared with scheme 1, the new optimization method becomes :

$$\begin{aligned} &\underset{\{\mathbf{e}_n,\mathbf{x}_n\}}{\text{minimize}} J\left(\mathbf{x}_n,\mathbf{e}_n\right) = F\left(\mathbf{x}_n\right) + G\left(\mathbf{e}_n\right), \quad n \in \chi \\ &\text{subject to } \begin{cases} \mathbf{x}_n = \mathbf{0}_{M_t\times 1}, & n \in \chi^c \\ \mathbf{e}_n = \mathbf{0}_{M_t\times 1}, & n \in \chi^c \end{cases} \end{aligned} \tag{3.13}$$

where $F\left(\mathbf{x}_n\right) = \|\mathbf{H}_n\mathbf{x}_n - \mathbf{s}_n\|_2^2, n \in \chi$ is the cost function of MUI elimination, while $G\left(\mathbf{e}_n\right)$ is given by equation (3.12) to obtain an optimal PAPR. $\mathbf{x}_n$ and $\mathbf{e}_n$ are the initialization of the algorithm.

By minimizing the objective function with respect to $\mathbf{x}_n$ and $\mathbf{e}_n$, a minimization strategy can be used to solve (3.13). We adopt an alternative minimization strategy, in which we first minimize $J\left(\mathbf{x}_n,\mathbf{e}_n\right)$ with respect to $\mathbf{x}_n$ while $\mathbf{e}_n$ being frozen and then we do the same minimization with respect to $\mathbf{e}_n$ while $\mathbf{x}_n$ being frozen. In the algorithm, the $(l+1)$-th iteration can be expressed as :

$$\begin{aligned} \mathbf{x}_n^{(l+1)} &= \underset{\{\mathbf{x}_n\}}{\text{argmin}} \quad J\left(\mathbf{x}_n,\mathbf{e}_n^{(l)}\right), \qquad n \in \chi \\ \mathbf{e}_n^{(l+1)} &= \underset{\{\mathbf{e}_n\}}{\text{argmin}} \quad J\left(\mathbf{x}_n^{(l+1)},\mathbf{e}_n\right), \quad n \in \chi \end{aligned} \tag{3.14}$$

This algorithm is referred to as MU-PP-GDm.

3.3.4 Scheme 3 : Joint precoding, PAPR reduction and DPD [2]

The above two schemes use clipping approach for PAPR reduction and a separate DPD module is added at branch for PA linearisation. The key idea of scheme 3 is to jointly perform MU precoding and PA non-linearity mitigation, including PAPR reduction and DPD, using the PA model directly at each transmit antenna. By using the well-known Bussgang theorem [110], one can decompose the nonlinear signal at the PA output into a linear function of the PA input and an uncorrelated distortion term, thanks to the fact that the transmitted signal is Gaussian. The amplified signal at each branch, can be given by $\mathbf{y}_{m_t}^t = q_{m_t}\mathbf{x}_{m_t}^t + \mathbf{d}_{m_t}^t$, where q_{m_t} is the m_t-th PA complex gain and $\mathbf{d}_{m_t}^t$ stands for the added zero-mean distortion noise with variance σ_d^2. Then, we can write the amplified signal in (3.4)

as :

$$\mathbf{y}_n = \mathbf{Q}_n \mathbf{x}_n + \mathbf{d}_n \tag{3.15}$$

where $\mathbf{Q}_n = \text{diag}\left([q_1, q_2, \ldots, q_{M_t}]\right)$ is the $M_t \times M_t$ square diagonal matrix with elements $\{q_{m_t}\}$ on the diagonal.

Substituting (3.15) in (3.5), we have :

$$\mathbf{r}_n = \mathbf{H}_n \mathbf{Q}_n \mathbf{x}_n + \mathbf{H}_n \mathbf{d}_n + \mathbf{b}_n \tag{3.16}$$

By minimizing the mean square error (MSE) of intended symbols $\mathbf{s}_n$ and the received signal $\mathbf{r}_n$, the problem can be formulated as :

$$\begin{aligned} &\text{minimize } J(\mathbf{x}_n) = \|\mathbf{H}_n \mathbf{Q}_n \mathbf{x}_n + \mathbf{H}_n \mathbf{d}_n - \mathbf{s}_n\|_2^2 \\ &\text{subject to } \|\mathbf{x}_n\|_2^2 \leq P_t \end{aligned} \tag{3.17}$$

Then, scheme 3 aims to search, iteratively, for the symbols $\{\mathbf{x}_{m_t}^t\}$ which can guarantee excellent transmission quality after amplification and channel propagation. The alternating procedure, in the $(l+1)$-th iteration, can be expressed as :

$$\begin{aligned} \mathbf{x}_n^{(l+1)} &= \underset{\{\mathbf{x}_n\}}{\text{argmin}} J\left(\mathbf{x}_n^{(l)}, \mathbf{d}_n^{(l)}\right) \\ \mathbf{d}_n^{(l+1)} &= P\left(\mathbf{x}_n^{(l+1)}\right) - \mathbf{Q}_n \mathbf{x}_n^{(l+1)} \end{aligned} \tag{3.18}$$

where $P(.)$ denotes the amplification operation, which has been given in equation (3.4). By doing so, the considered optimization problem lends itself to efficient, yet flexible implementation for MU-MIMO systems by avoiding the use of DPD. The proposed algorithm is referred as MU-PNL-GDm.

The three existing schemes introduced before proposed several solutions for PA non-linearity mitigation. However, they have high computational complexity. For scheme 1, the high complexity comes from the matrix inversion of ZF precoding and the computation of the channel null-space. For scheme 2, although the ZF precoding matrix computation is removed, the rest computation of null-space remains high. What is more is that, in scheme 3, there are PA models used at each iteration, leading to a high computational complexity.

3.4 Proposed new scheme : Joint MU precoding and PAPR reduction without null-space + DPD

In this section, we will propose a new scheme which can jointly perform MU precoding and PAPR reduction without using the high-complexity required by the null-space matrix computation. The new

scheme corresponds to Scheme 4 on FIGURE 3.2. The key idea is to iteratively reduce jointly the MUI and the clipping noise $\{\mathbf{d}_{m_t}^t\}$, which is obtained by clipping the time-domain signals $\{\mathbf{a}_{m_t}^t\}$. Given the clipping threshold λ, the clipped signal $\bar{\mathbf{a}}_n$ can be obtained by :

$$\bar{\mathbf{a}}_{m_t}(n) = \begin{cases} \mathbf{a}_{m_t}(n), \text{ if } |\mathbf{a}_{m_t}(n)| < \lambda \\ \lambda e^{j\phi(n)}, \text{ if } |\mathbf{a}_{m_t}(n)| > \lambda \end{cases} \tag{3.19}$$

where $\mathbf{a}_{m_t}(n) = |\mathbf{a}_{m_t}(n)|\, e^{j\phi(n)}$ and $\phi(n)$ is the phase of $a_{m_t}(n)$. In order to obtain the best PAPR, the optimal clipping threshold λ is closely related to the mean power of the OFDM signal σ_a^2 and the ratio of the used subcarriers [1] $\frac{N}{|\chi|}$ and is given by [111] :

$$\lambda = \sigma_a \sqrt{\ln\left(\frac{N}{|\chi|}\right)} \tag{3.20}$$

The clipping noise of the m_t-th transmit antenna, in frequency-domain, is presented as $\mathbf{d}_{m_t}^t = FFT(\bar{\mathbf{a}}_{m_t}^t - \mathbf{a}_{m_t}^t)$. Then, in the frequency domain, signal on the n-th subcarrier of the transmitting signal $\tilde{\mathbf{x}}_n$ is represented as :

$$\tilde{\mathbf{x}}_n = \mathbf{x}_n + \mathbf{d}_n, \quad n \in \chi \tag{3.21}$$

Compared with scheme 3, here, $\mathbf{d}_n$ represents the clipping noise instead of PA distortion noise in equation (3.15). It is worth mentioning that the computational complexity of the clipping noise $\mathbf{d}_n$ of the equation (3.21) is less than that of the distortion noise of the scheme 3.

Algorithm 1 Joint precoding + PAPR reduction algorithm

1: **Initialize**
$\mathbf{x}_n^{(1)} = \mathbf{0}_{M_t \times 1}$,
$\Delta\mathbf{x}_n^{(0)} = \mathbf{0}_{M_t \times 1}$
set the maximal iteration number $maxIter$
set learning rate λ
2: **for** $l = 1, ..., maxIter$ **do**
3: $\Delta\mathbf{x}_n^{(l)} = 2\mu\mathbf{H}_n^H \left(\mathbf{H}_n\mathbf{x}_n^{(l)} + \mathbf{H}_n\mathbf{d}_n^{(l)} - \mathbf{s}_n\right), \forall n \in \chi$
4: $\mathbf{x}_n^{(l+1)} = \mathbf{x}_n^{(l)} - \Delta\mathbf{x}_n^{(l)}, \quad \forall n \in \chi$
5: $\mathbf{a}_{m_t}^{t(l+1)} = IFFT\left(\mathbf{x}_{m_t}^{t(l+1)}\right), \quad \forall m_t = 1 \dots M_t$
6: $\mathbf{d}_{m_t}^{t(l+1)} = FFT\left(\bar{\mathbf{a}}_{m_t}^{t(l+1)} - \mathbf{a}_{m_t}^{t(l+1)}\right)$
7: **end for**
8: **return** $\tilde{\mathbf{x}}_n = \mathbf{x}_n^{(\text{maxlter}+1)}$

In order to ensure that the transmitted OFDM signal has ideal spectral characteristics, the inactive OFDM subcarriers (indexed by χ^c) must meet the following shaping constraints :

$$\tilde{\mathbf{x}}_n = \mathbf{0}_{M_t \times 1}, \quad n \in \chi^c \tag{3.22}$$

Therefore, the total energy of MUI is determined by $\|\mathbf{H}_n\tilde{\mathbf{x}}_n - \mathbf{s}_n\|_2^2$. By minimizing the MSE of the expected symbol $\mathbf{s}_n$ through the channel $\mathbf{H}$, the problem can be then formulated as follows

$$\begin{aligned} &\text{minimize } J(\mathbf{x}_n) = \|\mathbf{H}_n\mathbf{x}_n + \mathbf{H}_n\mathbf{d}_n - \mathbf{s}_n\|_2^2, n \in \chi \\ &\text{subject to } \mathbf{x}_n = \mathbf{0}_{M_t\times 1}, \quad n \in \chi^c \end{aligned} \tag{3.23}$$

In order to solve (3.23), a minimization strategy is used, in which the objective function is iteratively minimized. Note that the clipping noise used in the $(l+1)$-th iteration is calculated from the previous l-th iteration. It can be expressed as :

$$\mathbf{x}_n^{(l+1)} = \underset{\{\mathbf{x}_n\}}{\text{argmin}} J\left(\mathbf{x}_n, \mathbf{d}_n^{(l)}\right) \tag{3.24}$$

Here, we use steepest GD method to search the direction at each iteration, where iteration $l+1$ is determined by the negative gradient of the cost function J at iteration l, which is given by :

$$\nabla_x J\left(\mathbf{x}_n^{(l)}, \mathbf{d}_n^{(l)}\right) = 2\mu\mathbf{H}_n^H\left(\mathbf{H}_n\mathbf{x}_n^{(l)} + \mathbf{H}_n\mathbf{d}_n^{(l)} - \mathbf{s}_n\right) \tag{3.25}$$

where μ is the step size. The new algorithm referred to as MU-2P.

