Conception de schémas et récepteurs avancés pour le SCMA et la transmission MIMO non cohérente dans les futurs réseaux sans fil

Xiaotian Fu

HESAM UNIVERSITÉ

le cnam

ÉCOLE DOCTORALE Sciences des Métiers de l'Ingénieur
Centre d'études et de recherche en information et communications

THÈSE

présentée par : **Xiaotian FU**
soutenue le : **09 Decembre 2021**

pour obtenir le grade de : **Docteur d'HESAM Université**

préparée au : **Conservatoire national des arts et métiers**

Discipline : **Génie Informatique, Automatique et Traitement du Signal**

Spécialité : **Radiocommunications**

Scheme Design and Advanced Receivers for SCMA and Non-coherent MIMO in Future Wireless Networks

THÈSE dirigée par :
[**M. LE RUYET Didier**] Professeur des Universités, CNAM

Jury		
Mme. Ghaya REKAYA	Professeur, Telecom Paris	Président
M. Matthieu CRUSSIERE	Professeur des Universités, INSA Rennes	Rapporteur
M. Jean-Pierre CANCES	Professeur des Universités, Université de Limoges	Rapporteur
Mme. Inbar FIJALKOW	Professeure des Universités, ENSEA, Cergy Paris Université	Examinateur
M. Jean-Marie GORCE	Professeur des Universités, INSA Lyon	Examinateur

THÈSE

Affidavit

Je soussignée, Xiaotian Fu, déclare par la présente que le travail présenté dans ce manuscrit est mon propre travail, réalisé sous la direction scientifique de Prof. Didier Le Ruyet, dans le respect des principes d'honnêteté, d'intégrité et de responsabilité inhérents à la mission de recherche. Les travaux de recherche et la rédaction de ce manuscrit ont été réalisés dans le respect de la charte nationale de déontologie des métiers de la recherche. Ce travail n'a pas été précédemment soumis en France ou à l'étranger dans une version identique ou similaire à un organisme examinateur.

Fait à Paris, le 28/09/2021

Signature
Xiaotian Fu

Affidavit

I, undersigned, Xiaotian Fu, hereby declare that the work presented in this manuscript is my own work, carried out under the scientific direction of Prof. Didier Le Ruyet, in accordance with the principles of honesty, integrity and responsibility inherent to the research mission. The research work and the writing of this manuscript have been carried out in compliance with the French charter for Research Integrity. This work has not been submitted previously either in France or abroad in the same or in a similar version to any other examination body.

Place Paris, date 28/09/2021

Signature
Xiaotian Fu

Acknowledgment

I would like to thank the following people, without whom I would not have been able to complete this research thesis, and without whom I would not have made it through my Ph.D. journey.

First and foremost, I would like to express my gratitude and appreciation for my supervisor Prof. Didier Le Ruyet whose professional guidance, support and encouragement has been invaluable over the last three years. His rigorous as well as optimistic attitudes toward research have motivated and helped me. Also, I want to thank Mylène Pischella for her guidance of my research and thoughtful caring to my work life at CNAM.

I thank my colleague Bruno Fontana da Silva who has collaborated with me. His expertise in wireless communications and passion for research really inspire me. Also, I am grateful to Prof. Bartolomeu Ferreira Uchôa-Filho who has helped a lot on my research with his expertise in mathematics and thoroughness of research. I thank all the colleagues that I have met at CEDRIC lab. I am very honored to work with so many smart and interesting people coming from all over the world.

Last but not least, I really appreciate the love and unconditional support from my parents throughout my entire journey of studying abroad. I am grateful for their understanding and caring. I also want to thank all my friends who have ever supported and encouraged me in my Ph.D. study.

Abstract

With the rapid development of multimedia and heterogeneous devices, the future wireless networks are anticipated to provide full connectivity, intelligent and flexible services. Therefore, there have been many discussions on the next generation of wireless networks, referred to as the sixth-generation (6G). Considering the challenges of 6G and the limitations of the current networks, we study three promising schemes in this thesis, which are sparse code multiple access (SCMA), non-coherent (NC) multiple-input multiple-output (MIMO) and single-user (SU) multidimensional transmission.

First of all, we focus on SCMA which enables more users than radio resources to transmit their signals. Due to the superposition of codewords, multiuser detection (MUD) algorithms play an significant role in SCMA system performance. In this thesis, we investigate two kinds of MUD algorithms for SCMA, iterative and sphere decoding (SD)-based. The message passing algorithm (MPA) can achieve the optimal performance when the number of iterations is large. However, the complexity of MPA increases dramatically with the number of users and overloading factor. First, we work on the iterative detectors. In a simple coded transmission scenario, we propose a low-complexity iterative MUD algorithm based on Graussian approximation which shows good performance and substantial low complexity. Moreover, we consider a highly overload SCMA transmission where successive interference cancellation and iterative MUD algorithm are employed jointly. To satisfy the trade-off between spectral efficiency and system performance, an optimization method of choosing channel code rates is proposed. After that, we review the sphere decoding (SD)-based detectors that have less or no constraint of SCMA codebooks. To guarantee low MUD complexity and good symbol error rate (SER) performance at the same time, we two pruning algorithms and introduce the simplified generalized SD for SCMA (SGSD-SCMA). An estimated error rate of the pruning algorithms is derived. The proposed detector performances particularly well with moderate overloading factor and large codebook size.

Secondly, we study the NC MIMO systems where no channel state information is required, favo-

rable to high mobility transmission. The system capacity analysis suggests that NC MIMO systems should adopt Grassmannian constellations. The design of constellations is essential to NC MIMO system performance. On the other hand, deep learning (DL) applications in communication systems have achieved great success. Thus, we take advantage of DL-based techniques and propose two autoencoders for Grassmannian constellation design. The proposed approaches are capable of learning better constellations than the conventional state-of-the-art.

Finally, we generalize several potential orthogonal frequency division multiplexing schemes to the SU multidimensional transmission. The associated multidimensional constellations (MDCs) design is studied. We discuss the important figure of merit (FoM) of MDCs. By taking advantage of the domain knowledge of MDCs, we propose a simple structure deep neural network (DNN)-based approach for MDC design. Simulation results show that the novel approach is capable of constructing MDC with good symbol error rate performance and FoMs.

Keywords : SCMA, EXIT chart, sphere decoding, NC MIMO, Grassmannian constellation, multidimensional codebook

Résumé de la thèse en Français

Chapitre 1 : Introduction

Avec l'évolution de la société et le développement rapide des technologies, les chercheurs et les ingénieurs ont porté leur attention sur la nouvelle génération de réseaux sans fil, la sixième génération (6G) de réseaux sans fil. Avec l'introduction de nouveaux scénarios et d'applications ainsi que l'exigence de flexibilité, de nouvelles technologies au niveau de la couche physique sont essentielles pour la 6G ; [1, 2]. Dans cette thèse, nous nous concentrons sur trois thématiques prometteuses pour la sixième génération de réseaux sans fil.

Comme les communications de type machine (MTC) sont en plein essor avec le développement de l'Internet des objets (IoT) en 6G, la connectivité massive et l'efficacité spectrale élevée deviennent deux exigences majeures. Ainsi, les techniques d'accès multiple non orthogonal (NOMA) sont proposées pour augmenter la connectivité des utilisateurs avec une quantité limitée de bande passante [3]. Parmi les techniques d'accès multiple non orthogonal, l'accès multiple par codes creux (SCMA) occupe une place de choix. Le schéma SCMA permet à un nombre d'utilisateurs supérieur aux nombre d'éléments de ressources (RE) orthogonaux de transmettre leurs signaux en exploitant les degrés de liberté dans le domaine du code. Plus précisément, dans la système SCMA, l'étalement creux réduit le nombre d'utilisateurs qui interfèrent sur le même élément de ressource. Au niveau du récepteur, des algorithmes efficaces se chargent de la détection multi-utilisateurs (MUD), généralement en exploitant la parcimonie intrinsèque de l'étalement des mots de code.

La technique de transmissions multi-antennas (MIMO) a obtenu un grand succès en raison de son efficacité spectrale significative sans nécessiter de puissance de transmission supplémentaire [4]. Étant donné que l'estimation précise de l'information sur l'état du canal (CSI) est irréaliste en présence d'un grand nombre antennes et en raison des exigences de connectivité de l'Internet of Everything

(IoE) en 6G, la transmission MIMO non-cohérente (NC), dans laquelle ni l'émetteur ni le récepteur n'ont besoin du CSI, est un système prometteur [5]. Dans cette thèse, nous considérons la transmission MIMO NC sur canaux à évanouissement par bloc. Suivant les propriétés des variétés Grassmanniennes, les constellations dans le système MIMO NC sont également appelées constellations Grassmanniennes [6]. Le détecteur optimal au niveau du récepteur est appelé récepteur GLRT (generalized likelihood ratio test) qui est basé sur la densité de probabilité conditionnelle des signaux reçus.

Durant cette dernière décennie, le multiplexage par répartition orthogonale de la fréquence (OFDM) a été retenu dans de nombreuses normes de communication sans fil car il est robuste au canal sélectif en fréquence et à l'interférence inter-symboles. Cependant, les performances de l'OFDM sont limitées car il n'exploite pas la diversité des canaux. La modulation OFDM-index (OFDM-IM) est une alternative intéressante à l'OFDM en termes d'efficacité spectrale et de ses performances [57]. D'autre part, le précodage par constellation linéaire groupé (GLCP) a été associé à l'OFDM afin d'exploiter la diversité due aux des trajets multiples et de réduire la complexité du décodage [7]. L'OFDM-IM ou le GLCP peuvent être vus comme une transmission multidimensionnelle à simple utilisateur(SU) où la conception optimale du dictionnaire du code/constellation multidimensionnelle (MDC) est un problème ouvert.

Alors que nous nous dirigeons vers la 6G, l'augmentation considérable du nombre de dispositifs, l'amélioration de l'efficacité et de l'adaptabilité du réseau mettent en question la manière conventionnelle de concevoir les réseaux sans fil [8]. D'autre part, l'apprentissage profond (DL) a obtenu des succès considérables dans d'autres domaines grâce à des avancées algorithmiques spécifiques et à une forte puissance de calcul matériel. Les techniques d'apprentissage profond devraient permettre de résoudre des problèmes tels que l'extraction d'informations sur les canaux inconnus ou l'optimisation globale du système. Récemment, les approches DL ont été appliquées à la couche physique dans le décodage de canal [10, 11], la détection de symbole [12, 13], l'estimation de canal [14], etc. En outre, il existe une tendance croissante à l'intégration des connaissances du domaine des communications sans fil dans les techniques DL [15, 16] pour la future 6G.

Par conséquent, nous étudions trois techniques prometteuses dans cette thèse qui sont la transmission multidimensionnelle SCMA, NC MIMO et SU pour les futurs systèmes sans fil 6G. Plus précisemment, nous nous concentrons sur la MUD dans la système SCMA, la construction de constellations Grassmanniennes dans la transmission NC MIMO et la conception MDC dans la transmission

multidimensionnelle SU. Par ailleurs, certaines applications DL efficaces pour les réseaux sans fil sont envisagées.

Chapitre 2 : État de l'art

SCMA transmission

Nous considérons un système SCMA pour liaison montante dans lequel J utilisateurs partagent K ($K < J$) REs orthogonaux pour la transmission du signal. Chaque utilisateur occupe d_v ($d_v < K$) REs. Nous supposons que le nombre d'utilisateurs liés à chaque RE est le même, noté $d_f = Jd_v/K$, et idéalement $d_f \ll J$ pour maintenir la caractéristique de parcimonie. Le facteur de charge de la transmission est noté $\lambda = J/K$. Au niveau de l'émetteur, la procédure d'association pour l'utilisateur $j \in \{1, 2, \cdots, J\}$ est exprimée par $g_j : \mathbf{b}_j \rightarrow \mathbf{d}_j$, où $\mathbf{b}_j \in \{0,1\}^{\log_2(M)\times 1}$ est le message binaire et $\mathbf{d}_j \in \mathbb{C}^{d_v \times 1}$ est le point de constellation de dimension d_v sélectionné dans une séquence $\mathcal{D}_j \subset \mathbb{C}^{d_v \times 1}$ de taille M. Ensuite, une matrice d'association binaire $\mathbf{S}_j$ de taille $K \times d_v$ répartit le point de constellation de d_v dimensions de l'utilisateur j sur les K REs, écrit sous la forme suivant :

$$\mathbf{x}_j = \mathbf{S}_j \mathbf{d}_j \tag{1}$$

qui est désigné par le terme " mot de code " de utilisateur j. Par conséquent, la taille du dictionnaire de code est de $|\mathcal{X}_j| = M$. $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_J$ désigne le dictionnaire de code multidimensionnel de la transmission SCMA. Le signal complexe reçu $\mathbf{y} \in \mathbb{C}^{K\times 1}$ au niveau du récepteur s'écrit comme suit

$$\mathbf{y} = \sum_{j=1}^{J} diag(\mathbf{h}_j)\mathbf{x}_j + \mathbf{n} \tag{2}$$

où $\mathbf{h}_j \in \mathbb{C}^{K\times 1}$ est le vecteur de canal de l'utilisateur j au récepteur, et $\mathbf{n} \in \mathbb{C}^{K\times 1}$ est le vecteur de bruit blanc gaussien additif (AWGN) dont les entrées suivent $\mathcal{N}_c(0, \sigma^2)$. En considérant l'équation (1), le signal reçu est réécrit comme suit

$$\mathbf{y} = \mathbf{G}\mathbf{d} + \mathbf{n} \tag{3}$$

où $\mathbf{G} = [\mathbf{g}_1, \cdots, \mathbf{g}_J] \in \mathbb{C}^{K\times Jd_v}$ est appelée matrice de gain de canal effectif, où $\mathbf{g}_j = diag(\mathbf{h}_j)\mathbf{S}_j$ est la matrice de gain de canal effectif de l'utilisateur j. $\mathbf{d} = [\mathbf{d}_1^\intercal, \cdots, \mathbf{d}_J^\intercal]^\intercal \in \mathcal{D}$ représente le vecteur de mots de code effectif transmis où $\mathcal{D} = \mathcal{D}_1 \times \cdots \times \mathcal{D}_J$ est le dictionnaire de code effectif. On remarque que l'équation (3) décrit la transmission SCMA sous la forme d'un système linéaire.

Dans l'hypothèse d'un CSI parfait au niveau du récepteur, la détection par maximum de vraisemblance (ML) de SCMA est donnée par la formule suivante

$$\hat{\mathbf{d}} = \arg\max_{\mathbf{d}\in\mathcal{D}} p(\mathbf{y}|\mathbf{d},\mathbf{G}) \tag{4}$$

$$= \arg\min_{\mathbf{d}\in\mathcal{D}} \|\mathbf{y}-\mathbf{G}\mathbf{d}\|^2 \tag{5}$$

Le détecteur ML peut fournir une performance MUD optimale en testant de manière exhaustive toutes les combinaisons des mots de code effectifs transmis. Cependant, il devient infaisable lorsque le nombre d'utilisateurs ou la taille du dictionnaire de codes est grande. En raison de la parcimonie de la structure, la detection conjointe peut être réalisée à l'aide de l'algorithme de propagation des messages (MPA). Après avoir mis à jour de manière itérative les messages de probabilité entre les utilisateurs et les REs, le MPA détermine les signaux décodés avec la probabilité *a posteriori* maximale. Le MPA avec un grand nombre d'itérations peut atteindre une performance MUD optimale mais une complexité de calcul élevée. Pour réduire la complexité, l'algorithme expectation-propagation (EPA) et l'algorithme d'approximation gaussienne (GAA) pour le SCMA sont proposés, qui montrent une performance proche du MPA mais avec une réduction significative de la complexité dans un système codé [17, 18].

D'autre part, les détecteurs basés sur le décodage de sphères (SD) pour SCMA ont récemment attiré beaucoup d'attention. Dans [19, 20], les auteurs appliquent le SD au SCMA MUD sous réserve d'une certaine structure du dictionnaire de codes. Afin de préserver ne pas avoir de restriction du dictionnaire de codes, ils proposent le SD pour SCMA (SD-SCMA), indépendant de la structure du dictionnaire de codes, mais celui-ci ne s'applique qu'aux dictionnaires de codes à module constant. Pour surmonter cette contrainte, les auteurs de [22] proposent un SD-SCMA général (GSD-SCMA) qui effectue une recherche par recherche exhaustive sur les mots de code effectifs transmis partiellement, ce qui résulte en une complexité de calcul très élevée lorsque la taille du dictionnaire de codes ou le facteur de charge est grand.

NC MIMO transmission

Nous considérons un système de communication MIMO NC où l'émetteur et le récepteur ont N_t et N_r antennes, respectivement. Nous faisons l'hypothèse que les coefficients du canal sont constants pendant un intervalle de cohérence de T et changent aléatoirement à chaque nouvelle réalisation. Nous supposons que $N_t = \min\{\lfloor \frac{T}{2} \rfloor, N_r\}$. Le signal transmis est représenté par la matrice de symboles

transmis $\mathbf{X} \in \mathbb{C}^{T \times N_t}$ tiré du dictionnaire de codes $\mathcal{C}$ ($|\mathcal{C}| = M$). Le signal reçu $\mathbf{Y} \in \mathbb{C}^{T \times N_r}$ est exprimé comme suit

$$\mathbf{Y} = \mathbf{X}\mathbf{H} + \sqrt{\frac{N_t}{\rho T}}\mathbf{W} \tag{6}$$

où $\mathbf{H} \in \mathbb{C}^{N_t \times N_r}$ est la matrice de canal, ρ est le rapport signal/bruit (SNR) par antenne de réception et $\mathbf{W} \in \mathbb{C}^{T \times N_r}$ est la matrice AWGN. Les éléments de la matrice $\mathbf{W}$ sont tirés indépendamment dans $\mathcal{N}_c(0, 1)$. La puissance du signal transmis est la suivante

$$\sum_{t=1}^{T} \mathbb{E}[|x_{t,n}|^2] = 1 \tag{7}$$

où $x_{t,n}$ est l'élément de la $t^{\text{ème}}$ ligne et de la $n^{\text{ème}}$ colonne de la matrice $\mathbf{X}$. Comme le récepteur de NC MIMO n'a aucune connaissance préalable de CSI, la détection ML est appelée GLRT, exprimée par la formule suivante [23] :

$$\hat{\mathbf{X}} = \arg\max_{\mathbf{X} \in \mathcal{C}} \mathrm{Tr}\{\mathbf{Y}^\dagger \mathbf{X}\mathbf{X}^\dagger \mathbf{Y}\} \tag{8}$$

Comme analysé dans [24], pour atteindre la capacité du système MIMO NC, la matrice de symbole transmise $\mathbf{X}$ doit respecter la contrainte

$$\mathbf{X}^\dagger \mathbf{X} = \mathbf{I}_{N_t} \tag{9}$$

Chaque matrice de symbole transmise $\mathbf{X}$ représente un point individuel sur la variété Grassmannienn. Par conséquent, les constellations dans les systèmes MIMO NC sont également appelées constellations Grassmanniennes.

Il existe deux approches pour concevoir des constellations grassmanniennes. La première approche consite à utiliser des outils d'optimisation numérique pour résoudre le problème d'empilement de sphère [6] en maximisant la distance minimale par paire de symboles [25, 26, 27, 28] ou en minimisant directement la borne supérieure de la probabilité d'erreur [29]. Dans la seconde approche, nous introduisons des structures particulières dans la conception de la constellation grassmannienne. Par exemple, les auteurs du document [30] proposent une association paramétrée de matrices unitaires pour la conception de constellations grassmanniennes.

Transmission multidimensionelle simple utilisateur

En simplifiant le modèle de transmission basé sur l'OFDM et en tirant parti du gain de diversité, nous considérons une transmission multidimensionnelle SU dans le domaine réel, exprimée comme suit

$$\mathbf{y} = diag(|\tilde{\mathbf{h}}|)\mathbf{x} + \mathbf{n} \tag{10}$$

où $\mathbf{x} \in \mathbb{R}^{2N\times 1}$ est le mot de code transmis tiré du dictionnaire de code $\mathcal{C}$ de taille M, $\tilde{\mathbf{h}} \in \mathbb{C}^{2N\times 1}$ est le vecteur de canal complexe suivant $\tilde{\mathbf{h}} \sim \mathcal{N}_c(\mathbf{0}_{2N\times 1}, \mathbf{I}_{2N})$, $|\cdot|$ représente l'opérateur valeur absolu par éléments pour les vecteurs, $\mathbf{y} \in \mathbb{R}^{2N\times 1}$ est le signal reçu et $\mathbf{n} \sim \mathcal{N}(\mathbf{0}_{2N\times 1}, \frac{N_0}{2}\mathbf{I}_{2N})$ est le vecteur AWGN. $\frac{N_0}{2}$ est la puissance du bruit par dimension réelle. La contrainte de puissance du mot de code est définie comme suit

$$E_s = \mathbb{E}\left[\frac{\|\mathbf{x}\|^2}{2N}\right] = \frac{1}{2} \tag{11}$$

Si nous supposons l'ICS parfaite au niveau du récepteur, le détecteur ML d'une transmission multidimensionnelle SU s'écrit comme suit

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}\in\mathcal{C}} p(\mathbf{y}|\mathbf{x}, \tilde{\mathbf{h}}) \tag{12}$$

La conception des MDCs est essentielle pour les performances du système et son optimisation reste une question ouverte. Les auteurs de [31] justifient l'importance d'exploiter la diversité de l'espace du signal (SSD) pour les MDCs afin d'obtenir des gains de diversité sur les canaux à évanouissement de Rayleigh. Pour exploiter les gains de diversité, la rotation de constellation a été prise en compte dans la conception des MDC, où une recherche exhaustive a été appliquée pour trouver l'angle de rotation optimal [31]. D'autres chercheurs combinent le schéma GLCP et la rotation pour obtenir une diversité totale. En outre, une approche combinatoire est proposée pour la construction des MDCs à diversité entière qui nécessite peu de stockage.

Chapitre 3 : Détecteurs itératifs pour systèmes SCMA

Scénario transmission SCMA simple

Si le mot de code de l'utilisateur j peut être déterminé par un vecteur générateur $\mathbf{v}_j = [v_{1,j}, v_{2,j}, \cdot, v_{K,j}]^T$ [36], il s'exprime sous la forme $\mathbf{x}_j = \mathbf{v}_j \cdot s_j$ où s_j est un symbole complexe tiré de la constellation sélectionnée $\mathcal{S}$ ($|\mathcal{S}| = M$). Par conséquent, le signal reçu dans Eq.(2) est reformulé en

$$\mathbf{y} = \sum_{j=1}^{J} diag(\mathbf{h}_j)\mathbf{v}_j \cdot s_j + \mathbf{n} \tag{13}$$

$$= \sum_{j=1}^{J} \widetilde{\mathbf{h}}_j \cdot s_j + \mathbf{n} \tag{14}$$

où $\widetilde{\mathbf{h}}_j = \mathbf{h}_j \odot \mathbf{v}_j = [h_{1,j}v_{1,j}, \ldots, h_{K,j}v_{K,j}]$.

Contrairement au détecteur MPA qui met à jour itérativement les messages de probabilité entre les utilisateurs et les REs, le détecteur GAA échange itérativement la moyenne et la variance des mes-

sages en approximant les probabilités discrètes des messages par une distribution gaussienne continue. Pour réduire encore la complexité de calcul du détecteur GAA, nous proposons un détecteur GAA avec approximation au premier ordre (FO-GAA) qui est résumée dans Algorithm 1. Le détecteur à faible complexité proposé calcule la moyenne et la variance par utilisateur et par RE, ce qui réduit considérablement sa complexité.

Algorithm 1 FO-GAA.

Initialisation : $\zeta_{v_j}^0 = 0$, $\nu_{f_k}^0 = 1000$, $z_{f_k}^0 = 0$, $\gamma_{v_j}^0 = 1000$, $t = 1$.
while $t \leqslant T$ **do**
  for $j = 1 \to V$ **do**
    Calculer le message $\mu_{v_j}^t(s_j) = \frac{\mu_{\phi_j \to v_j}(s_j)\mathcal{N}_c(s_j;\zeta_{v_j}^{t-1},\gamma_{v_j}^{t-1})}{\sum_{s_j \in \mathcal{S}} \mu_{\phi_j \to v_j}(s_j)\mathcal{N}_c(s_j;\zeta_{v_j}^{t-1},\gamma_{v_j}^{t-1})}$.
    Calculer $\hat{x}_{v_j}^t = \mathbb{E}_{\mu_{v_j}^t}[s_j]$ et $\hat{\tau}_{v_j}^t = \mathbb{E}_{\mu_{v_j}^t}[|s_j|^2] - |\hat{x}_{v_j}^t|^2$.
  end for
  for $k = 1 \to F$ **do**
    Calculer $\nu_{f_k}^t = \sigma_n^2 + \sum_{j \in F(k)} |\widetilde{h}_{k,j}|^2 \cdot \hat{\tau}_{v_j}^t$ et $z_{f_k}^t \approx y_k - \sum_{j \in F(k)} \widetilde{h}_{k,j}\hat{x}_{v_j}^t + z_{f_k}^{t-1} \frac{\sum_{j \in F(k)} \hat{\tau}_{v_j}^t |\widetilde{h}_{k,j}|^2}{\nu_{f_k}^{t-1}}$.
  end for
  for $j = 1 \to V$ **do**
    Calculer $\gamma_{v_j}^t = \left(\sum_{k \in V(j)} \frac{|\widetilde{h}_{k,j}|^2}{\nu_{f_k}^t}\right)^{-1}$ et $\zeta_{v_j}^t \approx \hat{x}_{v_j}^t + \gamma_{v_j}^t \sum_{k \in V(j)} \frac{\widetilde{h}_{k,j}^* z_{f_k}^t}{\nu_{f_k}^t}$.
  end for
  $t = t + 1$.
end while
for $j = 1 \to V$ **do**
  Calculer les LLR basés sur $\Lambda_1^p(c_j^i) = \ln \frac{\sum_{s_j \in \mathcal{S}^+} \mu_{\phi_j \to v_j}(s_j)\mathcal{N}_c(s_j;\zeta_{v_j}^T,\gamma_{v_j}^T)}{\sum_{s_j \in \mathcal{S}^-} \mu_{\phi_j \to v_j}(s_j)\mathcal{N}_c(s_j;\zeta_{v_j}^T,\gamma_{v_j}^T)}$.
end for

Fig. 1 montre la performance du taux d'erreur binaire (BER) des différents détecteurs SCMA itératifs après convergence. Pour montrer l'influence du nombre d'itérations sur le détecteur GAA, nous présentons également les courbes de BER sans atteindre la convergence, marqué "GAA nc" dans la figure. Nous voyons que le FO-GAA proposé a seulement 0.2dB de perte de performance comparée au MPA à SNR = 2dB.

Fig. 2 compare la complexité de calcul des quatre détecteurs itératifs évalués par le nombre d'opérations en virgule flottante (FLOPs) avec $J = 6$, $K = 4$, $d_v = 2$, $d_f = 3$. Comme le détecteur GAA nécessite un plus grand nombre d'itérations pour atteindre la convergence, il a une complexité de

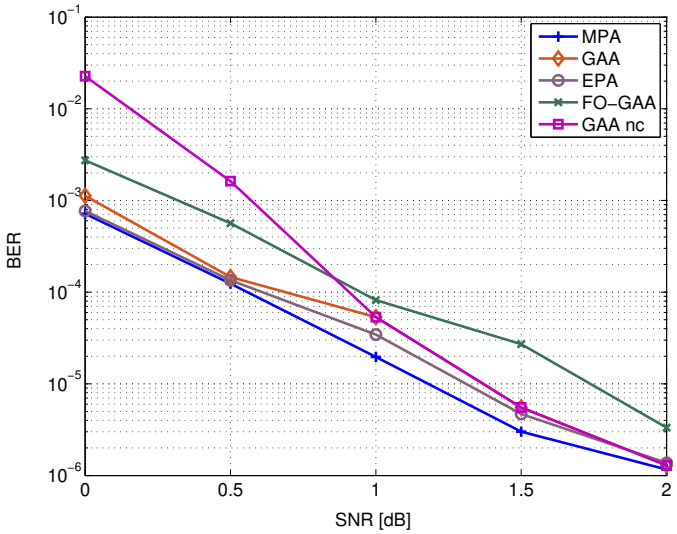


FIGURE 1 – Comparaison des performances BER.

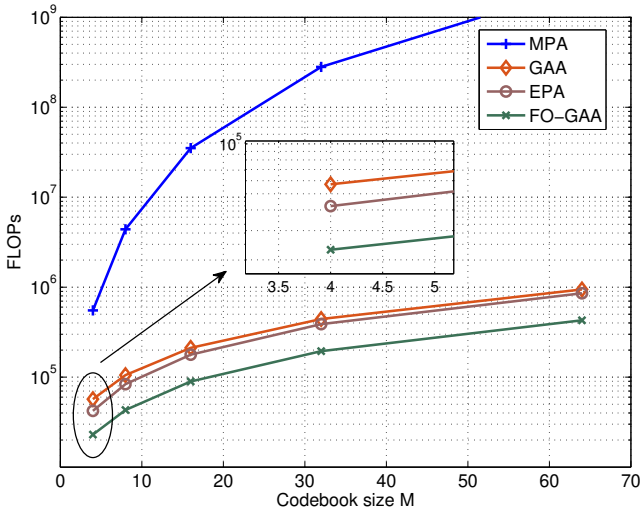


FIGURE 2 – Comparaison du nombre de FLOPs lorsque SNR = 0dB.

calcul plus élevée que les détecteurs EPA et FO-GAA. Le nombre de FLOPs du FO-GAA proposé est le plus faible et tombe à 54.6% de celui du détecteur EPA et 4.2% de celui du détecteur MPA quand $M = 4$. La faible complexité de calcul et le bon comportement de convergence du FO-GAA proposé compensent la perte de performance en terme de BER.

Scénario de transmission SCMA avec deux groupes

Afin de satisfaire les demandes de connectivité massive et de haute efficacité spectrale pour les futurs réseaux sans fil, nous avons proposé un système de transmission à deux groupes sur la liaison montante où la technique NOMA dans le domaine de la puissance (PD-NOMA) et le SCMA sont utilisés en même temps. Nous avons considéré un système à cellule unique dans lequel deux groupes d'utilisateurs, au total $2J$ utilisateurs, transmettent leurs signaux à la station de base (BS) au centre en partageant K REs. Ainsi, le facteur de charge de cette transmission est égal à $\frac{2J}{K}$. Le groupe dont les utilisateurs sont les plus proches de la station de base est désigné par *Group near* et l'autre par *Group far*. Les SNR des groupes représentent les SNR de tous les utilisateurs des groupes correspondants, notés SNR^{near} et SNR^{far}. Les deux groupes partagent le même dictionnaire de codes multidimensionnels : $\mathcal{C} = \mathcal{C}_1 \times \mathcal{C}_2 \cdots \times \mathcal{C}_J$ ($|\mathcal{C}_j| = M$). Le mot de code de l'utilisateur j_1 dans le *Group near* est $\mathbf{x}_{j_1}^{near} \in \mathcal{C}_{j_1}$ et celui de l'utilisateur j_2 dans le *Group far* est $\mathbf{x}_{j_2}^{far} \in \mathcal{C}_{j_2}$. Les évanouissements à petite échelle entre l'utilisateur j_1 et l'utilisateur j_2 vers la station de base sont respectivement de $\mathbf{h}_{j_1}^{near}$ et $\mathbf{h}_{j_2}^{far}$. Le signal reçu par la station de base s'exprime comme suit

$$\mathbf{y} = \sum_{j_1=1}^{J} diag(\mathbf{h}_{j_1}^{near})\sqrt{P_n}\mathbf{x}_{j_1}^{near} + \sum_{j_2=1}^{J} diag(\mathbf{h}_{j_2}^{far})\sqrt{P_f}\mathbf{x}_{j_2}^{far} + \mathbf{n} \tag{15}$$

$$= \mathbf{y}^{near} + \mathbf{y}^{far} + \mathbf{n} \tag{16}$$

où P_n et P_f sont la puissance reçue par la station de base de la part des utilisateurs du *Groupe near* et du *Groupe far* respectivement, qui prennent en compte la puissance d'émission et l'évanouissement à grande échelle, et $\mathbf{n} = [n_1, n_2, \ldots, n_K]^\intercal$ est le vecteur AWGN ayant des entrées $n_k \sim \mathcal{CN}(0, \sigma^2)$.

La structure du récepteur du système proposé est illustrée dans Fig. 3 où le détecteur SCMA supprime l'interférence intra-groupe et le SIC effectue l'annulation de l'interférence inter-groupe. Comme analysé dans [18], la performance BER du système proposé dépend de la détection SCMA, du rendement du codage canal ainsi que du nombre d'itérations de la boucle interne et externe. Plus précisément, le choix d'un code canal avec un rendement plus faible ou un plus grand nombre d'itérations accélère le processus de convergence du système proposé, mais diminue inévitablement l'efficacité spectrale.

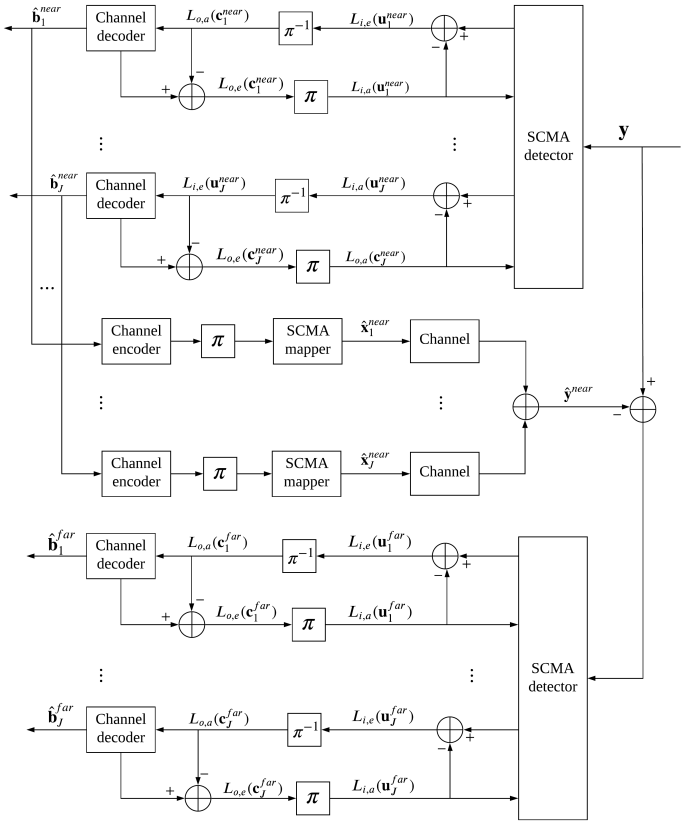


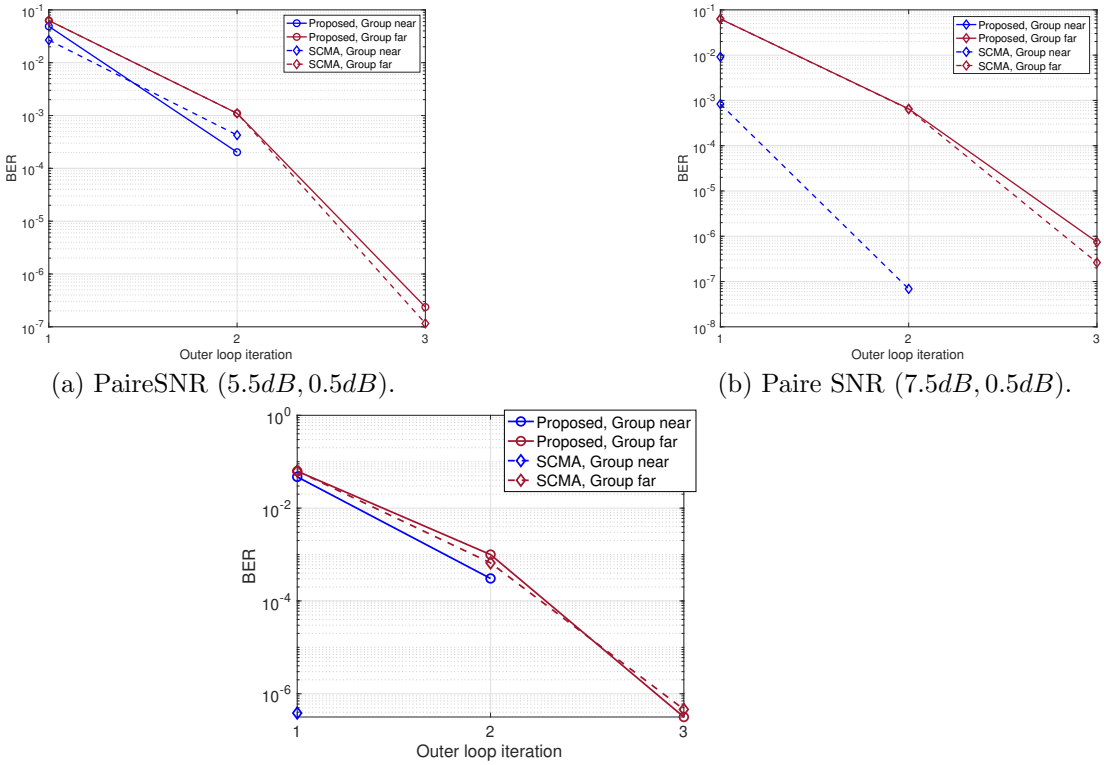
FIGURE 3 – Structure du récepteur du système proposé.

Compte tenu de ce compromis essentiel entre le comportement de convergence du système et l'efficacité spectrale, nous proposons une méthode d'optimisation pour choisir les rendements optimaux des codes des deux groupes individuellement. Comme le récepteur du système de transmission à deux groupes peut être considéré comme une structure de démapping et de décodage itératif (IDD), les graphiques de transfert d'information extrinsèque (EXIT) peuvent être utilisés comme un outil pour analyser son comportement de convergence MUD. Par conséquent, la stratégie de la méthode d'optimisation que nous proposons consiste, premièrement, à fixer l'objectif de convergence (c'est-à-dire le nombre d'itérations de la boucle externe), deuxièmement, à obtenir les diagrammes EXIT des deux groupes avec différents SNR et rendement de code de canal, et enfin, à choisir les rendements de code pour les groupes en fonction de l'objectif de convergence.

Dans la simulation, nous utilisons les codes LDPC (low-density parity-check) suivant le standard 5G [38] et le détecteur MPA. Pour évaluer la performance du système, une comparaison de deux systèmes, le système proposé et le système SCMA conventionnel, est étudiée. En appliquant la méthode d'optimisation proposée, les rendements des codes LDPC choisis des deux groupes dans les deux

TABLE 1 – Les rendements des codes LDPC choisis pour les deux systèmes.

	Système proposé		Système SCMA conventionnel	
Paires SNR/dB	R^{near}	R^{far}	R^{near}	R^{far}
(5.5, 0.5)	1/3	1/2	11/13	1/2
(7.5, 0.5)	2/5	1/2	8/9	1/2
(9.5, 0.5)	11/18	1/2	8/9	1/2



(a) PaireSNR $(5.5dB, 0.5dB)$. (b) Paire SNR $(7.5dB, 0.5dB)$.

(c) Paire SNR $(9.5dB, 0.5dB)$.

FIGURE 4 – Résultats BER des deux groupes dans différents scénarios SNR.

systèmes pour différents scénarios de SNR sont donnés dans Table 1. Sur la base de ces résultats, Fig. 4 compare la performance BER des deux groupes dans différents scénarios SNR. Nous pouvons voir que la performance BER du groupe de loin est presque la même dans les deux systèmes. La méthode d'optimisation proposée peut être considérée comme efficace.

L'efficacité spectrale du système est définie par $SE = \lambda(R^{near} + R^{far}) \log_2(M)$ bits/ton [39]. Le facteur de charge du système proposé est de $\lambda^P = 300\%$, tandis que celui du système SCMA est de $\lambda^S = 150\%$. Fig. 5 illustre l'efficacité spectrale des différents systèmes de manière distincte par groupes. Il est remarquable que le système proposé présente une amélioration substantielle de l'efficacité spectrale par rapport au système SCMA conventionnel. L'efficacité spectrale du système SCMA conventionnel

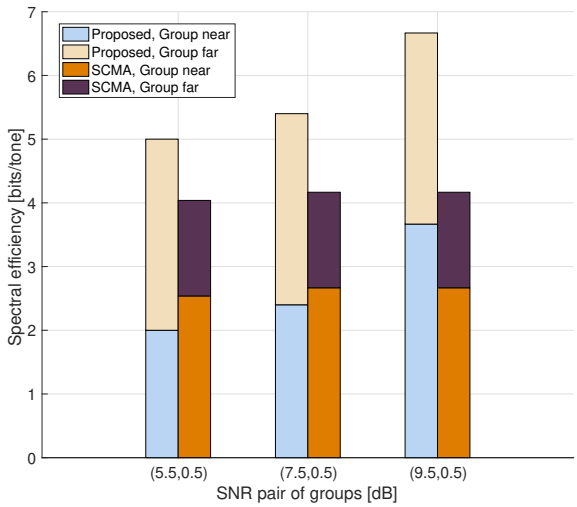


FIGURE 5 – Comparaison de l'efficacité spectrale de deux systèmes.

sature lorsque la paire de groupes SNR est $(9,5\text{dB}, 0,5\text{dB})$ alors que celle du système proposé continue d'augmenter. Cela implique que le système proposé avec la méthode d'optimisation améliore davantage l'efficacité spectrale par rapport au schéma SCMA conventionnel avec les paramètres considérés.

Chapitre 4 : Détecteurs SD pour systèmes SCMA

Dans ce chapitre, nous nous concentrons sur les détecteurs basés sur le SD pour les systèmes SCMA. Pour faciliter l'application de SD pour SCMA, nous récrivons le modèle de système à valeurs complexes exprimé par l'équation (3) dans le domaine réel comme suit

$$\bar{\mathbf{y}} = \bar{\mathbf{G}}\bar{\mathbf{d}} + \bar{\mathbf{n}} \tag{17}$$

où $\bar{\mathbf{y}} \in \mathbb{R}^{2K\times 1}$, $\bar{\mathbf{G}} \in \mathbb{R}^{2K\times 2Jd_v}$, $\bar{\mathbf{d}} \in \mathbb{R}^{2Jd_v\times 1}$ et $\bar{\mathbf{n}} \in \mathbb{R}^{2K\times 1}$ sont respectivement le vecteur de signal reçu à valeur réelle, la matrice de canal effective, le vecteur de symbole transmis effectif et le vecteur AWGN qui sont construits à partir de leurs valeurs complexes. Pour la simplicité, nous appellerons $\bar{\mathbf{G}}$ la matrice de canal et $\bar{\mathbf{d}}$ le vecteur de symbole transmis. En outre, nous définissons un dictionnaire de codes intermédiaires à valeurs réelles, $\bar{\mathcal{D}}_j$, dont les entrées sont construites à partir des entrées complexes de $\mathcal{D}_j$. Par conséquent, il existe $\bar{\mathbf{d}} \in \bar{\mathcal{D}} = \bar{\mathcal{D}}_1 \times \bar{\mathcal{D}}_2 \cdots \times \bar{\mathcal{D}}_J$. En conséquence, la détection ML pour SCMA dans le domaine réel est formulée de la manière suivante

$$\hat{\bar{\mathbf{d}}} = \arg\min_{\bar{\mathbf{d}}\in\bar{\mathcal{D}}} \|\bar{\mathbf{y}} - \bar{\mathbf{G}}\bar{\mathbf{d}}\|^2 \tag{18}$$

Comme pour un détecteur de SCMA basé sur le SD, nous réécrivons la matrice de canal comme

suit

$$\begin{aligned}\bar{\mathbf{G}} &= \begin{bmatrix}\bar{\mathbf{G}}^{(1)} & \bar{\mathbf{G}}^{(2)}\end{bmatrix} \\ &= \mathbf{Q}_1 \begin{bmatrix}\mathbf{R}_1\mathbf{P}_1^{-1} & \mathbf{Q}_1^{-1}\bar{\mathbf{G}}^{(2)}\end{bmatrix} \\ &= \mathbf{Q}_1 \begin{bmatrix}\mathbf{R}_1\mathbf{P}_1^{-1} & \mathbf{R}_2\end{bmatrix}\end{aligned} \tag{19}$$

où la matrice $\bar{\mathbf{G}}^{(1)}$ contient les $2K$ premières colonnes de la matrice $\bar{\mathbf{G}}$ de taille $2K \times 2K$ et la matrice $\bar{\mathbf{G}}^{(2)}$ contient la matrice $\bar{\mathbf{G}}$ de la $(2K+1)^{\text{ème}}$ colonne à la dernière colonne de taille $2K \times (2Jd_v - 2K)$. La même partition est appliquée au vecteur $\bar{\mathbf{d}}$, ce qui signifie que $\bar{\mathbf{d}}^{(1)}$ possède les $2K$ premiers éléments du vecteur $\bar{\mathbf{d}}$ et que $\bar{\mathbf{d}}^{(2)}$ possède les $2Jd_v - 2K$ derniers éléments. Par conséquent, le problème de décodage devient

$$\hat{\bar{\mathbf{d}}} = \arg \min_{\substack{\bar{\mathbf{d}}^{(1)} \in \bar{\mathcal{D}}^{(1)} \\ \bar{\mathbf{d}}^{(2)} \in \bar{\mathcal{D}}^{(2)}}} \|\mathbf{Q}_1^{\intercal}\bar{\mathbf{y}} - \mathbf{R}_1\mathbf{P}_1^{-1}\bar{\mathbf{d}}^{(1)} - \mathbf{R}_2\bar{\mathbf{d}}^{(2)}\|^2 \tag{20}$$

$$\text{soumis à } \|\mathbf{Q}_1^{\intercal}\bar{\mathbf{y}} - \mathbf{R}_1\mathbf{P}_1^{-1}\bar{\mathbf{d}}^{(1)} - \mathbf{R}_2\bar{\mathbf{d}}^{(2)}\|^2 \leqslant d^2 \tag{21}$$

où $\bar{\mathbf{y}}' = \mathbf{Q}_1^{\intercal}\bar{\mathbf{y}} - \mathbf{R}_2\bar{\mathbf{d}}^{(2)}$. Bien que le détecteur GSD-SCMA soit applicable à tous les dictionnaires de codes SCMA, il présente une complexité de calcul trés élevée en raison de la recherche exhaustive de toutes les valeurs de $M^{J'}$ du vecteur $\bar{\mathbf{d}}^{(2)}$ avec $J' = J(1 - 1/d_f)$.