The details of the proposed algorithm is summarized in **Algorithm 1**

3.5 Complexity analysis

TABLE 3.1 – Complexity comparison

scheme	scheme 1	scheme 2	scheme 3	scheme 4
ZF computation	$N(2M_tM_r^2 + M_r^3)$	-	-	-
ZF implementation	$N(M_tM_r)$	-	-	-
Null-space computation	$2M_t^2M_r^2 + M_rM_t^2 + M_r^3$	$2M_t^2M_r^2 + M_rM_t^2 + M_r^3$	-	-
Per iterarion	$2\lvert\chi\rvert M_t^2 + 2M_tN\log(N)$	$2\lvert\chi\rvert M_tM_r + 2\lvert\chi\rvert M_t^2 + 2M_tN\log(N)$	$3\lvert\chi\rvert M_tM_r + 2\lvert\chi\rvert M_t^2 + 2M_tN\log(N)$	$3\lvert\chi\rvert M_tM_r + 2M_tN\log(N)$
DPD	$13NM_t$	$13NM_t$	-	$13NM_t$

We adopt the number of complex multiplications as a complexity measure to compare the computational complexity of the proposed scheme with that of the existing scheme 1, 2 and 3. For the

proposed algorithm MU-2P, the computational cost is dominated by the gradient $\Delta \mathbf{x}_n^{(l)} J$ (Step 3) and the N-point IFFT/FFT (Step 5 and 6). The gradient in step 3 involves $\mathcal{O}\left(3|\chi| M_t M_r\right)$ complex multiplications. The IFFT/FFT costs $\mathcal{O}\left(M_t N \log(N)\right)$ complex multiplications. Therefore, the proposed MU-2P requires a total of $\mathcal{O}\left(3|\chi| M_t M_r + 2 M_t N \log(N)\right)$ complex multiplications at each iteration.

For ZF-PAPRnullspace, it needs $\mathcal{O}\left(N(2M_t M_r^2 + M_r^3)\right)$, $\mathcal{O}\left(N(M_t M_r)\right.$ and $\mathcal{O}\left(2M_t^2 M_r^2 + M_r M_t^2 + M_r^3\right)$ for ZF procoding matrix computation, ZF implementation and null-space computation. This leads to $\mathcal{O}\left(2|\chi| M_t^2 + 2 M_t N \log(N)\right)$ multiplications per iteration.

MU-PP-GDm needs $\mathcal{O}\left(2M_t^2 M_r^2 + M_r M_t^2 + M_r^3\right)$ for the null-space computation as initialization step and $\mathcal{O}\left(2|\chi| M_t M_r + 2|\chi| M_t^2 + 2 M_t N \log(N)\right)$ multiplications for each iteration. The complexity of MU-PNL-GDm is $\mathcal{O}\left(3|\chi| M_t M_r + 2|\chi| M_t^2 + 2 M_t N \log(N)\right)$ per iteration.

The complexities of different schemes are summarized in TABLE 3.1. For scheme 1, scheme 2 and the proposed scheme 4, a dedicated DPD block per antenna branch is used. The DPD is based on the well-known multi-layer perception (MLP) NN model. The NN predistorter has two inputs and two outputs, which represent the I and Q of, respectively, the input and output signals. It has one hidden layer with 10 neurons. A sigmoid activation function is used for hidden neurons while a linear one is used for output neurons. Then, the complexity of the NN DPD is easy to compute, it corresponds to the number of real multiplications to process one sample. The complexity of the NN DPD is the same for all the three algorithms.

Complexity comparison between different schemes will be given in the next section under certain system configurations.

3.6 Simulation results

3.6.1 System parameters

In this section, we evaluate the performance of the proposed scheme for joint MU precoding and PAPR reduction. All the simulation are conducted in a massive MIMO-OFDM downlink system. The BS is equipped with $M_t = 100$ antennas serving $M_r = 10$ single-antenna users simultaneously . We consider OFDM with $N = 512$ subcarriers, in which $|\chi| = 128$ subcarriers are used for data transmission. 16-quadrature amplitude modulation (QAM) is considered with Gray mapping. Complementary cumulative distribution function (CCDF), which indicates the probability that the PAPR

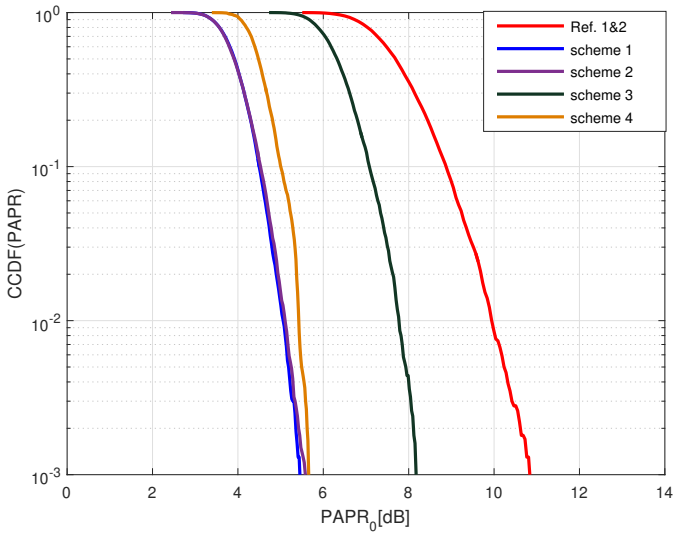


FIGURE 3.3 – PAPR performance. $M_t = 100, M_r = 10$

of the estimated signal exceeds a specified threshold $PAPR_0$, is used to evaluate the PAPR degradation performance, .

The wireless channel is set to be frequency-selective. It is modelled by a tap delay line with $D = 8$ taps. The channel response matrices in time domain, $\mathbf{H}_t, t = 1, ..., D$, consists of i.i.d. circularly symmetric distribution entries with zero mean and unit variance. The equivalent frequency-domain response $\mathbf{H}_n, n = 1, ..., N$ is obtained by FFT.

Multiple operators in a single cell are also considered as a scenario. As shown on FIGURE 3.6 (a), in the same cell, two Operators are serving two groups of users with adjacent channels in the frequency-domain. Under this scenario, inter-operator interference occurs, for example, between user m for operator No. 1 and user 1 for operator No. 2 due to spectral regrowth caused by nonlinear amplification.

We compare our proposed scheme with the existing three schemes and the reference schemes. In each algorithm, the learning rate μ are automatically adjusted to provide the best convergence rate to the optimal PAPR. For each scheme, the value of $maxIter$ is chosen to reach the MUI of -40 dB in order to create the fair comparisons between all the algorithms.

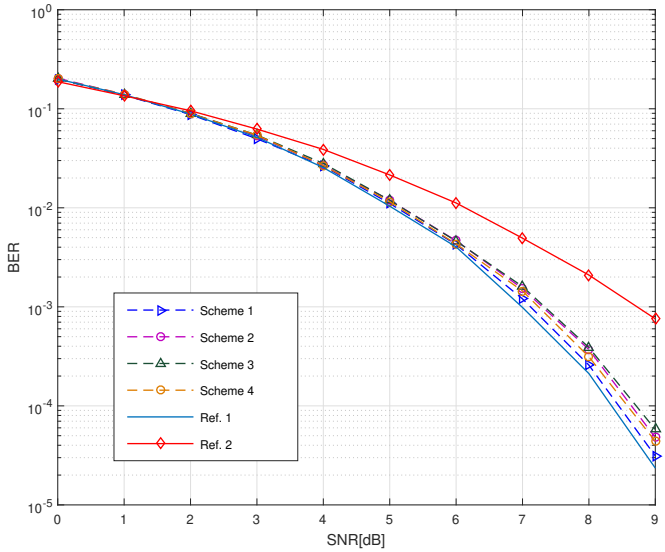


FIGURE 3.4 – BER performance. $M_t = 100, M_r = 10$, IBO = 3dB

3.6.2 Performance results

In order to evaluate the PAPR reduction efficiency of our proposed scheme, FIGURE 3.3 shows the CCDF of the PAPR of different methods. We recall that the schemes are shown in detail on FIGURE 3.2, where Ref.1 represents the best scenario with the perfect PA and Ref.2 represents the worst scenario without any PAPR reduction. Scheme 1, 2 and 3 refer to the existing schemes introduced in section 3.3 and scheme 4 is the proposed scheme presented in section 3.4. Note that PAPRs of all M_t antennas are taken into account in order to compute the empirical CCDF. The number of simulations is 1000. From FIGURE 3.3, we can see that, for the same level of MUI results, the PAPR achieves a gain of 5.4 dB, 5.2 dB, 2.6 dB and 5.1 dB, when, respectively, scheme 1, scheme 2, scheme 3, scheme 4 are considered (at CCDF of 10^{-3}).

We plot the bit error rate (BER) vs. SNR of all the schemes in FIGURE 3.4, where the SNR is defined as $SNR = \mathbb{E}\{\|\mathbf{x}_n\|_2^2\}/N_0$ and the PA is operated at IBO of 3 dB. All schemes reduce the in-band distortion compared with ZF scheme without PAPR reduction. We can also observe that there are little SNR performance (less than 0.5 dB) losses between all the schemes compared with ideal Ref. 1 at the level of 10^{-4} . This performance difference is primarily due to the fact that we add different power on the transmitter side with different schemes.

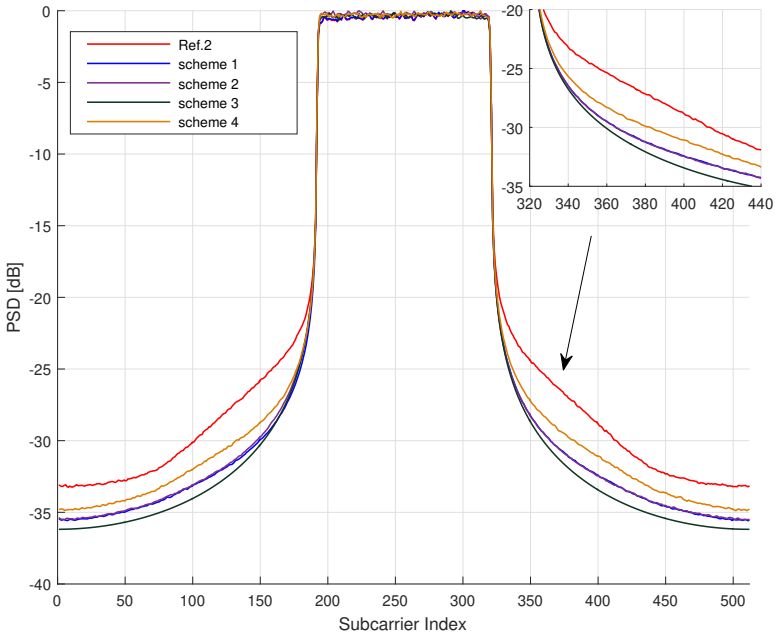
FIGURE 3.5 (a) and FIGURE 3.5 (b) show the PSD at the receiver side for the different schemes at IBO = 0 dB and IBO = 3 dB. The PSD spectrums are the mean of all M_r receiving antennas. Note that for the ZF scheme, the NN based DPD is also applied on each branch before the PA. From these results we can observe, first, when PAs are operating at IBO = 0 dB, scheme 3 has the best performance while all the schemes reduce the OOB radiation. Second, when PAs are operating at IBO = 3 dB, all schemes perform at nearly the same level, with a gain of 3 dB compared with reference 2.

As shown in [15], the uncorrelated distortion can only be omnidirectional if the number of directions is greater than the number of dimensions, i.e. when $(M_r^3 + M_r^2)/2 > M_t$. With a $M_t = 100$ MU-MIMO system, the PA distortion becomes omnidirectional at $M_r \geq 6$ users. Based on this knowledge, it is important to see if our proposed scheme have impact on other users using adjacent frequency bands. FIGURE 3.6 (b) shows the PSD at the receiver side of secondary users (dash curves) which are considered to be served by another operator, coexist in the same area as the users of interest (UoI) (solid curves). In order to better compare the OOB radiation, we shift the PSD of the two users. The in-band difference is due to the antenna gain of the massive MIMO system. We can observe from FIGURE 3.6 (b), that, in terms of OOB radiation, secondary users, as shown in dash curves, do not bring negative impact on the adjacent frequency bands, no matter in scheme Ref. 2 or our proposed scheme 4. Furthermore, our proposed scheme can reduce the OOB radiation compared with Ref.2.

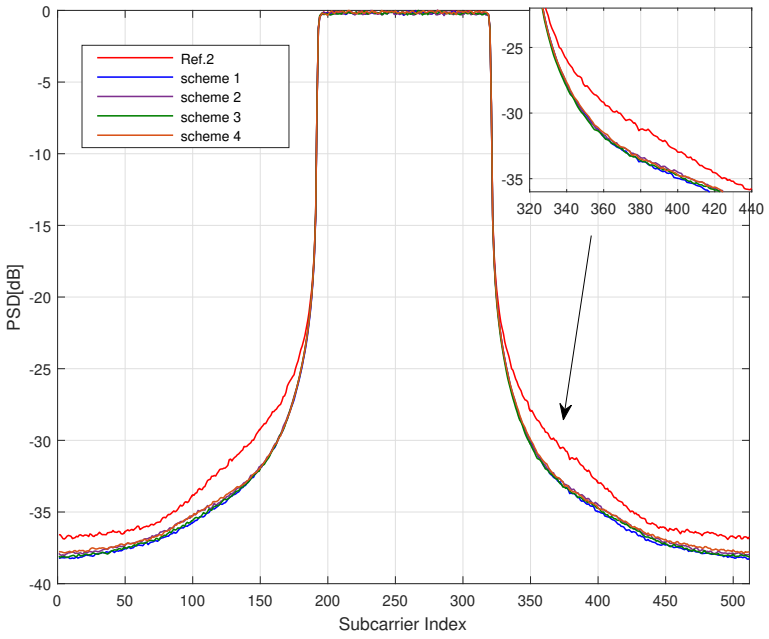
3.6.3 Complexity comparison

According to the given closed-form expressions in section 3.3 and section 3.4 as well as the configuration described in section 3.6.1, it is possible to compute the numerical complexity of all the schemes. We consider to compute the complexity required by each scheme through the entire transmission system to achieve a MUI at receiver side of -40dB, i.e. including all the initialization stages, the iterations and the NN DPDs. In order to achieve a MUI of -40 dB, the ZF-PAPRnullspace,MU-PP-GDm, MU-PNL-GDm and the proposed MU-2P require 35, 32,14 and 44 iterations, respectively.

First, we examine the efficiency of our proposed scheme with $M_t = 100, M_r = 10, N = 512$ and $\mid \chi \mid = 128$ & 256. TABLE II shows the numerical results of the complexity of different schemes. From this table we can observe, first, when the number of active subcarriers is 128, the complexity of the proposed scheme 4 is substantially less than the other schemes. When all the schemes achieve the MUI performance of -40 dB, the complexity of the proposed scheme 4 requires 36.86 %, 37.44 %

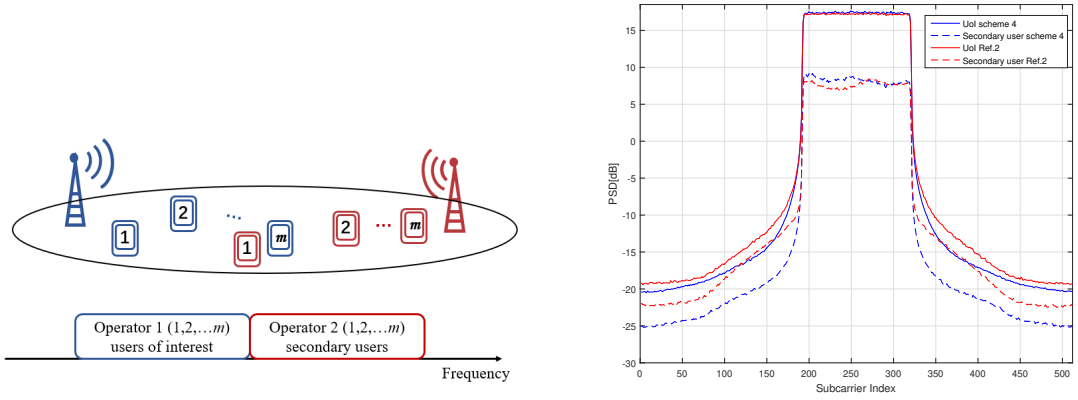


(a) PSD performance after PA at receiver side. IBO =0dB



(b) PSD performance after PA at receiver side. IBO =3dB

FIGURE 3.5 – PSD performance



(a) System of two operation systems in one single cell (b) PSD performance at receiver side for UoI and secondary users at IBO = 3dB

FIGURE 3.6 – Inter-operation interference

TABLE 3.2 – Complexity comparison to achieve MUI=-40 dB for each scheme with $M_t = 100, M_r = 10, N = 512$ and $|\chi| = 128$ & 256

scheme	$\|\chi\| = 128$	$\|\chi\| = 256$
scheme 1	1.2389×10^8	2.1349×10^8
scheme 2	1.2197^8	2.0133×10^8
scheme 3	5.0159^7	9.1375×10^7
scheme 4	4.5669×10^7	6.2565×10^7
scheme 4/scheme 1	36.86 %	29.31 %
scheme 4/scheme 2	37.44 %	29.5 %
scheme 4/scheme 3	91.05 %	68.47 %

and 91.05 % of the complexity needed by scheme 1, scheme 2 and scheme 3. What is more, when we active more subcarriers, i.e. increase the spectrum efficiency, the proposed scheme 4 has much lower complexity compared with other schemes. From TABLE 3.2 we can find the percentages drop to 29.31 %, 29.5 % and 68.47 %, respectively. This is due to the fact that ZF-PAPRnullspace and MU-PP-GDm have the initialization of null-space of ZF matrix computation and for MU-PNL-GDm, there is PA implementation in each iteration.

We move now to the complexity comparison considering channel variations. Since ZF precoding matrix computation, null-space computation, the iterative algorithms and NN DPD are all sensitive to the channel change, the channel variation condition plays an important role in the total complexity. FIGURE 3.7 shows the number of complex multiplication versus different number of τ with $M_t = 100$,

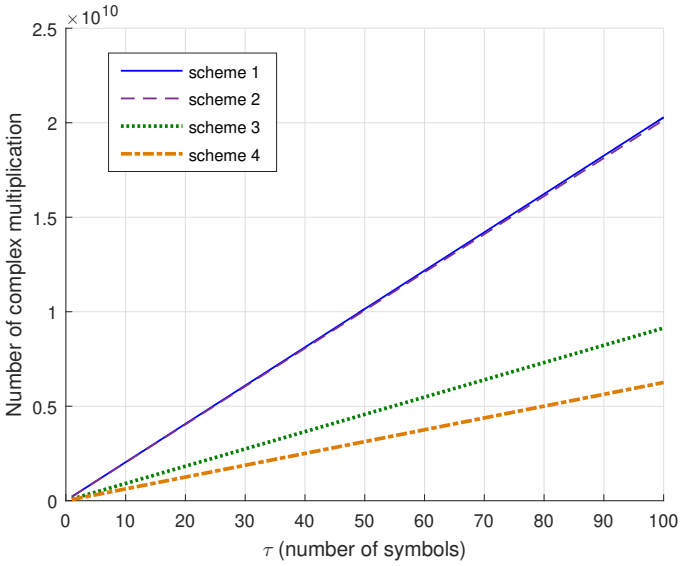


FIGURE 3.7 – Performance complexity vs.τ

$M_r = 10$ and $\mid \chi \mid = 256$, where τ represents the number of OFDM symbols within which the channel stays the same. Similar to the previous figure, our proposed scheme has the best performance, in terms of complexity, compared with all the other schemes while scheme 1 and scheme 2 are the most sensitive among all.

All the comparison and analysis in complexity illustrate the improvement of our proposed scheme compared to the other methods.

3.7 Conclusion

In this chapter, first, we presented some pertinent existing solutions for massive MU-MIMO under PA nonlinearities. Then, we introduced a low complexity downlink transmission scheme for an OFDM based massive MU-MIMO downlink scenario under PA nonlinearities. The key idea is to take advantage of the DoFs offered by massive MIMO to reduce the PAPR of transmitted signals in massive MU-MIMO, while maintaining a satisfying transmission quality and spectrum purity. Here, we emphasize the computational complexity enabling efficient deployment of real-time massive MIMO. The proposed joint MU precoding and PAPR reduction method, formulated as a simple convex optimization problem, is designed with a steepest GD-based algorithm. The simulation results showed that the proposed transmitting scheme allows a satisfying performance while it requires lower computational complexity

compared with the existing solutions.

Chapitre 4

End-to-end learning based massive MU-MIMO Downlink via deep autoprecoder

Contents

4.1 Introduction

In this chapter, we propose a new efficient autoprecoder (AP) based deep learning approach for massive MU-MIMO downlink systems in which the BS is equipped with a large number of antennas with energy-efficient PAs and serves multiple user terminals. The proposed method, (called AP-mMIMO) aims at compensating the severe nonlinear PA distortions. Unlike the other proposed solutions introduced in the previous chapter, the AP-mMIMO requires lower computational complexity.