Pour réduire efficacement la complexité de calcul du détecteur GSD-SCMA, nous proposons le détecteur GSD-SCMA simplifié (SGSD-SCMA) qui bénéficie de la séquence réduite $\mathcal{L}_2^*$ de valeurs $\bar{\mathbf{d}}^{(2)}$. Le détecteur SGSD-SCMA proposé se compose de deux étapes. Dans l'étape 1, un algorithme d'élagage obtient la séquence réduite $\mathcal{L}_2^*$, qui contient des valeurs potentiellement bonnes du vecteur $\bar{\mathbf{d}}^{(2)}$. Ensuite, à l'étape 2, le SD de $\bar{\mathbf{d}}^{(1)}$ est effectué sur la base de la séquence $\mathcal{L}_2^*$ semblable à celle du détecteur GSD-SCMA. Par conséquent, le problème de décodage du détecteur SGSD-SCMA est le suivant

$$\hat{\bar{\mathbf{d}}} = \arg \min_{\bar{\mathbf{d}}^{(2)} \in \mathcal{L}_2^*} \min_{\bar{\mathbf{d}}^{(1)\mathtt{p}} \in \bar{\mathcal{D}}^{(1)\mathtt{p}}} \|\bar{\mathbf{y}}' - \mathbf{R}_1\bar{\mathbf{d}}^{(1)\mathtt{p}}\|^2 \tag{22}$$

$$\text{soumis à } \|\bar{\mathbf{y}}' - \mathbf{R}_1\bar{\mathbf{d}}^{(1)\mathtt{p}}\|^2 \leqslant d^2 \tag{23}$$

Puisque $\mathbf{Q}$ est une matrice diagonale et que $\bar{\mathbf{G}}^{(2)}$ est une matrice creuse, $\mathbf{R}_2$ est également une

matrice creuse. Par exemple, la matrice $\mathbf{R}_2$ peut avoir la structure suivante

$$\mathbf{R}_2 = \begin{bmatrix} 0\,0\,0\,0\,\mathtt{x}\,\mathtt{x}\,0\,0\,\mathtt{x}\,\mathtt{x}\,0\,0\,0\,0\,0\,0 \\ 0\,0\,0\,0\,\mathtt{x}\,\mathtt{x}\,0\,0\,\mathtt{x}\,\mathtt{x}\,0\,0\,0\,0\,0\,0 \\ \mathtt{x}\,\mathtt{x}\,0\,0\,0\,0\,0\,0\,0\,0\,0\,\mathtt{x}\,\mathtt{x}\,0\,0 \\ \mathtt{x}\,\mathtt{x}\,0\,0\,0\,0\,0\,0\,0\,0\,0\,\mathtt{x}\,\mathtt{x}\,0\,0 \\ 0\,0\,0\,0\,0\,0\,\mathtt{x}\,\mathtt{x}\,0\,0\,0\,0\,0\,0\,\mathtt{x}\,\mathtt{x} \\ 0\,0\,0\,0\,0\,0\,\mathtt{x}\,\mathtt{x}\,0\,0\,0\,0\,0\,0\,\mathtt{x}\,\mathtt{x} \\ 0\,0\,\mathtt{x}\,\mathtt{x}\,0\,0\,0\,0\,0\,0\,\mathtt{x}\,\mathtt{x}\,0\,0\,0\,0 \\ 0\,0\,\mathtt{x}\,\mathtt{x}\,0\,0\,0\,0\,0\,0\,\mathtt{x}\,\mathtt{x}\,0\,0\,0\,0 \end{bmatrix} \tag{24}$$

où $\mathtt{x}$ représente un élément non nul de la matrice. Grâce à la parcimonie de la matrice $\mathbf{R}_2$ illustrée ci-dessus, pour une ligne donnée, on ne considère que les éléments de $\bar{\mathbf{d}}^{(2)}$ qui sont associés à l'élément non nul de $\mathbf{R}_2$. Il y a toujours $2(d_f - 1)$ éléments non nuls dans chaque vecteur de ligne $(\mathbf{r}_2)_i$ ($i \in \{1, \cdots, 2K\}$) qui correspondent au même utilisateur par paire. Par conséquent, au lieu de tester chaque valeur de $\bar{\mathbf{d}}^{(2)}$ ($\bar{\mathbf{d}}^{(2)} \in \bar{\mathcal{D}}^{(2)}$), nous pouvons tester uniquement M^{d_f-1} combinaisons des éléments de $\bar{\mathbf{d}}^{(2)}$ qui sont associés à l'élément non nul de $\mathbf{R}_2$ à chaque rang. De plus, les $(2k-1)^{\text{ème}}$ et $2k^{\text{ème}}$ ($1 \leqslant k \leqslant K$) lignes de $\mathbf{R}_2$ ont les mêmes indices d'éléments non nuls. Ainsi, nous effectuons l'opération d'élagage de deux rangées, la $(2k-1)^{\text{ème}}$ et la $2k^{\text{ème}}$, en une seule fois. En conséquence, nous proposons l'algorithme PRUN1 qui effectue toujours K opérations d'élagage présentées dans Algorithm 2. Dans l'algorithme d'élagage, δ est déterminé par le bruit additif et la probabilité de détection d'erreur prédéfinie p est calculée comme suit

$$\delta = \mathcal{F}^{-1}(1 - p, 0, \sigma) \tag{25}$$

où $\mathcal{F}^{-1}(x, \mu, \sigma')$ désigne la fonction de distribution cumulative inverse de la distribution gaussienne avec une moyenne de μ et un écart type de σ'. Le détecteur proposé avec l'algorithme PRUN1 est nommé SGSD1-SCMA.

Nous analysons les performances du détecteur SGSD1-SCMA en considérant la probabilité de détection des erreurs p comme paramètre. Le taux d'erreurs sur les mots de code (CER) du détecteur SGSD1-SCMA peut être approximé comme suit

$$Pr\{\hat{\bar{\mathbf{d}}} \neq \bar{\mathbf{d}}\} \approx P_{\text{ML}} + Pr\{\bar{\mathbf{d}}^{(2)} \notin \mathcal{L}_2^*\} \tag{26}$$

où P_{ML} est le CER du détecteur ML Eq.(18) et $Pr\{\bar{\mathbf{d}}^{(2)} \notin \mathcal{L}_2^*\}$ est le taux d'erreur de l'algorithme PRUN1. Pour des valeurs SNR élevées, puisque P_{ML} s'approche de zéro, le CER peut s'écrire comme suit

$$\lim_{\text{SNR}\to\infty} Pr\{\hat{\bar{\mathbf{d}}} \neq \bar{\mathbf{d}}_{\text{T}}\} = Pr\{\bar{\mathbf{d}}_{\text{T}}^{(2)} \notin \mathcal{L}_2^*\} \tag{27}$$

Algorithm 2 L'algorithme PRUN1.

Input : $\bar{\mathbf{y}}$, $\mathbf{R}_1$, $\mathbf{R}_2$, $\mathcal{D}^{(1)}$ et $\mathcal{D}^{(2)}$
Output : Υ

Initialisation : $\mathcal{V} = \emptyset$, $\Upsilon = (\mathbf{0}_{J' \times 1})$ et $L_1 = 1$.

for $k = K, K-1, \cdots, 1$ **do**
  $\mathcal{U}$, $\mathcal{W}$, $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{B}_1$, $\mathcal{B}_2 = \text{UserPart}((\mathbf{r}_2)_{2k}, \mathcal{V})$
  $l_1 = \text{size}(\mathcal{U})$
  $l_2 = d_f - 1 - l_1$
  if $\mathcal{W} \neq \emptyset$ **then**
    for $i_1 = 1, 2, \cdots, L_1$ **do**
      Obtenir $\mathbf{q} = \Upsilon^{\mathcal{W}}(i_1)$.
      Calculer $t_1(i_1) = \sum_{l=1}^{2l_2} (r_2)_{2k-1, \mathcal{A}_2(l)} \bar{\mathcal{D}}^{\mathcal{A}_2(l) - 2d_v(\mathcal{B}_2(l)-1)}_{\frac{K}{d_v} + \mathcal{B}_2(l)}(q_l)$
      et $t_2(i_1) = \sum_{l=1}^{2l_2} (r_2)_{2k, \mathcal{A}_2(l)} \bar{\mathcal{D}}^{\mathcal{A}_2(l) - 2d_v(\mathcal{B}_2(l)-1)}_{\frac{K}{d_v} + \mathcal{B}_2(l)}(q_l)$.
    end for
  else
    $t_1(1) = t_2(1) = 0$
  end if
  $\Upsilon' = \emptyset$
  $\mathbf{P} = \mathcal{F}_1(M, l_1)$
  $c_1 = 0$
  for $i_2 = 1, 2, \cdots, M^{l_1}$ **do**
    Calculer $s_1(i_2) = \bar{y}_{2k-1} - \sum_{l=1}^{2l_1} (r_2)_{2k-1, \mathcal{A}_1(l)} \bar{\mathcal{D}}^{\mathcal{A}_1(l) - 2d_v(\mathcal{B}_1(l)-1)}_{\frac{K}{d_v} + \mathcal{B}_1(l)}((p_{i_2})_{\lceil \frac{l}{2} \rceil})$
    et $s_2(i_2) = \bar{y}_{2k} - \sum_{l=1}^{2l_1} (r_2)_{2k, \mathcal{A}_1(l)} \bar{\mathcal{D}}^{\mathcal{A}_1(l) - 2d_v(\mathcal{B}_1(l)-1)}_{\frac{K}{d_v} + \mathcal{B}_1(l)}((p_{i_2})_{\lceil \frac{l}{2} \rceil})$.
    for $i_1 = 1, 2, \cdots, L_1$ **do**
      Calculer $\xi_1 = s_1(i_2) - t_1(i_1)$ et $\xi_2 = s_2(i_2) - t_2(i_1)$.
      if
      $(r_1)_{2k-1,2k-1} \min\left((\bar{\mathcal{D}}_j^{(1)\mathtt{p}})^{2k-1-2d_v(j-1)}\right) - \delta \leqslant \xi_1 \leqslant (r_1)_{2k-1,2k-1} \max\left((\bar{\mathcal{D}}_j^{(1)\mathtt{p}})^{2k-1-2d_v(j-1)}\right) + \delta$ et $(r_1)_{2k,2k} \min\left((\bar{\mathcal{D}}_j^{(1)\mathtt{p}})^{2k-2d_v(j-1)}\right) - \delta \leqslant \xi_2 \leqslant (r_1)_{2k,2k} \max\left((\bar{\mathcal{D}}_j^{(1)\mathtt{p}})^{2k-2d_v(j-1)}\right) + \delta$ sont satisfaites **then**
        $\mathbf{v} = \Upsilon(i_1)$
        for $i = 1, 2, \cdots, l_1$ **do**
          $v_{\mathcal{B}_1(i)} = (p_{i_2})_i$
        end for
        $c_1 = c_1 + 1$
        $\Upsilon'_i(c_1) = \mathbf{v}$
      end if
    end for
  end for
  $\Upsilon = \Upsilon'_i$
  $L_1 = c_1$,
  $\mathcal{V} = \mathcal{V} \bigcup \mathcal{U}$
end for

Fonction : $\mathcal{U}$, $\mathcal{W}$, $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{B}_1$, $\mathcal{B}_2$ =UserPart $((\mathbf{r}_2)_{2k}, \mathcal{V})$
$\mathcal{C} = \emptyset$
$i = 0$
for $j = 1, \cdots, 2Jdv - 2K$ **do**
  if $(r_2)_{2k,j} = 1$ **then**
    $i = i + 1$
    $\mathcal{A}(i) = j$
    $\mathcal{B}(i) = \lceil \frac{\mathcal{A}(i)}{2d_v} \rceil$
    $\mathcal{C} = \mathcal{C} \cup \{\mathcal{B}(i)\}$
  end if
end for
$\mathcal{U} = \mathcal{C} \setminus \mathcal{V}$
$\mathcal{W} = \mathcal{C} \setminus \mathcal{U}$
$i_1 = i_2 = 0$
for $i = 1, 2, \cdots, 2(d_f - 1)$ **do**
  if $\mathcal{B}(i) \in \mathcal{U}$ **then**
    $i_1 = i_1 + 1$
    $\mathcal{B}_1(i_1) = \mathcal{B}(i)$
    $\mathcal{A}_1(i_1) = \mathcal{A}(i)$
  else
    $i_2 = i_2 + 1$
    $\mathcal{B}_2(i_2) = \mathcal{B}(i)$
    $\mathcal{A}_2(i_2) = \mathcal{A}(i)$
  end if
end for

Pour l'algorithme d'élagage proposé, à la $k^{\text{ème}}$ opération d'élagage, le taux d'erreur est calculé comme suit

$$p_k = \frac{2p}{M_1}\left(1 - \frac{2p}{M_1}\right)^{k-1} \tag{28}$$

où M_1 est le nombre moyen de projections distinctes dans une dimension réelle du dictionnaire de codes SCMA. Nous supposons que la valeur de M_1 est approximativement égale à $\sqrt{M}$. Comme il y a toujours K opérations d'élagage séquentielles, le taux d'erreur de l'algorithme PRUN1 est calculé comme suit

$$\begin{aligned} Pr\{\bar{\mathbf{x}}_{\mathrm{T}}^{(2)} \notin \mathcal{L}_2^*\}_1 &= \sum_{k=1}^{K} p_k \\ &= \sum_{k=1}^{K} \frac{2p}{M_1}\left(1 - \frac{2p}{M_1}\right)^{k-1} \\ &= 1 - \left(1 - \frac{2p}{M_1}\right)^{K} \end{aligned} \tag{29}$$

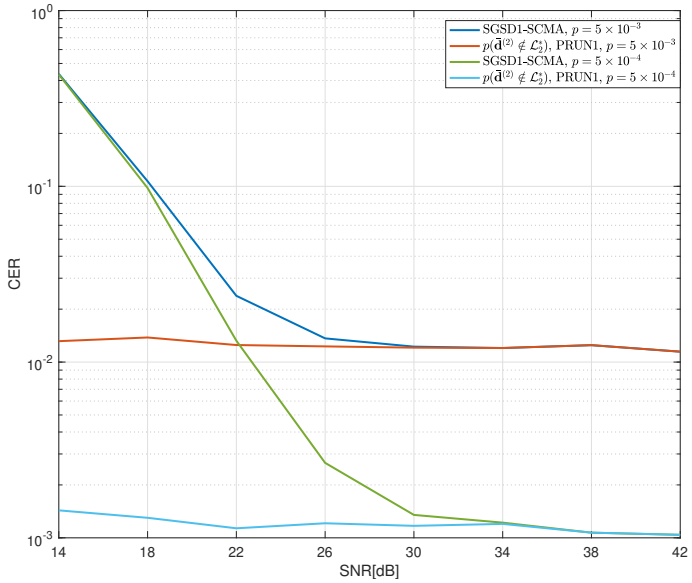


FIGURE 6 – Comparaison des performances CER de différents détecteurs.

Dans la simulation, nous considérons le dictionnaire avec $J = 6$, $K = 4$, $M = 16$ introduit dans [40]. Fig. 6 montre la performance simulée avec le dictionnaire de code étudié. Il est évident que le CER obtenu par simulation suit la limite démontrée dans l'équation (27). Table 2 compare les taux d'erreurs obtenus par simulation et théoriques de l'algorithme PRUN1 proposé. Elle montre que les résultats de la simulation correspondent à l'expression du taux d'erreur théorique Eq.(29). Fig. 7 compare les performances BER des détecteurs MPA et SGSD1-SCMA. Plus p est faible, meilleure est la performance BER du SGSD1-SCMA. Le détecteur SGSD1-SCMA avec $p = 5 \times 10^{-3}$ surpasse le détecteur MPA dans la gamme de 14dB $\leqslant$ SNR $\leqslant$ 22dB et le détecteur SGSD1-SCMA avec $p = 5{\times}10^{-4}$ est meilleur que le détecteur MPA dans la plage de 14dB $\leqslant$ SNR $\leqslant$ 24dB. Fig. 8 compare le nombre de FLOPs des deux étapes du détecteur SGSD1-SCMA avec différentes valeurs de p. Pour le MPA avec 5 itérations, le nombre de FLOPs est de 8.37×10^6. Quand le SNR est de 14dB, le nombre de FLOPs du détecteur SGSD1-SCMA avec $p = 5 \times 10^{-3}$ est d'environ $16,70\%$ de ceux du MPA avec 5 itérations. Dans un régime SNR modéré, le détecteur SGSD1-SCMA proposé surpasse le détecteur MPA avec 5 itérations en termes de performance BER et de complexité de calcul lors de la simulation avec le dictionnaire [40].

Chapitre 5 : Applications DL dans les systèmes NC MIMO

Dans ce chapitre, nous nous concentrons sur la transmission MIMO NC. Plus précisément, nous

TABLE 2 – Comparaison entre la simulation et le taux d'erreur théorique de l'algorithme PRUN1.

p	Simulation	Théorique
5×10^{-3}	1.1×10^{-2}	1.0×10^{-2}
5×10^{-4}	1.1×10^{-3}	1.0×10^{-3}

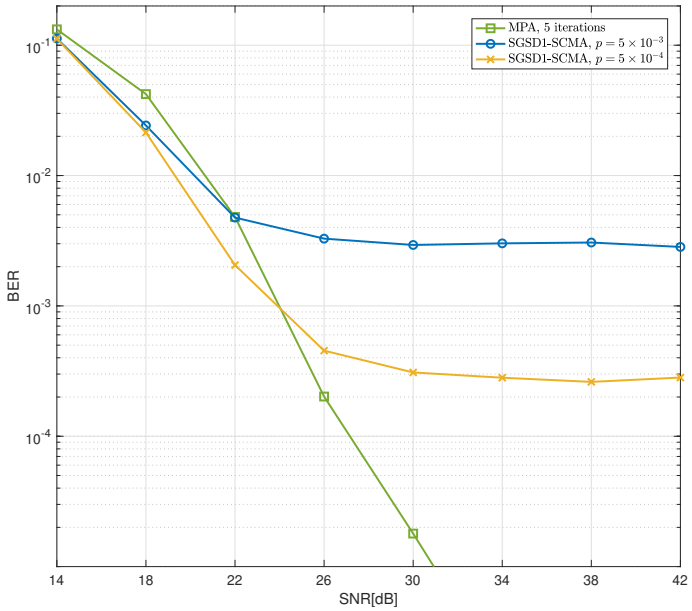


FIGURE 7 – Performance BER de différents détecteurs.

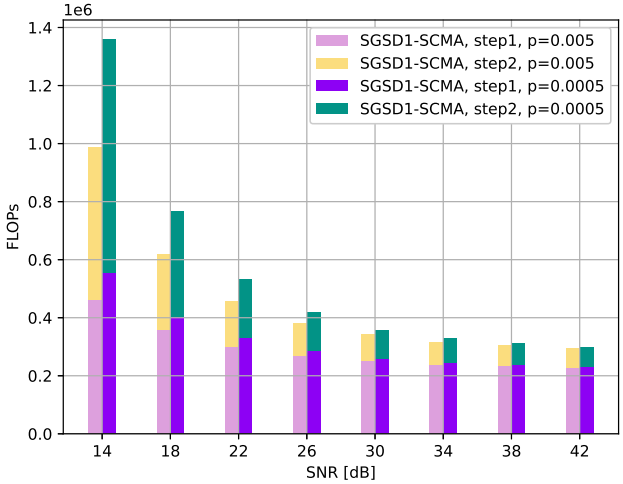


FIGURE 8 – Nombre de FLOPs des deux étapes du détecteur SGSD1-SCMA.

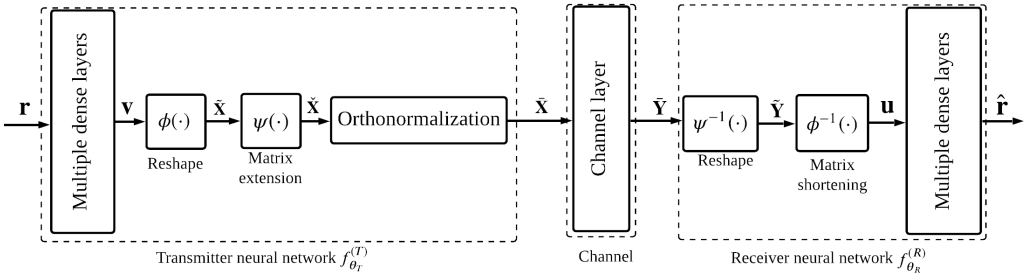


FIGURE 9 – Structure de l'AE-FC proposé.

tirons parti des techniques DL pour construire une constellation grassmannienne.

Pour faciliter l'application des réseaux neuronaux, l'équation de la transmission à valeurs complexes donnée dans l'équation (6) est réécrite en utilisant la notation équivalente à valeurs réelles :

$$\bar{\mathbf{Y}} = \bar{\mathbf{X}}\bar{\mathbf{H}} + \sqrt{\frac{N_t}{\rho T}}\bar{\mathbf{W}} \tag{30}$$

où $\bar{\mathbf{X}} = \begin{bmatrix} \Re\{\mathbf{X}\} & -\Im\{\mathbf{X}\} \\ \Im\{\mathbf{X}\} & \Re\{\mathbf{X}\} \end{bmatrix}$, $\bar{\mathbf{H}} = \begin{bmatrix} \Re\{\mathbf{H}\} & -\Im\{\mathbf{H}\} \\ \Im\{\mathbf{H}\} & \Re\{\mathbf{H}\} \end{bmatrix}$, $\bar{\mathbf{Y}} = \begin{bmatrix} \Re\{\mathbf{Y}\} & -\Im\{\mathbf{Y}\} \\ \Im\{\mathbf{Y}\} & \Re\{\mathbf{Y}\} \end{bmatrix}$ and $\bar{\mathbf{W}} = \begin{bmatrix} \Re\{\mathbf{W}\} & -\Im\{\mathbf{W}\} \\ \Im\{\mathbf{W}\} & \Re\{\mathbf{W}\} \end{bmatrix}$. $\mathbf{X}$, $\mathbf{H}$, $\mathbf{Y}$ et $\mathbf{W}$ sont les matrices à valeurs complexes utilisées dans Eq.(6). De même, la contrainte de constellation grassmannienne dans l'équation (9) peut également être écrite dans le domaine réel comme suit

$$\bar{\mathbf{X}}^{\dagger}\bar{\mathbf{X}} = \mathbf{I}_{2N_t} \tag{31}$$

Fig. 9 illustre la structure détaillée du premier AE proposé, appelé AE-fully connected (AE-FC), dans lequel des réseaux neuronaux profonds (DNN) FC sont utilisés dans le réseau neuronal émetteur et récepteur. L'entrée de l'AE-FC est un vecteur one hot $\mathbf{r}$ qui passe d'abord par les multiples couches denses, puis par des opérations matricielles et enfin par un processus d'orthonormalisation. Le vecteur $\mathbf{v} \in \mathbb{R}^{2TN_t \times 1}$ est la sortie des couches denses multiples au niveau de l'émetteur avec l'entrée $\mathbf{r}$. Comme le signal transmis est une matrice, il est nécessaire d'adapter le vecteur $\mathbf{v}$. Nous introduisons les fonctions $\boldsymbol{\phi}(\cdot)$ et $\boldsymbol{\psi}(\cdot)$ pour convertir le vecteur $\mathbf{v}$ en matrice $\check{\mathbf{X}}$ comme suit

$$\check{\mathbf{X}} = \boldsymbol{\psi}(\boldsymbol{\phi}(\mathbf{v})) = \boldsymbol{\psi}(\tilde{\mathbf{X}}) = \boldsymbol{\psi}\left(\begin{bmatrix} \tilde{\mathbf{X}}_1 \\ \tilde{\mathbf{X}}_2 \end{bmatrix} \right) = \begin{bmatrix} \tilde{\mathbf{X}}_1 & -\tilde{\mathbf{X}}_2 \\ \tilde{\mathbf{X}}_2 & \tilde{\mathbf{X}}_1 \end{bmatrix} \tag{32}$$

Pour satisfaire à l'exigence des constellations Grassmanniannes, l'orthonormalisation est adoptée pour garantir que les matrices de sortie du réseau neuronal émetteur sont unitaires, comme le montre l'équation (31). Dans notre implémentation, l'orthonormalisation est effectuée sur la base de la racine

carrée de $\check{\mathbf{X}}^{\intercal}\check{\mathbf{X}}$ qui s'écrit comme suit

$$\bar{\mathbf{X}} = \check{\mathbf{X}}(\check{\mathbf{X}}^{\intercal}\check{\mathbf{X}})^{-1/2} \tag{33}$$

Afin de bien entraîner l'AE-FC proposé en utilisant l'algorithme de descente de gradient stochastique, nous mettons en oeuvre la fonction de rétropropagation de l'opérateur de racine carrée inverse de la matrice dans l'équation ci-dessus.

Dans le réseau neuronal récepteur de AE-FC, $\boldsymbol{\phi}^{-1}(\cdot)$ et $\boldsymbol{\psi}^{-1}(\cdot)$ sont les fonctions inverses de $\boldsymbol{\phi}(\cdot)$ et $\boldsymbol{\psi}(\cdot)$ respectivement. Elles sont utilisées pour convertir la matrice du signal reçu à valeur réelle $\bar{\mathbf{Y}}$ en vecteur $\mathbf{u}$ afin que les multiples couches denses puissent traiter correctement le signal reçu. La sortie est le vecteur $\hat{\mathbf{r}} \in [0,1]^{M\times 1}$ suivant $|\hat{\mathbf{r}}||_1 = 1$, qui contient les probabilités de tous les messages correspondants. Par conséquent, la fonction d'activation $Softmax$ est appliquée à la dernière couche dense du réseau neuronal du récepteur. Dans l'AE-FC proposé, la fonction de perte est définie comme une entropie croisée catégorique pour pénaliser la différence entre $\mathbf{r}$ et $\hat{\mathbf{r}}$. Elle est définie comme suit

$$\mathcal{L} = -\sum_{i=1}^{S} \|diag(\mathbf{r}^{(i)}) \log(\hat{\mathbf{r}}^{(i)})\|_1 \tag{34}$$

où $\log(\cdot)$ est le logarithme par éléments des vecteurs et S désigne la taille du lot d'apprentissage.

Dans le domaine réel, le détecteur GLRT s'exprime comme suit

$$\hat{\bar{\mathbf{X}}} = \arg\max_{\bar{\mathbf{X}}\in\bar{\mathcal{C}}} \mathrm{Tr}\{\bar{\mathbf{Y}}^{\intercal}\bar{\mathbf{X}}\bar{\mathbf{X}}^{\intercal}\bar{\mathbf{Y}}\} \tag{35}$$

où $\bar{\mathcal{C}}$ est le dictionnaire de code correspondant à valeur réelle. En définissant $\mathbf{S} = \bar{\mathbf{X}}^{\intercal}\bar{\mathbf{Y}}$, le détecteur GLRT dans le domaine réel peut être reformulé comme suit

$$\hat{\bar{\mathbf{X}}} = \arg\max_{\bar{\mathbf{X}}\in\bar{\mathcal{C}}} \mathrm{Tr}\{\mathbf{S}^{\intercal}\mathbf{S}\} \tag{36}$$

Nous pouvons voir que le détecteur GLRT à valeur réelle dépend de la matrice $\mathbf{S}$. En exploitant la multiplication et la transposition de matrice, l'élément de la $i^{\text{ème}}$ $(1 \leqslant i \leqslant 2N_r)$ ligne et de la $j^{\text{ème}}$ $(1 \leqslant j \leqslant 2N_t)$ colonne de la matrice $\mathbf{S}$, $s_{i,j}$, est calculé comme suit

$$s_{i,j} = \sum_{n=1}^{2T} \bar{x}(n,i)\bar{y}(n,j) \tag{37}$$

où $\bar{x}(n,i)$ indique l'élément de la $n^{\text{ème}}$ ligne et de la $i^{\text{ème}}$ colonne de la matrice $\bar{\mathbf{X}}$ et $\bar{y}(n,j)$ indique l'élément de la n$i^{\text{ème}}$ ligne et de la $j^{\text{ème}}$ colonne de la matrice $\bar{\mathbf{Y}}$. Pour une couche convolutive bidimensionnelle (2D), si l'entrée a une taille de $(2T, 2N_r)$, le noyau a une taille de $(2T, 1)$ et le stride est

de $(2T, 1)$, la taille de la sortie est de $(1, 2N_r)$. Dans ce cas, la valeur du $q^{\text{ème}}$ élément $(1 \leqslant q \leqslant 2N_r)$ de la sortie de la couche convolutif est calculée comme suit

$$b_{1,q} = \sum_{k_1=1}^{2T} k_{k_1,1} a_{k_1,q} \tag{38}$$

Si le noyau de cette couche convolutive 2D est égal à la $i^{\text{ème}}$ $(1 \leqslant i \leqslant 2N_t)$ colonne de la matrice $\bar{\mathbf{X}}$ et que l'entrée est le signal reçu $\bar{\mathbf{Y}}$, la sortie de cette couche est équivalente au $i^{\text{ème}}$ vecteur de ligne de la matrice $\mathbf{S}$. Cela implique que le calcul de la matrice $\mathbf{S}$ peut être effectué par une couche convolutive 2D si nous pouvons obtenir les noyaux de manière appropriée à partir du dictionnaire de code donné. Par conséquent, nous proposons le détecteur CNN-GLRT qui implémente la détection GLRT à valeur réelle par un réseau de neurones convolutif (CNN) avec une couche convolutive 2D. Plus précisément, pour un mot de code donné $\bar{\mathbf{X}}$ $(\bar{\mathbf{X}} \in \bar{\mathcal{C}})$, il est divisé en $2N_t$ différents noyaux de taille $(2T, 1)$. En appliquant la couche convolutive 2D décrite ci-dessus, nous pouvons obtenir une valeur de la matrice $\mathbf{S}$ en empilant la sortie des $2N_t$ kernels. Après avoir obtenu toutes les valeurs de la matrice $\mathbf{S}$, le détecteur CNN-GLRT continue à reproduire l'opération indiquée dans l'équation (36). Il est intéressant de noter que, puisque tous les noyaux utilisés dans la couche convolutive sont obtenus à partir du dictionnaire de codes, il n'y a pas de paramètre à estimer et il n'est pas nécessaire d'entraîner le CNN-GLRT.

Nous proposons donc l'AE-GLRT pour concevoir des constellations Grassmanniennes qui tirent avantage du détecteur CNN-GLRT. Plus précisément, AE-GLRT consiste en un DNN entièrement connecté comme codeur et CNN-GLRT comme décodeur. Le décodeur de l'AE-GLRT devrait offrir des performances de détection quasi-GLRT. Fig. 10 décrit la structure détaillée de l'AE-GLRT proposé. Il est intéressant de noter que le décodeur de AE-GLRT peut traiter directement le signal reçu à valeur réelle $\bar{\mathbf{Y}}$. Comme pour la première proposition AE-FC, les entrées de AE-GLRT sont des vecteurs à un coup et les sorties sont des vecteurs contenant les probabilités de tous les mots de code correspondants obtenus en appliquant la fonction d'activation $Softmax$ à la sortie du réseau neuronal du décodeur. La fonction de perte de AE-GLRT est identique à l'équation suivante : Eq.(34).

Fig. 11 compare la performance du taux d'erreur symbole (SER) du détecteur CNN-GLRT à celui du détecteur GLRT conventionnel. Nous montrons que la performance de la constellation $G_{4,2}$ [30] ainsi que les constellations conçues par l'algorithme glouton [26] et l'AE-FC proposé. Comme le montre Fig. 11, CNN-GLRT a la même efficacité de détection que le GLRT conventionnel. Fig. 12 compare la performance SER de différentes constellations pour un canal de Rayleigh i.i.d.. Lorsque le détecteur

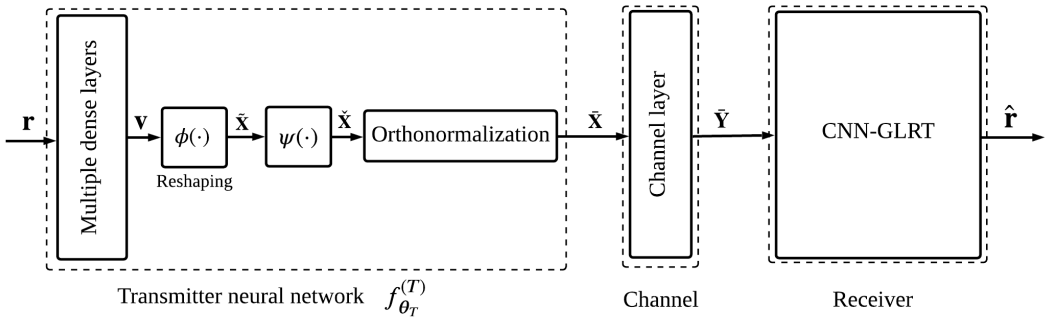


FIGURE 10 – Structure de l'AE-GLRT proposé.

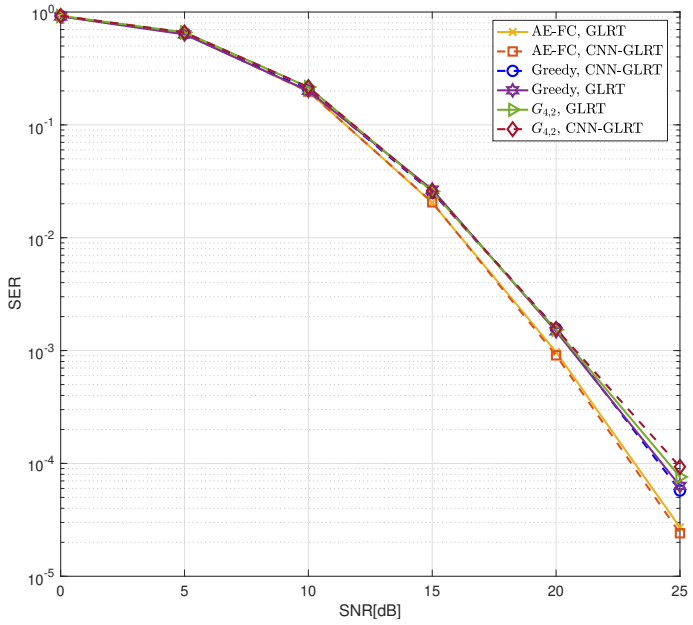


FIGURE 11 – Performance SER du détecteur CNN-GLRT.

CNN-GLRT est appliqué au récepteur, les constellations apprises par AE-FC et AE-GLRT surpassent les constellations générées par les approches de la littérature [30, 26]. De plus, le récepteur entièrement connecté d'AE-FC donne une mauvaise performance SER car il ne peut pas être compétitif avec le détecteur GLRT optimal.

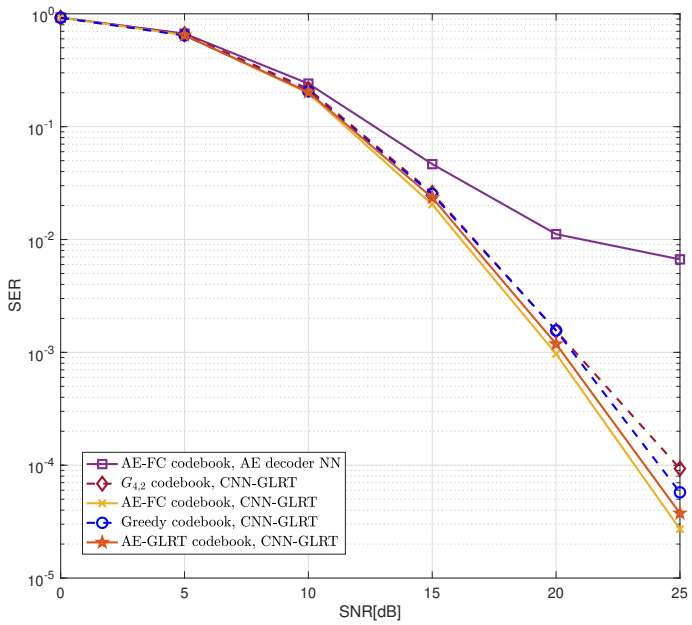


FIGURE 12 – Performance SER de différentes constellations dans un canal de Rayleigh i.i.d..

Chapitre 6 : Approche basée sur DL pour la conception SU-MDC

Dans ce chapitre, nous nous concentrons sur la conception de MDC pour la transmission multidimensionnelle SU qui est motivée par les schémas pratiques basés sur l'OFDM. Le modèle de système considéré est donné dans l'équation (10). Nous proposons une approche composée d'un DNN fully-connected avec plusieurs couches, illustrée dans Fig. 13, où la fonction $f(\cdot)$ représente la structure DNN proposée. Les entrées de $f(\cdot)$ sont les vecteurs one-hot $\mathbf{r}_m \in \mathbb{R}^{M\times 1}$ ($m \in \{1, \cdots, M\}$), dans lesquels l'indice m indique l'index du message transmis. La sortie correspondante du DNN $f(\cdot)$ avec l'entrée $\mathbf{r}_m$ est le mot code $\mathbf{x}_m \in \mathbb{R}^{2N\times 1}$, formulé comme suit

$$\mathbf{x}_m = f(\mathbf{r}_m) \tag{39}$$

Le dictionnaire de codes appris par l'approche proposée comprend tous les mots de codes $\mathbf{x}_m$ ($m \in \{1, \cdots, M\}$), formulés comme suit

$$\mathcal{C} = f(\mathbf{R}) \tag{40}$$

où la matrice $\mathbf{R} = [\mathbf{r}_1, \cdots, \mathbf{r}_M]$ concatène tous les vecteurs one-hot. À la différence des AEs, l'approche proposée s'appuie de manière significative sur sa fonction de perte pour apprendre des dictionnaires de codes avec de bonnes performances, comme l'illustre la figure Fig. 13.

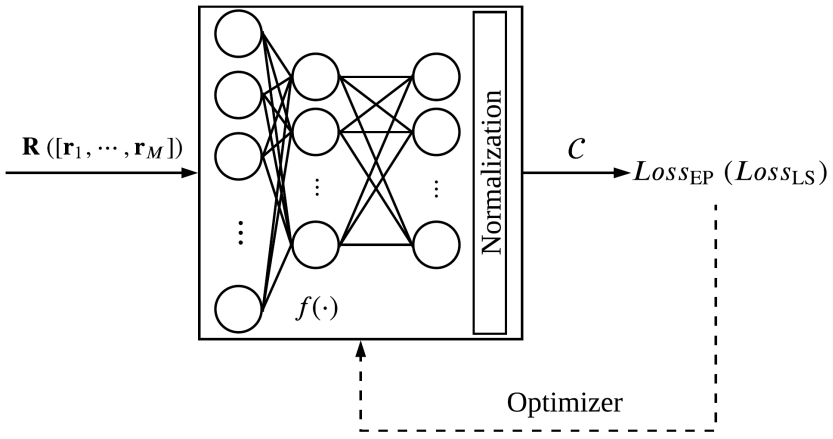


FIGURE 13 – Schéma d'apprentissage du DNN proposé.

La première fonction de perte de l'approche proposée a pour objet de maximiser la distance euclidienne minimale (MED) et la distance produit minimale (MPD) conjointement afin de résoudre le problème d'optimisation MDC. Elle est dénotée $Loss_{\mathrm{EP}}$. Par conséquent, $f_{\mathrm{EP}}(\cdot)$ fait référence au DNN proposé utilisant la fonction de perte $Loss_{\mathrm{EP}}$ pour la formation et son dictionnaire de codes résultant est désigné par $\mathcal{C}_{\mathrm{EP}}$. Compte tenu des caractéristiques de la descente de gradient, la fonction de perte est définie comme l'opposé de la somme de MED et MPD, qui est formulée comme suit

$$Loss_{\mathrm{EP}} = -\Big(d_E(\mathcal{C}_{\mathrm{EP}}) + \alpha d_P(\mathcal{C}_{\mathrm{EP}})\Big) \tag{41}$$

où la fonction $d_E(\cdot)$ calcule le dictionnaire de codes MED, la fonction $d_P(\cdot)$ calcule le dictionnaire de codes MPD, α est un coefficient choisi pour coordonner l'échelle de valeurs de MED et MPD afin de faciliter l'apprentissage. α est un coefficient positif et peut être optimisé séparément comme un hyper-paramètre.

Parce que l'objectif de l'approche proposée est de concevoir des MDCs avec de bonnes performances, par exemple de faibles probabilités d'erreur, il est intuitif d'exploiter la probabilité d'erreur de symbole (SEP) théorique dans la fonction de perte. Comme la valeur de la SEP est bien inférieure à 1, elle pourrait probablement devenir très faible pendant la formation, ce qui pourrait faire converger le DNN vers un optimum local. Comme la fonction logarithme est monotone et peut transformer le SEP en un nombre réel négatif, elle est appliquée dans la fonction de perte afin de permettre un apprentissage plus efficace de l'approche proposée. Par conséquent, le deuxième type de fonction de perte est désigné par $Loss_{\mathrm{LS}}$ et son dictionnaire de codes résultant est désigné par $\mathcal{C}_{\mathrm{LS}}$. La fonction de perte $Loss_{\mathrm{LS}}$

s'écrit comme suit

$$Loss_{\mathrm{LS}} = \log\left(\frac{1}{M}\sum_{m=1}^{M}\sum_{\substack{m_1=1\\ m_1\neq m}}^{M} P(f_{\mathrm{LS}}(\mathbf{r}_m) \to f_{\mathrm{LS}}(\mathbf{r}_{m_1}))\right) \tag{42}$$

où $f_{\mathrm{LS}}(\mathbf{r}_m)$ et $f_{\mathrm{LS}}(\mathbf{r}_{m_1})$ représentent respectivement le $m^{\text{ème}}$ et le $m_1^{\text{ème}}$ mot de code appris par le DNN proposé avec la fonction de perte $Loss_{\mathrm{LS}}$. $P(f_{\mathrm{LS}}(\mathbf{r}_m) \to f_{\mathrm{LS}}(\mathbf{r}_{m_1}))$ est la probabilité d'erreur par paire (PEP) entre les deux mots de code différents. En supposant que la distance dimensionnelle réelle entre deux mots de code est unique, la PEP entre le mot de code $\mathbf{a}$ et le mot de code $\mathbf{b}$ peut être approximée comme suit [41] :

$$P(\mathbf{a} \to \mathbf{b}) = \sum_{n=1}^{2N}\left(\frac{1-\mu_n}{2}\right)\prod_{\substack{n'=1\\ n'\neq n}}^{2N}\frac{\delta_n^2}{\delta_n^2 - \delta_{n'}^2} \tag{43}$$

où δ_n est le $n^{\text{ème}}$ élément du vecteur $\boldsymbol{\delta} = |\mathbf{a} - \mathbf{b}|$ défini comme vecteur de distance absolue entre $\mathbf{a}$ et $\mathbf{b}$, et $\mu_n = \sqrt{\frac{\delta_n^2}{4N_0+\delta_n^2}}$.