As introduced in the previous chapters, massive MIMO technology is emerging as a core feature in the future wireless communication systems. In the massive MIMO downlinks systems, linear signal processing methods, such as MF and ZF are considered as widely used precoding techniques to reduce transmitter power consumption, to realize signal transmission and to eliminate MUI [112]. These linear precoders, on the other hand, carry a high PAPR to transmit signals [13]. As a result of the high PAPR, the main hardware distortion is the non-linearity of the RF PA, which is inevitable in a typical transmission chain. When PA is present in the transmission chain, signal distortion and phase rotation emerge, resulting in a transmission performance degradation.

In chapter 3, we introduced some algorithm-based solutions to compensate the PA distortion and improve the system performance [1] [2]. Also, we proposed the new algorithm MU-2P, which jointly perform the linear precoding and PAPR reduction and has less computational complexity while maintaining satisfactory system performance.

In the other hand, machine learning (ML)/deep learning (DL) algorithms have proven success in performing the real-time optimization tasks required for massive MU-MIMO. As most signal processing algorithms in communications systems have strong foundations in statistics and information theory, DL has the potential to outperform conventional physical layer algorithms. Deep neural networks (NNs), and in particular the autoencoder idea, have been proposed for end-to-end learning of communications systems [113]. Unlike traditional communications systems, this configuration allows for combined optimization of the transmitter and receiver for different channel model, rather than being limited to component-wise optimization.

What is more, in a practical system includes PA nonlinearities, as we are focusing on, models can only be estimated approximately. As a result, a DL-based communications system may be more adequate to compensate for such imperfections.

Sohrabi et al. [16] proposed an autoencoder to enhance the performance of one-bit massive MU-MIMO systems : It has been shown to provide satisfactory performance, it is limited to multicasting communications and its computational complexity is still challenging due to the large size of the deployed neural networks (NN). It's worth noting that the concept of a deep NN-based autoencoder has been extensively researched for SISO systems and has been demonstrated to cope with hardware imperfections. [114] [115]. In [116], authors proposed, in MU-MIMO system, a way to adjust the algorithm parameters of the nonlinear biConvex 1-bit PrecOding (C2PO) algorithm developed in [117] using NN optimization tools. However, the exact channel knowledge is needed at the BS side. To the best of our knowledge, no previous work has looked into the use of DL to deal with PA non-linearities in massive MU-MIMO downlink systems over varying fading channels.

Regarding the issues discussed above and the motivations behind this work, in this chapter, our main contributions are summarized as follows :

- We investigate the use of DL in massive MU-MIMO downlink systems with varying fading channels to deal with PA non-linearities. The important concept is to model the transmit precoder, PAs, channel, and receiver as one deep NN, i.e. as an end-to-end reconstruction task with only the transmit precoder and receive decoders are trainable.
- Unlike previous work [16][114][115], our new AP-mMIMO is suitable for varying channel scenarios, requiring less computational complexity for adaptability. We propose a two-stage precoding scheme for this : (1) an off-line trained NN-precoder and a NN-decoder (see FIGURE 4.1) that fits any channel, and (2) a channel-dependent linear precoder that can be designed using the widely used ZF precoding scheme.
- We use matrix polynomials (MP)-based form for data precoding to reduce the significant computational complexity of the large-scale matrix inversion needed for ZF precoding [118]. Numerical simulations suggest that this latter can match the performance of the ZF precoder. It's worth noting that, unlike [16], the user's power allocation in our proposed AP-mMIMO is straightforward.

4.2 System model

We consider, in this chapter, a single-cell massive MIMO downlink system, where M_r single-antenna users are served simultaneously by the BS equipped with M_t antennas, over a fading channel

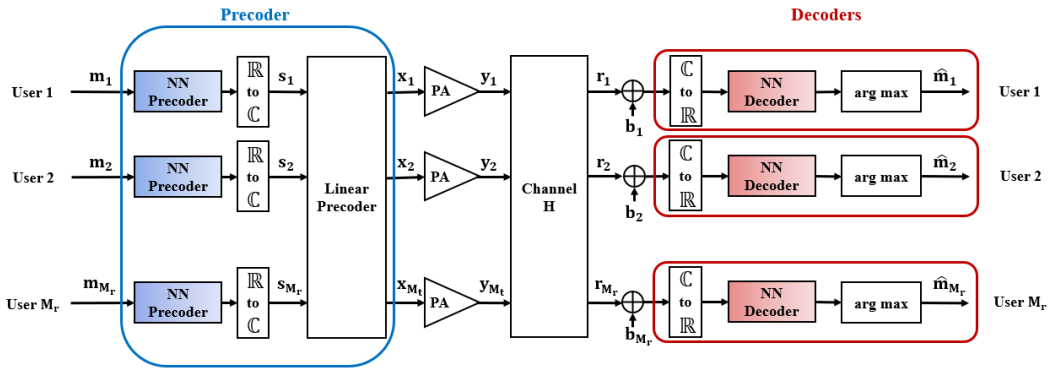


FIGURE 4.1 – Proposed AP-mMIMO structure : Generalization phase.

$\mathbf{H} \in \mathbb{C}^{M_r \times M_t}$ which has random complex Gaussian entries ($M_t \gg M_r$). The symbols for the M_r users are included in the $\mathbb{C}^{M_r \times 1}$ signal vector $\mathbf{s}$, where s_{m_r} is chosen from a M-quadrature amplitude modulation (QAM) constellation. The BS uses a linear precoding stage, transforming the vector $\mathbf{s}$ into a M_t-dimensional vector by using a $\mathbb{C}^{M_t \times M_r}$ matrix $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{M_r}]$, using the CSI. This can be written as

$$\mathbf{x} = \frac{1}{\sqrt{\varsigma_W}} \mathbf{W}\mathbf{s} \tag{4.1}$$

where $\frac{1}{\sqrt{\varsigma_W}}$ is the power normalization factor. Here, the precoding matrix can be obtained by using the ZF technique [119] or by the matrix polynomials technique [118], which are calculated by $\mathbf{W} = \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H\right)^{-1}$ and $\mathbf{W} = \mathbf{H}^H \sum_{j=0}^{\mathcal{J}} \mu_j \left(\mathbf{H}\mathbf{H}^H\right)^j$, respectively, where μ is a vector containing the real-valued coefficients of the MP precoder. The coefficients in μ can be optimized theoretically and they do not depend on the instantaneous channel estimates. Note that the Horner's implementation rule [118] should be adopted for seeking low complexity. Interested readers are referred to [118] for more details concerning the μ coefficients optimization and the Horner's rule implementation.

For the case of massive MIMO with non-linear PAs, the precoded symbols in equation (4.1) are fed, towards the BS antennas, through M_t parallel transmit chains with PAs (see FIGURE 4.1). The resulting amplified symbols are presented by $\mathbf{y}$. The amplification and channel propagation are illustrated in equations (3.4) and (3.5).

The principal sources of distortions in the transmitter for our massive MU-MIMO downlink system, as in the previous chapter, are the nonlinear PAs, especially when they are operated close to their

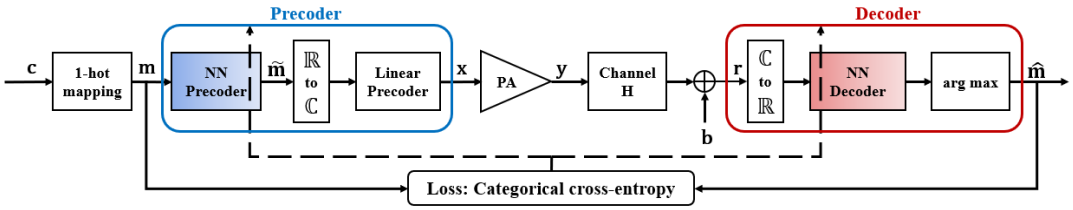


FIGURE 4.2 – Structure of the precoder and decoder. Two NNs are trained at the transmitter side and the receiver side, in which the cross entropy is used as loss function.

saturation regions to maximize their power-efficiency. Note that the nonlinear PAs are also modeled in this chapter using the 3GPP's modified Rapp model for the New Radio (NR) evaluation [109]. Its corresponding output signal, AM/AM conversion and AM/PM conversion are given in equations (3.10) and (3.11).

4.3 Proposed autoprecoder structure and learning solution

In this section, we present our autoprecoder based end-to-end mMIMO downlink communication structure and the training/generalization details.

4.3.1 Autoprecoder structure

The structure of the proposed AP-mMIMO is shown in FIGURE 4.2. Most DL libraries only offer real-value operations, as a result, we should map the complex vectors to real ones. Since the transmitted signal is complex, in the block diagram, there is a $\mathbb{R}$-to-$\mathbb{C}$ block and a $\mathbb{C}$-to-$\mathbb{R}$ block, representing the real-to-complex and complex-to-real mappings, respectively.

The symbol $\mathbf{c} \in \{1, ..., M\}$ represents the index of the intended message for a specific user, and $\mathbf{m}$ indicates its one-hot representation, i.e., an M-dimensional vector with one element for the m-th element and zero for the remaining elements. This vector is passed into the transmit NN, which has numerous dense nonlinear layers and one last linear layer. A dense layer is a matrix multiplication with learnable weights followed by a non-linear activation function in this scenario. Two linear neurons representing the outputs I and Q of the base-band complex signal belonging to one user should be present in the linear layer.

The 2-dimensional real-valued signal at the output of the NN-precoder $\widetilde{\mathbf{m}}$ in this model can be

written as

$$\widetilde{\mathbf{m}} = \mathbf{K}_{\mathrm{T}} \cdot \sigma_{\mathrm{T}-1} \left(\cdots \mathbf{K}_2 \cdot \sigma_1 \left(\mathbf{K}_1 \cdot \mathbf{1}_m + \mathbf{n}_1 \right) + \cdots \mathbf{n}_{\mathrm{T}-1} \right) + \mathbf{n}_{\mathrm{T}} \tag{4.2}$$

where T is the number of layers (i.e., $\mathrm{T}-1$ hidden layers and one output layer), σ_t is the activation function for the t-th layer. $\mathbf{\Theta}_{\mathrm{T}} = \{\mathbf{K}_{\mathrm{t}}, \mathbf{n}_{\mathrm{t}}\}_{\mathrm{t}=1}^{\mathrm{T}}$ represents the set of the NN parameters (matrices and bias). The dimensions of the NN parameters are

$$\begin{aligned} \dim\left(\mathbf{K}_{\mathrm{t}}\right) &= \begin{cases} \ell_{\mathrm{t}} \times M, & t = 1 \\ \ell_{\mathrm{t}} \times \ell_{\mathrm{t}-1}, & t = 2, \ldots, T-1, \\ 2 \times \ell_{\mathrm{t}-1}, & t = T, \end{cases} \\ \dim\left(\mathbf{n}_{\mathrm{t}}\right) &= \begin{cases} \ell_{\mathrm{t}} \times 1, & t = 1, \ldots, T-1, \\ 2 \times 1, & t = T \end{cases} \end{aligned} \tag{4.3}$$

where ℓ_{t} is the number of neurons in the t-th hidden layer.

During the learning phase (as illustrated in FIGURE 4.2), only one NN-precoder is trained, and it is used for the M_r users at test time. We generate the $2M_r$ real-valued signal corresponding to the M_r users (as shown in FIGURE 4.1) during the generalization phase, which serves as the initial stage of precoding. Then, a traditional linear precoder, such as a ZF one, is used in the second step. It is worth noting that the transmit signal at each antenna is adjusted to the required IBO at the linear precoder's output (for the sake of simplicity, it is not presented in the FIGURE 4.1 and FIGURE 4.2).

On the receiver side, another NL NN with R thick layers is utilized as a decoder, similar to the transmitting NN-precoder. The r-th layer of the NN decoder has ℓ_{r} neurons and represents the receivers' activities. All users' decoding operations in our system employ the same NN-decoder. Only one NN-decoder is trained during the off-line learning phase, similar to the precoder, however all users employed the same decoder during the on-line testing phase.

It IS worth noting that utilizing this method reduces receiver training parameter measurements, theoretically leading to a speedier training process and low-complexity precoder/decoder on-line execution. In order to construct the probability vector at user m_r, $p_k \in (0,1)^{|\mathbf{M}|}$, the last layer's activation function in the receivers' NN is set to softmax activation, where the i-th element represents the probability for the desired symbol's index to be i. Finally, user m_r declares $\hat{m}$ as the decoded index of the intended symbol, which corresponds to the index of the highest probability element in p_k. Here, we may use cross-entropy to train the network so that it learns to map the proper output to the proper input.

TABLE 4.1 – Complexity analysis

Method	MU-PNL-GDm [2]	AP-mMIMO (ZF)	AP-mMIMO (MP)
Linear precoder computation	-	$8M_tM_r^2 + M_r^3$	–
NN-precoder/decoder	-	$(M+2)\ell_{\mathrm{t}}M_r\tau + (M+2)\ell_{\mathrm{r}}M_r\tau$	$(M+2)\ell_{\mathrm{t}}M_r\tau + (M+2)\ell_{\mathrm{r}}M_r\tau$
Precoding vector update	$\tau N_{iter}(12M_tM_r + 128M_t)$	$4\tau M_tM_r$	$4\tau(2\mathcal{J}+1)M_tM_r$ [118]

4.3.2 Implementation details

TensorFlow [120], an open source Python-based machine learning framework, is used to create the AP-mMIMO based deep NN in FIGURE 4.2. The Adamax optimizer [121], a variant of the stochastic gradient descent methodology for neural network optimization, is used to train the NN-precoder and NN-decoder. As a loss function, we employ categorical cross-entropy. There is only one hidden layer in each of the NN precoder and decoder (i.e., $T = 2$ and $R = 2$) with, respectively, ℓ_{t} and ℓ_{r} hidden neurons. Furthermore, we use rectified linear unit (ReLU) as the hidden layer activation function [122].

FIGURE 4.1 illustrates the generalization phase of our proposed AP-mMIMO illustrating a typical massive MIMO Downlink system. We recall that the NN-precoder and NN-decoder blocks are identical for all users and are trained with only one given channel matrix $\mathbf{H}$, which is chosen randomly.

Finally, we use the categorical cross-entropy loss as shown in FIGURE 4.2 to train both the NN-precoder and the NN-decoder simultaneously and through various mMIMO chain operations, which is appropriate to our purposes (detection of changes in the probability distribution) and has been demonstrated to lead to quick training as well as extremely good generalization [123].

4.4 Computational Complexity Analysis

In this section, we examine the proposed AP-online mMIMO's computational complexity and compare it to the algorithm-based solution MU-PNL-GDm [2], which combines linear precoding, PAPR reduction, and DPD. To our knowledge, the MU-PNL-GDm provides the highest performance in terms of symbol error rate (SER) among the current literature, but its computational complexity and latency are sufficiently high, making it unsuitable for real-time massive MU-MIMO Downlink systems.

We recall that the linear precoding in the proposed AP-mMIMO can be either the ZF or its low-

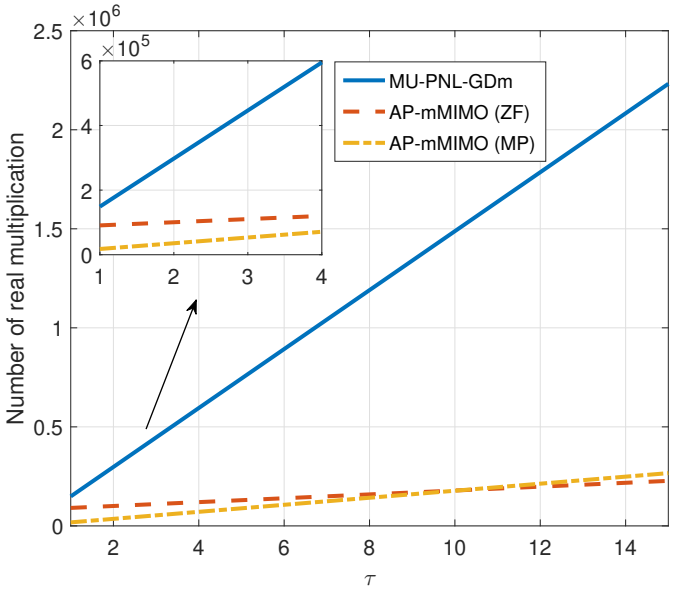


FIGURE 4.3 – Complexity vs. τ.

complexity version based on MPs. As a result, there are two AP-mMIMO schemes : 1) AP-mMIMO (ZF) and 2) AP-mMIMO (MP).

The complexity of implementing the three different approaches (AP-mMIMO (ZF), AP-mMIMO (MP), and MU-PNL-GDm) are summarized in TABLE 4.1. The number of real multiplications is computed here, where τ is related to the channel coherence interval, which indicates that the channel properties remain constant throughout τ symbols.

Because the NN-precoder and NN-decoder are trained off-line, the corresponding complexity is not taken into account. As a result, the proposed AP-online mMIMO's computational complexity involves performing NL precoding, linear precoding (ZF or MP), and NL decoding within a channel coherence interval (i.e., processing τ symbols).

4.5 Simulation Results

In this section, we evaluate the performance of the proposed AP-mMIMO based end-to-end massive MU-MIMO Downlink systems under PA non-linearities. For all the simulations, we consider the BS to be equipped with $M_t = 100$ antennas and serving $M_r = 10$ single-antenna users simultaneously. 16-quadrature amplitude modulation (QAM) is considered (i.e., $M = 16$). The memory-less modified

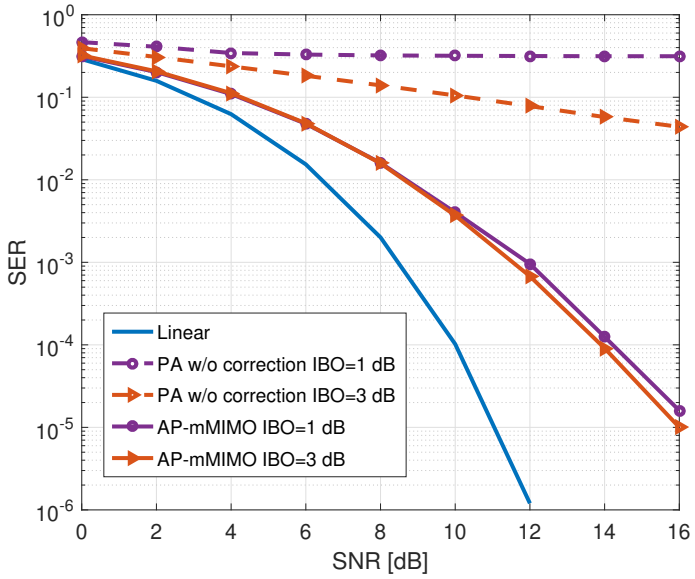


FIGURE 4.4 – SER vs. SNR comparison when PA is operated at IBO = 1 dB and 3 dB

Rapp model is adopted to determine the PA input and output, with parameters $G = 16$, $V_{sat} = 1.9$, $p = 1.1$, $A = -345$, $B = 0.17$, and $q = 4$ [109]. Note that new test data over 10^5 channel realizations were considered to evaluate the SER performance.

FIGURE 4.4 illustrates the SER vs. signal-to-noise ratio (SNR) of the proposed AP-mMIMO when PAs are operating at IBO = 1 dB and 3 dB. As references, we plot the scenario that there is no PA distortion correction ("PA w/o correction") and the linear case where the PA is considered as an ideal one ("Linear"). Each of the NN-precoder and NN-decoder has one hidden layer of 16 neurons and the SNR is defined as SNR $= \mathbb{E}\{\|y_n\|_2^2\}/\sigma_b^2$. From the FIGURE 4.4, we can clearly note that our proposed AP-mMIMO scheme can provide significant improvement compared to the case without correction, which is quite close to the one provided by an ideal massive MU-MIMO system. It is worth mentioning that the two options of the proposed AP-mMIMO, either with ZF or MP, provide the same performance in terms of SER when $\mathcal{J} = 5$, it was also mentioned in [118]. Then, for the sake of simplicity, only one AP-mMIMO SER curve is given for each IBO value.

Most importantly, the advantage of our proposed method is the low complexity needed in online, which makes it interesting for real-time massive MU-MIMO downlink systems. Indeed, according to the results illustrated in FIGURE 4.3, the computational complexity vs. τ, one can note that our proposed method, the AP-mMIMO ZF/MP, has much lower complexity for all range of τ compared

to the MU-PNL-GDm algorithm [2]. Here, $N_{iter} = 6$ and $\mathcal{J} = 5$, as presented, respectively, in [2] and [118]. When $\tau = 5$, the complexity of AP-mMIMO (ZF) and AP-mMIMO (MP) are only 17.85% and 11.94% of that of MU-PNL-GDm, respectively.

In addition, we can note that when τ is small (≤ 10), the AP-mMIMO (MP) scheme requires lower complexity than the AP-mMIMO (ZF) one. However, this latter performs better, in terms of complexity, when τ becomes larger.

4.6 Conclusion

We presented a novel DL Autoprecoder-based massive MU-MIMO downlink transmission method, which aims to minimize MUI and compensate PA impairments over varying fading channels. In comparison to current literature, numerical results clearly showed the capabilities of the proposed AP-mMIMO to provide competitive performance with a competitive computational complexity. The following are some of the chapter's most interesting findings : (1) The proposed AP-mMIMO is suitable for varying fading channels due to the two-stage precoding approach used. As a result, when the channel changes, just the linear precoding is changed. (2) The NN precoder/decoder can be trained off-line and do not rely on real-time channel estimations. (3) For fast-changing channel conditions, the AP-mMIMO (MP) is preferred, whereas the AP-mMIMO (ZF) is suitable for slow-changing channels.