Fig. 14 illustre les performances SER de différents dictionnaires dans le système avec $M = 256$ et $N = 4$. La performance SER du dictionnaire $\mathcal{C}_{\mathrm{EP}}$ et $\mathcal{C}_{\mathrm{LS}}$ est presque la même. Ils surpassent le dictionnaire de codes résultant de l'AE model-free [42] 1.6dB et celui de l'AE model-aware [43] 2dB, lorsque le SNR est égal à 20dB. Pour les valeurs élevées de SNR (20dB $\leqslant$ SNR $\leqslant$ 25dB), la performance des dictionnaires appris par la nouvelle approche est très proche de celle du dictionnaire rot. GLCP $\mathcal{C}_{E8}$ [35]. En considérant que le codebook rot.GLCP $\mathcal{C}_{E8}$ montre la meilleure performance SER dans l'évaluation de performance du papier [35], le schéma proposé est une approche efficace.

Table 3 liste les figures de mérite (FoMs) de différents dictinnaires dans le système avec $M = 256$ et $N = 4$. Les MPDs des dictinnaires $\mathcal{C}_{\mathrm{EP}}$ et $\mathcal{C}_{\mathrm{LS}}$ sont considérablement plus élevés que ceux des dictionnaires appris par les AEs. La raison pourrait être que l'apprentissage des AE ne tient pas compte des exigences spécifiques de la conception des MDCs sur les canaux à évanouissement de Rayleigh. Comme le dictionnaire $\mathcal{C}_{\mathrm{LS}}$ a une SSD plus élevée que le dictionnaire $\mathcal{C}_{\mathrm{EP}}$, le premier a un SER plus faible, en particulier pour des valeurs SNR élevées.

Chapitre 7 : Conclusions

Tout d'abord, nous nous sommes concentrés sur les algorithmes MUD itératifs pour les systèmes SCMA, où deux scénarios de transmission sont considérés. Dans le premier scénario, nous avons proposé un algorithme itératif de faible complexe appelé FO-GAA qui a une bonne vitesse de convergence et

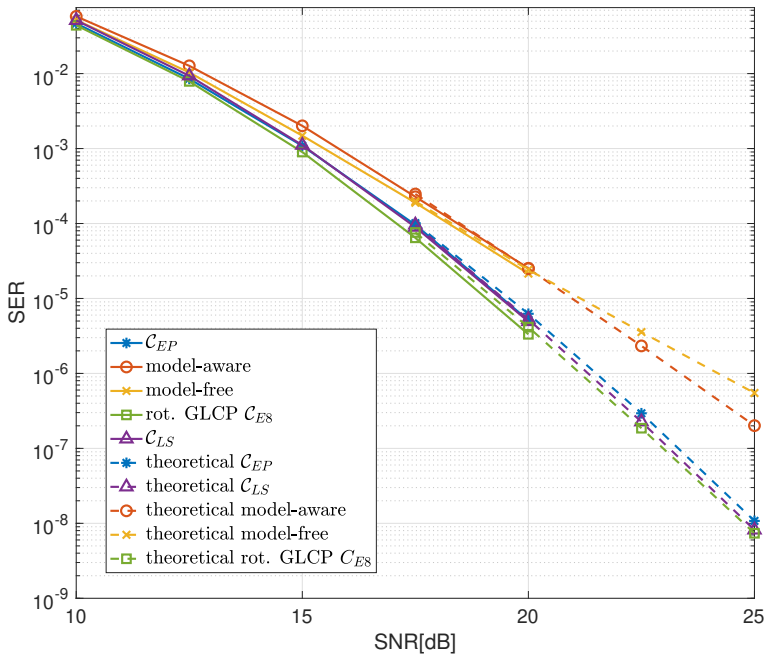


FIGURE 14 – SER des différents dictionnaires dans le système avec $M = 256$ et $N = 4$.

TABLE 3 – FoM des dictionnaires dans le système avec $M = 256$ et $N = 4$.

MDC	MED	MPD	SSD
$f_{\mathrm{EP}}^{(T)}$	1.660	5.386×10^{-8}	6
$f_{\mathrm{LS}}^{(T)}$	1.025	1.442×10^{-8}	7
Model-aware	1.055	1.339×10^{-9}	6
Model-free	1.359	5.873×10^{-9}	7
rot. GLCP $\mathcal{C}_{E8}$	1.952	3.308×10^{-5}	7

réduit considérablement la complexité tout en montrant une dégradation de performance BER limitée par rapport au décodage MPA dans un système SCMA codé. Dans le second scénario, on introduit une transmission SCMA fortement chargée où deux groupes de SCMA partagent les mêmes REs. Au niveau du récepteur, nous avons employé simultanément le MPA et le SIC. Compte tenu du compromis entre les performances MUD et l'efficacité spectrale, nous avons proposé une méthode d'optimisation pour sélectionner les taux de code des canaux en analysant le graphique EXIT de la transmission. Le schéma à forte charge montre une amélioration substantielle de l'efficacité spectrale et une bonne performance BER par rapport au schéma SCMA conventionnel.

Dans un deuxième temps, nous étudions les détecteurs basés sur le SD pour les systèmes SCMA. Pour résoudre le problème de complexité du détecteur GSD-SCMA, nous avons proposé le détecteur

SGSD-SCMA qui utilise la méthodologie d'élagage. En outre, nous avons calculé les taux d'erreur théoriques de l'algorithme d'élagage proposé en supposant le nombre moyen de projections du dictionnaire de codes dans chaque dimension réelle. Nous avons montré que le détecteur SGSD-SCMA proposé est avantageux pour les dictionnaires de codes avec un facteur de charge modéré et une grande taille.

Ensuite, nous nous concentrons sur la construction de constellations grassmanniennes pour les systèmes MIMO NC. Nous avons proposé deux AE, à savoir AE-FC et AE-GLRT. L'AE-FC se compose de réseaux neuronaux DNN entièrement connectés pour le codage et le décodage. Pour garantir la génération de constellations grassmanniennes, nous avons introduit un bloc d'orthonormalisation dans le réseau neuronal du codeur afin de réaliser avec succès la rétropropagation pour la formation. D'autre part, la deuxième proposition AE-GLRT profite du récepteur CNN-GLRT qui met en œuvre la détection GLRT à valeur réelle sur un CNN. Les approches proposées basées sur l'AE surpassent les approches conventionnelles en termes de performance SER des constellations Grassmanniennes construites.

Enfin, nous travaillons sur la conception de MDC dans la transmission multidimensionnelle SU. Nous avons proposé une nouvelle approche basée sur le DNN pour concevoir les MDC et nous avons conçu deux fonctions de perte dédiées à cette nouvelle approche. L'approche proposée a une structure beaucoup plus simple que les approches basées sur l'AE et elle nécessite peu de données. Les MDCs obtenus par l'approche proposée sont beaucoup plus performants que les MDCs conçus par les AEs et ils s'approchent des MDCs optimaux de la littérature.

Table des matières

Liste des tableaux

Table des figures

Nomenclature

Abbreviations and Acronyms

2D	2-dimensional
4G	Fourth Generation
5G	Fifth Generation
6G	Sixth Generation
AE	Autoencoder
AR	Augment Reality
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BG	Base Graph
BICM	Bit-Interleaved Coded Modulation
BP	Belief Propagation
BS	Base Station
CD-NOMA	Code-Domain Non-Orthogonal Multiple Access
CER	Codeowrd Error Rate
CNN	Convolutional Neural Network
CP	Cyclic Prefix
CSI	Channel State Information
DL	Deep learning
DLT	Distributed Ledger Technology
DNN	Deep Neural Network
eMBB	Enhanced Mobile Broadband
EPA	Expectation Propagation Algorithm
EXIT	Extrinsic Information Transfer
FFT	Fast Fourier Transform
FLOP	Floating-Point Operation
FN	Function Node
FoM	Figure of Merit
GAA	Gaussian Approximation Algorithm
GLCP	Grouped Linear Constellation-Precoding
GLRT	Generalized Likelihood Ratio Test
IDD	Iterative Demapping and Decoding
IFFT	Inverse Fast Fourier Transform

IM	Index Modulation
IoE	Internet of Everything
IoT	Internet of Things
ITU	International Telecommunication Union
LDPC	Low-Density Parity-Check
LDS	Low-Density Signature
LLR	Log-Likelihood Ratio
LTE	Long-Term Evolution
MAP	Maximum a Posteriori
MDC	Multidimensional Constellation (Codebook)
MED	Minimum Euclidean Distance
MI	Mutual information
MIMO	Multiple-Input Multiple-Output
ML	Maximum-Likelihood
mMTC	Massive Machine Type Communications
mmWave	Millimeter Wave
MPA	Message Passing Algorithm
MPD	Minimum Product Distance
MR	Mixed Reality
MTC	Machine Type Communication
MUD	Multiuser Detection
NC MIMO	Non-Coherent Multiple-Input Multiple-Output
NN	Neural Network
NOMA	Non-Orthogonal Multiple Access
OFDM	Orthogonal Frequency Division multiplexing
OMA	Orthogonal Multiple Access
PD-NOMA	Power-Domain Non-Orthogonal Multiple Access
PDF	Probability Density Function
PEP	Pairwise Error Probability
PSMA	Power-Domain Sparse Code Multiple Access
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
QPSK	Quadrature Phase Shift Keying
RE	Resource Element
ReLU	Rectified Linear Unit
RNN	Recurrent Neural Network
SCMA	Sparse Code Multiple Access
SD	Sphere Decoding
SEP	Symbol Error Probability
SER	Symbol Error Rate
SGD	Stochastic Gradient Descent
SIC	Successive Interference Cancellation
SNR	Signal-to-Noise Ratio
SQRD	Sorted QR Decomposition
SSD	Signal Space Diversity

SU	Single-User
SVD	Singular Value Decomposition
UAV	Unmanned Aerial Vehicle
UL	Uplink
URLLC	Ultra-Reliable and Low-Latency Communication
USTM	Unitary Space-Time Modulation
VN	Variable Node
VR	Virtual Reality
XR	Extended Reality

Notations

x	Scalar.
$\mathbf{x}$	Column vector.
$\mathbf{X}$	Matrix.
$\mathcal{X}$	Set (sequence).
$\mathbf{I}_q$	The $q \times q$ identity matrix.
$\mathbf{0}_{m \times n}$	The $m \times n$ zero matrix.
$\mathbb{V}_{p,q}$	The space of $p \times q$ $(p \geqslant q)$ unitary matrices.
$(\cdot)^\intercal$	The transpose operator of matrices.
$(\cdot)^\dagger$	The Hermitian transpose operator of matrices.
$\odot$	The Hadamard product operator of vectors.
$diag(\mathbf{x})$	The square diagonal matrix with the elements of vector $\mathbf{x}$ on the main diagonal.
$Tr(\cdot)$	The trace operator of matrices.
$\|\cdot\|_2$	l_2-norm of vector.
$\lceil x \rceil$	The ceiling operator.
$\mathcal{N}_c(\tau, \sigma^2)$	Complex Gaussian distribution function with mean τ and variance σ^2.
$\mathcal{N}(\tau, \sigma^2)$	Real-valued Gaussian distribution function with mean τ and variance σ^2.
$\mathbf{x} \backslash x_i$	The symbols in $\mathbf{x}$ excluding x_i.
$\mathbb{E}_\rho[x]$	The expectation of variable x following the distribution of ρ.

Chapitre 1

Introduction

Contenu

1.1 Motivation

Over the last decade, the fifth generation (5G) of wireless communication systems are becoming a commercial reality. In 5G systems, enhanced mobile broadband (eMBB), ultra-reliable and low-latency communications (URLLC) and massive machine type communications (mMTC) are three usage scenarios for the applications, pinpointed by the International Telecommunication Union (ITU) [44]. With the rapid development of multimedia, applications in the form of high-fidelity holograms, extended reality (XR) services (including augment, mixed and virtual reality (AR/MR/VR)) and haptic/tactile-based communications requires extremely high throughput in 2030 and hereafter [45]. Besides, the unmanned mobility, such as unmanned aerial vehicles (UAVs) and autonomous driving, and the deployment of smart cities and homes will introduce millions of sensors in the wireless communication systems [46]. Therefore, researchers and engineers have been turning their attention to a new generation of wireless networks, a systems beyond 5G which is referred to as the sixth generation (6G) of wireless systems in this thesis. According to the white paper [47], three most important key drivers of 6G systems are : high fidelity holographic society, connectivity for all things and time sensitive and time engineering applications. With the shift of driving scenarios and applications as well as the flexibility requirement, new enabling technologies at the physical layer is essential for the 6G wireless networks [1, 2]. The enablers of 6G can either enhance or adapt the 5G design or innovate new communication paradigms.

In the future 6G systems, machine-to-machine and machine-to-human communications are anticipated to take place alongside the people-to-people communications. Wireless connectivity will become an ubiquitous utility, like water and electricity. The devices, ranging from sensors in the industry, distributed ledger technology (DLT) devices to wearable equipment and smart implants, need wireless interconnection without human intervention [1]. Concerning this exponentially increase of machine type communication (MTC), massive connectivity and high spectral efficiency are two major demands. Thus, non-orthogonal multiple access (NOMA) techniques have been proposed as potential enablers to meet these demands [3]. Compared with conventional orthogonal multiple access (OMA), NOMA enables more users than radio (time/frequency) resources to transmit their signals by leveraging degrees of freedom in the power or code domain. Thus, NOMA allows reliable overload transmission which increase the capacity of multiple access schemes.

NOMA techniques are categorized into power-domain NOMA (PD-NOMA) and code-domain NOMA (CD-NOMA). CD-NOMA is motivated by the idea of non-orthogonal spreading sequences. Among all the CD-NOMA techniques, sparse code multiple access (SCMA) has figured prominently. In SCMA, users transmit signals over orthogonal resource elements (REs) and the sparse spreading reduces the number of interfering users at the same RE. SCMA can be regarded as generalized low-density signature (LDS) [48] which combines binary mapping and constellation spreading jointly at the encoder. In other words, SCMA maps users' incoming bits directly to multidimensional codewords according to a pre-designed codebook set [40]. At the receiver, efficient algorithms take charge of multiuser detection (MUD), generally by exploiting the intrinsic sparsity of the signature design.

As one of the defining features of 5G systems, the deployment of multiple antenna techniques has achieved great success. Transmission relying on multiple antennas at both transmitter and receiver, known as multiple-input multiple-output (MIMO) technique, achieves substantially large spectral efficiency without requiring additional transmission power [4]. As illustrated in [49], the capacity of a point-to-point MIMO system without any channel correlation increases proportionally with the minimum of the transmit and receive antennas. However, this result is based on the perfect knowledge of channel state information (CSI) at the receiver and/or the transmitter. This kind of MIMO systems are called coherent MIMO systems. In practice, estimation of CSI is carried out by sending reference symbols, so-called pilots, periodically from the transmitter. Since the pilot symbols are known to the receiver, they do not carry any information and occupy transmission resource. Due to the pilot overhead and CSI estimation error, the coherent scheme always suffer from capacity loss [50]. Therefore, to circumvent the CSI estimation problem and save valuable communication resources, the non-coherent (NC) MIMO transmission scheme, where neither the transmitter nor the receiver needs CSI knowledge, has been proposed [24, 51]. Paper [24] reveals that when the coherence interval grows to infinity, the capacity of NC MIMO approaches to the Shannon capacity with perfect CSI. As the unprecedented Internet-of-everything (IoE) connectivity demands in 6G makes pilot-based CSI estimation infeasible, NC MIMO stands as a promising technique [5]. In addition, the non-coherent schemes can achieve high capacity in the high-mobility scenarios and requires less bandwidth than coherent schemes in future wireless applications, for example UAVs [50].

In NC MIMO systems, none of the channel coefficient is known at either the transmitter or receiver and the channel coefficients remain constant for a coherence interval. Based on the capacity analysis,

Hochwald and Marzetta have proposed the use of unitary space–time modulation (USTM), in which the signals of transmit antennas are mutually orthogonal, for Rayleigh flat-fading channels [24]. On the other hand, for any consistent or slowly varying channel, the authors of [52] propose differential unitary space-time modulation for NC MIMO systems inspired by the standard single-antenna differential phase-shift keying. In this thesis, we focus on the NC MIMO systems taking into account the block-fading channels which is referred to as NC MIMO for simplicity. Following the principles of (compact) Grassman manifold, the constellations of NC MIMO systems are also called Grassmannian constellations [6]. At the receiver, the optimal detector for NC MIMO is called generalized likelihood ratio test (GLRT) which is based on the conditional probability density of the received signals.

Over the last decade, orthogonal frequency-division multiplexing (OFDM), as a multi-carrier modulation technique, has been adopted in many wireless communications standards, such as the fourth generation (4G) and 5G, due to its robustness to the frequency selectivity and inter-symbol interference. However, the performance of OFDM suffers since it does not exploit the channel diversity. To recover the diversity, bit-interleaved coded modulation OFDM (BICM-OFDM) has been proposed as a promising approach which is inspired by the BICM scheme [53] and driven by the development of channel codes [54, 55]. Still, with the increasing multiplicity of future wireless devices, coded conventional OFDM will not be able to meet all the requirements simultaneously. For instance, the coded OFDM schemes, such as BICM-OFDM, may be impossible to provides URLLC services because of the high complexity at the receiver. Therefore, it is expected that different techniques can be used and integrated with OFDM for the increased heterogeneity of applications and flexibility of wireless networks [5, 56, 57].

In [57], the authors evaluate and compare various modulation options for OFDM that can improve spectral efficiency and reliability for future wireless communication challenges. Among them, OFDM-index modulation (OFDM-IM) is one of the promising candidates in terms of spectral efficiency and decoding performance. In OFDM-IM, the subcarriers are divided into groups in each of which an information-carrying index is associated with the active subcarrier subset [58, 59]. In this scheme, the information is conveyed not only by the signal constellations but also by the indices of the subcarriers. On the other hand, the authors of paper [7] propose grouped linear constellation-precoding (GLCP) for OFDM in order to recover the multipath loss diversity and reduce decoding complexity. In both OFDM-IM and GLCP, OFDM subcarriers are split into groups and the signals of the subcarriers in a

same group is modulated by multidimensional constellations (MDCs). Moreover, without adopting any channel code, these techniques exploit signal space diversity (SSD) which is deemed to provided low-latency service and high reliability. The scheme design for either OFDM-IM or GLCP can be simplified to the MDC design problem in a generalized single-user (SU) multidimensional transmission system.

As we are moving toward 6G, explosively increasing devices, improved service coverage, better cost efficiency and network adaptability are challenging the conventional way of wireless network design [8]. On the other hand, deep learning (DL) has achieved tremendous success in computer vision, natural language processing and autonomous driving thanks to the dedicated algorithmic advances and rapid hardware computing power growth. As a class of machine learning algorithms, DL consists in employing multiple layers in an artificial neural network. Contrary to conventional engineering approaches, DL enables computers to learn and perform certain tasks that are not explicitly programmed. DL-based approaches are referred to as data-driven approaches because they are usually trained with sufficiently large number of examples with desired behaviors. Potential benefits of applying DL into 6G are as follows. First, since large amount of data are available, DL techniques can extract the unknown channel information or latent model patterns by leaning from the training [60, 61]. Secondly, sophisticated DL techniques may be able to solve the difficult tasks such as global optimization and fine tuning of system settings for the wireless communications [60]. Last, DL-based approaches are applicable to model and/or algorithm deficit scenarios where no physical-based mathematical models exist and/or model-based algorithms are rather complicated [8, 61]. The significant and positive impact of DL approaches on future wireless communications has been illustrated [9]. Specifically, at the physical layer, DL could play a major role in channel decoding [10, 11], symbol detection [12, 13], channel estimation [14], and etc. Furthermore, there is a growing trend of integrating domain knowledge of wireless communications in DL [15, 16] for the future 6G networks.

In this thesis, we study three promising enabling techniques introduced above, namely SCMA, NC MIMO and SU multidimensional transmission, for the future 6G wireless systems. To be more specific, we focus on MUD in SCMA systems, Grassmannian constellation construction in NC MIMO and MDC design in SU multidimensional transmission. Besides, this thesis considers some effective DL applications in the wireless networks.

1.2 Contributions

The major contributions of this thesis are as follows :

(i) We study the iterative MUD algorithms for both simple SCMA transmission and highly overload two-groups SCMA transmission. In the simple scenario, we propose a low-complexity iterative MUD receiver which achieves good performance in the coded SCMA transmission. In the second scenario, we introduce an uplink transmission scheme where PD-NOMA and SCMA techniques are adopted jointly in order to enable massive user connectivity. Considering the trade-off between MUD performance and spectral efficiency, an optimization method by exploiting extrinsic information transfer (EXIT) charts to choose the code rates for the error-correcting channel codes is proposed. An example of analyzing EXIT chart for the proposed optimization method is provided.

(ii) We survey the sphere decoding (SD)-based detectors for SCMA. Without loss of generality, we thoroughly review two state-of-the-art detectors which can be applied to the generalized SCMA transmission model. As a novel contribution, we leverage the sorted QR decomposition (SQRD) and Schnorr-Euchner (SE) enumeration in the considered state-of-the-art detectors to futher accelerate the tree search speed. We propose two pruning algorithms and introduce a low-complexity SD-based detector. Additionally, theoretical error rates of the proposed algorithms are derived. The impact of system parameters on the MUD performance and computational complexity of different SD-based detectors is discussed.

(iii) We propose two autoencoders (AEs) to construct Grassmannian constellations for NC MIMO systems. The first proposed scheme comprises fully-connected deep neural network (DNN) as both encoder and decoder. In addition, we introduce an orthonormalization process to enable the back propagation of the training. We propose the so-called convolutional neural network (CNN)-GLRT which implements the GLRT detection on a CNN. By applying CNN-GLRT as the decoder, we propose AE-GLRT which only has trainable parameters at the decoder. The resulting constellations of the proposed AE schemes are compared with other constellations constructed by conventional approaches.

(iiii) A novel DL-based scheme to design MDCs for multidimentional transmission is proposed and two classes of loss functions are devised for the proposed scheme. Unlike AE-based MDC design approaches, the proposed DL-based scheme has significantly simpler structure and is

not training data-hungry. Besides, it can be regarded as a standalone optimization method. A comparative study of the performance and figures of merit (FoM) of the novel scheme and sate-of-the-art approaches is presented.

1.3 Thesis structure

The outline of the thesis is given as follows :

Chapter 2 is an introduction of the technical backgrounds and the state-of-the-art of the three promising enabling techniques covered in this thesis, which are SCMA, NC MIMO and SU multidimensional transmission. Please note that basic system models of the three techniques are provided in this chapter. Furthermore, we present the DL techniques that are used in the thesis.

In chapter 3, we focus on iterative detectors for SCMA systems. Two different transmission scenarios, namely the simple and two-groups scenarios, are considered. In the first scenario where a typical coded SCMA system is considered, we propose a low-complexity iterative detector. In the second scenario, there are two groups of SCMA users sharing the same REs and we optimize the channel codes in order to enhance the spectral efficiency while guaranteeing the good system performance. The two groups transmission system with the proposed optimization method is evaluated by comparing it with a conventional system with two independent SCMA transmission.

In chapter 4, we first review two promising sphere decoding (SD)-based SCMA detector proposed in the literature. We not only provide flowcharts and pseudo-codes of the state-of-the-art, but also introduce SQRD and Schnorr-Euchner enumeration to further increase the decoding efficiency and tree search speed. Furthermore, we propose two pruning algorithms, PRUN1 and PRUN2, and introduce the simplified generalized SD for SCMA (SGSD-SCMA) detector. Also, we derive theoretical error rates of the proposed pruning algorithms. We conduct a performance investigation of the SD-based detectors regarding various SCMA parameter settings.

Chapter 5 sheds light on DL techniques for NC MIMO systems. Two AEs with distinct structures are proposed for Grassmannian constellation design. Besides, we introduce an orthonormalization block which is compatible with the training of the proposed AEs and satisfies the restriction of Grassmannian constellations at the same time. The details of the proposed AE schemes and necessary operations are well elaborated. In the performance evaluation, both the resulting codebooks of the proposed AE

approaches and the conventional approaches are investigated.

In chapter 6, we concentrate on MDC design for the SU multidimensional transmission. We investigate the essential FoMs of designing MDCs. A novel and much simpler DNN scheme is proposed for MDC design, which can be regarded as a generic approach. Besides, we introduce two effective loss functions devised for the training of the proposed scheme. The proposed DNN-based approach is compared with some state-of-the-art approaches in the performance evaluation.

1.4 Related publications

The content of this dissertation is mainly based on the following;

1 X.Fu, M.Pischella and D.Le Ruyet, "On Gaussian Approximation Algorithms for SCMA", 2019 16th International Symposium on Wireless Communication Systems (ISWCS), pp.155-160, 2019.

2 X.Fu, M.Pischella and D.Le Ruyet, "Joint uplink PD-NOMA and SCMA for future multiple access systems", 2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring), 2021.

3 X.Fu, B.F.da Silva and D.Le Ruyet, "Multidimensional Codebook Design Using Deep Learning Techniques for Rayleigh Fading Channels", IEEE Wireless Communications Letters, 2021.

4 X.Fu, and D.Le Ruyet, "Grassmannian constellation design for noncoherent mimo systems using autoencoders", submitted for publication.

5 X.Fu, and D.Le Ruyet, "On deep learning techniques for Noncoherent MIMO systems", submitted for publication.

6 X.Fu, D. Le Ruyet, B.F.da Silva and B.F.Uchôa-Filho, "A Simplified Sphere Decoding-Based Detector for Generalized SCMA Codebooks", in preparation.

Chapitre 2

State-of-the-art

Contenu

In this chapter, we give the backgrounds and major state-of-the-art that are relevant to our research studies in this thesis. The first three sections, section 2.1 2.2, 2.3, introduce the technical basics and system model of SCMA, NC MIMO and SU multidimensional transmission, respectively. Section 2.4 sheds light on the DL techniques used in this thesis.

2.1 SCMA systems

In this section, we will introduce SCMA regarding system model, multidimensional codebook design and MUD algorithms in order.

2.1.1 SCMA system model

We consider an uplink overload SCMA system where J users share K orthogonal REs for the signal transmission with $J > K$. Each user occupies d_v ($d_v < K$) REs. We assume that the number of users linked to each RE is the same, denoted as $d_f = Jd_v/K$, and ideally $d_f \ll J$ to maintain the sparsity feature. We introduce a general uncoded SCMA system in this section while coded scenarios will be discussed in chapter 3

At the transmitter, an SCMA mapper consists of codeword mapping and spreading. The mapping procedure for user $j \in \{1, 2, \cdots, J\}$ is expressed as $g_j : \mathbf{b}_j \rightarrow \mathbf{d}_j$, where $\mathbf{b}_j \in \{0, 1\}^{\log_2(M)\times 1}$ is the binary message and $\mathbf{d}_j \in \mathbb{C}^{d_v \times 1}$ is the d_v-dimensional constellation point selected from an sequence $\mathcal{D}_j \subset \mathbb{C}^{d_v \times 1}$ with the size M. Then a $K \times d_v$ binary mapping matrix $\mathbf{S}_j$ spreads the d_v-dimensional constellation point of user j over the K REs, written as

$$\mathbf{x}_j = \mathbf{S}_j \mathbf{d}_j \tag{2.1}$$

$\mathbf{x}_j \in \mathcal{X}_j$ is referred to as the transmitted codeword of user j. Accordingly, the size of the codebook is $|\mathcal{X}_j| = M$. $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_J$ denotes the multidimensional codebook of the SCMA transmission. The complex-valued received signal $\mathbf{y} \in \mathbb{C}^{K\times 1}$ at the receiver is written as

$$\mathbf{y} = \sum_{j=1}^{J} diag(\mathbf{h}_j)\mathbf{x}_j + \mathbf{n} \tag{2.2}$$

where $\mathbf{h}_j \in \mathbb{C}^{K\times 1}$ is the channel vector from user j to the receiver, and $\mathbf{n} \in \mathbb{C}^{K\times 1}$ is the additive white Gaussian noise (AWGN) vector whose entries follow $\mathcal{N}_c(0, \sigma^2)$. If the mapping procedure in Eq.(2.1)

is considered in the received signal, Eq.(2.2) can be rewritten as

$$\mathbf{y} = \sum_{j=1}^{J} diag(\mathbf{h}_j)\mathbf{S}_j\mathbf{d}_j + \mathbf{n} \tag{2.3}$$

$$= \mathbf{G}\mathbf{d} + \mathbf{n} \tag{2.4}$$

where $\mathbf{G} = [\mathbf{g}_1, \cdots, \mathbf{g}_J] \in \mathbb{C}^{K \times Jd_v}$ is called effective channel gain matrix, where $\mathbf{g}_j = diag(\mathbf{h}_j)\mathbf{S}_j$ is the effective channel gain matrix of user j. $\mathbf{d} = [\mathbf{d}_1^\intercal, \cdots, \mathbf{d}_J^\intercal]^\intercal \in \mathcal{D}$ represents the transmitted effective codeword vector where $\mathcal{D} = \mathcal{D}_1 \times \cdots \times \mathcal{D}_J$ is the effective codebook. It is noticeable that Eq.(2.4) represents the SCMA transmission as a linear system.

Every column of the binary mapping matrix $\mathbf{S}_j$ has only one non-zero element. Usually, the mapping relation of $\mathbf{S}_j$ can also be represented by the j-th column of the factor graph matrix $\mathbf{F} = [\mathbf{f}_1, \cdots, \mathbf{f}_J] \in \{0,1\}^{K \times J}$ which indicates the resource allocation of the transmission. The relation between $\mathbf{f}_j$ and $\mathbf{S}_j$ can be expressed as $diag(\mathbf{f}_j) = \mathbf{S}_j\mathbf{S}_j^\intercal$. A benchmark example of $\mathbf{F}$ is given below :

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \end{bmatrix} \tag{2.5}$$

which has $J = 6$, $K = 4$, $d_v = 2$ and $d_f = 3$. Fig. 2.1 depicts the example matrix $\mathbf{F}$ in Eq.(2.5) graphically, in which the circular nodes represent the users and the square nodes represent the REs. The overlapping of users' codewords is illustrated Fig. 2.1 also. Thus the mapping matrix of the first user is

$$\mathbf{S}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{2.6}$$

By concatenating the mapping matrix of all users, we can obtain the mapping matrix of the transmission system

$$\mathbf{S} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \tag{2.7}$$

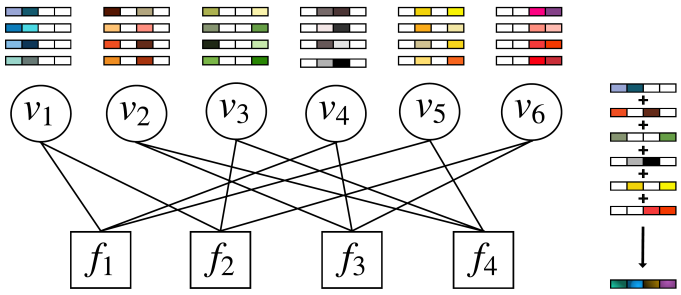


FIGURE 2.1 – Factor graph example of an SCMA system.

2.1.2 Multidimensional codebooks

The pre-designed multidimensional codebooks play an essential role in SCMA system performance [62]. The design problem of SCMA codebook can be formulated as [40]

$$\mathbf{S}^*, \mathcal{D}^* = \arg\max_{\mathbf{S},\mathcal{D}} \mathscr{C}(\mathcal{S}(\mathbf{S}, \mathcal{D}; J, K, M, d_v)) \tag{2.8}$$

where $\mathscr{C}$ is a given design criterion and $\mathcal{S}(\mathbf{S}, \mathcal{D}; J, K, M, d_v)$ refers to the SCMA codebook structure. The design problem in Eq. (2.8) is equivalent to a joint optimization of $\mathbf{S}$ and $\mathcal{D}$. In one of the first works on SCMA codebook design, authors of paper [40] introduce a multistage method in order to simplify the design steps of the set of users' codebooks. The sophisticated task of multidimensional codebook design is divided into the design of factor graph matrix, mother constellation and rotation [40, 63]. Usually, it is assumed that the factor graph is given, which means that the transmission mapping matrix $\mathbf{S}$ is given as well. As a result, the design problem becomes

$$\mathbf{S}, \mathcal{D}^* = \arg\max_{\mathcal{D}} \mathscr{C}(\mathcal{S}(\mathbf{S}, \mathcal{D}; J, K, M, d_v)) \tag{2.9}$$

which is equivalent to the optimization of $\mathcal{D}$ based on the given SCMA codebook structure.

Since the mother constellation is critical to SCMA codebook design, papers [64, 65] focus on optimizing the mother constellation by maximizing figures of merit (FoMs), such as minimum Euclidean distance (MED) and minimum product distance (MPD). Paper [62] illustrates the benefits of exploiting good FoMs in the design of SCMA codebooks by performing a comprehensive performance investigation. Besides, authors in [66] consider system capacity as metric in the criterion of SCMA codebook design.

For coded SCMA systems, in paper [67], the authors take advantage of EXITcharts in the codebook design and bit labeling optimization. Jiang and Wang in paper [68] combine probabilistic shaping and

geometric shaping simultaneously in uplink SCMA codebook design which shows the superiority in terms of simulation performance and average mutual information.

2.1.3 MUD algorithms

Based on the SCMA transmission expressed in Eq.(2.4), assume that the channel coefficients are estimated perfectly at the receiver, the maximum likelihood (ML) detection of SCMA is given as

$$\hat{\mathbf{d}} = \arg\max_{\mathbf{d}\in\mathcal{D}} p(\mathbf{y}|\mathbf{d},\mathbf{G}) \tag{2.10}$$

$$= \arg\min_{\mathbf{d}\in\mathcal{D}} \|\mathbf{y}-\mathbf{G}\mathbf{d}\|^2 \tag{2.11}$$

The ML algorithm can provide optimal MUD performance by joint decoding at the cost of exhaustively testing all combinations of the transmitted effective codewords. The complexity order of the ML algorithm is $\mathcal{O}(M^J)$. The ML receiver becomes infeasible when the user number or codebook size is large.

Due to the sparsity of the codewords, the joint MUD can be performed using the message passing algorithm (MPA). After iteratively updating the probability messages between users and REs, MPA outputs the decoded signals which have the maximum *a posteriori* probability. In the MPA, the circular nodes and square nodes shown in Fig. 2.1 are known as variable nodes (VNs) and function nodes (FNs), respectively. d_f and d_v are also known as degrees of function node and variable node, respectively. v_j ($j \in \{1,\cdots,J\}$) represents the j-th VN and f_k ($k \in \{1,\cdots,K\}$) represents the k-th FN. Furthermore, $V(j)$ denotes the neighboring FNs of v_j and $F(k)$ denotes the neighboring VNs of f_k. The number of iterations is denoted as T. Following the message-update rules, the message from v_j to f_k and the message from f_k to v_j at the t-th iteration are defined as [69, 70]

$$\mu^t_{v_j\to f_k}(\mathbf{x}_j) = \frac{\mu_{\phi_j\to v_j}(\mathbf{x}_j)\cdot\prod\limits_{l\in V(j)\backslash k}\mu^{t-1}_{f_l\to v_j}(\mathbf{x}_j)}{\sum\limits_{\mathbf{x}_j\in\mathcal{X}_j}\mu_{\phi_j\to v_j}(\mathbf{x}_j)\cdot\prod\limits_{l\in V(j)\backslash k}\mu^{t-1}_{f_l\to v_j}(\mathbf{x}_j)} \tag{2.12}$$

and

$$\mu^t_{f_k\to v_j}(\mathbf{x}_j) = \sum_{\mathbf{x}'_j\in\tilde{\mathcal{X}}_j} f_k(y_k|\mathbf{x}'_j)\cdot\prod_{q\in F(k)\backslash j}\mu^t_{v_q\to f_k}(\mathbf{x}_q) \tag{2.13}$$

where $\mu_{\phi_j\to v_j}(\mathbf{x}_j)$ denotes the *a priori* information of codeword $\mathbf{x}_j$ and $\tilde{\mathcal{X}}_j = \mathcal{X}_1\times\cdots\times\{\mathbf{x}_j\}\times\cdots\times\mathcal{X}_J$.

The function $f_k(y_k|\mathbf{x}'_j)$ calculates channel transition probability, defined by

$$f_k(y_k|\mathbf{x}'_j) = \frac{1}{\pi\sigma^2}\exp\left(-\frac{|y_k - \sum_{j\in F(k)} h_{k,j}x'_{k,j}|^2}{\sigma^2}\right) \tag{2.14}$$

where $x'_{k,j}$ represents the k-th element of vector $\mathbf{x}'_j$.

After T iterations, the output of MPA depends on the system scenario. In uncoded SCMA scenario, the output is hard decision, expressed as

$$\hat{b}_j^i = \begin{cases} 1, & \Lambda_1(\hat{b}_j^i) \geqslant 0 \\ 0, & \Lambda_1(\hat{b}_j^i) < 0 \end{cases} \tag{2.15}$$

where

$$\Lambda_1(\hat{b}_j^i) = \ln \frac{\sum\limits_{\mathbf{x}_j\in\mathcal{X}_j^+} \mu_{\phi_j\to v_j}(\mathbf{x}_j)\cdot \prod\limits_{k\in V(j)} \mu^T_{f_k\to v_j}(\mathbf{x}_j)}{\sum\limits_{\mathbf{x}_j\in\mathcal{X}_j^-} \mu_{\phi_j\to v_j}(\mathbf{x}_j)\cdot \prod\limits_{k\in V(j)} \mu^T_{f_k\to v_j}(\mathbf{x}_j)} \tag{2.16}$$

is the *a posteriori* log-likelihood ratio (LLR) of $\hat{b}_j^i$. $\mathcal{X}_j^+$ and $\mathcal{X}_j^-$ denote the subset of $\mathbf{x}_j$ in which the $\hat{b}_j^i$ bit has the value of 1 and 0, respectively. The detection procedure of MPA is concluded in Algorithm 3

Algorithm 3 MPA.

Initialization : $\mu^0_{v_j\to f_k} = 0$, $\mu^0_{f_k\to v_j} = 1$, $t = 1$.
while $t \leqslant T$ **do**
  for $j = 1 \to J$ **do**
    for $k \in V(j)$ **do**
      Calculate message $\mu^t_{v_j\to f_k}(\mathbf{x}_j)$ using Eq.(2.12).
    end for
  end for
  for $k = 1 \to K$ **do**
    for $j \in F(k)$ **do**
      Calculate message $\mu^t_{f_k\to v_j}(\mathbf{x}_j)$ using Eq.(2.13) and Eq.(2.14).
    end for
  end for
  $t = t + 1$.
end while
for $j = 1 \to J$ **do**
  Calculate hard (soft) output using Eq.(2.15) (Eq.(2.16)).
end for

As demonstrated above, MUD performance of MPA is determined by the number of iterations, that is the larger number of iterations is the better MUD performance it has. Still, the computational complexity of MPA increases dramatically with the codebook size, number of users and overloading

factor. To reduce the complexity of MPA, some researchers propose Log-MPA and Max-Log-MPA which calculate the probability messages in the logarithm domain [71], while others propose edge selection criteria to decrease the number of interfering users considered in the message calculation [72, 73]. Other iterative algorithms, such as expectation propagation algorithm (EPA) and Gaussian approximation algorithm (GAA), show near MPA performance and substantial complexity decrease in coded SCMA systems [17, 18].

Meanwhile, sphere decoding (SD) has recently been considered in the MUD of SCMA systems. SD, firstly proposed as a resolution of the closest lattice point search problem [74, 75], has been regarded as an effective scheme in MIMO detection [76]. The core of SD is to search the nearest lattice point to the received signal within a predefined squared sphere radius. By leveraging matrix operation, e.g. QR decomposition, SD obtains an upper triangular matrix and transforms the detection into sequential decoding by the tree search process. The computational complexity of SD is proportional to the number of nodes visited in the tree search. We refer the interested readers to AnnexA where the details of the SD for a simple MIMO transmission example is provided.

The following is a brief literature review on SD-based detector for SCMA. Authors in [77] adopt list SD in the MPA for the purpose of pruning the candidates that are outside of the search space when calculating the probability messages. To employ SD to the entire SCMA system, in papers [19, 20], the authors reformulate some multidimensional codebooks with special structure and transform SCMA detection into binary lattice constellation detection. Although the proposed detectors in [19, 20] show better performance and lower complexity than the MPA, they can not be compatible with generalized SCMA multidimensional codebook. Aiming at improving the generality, SD for SCMA (SD-SCMA) is proposed in paper [21], which does not require specific codebook reformulation. Limited by the overload feature of SCMA transmission, SD-SCMA constrains the transmitted codewords to have constant modulus. To overcome this constraint, authors in [22] propose general SD-SCMA (GSD-SCMA) which splits the detection into two parts. As GSD-SCMA needs to do brute-force search on partial transmitted effective codewords, the complexity may surge when codebook size or the number of users is large.

2.2 NC MIMO transmission

This section sheds light on the transmission model and constellation design in NC MIMO systems.

2.2.1 NC MIMO system model

We consider a NC MIMO communication system where the transmitter and receiver have N_t and N_r antennas, respectively. The channel coefficients are assumed to remain constant for a coherence interval of T and change to a new independent realization in the next time period. This channel model is also called block-fading channel. We assume that $N_t = \min\{\lfloor \frac{T}{2} \rfloor, N_r\}$. The transmitted signal is represented by the transmitted symbol matrix $\mathbf{X} \in \mathbb{C}^{T \times N_t}$ drawn from the codebook $\mathcal{C}$ with cardinal $|\mathcal{C}| = M$. The received signal $\mathbf{Y} \in \mathbb{C}^{T \times N_r}$ is expressed as

$$\mathbf{Y} = \mathbf{X}\mathbf{H} + \sqrt{\frac{N_t}{\rho T}}\mathbf{W} \tag{2.17}$$

where $\mathbf{H} \in \mathbb{C}^{N_t \times N_r}$ is the channel matrix, ρ is the signal-to-noise ratio (SNR) per receive antenna and $\mathbf{W} \in \mathbb{C}^{T \times N_r}$ is the AWGN matrix. The entries of $\mathbf{W}$ are drawn independently from $\mathcal{N}_c(0, 1)$. Accordingly, the power of the transmitted signal follows

$$\sum_{t=1}^{T} \mathbb{E}[|x_{t,n}|^2] = 1 \tag{2.18}$$

where $x_{t,n}$ is the element in the t-th row and n-th column of matrix $\mathbf{X}$.

In NC MIMO systems, although CSI, instantaneous values of the channel coefficients, are not known at any node, the channel distribution information which can be tracked with negligible transmission resource is known at any node. Kronecker model [78] is used to describe the considered channel model, which is written as

$$\mathbf{H} = (\mathbf{R}^t)^{1/2}\mathbf{G}(\mathbf{R}^r)^{1/2} \tag{2.19}$$

where the elements of the matrix $\mathbf{G} \in \mathbb{C}^{N_t \times N_r}$ are i.i.d. and drawn from $\mathcal{N}_c(0, 1)$, and $\mathbf{R}^t$ and $\mathbf{R}^r$ are called the transmit and receive covariance matrix, respectively. The Kronecker model is generalized to represent both correlated fading and i.i.d. Rayleigh channel. For the correlated fading channel, we assume that only the receive antennas are correlated and the receive covariance matrix is modelled using the exponential model. Therefore, the transmit covariance matrix is $\mathbf{R}^t = \mathbf{I}_{N_t}$ and the entries of

the receive covariance matrix are defined as

$$R_{i,j}^{r} = \{r^{|i-j|}\}, \quad r \in [0,1) \tag{2.20}$$

where $i, j \in \{1, 2 \cdots, N_r\}$ are the row and column indices of matrix $\mathbf{R}^r$ respectively. For the i.i.d. Raleigh channel, both $\mathbf{R}^t$ and $\mathbf{R}^r$ are identity matrices.