Chapitre 5

Conclusion and Perspectives

Contents

5.1 Conclusion

This thesis focuses on the efficient transmission in 5G and beyond wireless communication systems when non linear PA and different numerologies are used in a massive MIMO context. When using different numerologies, we have established closed form expressions for the inter numerology interference (INI) and we have proposed an INI cancellation system for both uplink and downlink when using CP-OFDM waveform. Concerning PA impairments, we have proposed a new low complexity algorithm able to realize the MU-MIMO precoding together with PAPR reduction. Performances of this proposed algorithm are very closed to the ones of the available algorithms in litterature while having a lower complexity. Finally we have proposed an auto-precoder based on DL approaches called AP-mMIMO scheme. This auto-precoder is able to correct the impairments brought by NL PA while having a reasonnable complexity.

In Chapter 1, we gave an overview of the development of wireless communication systems as well as its main characteristics. The recent massive MIMO technology was also introduced and the PA imperfections, which is an essential problem for real transmission, are presented.

In Chapter 2, we have focused on improving the spectral efficiency in massive MIMO-OFDM systems. The mixed-numerologies method is applied to investigate the use of spatial multiplexing of users who are sharing the same bandwidth. The downlink and uplink transmission schemes are propoed and the INI models were built to analyse the inteference introduced between different users. The results show that the proposed cancellation method can reduce the INI and improve the transmission quality. We showed that, in the massive MIMO-OFDM downlink scenario, INI is only created in frequency selective channels and only when users with large SCS communicate with users with small SCS. However, in uplink scenario, all users suffer from INI in frequency selective channels.

In Chapter 3, we have turned our view on the energy efficiency in the real transmission chain. In this chapter, a special interest was given in RF PAs. The PA nonlinear characteristics showed its negative impact on the transmission quality as it brings hardware distortion. In order to compensate this imperfection while maintaining the system efficiency, we analysed several existing solutions and then proposed our new joint MU precoding and PAPR reduction scheme. Here, we designed the scheme in a massive MU-MIMO downlink system. A steepest GD-based technique is used to build the proposed joint MU precoding and PAPR reduction method, which is formulated as a simple convex

optimization problem. The simulation results indicated that, compared to the previous methods, the suggested transmitting scheme provides satisfactory performance while requiring lower computing complexity.

In Chapter 4, we continue to compensate the PA nonlinearities, but with the deep NN tool. Different from the classical optimization method, which separates the transmission chain into several blocks, we have tried to design an end-to-end solution. In our proposed AP-mMIMO scheme, two NNs are trained off-line and are used for all antenna branches and users. In comparison to current literature, numerical results clearly showed the capabilities of the proposed AP-mMIMO to provide competitive performance with even lower computational complexity, compared with the classical algorithms-based solutions.

5.2 Perspectives

Despite the 5G research interests in recent years, researchers have begun working on the sixth generation of the radio-mobile network. 6G is anticipated to arrive in 10 years and should be 100 times faster than 5G, with better coverage and availability [124]. It should also complete the dependability of 5G services (eMBB, mMTC, URLLC) while taking into account extremely low latency and low energy consumption. Communication systems that are autonomous and intelligent have the potential capability to enhance the overall system performance while reducing the workload. 6G will open up new avenues for digital innovations. Under this, we could have some points for future work.

- *Cell-free massive MIMO and Intelligent Reflecting Surface (IRS)*
  Cell-free massive MIMO [125] is a promising solution for meeting the demand for an increasing number of users, higher data rates, and stringent quality-of-service in 6G networks. The basic concept behind cell-free massive MIMO is to use a great number of access points that are randomly placed around the coverage area and attached to a central processing unit (CPU). The access points serve all UEs on the same time-frequency resource with the CPU's coordination and computational help. Therefore, we can consider to extend our studied schemes under the concept of cell-free massive MIMO. What is more, the use of IRS allows for controllable anomalous scattering of incident electromagnetic waves [126] [127]. This allows for active control of the propagation environment and adds another level of optimization to future wireless com-

munication networks. It could be also interesting to investigate energy-efficient schemes with the new IRS concept and channel model.

- *Other hardware imperfections in 6G systems*

  Tx Leakage due to a nearby antenna, mirror-frequency interference due to IQ mismatch, oscillator phase noise, and non-linear distortion due to PA could all occur in the radio transceiver, which is called "Dirty-RF," or radio imperfections caused by the RF chain [128]. Furthermore, cross-talk between antennas may occur, lowering broadcast quality. These imperfections, which are frequently overlooked, have a negative impact on massive MIMO communications. In this thesis, the study was limited to the PA non linearity and its compensation. For future work, we could first investigate the influence of other Dirty-RF imperfections on massive MIMO performance. Then, the next step is to provide dependable DSP-based solutions to reduce the studied RF impairments in 6G communication systems on both the transmitter and receiver sides. This will ensure low-cost and low-power components more widely available.

- *Deep learning enabled 6G*

  In Chapter 4, we have already introduced the deep learning concept in the massive MIMO system and applied it by proposing an AP-based deep learning approach for massive MU-MIMO downlink systems. Apart from that, We could look at other new approaches, such as meta-learning [129], which allows for lesser complexity and complements standard machine learning approaches. The goal is to create a meta-learning model that can be extended to a new configuration that was never learnt during the learning process. We propose to employ two neural networks. The first one, the NN Precoder, creates a precoded data vector, for a certain channel configuration, to compensate the PA nonlinearities or other hardware imperfection. The weights of the NN-Precoder associated to a channel configuration are generated by the second NN, the Meta-NN, which may be extended to any channel configuration. By doing so, we eliminate the NN Precoder's adaption and the associated computing complexity when the channel changes.

Bibliographie

[1] R. Zayani, H. Shaiek et D. Roviras, "Papr-aware massive mimo-ofdm downlink," *IEEE Access*, vol. 7, p. 25 474–25 484, 2019.

[2] R. Zayani, H. Shaïek et D. Roviras, "Efficient precoding for massive mimo downlink under pa nonlinearities," *IEEE Communications Letters*, vol. 23, n°. 9, p. 1611–1615, 2019.

[3] Qualcomm, "Making 5g nr a reality : leading the technology inventions for a unified, more capable 5g air interface," *White paper*, 2016.

[4] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

[5] S. Parkvall, A. Furuskär et E. Dahlman, "Evolution of lte toward imt-advanced," *IEEE Communications Magazine*, vol. 49, n°. 2, p. 84–91, 2011.

[6] A. Osseiran *et al.*, "Scenarios for 5g mobile and wireless communications : the vision of the metis project," *IEEE communications magazine*, vol. 52, n°. 5, p. 26–35, 2014.

[7] S.-Y. Lien *et al.*, "5g new radio : Waveform, frame structure, multiple access, and initial access," *IEEE communications magazine*, vol. 55, n°. 6, p. 64–71, 2017.

[8] J. G. Andrews *et al.*, "What will 5g be ?" *IEEE Journal on selected areas in communications*, vol. 32, n°. 6, p. 1065–1082, 2014.

[9] I. Chih-Lin *et al.*, "New paradigm of 5g wireless internet," *IEEE Journal on Selected Areas in Communications*, vol. 34, n°. 3, p. 474–482, 2016.

[10] A. Yazar, F. A. Onat et H. Arslan, "New generation waveform approaches for 5g and beyond," dans *2016 24th Signal Processing and Communication Application Conference (SIU)*. IEEE, 2016, p. 961–964.

[11] "Technical specification group radio access network, 3gpp tr 38.201 v15.3.0 general aspects for physical layer for nr," 3rd Generation Partnership Project, Rapport technique, 2017.

[12] J. Choi *et al.*, "A transceiver design for spectrum sharing in mixed numerology environments," *IEEE Transactions on Wireless Communications*, vol. 18, n°. 5, p. 2707–2721, May 2019.

[13] J. Guerreiro, R. Dinis et P. Montezuma, "Massive mimo with nonlinear amplification : Signal characterization and performance evaluation," dans *2016 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2016, p. 1–6.

[14] A. A. Khan *et al.*, "Energy efficient design of massive mimo by considering the effects of nonlinear amplifiers," *Energies*, vol. 11, n°. 5, p. 1045, 2018.

[15] C. Mollén *et al.*, "Spatial characteristics of distortion radiated from antenna arrays with transceiver nonlinearities," *IEEE Transactions on Wireless Communications*, vol. 17, n°. 10, p. 6663–6679, 2018.

[16] F. Sohrabi et W. Yu, "One-bit precoding constellation design via autoencoder-based deep learning," dans *2019 53rd Asilomar Conference on Signals, Systems, and Computers*. IEEE, 2019, p. 754–758.

[17] A. Kaye et D. George, "Transmission of multiplexed pam signals over multiple channel and diversity systems," *IEEE Transactions on Communication Technology*, vol. 18, n°. 5, p. 520–526, 1970.

[18] L. Brandenburg et A. Wyner, "Capacity of the gaussian channel with memory : The multivariate case," *Bell System Technical Journal*, vol. 53, n°. 5, p. 745–778, 1974.

[19] W. van Etten, "Maximum likelihood receiver for multiple channel transmission systems," *IEEE Transactions on Communications*, vol. 24, n°. 2, p. 276–283, 1976.

[20] J. Salz, "Digital transmission over cross-coupled linear channels," *At&T Technical Journal*, vol. 64, n°. 6, p. 1147–1159, 1985.

[21] V. Jones et G. C. Raleigh, "Channel estimation for wireless ofdm systems," dans *IEEE GLOBECOM 1998 (Cat. NO. 98CH36250)*, vol. 2. IEEE, 1998, p. 980–985.

[22] G. D. Golden *et al.*, "Detection algorithm and initial laboratory results using v-blast space-time communication architecture," *Electronics letters*, vol. 35, n°. 1, p. 14–16, 1999.

[23] L. Zheng et D. N. C. Tse, "Diversity and multiplexing : A fundamental tradeoff in multiple-antenna channels," *IEEE Transactions on information theory*, vol. 49, n°. 5, p. 1073–1096, 2003.

[24] D. Tse et P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.

[25] A. F. Molisch *et al.*, "Hybrid beamforming for massive mimo : A survey," *IEEE Communications magazine*, vol. 55, n°. 9, p. 134–141, 2017.

[26] A. Van Zelst et T. C. Schenk, "Implementation of a mimo ofdm-based wireless lan system," *IEEE Transactions on signal processing*, vol. 52, n°. 2, p. 483–494, 2004.

[27] Q. Li *et al.*, "Mimo techniques in wimax and lte : a feature overview," *IEEE Communications magazine*, vol. 48, n°. 5, p. 86–92, 2010.

[28] "Technical specification group radio access network, 3gpp tr 38.817-02 v15.3.0 general aspects for base station radio frequency for nr," 3rd Generation Partnership Project, Rapport technique, 2019.

[29] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE transactions on wireless communications*, vol. 9, n°. 11, p. 3590–3600, 2010.

[30] M. Costa, "Writing on dirty paper (corresp.)," *IEEE transactions on information theory*, vol. 29, n°. 3, p. 439–441, 1983.

[31] B. M. Hochwald, C. B. Peel et A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication-part ii : Perturbation," *IEEE Transactions on Communications*, vol. 53, n°. 3, p. 537–544, 2005.

[32] C. Windpassinger, R. F. Fischer et J. B. Huber, "Lattice-reduction-aided broadcast precoding," *IEEE Transactions on Communications*, vol. 52, n°. 12, p. 2057–2060, 2004.

[33] F. Rusek *et al.*, "Scaling up mimo : Opportunities and challenges with very large arrays," *IEEE signal processing magazine*, vol. 30, n°. 1, p. 40–60, 2012.

[34] T. K. Lo, "Maximum ratio transmission," dans *1999 IEEE International Conference on Communications (Cat. No. 99CH36311)*, vol. 2. IEEE, 1999, p. 1310–1314.

[35] C. B. Peel, B. M. Hochwald et A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication-part i : channel inversion and regularization," *IEEE Transactions on Communications*, vol. 53, n°. 1, p. 195–202, 2005.

[36] E. Björnson, M. Bengtsson et B. Ottersten, "Optimal multiuser transmit beamforming : A difficult problem with a simple solution structure [lecture notes]," *IEEE Signal Processing Magazine*, vol. 31, n°. 4, p. 142–148, 2014.

[37] W. H. Greub, *Linear algebra*. Springer Science & Business Media, 2012, vol. 23.

[38] H. Sun, J. Guo et L. Fang, "Improved singular value decomposition (topsvd) for source number estimation of low snr in blind source separation," *IEEE Access*, vol. 5, p. 26 460–26 465, 2017.

[39] Y. Rahulamathavan *et al.*, "Adaptive subcarrier and bit allocation techniques for mimo-ofdma based uplink cognitive radio networks," dans *2009 First UK-India International Workshop on Cognitive Wireless Systems (UKIWCWS)*. IEEE, 2009, p. 1–5.

[40] Z. Xiong *et al.*, "Grassmannian beamforming and null space broadcasting protocols for cognitive radio networks," *IET signal processing*, vol. 5, n°. 5, p. 451–460, 2011.

[41] G. J. Foschini et M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless personal communications*, vol. 6, n°. 3, p. 311–335, 1998.

[42] A. Sahin et H. Arslan, "Multi-user aware frame structure for ofdma based system," dans *2012 IEEE Vehicular Technology Conference (VTC Fall)*. IEEE, 2012, p. 1–5.

[43] P. Guan *et al.*, "5g field trials : Ofdm-based waveforms and mixed numerologies," *IEEE Journal on Selected Areas in Communications*, vol. 35, n°. 6, p. 1234–1243, 2017.

[44] L. Zhang *et al.*, "Subband filtered multi-carrier systems for multi-service wireless communications," *IEEE Transactions on Wireless Communications*, vol. 16, n°. 3, p. 1893–1907, 2017.

[45] J. Choi *et al.*, "A transceiver design for spectrum sharing in mixed numerology environments," *IEEE Transactions on Wireless Communications*, vol. 18, n°. 5, p. 2707–2721, 2019.

[46] X. Cheng *et al.*, "Inter-numerology interference analysis and cancellation for massive mimo-ofdm downlink systems," *IEEE Access*, vol. 7, p. 177 164–177 176, 2019.

[47] T. S. Rappaport, "Wireless communications–principles and practice, (the book end)," *Microwave Journal*, vol. 45, n°. 12, p. 128–129, 2002.

[48] H. Jafarkhani, *Space-time coding : theory and practice*. Cambridge university press, 2005.

[49] J. Proakis et M. Salehi, "Digital communication 5th edition," 2007.

[50] G. T. Zhou et R. Raich, "Spectral analysis of polynomial nonlinearity with applications to rf power amplifiers," *EURASIP Journal on Advances in Signal Processing*, vol. 256395, p. 1840–1931, 2004.

[51] W. Bosch et G. Gatti, "Measurement and simulation of memory effects in predistortion linearizers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 37, n°. 12, p. 1885–1890, 1989.

[52] P. Colantonio, F. Giannini et E. Limiti, "High efficiency rf and microwave solid state power amplifiers," *Wiley*, 2009.

[53] A. A. M. Saleh, "Frequency independent and frequency dependent nonlinear model of twt amplifier," *IEEE Transactions on Communications*, vol. 29, n°. 11, p. 1715–1720, 1981.

[54] A. L. Berman et C. E. Mahle, "Nonlinear phase shift in traveling-wave tubes as applied to multiple access communications satellites," *IEEE Transactions on Communication Technology*, vol. 18, 1970.

[55] S. Boumaiza *et al.*, "Adaptive digital/rf predistortion using a nonuniform lut indexing function with built-in dependence on the amplifier nonlinearity," *IEEE Transactions on Microwave Theory Techniques*, vol. 52, n°. 12, p. 2670–2677, 2004.

[56] J. Kim et K. Konstantinou, "Digital predistortion of wideband signals based on power amplifier model with memory," *Electronics Letters*, vol. 37, n°. 23, p. 1417–1418, 2002.

[57] D. R. Morgan *et al.*, "A generalized memory polynomial model for digital predistortion of rf power amplifiers," *IEEE Transactions on Signal Processing*, vol. 54, p. p.3852–3860, 2006.

[58] A. Zhu, J. C. Pedro et T. J. Brazil, "Dynamic deviation reduction-based volterra behavioral modeling of rf power amplifiers," *IEEE Transactions on Microwave Theory Techniques*, vol. 54, p. 4323–4332, 2006.

[59] O. Hammi, M. Younes et F. M. Ghannouchi, "Metrics and methods for benchmarking of rf transmitter behavioral models with application to the development of a hybrid memory polynomial model," *IEEE Transactions on Broadcasting*, vol. 56, n°. 3, p. 350–357, 2010.

[60] M. Ibukahla et J. Sombria, "Neural networks for modeling nonlinear memoryless communication channels," *IEEE Transactions on Communications*, vol. 45, n°. 7, p. 768–771, 1997.

[61] M. Ibnkahla et F. Castanie, "Vector neural networks for digital satellite communications," dans *IEEE International Conference on Communications*, 1995.

[62] T. Liu, S. Boumaiza et F. M. Ghannouchi, "Dynamic behavioral modeling of 3g power amplifiers using real-valued time-delay neural networks," *IEEE Transactions on Microwave Theory Techniques*, vol. 52, n°. 3, p. 1025–1033, 2004.

[63] H. Leung et S. Haykin, "The complex backpropagation algorithm," *IEEE Transactions on Signal Processing*, vol. 39, n°. 9, p. P.2101–2104, 1991.

[64] "Realistic power amplifier model for the new radio evaluation, document r4-163314, 3gpp tsg-ran wg4 meeting," 3rd Generation Partnership Project, Rapport technique, 2016.

[65] "Rp-160671 : Study on new radio access technology, ntt docomo, inc," Tokyo, Japan, Rapport technique, 2016.

[66] J. K. Cavers, "Amplifier linearization by adaptive predistortion," 1991.

[67] J. Wood, "System-level design considerations for digital pre-distortion of wireless base station transmitters," *IEEE Transactions on Microwave Theory Techniques*, vol. 65, n°. 5, p. 1880–1890, 2017.

[68] J. Tao et Y. Wu, "An overview : Peak-to-average power ratio reduction techniques for ofdm signals," *IEEE Transactions on Broadcasting*, vol. 54, n°. 2, p. 257–268, 2008.

[69] S. H. Han et J. H. Lee, "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission," *IEEE Wireless Communications*, vol. 12, n°. 2, p. 56–65, 2005.

[70] T. A. Wilkinson et A. E. Jones, "Minimisation of the peak to mean envelope power ratio of multicarrier transmission schemes by block coding," dans *IEEE Vehicular Technology Conference*, 1995.

[71] S. H. Müller *et al.*, "Ofdm with reduced peak-to-average power ratio by multiple signal representation," *Annales Des Télécommunications*, vol. 52, n°. 1, p. 58–67, 1997.

[72] S. H. Muller et J. B. Huber, "Ofdm with reduced peak-to-average power ratio by optimum combination of partial transmit sequences," *Electronics Letters*, vol. 33, n°. 5, p. 368–369, 1997.

[73] R. Bauml et R. F. H. Fischer, "Reducing the peak-to-average power ratio of multicarrier modulation by selected mapping," *Electronics Letters*, vol. 32, n°. 22, p. 2056–2057, 1996.

[74] J. Tellado, "Peak to average power reduction for multicarrier modulation," *Thesis*, 1999.

[75] R. O'Neill et L. B. Lopes, "Envelope variations and spectral splatter in clipped multicarrier signals," dans *IEEE Conference Pmirc*, 1995.

[76] S. H. Han, J. Cioffi et J. H. Lee, "Tone injection with hexagonal constellation for peak-to-average power ratio reduction in ofdm," *Communications Letters IEEE*, vol. 10, n°. 9, p. 646–648, 2006.

[77] A. Sahin et H. Arslan, "Multi-user aware frame structure for ofdma based system," dans *2012 IEEE Vehicular Technology Conference (VTC Fall)*, Sep. 2012, p. 1–5.

[78] S. Lien *et al.*, "5g new radio : Waveform, frame structure, multiple access, and initial access," *IEEE Communications Magazine*, vol. 55, n°. 6, p. 64–71, Jun. 2017.

[79] P. Guan *et al.*, "5g field trials : Ofdm-based waveforms and mixed numerologies," *IEEE Journal on Selected Areas in Communications*, vol. 35, n°. 6, p. 1234–1243, Jun. 2017.

[80] L. Zhang *et al.*, "Subband filtered multi-carrier systems for multi-service wireless communications," *IEEE Transactions on Wireless Communications*, vol. 16, n°. 3, p. 1893–1907, Mar. 2017.

[81] ——, "Multi-service system : An enabler of flexible 5g air interface," *IEEE Communications Magazine*, vol. 55, n°. 10, p. 152–159, Oct. 2017.

[82] A. Yazar et H. Arslan, "Reliability enhancement in multi-numerology-based 5g new radio using ini-aware scheduling," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, n°. 1, p. 110, May 2019. [En ligne]. Disponible : https://doi.org/10.1186/s13638-019-1435-z

[83] X. Zhang *et al.*, "Mixed numerologies interference analysis and inter-numerology interference cancellation for windowed ofdm systems," *IEEE Transactions on Vehicular Technology*, vol. 67, n°. 8, p. 7047–7061, Aug. 2018.

[84] A. B. Kihero *et al.*, "Inter-numerology interference analysis for 5g and beyond," dans *2018 IEEE Globecom Workshops (GC Wkshps)*, Dec. 2018, p. 1–6.