As the receiver of NC MIMO has no prior CSI knowledge, the ML detection is based on the conditional probability density of the received signals, given as [23]

$$\begin{aligned}
\hat{\mathbf{X}} &= \arg\max_{\mathbf{X}\in\mathcal{C}} p(\mathbf{Y}|\mathbf{X}) \\
&= \arg\max_{\mathbf{X}\in\mathcal{C}} \frac{\exp\left(-\mathrm{Tr}\left\{(\mathbf{I}_T + \mathbf{X}\mathbf{X}^\dagger)^{-1}\mathbf{Y}\mathbf{Y}^\dagger\right\}\right)}{\pi^{TN_r}\det^{N_r}(\mathbf{I}_T + \mathbf{X}\mathbf{X}^\dagger)} \\
&= \arg\max_{\mathbf{X}\in\mathcal{C}} \frac{\exp\left(-\mathrm{Tr}\left\{(\mathbf{I}_T - \frac{1}{2}\mathbf{X}\mathbf{X}^\dagger)\mathbf{Y}\mathbf{Y}^\dagger\right\}\right)}{\pi^{TN_r}2^{MN_r}} \\
&= \arg\max_{\mathbf{X}\in\mathcal{C}} \mathrm{Tr}\{\mathbf{Y}^\dagger\mathbf{X}\mathbf{X}^\dagger\mathbf{Y}\}
\end{aligned} \tag{2.21}$$

The above detection is also called GLRT which is the optimal receiver for NC MIMO.

2.2.2 Constellation design for NC MIMO systems

As derived in [24], the capacity-achieving signal of NC MIMO with flat-fading channels is represented as

$$\mathbf{X} = \mathbf{\Theta}\mathbf{D} \tag{2.22}$$

where $\mathbf{\Theta}$ is a $T \times T$ isotropically distributed unitary matrix, $\mathbf{\Theta}^\intercal\mathbf{\Theta} = \mathbf{I}_{N_t}$, and $\mathbf{D}$ is a $T \times N_t$ independent real, non-negative diagonal matrix. Considering $N_t = \min\{\lfloor \frac{T}{2} \rfloor, N_r\}$, it is optimal to let $\mathbf{D}$ become an identity matrix [24]. Thus, the transmitted signal is $\mathbf{X} = \mathbf{\Theta}$ following the constraint

$$\mathbf{X}^\dagger\mathbf{X} = \mathbf{I}_{N_t} \tag{2.23}$$

In other words, the transmitted signals are mutually orthogonal with respect to time among the transmitter antennas. The transmitted symbol matrix $\mathbf{X}$ is referred to as unitary space-time modulation (USTM) [51]. The critical observation of USTM in NC MIMO is that the channel matrix $\mathbf{H}$ only scales and rotates the bases of $\mathbf{X}$ and leaves the subspace $\Omega_{\mathbf{X}}$, spanned by the column vectors of $\mathbf{X}$, unchanged. This reveals that the information-carrying object in NC MIMO is the random subspace

$\Omega_{\mathbf{X}}$ rather than the input signal $\mathbf{X}$ itself in coherent MIMO [6]. The number of degree of freedom is $N_t(T-N_t)$ in the noncoherent case which is lower than that in the coherent case equal to N_tT. From this observation, a constellation of NC MIMO systems is a set of N_t-dimensional linear subspace of T-dimensional complex Euclidean space, denoted as $\mathbb{C}^T$ [6]. Since Grassmann manifold $\mathbb{G}_{N_t}(\mathbb{C}^T)$ is the set of all N_t-dimensional subspace of $\mathbb{C}^T$, each of the N_t-dimensional linear subspace of $\mathbb{C}^T$ represented by a $T \times N_t$ unitary matrix can be deemed as a single point on the Grassmann manifold. Accordingly, each transmitted symbol matrix $\mathbf{X}$ represents a single point on the compact Grassmann manifold, expressed as $\mathbf{X} \in \mathbb{G}_M(\mathbb{C}^T)$ [6]. Constellations in NC MIMO systems are also called Grasmannian constellations.

In [6], the authors provide a geometric interpretation of ergodic capacity in NC MIMO which is sphere packing in the Grassmann manifold. The optimal packing aims at minimizing the detection error under noise observation. Based on this Grassmannian packing interpretation, many researchers have proposed different Grassmannian constellations. There have been two classes of approaches to design Grassmannian constellations. The first approach adopts numerical optimization tools to solve the sphere packing problem by maximizing the minimum symbol pairwise distance [25, 26, 27, 28] or minimizing directly the error probability upper bound [29]. As discussed in [25], a fundamental issue of designing Grassmannian constellations is to find a sensible metric that is used to measure distances between any two constellation points. As the combined effect of noise and channel fading brings about a perturbation of the subspace $\Omega_{\mathbf{X}}$, paper [26] analyzes the subspace perturbation of the received signal in NC MIMO systems. It illustrates that chordal Frobenius distance (norm) is the appropriate metric for the Grassmannian constellation design regarding the ergodic rate of the transmission. The chordal Frobenius distance between two constellation points, $\mathbf{X}_1 \in \mathcal{C}$ and $\mathbf{X}_2 \in \mathcal{C}$ ($\mathbf{X}_2 \neq \mathbf{X}_1$), is expressed as

$$d(\mathbf{X}_1, \mathbf{X}_2) = \sqrt{2N_t - 2\,\mathrm{Tr}(\mathbf{\Sigma}_{\mathbf{X}_1,\mathbf{X}_2})} \tag{2.24}$$

where $\mathbf{\Sigma}_{\mathbf{X}_1,\mathbf{X}_2}$ is the diagonal matrix containing the singular values of $\mathbf{X}_1^\dagger \mathbf{X}_2$. Therefore, the problem of design a Grassmannian constellation becomes an optimization problem given as

$$\min_{\mathcal{C}} \max_{1 \leqslant i < j \leqslant M} \mathrm{Tr}(\mathbf{\Sigma}_{\mathbf{X}_i,\mathbf{X}_j}) \tag{2.25}$$

$$\text{subject to } \mathbf{X}_i \in \mathbb{G}_M(\mathbb{C}^T), \mathbf{X}_j \in \mathbb{G}_M(\mathbb{C}^T) \tag{2.26}$$

where $\mathbf{X}_m$ denotes the m-th codeword of the codebook $\mathcal{C}$. Based on this optimization problem, three different approaches are proposed in [26]. The resulting constellations of the first class of approaches

Algorithm 4 Lloyd algorithm.

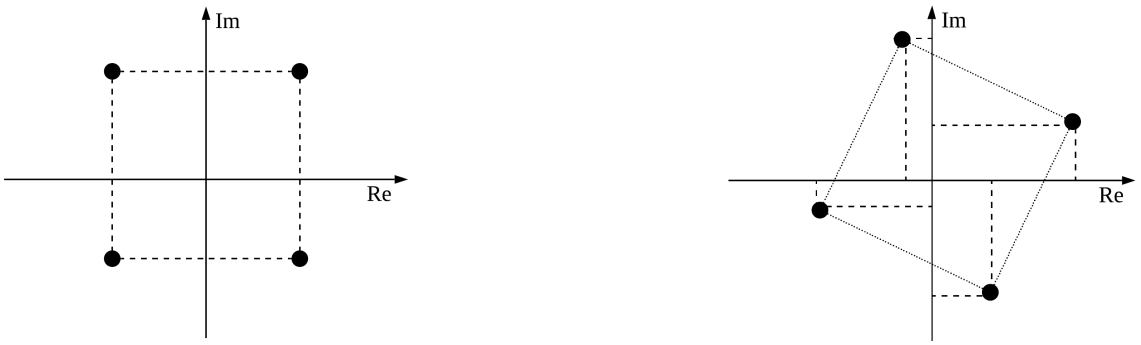
Initialization : Generate M random codeword $\mathbf{X} \in \mathbb{V}_{T,N_t}$. Generate a training sequence consisting of I number of source matrices $\mathbf{V} \in \mathbb{V}_{T,N_t}$. Given the number of iterations L. $l = 0$.
while $l < L$ **do**
  for $i \in 1, 2, \cdots, I$ **do**
    if $d(\mathbf{V}_i, \mathbf{X}_m) < d(\mathbf{V}_i, \mathbf{X}_{m'})$, $\forall m' \neq m$ **then**
      $\mathbf{V}_i \in S_m$, S_m is the neighborhood of $\mathbf{X}_m$
    end if
  end for
  for $m \in 1, 2, \cdots, M$ **do**
    $\Sigma_m = \frac{1}{|S_m|} \sum_{\mathbf{V} \in S_m} \mathbf{V}\mathbf{V}^\dagger$
    $\mathbf{X}_m = \mathbf{U}_{N_t}$, $\mathbf{U}_{N_t}$ is built from the N_t eigenvectors corresponding to the largest eigenvalue of Σ_m
  end for
  $l = l + 1$
end while

can reach a good distance spectrum while not having any particular structure. When the projection Frobenius distance (norm) is chosen as the distance measure for the sphere packing, Lloyd algorithm can be employed to construct Grassmannian constellations as well [28]. The generalized Lloyd algorithm is concluded in Algorithm 4,where $d(\mathbf{X}_1, \mathbf{X}_2)$ is the projection Frobenius distance between constellation points $\mathbf{X}_1$ and $\mathbf{X}_2$.

On the other hand, the second class of approaches introduces particular structure into the Grassmannian constellation design. For instance, the authors in paper [30] propose a parameterized mapping of unitary matrices for Grassmannian constellations in NC MIMO systems. In contrast, the second class has simplicity in terms of design process but stringent structure of Grassmannian constellations, which cannot become a genetic solution.

2.3 SU multidimensional transmission

In this section, we introduce the simplified base system model, called SU multidimensional transmission, for some promising OFDM-based techniques, such as OFDM-IM and GLCP, and the associated MDC design.



(a) Constellation with diversity 1.

(b) Constellation with diversity 2.

FIGURE 2.2 – Illustration of constellation diversity.

2.3.1 Simplified base system model

In the following, we first present the considered base system model, then introduce some practical multicarrier transmission schemes that can be adapted to fit the generalized base system model.

Before diving into the base system model, we would like to talk about the diversity order, also known as SSD which is essential to MDC performance. The diversity of a MDC is the minimum number of distinct (real) components between any two constellation points, which is illustrated in Fig. 2.2 taking the quadrature phase shift keying (QPSK) constellation an an instance. Full diversity, exemplified in Fig. 2.2b, is suggested to obtain diversity gain in fading channels without sacrificing spectral efficiency [31]. In the example, all the points of the constellation shown in Fig. 2.2b can be distinguished at the receiver when one (real) dimension is affected by a deep fading in the absence of noise.

Therefore, to obtain the gain of diversity for MDC design, the base system model considers a component-wise interleaving which destroys the correlation between in-phase and quadrature channel fading coefficients [31]. Assuming ideal interleaving, the channel is characterized by $2N$ orthogonal real resources, and the i.i.d. channel coefficients follow a Rayleigh distribution with unit second moment. A block of N complex symbols is treated as a $2N$-dimensional point in the (real) Euclidean space. As a result of the above assumptions, the system model is expressed as

$$\mathbf{y} = diag(|\tilde{\mathbf{h}}|)\mathbf{x} + \mathbf{n} \tag{2.27}$$

where $\mathbf{x} \in \mathbb{R}^{2N\times 1}$ is the transmitted codeword drawn from the codebook $\mathcal{C}$ of size M, $\tilde{\mathbf{h}} \in \mathbb{C}^{2N\times 1}$ is complex channel vector, following $\tilde{\mathbf{h}} \sim \mathcal{N}_c(\mathbf{0}_{2N\times 1}, \mathbf{I}_{2N})$, $|\cdot|$ represents the element-wise absolute

operator for vectors and $\mathbf{y} \in \mathbb{R}^{2N\times 1}$ is the received signal. and $\mathbf{n} \sim \mathcal{N}(\mathbf{0}_{2N\times 1}, \frac{N_0}{2}\mathbf{I}_{2N})$ is the AWGN vector. $\frac{N_0}{2}$ is the power of noise per real dimension. The power constraint of the codebook is defined as

$$E_s = \mathbb{E}\left[\frac{\|\mathbf{x}\|^2}{2N}\right] = \frac{1}{2} \tag{2.28}$$

and the system SNR is defined as

$$\text{SNR} = \frac{\mathbb{E}[\|\mathbf{x}\|^2/2N]}{N_0/2} = \frac{1}{N_0} \tag{2.29}$$

The ML algorithm achieves optimal decoding performance in the SU multidimensional transmission. We assume that the CSI is known perfectly at the receiver, the ML receiver is written as

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}\in\mathcal{C}} p(\mathbf{y}|\mathbf{x}, \tilde{\mathbf{h}}) \tag{2.30}$$

The above base system model can be interpreted by a mother OFDM-based system model as illustrated in Fig. 2.3. In the transmission, N_s number of subcarriers are split into G groups, where each group contains $N = N_s/G$ subcarriers. The signal of each group is modulated by a M-ary multidimensional modulation. Accordingly, total number of $G\log_2 M$ information bits are divided into G groups of $\log_2 M$ bits. At the transmitter, in each group, a bit vector $\mathbf{b}_g$ $(1 \leqslant g \leqslant G)$ is mapped into a real-valued vector $\mathbf{x}_g \in \mathbb{R}^{2N\times 1}$. After parallel-to-serial conversion, the $2GN$ real-valued components are interleaved based on the strategy proposed in [33]. Next, real-to-complex conversion and serial-to-parallel conversion are realized. The following five blocks, inverse fast Fourier transform (IFFT) & cyclic prefix (CP) addition, parallel-to-serial conversion, fading channel, serial-to-parallel conversion and CP removal & IFFT are standard processes in an OFDM transmission. Afterwards, the complex-valued received symbols are serialized and converted to real-valued symbols. After the deinetrleaving and serial-to-parallel conversion, the received signals are split back into G real-valued vectors, $\mathbf{y}_g \in \mathbb{R}^{2N\times 1}$ $(1 \leqslant g \leqslant G)$. The signal of each group, $\mathbf{y}_g$, fit in the generalized system model in Eq.(2.27). When G is equal to N_s, the mother OFDM-based system is equivalent to a conventional OFDM system. The mother OFDM-based system can be deemed as grouped OFDM. One of the advantages of the grouped OFDM is that the detection complexity drops from $\mathcal{O}(N_s^M)$ to $\mathcal{O}(GN^M)$.

In the following, we will instance two practical OFDM-based schemes which can be adapted to fit in the mother system.

OFDM-IM : Different from the conventional OFDM, OFDM-IM allows the information bits to be carried not only by M-ary signal constellations but also by the indices of the subcarriers [58].

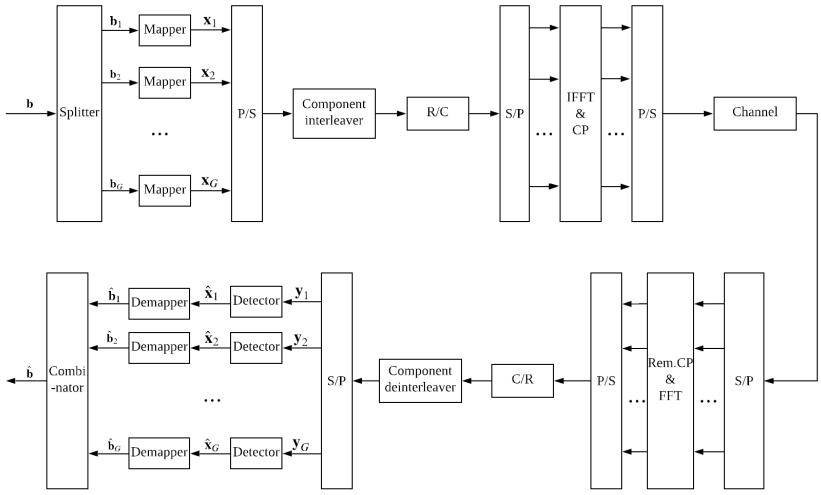


FIGURE 2.3 – Mother OFDM-based system diagram.

In every group, $N'(N' < N)$ out of N OFDM subcarriers, are activated each time an OFDM symbol is transmitted. The rest of the subcarriers in the group remain inactive sending zeros. Those N' indices are selected by a predefined procedure, based on the first $p_1 = \lfloor \log_2 \binom{N}{N'} \rfloor$ number of bits. The remaining $p_2 = \log_2 M$ bits are mapped into the M-ary real-valued vector $\mathbf{x}_g \in \mathbb{R}^{2N' \times 1}$ that modulates the active subcarriers. Therefore, each group can transmit $p = p_1 + p_2$ bits information in total.

GLCP : To circumvent the loss of multipath diversity, authors in [79] introduce GLCP for OFDM over frequency-selective fading channels. An optimal subcarrier grouping is proposed to divide the subcarriers into subsets which can be seen as complex version of the component mentioned interleaving in the mother system. The symbol vector of a group $\mathbf{x}_g$ is multiplied by a precoding matrix with shape $2N \times 2N$. Consequently, each subcarrier transmits the linear combination of the symbols of all the subcarriers within the group, which helps to maximize the diversity gain.

2.3.2 MDC design

Conventionally, MDCs can either be constructed through the Cartesian product of classical modulations, for example quadrature amplitude modulation (QAM), or be carved from lattices [32]. The problem of designing good MDCs is dependent on the transmission channel. In other words, the codebooks optimized for the Gaussian channel (may) have poor performance when adopted over the Rayleigh fading channel, and vice versa. For example, paper [80] studies *constellation figure of merit*

based on MED for MDC design in optical transport, as the channel is similar with the AWGN channel. The authors in [31] justify the significance of exploiting SSD for MDCs in order to achieve diversity gains over Rayleigh fading channels. In this thesis, we concentrate only on MDC design over Rayleigh fading channels.

To leverage the diversity gains, constellation rotation have been considered in the design of MDCs, where an exhaustive search have been applied to find the optimal rotation angle [31]. The methodology of combining GLCP scheme and rotation is proposed and studied in [33, 34] in order to achieve the full diversity. This methodology can be applied to OFDM-IM scheme as a special case. On the other hand, the author of paper [35] propose a combinatorial approach to construct full diversity MDCs. This approach provides a generic methodology and requires little storage. The combinatorial approach consists in finding good permutations of the labels of points in a fixed grid. Therefore, as long as there are M distinct points in the grid, the constructed MDCs always have full diversity. Additionally, the next-generation handheld digital video broadcasting standards [81], as a practical system, exploits diversity gains when constructing MDCs and achieves superior capacity.

2.4 DL techniques used in the thesis

In the following, we introduce three DL techniques, fully-connected DNNs, AEs and 2-dimensional (2D) convolutional layer, that are involved in this thesis.

2.4.1 Fully-connected DNNs

A fully-connected DNNs refers to a neural network model with several successive nonlinear layers of neurons where all the neurons in one layer are connected to all the neurons in the next layer. Fig. 2.4 demonstrates a fully-connected DNN example with 2 hidden layers where round nodes represent neurons. The number of neurons at the k-th ($k = \{1, \cdots, K+1\}$) layer of a fully-connected DNN is denoted by N_k, where K is the number of hidden layers. N_0 denotes the number of neurons at the input layer. The input at the input layer is denoted as $\mathbf{s}_0 \in \mathbb{R}^{N_0 \times 1}$. If $\mathbf{s}_{k-1} \in \mathbb{R}^{N_{k-1} \times 1}$ is the output at the $(k-1)$-th layer (also the input at the k-th layer), the output vector at the k-th layer is written as

$$\mathbf{s}_k = \alpha_k(\mathbf{W}_k \mathbf{s}_{k-1} + \mathbf{b}_k) \tag{2.31}$$

$$= f_k(\mathbf{s}_{k-1}) \tag{2.32}$$

where α_k, $\mathbf{W}_k \in \mathbb{R}^{N_k \times N_{k-1}}$ and $\mathbf{b}_k \in \mathbb{R}^{N_k \times 1}$ are activation function at the k-th layer, weight matrix and bias matrix respectively. Function $f_k(\cdot)$ represents the calculation at the k-th layer. The parameter set of the k-th layer is denoted as $\boldsymbol{\theta}_k = \{\mathbf{W}_k, \mathbf{b}_k\}$ and the parameter set of the fully-connected DNN is denoted by $\boldsymbol{\theta} = \{\boldsymbol{\theta}_1, \cdots, \boldsymbol{\theta}_{K+1}\}$. Accordingly, the output of the example shown in Fig. 2.4 can be formulated as $\mathbf{s}_3 = f_3(f_2(f_1(\mathbf{s}_0))) = f_{\text{DNN}}(\mathbf{s}_0)$, where function $f_{\text{DNN}}(\cdot)$ represents the computation of the example fully-connected DNN. For a supervised learning, the training objective of a neural network is to minimize the loss function. If the loss function is defined as cross-entropy, it is computed as

$$\mathcal{L}(\boldsymbol{\theta}) = -\frac{1}{|\mathcal{S}|} \sum_{(\mathbf{s}_0^{(i)}, \tilde{\mathbf{s}}_{K+1}^{(i)}) \in \mathcal{S}} diag(f_{\text{DNN}}(\mathbf{s}_0^{(i)})) \log(\tilde{\mathbf{s}}_{K+1}^{(i)}) \tag{2.33}$$

where $\mathcal{S}$ is the training input, $|\mathcal{S}|$ represents the cardinality of the training set and $(\mathbf{s}_0^{(i)}, \tilde{\mathbf{s}}_{K+1}^{(i)})$ is a pair of input and desired output data. The common approach to optimize the parameter set $\boldsymbol{\theta}$ is to operate on a so-called mini-batch $\underline{\mathcal{S}}$ acting as random subset of $\mathcal{S}$ and perform the stochastic gradient descent (SGD) algorithm which updates the parameter set iteratively formulated as

$$\boldsymbol{\theta}_t = \boldsymbol{\theta}_{t-1} - \eta \nabla \underline{\mathcal{L}}(\boldsymbol{\theta}_{t-1}) \tag{2.34}$$

where subscript of $\boldsymbol{\theta}$ refers to the index of SGD performing, η is the learning rate often being a small decimal and $\nabla \underline{\mathcal{L}}(\boldsymbol{\theta})$ is the gradient of the loss function based on a mini-batch. For instance, if loss function is cross-entropy, $\underline{\mathcal{L}}_{\text{DNN}}(\boldsymbol{\theta})$ is calculated by

$$\underline{\mathcal{L}}(\boldsymbol{\theta}) = -\frac{1}{|\underline{\mathcal{S}}|} \sum_{(\mathbf{s}_0^{(i)}, \tilde{\mathbf{s}}_{K+1}^{(i)}) \in \underline{\mathcal{S}}} diag(f(\mathbf{s}_0^{(i)})) \log(\tilde{\mathbf{s}}_{K+1}^{(i)}) \tag{2.35}$$

In modern DL, back-propagation is an efficient approach to compute the gradients in Eq.(2.34). To facilitate the implementation of neural networks, different DL libraries have been developed. In this thesis, all the neural networks are built on the DL framework PyTorch [82].

Recently, full-connected DNNs have been involved into many wireless communications research for signal detection, channel decoding or prediction work in communication system. In [12], fully-connect DNN is trained for MIMO detection in a fixed channel case which shows near optimal performance. Gruber *et al.* [10] propose a fully-connected DNN for channel decoder which achieves maximum a posteriori bit error rate (BER) performance for short codeword lengths and easily learns structured codes. A full-connected DNN is proposed to predict initial radius for SD which effectively reduce complexity by 60% with marginal performance loss compared to conventional SD detector [83].

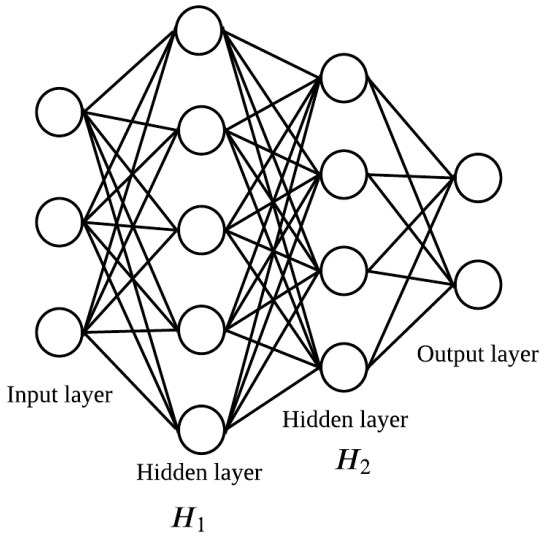


FIGURE 2.4 – An example of fully-connected DNN.

On the over hand, sparse DNNs, regarded as adapted fully-connected DNN, have drawn much attention due to its capability to implement sum-product algorithm. In contrary to fully-connected DNN, sparse DNNs allow a neuron to connect with only few or none of the neurons in the next layers. In paper [84], the researchers construct a sparse DNN based on the trellis representation of the standard belief propagation (BP) decoder for high density parity-check codes, which is motivated by the idea of implementing model-based BP algorithm on DNN. Thus, this sparse DNN can be regarded as parameterized BP decoder with the optimized parameters by the training and has comparable BER performance and substantially lower complexity with conventional BP decoder. Similarly, Lu *et al.* [85] propose a sparse DNN carrying out MPA which achieve a better MUD performance in SCMA systems. The application of sparse DNNs is an example of merging expertise in wireless communications into DL techniques.

2.4.2 AE techniques

Typically, an AE, consisting of an encoder and decoder neural network (NN), aims at reconstructing the input at the output. The encoder neural network maps the input into a low-dimensional representation and the decoder neural network maps the representation to a reconstruction of the input. Since 1980, AEs have been applied in dimensionality reduction and feature learning. Recently,

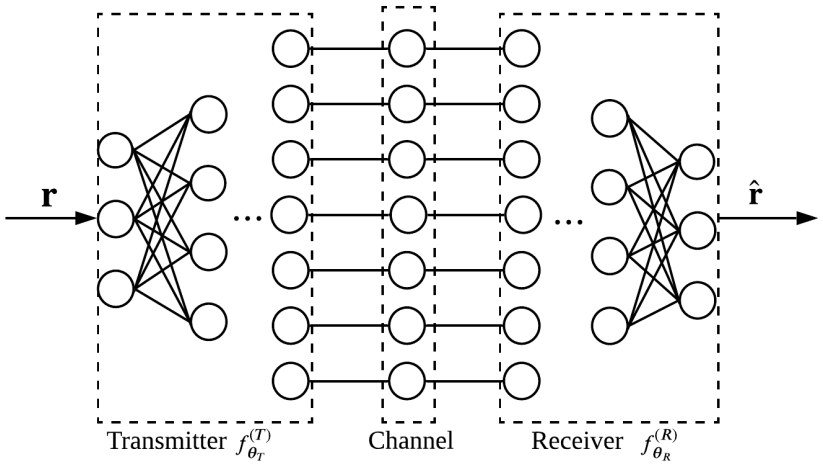


FIGURE 2.5 – A general autoencoder-based communication system.

the idea of AE for communication systems have been proposed because a communication system can be seen as a particular type of AE [86, 43]. In a communication system, the transmitter maps the transmitted messages to symbols and the receiver seeks to recover the transmitted messages with low probabilities of error. Different from traditional AEs compressing the input data into low-dimensional representations, AEs for communication systems aims at learning intermediate representations robust with respect to the channel and noise impairment. For simplicity, the AE for communication systems is called AE in the following.

AEs comprise encoder (transmitter) NN, channel layer and decoder (receiver) NN as demonstrated in Fig. 2.5. During the training, an AE learns to recover the transmitted messages with the channel and noise impairments. As the objective of an AE is to optimize the transmitter and receiver NN jointly for a particular channel, AE techniques are also called end-to-end learning in communication systems. An example cross entropy loss function of an AE is given as

$$\mathcal{L}_{\mathrm{AE}}(\boldsymbol{\theta}_T, \boldsymbol{\theta}_R) = -\frac{1}{|\mathcal{S}|} \sum_{\mathbf{r}^{(i)} \in \mathcal{S}} diag(\hat{\mathbf{r}}) \log(\mathbf{r}) \tag{2.36}$$

where $\mathcal{S}$ is the training dataset, $\boldsymbol{\theta}_T$ and $\boldsymbol{\theta}_R$ denote parameter set of the transmitter NN and receiver NN of the AE respectively.

AE can be employed to resolve different optimization problems in wireless communication systems. Since the output of the encoder NN can be regarded as modulated symbols, AE is applied for constellation design and shows its superiority over the constellations designed by block-based methods [43]. To further improve the system capacity, geometric and probabilistic shaping are performed jointly in the

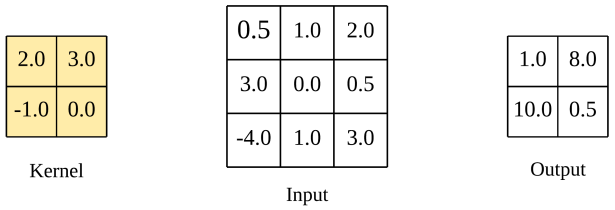
AE by leveraging the Gumbel-Softmax trick [87]. On the other hand, since AEs enable optimization of transmitter and receiver NN jointly, they have been proposed to optimize transceivers in optical fiber [88], radar and wireless coexistence [89] and optical wireless [90] communications. J. Tao *et al.* [91] propose an AE to address the hybrid beamforming design and decoding for multiuser millimeter wave (mmWave) MIMO systems, which exhibits better performance than the traditional linear processing methods.

To further understand and investigate AEs in practice, many researchers have integrated or built AEs in experimental physical layers of wireless systems. As the first work on this topic, authors in [92] build and run a complete communication system by implementing the transmitter and receiver NN of a trained AE, which demonstrates the possibility of AE over-the-air transmission. A two-phase training strategy is proposed [92], including an offline-training based on the concept of transfer learning [43] and an online fine-tuning of the receiver NN, to compensate the mismatch between channel model and actual channel. To remedy the problem of unknown channels or channels with non-differentiable components, a novel AE training approach called model-free approach is proposed [42]. Contrary to the conventional training method referred to as model-aware, model-free training performs the gradient descent of the transmitter and receiver NN separately with the help of a reliable feedback channel inspired by the concept of reinforcement learning. In [42], model-free AEs show good performance over actual channels in the experimental simulation. Furthermore, AE techniques have been extended to iterative demapping and decoding (IDD) structure which outperforms the conventional baseline in over-the-air experiment by applying model-free training method [93].

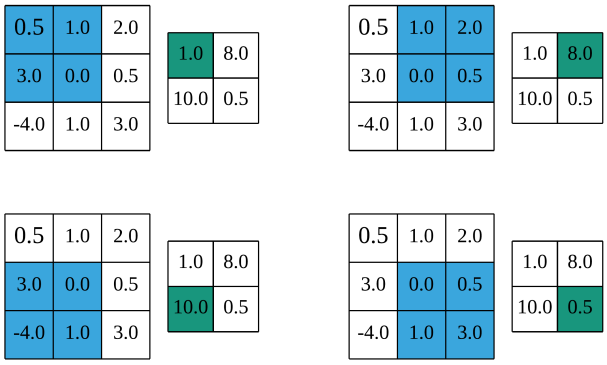
2.4.3 2D convolutional layer

Convolutional layers are essential components in convolutional neural networks (CNNs) which have achieved great success in image processing and natural language processing. Generally, (discrete) convolution is an orderly linear transformation that applying a filter also called kernel to the input data (signal) and obtaining the output data (signal).

Here, we focus on the 2D convolutional layer and explain the details of 2D convolution. As a common type of convolutional layer, a 2D convolutional layer allows a 2D kernel to slide over the 2D input. The size of the output depends on the size of the input as well as the choice of kernel size, zero padding, stride, dilation and number of groups. In this work, we only care about the kernel size and



(a) Example data.



(b) Computing the output values of a 2D convolution.

FIGURE 2.6 – Example of 2D convolutional layer.

stride. Other parameters are set by default. For further information about other parameters, please refer to [94].

To better illustrate the 2D convolutional layer and considered parameters, we take an example of a 2D convolution with kernel size $(2, 2)$ and stride $(1, 1)$ shown in Fig. 2.6. The kernel, input and output of the example are presented in Fig. 2.6a, and the computing process of the output values of 2D convolution is demonstrated in Fig. 2.6b. Denoting the size of the input as (J, R), the kernel as (K_1, K_2), and the stride as (S_1, S_2), the size of the output, (P, Q), is determined according to

$$P = \frac{J - (K_1 - 1) - 1}{S_1} + 1 \tag{2.37}$$

$$Q = \frac{R - (K_2 - 1) - 1}{S_2} + 1 \tag{2.38}$$

In the example of $(J, R) = (3, 3)$, $(K_1, K_2) = (2, 2)$ and $(S_1, S_2) = (1, 1)$, the size of the output is $(P, Q) = (2, 2)$. We denote the input as $\mathbf{A}$ and the kernel as $\mathbf{K}$. At a given location (p, q) $(1 \leqslant p \leqslant P,\ 1 \leqslant q \leqslant Q)$, the output value $b_{p,q}$ is calculated as

$$b_{p,q} = \sum_{k_1=1}^{K_1} \sum_{k_2=1}^{K_2} k_{k_1,k_2} a_{S_1(p-1)+k_1, S_2(q-1)+k_2} \tag{2.39}$$

where $k_{i,j}$ refers the element in the i-th row and j-th column of the $\mathbf{K}$ and the same notation is applied

to **A**. All the output values $b_{p,q}$ $(1 \leqslant p \leqslant P,\ 1 \leqslant q \leqslant Q)$ constitute the output **B**. In practice, there can be multiple kernels for a 2D convolutional layer.

2.5 Conclusion

In this chapter, we provide the system models of the three promising enabling techniques at the physical layer for the future wireless networks in this thesis which are SCMA, NC MIMO and SU multidimensional transmission. The following chapters will consider these system models. By reviewing the state-of-the-art for the three involved techniques, we point out the current technical challenges of the techniques. Among them, we highlight the significance of MUD for SCMA systems, Grassmannian constellation construction in NC MIMO systems and MUC design in SU multidimensional transmission which are our contributions of this thesis. Besides, as some DL techniques are applied in the thesis, they are introduced in this chapter.

Chapitre 3

Iterative detectors for SCMA systems

Contenu

3.1 Introduction

In this chapter, we focus on iterative detectors for SCMA transmission. Two different transmission scenarios are studied. In the first scenario where a simple coded SCMA transmission is considered, we propose a low-complexity iterative MUD algorithm for SCMA. In the second scenario, we consider a highly overload SCMA transmission in order to resolve the stringent spectrum resource challenge for the future wireless networks. We introduce an uplink two-groups SCMA transmission scheme and apply iterative MUD algorithm and successive interference cancellation (SIC) jointly at the receiver. To guarantee good BER performance and reach high spectral efficiency, we propose an optimization method for channel code rates.

3.2 Simple SCMA transmission scenario

3.2.1 Coded SCMA system model

A coded SCMA transmission system is considered, which is illustrated in Fig. 3.1. The binary bit sequence $\mathbf{b}_j$ of user j is coded to $\mathbf{c}_j$ by a channel encoder and then interleaved into sequence $\mathbf{u}_j$. The interleaver is denoted as π in the figure. Afterwords, every $\log_2(M)$ bits u_j^i $(1 \leqslant i \leqslant \log_2(M))$ are mapped into a complex codeword $\mathbf{x}_j$ according to the codebook $\mathcal{X}_j$ by an SCMA mapper. If the codeword of user j is determined by a generator vector $\mathbf{v}_j = [v_{1,j}, v_{2,j}, \ldots, v_{K,j}]^T$ [36], it can be expressed as

$$\mathbf{x}_j = \mathbf{v}_j \cdot s_j \tag{3.1}$$

where s_j is a complex symbol from the selected constellation named set $\mathcal{S}$ $(|\mathcal{S}| = M)$. Therefore, the received signal in Eq.(2.2) can be rewritten as

$$\mathbf{y} = \sum_{j=1}^{J} diag(\mathbf{h}_j)\mathbf{v}_j \cdot s_j + \mathbf{n} \tag{3.2}$$

$$= \sum_{j=1}^{J} \widetilde{\mathbf{h}}_j \cdot s_j + \mathbf{n} \tag{3.3}$$

where $\widetilde{\mathbf{h}}_j = \mathbf{h}_j \odot \mathbf{v}_j = [h_{1,j}v_{1,j}, \ldots, h_{K,j}v_{K,j}]$.

At the receiver, the iterative MUD scheme is composed of an SCMA detector and J parallel channel decoders along with interleavers and deinterleavers [95, 96], as shown in Fig. 3.1. T_{ex} is used to denote

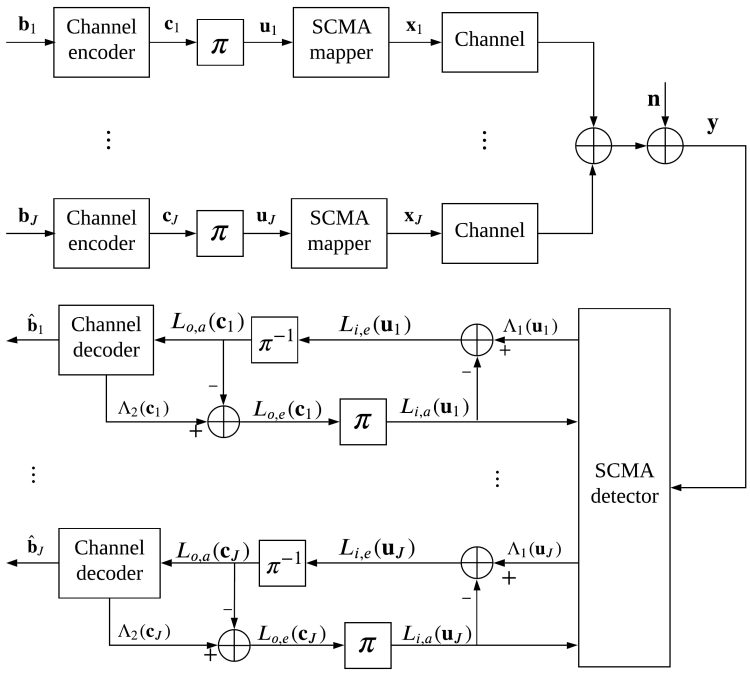


FIGURE 3.1 – Transmission structure of the coded SCMA system.

the number of outer loops, the iterative loops between the SCMA detector and the channel decoders. On the contrary, T is used to denote the number of inner loops, the number of iterations of an SCMA iterative MUD algorithm. At the t_{ex}-th ($1 \leqslant t_{ex} \leqslant T_{ex}$) outer loop, the output of the SCMA detector is the soft decision, namely *a posteriori* LLRs. The *a posteriori* LLR of every bit u_j^i delivered by SCMA detector is given by

$$\Lambda_1(u_j^i) = \ln \frac{P(u_j^i = 1|\mathbf{y})}{P(u_j^i = 0|\mathbf{y})} \tag{3.4}$$

According to the Baye's rule, the extrinsic LLR $L_{i,e}(u_j^i)$ of the SCMA detector is given as

$$L_{i,e}(u_j^i) = \Lambda_1(u_j^i) - L_{i,a}(u_j^i) \tag{3.5}$$

where $L_{i,a}(u_j^i)$ is the *a priori* LLR of the corresponding channel decoder in the previous outer iteration. The *a priori* information of s_j can be calculated by the associated *a priori* LLR $L_{i,a}(u_j^i)$, formulated as

$$\mu_{\phi_j \to v_j}(s_j) = \prod_{i=1}^{\log_2(M)} \frac{\exp\left(u_j^i L_{i,a}(u_j^i)\right)}{1 + \exp\left(L_{i,a}(u_j^i)\right)} \tag{3.6}$$

For the first outer iteration, we assume $L_{i,a}(u_j^i) = 0$. Similarly, the *a posteriori* LLR delivered by channel decoder can be written as

$$\Lambda_2(c_j^i) = L_{o,a}(c_j^i) + L_{o,e}(c_j^i) \tag{3.7}$$

where $L_{o,a}(c_j^i)$ represents the *a priori* LLR from the SCMA detector and $L_{o,e}(c_j^i)$ represents the extrinsic information of user j's channel decoder.

3.2.2 State-of-the-art iterative SCMA MUD algorithms

In this subsection, we revisit some of the low-complexity algorithms, GAA [97] and EPA [98], which tackle the complexity issue of MPA in the literature.

3.2.2.1 GAA for the SCMA system

Considering the representation of the users' codewords in Eq.(3.1), the discrete probabilities (message) from v_j to f_k, $\mu_{v_j \to f_k}(\mathbf{x}_j)$ shown in Eq.(2.12), can be approximated by a continuous Gaussian distribution given as

$$\hat{\mu}^t_{v_j \to f_k}(s_j) = \mathcal{N}_c(s_j; \hat{x}^t_{v_j \to f_k}, \hat{\tau}^t_{v_j \to f_k}) \tag{3.8}$$

where s_j follow the rule in Eq.(3.1), in order to decrease the computational complexity. By leveraging the minimization of the Kullback-Leibler divergence [70] of the two distribution, $D(\mu_{v_j \to f_k}(s_j) \| \hat{\mu}_{v_j \to f_k}(s_j))$, the mean and variance of $\hat{\mu}^t_{v_j \to f_k}(s_j)$ are calculated by

$$\hat{x}^t_{v_j \to f_k} = \mathbb{E}_{\mu^t_{v_j \to f_k}}[s_j] \quad s_j \in \mathcal{S} \tag{3.9}$$

and

$$\hat{\tau}^t_{v_j \to f_k} = \mathbb{E}_{\mu^t_{v_j \to f_k}}[|s_j|^2] - |\hat{x}^t_{v_j \to f_k}|^2 \quad s_j \in \mathcal{S} \tag{3.10}$$

Taking the distribution $\mathcal{N}_c(s_j; \hat{x}^t_{v_j \to f_k}, \hat{\tau}^t_{v_j \to f_k})$ in the calculation of the message from f_k to v_j, $\mu^t_{f_k \to v_j}(s_j)$ is approximated as

$$\begin{aligned} \mu^t_{f_k \to v_j}(s_j) &= \sum_{s'_j \in \mathcal{S} \setminus s_j} f_k(y_k | s'_j) \prod_{q \in F(k) \setminus j} \mathcal{N}_c(s_q; \hat{x}^t_{v_q \to f_k}, \hat{\tau}^t_{v_q \to f_k}) \\ &\approx \mathcal{N}_c(\widetilde{h}_{k,j} s_j; z^t_{f_k \to v_j}, \nu^t_{f_k \to v_j}) \end{aligned} \tag{3.11}$$

According to the expression of the received signal, Eq.(3.3), $z^t_{f_k \to v_j}$ and $\nu^t_{f_k \to v_j}$ are computed by

$$z^t_{f_k \to v_j} = y_k - \sum_{q \in F(k) \setminus j} \widetilde{h}_{k,q} \hat{x}^t_{v_q \to f_k} \tag{3.12}$$

and

$$\nu^t_{f_k \to v_j} = \sigma^2 + \sum_{q \in F(k) \setminus j} |\widetilde{h}_{k,q}|^2 \hat{\tau}^t_{v_q \to f_k} \tag{3.13}$$

Recall the principles of joint Gaussian distribution [99], the message from v_j to f_k can be rewritten as

$$\mu^t_{v_j\to f_k}(s_j) = \frac{\mu_{\phi_j\to v_j}(s_j)\prod\limits_{l\in V(j)\backslash k}\mu^{t-1}_{f_l\to v_j}(s_j)}{\sum\limits_{s_j\in\mathcal{S}}\mu_{\phi_j\to v_j}(s_j)\prod\limits_{l\in V(j)\backslash k}\mu^{t-1}_{f_l\to v_j}(s_j)} \tag{3.14}$$

$$= \frac{\mu_{\phi_j\to v_j}(s_j)\mathcal{N}_c(s_j;\zeta^{t-1}_{v_j\to f_k},\gamma^{t-1}_{v_j\to f_k})}{\sum\limits_{s_j\in\mathcal{S}}\mu_{\phi_j\to v_j}(s_j)\mathcal{N}_c(s_j;\zeta^{t-1}_{v_j\to f_k},\gamma^{t-1}_{v_j\to f_k})} \tag{3.15}$$

where

$$\gamma^{t-1}_{v_j\to f_k} = \left(\sum_{l\in V(j)\backslash k}\frac{|\widetilde{h}_{l,j}|^2}{\nu^{t-1}_{f_l\to v_j}}\right)^{-1} \tag{3.16}$$

and

$$\zeta^{t-1}_{v_j\to f_k} = \gamma^{t-1}_{v_j\to f_k}\sum_{l\in V(j)\backslash k}\frac{\widetilde{h}^*_{l,j}z^{t-1}_{f_l\to v_j}}{\nu^{t-1}_{f_l\to v_j}} \tag{3.17}$$

where $\widetilde{h}^*_{l,j}$ denotes the complex conjugate of the complex value $\widetilde{h}_{l,j}$. The values of $\mu^t_{v_j\to f_k}$ are then used to calculated the mean $\hat{x}^t_{v_j\to f_k}$ and variance $\hat{\tau}^t_{v_j\to f_k}$ for the current iteration. Similarly, after T iterations, the *a posteriori* LLR can be computed by

$$\Lambda^p_1(c^i_j) = \ln\frac{\sum\limits_{s_j\in\mathcal{S}^+}\mu_{\phi_j\to v_j}(s_j)\mathcal{N}_c(s_j;\zeta^T_{v_j},\gamma^T_{v_j})}{\sum\limits_{s_j\in\mathcal{S}^-}\mu_{\phi_j\to v_j}(s_j)\mathcal{N}_c(s_j;\zeta^T_{v_j},\gamma^T_{v_j})} \tag{3.18}$$

where the mean $\gamma^T_{v_j}$ and variance $\zeta^T_{v_j}$ are defined by

$$\gamma^T_{v_j} = \left(\sum_{k\in V(j)}\frac{|\widetilde{h}_{k,j}|^2}{\nu^T_{f_k\to v_j}}\right)^{-1} \tag{3.19}$$

and

$$\zeta^T_{v_j} = \gamma^T_{v_j}\sum_{k\in V(j)}\frac{\widetilde{h}^*_{k,j}z^T_{f_k\to v_j}}{\nu^T_{f_k\to v_j}} \tag{3.20}$$

The details of GAA is concluded in Algorithm 5

3.2.2.2 EPA for the SCMA system

As described in [100], expectation propagation approximates the observation term based on some term approximations and then uses an exact posterior with the term approximations until all the term

Algorithm 5 GAA.