[85] A. Yazar, B. Pekoz et H. Arslan, "Flexible multi-numerology systems for 5g new radio," 05 2018.

[86] S.Rajagopal et M. S. Rahman, "Multi-user MIMO with flexible numerology for 5g," *CoRR*, vol. abs/1610.03056, 2016. [En ligne]. Disponible : http://arxiv.org/abs/1610.03056

[87] Q. H. Spencer, A. L. Swindlehurst et M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser mimo channels," *IEEE Transactions on Signal Processing*, vol. 52, n°. 2, p. 461–471, Feb. 2004.

[88] Q. Qin *et al.*, "Sparse channel estimation for massive mimo-ofdm systems over time-varying channels," *IEEE Access*, vol. 6, p. 33 740–33 751, 2018.

[89] L. Lu *et al.*, "An overview of massive mimo : Benefits and challenges," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, n°. 5, p. 742–758, Oct. 2014.

[90] J. Vieira *et al.*, "Reciprocity calibration for massive mimo : Proposal, modeling, and validation," *IEEE Transactions on Wireless Communications*, vol. 16, n°. 5, p. 3042–3056, May 2017.

[91] E. G. Larsson *et al.*, "Massive mimo for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, n°. 2, p. 186–195, Feb. 2014.

[92] S. on channel model for frequencies from 0.5 to 100 GHz (Release 14), "V14.3.0," *document*, vol. Tr, n°. 38, p. 901, 3GPP 2018.

[93] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, n°. 11, p. 3590–3600, Nov. 2010.

[94] P. Pan *et al.*, "How many antenna arrays are dense enough in massive mimo systems," *IEEE Transactions on Vehicular Technology*, vol. 67, n°. 4, p. 3042–3053, Apr. 2018.

[95] A. S. Y. Poon, R. W. Brodersen et D. N. C. Tse, "Degrees of freedom in multiple-antenna channels : a signal space approach," *IEEE Transactions on Information Theory*, vol. 51, n°. 2, p. 523–536, Feb. 2005.

[96] H. Q. Ngo, E. G. Larsson et T. L. Marzetta, "Energy and spectral efficiency of very large multiuser mimo systems," *IEEE Transactions on Communications*, vol. 61, n°. 4, p. 1436–1449, Apr. 2013.

[97] P. Patcharamaneepakorn *et al.*, "Spectral, energy, and economic efficiency of 5g multicell massive mimo systems with generalized spatial modulation," *IEEE Transactions on Vehicular Technology*, vol. 65, n°. 12, p. 9715–9731, Dec. 2016.

[98] Y. Wang et J. Lee, "A zf-based precoding scheme with phase noise suppression for massive mimo downlink systems," *IEEE Transactions on Vehicular Technology*, vol. 67, n°. 2, p. 1158–1173, Feb. 2018.

[99] S. F. Mohammed, A. Chockalingam et B. S. Rajan, "A low-complexity precoder for large multiuser miso systems," dans *VTC Spring 2008-IEEE Vehicular Technology Conference*. IEEE, 2008, p. 797–801.

[100] I. Iofedov, I. Gutmany et D. Wulich, "Iterative precoding of ofdm-miso with nonlinear power amplifiers," dans *2015 IEEE International Conference on Communications (ICC)*. IEEE, 2015, p. 2399–2404.

[101] C. Siegl et R. F. Fischer, "Selected basis for par reduction in multi-user downlink scenarios using lattice-reduction-aided precoding," *EURASIP Journal on advances in signal processing*, vol. 2011, n°. 1, p. 1–11, 2011.

[102] J.-C. Chen, "Low-papr precoding design for massive multiuser mimo systems via riemannian manifold optimization," *IEEE Communications Letters*, vol. 21, n°. 4, p. 945–948, 2017.

[103] M. Yao *et al.*, "Semidefinite relaxation-based papr-aware precoding for massive mimo-ofdm systems," *IEEE Transactions on Vehicular Technology*, vol. 68, n°. 3, p. 2229–2243, 2018.

[104] C. Ni, Y. Ma et T. Jiang, "A novel adaptive tone reservation scheme for papr reduction in large-scale multi-user mimo-ofdm systems," *IEEE Wireless Communications Letters*, vol. 5, n°. 5, p. 480–483, 2016.

[105] H. Prabhu *et al.*, "A low-complex peak-to-average power reduction scheme for ofdm based massive mimo systems," dans *2014 6th international symposium on communications, control and signal processing (ISCCSP)*. IEEE, 2014, p. 114–117.

[106] H. Bao *et al.*, "An admm approach for papr reduction for large-scale mimo-ofdm systems," *IEEE Transactions on Vehicular Technology*, vol. 67, n°. 8, p. 7407–7418, 2018.

[107] H. Bao *et al.*, "An efficient bayesian papr reduction method for ofdm-based massive mimo systems," *IEEE Transactions on Wireless Communications*, vol. 15, n°. 6, p. 4183–4195, 2016.

[108] M. Yao *et al.*, "A digital predistortion scheme exploiting degrees-of-freedom for massive mimo systems," dans *2018 IEEE International Conference on Communications (ICC)*. IEEE, 2018, p. 1–5.

[109] "Realistic power amplifier model for the new radio evaluation," *document*, vol. R4-163314, n°. 3GPP TSG-RAN WG4 Meeting 79, p. Nokia, 2016.

[110] H. Shaiek *et al.*, "Analytical analysis of ser for beyond 5g post-ofdm waveforms in presence of high power amplifiers," *IEEE Access*, vol. 7, p. 29 441–29 452, 2019.

[111] T. Jiang *et al.*, "Curve fitting based tone reservation method with low complexity for papr reduction in ofdm systems," *IEEE Communications Letters*, vol. 18, n°. 5, p. 805–808, 2014.

[112] Y.-F. Wang et J.-H. Lee, "A zf-based precoding scheme with phase noise suppression for massive mimo downlink systems," *IEEE Transactions on vehicular Technology*, vol. 67, n°. 2, p. 1158–1173, 2017.

[113] T. O'Shea et J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, n°. 4, p. 563–575, 2017.

[114] S. Dörner *et al.*, "Deep learning based communication over the air," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, n°. 1, p. 132–143, 2017.

[115] A. Felix *et al.*, "Ofdm-autoencoder for end-to-end learning of communications systems," dans *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2018, p. 1–5.

[116] A. Balatsoukas-Stimming *et al.*, "Neural-network optimized 1-bit precoding for massive mu-mimo," dans *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2019, p. 1–5.

[117] O. Castañeda *et al.*, "1-bit massive mu-mimo precoding in vlsi," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 7, n°. 4, p. 508–522, 2017.

[118] J. Zhu, R. Schober et V. K. Bhargava, "Linear precoding of data and artificial noise in secure massive mimo systems," *IEEE Transactions on Wireless Communications*, vol. 15, n°. 3, p. 2245–2261, 2015.

[119] Q. H. Spencer, A. L. Swindlehurst et M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser mimo channels," *IEEE transactions on signal processing*, vol. 52, n°. 2, p. 461–471, 2004.

[120] M. Abadi *et al.*, "Tensorflow : Large-scale machine learning on heterogeneous distributed systems," *arXiv preprint arXiv :1603.04467*, 2016.

[121] D. P. Kingma et J. Ba, "Adam : A method for stochastic optimization," *arXiv preprint arXiv :1412.6980*, 2014.

[122] B. Xu *et al.*, "Empirical evaluation of rectified activations in convolutional network," *arXiv preprint arXiv :1505.00853*, 2015.

[123] I. J. Goodfellow, Y. Bengio et A. Courville, *Deep Learning*. Cambridge, MA, USA : MIT Press, 2016, http://www.deeplearningbook.org.

[124] T. Nakamura, "5g evolution and 6g," dans *2020 IEEE Symposium on VLSI Technology*. IEEE, 2020, p. 1–5.

[125] S. Chen *et al.*, "Structured massive access for scalable cell-free massive mimo systems," *IEEE Journal on Selected Areas in Communications*, vol. 39, n°. 4, p. 1086–1100, 2020.

[126] Q. Wu et R. Zhang, "Towards smart and reconfigurable environment : Intelligent reflecting surface aided wireless network," *IEEE Communications Magazine*, vol. 58, n°. 1, p. 106–112, 2019.

[127] B. Matthiesen *et al.*, "Intelligent reflecting surface operation under predictable receiver mobility : A continuous time propagation model," *IEEE Wireless Communications Letters*, vol. 10, n°. 2, p. 216–220, 2020.

[128] S. Gunturi et J. Balakrishnan, "Mitigation of narrowband interference in differentially modulated communication systems," dans *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2009, p. 2549–2552.

[129] J. Vanschoren, "Meta-learning : A survey," *arXiv preprint arXiv :1810.03548*, 2018.

Xinying CHENG

Study and mitigation techniques of RF impairments for beyond 5G multi-carrier waveforms

Résumé : La recherche fondamentale vers les réseaux cellulaires au-delà de la 5G est en cours et la vision de 2020 et au-delà comprend un nombre important de cas d'utilisation compte tenu d'un grand nombre d'appareils avec un large éventail de caractéristiques et de demandes. D'une part, la bonne efficacité spectrale donne plus de capacité pour la transmission tandis que d'autre part, la construction d'équipements radio compacts et peu coûteux, flexibles et de haute qualité est une tâche très difficile. Le contexte des travaux de cette thèse est l'étude des techniques de type massive MIMO en présence d'imperfections RF, notamment celles apportées par les amplificateurs de puissance (PA) non linéaires (NL). Le nombre important de PA dans une station de base de type massive-MIMO crée différentes contraintes et les résultat de ces contraintes entraine fortes dégradation des signaux émis tant dans la bande que hors bande. D'autre part, la normalisation 5G a introduit le concept de différentes numérologies conjointement à la technique massive MIMO. L'influence des différentes numérologies utilisées par les différents utilisateurs sera aussi étudiée dans cette thèse.

Mots clés : Massive MIMO, 5G, Numérologies, Imperfection RF, Amplificateurs de Puissance (PA), MU precoding

Abstract : Fundamental research towards beyond 5G cellular networks is ongoing and the vision of 2020 and beyond includes a significant amount use cases considering a massive number of devices with a wide range of characteristics and demands. On the one hand, the good spectral efficiency gives more capacity for the transmission while on the other hand, building compact and low-cost flexible and high-quality radio equipment is a very challenging task. The context of this thesis is the study of massive MIMO techniques with the presence of radio frequency (RF) imperfections, in particular, the non-linear (NL) power amplifiers (PA). The large number of PAs equipped in the base station (BS) creates various constraints which lead to strong degradation of the transmission quality both in the band and out of band. On the other hand, 5G standardization introduced the concept of different numerologies together with the massive MIMO technique. The influence of the different numerologies used by the different users will also be studied in this thesis.

Keywords : Massive MIMO, 5G, Numerologies, RF Imperfection, Power Amplifier (PA), MU precoding