Initialization : $\zeta^0_{v_j \to f_k} = 0$, $\gamma^0_{v_j \to f_k} = 1000$, $t = 1$.
while $t \leqslant T$ **do**
  for $j = 1 \to J$ **do**
    for $k \in V(j)$ **do**
      Calculate message $\mu^t_{v_j \to f_k}(s_j)$ using Eq.(3.15).
      Calculate $\hat{x}^t_{v_j \to f_k}$ and $\hat{\tau}^t_{v_j \to f_k}$ using Eq.(3.9) and Eq.(3.10), respectively.
    end for
  end for
  for $k = 1 \to K$ **do**
    for $j \in F(k)$ **do**
      Calculate $z^t_{f_k \to v_j}$ and $\nu^t_{f_k \to v_j}$ using Eq.(3.12) and Eq.(3.13), respectively.
    end for
  end for
  for $j = 1 \to J$ **do**
    Calculate $\gamma^t_{v_j \to f_k}$ and $\zeta^t_{v_j \to f_k}$ using Eq.(3.16) and Eq.(3.17), respectively.
  end for
  $t = t + 1$.
end while
for $j = 1 \to V$ **do**
  Calculate $\gamma^T_{v_j}$ and $\zeta^T_{v_j}$ using Eq.(3.19) and Eq.(3.20), respectively.
  Calculate LLR using Eq.(3.18).
end for

approximations converge. In a nutshell, expectation propagation can approximate a distribution with another distributions. Since the approximate posterior and the term approximations are in the same family of distribution, they are constrained to lie in a complex Gaussian distribution family in SCMA context.

In the MUD of SCMA, we define the symbol belief of s_j as [98]

$$\beta^t_{v_j}(s_j) = \frac{\mu_{\phi_j \to v_j}(s_j) \prod\limits_{k \in V(j)} \mu^{t-1}_{f_k \to v_j}(s_j)}{\sum\limits_{s_j \in \mathcal{S}} \mu_{\phi_j \to v_j}(s_j) \prod\limits_{k \in V(j)} \mu^{t-1}_{f_k \to v_j}(s_j)} \tag{3.21}$$

which is applied to approximate the complex Gaussian distribution $\mathcal{N}_c(s_j; \hat{x}^t_{v_j \to f_k}, \hat{\tau}^t_{v_j \to f_k})$. The symbol belief $\beta^t_{v_j}(s_j)$ is approximated by $\hat{\beta}^t_{v_j}(s_j) = \mathcal{N}_c(s_j; \hat{x}^t_{v_j}, \hat{\tau}^t_{v_j})$ using the approaches in [70] and [101]. Consequently, $\hat{x}^t_{v_j}$ and $\hat{\tau}^t_{v_j}$ are calculated by

$$\hat{x}^t_{v_j} = \mathbb{E}_{\beta^t_{v_j}}[s_j] \qquad\qquad s_j \in \mathcal{S} \tag{3.22}$$

and

$$\hat{\tau}_{v_j}^t = \mathbb{E}_{\beta_{v_j}^t}[|s_j|^2] - |\hat{x}_{v_j}^t|^2 \qquad s_j \in \mathcal{S} \tag{3.23}$$

By applying the principles of expectation propagation [100], the message $\mu_{v_j \to f_k}^t(s_j)$ can be rewritten as

$$\begin{aligned} \mu_{v_j \to f_k}^t(s_j) &\propto \frac{\beta_{v_j}^t(s_j)}{\mu_{f_k \to v_j}^{t-1}(s_j)} \\ &\approx \frac{\hat{\beta}_{v_j}^t(s_j)}{\mu_{f_k \to v_j}^{t-1}(s_j)} \\ &= \mathcal{N}_c(s_j; \hat{x}_{v_j \to f_k}^t, \hat{\tau}_{v_j \to f_k}^t) \end{aligned} \tag{3.24}$$

Following the rules of Bayesian networks [99], $\hat{x}_{v_j \to f_k}^t$ and $\hat{\tau}_{v_j \to f_k}^t$ are computed by

$$\hat{\tau}_{v_j \to f_k}^t = \left(\frac{1}{\hat{\tau}_{v_j}^t} - \frac{|\widetilde{h}_{k,j}|^2}{\nu_{f_k \to v_j}^{t-1}} \right)^{-1} \tag{3.25}$$

and

$$\hat{x}_{v_j \to f_k}^t = \hat{\tau}_{v_j \to f_k}^t \left(\frac{\hat{x}_{v_j}^t}{\hat{\tau}_{v_j}^t} - \frac{\widetilde{h}_{k,j}^* z_{f_k \to v_j}^{t-1}}{\nu_{f_k \to v_j}^{t-1}} \right) \tag{3.26}$$

It is worth noticing that the parameter update in Eq.(3.25) might return a negative value of $\hat{\tau}_{v_j \to f_k}^t$, although it should be positive. In this case, we just keep the previous values for variance $\hat{\tau}_{v_j \to f_k}^t$ and mean $\hat{x}_{v_j \to f_k}^t$.

For mean and variance calculation of the message $\mu_{f_k \to v_j}^t(s_j)$ in EPA, they are the same as in the GAA, as expressed in Eq.(3.12)-(3.13). The *a posteriori* LLR is calculated using Eq.(2.16), where $\mu_{f_k \to v_j}^T(s_j)$ is computed according to Eq.(3.11).

The detection procedures of the EPA is summarized in Algorithm 6.

3.2.3 Proposed SCMA detection algorithm

We introduce the proposed first-order approximated Gaussian approximation algorithm (FO-GAA) in this section.

To further reduce the computational complexity, we propose to neglect the term $\frac{|\widetilde{h}_{k,j}|^2}{\nu_{f_k \to v_j}^{t-1}}$ in Eq.(3.25), expressed as

$$\hat{\tau}_{v_j \to f_k}^t \approx \hat{\tau}_{v_j}^t \tag{3.27}$$

Algorithm 6 EPA.

Initialization : $z^0_{f_k \to v_j} = 0$, $\nu^0_{f_k \to v_j} = 1000$, $t = 1$.
while $t \leqslant T$ **do**
  for $j = 1 \to V$ **do**
    for $k \in V(j)$ **do**
      Calculate symbol belief $\beta^t_{v_j}(s_j)$ using Eq.(3.21).
      Calculate $\hat{x}^t_{v_j}$ and $\hat{\tau}^t_{v_j}$ using Eq.(3.22) and Eq.(3.23), respectively.
    end for
  end for
  for $j = 1 \to V$ **do**
    for $k \in V(j)$ **do**
      if $\frac{1}{\hat{\tau}^t_{v_j}} > \frac{|\widetilde{h}_{k,j}|^2}{\nu^{t-1}_{f_k \to v_j}}$ **then**
        Calculate $\hat{\tau}^t_{v_j \to f_k}$ and $\hat{x}^t_{v_j \to f_k}$ using Eq.(3.25) and Eq.(3.26), respectively.
      end if
    end for
  end for
  for $k = 1 \to F$ **do**
    for $j \in F(k)$ **do**
      Calculate $z^t_{f_k \to v_j}$ and $\nu^t_{f_k \to v_j}$ using Eq.(3.12) and Eq.(3.13), respectively.
    end for
  end for
  $t = t + 1$.
end while
for $j = 1 \to V$ **do**
  Calculate $\mu^T_{f_k \to v_j}$ using Eq.(3.11).
  Calculate LLR using Eq.(2.16).
end for

Hence, we define $\nu^t_{f_k}$ using $\hat{\tau}^t_{v_j}$ as follows

$$\nu^t_{f_k} = \sigma_n^2 + \sum_{j \in F(k)} |\widetilde{h}_{k,j}|^2 \cdot \hat{\tau}^t_{v_j} \tag{3.28}$$

Fig. 3.2 illustrates the probability density function (PDF) of $\nu^t_{f_k \to v_j}$ and $\nu^t_{f_k}$, showing that it $\nu^t_{f_k \to v_j}$ can be substituted by $\nu^t_{f_k}$. For the same reason, $\gamma^t_{v_j}$ can be the approximation of $\gamma^t_{v_j \to f_k}$, which is defined as

$$\gamma^t_{v_j} = \left(\sum_{k \in V(j)} \frac{|\widetilde{h}_{k,j}|^2}{\nu^t_{f_k}} \right)^{-1} \tag{3.29}$$

Therefore, the mean of the messages from v_j to f_k, Eq.(3.26), can be rewritten as

$$\hat{x}^t_{v_j \to f_k} \approx \hat{x}^t_{v_j} - \frac{\hat{\tau}^t_{v_j} \widetilde{h}^*_{k,j} z^{t-1}_{f_k \to v_j}}{\nu^{t-1}_{f_k}} \tag{3.30}$$

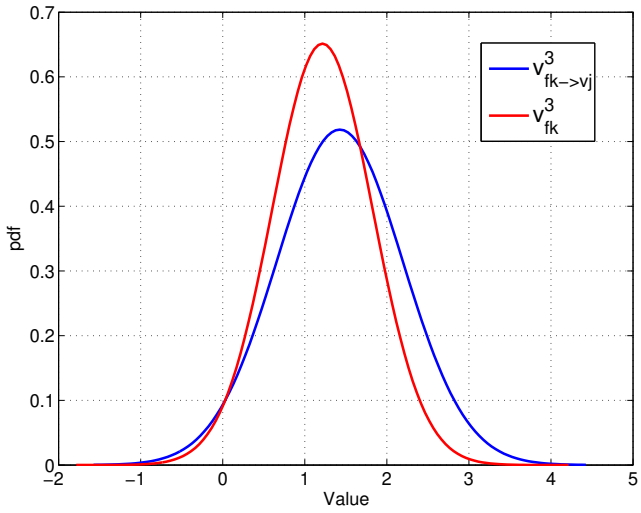


FIGURE 3.2 – PDF comparison for $\nu^t_{f_k \to v_j}$ and $\nu^t_{f_k}$ with $J = 6$, $K = 4$, $M = 4$, $d_v = 2$, $d_f = 3$ and QPSK constellation.

Similarly, $z^t_{f_k}$ and $\zeta^t_{v_j}$ substitute for $z^t_{f_k \to v_j}$ and $\zeta^t_{v_j \to f_k}$ respectively. Following the approximation rules in [102], $z^t_{f_k \to v_j}$ and $\zeta^t_{v_j \to f_k}$ are computed as

$$z^t_{f_k} \approx y_k - \sum_{j \in F(k)} \widetilde{h}_{k,j} \hat{x}^t_{v_j} + z^{t-1}_{f_k} \frac{\sum_{j \in F(k)} \hat{\tau}^t_{v_j} |\widetilde{h}_{k,j}|^2}{\nu^{t-1}_{f_k}} \tag{3.31}$$

and

$$\zeta^t_{v_j} \approx \hat{x}^t_{v_j} + \gamma^t_{v_j} \sum_{k \in V(j)} \frac{\widetilde{h}^*_{k,j} z^t_{f_k}}{\nu^t_{f_k}} \tag{3.32}$$

Consequently, the message $\mu^t_{v_j}(s_j)$ replaces the $\mu^t_{v_j \to f_k}(s_j)$ in MPA, defined as

$$\mu^t_{v_j}(s_j) = \frac{\mu_{\phi_j \to v_j}(s_j) \mathcal{N}_c(s_j; \zeta^{t-1}_{v_j}, \gamma^{t-1}_{v_j})}{\sum_{s_j \in \mathcal{S}} \mu_{\phi_j \to v_j}(s_j) \mathcal{N}_c(s_j; \zeta^{t-1}_{v_j}, \gamma^{t-1}_{v_j})} \tag{3.33}$$

The mean of variance of $\mu^t_{v_j}(s_j)$ are computed as

$$\hat{x}^t_{v_j} = \mathbb{E}_{\mu^t_{v_j}}[s_j] \tag{3.34}$$

and

$$\hat{\tau}^t_{v_j} = \mathbb{E}_{\mu^t_{v_j}}[|s_j|^2] - |\hat{x}^t_{v_j}|^2 \tag{3.35}$$

Algorithm 7 FO-GAA.

Initialization : $\zeta_{v_j}^0 = 0$, $\nu_{f_k}^0 = 1000$, $z_{f_k}^0 = 0$, $\gamma_{v_j}^0 = 1000$, $t = 1$.
while $t \leqslant T$ **do**
  for $j = 1 \rightarrow V$ **do**
    Calculate message $\mu_{v_j}^t(s_j)$ using Eq.(3.33).
    Calculate $\hat{x}_{v_j}^t$ and $\hat{\tau}_{v_j}^t$ using Eq.(3.34) and Eq.(3.35), respectively.
  end for
  for $k = 1 \rightarrow F$ **do**
    Calculate $\nu_{f_k}^t$ and $z_{f_k}^t$ using Eq.(3.28) and Eq.(3.31), respectively.
  end for
  for $j = 1 \rightarrow V$ **do**
    Calculate $\gamma_{v_j}^t$ and $\zeta_{v_j}^t$ using Eq.(3.29) and Eq.(3.32), respectively.
  end for
  $t = t + 1$.
end while
for $j = 1 \rightarrow V$ **do**
  Calculate LLR using Eq.(3.18).
end for

respectively. The *a posteriori* LLR calculation of the proposed FO-GAA is the same of GAA, which is Eq.(3.18).

The proposed FO-GAA is summarized in Algorithm 7.

3.2.4 Performance analysis

The proposed algorithm and the state-of-the-art SCMA MUD algorithms, including MPA, GPA and EPA, are investigated in terms of the convergence behavior and BER performance in the simulation. The codebook used in this simulation is the one proposed in [36] with $J = 6$, $K = 4$, $M = 4$, $d_v = 2$, $d_f = 3$ and QPSK constellation. The channel model is long-term evolution (LTE) pedestrian A with 3km/h moving speed. Convolutional code with code rate $R = 1/2$ and S random interleaver are applied in iterative multiuser receiver. Additionally, each frame has 8000 bits.

Fig. 3.3 compares the convergence behavior of the four considered algorithms for SNR = 0dB and $T = 3$. Furthermore, Table 3.1 shows the number of outer loop iterations required for different algorithms' convergence. EPA and FO-GAA convergence behaviors resemble that of MPA, while GAA converges more slowly. This indicates that GAA might not get convergence when the other algorithms have converged.

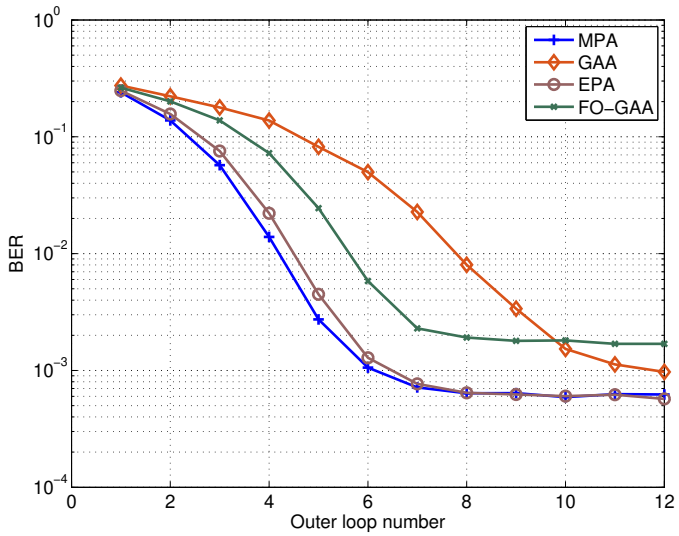


FIGURE 3.3 – Convergence behavior comparison for different detection algorithms when SNR = 0dB, $T = 3$.

TABLE 3.1 – Number of outer loop iterations to reach convergence.

SNR [dB]	0	0.5	1	1.5	2
T_{ex} for MPA, EPA and FO-GAA	7	5	5	5	5
T_{ex} for GAA	11	6	5	5	5

Fig. 3.4 illustrates the BER performance of the four algorithms when they converge. In other words, the BER of MPA, GAA, EPA and FO-GAA are obtained by simulating with the T_{ex} required for convergence given in Table 3.1. Furthermore, in order to show the influence of T_{ex} on GAA, we also present its BER without reaching the convergence. The label "GAA nc" in the figure refers to the BER performance of GAA simulated with the T_{ex} required for the convergence for all the other algorithms in Table 3.1. The curves are not smooth due to the T_{ex} configuration depending on the SNR values. EPA outperforms sightly GAA when they both converge. Besides, FO-GAA has a $0.2dB$ performance loss compared to MPA at SNR = 2dB.

Next, we study the computational complexity of the studied algorithms. The computation complexity of the four algorithms is evaluated in terms of floating-point operations (FLOPs). Due to the structure of the receiver shown in Fig. 3.1, the complexity incorporates the iteration part and the *a posteriori* LLR part, as listed in Table 3.2. This means that the overall complexity of an algorithm depends on not only the number of inner loops T but also the number of outer loops T_{ex}.

It is assumed that the operation of $\exp(\cdot)$ and $\ln(\cdot)$ can be implemented by a look-up table, needing

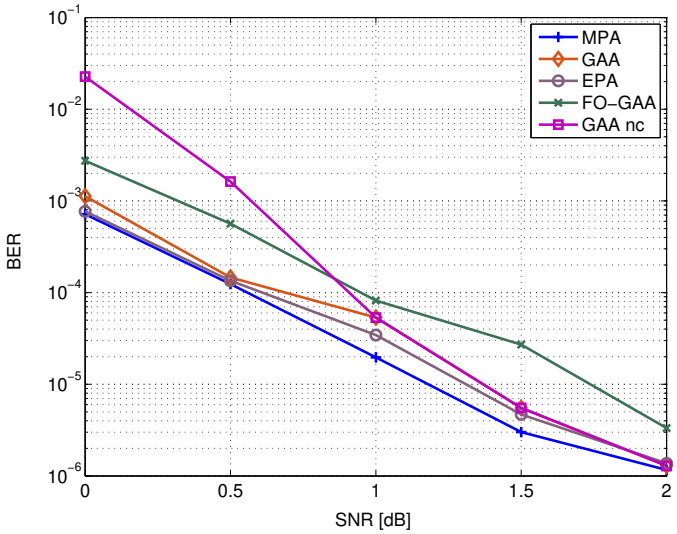


FIGURE 3.4 – BER performance comparison.

TABLE 3.2 – Complexity of Different Detection algorithms.

Algorithm	Iteration part	LLR part
MPA	$T\big((9d_f+7)M^{d_f}Kd_f + MJ{d_v}^2 - Kd_f + 5(\log_2(M)-1)MJ\big)$	$(d_v+1)MJ\log_2(M)$
GAA	$T\big((21M+15d_v+13d_f-28)Kd_f + 5(\log_2(M)-1)MJ\big)$	$\big(15d_v+11M\log_2(M)\big)J$
EPA	$T\big((16M+13d_f+8)Kd_f + (5\log_2(M)+6)MJ\big)$	$(16d_v+1)MJ\log_2(M)$
FO-GAA	$T\big((33d_f+4)K + 2J + (5\log_2(M)+16)MJ\big)$	$11MJ\log_2(M)$

one FLOP. Note that the addition of two complex numbers requires two FLOPs, the multiplication of a real number and a complex number needs two FLOPs. Besides, the multiplication of two conjugated and two non-conjugated complex numbers require three and six FLOPs, respectively.

We hereafter provide the details of the number of FLOPs of the proposed FO-GAA. The computational complexity of the other three algorithms is not detailed due to a lack of space. First of all, the *a priori* information $\{\mu_{\phi_j\to v_j}(s_j), \forall j\}$ is calculated with $5(\log_2(M)-1)MJ$ FLOPs. Secondly, $(12M-1)J$ FLOPs are required to compute the message $\{\mu^t_{v_j}(s_j), \forall j\}$. Once it is done, $\{\hat{x}^t_{v_j}, \hat{\tau}^t_{v_j}, \forall j\}$ can be calculated with $(9M+1)J$ FLOPs. After, $(18d_f+4)K$ FLOPs are required to calculate $\{z^t_{f_k}, \nu^t_{f_k}, \forall k\}$. Then, to calculate $\{\gamma^t_{v_j}, \zeta^t_{v_j}, \forall j\}$, $(15d_v+2)J$ FLOPs are required. Considering the relationship of $Kd_f = Jd_v$ in the factor graph, Jd_v is substituted by Kd_f in the expression. Therefore, at the t-th inner loop iteration of the FO-GAA, $(33d_f+4)K + 2J + (5\log_2(M)+16)MJ$ FLOPs are required. Finally, the complexity of calculating *a posteriori* LLR $\{\Lambda^p_1(c^i_j), \forall j, \forall i\}$ is $11MJ\log_2(M)$ FLOPs.

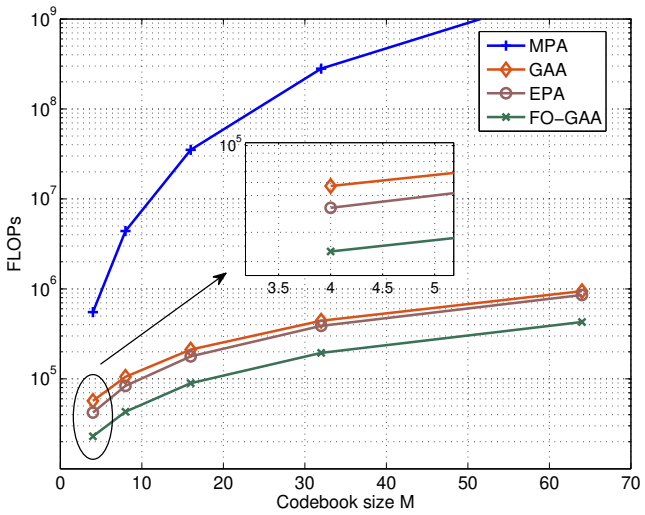


FIGURE 3.5 – Number of FLOPs comparison when SNR = 0dB.

Fig. 3.5 compares the number of FLOPs of the four algorithms with $J = 6$, $K = 4$, $d_v = 2$, $d_f = 3$ and variant codebook sizes. The figure reveals the number of FLOPs of different algorithms with BER convergence at SNR = 0dB. Because the GAA requires a larger T_{ex} for the convergence, its computational complexity is higher than that of EPA and FO-GAA. The number of FLOPs of the proposed FO-GAA is the lowest and falls to 54.6% of that of the EPA and 4.2% of that of the MPA for the codebook size $M = 4$. The low computational complexity and good convergence behavior of the proposed FO-GAA compensate for its BER performance loss.

3.3 Two-groups SCMA transmission scenario

3.3.1 Motivation of the two groups transmission

Considering the limited amount of bandwidth and the demand of accommodating increasing number of devices in IoE, to increase connectivity and spectral efficiency are essential for the 6G. As PD-NOMA and SCMA superpose multiuser signals in the power and code domain respectively, the idea of combining PD-NOMA and SCMA to further increase the connectivity has recently emerged. Authors in [103] propose a downlink hybrid multiple access system, in which two groups of users, one using SCMA mapping and the other using binary phase-shift keying modulation, are superposed by applying PD-NOMA techniques. By adopting power allocation at the base station (BS) and SIC along with MPA detection at the receiver, the system in [103] has better spectral efficiency and higher

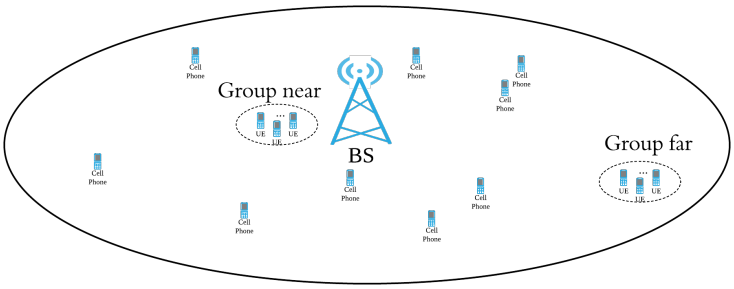


FIGURE 3.6 – Illustration of the proposed system.

connectivity than the conventional SCMA system. Paper [104] extends the modulation of both two groups to SCMA codebook in a downlink multiple access system and obtains larger connectivity than [103]. In [105], the authors propose a new multiple access technique called power domain sparse code multiple access for downlink heterogeneous cellular network. The work in [105] concentrates on the optimization of the downlink resource allocation for different users with some system level and quality of service constraints.

Different from the above mentioned work contributes to downlink systems, in the chapter we focus on the uplink two-groups transmission scenario where PD-NOMA and SCMA are jointly employed. Hence, no power allocation is considered in the proposed transmission scheme. In the system, two groups of SCMA users with different average transmission channel power share the same orthogonal resources in order to improve the system spectral efficiency. In the following, we will introduce the proposed system and optimization method for good coded BER performance and spectral efficiency.

3.3.2 The proposed system

We consider a single-cell uplink system, in which a BS equipped with one single antenna is located at the center, as illustrated in Fig. 3.6. A large number of users having the same transmit power are located randomly in the cell. Assume that there are two groups each containing J users located exceedingly close to each other in the cell. All users in the same group are supposed to have the same average channel gain over time. Thus, the SNR of groups represent the SNR of all the users in the corresponding groups. For the sake of simplicity, the group with users located nearer to the BS is denoted as *Group near* and the other one is denoted as *Group far*. Therefore, the received power at the BS from *Group near* is greater than that from *Group far*.

The proposed system superposes the signals of two groups, allowing all $2J$ users to share K

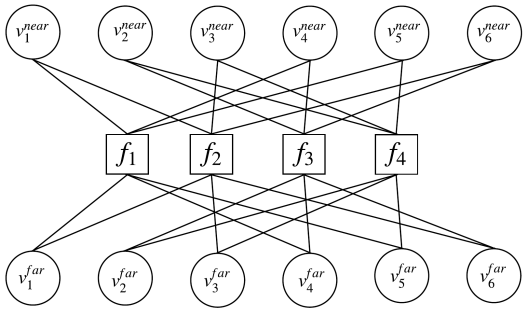


FIGURE 3.7 – Factor graph of the proposed system.

REs. The overlapping of the users' signals is illustrated by the factor graph as shown in Fig. 3.7. As illustrated in the figure, function node degree d_f in the proposed system is doubled compared to that of a conventional SCMA system [18], while the variable node degree d_v remains the same. Therefore, the overloading factor of this system is $\lambda = \frac{2J}{K}$.

Fig. 3.8 shows the transmitter structure of the proposed system, where channel coding is involved. $\mathbf{c}_j^{near(far)}$ denotes the coded bit sequence and $\mathbf{u}_j^{near(far)}$ denotes the interleaved bit sequence. Two groups share the same multidimensional codebook $\mathcal{C} = \mathcal{C}_1 \times \mathcal{C}_2 \cdots \times \mathcal{C}_J$ ($|\mathcal{C}_j| = M$). The SCMA codeword of user j_1 in *Group near* is $\mathbf{x}_{j_1}^{near} = [x_{1,j_1}, x_{2,j_1}, \ldots, x_{K,j_1}]^\intercal$, $\mathbf{x}_{j_1}^{near} \in \mathcal{C}_{j_1}$ and that of user j_2 in *Group far* is $\mathbf{x}_{j_2}^{far} = [x_{1,j_2}, x_{2,j_2}, \ldots, x_{K,j_2}]^\intercal$, $\mathbf{x}_{j_2}^{far} \in \mathcal{C}_{j_2}$. The power of codeword is subject to $\mathbb{E}_j[\frac{\sum_{\mathcal{C}_j} |\mathbf{x}_j|^2}{dv*M}] = 1$. The CSI between users and the BS is perfectly estimated by the BS. The small-scale fading between user j_1 and user j_2 to the BS are $\mathbf{h}_{j_1}^{near} = [h_{1,j_1}^{near}, h_{2,j_1}^{near}, \ldots, h_{K,j_1}^{near}]^\intercal$ and $\mathbf{h}_{j_2}^{far} = [h_{1,j_2}^{far}, h_{2,j_2}^{far}, \ldots, h_{K,j_2}^{far}]^\intercal$, respectively. Therefore, the received signal at the BS is expressed as

$$\mathbf{y} = \sum_{j_1=1}^{J} diag(\mathbf{h}_{j_1}^{near})\sqrt{P_n}\mathbf{x}_{j_1}^{near} + \sum_{j_2=1}^{J} diag(\mathbf{h}_{j_2}^{far})\sqrt{P_f}\mathbf{x}_{j_2}^{far} + \mathbf{n} \tag{3.36}$$

$$= \mathbf{y}^{near} + \mathbf{y}^{far} + \mathbf{n} \tag{3.37}$$

where P_n and P_f are the received power at the BS from users in *Group near* and *Group far*, respectively. They take into account transmit power and large-scale fading. $\mathbf{n} = [n_1, n_2, \ldots, n_K]^\intercal$, having entries $n_k \sim \mathcal{CN}(0, \sigma^2)$, is the AWGN vector.

The SNR of *Group near* and *Group far* are defined as

$$\text{SNR}^{near} = \frac{\mathbb{E}[|h_{k,j_1}^{near}|^2]P_n}{\sigma^2},\ \forall (k, j_1) \in \{1, .., K\} \times \{1, .., J\} \tag{3.38}$$

and

$$\text{SNR}^{far} = \frac{\mathbb{E}[|h_{k,j_2}^{far}|^2]P_f}{\sigma^2},\ \forall (k, j_2) \in \{1, .., K\} \times \{1, .., J\} \tag{3.39}$$

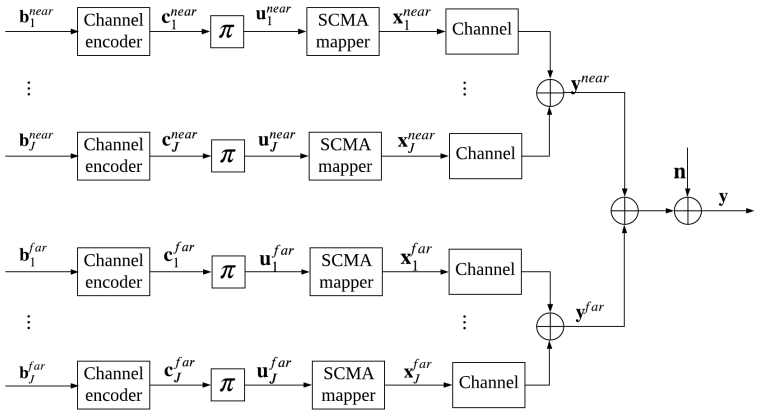


FIGURE 3.8 – Transmitter structure of the proposed system.

respectively.

As both groups use the same REs, the received signal of each user at the BS suffers from two kinds of interference : i) intra-group interference ; and ii) inter-group interference. The intra-group interference is accounted for the SCMA multiplexing on the same subcarriers, coming from the users in the same group. The inter-group interference, however, is caused by the superposition of the transmission signals of two groups, coming from the users in the other group. The receiver structure of the proposed system is depicted in Fig. 3.9. The SIC and iterative SCMA detection structure are combined at the receiver to eliminate the two types of interference and facilitate the MUD. The SCMA MUD algorithm removes the intra-group interference and the SIC performs inter-group interference cancellation. As shown in Fig. 3.9, there are two types of loops within the receiver which are the inner loops, the loops inside the SCMA detection algorithm, and the outer loops, the loops between the SCMA detector and the channel decoder. According to the rules of the uplink SIC, the BS decodes the received signal $\mathbf{y}$ first for *Group near*, by considering the signal of *Group far* as noise and adopting SCMA detection. Once the decoding of *Group near* has finished, the BS recovers the estimated signal $\hat{\mathbf{y}}^{near}$ and subtracts it from the signal $\mathbf{y}$. Finally, the signals of users in *Group far* is decoded based on the signal $\mathbf{y} - \hat{\mathbf{y}}^{near}$ by the SCMA detector.

3.3.3 System performance optimization

As expatiated in [18], performance of the proposed system depends on the SCMA detection algorithm, the channel code rate as well as the numbers of inner and outer loop iterations. Specifically, choosing either a channel code with lower code rate or a larger number of iterations improves the

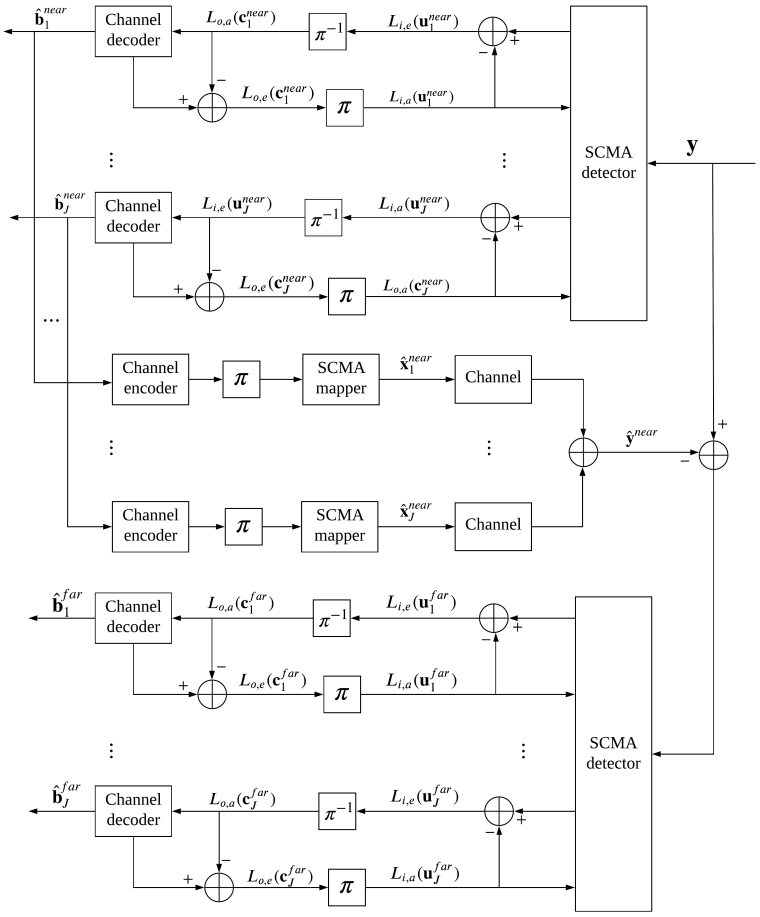


FIGURE 3.9 – Receiver structure of the proposed system.

convergence behavior of the proposed system, but inevitably decreases spectral efficiency. Therefore, in the following, we introduce a method to optimize the system performance considering this essential trade-off between the system convergence behavior and spectral efficiency.

3.3.3.1 Optimization method for choosing the channel code rates

Considering all the factors that mentioned above, we fix the detection algorithm as well as the number of inner loop iterations of the system, and propose a method to tune the other two parameters. The method aims at choosing the optimal channel code rates of the two groups individually in order to have higher system spectral efficiency while reaching the required convergence behavior.

Because the receiver of the two-groups transmission system can be regarded as an IDD structure as shown in Fig. 3.9, the EXIT charts can be used as a tool to analyze and design the proposed system [37]. To perform the EXIT charts analysis, we need to examine the evolution of the input/output mutual

information (MI) between the SCMA detector and the channel decoder. Following the deduction in [106], the MI $I_{i,e}^{near}$ between the interleaved bit sequence $\mathbf{u}_{j_1}^{near}$ and the extrinsic LLR of SCMA detector $L_{i,e}(\mathbf{u}_{j_1}^{near})$ is expressed as

$$I_{i,e}^{near} = \frac{1}{J}\sum_{j_1=1}^{J}\left(1 - \mathbb{E}_{u_{j_1}^{near}=+1}\big[\log_2\left(1+e^{-L_{i,e}(u_{j_1}^{near})}\right)\big]\right) \tag{3.40}$$

where $L_{i,e}(u_{j_1}^{near})$ is the corresponding output LLR of the SCMA detector with the input *a priori* LLR $L_{i,a}(u_{j_1}^{near})$. The input LLR is modeled as an independent Gaussian variable $n_A \sim (0, \sigma_A^2)$, which is defined as

$$L_{i,a}(u_{j_1}^{near}) = \frac{\sigma_A^2}{2}u_{j_1}^{near} + n_A = \frac{J^{-1}(I_{i,a}^{near})^2}{2}u_{j_1}^{near} + n_A \tag{3.41}$$

where $u_{j_1}^{near} \in \{-1, +1\}$ and $J^{-1}(\sigma)$ is the inverse of function

$$J(\sigma) = 1 - \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}\exp\big(-\frac{(\zeta-\frac{\sigma^2}{2})^2}{2\sigma^2}\big)\log_2(1+e^{-\zeta})d\zeta \tag{3.42}$$

Therefore, Eq. (3.40) can be reformulated as

$$I_{i,e}^{near} = T_i^{near}[I_{i,a}^{near}, SNR^{near}] \tag{3.43}$$

where T_i^{near} denotes the EXIT characteristics of the SCMA detector in the transmission environment with SNR^{near}. Accordingly, the MI $I_{o,e}^{near}$ between the coded bit sequence $\mathbf{c}_{j_1}^{near}$ and the channel decoder's extrinsic LLR $L_{o,e}(\mathbf{c}_{j_1}^{near})$ is formulated as

$$I_{o,e}^{near} = \frac{1}{J}\sum_{j_1=1}^{J}\left(1 - \mathbb{E}_{c_{j_1}^{near}=+1}\big[\log_2\left(1+e^{-L_{o,e}(c_{j_1}^{near})}\right)\big]\right) \tag{3.44}$$

where $L_{o,e}(c_{j_1}^{near})$ is the corresponding output LLR of the channel decoder with the input *a priori* LLR $L_{o,a}(c_{j_1}^{near})$. The input LLR is generated by

$$L_{i,a}(c_{j_1}^{near}) = \frac{\sigma_A^2}{2}c_{j_1}^{near} + n_A = \frac{J^{-1}(I_{o,a}^{near})^2}{2}c_{j_1}^{near} + n_A \tag{3.45}$$

Correspondingly, the EXIT characteristic of the channel decoder is independent of the SNR^{near}. Hence, it can be written as

$$I_{o,e}^{near} = T_o^{near}[I_{o,a}^{near}] \tag{3.46}$$

Therefore, two EXIT characteristics T_o^{near} and T_i^{near} constitute the EXIT chart for the *Group near*. The same calculation is also carried out for the *Group far* with superscripts far and subscripts j_2.

The decoding trajectory, depending on a so-called open-convergence tunnel which is shaped by the two characteristics, illustrates convergence behavior of the proposed system. The wider the tunnel is, the less number of outer iterations are required to reach the convergence. The width of the open-convergence tunnel varies with the code rates of the error-correcting channel code. The low-density parity-check (LDPC) codes unified by 5G standards [38] is utilized in the proposed system.

Therefore, the strategy of our proposed optimization method consists in, firstly, setting the convergence target (i.e., the number of outer loop iterations), secondly, obtaining the EXIT charts of the two groups with different SNRs and channel code rates, finally, choosing the code rates for the groups according to the convergence target.

3.3.3.2 Analysis of the EXIT charts

To evaluate the system performance, a comparison of two systems, the proposed system and the conventional SCMA system, is studied. For the sake of simplicity, the latter system is denoted as SCMA system in the rest of this chapter. For fair comparison, the SCMA system accommodates the same number of users as in the proposed system. In the SCMA system, there are also two groups of users, with each of them having J users occupying K REs as demonstrated in [18]. For the SCMA system, each group detects their signals independently and uses the same SCMA codebook as in the proposed system. The SCMA codebook introduced in [107] with $M = 4$ is adopted in the simulation. The SCMA detection algorithm is MPA and the number of inner iterations is 3. The base graph (BG) of the LDPC code used in the simulation is BG_1, which means that the code rates can be selected from 1/3 to 8/9 [38]. The lifting size is $Z_c = 16$. The decoder of the LDPC code adopts the sum-product algorithm [108, Chapter 6] with 10 iterations. In the performance comparison, three SNR pairs $(\mathrm{SNR}^{near}, \mathrm{SNR}^{far})$, $(5.5dB, 0.5dB)$, $(7.5dB, 0.5dB)$ and $(9.5dB, 0.5dB)$, are considered. Last but not least, the convergence target is specified to 3 outer loop iterations.

The EXIT charts of the SCMA detector and the LDPC decoder considering different SNR pairs in both systems are shown in Fig. 3.10. By carrying out the proposed optimization method with the parameters mentioned above, the chosen LDPC code rates of both groups in two systems for different SNR scenarios are given in Table 3.3. R^{near} and R^{far} denote the code rate for the *Group near* and *Group far*, respectively. Fig. 3.10 also demonstrates the corresponding decoding trajectories of the *Group Near* for different SNR scenarios. We can see that there is no further gain after three outer

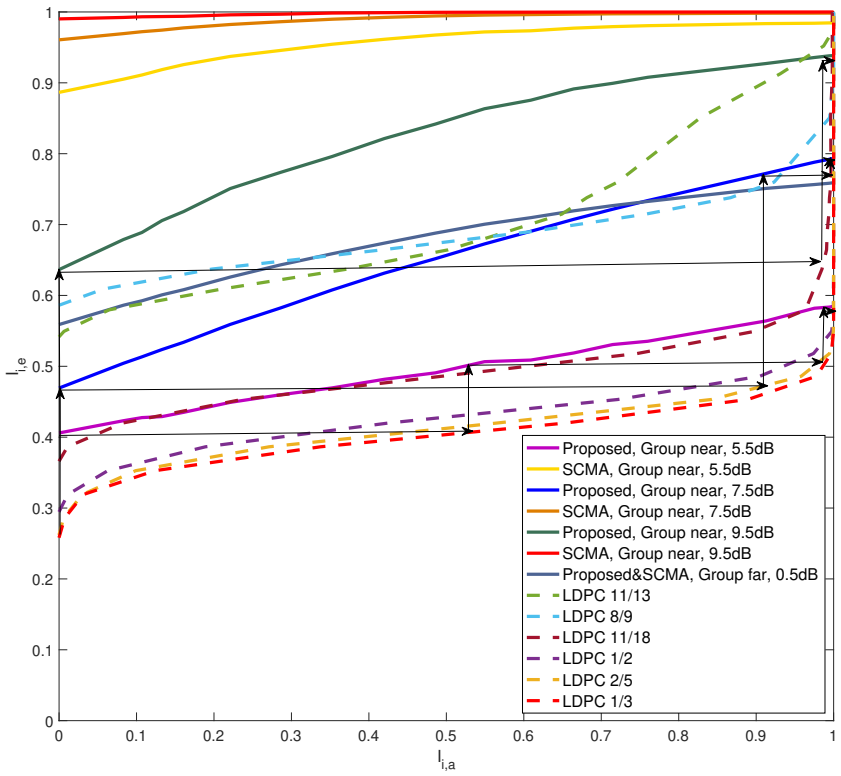


FIGURE 3.10 – EXIT characteristics of SCMA in both systems.

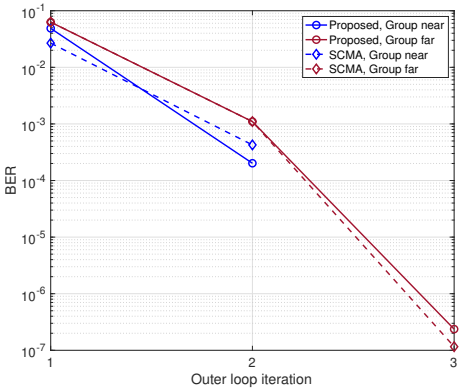
TABLE 3.3 – LDPC code rate parameters.

	Proposed system		SCMA system	
SNR pair/dB	R^{near}	R^{far}	R^{near}	R^{far}
(5.5, 0.5)	1/3	1/2	11/13	1/2
(7.5, 0.5)	2/5	1/2	8/9	1/2
(9.5, 0.5)	11/18	1/2	8/9	1/2

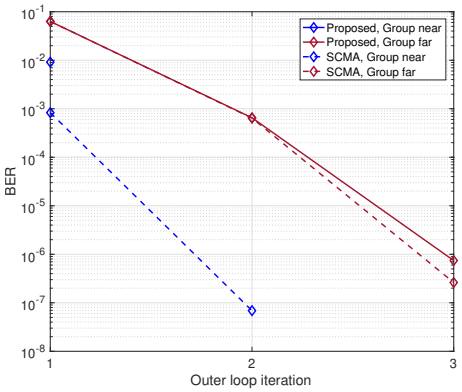
loop iterations with the chosen code rates, which meets the target. Moreover, Table 3.3 shows that the *Group near* requires a stronger code in the proposed system than in the SCMA system because of the inter-group interference.

3.3.4 Performance analysis

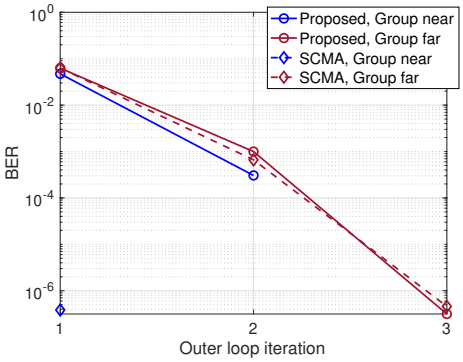
In this subsection, we evaluate the proposed system along with the optimization method by comparing it with the SCMA system in terms of coded BER performance and spectral efficiency. The system parameters including the code rates are those provided in section 3.3.3.2.



(a) Group SNR pair $(5.5dB, 0.5dB)$.



(b) Group SNR pair $(7.5dB, 0.5dB)$.



(c) Group SNR pair $(9.5dB, 0.5dB)$.

FIGURE 3.11 – BER results of two groups in different SNR scenarios.

3.3.4.1 BER results

Fig. 3.11 compares the BER performance of the *Group near* and *Group far* in different SNR scenarios. We have not shown the points corresponding to the BER less than 10^{-8} in the figure. We can see that the BER performance of the *Group far* is almost the same in both systems. It is notable that the BER performance of both groups is lower than 10^{-6} after 3 outer loop iterations in both systems. Therefore, the proposed optimization method can be deemed to be effective.

3.3.4.2 Spectral efficiency

As mentioned in [39], the system spectral efficiency is defined as $SE = \lambda(R^{near} + R^{far}) \log_2(M)$ bits/tone. The overloading factor of the proposed system is $\lambda^P = 300\%$, while that of the SCMA system is $\lambda^S = 150\%$. Fig. 3.12 illustrates the spectral efficiency in different systems distinctly by groups. First of all, it is notable that the proposed system has 23.8%, 29.6% and 60% of spectral efficiency improvement than the SCMA system when the SNR pairs are (5.5dB, 0.5dB), (7.5dB, 0.5dB)

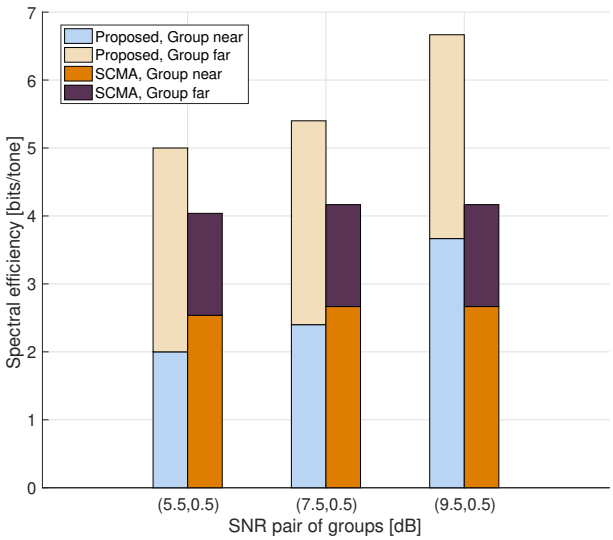


FIGURE 3.12 – Spectral efficiency comparison of two systems.

and (9.5dB, 0.5dB), respectively. It means that the proposed optimization method is effective in terms of augmenting spectral efficiency. Secondly, in the proposed system, *Group far* is the main contributor in the spectral efficiency increase when the SNR difference between both groups is equal to 5dB or 7dB. One reason is that the chosen LDPC code rates for the *Group near* in the proposed system coincide with half of those in SCMA system when SNR differences are 5dB and 7dB. Moreover, the proposed system doubles the spectral efficiency of the *Group far* thanks to the SIC. Last but not least, the spectral efficiency of the SCMA system is saturated when the SNR group pair is (9.5dB, 0.5dB) while that of the proposed system keeps increasing. As it is known from [63] that codebooks with higher values of M degrade the BER performance, the proposed system along with the optimization method further improves the spectral efficiency compared to the conventional SCMA with the provided parameters.

3.4 Conclusion

In this chapter, we focus on the iterative MUD algorithms for SCMA systems. First, the simple coded SCMA transmission scenario is considered. We review two sate-of-the-art SCMA MUD algorithms built on a linear expression of the codebook. To further decrease the computational complexity, we proposed a novel low-complexity SCMA MUD algorithm, named FO-GAA. The proposed algorithm shows substantial complexity decrease as well as fast convergence speed at the cost of slight BER degradation. Next, we introduce an uplink two-groups SCMA transmission scheme to meet the chal-

lenge of massive connectivity considering the limited amount of bandwidth in the 6G. In the proposed scheme, two groups of users share the same REs which augments the system overloading factor but inevitably results in the inter-group interference. Therefore, we apply the iterative algorithm and the SIC jointly at the receiver. In such a transmission system, there is a trade-off between the system convergence behavior and spectral efficiency. Thus, we propose an optimization method for choosing the code rates of the channel code with the purpose of guaranteeing the desired convergence behavior and maximizing the spectral efficiency at the same time. We take advantage of analyzing the EXIT charts in the optimization method and give an example of analyzing the EXIT charts in this chapter. The proposed scheme is compared with the conventional SCMA scheme in the performance evaluation. We show that the novel highly overload scheme along with the proposed optimization method can reach significantly high spectral efficiency and good BER performance.

Chapitre 4

SD-based detectors for SCMA systems

Contenu

4.1 Introduction

We have seen in the previous chapter that iterative detectors, for example MPA, can approach optimal MUD performance but become infeasible if the codebook size or overloading factor increases. On the other hand, SD has been recently considered in the MUD of SCMA systems by rewriting the generalized transmission into a linear system. This chapter is dedicated to the study of SD-based detectors for SCMA systems. In this chapter, we first review the state-of-the-art SD-based detectors for SCMA : sphere decoding for SCMA (SD-SCMA) and generalized SD-SCMA (GSD-SCMA). We not only explain the state-of-the-art in a comprehensive way, but also exploit the sorted QR decomposition and Schnorr-Euchner enumeration to accelerate the tree search. Although GSD-SCMA overcomes the codebook constraint of SD-SCMA, its computational complexity is extremely sensitive to the overloading factor. To satisfy the trade-off between complexity and MUD performance, we propose two pruning algorithms, PRUN1 and PRUN2, and introduce the simplified GSD-SCMA (SGSD-SCMA). An estimated error rate of the pruning algorithms is derived. We examine the MUD performance as well as computational complexity of various SD-based detectors by simulating with three setting of system parameters.

4.2 The rewritten system model in the real domain

In this chapter, we consider an uncoded SCMA system. In order to facilitate the application of SD-based detectors for SCMA, we rewrite the complex-valued system model expressed in Eq.(2.4) in the real domain. Please note that this rewriting is generalized for any type of SCMA codebooks.

To express the SCMA transmission Eq.(2.4) in the real domain, we separate the real and imaginary parts. The received signal in the real domain is defined as

$$\bar{\mathbf{y}} = \bar{\mathbf{G}}\bar{\mathbf{d}} + \bar{\mathbf{n}} \tag{4.1}$$

where $\bar{\mathbf{y}} \in \mathbb{R}^{2K\times 1}$, $\bar{\mathbf{G}} \in \mathbb{R}^{2K\times 2Jd_v}$, $\bar{\mathbf{d}} \in \mathbb{R}^{2Jd_v\times 1}$ and $\bar{\mathbf{n}} \in \mathbb{R}^{2K\times 1}$ are real-valued received signal vector, effective channel matrix, effective transmitted symbol vector and AWGN vector, respectively. For simplicity, in the remainder of this chapter we will refer to $\bar{\mathbf{G}}$ as the channel matrix and to $\bar{\mathbf{d}}$ as the transmitted symbol vector. The real-valued channel gain matrix $\bar{\mathbf{G}}$ is built from the complex matrix $\mathbf{G}$ by replacing each of its entries g_{il} $(1 \leqslant i \leqslant K,\ 1 \leqslant l \leqslant Jd_v)$ with $\begin{bmatrix} \Re\{g_{il}\} & -\Im\{g_{il}\} \\ \Im\{g_{il}\} & \Re\{g_{il}\} \end{bmatrix}$. For $\bar{\mathbf{d}}$ as a

column vector, it is constructed from its associated complex vector $\mathbf{d}$ by substituting every element d_{i_1} $(1 \leqslant i_1 \leqslant Jd_v)$ with $\begin{bmatrix}\Re\{d_{i_1}\} & \Im\{d_{i_1}\}\end{bmatrix}^\intercal$. The same real column conversion is also applied to $\mathbf{y}$ and $\mathbf{n}$ for obtaining $\bar{\mathbf{y}}$ and $\bar{\mathbf{n}}$, respectively. Besides, we define real-valued intermediate codebook $\bar{\mathcal{D}}_j$ whose entries are constructed from the complex entries in $\mathcal{D}_j$. Therefore, there is $\bar{\mathbf{d}} \in \bar{\mathcal{D}} = \bar{\mathcal{D}}_1 \times \bar{\mathcal{D}}_2 \cdots \times \bar{\mathcal{D}}_J$. Consequently, the ML detection for SCMA in the real domain is formulated as

$$\hat{\bar{\mathbf{d}}} = \arg\min_{\bar{\mathbf{d}} \in \bar{\mathcal{D}}} \|\bar{\mathbf{y}} - \bar{\mathbf{G}}\bar{\mathbf{d}}\|^2 \tag{4.2}$$

4.3 State-of-the-art SD-based SCMA detectors

In this section, we thoroughly review the two detectors, SD-SCMA and GSD-SCMA, which have less or no constraint of the codebook mentioned in a brief literature review provided in section 2.1.3. A comprehensive description of the two detectors is provided. We not only elaborate on them by illustrating with flowcharts and detailed pseudo-code algorithms, but also enhance the original scheme by adopting SQRD.

4.3.1 SD-SCMA

In stead of performing exhaustive search shown in Eq.(4.2), the SD can be deemed as ML with the conditions which is expressed as

$$\hat{\bar{\mathbf{d}}} = \arg\min_{\bar{\mathbf{d}} \in \bar{\mathcal{D}}} \|\bar{\mathbf{y}} - \bar{\mathbf{G}}\bar{\mathbf{d}}\|^2 \tag{4.3}$$

$$\text{subject to } \|\bar{\mathbf{y}} - \bar{\mathbf{G}}\bar{\mathbf{d}}\|^2 \leqslant d^2 \tag{4.4}$$

where constant d^2 is the squared sphere radius which defines the hyper-sphere for the tree search process. However, the channel matrix $\bar{\mathbf{G}}$ is a "flat" matrix since its number of rows is less than number of columns. The fact that $\bar{\mathbf{G}}$ is not a square matrix makes it difficult to solve Eq.(4.3)-(4.4) in an efficient way. To resolve this problem and properly implement SD in MUD of SCMA, SD-SCMA takes advantage of the method inspired by the Tikhonov regularization [109]. We define the real-valued modified channel matrix as

$$\tilde{\mathbf{G}} = \begin{bmatrix}\bar{\mathbf{G}}^{(1)} & \bar{\mathbf{G}}^{(2)} \\ \mathbf{0} & \mathbf{I}\end{bmatrix} \tag{4.5}$$

where $\bar{\mathbf{G}} = \begin{bmatrix}\bar{\mathbf{G}}^{(1)} & \bar{\mathbf{G}}^{(2)}\end{bmatrix}$. We remark that matrices $\bar{\mathbf{G}}^{(1)}$ and $\bar{\mathbf{G}}^{(2)}$ have sizes $2K \times 2K$ and $2K \times (2Jd_v - 2K)$ while the null matrix has $(2Jd_v - 2K) \times 2K$ and the identity matrix has size order

$2Jd_v - 2K$. The received signal in the real domain can be modified as

$$\tilde{\mathbf{y}} = \tilde{\mathbf{G}}\bar{\mathbf{d}} + \tilde{\mathbf{n}} \tag{4.6}$$

$$\begin{bmatrix}\bar{\mathbf{y}}\\ \mathbf{0}\end{bmatrix} = \tilde{\mathbf{G}}\begin{bmatrix}\bar{\mathbf{d}}^{(1)}\\ \bar{\mathbf{d}}^{(2)}\end{bmatrix} + \begin{bmatrix}\mathbf{n}\\ -\bar{\mathbf{d}}^{(2)}\end{bmatrix} \tag{4.7}$$

where $\bar{\mathbf{d}} = \begin{bmatrix}\bar{\mathbf{d}}^{(1)}\\ \bar{\mathbf{d}}^{(2)}\end{bmatrix}$. We remark that the vector $\bar{\mathbf{d}}^{(1)} \in \bar{\mathcal{D}}^{(1)}$ has size $2K \times 1$ and vector $\bar{\mathbf{d}}^{(2)} \in \bar{\mathcal{D}}^{(2)}$ has size $(2Jd_v - 2K) \times 1$, where $\bar{\mathcal{D}}^{(1)} = \bar{\mathcal{D}}_1 \times \cdots \times \bar{\mathcal{D}}_{\frac{K}{d_v}}$ and $\bar{\mathcal{D}}^{(2)} = \bar{\mathcal{D}}_{\frac{K}{d_v}+1} \times \cdots \times \bar{\mathcal{D}}_J$ while $\mathbf{0}$ in this case has size $(2Jd_v - 2K) \times 1$. Thus, the SD is rewritten as

$$\hat{\bar{\mathbf{d}}} = \arg\min_{\bar{\mathbf{d}} \in \bar{\mathcal{D}}} \|\tilde{\mathbf{y}} - \tilde{\mathbf{G}}\bar{\mathbf{d}}\|^2 - \bar{\mathbf{d}}^{(2)\intercal}\bar{\mathbf{d}}^{(2)} \tag{4.8}$$

$$\text{subject to } \|\tilde{\mathbf{y}} - \tilde{\mathbf{G}}\bar{\mathbf{d}}\|^2 - \bar{\mathbf{d}}^{(2)\intercal}\bar{\mathbf{d}}^{(2)} \leqslant d^2 \tag{4.9}$$

The prerequisite of applying the Tikhonov regularization in SD-SCMA is $\bar{\mathbf{d}}^{(2)}$ of constant modulus, which guarantees the equivalence between Eq.(4.3)-(4.4) and Eq.(4.8)-(4.9). As $\bar{\mathbf{d}}^{(2)}$ has constant modulus, Eq.(4.8)-(4.9) can be simplified to

$$\hat{\bar{\mathbf{d}}} = \arg\min_{\bar{\mathbf{d}} \in \bar{\mathcal{D}}} \|\tilde{\mathbf{y}} - \tilde{\mathbf{G}}\bar{\mathbf{d}}\|^2 \tag{4.10}$$

$$\text{subject to } \|\tilde{\mathbf{y}} - \tilde{\mathbf{G}}\bar{\mathbf{d}}\|^2 \leqslant d'^2 \tag{4.11}$$

where $d'^2 = d^2 - \|\bar{\mathbf{d}}^{(2)}\|^2$. This prerequisite can be regarded as a constraint of the SD-SCMA detector. This means that the SD-SCMA can be only applied to SCMA codebooks which guarantee $\bar{\mathbf{d}}^{(2)}$ to have constant modulus.

We adopt SQRD [110] to obtain an upper triangular matrix for the tree search in SD-SCMA. SQRD can be seen as a modified Gram-Schmidt algorithm which is famous for efficiently improving MIMO detection [111]. In the MIMO scenario, it minimizes the diagonal element modulus of the output upper triangular matrix from upper left to bottom right by reordering the columns of the input matrix, usually the channel matrix. As a result, SQRD is able to enhance the efficiency of the earlier layers in the tree search, which further reduces the complexity by decreasing the number of visited nodes. Moreover, SQRD guarantees output upper triangular matrix to have positive diagonal elements. The SQRD for SD-SCMA is written as

$$\text{sqrd}(\tilde{\mathbf{G}}) : \mathbf{QR} = \tilde{\mathbf{G}}\mathbf{P} \tag{4.12}$$

Algorithm 8 SQRD for SCMA.

Input : $\tilde{\mathbf{G}}$
Output : $\mathbf{R}$, $\mathbf{Q}$ and $\mathbf{P}$
Initialization : $\mathbf{Q} = \tilde{\mathbf{G}}$, $\mathbf{R} = \mathbf{0}^{2Jd_v \times 2Jd_v}$, $\mathbf{P} = \mathbf{I}_{2Jd_v}$ and $\mathbf{m} = \mathbf{0}^{J \times 1}$
for $j = 1, 2, \cdots, J$ **do**
  $m_j = \sum_{l=0}^{2d_v-1} \|\mathbf{q}_{j+l}\|_2^2$
end for
for $j = 1, 2, \cdots, J$ **do**
  $i_j = \arg\min_{l \in \{j, j+1 \cdots, J\}} m_l$
  Exchange columns $2(j-1)d_v + 1$ to $2jd_v$ and $2(i_j - 1)d_v + 1$ to $2i_j d_v$ one-by-one in $\mathbf{R}$, $\mathbf{Q}$ and $\mathbf{P}$
  for $i_1 = j, j+1, \cdots, j + (2d_v - 1)$ **do**
    $r_{i_1,i_1} = \|\mathbf{q}_{i_1}\|_2^2$
    $\mathbf{q}_{i_1} = \mathbf{q}_{i_1} / \sqrt{r_{i_1,i_1}}$
    for $i_2 = i_1 + 1, i_1 + 2, \cdots, 2Jd_v$ **do**
      $r_{i_1,i_2} = \mathbf{q}_{i_1}^{\intercal} \mathbf{q}_{i_2}$
      $\mathbf{q}_{i_2} = \mathbf{q}_{i_2} - r_{i_1,i_2} \mathbf{q}_{i_1}$
      $m_{\lceil i_2/(2d_v) \rceil} = m_{\lceil i_2/(2d_v) \rceil} - r_{i_1,i_2}$
    end for
  end for
end for

where $\mathbf{Q} \in \mathbb{R}^{2Jd_v \times 2Jd_v}$ is an orthogonal matrix, $\mathbf{R} \in \mathbb{R}^{2Jd_v \times 2Jd_v}$ is an upper triangular matrix and $\mathbf{P} \in \mathbb{R}^{2Jd_v \times 2Jd_v}$ is a permutation matrix. Considering that the columns with the indices $[2d_v j - (2d_v - 1), \cdots, 2d_v j]$ ($j \in \{1, 2, \cdots, J\}$) belong to the same user, the SQRD for SCMA should reorder the columns of $\tilde{\mathbf{G}}$ in groups of $2d_v$. To achieve this, SQRD for SCMA calculates the squared modulus sum of the $2d_v$ columns that are associated with each user. To better understand the SQRD for SCMA, please refer to Algorithm 8.

As SQRD introduces permutation among columns of $\tilde{\mathbf{G}}$, the decoding problem in Eq.(4.10)-(4.11) becomes

$$\hat{\bar{\mathbf{d}}} = \arg \min_{\bar{\mathbf{d}} \in \bar{\mathcal{D}}} \|\mathbf{Q}^{\intercal} \tilde{\mathbf{y}} - \mathbf{R}\mathbf{P}^{-1} \bar{\mathbf{d}}\|^2 \tag{4.13}$$

$$\text{subject to } \|\mathbf{Q}^{\intercal} \tilde{\mathbf{y}} - \mathbf{R}\mathbf{P}^{-1} \bar{\mathbf{d}}\|^2 \leqslant d'^2 \tag{4.14}$$

By defining $\tilde{\mathbf{y}}' = \mathbf{Q}^{\intercal} \tilde{\mathbf{y}}$ and $\bar{\mathbf{d}}^{\mathrm{P}} = \mathbf{P}^{-1} \bar{\mathbf{d}}$, the above decoding problem is simplified to

$$\hat{\bar{\mathbf{d}}} = \mathbf{P} \arg \min_{\bar{\mathbf{d}}^{\mathrm{P}} \in \bar{\mathcal{D}}^{\mathrm{P}}} \|\tilde{\mathbf{y}}' - \mathbf{R}\bar{\mathbf{d}}^{\mathrm{P}}\|^2 \tag{4.15}$$

$$\text{subject to } \|\tilde{\mathbf{y}}' - \mathbf{R}\bar{\mathbf{d}}^{\mathrm{P}}\|^2 \leqslant d'^2 \tag{4.16}$$

where $\bar{\mathcal{D}}^{\mathrm{P}}$ represents the permuted real-valued codebook collection. Since $\mathbf{R}$ is an upper triangular

matrix, Eq.(4.16) can be rewritten as

$$d'^2 \geqslant \sum_{i=1}^{2Jd_v} \left(\tilde{y}'_i - \sum_{l=i}^{2Jd_v} r_{i,l} \bar{d}_l^{\mathrm{P}} \right)^2 \tag{4.17}$$

$$= (\tilde{y}'_{2Jd_v} - r_{2Jd_v,2Jd_v} \bar{d}_{2Jd_v}^{\mathrm{P}})^2 + \left(\tilde{y}'_{2Jd_v-1} - \sum_{l=2Jd_v-1}^{2Jd_v} r_{2Jd_v-1,l} \bar{d}_l^{\mathrm{P}} \right)^2 + \cdots \tag{4.18}$$

Therefore, it is natural to start the tree search from the bottom and trace upwards. The layers with the indices $i \in \{2d_v j - (2d_v - 1), \cdots, 2d_v j\}$ ($j \in \{1, 2, \cdots, J\}$) correspond to user j. Because each user spreads the same information bits over d_v REs, $2d_v$ layers should be correlated in the tree search of SD-SCMA. The layers with indices i multiple of $2d_v$ have M branches and the rest has only one branch. In other words, the layers having indices $2d_v j$ ($j \in \{1, 2, \cdots, J\}$) determine the layers with indices from $2d_v j - (2d_v - 1)$ to $2d_v j - 1$. As a result, branch selection calculation is performed only in J layers. Details of the SD-SCMA tree search is illustrated by a flowchart as shown in Fig. 4.1.

At the $2Jd_v$-th layer, for the first branch selection calculation, the necessary condition of Eq.(4.17) is

$$(\tilde{y}'_{2Jd_v} - r_{2Jd_v,2Jd_v} \bar{d}_{2Jd_v}^{\mathrm{P}})^2 \leqslant d'^2 \tag{4.19}$$

The corresponding range of $\bar{d}_{2Jd_v}^{\mathrm{P}}$ is

$$\frac{\tilde{y}'_{2Jd_v} - d'}{r_{2Jd_v,2Jd_v}} \leqslant \bar{d}_{2Jd_v}^{\mathrm{P}} \leqslant \frac{\tilde{y}'_{2Jd_v} + d'}{r_{2Jd_v,2Jd_v}} \tag{4.20}$$

To extend this to the i-th ($i = 2d_v j,\ j \in \{1, 2, \cdots, J\}$) layer, we define the lower and upper bounds of $\bar{d}_i^{\mathrm{P}}$ as

$$l_i = \frac{1}{r_{i,i}} \left(\tilde{y}'_i - \sum_{l=i+1}^{2Jd_v} r_{i,l} \bar{d}_l^{\mathrm{P}} - \sqrt{d'^2 - p_{i+2d_v}} \right) \tag{4.21}$$

and

$$u_i = \frac{1}{r_{i,i}} \left(\tilde{y}'_i - \sum_{l=i+1}^{2Jd_v} r_{i,l} \bar{d}_l^{\mathrm{P}} + \sqrt{d'^2 - p_{i+2d_v}} \right) \tag{4.22}$$

respectively, where p_i is the path metric in the i-th layer which will be introduced later. Thus, the corresponding range of $\bar{d}_i^{\mathrm{P}}$ is expressed as

$$l_i \leqslant \bar{d}_i^{\mathrm{P}} \leqslant u_i \tag{4.23}$$

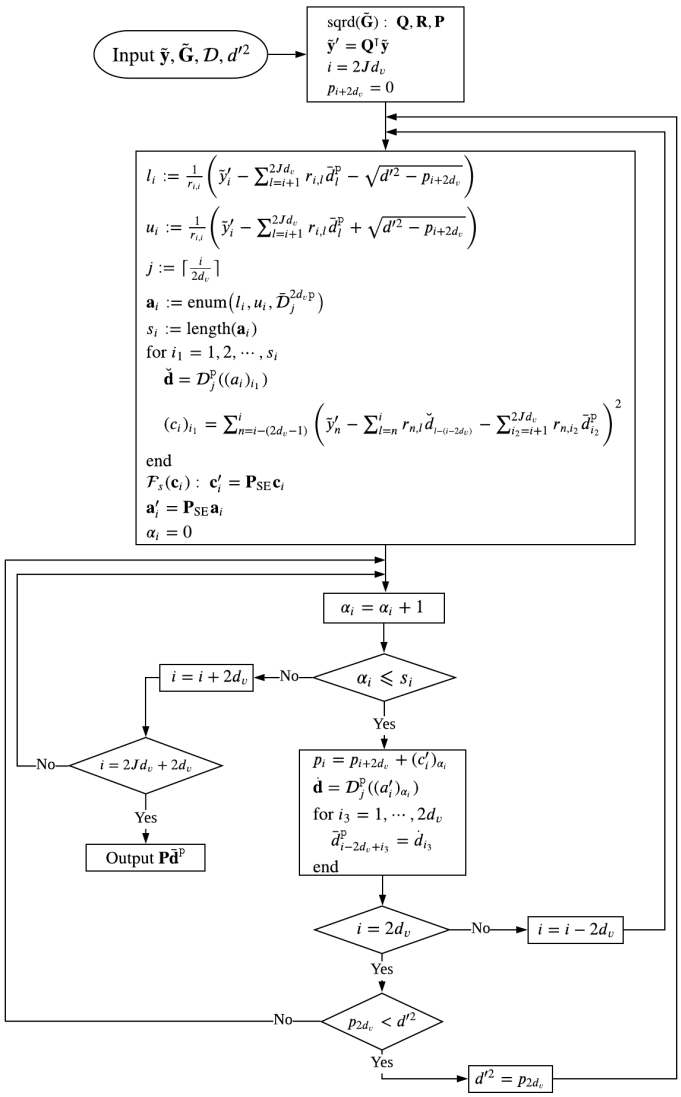


FIGURE 4.1 – Flowchart of SD-SCMA.

We use vector $\mathbf{a}_i$ to record the indices of the candidate codewords, i.e., the codewords that satisfy the condition Eq.(4.23), which can be expressed as

$$\mathbf{a}_i = \mathrm{enum}(l_i, u_i, \bar{\mathcal{D}}_j^{2d_v\mathrm{P}}) \tag{4.24}$$

where $j = \lceil \frac{i}{2d_v} \rceil$, $\mathrm{enum}(v_1, v_2, \mathcal{S})$ denotes the function enumerating all the positions of elements in the real sequence $\mathcal{S}$ having values between v_1 and v_2 and $\bar{\mathcal{D}}_l^k$ represents the real sequence consisting of all the k-th entries of the elements in the sequence $\bar{\mathcal{D}}_l$. s_i is the number of valid branches associated with the candidate codewords at the i-th layer, obtained by

$$s_i = \mathrm{length}(\mathbf{a}_i) \tag{4.25}$$

According to the tree search rules for SCMA, $\mathbf{a}_i$ and s_i with $i = 2d_v j,\ j \in \{1, 2, \cdots, J\}$ are the vector of candidate codewords indices and the number of candidate codewords of user j, respectively. Hence, to visit the next branch selection layer, the layer index is updated by $i = i - 2d_v$ as shown in Fig. 4.1.

To accelerate the tree search in SD-SCMA, we adopt the SE strategy [112] to order the codeword candidates. In Schnorr-Euchner enumeration, the candidates are examined and sorted based on their path metric values. As the path metric is identical for all child nodes having the same parent in tree search, the ordering is based on the branch metrics of child nodes. At the i-th layer ($i = 2d_v j,\ j \in \{1, 2, \cdots, J\}$), all the candidate codewords

$$\breve{\mathbf{d}} = \mathcal{D}_j^{\mathrm{p}}((a_i)_{i_1}), \quad 1 \leqslant i_1 \leqslant s_i \tag{4.26}$$

are considered, where $(a_i)_{i_1}$ refers to the i_1-th elements of vector $\mathbf{a}_i$. For a given candidate codeword $\breve{\mathbf{d}}$, its branch metric is calculated as

$$(c_i)_{i_1} = \sum_{n=i-(2d_v-1)}^{i} \left(\tilde{y}'_n - \sum_{l=n}^{i} r_{n,l} \breve{d}_{l-(i-2d_v)} - \sum_{i_2=i+1}^{2Jd_v} r_{n,i_2} \bar{d}_{i_2}^{\mathrm{p}} \right)^2 \tag{4.27}$$

where $\breve{d}_k$ is the k-th entry of vector $\breve{\mathbf{d}}$. Consequently, vector $\mathbf{c}_i = [(c_i)_1, (c_i)_2, \cdots, (c_i)_{s_i}]^{\intercal}$ stores all the branch metrics of the candidate codewords at i-th layer. Then, we sort the elements of vector $\mathbf{c}_i$ to satisfy the condition $(c_i)_n \leqslant (c_i)_{n+1}$ where $1 \leqslant n \leqslant s_i - 1$. This sorting process can be formulated as

$$\mathcal{F}_s(\mathbf{c}_i):\ \mathbf{c}'_i = \mathbf{P}_{\mathrm{SE}} \mathbf{c}_i \tag{4.28}$$

where function $\mathcal{F}_s(\cdot)$ refers to the aforementioned sorting criterion, $\mathbf{c}'_i$ denotes the sorted vector and $\mathbf{P}_{\mathrm{SE}}$ is the permutation matrix. Then, the same element-wise permutation is applied to obtain the sorted codeword indices vector $\mathbf{a}'_i$, denoted as

$$\mathbf{a}'_i = \mathbf{P}_{\mathrm{SE}} \mathbf{a}_i \tag{4.29}$$

α_i is used to refer to the index of the chosen valid branch in the i-th layer, which follows $1 \leqslant \alpha_i \leqslant s_i$. If the branch index α_i is chosen, the cumulative path metric p_i ($i = 2d_v j,\ j \in \{1, 2, \cdots, J\}$) is computed by

$$p_i = p_{i+2d_v} + (c'_i)_{\alpha_i} \tag{4.30}$$

where $(c_i')_{\alpha_i}$ is the α_i-th element of vector $\mathbf{c}_i'$. At the beginning of the tree search, $p_{2Jd_v+2d_v} = 0$ is initialized. The decoding results are partially updated by the chosen branch, expressed as

$$\bar{d}^{\mathrm{P}}_{i-2d_v+i_3} = \dot{d}_{i_3}, \quad 1 \leqslant i_3 \leqslant 2d_v \tag{4.31}$$

where $\dot{\mathbf{d}}$ is an intermediate vector $\dot{\mathbf{d}} = \mathcal{D}_j^{\mathrm{P}}((a_i')_{\alpha_i})$. Every time, when a temporary best solution is found, the squared radius is updated by

$$d'^2 = p_{2d_v} \tag{4.32}$$

The tree search stops when no better solution can be found. The output of the algorithm is the final result $\mathbf{P}\bar{\mathbf{d}}^{\mathrm{P}}$.

4.3.2 GSD-SCMA

As discussed above, SD-SCMA is only compatible with codebooks having $\bar{\mathbf{d}}^{(2)}$ of constant modulus. To circumvent this codebook power constraint, GSD-SCMA is proposed for generalized SCMA codebooks [22].

Since the channel matrix can be expressed as $\bar{\mathbf{G}} = \begin{bmatrix}\bar{\mathbf{G}}^{(1)} & \bar{\mathbf{G}}^{(2)}\end{bmatrix}$, it can be rewritten as

$$\bar{\mathbf{G}} = \begin{bmatrix}\bar{\mathbf{G}}^{(1)} & \bar{\mathbf{G}}^{(2)}\end{bmatrix} \tag{4.33}$$

$$= \mathbf{Q}_1 \begin{bmatrix}\mathbf{R}_1\mathbf{P}_1^{-1} & \mathbf{Q}_1^{-1}\bar{\mathbf{G}}^{(2)}\end{bmatrix} \tag{4.34}$$

$$= \mathbf{Q}_1 \begin{bmatrix}\mathbf{R}_1\mathbf{P}_1^{-1} & \mathbf{R}_2\end{bmatrix} \tag{4.35}$$

where $\mathbf{Q}_1$ and $\mathbf{R}_1$ are outputs of SQRD of $\bar{\mathbf{G}}^{(1)}$ expressed as

$$\mathrm{sqrd}(\bar{\mathbf{G}}^{(1)}) : \mathbf{Q}_1\mathbf{R}_1 = \bar{\mathbf{G}}^{(1)}\mathbf{P}_1 \tag{4.36}$$

and $\mathbf{R}_2 \in \mathbb{R}^{2K\times(2Jd_v-2K)}$ is the multiplication of matrix $\mathbf{Q}_1^{-1}$ and $\bar{\mathbf{G}}^{(2)}$. Please note that because the first $2K$ columns of the transmission mapping matrix $\mathbf{S}$ are orthogonal, $\mathbf{R}_1$ is a positive diagonal matrix which simplifies the calculation in the tree search of GSD-SCMA.

Based on the channel matrix expression, Eq.(4.35), the decoding problem in Eq.(4.3)-(4.4) becomes

$$\hat{\bar{\mathbf{d}}} = \arg \min_{\substack{\bar{\mathbf{d}}^{(1)}\in\bar{\mathcal{D}}^{(1)} \\ \bar{\mathbf{d}}^{(2)}\in\bar{\mathcal{D}}^{(2)}}} \|\mathbf{Q}_1^{\mathsf{T}}\bar{\mathbf{y}} - \mathbf{R}_1\mathbf{P}_1^{-1}\bar{\mathbf{d}}^{(1)} - \mathbf{R}_2\bar{\mathbf{d}}^{(2)}\|^2 \tag{4.37}$$

$$\text{subject to } \|\mathbf{Q}_1^{\mathsf{T}}\bar{\mathbf{y}} - \mathbf{R}_1\mathbf{P}_1^{-1}\bar{\mathbf{d}}^{(1)} - \mathbf{R}_2\bar{\mathbf{d}}^{(2)}\|^2 \leqslant d^2 \tag{4.38}$$

which is known as the GSD-SCMA decoding problem [22]. Defining $\bar{\mathbf{y}}' = \mathbf{Q}_1^{\intercal}\bar{\mathbf{y}} - \mathbf{R}_2\bar{\mathbf{d}}^{(2)}$ and $\bar{\mathbf{d}}^{(1)\mathtt{p}} = \mathbf{P}_1^{-1}\bar{\mathbf{d}}^{(1)}$, Eq.(4.37)-(4.38) become

$$\hat{\bar{\mathbf{d}}} = \arg \min_{\bar{\mathbf{d}}^{(2)} \in \bar{\mathcal{D}}^{(2)}} \min_{\bar{\mathbf{d}}^{(1)\mathtt{p}} \in \bar{\mathcal{D}}^{(1)\mathtt{p}}} \|\bar{\mathbf{y}}' - \mathbf{R}_1\bar{\mathbf{d}}^{(1)\mathtt{p}}\|^2 \tag{4.39}$$

$$\text{subject to } \|\bar{\mathbf{y}}' - \mathbf{R}_1\bar{\mathbf{d}}^{(1)\mathtt{p}}\|^2 \leqslant d^2 \tag{4.40}$$

where $\bar{\mathcal{D}}^{(1)\mathtt{p}}$ represents the codebook collection $\bar{\mathcal{D}}^{(1)}$ permuted by $\mathbf{P}_1$. Since $\bar{\mathbf{y}}'$ is dependent on $\bar{\mathbf{d}}^{(2)}$, the latter impacts on the decoding results of $\bar{\mathbf{d}}^{(1)\mathtt{p}}$. Every choice of $\bar{\mathbf{d}}^{(2)}$ must be considered in the detection problem in Eq.(4.39) since there is no codebook power constraint. As K is an integer multiple of d_v, vector $\bar{\mathbf{d}}^{(2)}$ encompasses the real-valued codewords of the last $J' = J - \frac{K}{d_v} = J(1 - 1/d_f)$ users. Vector $\bar{\mathbf{d}}^{(2)}$ can take on $M^{J'}$ different values that will be tested during the decoding. It is noticeable that as the overloading factor increases (hence, d_f increases), J' gets larger and larger. For heavy overloading scenarios, the complexity of the exhaustive search in GSD-SCMA will approach the one of ML detector.

Fig. 4.2 shows the flowchart of GSD-SCMA. We initialize a sequence $\mathcal{L}_2 = (\mathbf{l}_1, \mathbf{l}_2, \cdots, \mathbf{l}_{M^{J'}})$ which contains all the $M^{J'}$ possible values of vector $\bar{\mathbf{d}}^{(2)}$. GSD-SCMA performs SD for $\bar{\mathbf{d}}^{(1)\mathtt{p}}$ based on every element in the sequence $\mathcal{L}_2$. For a given element in the sequence $\mathcal{L}_2$, namely $\mathbf{l}_t$ $(1 \leqslant t \leqslant M^{J'})$, if the tree search outputs $\mathbf{v}$ based on $\bar{\mathbf{y}}' = \mathbf{Q}_1^{\intercal}\bar{\mathbf{y}} - \mathbf{R}_2\mathbf{l}_t$, the value of $\bar{\mathbf{d}}^{(1)\mathtt{p}}$ and $\bar{\mathbf{d}}^{(2)}$ are updated as

$$\bar{\mathbf{d}}^{(1)\mathtt{p}} = \mathbf{v} \tag{4.41}$$

and

$$\bar{\mathbf{d}}^{(2)} = \mathbf{l}_t \tag{4.42}$$

respectively. As the squared radius d^2 gets smaller every time a solution of $\bar{\mathbf{d}}^{(1)\mathtt{p}}$ is found, the current values of $\bar{\mathbf{d}}^{(1)\mathtt{p}}$ and $\bar{\mathbf{d}}^{(2)}$ are better than the previous ones. Therefore, after testing on all $M^{J'}$ values of vector $\bar{\mathbf{d}}^{(2)}$, the best decoding results according to Eq.(4.39)-(4.38) can be found.

Specifically, for a given vector $\mathbf{l}_t$ $(1 \leqslant t \leqslant M^{J'})$, since we have $\bar{\mathbf{y}}' = \mathbf{Q}_1^{\intercal}\bar{\mathbf{y}} - \mathbf{R}_2\mathbf{l}_t$ and the fact that $\mathbf{R}_1$ is diagonal, Eq.(4.40) can be rewritten as

$$d^2 \geqslant \sum_{i=1}^{2K} \left(\bar{y}_i' - (r_1)_{i,i}\bar{d}_i^{(1)\mathtt{p}}\right)^2 \tag{4.43}$$

$$= (\tilde{y}_{2K}' - (r_1)_{2K,2K}\bar{d}_{2K}^{(1)\mathtt{p}})^2 + (\tilde{y}_{2K-1}' - (r_1)_{2K-1,2K-1}\bar{d}_{2K}^{(1)\mathtt{p}})^2 + \cdots \tag{4.44}$$

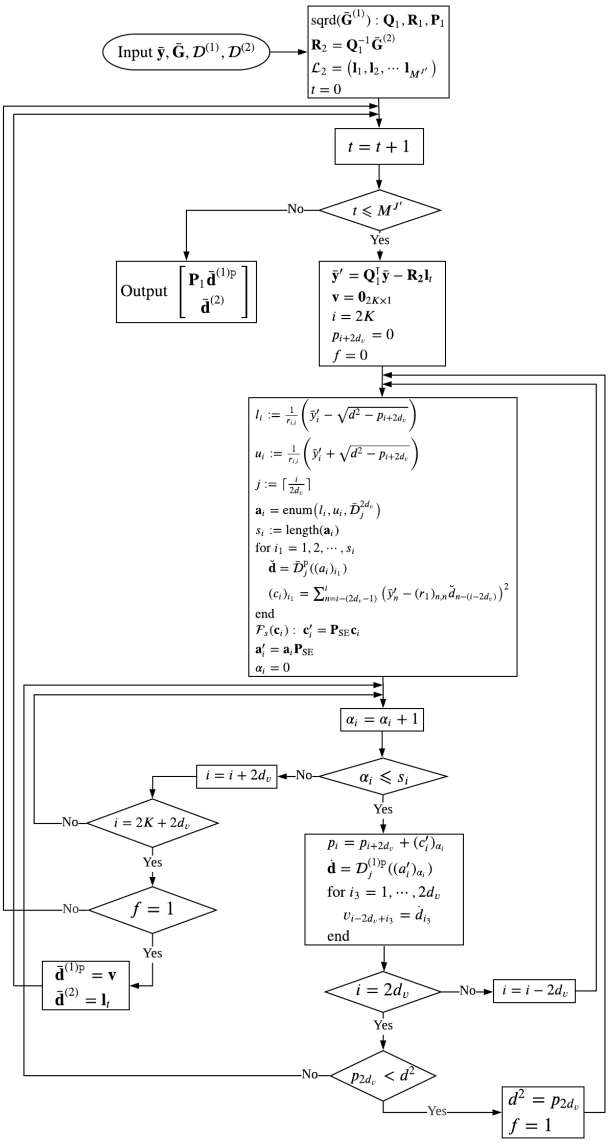


FIGURE 4.2 – Flowchart of GSD-SCMA.

where $(r_1)_{i,l}$ is the element on the i-th row and l-th column of matrix $\mathbf{R}_1$. The tree search for $\bar{\mathbf{d}}^{(1)\mathtt{p}}$ is similar to what have been explained in the previous subsection. Similarly, at the i-th ($i = 2d_v j,\ j \in \{1, \cdots, \lceil \frac{K}{2d_v} \rceil\}$) layer, the lower and upper bounds of $\bar{d}_i^{(1)\mathtt{p}}$ are computed as

$$l_i = \frac{1}{r_{i,i}}\left(\bar{y}'_i - \sqrt{d^2 - p_{i+2d_v}}\right) \tag{4.45}$$

and

$$u_i = \frac{1}{r_{i,i}}\left(\bar{y}'_i + \sqrt{d^2 - p_{i+2d_v}}\right) \tag{4.46}$$

respectively. Accordingly, in GSD-SCMA the branch metric is computed by

$$(c_i)_{i_1} = \sum_{n=i-(2d_v-1)}^{i} \left(\bar{y}'_n - (r_1)_{n,n}\breve{d}_{n-(i-2d_v)}\right)^2 \tag{4.47}$$

where $\breve{d}_k$ is the k-th entry of vector $\breve{\mathbf{d}}$ defined in Eq.(4.26). Considering the permutation in SQRD of $\bar{\mathbf{G}}^{(1)}$, the final decoding results of GSD-SCMA is obtained by

$$\hat{\bar{\mathbf{d}}} = \begin{bmatrix} \mathbf{P}_1 \bar{\mathbf{d}}^{(1)\mathrm{p}} \\ \bar{\mathbf{d}}^{(2)} \end{bmatrix} \tag{4.48}$$

4.4 Introducing pruning algorithm : a simple example

In this section, we take a toy example to explain the idea of pruning algorithm, which helps to introduce the proposed pruning algorithms for SCMA in the next section. In this toy example, we consider a real-valued multiuser uplink overloading transmission. Specifically, there are two users transmitting real-valued signals to the same base station, where each end has single antenna. Thus, the received signal is expressed as

$$y = \mathbf{h}\mathbf{x} + n = [h_1 \ h_2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + n \tag{4.49}$$

where $\mathbf{h} \in \mathbb{R}^{1\times 2}$ is the channel vector, $\mathbf{x} \in \mathbb{R}^{2\times 1}$ is the transmitted signal vector and $n \in \mathcal{N}(0, \sigma_1^2)$ is AWGN. The transmitted symbols x_1 and x_2 are drawn independently from an identical codebook $\mathcal{C}$ whose cardinality is M_1. We assume that the CSI is known perfectly at the receiver. By applying the GSD rules, the decoding problem of this example system is written as

$$\hat{\mathbf{x}} = \arg\min_{x_2 \in \mathcal{C}} \arg\min_{x_1 \in \mathcal{C}} \|y' - h_1 x_1\|^2 \tag{4.50}$$

$$\text{subject to } \|y' - h_1 x_1\|^2 \leqslant d^2 \tag{4.51}$$

where $y' = y - h_2 x_2$, which is equivalent to an SD problem of x_1 based on an exhaustive test of x_2. To reduce the decoding complexity, we adopt a pruning algorithm aiming at reducing the number of tested x_2. We denote $\mathcal{C}^*$ as the reduced candidate sequence of x_2. Without considering the additive noise, the pruning algorithm builds the sequence $\mathcal{C}^*$ in an efficient way, given as

$$x_2 \in \mathcal{C} \tag{4.52}$$

$$\text{subject to} \quad \min(h_1 x_1) \leqslant |y - h_2 x_2| \leqslant \max(h_1 x_1) \tag{4.53}$$

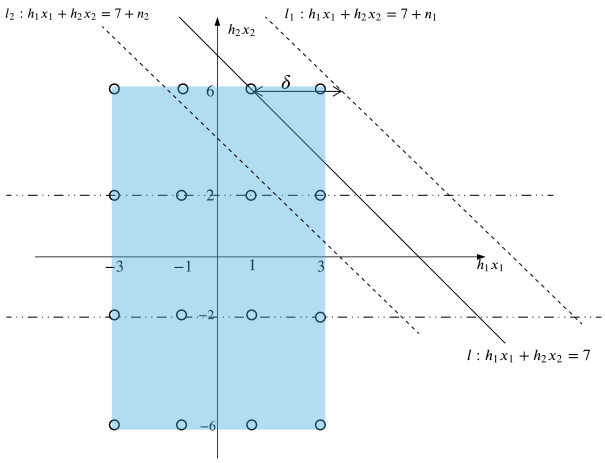


FIGURE 4.3 – Decoding diagram of the transmission example.

Considering the sign of h_1, Eq.(4.53) can be rewritten as

$$\begin{cases} h_1 \min(x_1) \leqslant |y - h_2x_2| \leqslant h_1 \max(x_1), \ h_1 > 0 \\ h_1 \max(x_1) \leqslant |y - h_2x_2| \leqslant h_1 \min(x_1), \ h_1 \leqslant 0 \end{cases} \tag{4.54}$$

After applying the pruning algorithm, the decoding problem becomes

$$\hat{\mathbf{x}} = \arg \min_{x_2 \in \mathcal{C}^*} \arg \min_{x_1 \in \mathcal{C}} \|y' - h_1x_1\|^2 \tag{4.55}$$

$$\text{subject to } \|y' - h_1x_1\|^2 \leqslant d^2 \tag{4.56}$$

In the following, we explain the pruning algorithm geometrically. As an example, consider that the channel vector is $\mathbf{h} = \begin{bmatrix} 1 & -2 \end{bmatrix}$ and the transmitted symbol vector is $\mathbf{x} = \begin{bmatrix} 1 & -3 \end{bmatrix}^\mathsf{T}$, whose entries are drawn independently from the PAM-4 codebook $\mathcal{C}$. The received constellation points transformed by the channel vector are illustrated in Fig. 4.3. The constellation region is marked blue where all points lie. The received signal without any noise is $\mathbf{hx} = 7$ depicted as line l in the figure. Ideally, there is no noise and the pruning criterion is that the value of h_2x_2 whose geometric representation does not have an intersection point with the line l in the constellation region is eliminated. For example, as the line of $h_2x_2 = 2$ has an intersection point with the line l that is outside of the constellation region, $x_2 = -1$ is pruned from the candidate sequence $\mathcal{C}^*$.

Because of the existence of noise n, the received signal in practice can be represented by line l_1 (l_2) which is horizontally shifted by the noise n_1 (n_2) as shown in Fig. 4.3. If the received signal can be represented as l_1, it has no intersection point with the line $h_2x_2 = 6$ in the constellation region, while line l meets the line $h_2x_2 = 6$ in the constellation region. This implies that when the original

transmission signal, line l, is impaired by noise becoming l_1, the value $x_2 = -3$, which is transmitted, will be pruned by accident. Considering the additive noise in practice, the sequence $\mathcal{C}^*$ built by the pruning algorithm is refined to

$$x_2 \in \mathcal{C} \tag{4.57}$$

$$\text{subject to } \begin{cases} h_1 \min(x_1) - \delta \leqslant |y - h_2 \hat{x}_2| \leqslant h_1 \max(x_1) + \delta, \ h_1 > 0 \\ h_1 \max(x_1) - \delta \leqslant |y - h_2 \hat{x}_2| \leqslant h_1 \min(x_1) + \delta, \ h_1 \leqslant 0 \end{cases} \tag{4.58}$$

where δ is a positive constant that is calculated based on the additive noise level. As the additive noise n follows $\mathcal{N}(0, \sigma_1^2)$, we define the error detection probability p guaranteeing that

$$p = Q\left(\frac{\delta}{\sigma_1}\right) \tag{4.59}$$

where $Q(\cdot)$ is the Q-function. Consequently, δ is computed by

$$\delta = \mathcal{F}^{-1}(1 - p, 0, \sigma_1) \tag{4.60}$$

where $\mathcal{F}^{-1}(x, \mu, \sigma')$ denotes the inverse cumulative distribution function of the Gaussian distribution with mean μ and standard deviation σ'. Since function $\mathcal{F}^{-1}(\cdot)$ is monotonically increasing, the smaller p is, the larger δ is. The increase of δ can bring decoding performance gain but high computational complexity. Thus, there is a trade-off between the decoding performance and complexity when choosing the value of p. Normally, p is set to a positive real number that is very close to 0.

In the following, we will discuss the impact of p on the symbol error rate (SER) performance of the transmission example. The SER is defined as $Pr\{\hat{\mathbf{x}}_\mathrm{T} \neq \mathbf{x}_\mathrm{T}\}$. We define an error probability $Pr\{x_{\mathrm{T}2} \notin \mathcal{C}^*\}$ describing the probability that the transmitted codeword $x_{\mathrm{T}2}$ is not included in the reduced sequence $\mathcal{C}^*$, also known as error rate. SER of the proposed pruning detector can be written as

$$Pr\{\hat{\mathbf{x}} \neq \mathbf{x}\} = P_\mathrm{ML}\big(1 - Pr\{x_{\mathrm{T}2} \notin \mathcal{C}^*\}\big) + Pr\{x_{\mathrm{T}2} \notin \mathcal{C}^*\} \tag{4.61}$$

where P_ML is the SER of the ML detector. If the term $Pr\{x_{\mathrm{T}2} \notin \mathcal{C}^*\}$ is very close to zero, the above equation can be approximated to

$$Pr\{\hat{\mathbf{x}} \neq \mathbf{x}\} \approx P_\mathrm{ML} + Pr\{x_{\mathrm{T}2} \notin \mathcal{C}^*\} \tag{4.62}$$

In high SNR regime, as P_ML approaches to 0, the SER can be formulated as

$$\lim_{\mathrm{SNR}\to\infty} Pr\{\hat{\mathbf{x}} \neq \mathbf{x}\} = Pr\{x_{\mathrm{T}2} \notin \mathcal{C}^*\} \tag{4.63}$$

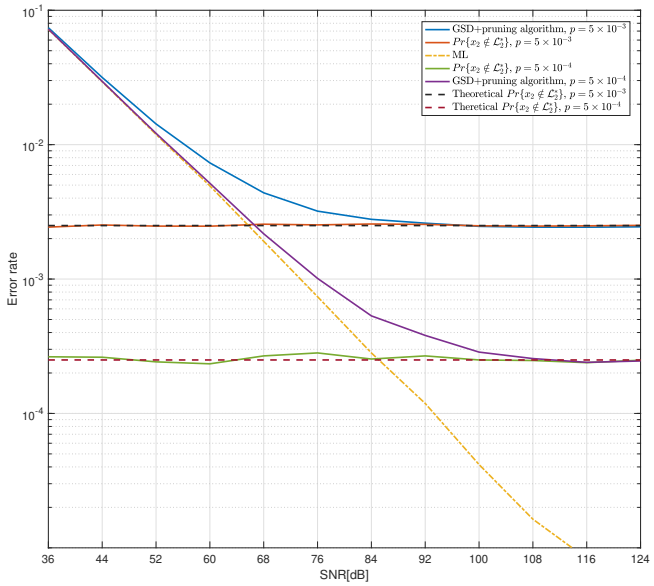


FIGURE 4.4 – SER comparison of the example transmission.

In this transmission example, each real dimension has M_1 distinct projections, half positive and half negative. For the $M_1/2$ positive projections, we assume that only the decoding accuracy of the projection on the edge, the maximum one, is affected by the pruning algorithm due to the additive noise. So do the negative projections. Since the pruning is performed only once in the example, $Pr\{x_{\mathrm{T}2} \notin \mathcal{C}^*\}$ is calculated by

$$Pr\{x_2 \notin \mathcal{L}_2^*\} = \frac{p}{M_1/2} = \frac{2p}{M_1} \tag{4.64}$$

which is proportional with p and inversely proportional with M_1.

To verify the accuracy of Eq.(4.64), we simulate the SER performance of the transmission example with codebook PAM-4. Fig. 4.4 shows the error rate of different detectors in the transmission example. The Monte Carlo simulation of SER and $Pr\{x_2 \notin \mathcal{C}^*\}$ is illustrated in the figure. In high SNR regime, the SER performance of the proposed pruning detector follows Eq.(4.63). That is, $Pr\{x_2 \notin \mathcal{C}^*\}$ is the error floor of the GSD detector with the proposed pruning algorithm. To verify the effectiveness of Eq.(4.64), we draw the theoretical performance of $Pr\{x_2 \notin \mathcal{C}^*\}$ in the same figure. As shown in Fig. 4.4, the simulation performance fits the theoretical analysis in Eq.(4.64) perfectly.

4.5 Proposed SGSD-SCMA and pruning algorithms

Recall that, since the GSD-SCMA detector performs an exhaustive test of $M^{J'}$ different values of $\bar{\mathbf{d}}^{(2)}$, where $J' = J(1 - 1/d_f)$, the computational complexity of GSD-SCMA approaches that of ML

when overloading factor is large. To efficiently decrease the computational complexity of GSD-SCMA, it is sensible to consider a reduced sequence $\mathcal{L}_2^*$ for $\bar{\mathbf{d}}^{(2)}$ values. Motivated by the pruning methodology demonstrated in section 4.4, we propose SGSD-SCMA, the GSD-SCMA detector simplified by applying a pruning algorithm for $\bar{\mathbf{d}}^{(2)}$ values.

The proposed SGSD-SCMA consists of two steps described as follows :

- Step1 : A pruning algorithm obtains the reduced sequence $\mathcal{L}_2^*$, which contains potentially good values of vector $\bar{\mathbf{d}}^{(2)}$. The pruning algorithm for SGSD-SCMA follows the same principle of the pruning criterion illustrated by the transmission example in section 4.4, but with the particularities of the SCMA transmission.
- Step2 : The SD of $\bar{\mathbf{d}}^{(1)}$ is performed based on the sequence $\mathcal{L}_2^*$, which is similar with that of GSD-SCMA. The flowchart of SGSD-SCMA is the same as in Fig. 4.2, except for $\mathcal{L}_2$ substituted with $\mathcal{L}_2^*$ and the condition $t \leqslant M^{J'}$ replaced by $t \leqslant L_1$ where L_1 is the size of $\mathcal{L}_2^*$ with $L_1 < M^{J'}$.

Therefore, the decoding problem of SGSD-SCMA becomes

$$\hat{\bar{\mathbf{d}}} = \arg \min_{\bar{\mathbf{d}}^{(2)} \in \mathcal{L}_2^*} \min_{\bar{\mathbf{d}}^{(1)\mathrm{p}} \in \bar{\mathcal{D}}^{(1)\mathrm{p}}} \|\bar{\mathbf{y}}' - \mathbf{R}_1 \bar{\mathbf{d}}^{(1)\mathrm{p}}\|^2 \tag{4.65}$$

$$\text{subject to } \|\bar{\mathbf{y}}' - \mathbf{R}_1 \bar{\mathbf{d}}^{(1)\mathrm{p}}\|^2 \leqslant d^2 \tag{4.66}$$

where $\bar{\mathbf{y}}' = \mathbf{Q}_1^{\intercal} \bar{\mathbf{y}} - \mathbf{R}_2 \bar{\mathbf{d}}^{(2)}$.

In the following, we propose two pruning algorithms, namely PRUN1 and PRUN2, for SGSD-SCMA. PRUN1 is a general algorithm applicable to any kind of SCMA codebooks, while PRUN2 algorithm requires a constraint on the codebook factor graph to work. In the sequel, SGSD-SCMA detector with the PRUN1 algorithm is referred as SGSD1-SCMA detector and the SGSD-SCMA detector with the PRUN2 algorithm is named SGSD2-SCMA detector. Besides, a theoretical analysis of the codeword error rate (CER) of the SGSD-SCMA is presented.

4.5.1 Proposed PRUN1 algorithm

As SCMA is a multiple access transmission, a similar pruning process explained in section 4.4 can be employed in every dimension of the received signal $\bar{\mathbf{y}}$. Recalling that matrix $\mathbf{R}_1$ has positive

diagonal, the pruning algorithm for SGSD-SCMA builds the sequence $\mathcal{L}_2^*$ given as

$$\bar{\mathbf{d}}^{(2)} \in \bar{\mathcal{D}}^{(2)} \tag{4.67}$$

subject to

$$(r_1)_{i,i} \min\left((\bar{\mathcal{D}}_{j_1}^{(1)\mathtt{p}})^{i-2d_v(j_1-1)}\right) - \delta \leqslant |\bar{y}_i - (\mathbf{r}_2)_i \bar{\mathbf{d}}^{(2)}| \leqslant (r_1)_{i,i} \max\left((\bar{\mathcal{D}}_{j_1}^{(1)\mathtt{p}})^{i-2d_v(j_1-1)}\right) + \delta, \tag{4.68}$$

$$\forall i \in \{1, \cdots, 2K\}$$

where $j_1 = \lceil \frac{i}{2d_v} \rceil$, $(r_1)_{i,i}$ denotes the i-th diagonal element of square matrix $\mathbf{R}_1$, $(\bar{\mathcal{D}}_j^{(1)\mathtt{p}})^k$ is the real sequence consisting of the k-th entry of all the elements of sequence $\bar{\mathcal{D}}_j^{(1)\mathtt{p}}$ and $(\mathbf{r}_2)_i$ denotes the i-th row vector of matrix $\mathbf{R}_2$. As explained in section 4.4, δ is determined by the additive noise and predefined error detection probability p. In this case, δ is calculated by

$$\delta = \mathcal{F}^{-1}(1-p, 0, \sigma) \tag{4.69}$$

Before diving into the details of the PRUN1 algorithm, we want to illustrate the sparse nature of matrix $\mathbf{R}_2$ which helps to simplify the algorithm. Because $\mathbf{Q}$ is diagonal and $\bar{\mathbf{G}}^{(2)}$ is sparse due to SCMA property, $\mathbf{R}_2$ is also sparse according to Eq.(4.34). For instance, when the codebook factor graph matrix is Eq.(2.5), the matrix $\mathbf{R}_2$ has the following structure

$$\mathbf{R}_2 = \begin{bmatrix}
0\ 0\ 0\ 0\ \mathtt{x}\ \mathtt{x}\ 0\ 0\ \mathtt{x}\ \mathtt{x}\ 0\ 0\ 0\ 0\ 0\ 0 \\
0\ 0\ 0\ 0\ \mathtt{x}\ \mathtt{x}\ 0\ 0\ \mathtt{x}\ \mathtt{x}\ 0\ 0\ 0\ 0\ 0\ 0 \\
\mathtt{x}\ \mathtt{x}\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ \mathtt{x}\ \mathtt{x}\ 0\ 0 \\
\mathtt{x}\ \mathtt{x}\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ \mathtt{x}\ \mathtt{x}\ 0\ 0 \\
0\ 0\ 0\ 0\ 0\ 0\ \mathtt{x}\ \mathtt{x}\ 0\ 0\ 0\ 0\ 0\ 0\ \mathtt{x}\ \mathtt{x} \\
0\ 0\ 0\ 0\ 0\ 0\ \mathtt{x}\ \mathtt{x}\ 0\ 0\ 0\ 0\ 0\ 0\ \mathtt{x}\ \mathtt{x} \\
0\ 0\ \mathtt{x}\ \mathtt{x}\ 0\ 0\ 0\ 0\ 0\ 0\ \mathtt{x}\ \mathtt{x}\ 0\ 0\ 0\ 0 \\
0\ 0\ \mathtt{x}\ \mathtt{x}\ 0\ 0\ 0\ 0\ 0\ 0\ \mathtt{x}\ \mathtt{x}\ 0\ 0\ 0\ 0
\end{bmatrix} \tag{4.70}$$

where $\mathtt{x}$ represents a non-zero element in the matrix. Thanks to the sparsity of $\mathbf{R}_2$ illustrated above, for a given row, we only consider the elements of $\bar{\mathbf{d}}^{(2)}$ that are associated with the non-zero element of $\mathbf{R}_2$. There are always $2(d_f - 1)$ non-zero elements of in each row vector $(\mathbf{r}_2)_i$ ($i \in \{1, \cdots, 2K\}$) which correspond to the same user in pairs. Therefore, instead of testing every value of $\bar{\mathbf{d}}^{(2)}$ ($\bar{\mathbf{d}}^{(2)} \in \bar{\mathcal{D}}^{(2)}$), we can test M^{d_f-1} combinations of those elements of $\bar{\mathbf{d}}^{(2)}$ that are associated with the non-zero element of $\mathbf{R}_2$ at each row. Moreover, another property of matrix $\mathbf{R}_2$ is that the $(2k-1)$-th and $2k$-th ($1 \leqslant k \leqslant K$) rows of $\mathbf{R}_2$ have the same indices of non-zero elements. Thus, we perform the pruning operation of two rows, the $(2k-1)$-th and $2k$-th, at once. Accordingly, there will be K pruning operations for the PRUN1 algorithm in total.

As $\bar{\mathbf{x}}^{(2)}$ contains the real-valued codewords of J' users, we define the vector $\mathbf{v} \in \{1, \cdots, M\}^{J' \times 1}$ to efficiently represent the values of vector $\bar{\mathbf{x}}^{(2)}$. Every entry of $\mathbf{v}$ is the codeword index of the corresponding user. To simplify the process, PRUN1 algorithm outputs the reduced sequence Υ of values of vector $\mathbf{v}$. The sequence $\mathcal{L}_2^*$ can be constructed based on the sequence Υ after the PRUN1 algorithm and the size of the two sequences are the same. The PRUN1 algorithm is summarized in Algorithm 9.

Algorithm 9 The PRUN1 algorithm.

Input : $\bar{\mathbf{y}}$, $\mathbf{R}_1$, $\mathbf{R}_2$, $\mathcal{D}^{(1)}$ and $\mathcal{D}^{(2)}$
Output : Υ
Initialization : $\mathcal{V} = \emptyset$, $\Upsilon = (\mathbf{0}_{J' \times 1})$ and $L_1 = 1$.
for $k = K, K-1, \cdots, 1$ **do**
  $\mathcal{U}, \mathcal{W}, \mathcal{A}_1, \mathcal{A}_2, \mathcal{B}_1, \mathcal{B}_2 = \text{UserPart}((\mathbf{r}_2)_{2k}, \mathcal{V})$
  $l_1 = \text{size}(\mathcal{U})$
  $l_2 = d_f - 1 - l_1$
  if $\mathcal{W} \neq \emptyset$ **then**
    for $i_1 = 1, 2, \cdots, L_1$ **do**
      Obtain $\mathbf{q}$ using Eq.(4.71).
      Calculate $t_1(i_1)$ and $t_2(i_1)$ based on Eq.(4.73) and Eq.(4.72), respectively.
    end for
  else
    $t_1(1) = t_2(1) = 0$
  end if
  $\Upsilon' = \emptyset$
  $\mathbf{P} = \mathcal{F}_1(M, l_1)$
  $c_1 = 0$
  for $i_2 = 1, 2, \cdots, M^{l_1}$ **do**
    Calculate $s_1(i_2)$ and $s_2(i_2)$ using Eq.(4.75) and Eq.(4.76), respectively.
    for $i_1 = 1, 2, \cdots, L_1$ **do**
      Calculate ξ_1 and ξ_2 using Eq.(4.80) and Eq.(4.81), respectively.
      if Eq.(4.77) and Eq.(4.78) are satisfied **then**
        $\mathbf{v} = \Upsilon(i_1)$
        for $i = 1, 2, \cdots, l_1$ **do**
          $v_{\mathcal{B}_1(i)} = (p_{i_2})_i$
        end for
        $c_1 = c_1 + 1$
        $\Upsilon'_i(c_1) = \mathbf{v}$
      end if
    end for
  end for
  $\Upsilon = \Upsilon'_i$
  $L_1 = c_1$,
  $\mathcal{V} = \mathcal{V} \bigcup \mathcal{U}$
end for

Function : $\mathcal{U}$, $\mathcal{W}$, $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{B}_1$, $\mathcal{B}_2$ =UserPart $((\mathbf{r}_2)_{2k}, \mathcal{V})$
$\mathcal{C} = \emptyset$
$i = 0$
for $j = 1, \cdots, 2Jdv - 2K$ **do**
  if $(r_2)_{2k,j} = 1$ **then**
    $i = i + 1$
    $\mathcal{A}(i) = j$
    $\mathcal{B}(i) = \lceil \frac{\mathcal{A}(i)}{2d_v} \rceil$
    $\mathcal{C} = \mathcal{C} \cup \{\mathcal{B}(i)\}$
  end if
end for
$\mathcal{U} = \mathcal{C} \setminus \mathcal{V}$
$\mathcal{W} = \mathcal{C} \setminus \mathcal{U}$
$i_1 = i_2 = 0$
for $i = 1, 2, \cdots, 2(d_f - 1)$ **do**
  if $\mathcal{B}(i) \in \mathcal{U}$ **then**
    $i_1 = i_1 + 1$
    $\mathcal{B}_1(i_1) = \mathcal{B}(i)$
    $\mathcal{A}_1(i_1) = \mathcal{A}(i)$
  else
    $i_2 = i_2 + 1$
    $\mathcal{B}_2(i_2) = \mathcal{B}(i)$
    $\mathcal{A}_2(i_2) = \mathcal{A}(i)$
  end if
end for

In the following, we elaborate on the PRUN1 algorithm. At the beginning of the PRUN1 algorithm, the set $\mathcal{V}$, which stores the indices of users that have been visited, is empty at the initialization. Sequence Υ is the output of PRUN1 algorithm, initialized as $\Upsilon = (\mathbf{0}_{J' \times 1})$ for the vector operations in the following of the algorithm. Accordingly, its size is initialized as $L_1 = 1$. Similarly to the tree search, the PRUN1 algorithm starts by performing pruning from the last two rows of $\mathbf{R}_2$ and moves upwards. At each pruning operation, $d_f - 1$ users associated with the non-zero elements are visited. These users can be divided into the revisited users and newly visited users. For simplicity, the newly visited users are referred to as new users in sequel. At the k-th ($1 \leqslant k \leqslant K$) pruning operation, function UserPart$((\mathbf{r}_2)_{2k}, \mathcal{V})$ outputs the partitioned users and their corresponding indices. Specifically, the sequence $\mathcal{A}$ stores the indices of non-zero elements of $(\mathbf{r}_2)_{2k}$ and the sequence $\mathcal{B}$ stores the indices of users associated with the elements in sequence $\mathcal{A}$. We define the set $\mathcal{C}$ to store the currently visited users. Thus, the set $\mathcal{C}$ consists of the $d_f - 1$ distinct elements in sequence $\mathcal{B}$. Sets $\mathcal{U}$ and $\mathcal{W}$ store the indices of new users and revisited users, respectively. The sequence $\mathcal{A}_1$ stores the indices in the

sequence $\mathcal{A}$ that are associated with the new users and the sequence $\mathcal{A}_2$ stores the indices in $\mathcal{A}$ that associated with the revisited users. Besides, sequences $\mathcal{B}_1$ and $\mathcal{B}_2$ contain the indices of users associated to the elements in $\mathcal{A}_1$ and $\mathcal{A}_2$, respectively. Afterwards, the number of new users l_1 and the number of revisited users l_2 are obtained.

For the sake of reducing computational complexity, the PRUN1 algorithm performs the computations concerning the two kinds of users separately. Because PRUN1 carries out the K pruning operations in sequence, the indices combinations of the revisited users are those stored in the sequence Υ. At the k-th ($1 \leqslant k \leqslant K$) pruning operation, if the set $\mathcal{W}$ is not empty, for the i_1-th ($1 \leqslant i_1 \leqslant L_1$) element in sequence Υ, the codeword indices of the revisited users constitute the vector $\mathbf{q}$, given as

$$\mathbf{q} = \Upsilon^{\mathcal{W}}(i_1) \tag{4.71}$$

where $\Upsilon^{\mathcal{W}}(i_1)$ denotes the entries with indices in set $\mathcal{W}$ of the i_1-th element of sequence Υ. Then, we calculate two intermediates regarding the revisited users, $t_1(i_1)$ and $t_2(i_1)$, as

$$t_1(i_1) = \sum_{l=1}^{2l_2} (r_2)_{2k-1,\mathcal{A}_2(l)} \bar{\mathcal{D}}^{\mathcal{A}_2(l)-2d_v(\mathcal{B}_2(l)-1)}_{\frac{K}{d_v}+\mathcal{B}_2(l)}(q_l) \tag{4.72}$$

and

$$t_2(i_1) = \sum_{l=1}^{2l_2} (r_2)_{2k,\mathcal{A}_2(l)} \bar{\mathcal{D}}^{\mathcal{A}_2(l)-2d_v(\mathcal{B}_2(l)-1)}_{\frac{K}{d_v}+\mathcal{B}_2(l)}(q_l) \tag{4.73}$$

respectively, where $(r_2)_{i,l}$ denotes the element in the i-th row and the l-th column of matrix $\mathbf{R}_2$, $\bar{\mathcal{D}}^a_j(b)$ refers to the a-th entry of the b-th element of the sequence $\bar{\mathcal{D}}_j$, which is a scalar. For the new users, all M^{l_1} combinations of the associated codewords are tested. To enable the codeword enumeration of l_1 new users, we utilize the function $\mathcal{F}_1(M, l_1)$ to generate the matrix $\mathbf{P} = [\mathbf{p}_1, \cdots, \mathbf{p}_{M^{l_1}}] \in \{1, 2, \cdots, M\}^{l_1 \times M^{l_1}}$. If $M = 4$ and $l_1 = 2$, the matrix $\mathbf{P}$ is written as

$$\begin{aligned} \mathbf{P} &= \mathcal{F}_1(4, 2) \\ &= \begin{bmatrix} 1 & 1 & 1 & 1 & \cdots & 4 & 4 & 4 & 4 \\ 1 & 2 & 3 & 4 & \cdots & 1 & 2 & 3 & 4 \end{bmatrix} \end{aligned} \tag{4.74}$$

Each column of $\mathbf{P}$ is a codeword indices combination. For the i_2-th ($1 \leqslant i_2 \leqslant M^{l_1}$) combination, $\mathbf{p}_{i_2}$ is used to calculate two intermediates, $s_1(i_2)$ and $s_2(i_2)$, as

$$s_1(i_2) = \bar{y}_{2k-1} - \sum_{l=1}^{2l_1} (r_2)_{2k-1,\mathcal{A}_1(l)} \bar{\mathcal{D}}^{\mathcal{A}_1(l)-2d_v(\mathcal{B}_1(l)-1)}_{\frac{K}{d_v}+\mathcal{B}_1(l)}((p_{i_2})_{\lceil \frac{l}{2} \rceil}) \tag{4.75}$$

and

$$s_2(i_2) = \bar{y}_{2k} - \sum_{l=1}^{2l_1} (r_2)_{2k,\mathcal{A}_1(l)} \bar{\mathcal{D}}^{\mathcal{A}_1(l)-2d_v(\mathcal{B}_1(l)-1)}_{\frac{K}{d_v}+\mathcal{B}_1(l)}((p_{i_2})_{\lceil \frac{l}{2} \rceil}) \tag{4.76}$$

respectively. Based on Eq.(4.67)-(4.68), at the k-th pruning operation, for a given i_1-th element in sequence Υ, the pruning criterion keeps the vectors $\mathbf{p}_{i_2}$ $(1 \leqslant i_2 \leqslant M^{l_1})$ that jointly satisfy the conditions

$$(r_1)_{2k-1,2k-1} \min\left((\bar{\mathcal{D}}_j^{(1)\mathtt{p}})^{2k-1-2d_v(j-1)}\right) - \delta \leqslant \xi_1 \leqslant (r_1)_{2k-1,2k-1} \max\left((\bar{\mathcal{D}}_j^{(1)\mathtt{p}})^{2k-1-2d_v(j-1)}\right) + \delta \tag{4.77}$$

and

$$(r_1)_{2k,2k} \min\left((\bar{\mathcal{D}}_j^{(1)\mathtt{p}})^{2k-2d_v(j-1)}\right) - \delta \leqslant \xi_2 \leqslant (r_1)_{2k,2k} \max\left((\bar{\mathcal{D}}_j^{(1)\mathtt{p}})^{2k-2d_v(j-1)}\right) + \delta \tag{4.78}$$

where

$$j = \lceil \frac{k}{d_v} \rceil \tag{4.79}$$

$$\xi_1 = s_1(i_2) - t_1(i_1) \tag{4.80}$$

$$\xi_2 = s_2(i_2) - t_2(i_1) \tag{4.81}$$

and δ is calculated according to Eq.(4.69). If vectors $\mathbf{p}_{i_2}$ and $\Upsilon(i_1)$ satisfy the conditions Eq.(4.77)-(4.78) jointly, they are used to construct the elements in sequence Υ.

We provide a performance analysis of SGSD-SCMA detector with the PRUN1 algorithm considering the error detection probability p as a factor. Similar to Eq.(4.61)-(4.63), CER of SGSD1-SCMA can be approximated to

$$Pr\{\hat{\bar{\mathbf{d}}} \neq \bar{\mathbf{d}}\} \approx P_{\mathrm{ML}} + Pr\{\bar{\mathbf{d}}^{(2)} \notin \mathcal{L}_2^*\} \tag{4.82}$$

where P_{ML} is the CER of the ML detector Eq.(4.2) and $Pr\{\bar{\mathbf{d}}^{(2)} \notin \mathcal{L}_2^*\}$ is the error rate of the PRUN1 algorithm. At high SNR values, since P_{ML} approaches to zero, CER can be written as

$$\lim_{\mathrm{SNR}\to\infty} Pr\{\hat{\bar{\mathbf{d}}} \neq \bar{\mathbf{d}}_{\mathrm{T}}\} = Pr\{\bar{\mathbf{d}}_{\mathrm{T}}^{(2)} \notin \mathcal{L}_2^*\} \tag{4.83}$$

At the k-th pruning operation, the error rate is computed as

$$p_k = \frac{2p}{M_1}\left(1 - \frac{2p}{M_1}\right)^{k-1} \tag{4.84}$$

where M_1 is the average number of distinct projections in a real dimension of the SCMA codebook. In this work, M_1 is approximated to $\sqrt{M}$. Because there are always K sequential pruning operations, the error rate of PRUN1 algorithm is calculated by

$$Pr\{\bar{\mathbf{x}}_{\mathrm{T}}^{(2)} \notin \mathcal{L}_2^*\}_1 = \sum_{k=1}^{K} p_k \tag{4.85}$$

$$= \sum_{k=1}^{K} \frac{2p}{M_1}\left(1 - \frac{2p}{M_1}\right)^{k-1} \tag{4.86}$$

$$= 1 - \left(1 - \frac{2p}{M_1}\right)^{K} \tag{4.87}$$

4.5.2 Proposed PRUN2 algorithm

In this subsection, we introduce another pruning algorithm called PRUN2 which simplifies the process by considering certain factor graph constraint of the codebook. The PRUN2 algorithm can be regarded as a simplified version of the PRUN1 algorithm.

A major difference between the two pruning algorithms for SCMA is that there is no revisited user at a given pruning operation in PRUN2 algorithm. To satisfy this condition, at any pruning operation, the set $\mathcal{C}$ containing the currently visiting users and the set $\mathcal{V}$ storing the users that have been visited must be disjoint. Besides, the PRUN2 algorithm stops the pruning operation once all J' users associated with $\mathbf{R}_2$ have been visited. Therefore, the PRUN2 algorithm only carries out $K' = \frac{J_i'}{d_f - 1} = \frac{K}{d_v}$ $(K' < K)$ pruning operations. If we can rewrite the factor graph matrix $\mathbf{F}$ as

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}^{(1)} & \mathbf{F}^{(2)} \end{bmatrix} \tag{4.88}$$

where $\mathbf{F}^{(1)} = [\mathbf{f}_1, \cdots, \mathbf{f}_{\frac{K}{d_v}}]$ and $\mathbf{F}^{(2)} = [\mathbf{f}_{\frac{K}{d_v}+1}, \cdots, \mathbf{f}_J]$, the requirement of the PRUN2 algorithm can be met by the fact that the last K' rows of matrix $\mathbf{F}^{(2)}$ are mutually orthogonal. Defining matrix $\mathbf{F}_s^{(2)}$ containing the last K' rows of matrix $\mathbf{F}^{(2)}$, the constraint of the PRUN2 algorithm can be expressed as

$$\mathbf{F}_s^{(2)} \mathbf{F}_s^{(2)\mathsf{T}} = \mathbf{I}_{K'} \tag{4.89}$$

Algorithm 10 describes the PRUN2 algorithm and we will introduce the details in the following. At a pruning operation, the algorithm obtains sequences $\mathcal{A}$ and $\mathcal{B}$ and the set $\mathcal{C}$ as in the PRUN1 algorithm. In the PRUN2 algorithm, as the set $\mathcal{C}$ only contains new users, we need to test $M^{d_f - 1}$

codeword combinations of the $d_f - 1$ new users. Thus, we use function $\mathcal{F}_1$ mentioned previously to generate indices matrix $\mathbf{P}$ as

$$\mathbf{P} = \mathcal{F}_1(M, d_f - 1) \tag{4.90}$$

The pruning criterion retains the vectors $\mathbf{p}_i$ $(1 \leqslant i \leqslant M^{d_f-1})$ satisfying the condition in Eq.(4.77)-(4.78) jointly. As $\mathcal{C}$ has only new users, the intermediates ξ_1 and ξ_2 in the PRUN2 algorithm are computed by

$$\xi_1 = \bar{y}_{2k-1} - \sum_{l=1}^{2(d_f-1)} (r_2)_{2k-1,\mathcal{A}(l)} \bar{\mathcal{D}}^{\mathcal{A}(l)-2d_v(\mathcal{B}(l)-1)}_{\frac{K}{d_v}+\mathcal{B}(l)}((p_i)_{\lceil \frac{l}{2} \rceil}) \tag{4.91}$$

and

$$\xi_2 = \bar{y}_{2k} - \sum_{l=1}^{2(d_f-1)} (r_2)_{2k,\mathcal{A}(l)} \bar{\mathcal{D}}^{\mathcal{A}(l)-2d_v(\mathcal{B}(l)-1)}_{\frac{K}{d_v}+\mathcal{B}(l)}((p_i)_{\lceil \frac{l}{2} \rceil}) \tag{4.92}$$

respectively. The qualified vectors $\mathbf{p}_i$ are used to create the elements of sequence

The CER of SGSD2-SCMA has the same expression as shown in Eq.(4.82) expect the error rate of the PRUN2 algorithm is computed as

$$Pr\{\bar{\mathbf{x}}_{\mathrm{T}}^{(2)} \notin \mathcal{L}_2^*\}_2 = \sum_{k'=1}^{K'} \frac{2p}{M_1}\left(1 - \frac{2p}{M_1}\right)^{k'-1} \tag{4.93}$$

$$= 1 - \left(1 - \frac{2p}{M_1}\right)^{K'} \tag{4.94}$$

Compared with the PRUN1 algorithm carrying out K pruning operations, the PRUN2 algorithm performs only K' pruning operations. For a codebook compatible with both algorithms, some values of $\mathbf{v}$ satisfying the conditions in the PRUN2 algorithm might be eliminated by the PRUN1 algorithm. As a result, the PRUN1 algorithm leads to a smaller size of $\mathcal{L}_2^*$ than the PRUN2 algorithm reducing the complexity of SGSD-SCMA but has higher error rate, $Pr\{\bar{\mathbf{d}}^{(2)} \notin \mathcal{L}_2^*\}_1 > Pr\{\bar{\mathbf{d}}^{(2)} \notin \mathcal{L}_2^*\}_2$. This property of the PRUN1 algorithm will be verified by the simulation performance in Eq.(5.4).

4.6 Complexity analysis

In this section, we study the computational complexity of different detectors, categorized to iterative detectors and SD-based detectors. For the first category, only the conventional MPA is considered. Three SD-based detectors are considered in the second category, which are SD-SCMA, GSD-SCMA

and SGSD-SCMA. The computational complexity is investigated in terms of floating point operations (FLOPs).

Algorithm 10 The PRUN2 algorithm.

Input : $\bar{\mathbf{y}}$, $\mathbf{R}_1$, $\mathbf{R}_2$, $\mathcal{D}^{(1)}$ and $\mathcal{D}^{(2)}$
Output : Υ
Initialization : $\mathcal{V} = \emptyset$, $\Upsilon = (\mathbf{0}_{J' \times 1})$ and $L_1 = 1$.
for $k = K, K-1, \cdots, K-K'+1$ **do**
  $\mathcal{C} = \emptyset$
  $i = 0$
  for $j = 1, \cdots, 2Jdv - 2K$ **do**
    if $(r_2)_{2k,j} = 1$ **then**
      $i = i + 1$
      $\mathcal{A}(i) = j$
      $\mathcal{B}(i) = \lceil \frac{\mathcal{A}(i)}{2d_v} \rceil$
      $\mathcal{C} = \mathcal{C} \cup \{\mathcal{B}(i)\}$
    end if
  end for
  if $\mathcal{C} \cap \mathcal{V} = \emptyset$ **then**
    $c_1 = 0$
    $\Phi = \emptyset$
    $\mathbf{P} = \mathcal{F}_1(M, d_f - 1)$
    for $i = 1, \cdots, M^{d_f - 1}$ **do**
      Calculate ξ_1 and ξ_2 using Eq.(4.91) and Eq.(4.92), respectively.
      if Eq.(4.77) and Eq.(4.78) are satisfied **then**
        $\mathbf{v} = \mathbf{0}_{J' \times 1}$
        for $n = 1, \cdots, d_f - 1$ **do**
          $v_{\mathcal{B}(2n)} = (p_i)_n$
        end for
        $c_1 = c_1 + 1$
        $\Phi(c_1) = \mathbf{v}$
      end if
    end for
    $\Upsilon' = \Upsilon$
    for $i_1 = 1, 2, \cdots, L_1$ **do**
      for $i_2 = 1, 2, \cdots, c_1$ **do**
        $\mathbf{v} = \Upsilon'(i_1) + \Phi(i_2)$
        $\Upsilon(i_1(L_1 - 1) + i_2) = \mathbf{v}$
      end for
    end for
    $L_1 = L_1 c_1$
    $\mathcal{V} = \mathcal{V} \cup \mathcal{C}$
  end if
end for

4.6.1 Iterative detector

MPA can be regarded as a sub-optimal detector for SCMA systems. It follows the message-update rules and exchanges messages from REs (FNs) to users (VNs) and vice versa. The number of iterations of MPA not only impacts the MUD performance but also the computational complexity. A larger number of iterations improves the MUD performance, but brings about higher computational complexity. As concluded in [18], the number of FLOPs of MPA is $T\big((9d_f+7)M^{d_f}Kd_f + MJ{d_v}^2 - Kd_f + 5(\log_2 M - 1)MJ\big) + (d_v+1)MJ\log_2(M)$, where T ($T \geqslant 1$) is the number of iterations.

4.6.2 SD-based detectors

In contrary to MPA, the numbers of FLOPs of SD-based detectors are stochastic and depend on the transmitted signals, the additive noise and the squared radius. As the complexity of SD-based detectors is relevant to the number of visited nodes in the tree search which is unpredictable, we provide the FLOPs expression of different constituting components of the detectors. The number of FLOPs of these detector are obtained by the Monte-Carlo simulation and will be presented in Eq.(5.4).

4.6.2.1 Number of FLOPs of SD-SCMA

As demonstrated in Fig. 4.1, the computational complexity of the SD-SCMA detector comes from the SQRD of the channel matrix $\tilde{\mathbf{G}}$, the matrix multiplication $\mathbf{Q}^{\intercal}\tilde{\mathbf{y}}$, the tree search and the re-permutation at the output. The number of FLOPs of sqrd$(\tilde{\mathbf{G}})$ are $2(2Jd_v)^3 + 2Jd_v(6Jd_v-1) - J$. The matrix multiplication needs $2Jd_v(4Jd_v-1)$ FLOPs. For the tree search, we list the intermediate calculation complexity in Table 4.1. For the re-permutation process, it requires $2Jd_v(4Jd_v-1)$ FLOPs.

Table 4.1 – The number of FLOPs of the calculations in tree search for SD-SCMA.

Variables	u_i	l_i	$\mathbf{c}_i$	p_i
FLOPs	$2(2Jd_v - i) + 4$	$2(2Jd_v - i) + 4$	$s_i\big[\big(4(2Jd_v - i) + 4d_v + 6\big) - 1\big]$	1

4.6.2.2 Number of FLOPs of GSD-SCMA

For GSD-SCMA, the complexity of the calculation in Eq.(4.35), the computation of $\bar{\mathbf{y}}' = \mathbf{Q}_1^{\intercal}\bar{\mathbf{y}} - \mathbf{R}_2\bar{\mathbf{d}}^{(2)}$, the tree search process and the re-permutation of the output constitute its computational complexity. The computation of sqrd$(\bar{\mathbf{G}}_1)$ needs $16K^3 + 2K(6K-1) - J'$ FLOPs. Considering the

complexity of square matrix inversion is $\mathcal{O}(n^3)$, the calculation of matrix $\mathbf{R}_2$ takes $(2K)^3 + 2K(4K-1)(2Jd_v - 2K)$ FLOPs. Consequently, the number of FLOPs of the calculation of Eq.(4.35) is $(2K)^3 + 2K(8K-1)(2Jd_v+1) - J'$. To avoid repeated computation and decrease complexity, $\mathbf{Q}_1^\intercal \bar{\mathbf{y}}$ is calculated once and stored for reuse. The number of FLOPs of $\mathbf{Q}_1^\intercal \bar{\mathbf{y}}$ are $2K(4K-1)$. For a given value of $\bar{\mathbf{d}}^{(2)}$, the computation of $\bar{\mathbf{y}}'$ needs $2K(4Jd_v - 4K)$ FLOPs. Since all $M^{J'}$ values of $\bar{\mathbf{x}}^{(2)}$ are considered in GSD-SCMA, computation of $\bar{\mathbf{y}}' = \mathbf{Q}_1^\intercal \bar{\mathbf{y}} - \mathbf{R}_2 \bar{\mathbf{d}}^{(2)}$ needs $2K\big[M^{J'}(4Jd_v - 4K) + (4K-1)\big]$ FLOPs in total. Because the matrix $\mathbf{R}_1$ is diagonal, the calculation of u_i and l_i is simplified. Therefore, the number of FLOPs of the intermediate calculations in tree search process for GSD-SCMA is shown in Table 4.2. For the re-permutation $\mathbf{P}_1 \bar{\mathbf{d}}^{(1)\mathtt{p}}$, the complexity is $2K(4K-1)$ FLOPs.

TABLE 4.2 – The number of FLOPs of the calculations in tree search for GSD-SCMA.

Variables	u_i	l_i	$\mathbf{c}_i$
FLOPs	4	4	$s_i(8d_v - 1)$

4.6.2.3 Number of FLOPs of SGSD-SCMA

The components of SGSD-SCMA's complexity are the same as those of the GSD-SCMA detector. For the SGSD-SCMA detector, the complexity of the calculation in Eq.(4.35) is the same as that in GSD-SCMA detector. The complexity of $\bar{\mathbf{y}}' = \mathbf{Q}_1^\intercal \bar{\mathbf{y}} - \mathbf{R}_2 \bar{\mathbf{d}}^{(2)}$ is affected by the size of $\mathcal{L}_2^*$, which is $2K\big[L_1(4Jd_v - 4K) + (4K-1)\big]$ FLOPs. The numbers of FLOPs required for the intermediate calculations of the PRUN1 and PRUN2 algorithms are listed in Table 4.3 and Table 4.4, respectively. Because the intermediates computations in the tree search for the SGSD-SCMA detector are the same as those for the GSD-SCMA detector, their complexity are given in Table 4.2.

TABLE 4.3 – The number of FLOPs of the PRUN1 algorithm.

Variables	USERPART $((\mathbf{r}_2)_{2k}, \mathcal{V})$	l_2	$t_1(i_1)(t_2(i_1))$	$s_1(i_1)(s_2(i_1))$	j	$\xi_1(\xi_2)$
FLOPs	$6(d_f - 1)$	2	$4l_2 - 1$	$2l_1$	1	1

TABLE 4.4 – The number of FLOPs of the PRUN2 algorithm.

Variables	$\mathcal{B}$	$\xi_1(\xi_2)$
FLOPs	$4(d_f - 1)$	$4(d_f - 1)$

TABLE 4.5 – Simulation and theoretical error rate comparison for the first codebook.

p	Algorithm	Simulation	Theoretical
5×10^{-4}	PRUN1	1.5×10^{-3}	2.0×10^{-3}
	PRUN2	1.1×10^{-3}	1.0×10^{-3}
1×10^{-4}	PRUN1	3.0×10^{-4}	4.0×10^{-4}
	PRUN2	2.0×10^{-4}	2.0×10^{-4}

4.7 Numerical results

This section investigates the performance of different SCMA detectors in terms of CER, BER and the computational complexity. The performance of the proposed SGSD-SCMA is compared with MPA and with SD-based state-of-the-art, i.e. SD-SCMA and GSD-SCMA. In the simulation, the squared radius is initialized as $d^2 = 50$. To better understand the capability of various detectors, three settings of SCMA system parameters are considered in the simulation. First, typical benchmark setting $J = 6$, $K = 4$ and $M = 4$ [113] is considered. The second setting has also typical factor graph $J = 6$ and $K = 4$ but has codebook size $M = 16$ [40]. The last one [107] has $J = 15$, $K = 6$ and $M = 4$, which results in an overloading factor of 250%. The first two settings meet the constraint of the PRUN2 algorithm and are compatible with both two proposed pruning algorithms. As the factor graph of the third setting does not satisfy the constraint in Eq.(4.89), only the PRUN1 algorithm is applicable to it. As the compatibility of investigated algorithms is different, the following results are organized by the setting of parameters.

4.7.1 Performance of the first codebook with $J = 6$, $K = 4$ and $M = 4$

We first study the CER performance of the codebook [113]. Fig. 4.5 shows the CER performance of various detectors. The ML performance is given by the GSD-SCMA detector. It is obvious that, the simulated CER follows the limit demonstrated in Eq.(4.63). Table 4.5 compares the simulation and theoretical error rate of codebook [113]. Apart from the case of the PRUN1 algorithm with $p = 1 \times 10^{-4}$, the rest of the simulation performance fits the theoretical expressions Eq.(4.87)-(4.94) for the proposed pruning algorithms.

Fig. 4.6 compares the BER performance of different detectors when using codebook [113]. All the studied detectors are compatible with this codebook. The GSD-SCMA and MPA with 5 iterations has almost the same BER performance. It is obvious that for a given pruning algorithm, the lower p is, the

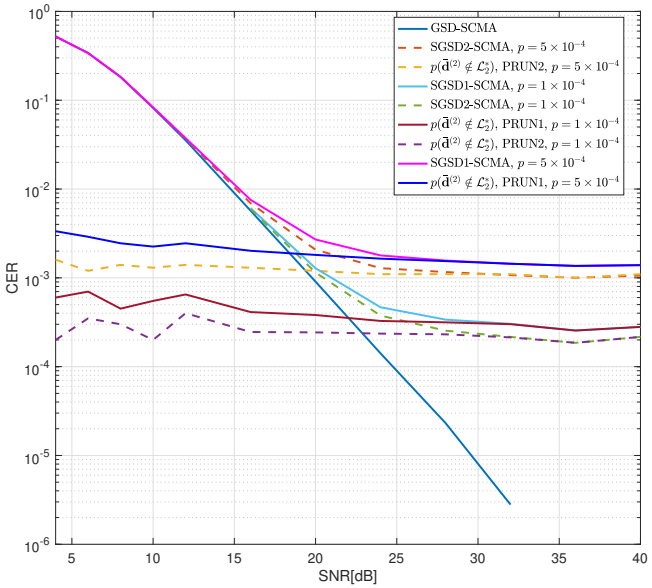


FIGURE 4.5 – CER performance of different detectors for the first codebook.

better BER performance SGSD-SCMA has. For a given value of p, SGSD2-SCMA has slightly better performance than SGSD1-SCMA. The proposed SGSD-SCMA has performance gap with GSD-SCMA and MPA when SNR $\geqslant$ 12dB for $p = 5 \times 10^{-4}$ and SNR$\geqslant$ 16dB for $p = 1 \times 10^{-4}$ respectively.

Lastly, we evaluate the computational complexity. Fig. 4.7 illustrates the FLOPs of MPA with 5 iterations and GSD-SCMA. Please note that FLOPs of the MPA is identical for all SNRs. We can see that GSD-SCMA has lower complexity than MPA with 5 iterations. However, there is no obvious decrease of GSD-SCMA FLOPs with the increase of SNR values. Fig. 4.8 compares the FLOPs evolution of different steps in SGSD-SCMA with different values of p. In contrary to GSD-SCMA, SGSD-SCMA shows significant decrement of FLOPs with the increment of SNR in Fig. 4.8. The FLOPs of SGSD-SCMA are inversely proportional to the error detection probability p at low to moderate SNR values. In high SNR regime, the number of FLOPs of SGSD-SCMA with different p values converge to the same level. The computational complexity of SGSD-SCMA consists of the complexity of the pruning algorithm, the step1, and that of the SD for $\bar{\mathbf{x}}^{(1)}$, the step2. Fig. 4.8 also compares the complexity of the two steps within SGSD-SCMA. The number of FLOPs of the PRUN1 algorithm is greater than that of the PRUN2 algorithm. However, SGSD1-SCMA has a significantly lower overall number of FLOPs in comparison to SGSD2-SCMA because the PRUN1 algorithm outputs a smaller size of $\mathcal{L}_2^*$ as shown in Table 4.6. It implies that the size of $\mathcal{L}_2^*$ plays a significant role in the overall complexity of SGSD-SCMA.

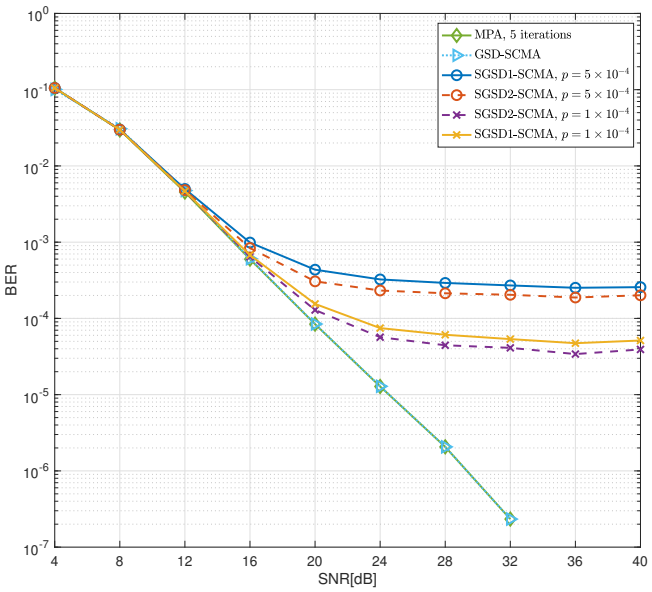


FIGURE 4.6 – BER performance of different detectors for the first codebook.

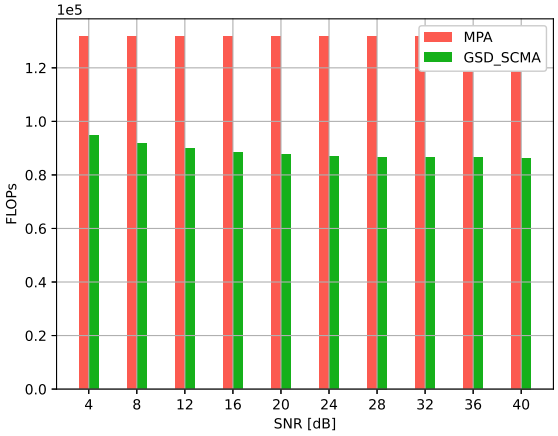


FIGURE 4.7 – FLOPs of MPA and GSD-SCMA for the first codebook.

For the codebook given in [113], when 4dB $\leqslant$ SNR $\leqslant$ 12dB, SGSD-SCMA with $p = 5 \times 10^{-4}$ has the same BER performance and much lower complexity than MPA and GSD-SCMA. When 4dB $\leqslant$ SNR $\leqslant$ 16dB, SGSD-SCMA with $p = 1 \times 10^{-5}$ is more efficient than MPA and GSD-SCMA in terms of complexity. Overall, SGSD-SCMA has a better trade-off than MPA and GSD-SCMA at low to moderate SNR values for the codebook [113].

4.7.2 Performance of the second codebook with $J = 6$, $K = 4$, $M = 16$

Fig. 4.9 shows the simulation CER performance of the codebook [40]. Table 4.7 compares the simulation and theoretical error rates of the proposed pruning algorithms regarding codebook [40].

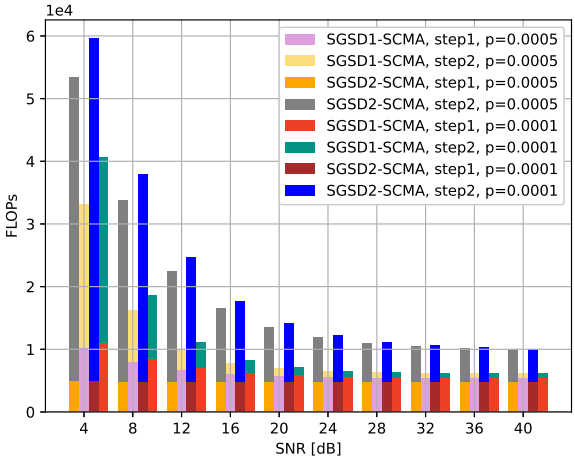


FIGURE 4.8 – FLOPs of two steps in SGSD-SCMA for the first codebook.

TABLE 4.6 – Size of sequence $\mathcal{L}_2^*$ of the pruning algorithms for the first codebook.

p	Algorithm	4dB	8dB	12dB	16dB	20dB	24dB	28dB	32dB	36dB	40dB
5×10^{-4}	PRUN1	63	23	9	4	3	2	2	2	2	2
	PRUN2	130	79	48	32	24	19	17	16	15	14
1×10^{-4}	PRUN1	82	31	11	5	3	2	2	2	2	2
	PRUN2	148	91	55	35	25	20	17	16	15	14

There is a mismatch of the PRUN2 algorithm between the simulation and theoretical performance. The mismatch might be brought by the value of $M_1 = \sqrt{M}$ because it is a coarse approximation derived from a classical modulation. In fact, the number of projections of SCMA multidimensional codebook in each real dimension is not necessarily identical.

Because SD-SCMA is not compatible with codebook [40], Fig. 4.10 compares BER performance of the MPA and SGSD-SCMA. SGSD-SCMA with $p = 5 \times 10^{-3}$ outperforms MPA at the range of 14dB $\leqslant$ SNR $\leqslant$ 22dB and SGSD-SCMA with $p = 5 \times 10^{-4}$ outperforms MPA at the range of 14dB $\leqslant$ SNR $\leqslant$ 24dB.

Fig. 4.11 compares the number of FLOPs of two steps in SGSD-SCMA with different p values. For MPA with 5 iterations, the number of FLOPs is 8.37×10^6. When SNR = 14dB, the number of FLOPs of SGSD2-SCMA is about 70.60% of those of MPA with 5 iterations. For a given value of p, the number of FLOPs of SGSD1-SCMA is always lower than one fifth of that of SGSD2-SCMA. The PRUN1 algorithm shows further substantial complexity advantage over the PRUN2 algorithm for this codebook $M = 16$ [40] than for the first codebook $M = 4$ [113].The reason is that the complexity of

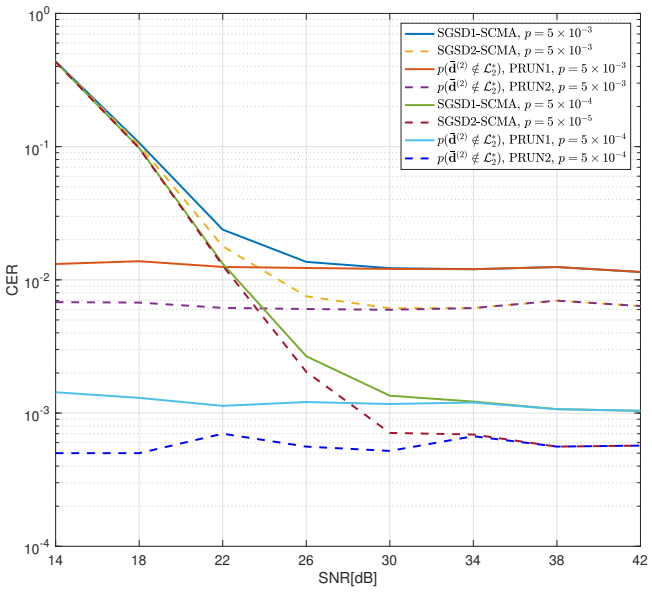


FIGURE 4.9 – CER performance comparison of different detectors for the second codebook.

TABLE 4.7 – Simulation and theoretical error rate comparison of the second codebook.

p	Algorithm	Simulation	Theoretical
5×10^{-3}	PRUN1	1.1×10^{-2}	1.0×10^{-2}
	PRUN2	6.0×10^{-3}	5.0×10^{-3}
5×10^{-4}	PRUN1	1.1×10^{-3}	1.0×10^{-3}
	PRUN2	6.0×10^{-4}	5.0×10^{-4}

tree search, the step2 of SGSD-SCMA, is sensitive to the codebook size.

4.7.3 Performance of the third codebook with $J = 15$, $K = 6$, $M = 4$

The codebook $J = 15$, $K = 6$, $M = 4$ [107] is only compatible with MPA and SGSD1-SCMA. Fig. 4.12 illustrates the CER of the SGSD1-SCMA detector and the simulation and theoretical error rate of the PRUN1 algorithm when simulating with the codebook [107]. The simulation error rate shows considerable performance gap with the theoretical error rate because the overlapping of the projections in each real dimension of the third codebook. This implies that the low value of M_1 of the codebook deteriorates the error rate of the proposed pruning algorithm.

Fig. 4.13 illustrates the BER performance of the two detectors when simulated with the third codebook. The SGSD1-SCMA detector outperforms MPA with 5 iterations at 8dB $\leqslant$ SNR $\leqslant$ 18.5dB when $p = 1 \times 10^{-4}$. Fig. 4.14 compares the number of FLOPs of the two detectors. The proposed detector has lower complexity than the MPA when SNR is larger than about 12dB. For such a high

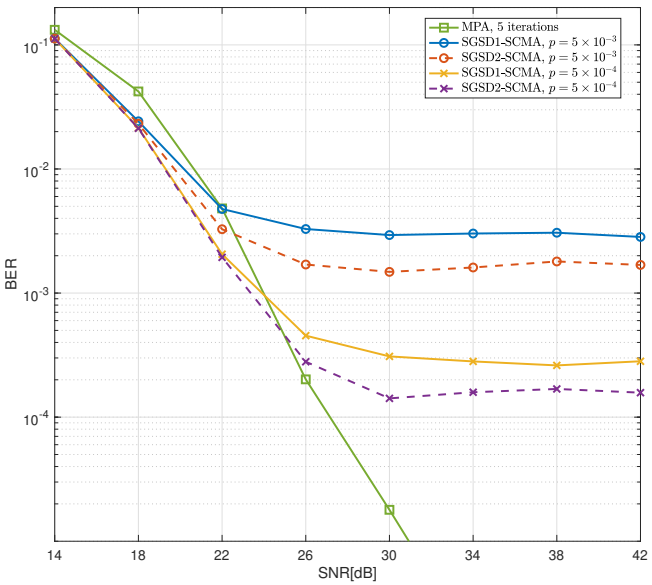


FIGURE 4.10 – BER performance of different detectors for the second codebook.

TABLE 4.8 – Size of sequence $\mathcal{L}_2^*$ of the proposed pruning algorithms for the second codebook.

p	Algorithm	14dB	18dB	22dB	26dB	30dB	34dB	38dB	40dB
5×10^{-3}	PRUN1	1255	638	398	288	234	198	184	172
	PRUN2	10552	7792	6200	5410	4941	4600	4393	4285
5×10^{-4}	PRUN1	1948	897	511	339	257	215	190	176
	PRUN2	12898	9047	7049	5842	5138	4724	4517	4349

overloading factor and small size codebook, substantial complexity of the SGSD1-SCMA detector comes from the PRUN1 algorithm, which is the opposite of two above codebooks.

Overall, the proposed SGSD-SCMA detector is significantly less complex than GSD-SCMA and MPA while reaching good MUD performance at low and moderate SNRs. The two proposed pruning algorithms, PRUN1 and PRUN2, have their own pros and cons. The PRUN1 algorithm is compatible with any SCMA codebook and has lower complexity cost than the PRUN2 algorithm. However, the PRUN1 algorithm has a slight relative performance degradation at moderate to high SNR values. On the other hands, the PRUN2 algorithm is subject to the factor graph constraint in Eq.(4.89). Furthermore, among the three investigated codebook settings, SGSD-SCMA detector is the most efficient when simulating with the second codebook. This implies that the proposed SGSD-SCMA detector is especially advantageous to the codebooks with moderate overloading factor and large codebook size in terms of both complexity and MUD performance. Regarding the error rate increase of the proposed pruning algorithms brought by the low codeword projection, the proposed detector

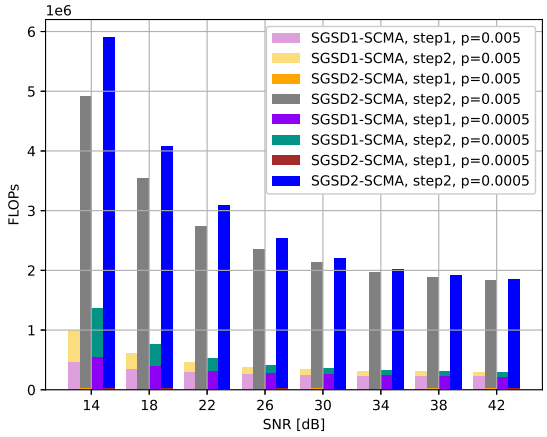


FIGURE 4.11 – FLOPs of two steps in SGSD-SCMA for the second codebook.

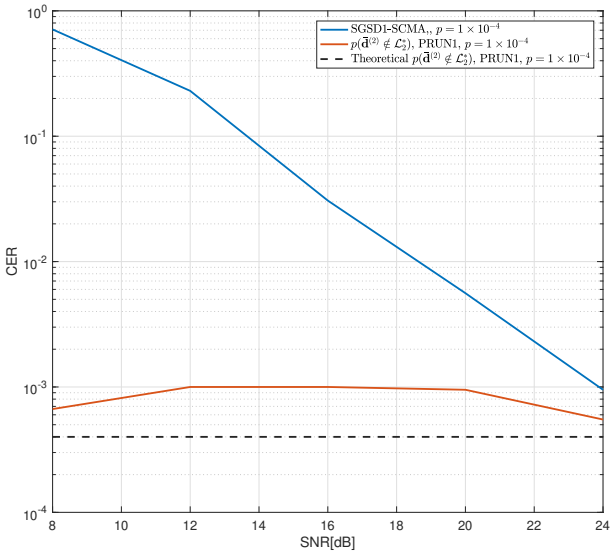


FIGURE 4.12 – CER performance of SGSD-SCMA for the third codebook.

performs better with full diversity SCMA codebooks.

4.8 Conclusion

In this chapter, we explore the SD-based detectors for SCMA systems. First, we review some of the SD-based state-of-the-art detectors which have less or no constraint of the codebook, namely SD-SCMA and GSD-SCMA. A comprehensive introduction of thr sate-of-the-art including flowchart and pseudo-code is presented. Moreover, as a novelty, we leverage the SQRD and Schnorr-Euchner enumeration to accelerate the tree search. Secondly, to remedy the computational complexity issue of

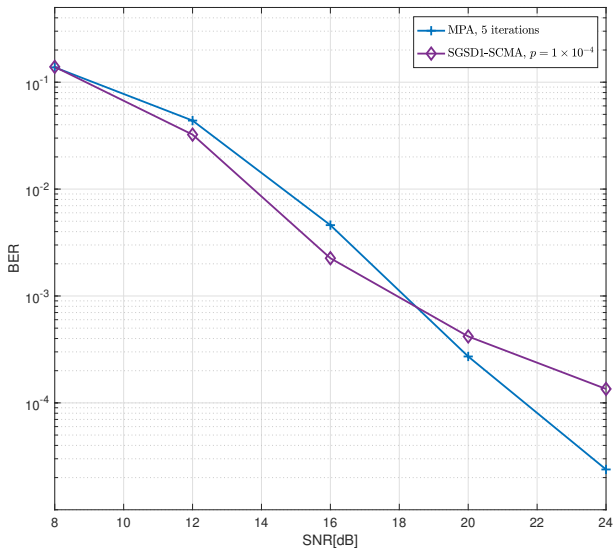


FIGURE 4.13 – BER performance of different detectors for the third codebook.

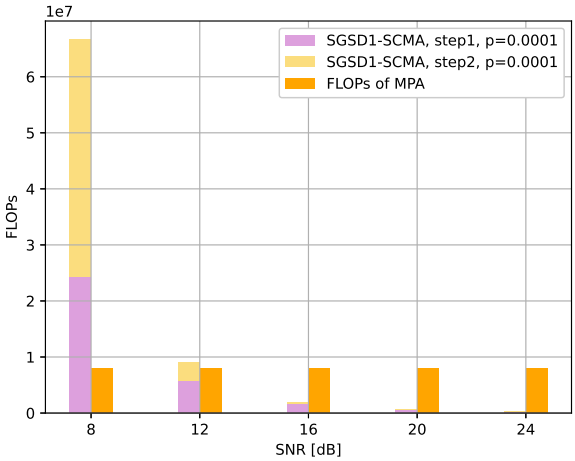


FIGURE 4.14 – FLOPs comparison of different detector for the third codebook.

GSD-SCMA, we propose two pruning algorithms, named PRUN1 and PRUN2, and introduce SGSD-SCMA. The PRUN1 algorithm is compatible to any SCMA codebook while the PRUN2 algorithm restricts the codebook factor graph. Beside, we derive a theoretical error rate expression of the two proposed pruning algorithms. In the performance evaluation, codebooks with three parameter settings are investigated. The simulated error rate of the pruning algorithms matches the theoretical one when the average number of projection in a real dimension approximates to the square root of the codebook size. The diversity of the codebook influences the error rate of the proposed pruning algorithms. The BER and complexity performance of the proposed SGSD-SCMA is compared with SD-based state-of-the-art and the conventional MPA. Although the efficiency of the proposed detector varies with

the codebook, generally it shows substantial BER performance gains and lower complexity than other detectors. By comparing the performance of different codebooks, we can conclude that SGSD-SCMA detector is especially advantageous for codebooks with moderate overloading factor and large size.

Chapitre 5

DL applications in NC MIMO systems

Contenu

5.1 Introduction

This chapter sheds light on NC MIMO systems. As introduced in section 2.2, Grassmannian constellations are theoretically verified to be the efficient space-time modulation for NC MIMO transmission. Constructing the optimal Grassmannian constellation is still an open question, although many conventional approaches have been proposed.

In this chapter, we take advantage of DL techniques and propose two AEs for Grassmannian constellation construction. The first proposed AE is called AE-FC which encompasses two fully-connected DNN as encoder and decoder, respectively. On the other hand, we integrate the domain knowledge of NC MIMO with DL techniques and introduce a CNN which carries out GLRT detection, called CNN-GLRT. By employing the CNN-GLRT at the decoder, we propose the second AE named AE-GLRT. To properly construct Grassmannian constellations, we introduce an orthonormalization block within the encoder of the proposed AEs which is also able to guarantee that the training algorithm, for example SGD, functions well. The proposed approaches is evaluated by comparing them to various state-of-the-art solutions.

5.2 Proposed AE-FC for Grassmannian constellation design

To facilitate the application of neural networks, the input-output equation given in Eq. (2.17) is rewritten using the equivalent real-valued notation :

$$\bar{\mathbf{Y}} = \bar{\mathbf{X}}\bar{\mathbf{H}} + \sqrt{\frac{N_t}{\rho T}}\bar{\mathbf{W}} \tag{5.1}$$

where

$$\bar{\mathbf{X}} = \begin{bmatrix} \Re\{\mathbf{X}\} & -\Im\{\mathbf{X}\} \\ \Im\{\mathbf{X}\} & \Re\{\mathbf{X}\} \end{bmatrix} \quad \bar{\mathbf{H}} = \begin{bmatrix} \Re\{\mathbf{H}\} & -\Im\{\mathbf{H}\} \\ \Im\{\mathbf{H}\} & \Re\{\mathbf{H}\} \end{bmatrix}$$
$$\bar{\mathbf{Y}} = \begin{bmatrix} \Re\{\mathbf{Y}\} & -\Im\{\mathbf{Y}\} \\ \Im\{\mathbf{Y}\} & \Re\{\mathbf{Y}\} \end{bmatrix} \quad \bar{\mathbf{W}} = \begin{bmatrix} \Re\{\mathbf{W}\} & -\Im\{\mathbf{W}\} \\ \Im\{\mathbf{W}\} & \Re\{\mathbf{W}\} \end{bmatrix}$$

$\mathbf{X}$, $\mathbf{H}$, $\mathbf{Y}$ and $\mathbf{W}$ are the complex-valued matrices used in Eq.(2.17). Consequently, the Grassmannian constellation constraint in Eq.(2.23) is rewritten as

$$\bar{\mathbf{X}}^{\dagger}\bar{\mathbf{X}} = \mathbf{I}_{2N_t} \tag{5.2}$$

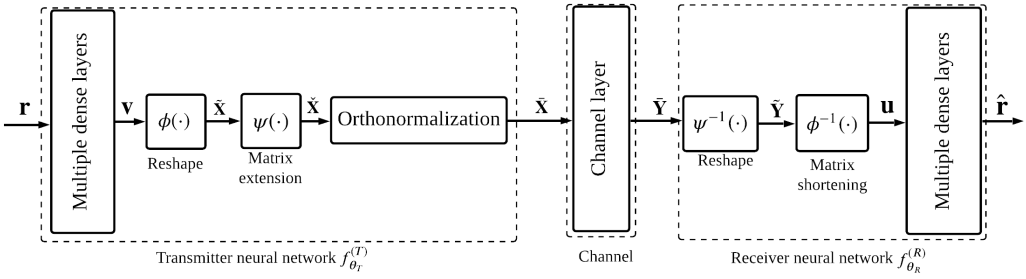


FIGURE 5.1 – Structure of the proposed AE-FC.

Fig. 5.1 illustrates the detailed structure of the proposed AE-FC where fully-connected DNNs are adopted in the transmitter and receiver neural network. Functions $f_{\boldsymbol{\theta}_T}^{(T)}(\cdot)$ and $f_{\boldsymbol{\theta}_R}^{(R)}(\cdot)$ are used to represent the transmitter and receiver neural network, respectively. Because $\bar{\mathbf{X}}$ and $\bar{\mathbf{Y}}$ are in a specific matrix format, both $f_{\boldsymbol{\theta}_T}^{(T)}(\cdot)$ and $f_{\boldsymbol{\theta}_R}^{(R)}(\cdot)$ contain not only multiple neural layers but also matrix operation blocks. $\boldsymbol{\theta}_T$ and $\boldsymbol{\theta}_R$ are the weight and bias set of fully-connected layers inside $f_{\boldsymbol{\theta}_T}^{(T)}(\cdot)$ and $f_{\boldsymbol{\theta}_R}^{(R)}(\cdot)$, respectively. The channel layer is built by a single neural layer where weights represent channel coefficients and biases represent additional noise. Stress that the channel layer is not trainable. Vector $\mathbf{O}_T = [O_{T,0}, O_{T,1}, \cdots, O_{T,l_T+1}]$ and $\mathbf{O}_R = [O_{R,0}, O_{R,1}, \cdots, O_{R,l_R+1}]$ denote the neuron number of each fully-connected layer at the transmitter and receiver, respectively. l_T is the number of hidden layers in $f_{\boldsymbol{\theta}_T}^{(T)}$ and l_R is the number of hidden layers in $f_{\boldsymbol{\theta}_R}^{(R)}$.

The input of the AE-FC is one-hot vector $\mathbf{r}$ with $M \times 1$ binary bits, in which only one element, indicating the transmitted message, is one and the others are zero. Notice that

$$r_n = \begin{cases} 1, & n = m \\ 0, & \text{otherwise} \end{cases} \quad (1 \leqslant n \leqslant M) \tag{5.3}$$

, where r_n is the n-th element of vector $\mathbf{r}$ and m is the transmitted message. The input vector is passed through the first multiple dense layers, then matrix operations and finally orthonormalization process. Vector $\mathbf{v} \in \mathbb{R}^{2TN_t \times 1}$ is the output of the multiple dense layers at the transmitter with input $\mathbf{r}$. Since the transmitted signal is a matrix, it is necessary to tailor vector $\mathbf{v}$. Firstly, reshaping function $\boldsymbol{\phi}(\cdot) : \mathbb{R}^{2TN_t \times 1} \to \mathbb{R}^{2T \times N_t}$ converts column vector $\mathbf{v}$ into matrix $\tilde{\mathbf{X}}$, formulated as

$$\tilde{\mathbf{X}} = \boldsymbol{\phi}(\mathbf{v}) = \begin{bmatrix} \tilde{\mathbf{X}}_1 \\ \tilde{\mathbf{X}}_2 \end{bmatrix} \tag{5.4}$$

where $\tilde{\mathbf{X}}_1 \in \mathbb{R}^{T \times N_t}$ and $\tilde{\mathbf{X}}_2 \in \mathbb{R}^{T \times N_t}$ are upper and lower half of $\tilde{\mathbf{X}}$, respectively. Then, function $\boldsymbol{\psi}(\cdot) : \mathbb{R}^{2T \times N_t} \to \mathbb{R}^{2T \times 2N_t}$ enlarges the matrix $\tilde{\mathbf{X}}$ to $\check{\mathbf{X}}$ by duplicating the entries in a particular way,

which is expressed as

$$\check{\mathbf{X}} = \boldsymbol{\psi}(\tilde{\mathbf{X}}) = \boldsymbol{\psi}\Big(\begin{bmatrix} \tilde{\mathbf{X}}_1 \\ \tilde{\mathbf{X}}_2 \end{bmatrix} \Big) = \begin{bmatrix} \tilde{\mathbf{X}}_1 & -\tilde{\mathbf{X}}_2 \\ \tilde{\mathbf{X}}_2 & \tilde{\mathbf{X}}_1 \end{bmatrix} \tag{5.5}$$

To satisfy the requirement of Grassmannian constellations, orthonormalization is adopted to ensure that the output matrices of $f_{\boldsymbol{\theta}_T}^{(T)}$ are unitary as demonstrated in Eq.(5.2). Two methods can be applied to carry out the orthonormalization.

Method 1. This is based on the square root of $\check{\mathbf{X}}^{\intercal}\check{\mathbf{X}}$ which is written as

$$\bar{\mathbf{X}} = \check{\mathbf{X}}(\check{\mathbf{X}}^{\intercal}\check{\mathbf{X}})^{-1/2} \tag{5.6}$$

Method 2. If the matrix $\check{\mathbf{X}}$ can be decomposed by applying the singular value decomposition (SVD), it is expressed as $\check{\mathbf{X}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\intercal}$. Therefore, Eq.(5.6) can be rewritten as

$$\bar{\mathbf{X}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\intercal}(\mathbf{V}\boldsymbol{\Sigma}^{-1}\mathbf{V}^{\intercal}) = \mathbf{U}\mathbf{V}^{\intercal} \tag{5.7}$$

In the receiver neural network, $\boldsymbol{\phi}^{-1}(\cdot)$ and $\boldsymbol{\psi}^{-1}(\cdot)$ are the inverse functions of $\boldsymbol{\phi}(\cdot)$ and $\boldsymbol{\psi}(\cdot)$, respectively. They are utilized to convert matrices to column vectors. The input of the multiple dense layers within $f_{\boldsymbol{\theta}_R}^{(R)}$ is $\mathbf{u} \in \mathbb{R}^{2TN_r \times 1}$, which is obtained by

$$\mathbf{u} = \boldsymbol{\phi}^{-1}(\tilde{\mathbf{Y}}) = \boldsymbol{\phi}^{-1}(\boldsymbol{\psi}^{-1}(\bar{\mathbf{Y}})) \tag{5.8}$$

The output of multiple dense layers in $f_{\boldsymbol{\theta}_R}^{(R)}(\cdot)$ is vector $\hat{\mathbf{r}} \in [0,1]^{M \times 1}$ following $||\hat{\mathbf{r}}||_1 = 1$, which contains the probabilities of all corresponding messages. Hence, the $Softmax$ activation function is applied at the last dense layer at the receiver. In the proposed AE-FC, loss function is defined as categorical cross entropy to penalize the difference between $\mathbf{r}$ and $\hat{\mathbf{r}}$. It is defined as

$$\mathcal{L} = -\sum_{i=1}^{S} \|diag(\mathbf{r}^{(i)}) \log(\hat{\mathbf{r}}^{(i)})\|_1 \tag{5.9}$$

where $\log(\cdot)$ is the element-wise logarithm operation of vectors and S denotes the training batch size. Please note that instead of maximizing the metric in Eq.(2.24), the proposed AE optimizes the Grassmannian constellations by minimizing the loss function values.

5.3 Proposed AE-GLRT for Grassmannian constellation design

5.3.1 Proposed CNN-GLRT receiver

Driven by the real-valued system model expressed in Eq.(5.1), the GLRT detector can also be rewritten in the real domain, as below

$$\hat{\bar{\mathbf{X}}} = \arg\max_{\bar{\mathbf{X}}\in\bar{\mathcal{C}}} \mathrm{Tr}\{\bar{\mathbf{Y}}^{\intercal}\bar{\mathbf{X}}\bar{\mathbf{X}}^{\intercal}\bar{\mathbf{Y}}\} \tag{5.10}$$

where $\bar{\mathcal{C}}$ is the corresponding real-valued codebook. Defining $\mathbf{S} = \bar{\mathbf{X}}^{\intercal}\bar{\mathbf{Y}}$, the GLRT detector in the real domain can be reformulated to

$$\hat{\bar{\mathbf{X}}} = \arg\max_{\bar{\mathbf{X}}\in\bar{\mathcal{C}}} \mathrm{Tr}\{\mathbf{S}^{\intercal}\mathbf{S}\} \tag{5.11}$$

We can see that the real-valued GLRT detector depends on the matrix $\mathbf{S}$. Exploiting the matrix multiplication and matrix transpose, the element in the i-th ($1 \leqslant i \leqslant 2N_r$) row and j-th ($1 \leqslant j \leqslant 2N_t$) column of matrix $\mathbf{S}$, $s_{i,j}$, is calculated by

$$s_{i,j} = \sum_{n=1}^{2T} \bar{x}(n,i)\bar{y}(n,j) \tag{5.12}$$

where $\bar{x}(n,i)$ denotes the element in the n-th row and i-th column of matrix $\bar{\mathbf{X}}$ and $\bar{y}(n,j)$ denotes the element in the n-th row and j-th column of matrix $\bar{\mathbf{Y}}$.

Recall the 2D convolutional layer presented in section 2.4.3, if the input has size $(2T, 2N_r)$, the kernel has size $(2T, 1)$ and the stride is $(2T, 1)$, the size of the output is $(1, 2N_r)$. According to Eq.(2.39), the value of the q-th ($1 \leqslant q \leqslant 2N_r$) element of the output is computed as

$$b_{1,q} = \sum_{k_1=1}^{2T} k_{k_1,1}a_{k_1,q} \tag{5.13}$$

Thus, if the kernel of this 2D convolutional layer equals to the i-th ($1 \leqslant i \leqslant 2N_t$) column of matrix $\bar{\mathbf{X}}$ and input is the received signal $\bar{\mathbf{Y}}$, the output of this layer is equivalent to the i-th row vector of matrix $\mathbf{S}$. This implies that the calculation of matrix $\mathbf{S}$ can be carried out by a 2D convolitional layers if we can obtain the kernels appropriately from the given codebook. As a result, we propose CNN-GLRT which implement the GLRT detection by a CNN with a 2D concolutional layer. Specifically, for a given codeword $\bar{\mathbf{X}}$ ($\bar{\mathbf{X}} \in \bar{\mathcal{C}}$), it is split into $2N_t$ different kernels with size $(2T, 1)$. By applying the 2D convolutional layer described above, we can obtain a value of matrix $\mathbf{S}$ by stacking the output of the $2N_t$

kernels. Since the GLRT detector tests all M codewords as shown in Eq.(5.11), the 2D convolutional layer adopts the kernels generated from all the codewords. That is the 2D convolutional layer has $2MN_t$ kernels in total. After obtaining all the values of matrix $\mathbf{S}$, the CNN-GLRT detector continues to reproduce the operation shown in Eq.(5.11). It is worth noticing that because all the kernels used in the convolutional layer are obtained from the codebook, there is no trainable parameter and no need of training the CNN-GLRT.

Here, we will explain how we build the proposed CNN-GLRT on the DL framework PyTorch [82]. Because there is no trainable parameter, we call the PyTorch function $torch.nn.functional.conv2d(\cdot)$ in our code. The weights of this function are the stacked kernels having size $(2MN_t, 1, 2T, 1)$ and the biases of this function are zeros. The input of the called function has size $(N_b, 1, 2T, 2N_r)$, where N_b is the size of the batch and 1 indicates the number of input channels. Following the rules of the function $torch.nn.functional.conv2d(\cdot)$ [114], the output has size $(N, 2MN_t, 1, 2N_r)$. By shaping the output properly, its size becomes $(N, M, 2N_t, 2N_r)$ which is the batch of the stacked M values of matrix $\mathbf{S}$ calculated based on the received signals and the codebook.

5.3.2 Structure of AE-GLRT

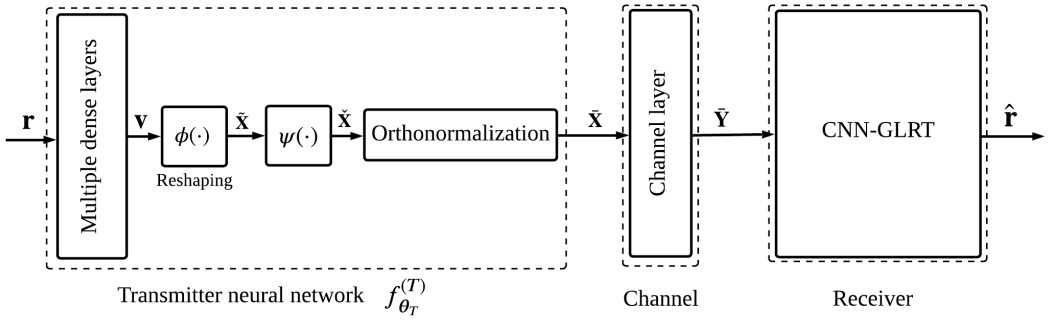


FIGURE 5.2 – Structure of the proposed AE-GLRT.

In the last subsection, we introduce a method to implement GLRT detection on a CNN. Thus, we propose the AE-GLRT by applying the CNN-GLRT as the decoder of an AE. Specifically, AE-GLRT consists of a fully-connected DNN as the encoder and CNN-GLRT as the decoder. The decoder of AE-GLRT is expected to give quasi-GLRT detection performance. It has no gives trainable parameter and aims at evaluating the performance of the generated codebook. Fig. 5.2 depicts the detailed structure of the proposed AE-GLRT. It is worth-noticing that the decoder of AE-GLRT can process

TABLE 5.1 – Structure and training hyper-parameters of the proposed AE-FC in NC MIMO.

Name	Proposed AE1
M	256
$\mathbf{O}_T$	$[M, 20M, 20M, 2TN_t]$
$\mathbf{O}_R$	$[2TN_r, 20M, 20M, M]$
S	2000
Nb of epochs	50
Nb of batches	50
Learning rate	0.0003
Optimizer	Adam
Channel	i.i.d. Rayleigh

the real-valued received signal $\bar{\mathbf{Y}}$ directly.

Similarly with the first proposed AE-FC, the inputs of AE-GLRT are one-hot vectors and the outputs are vectors containing the probabilities of all the corresponding codewords obtained by applying the $Softmax$ activation function at the output of the decoder. The encoder DNN structure of AE-GLRT is denoted as $\mathbf{O}_T$ as well. The loss function is defined as the categorical cross entropy between the input and output as in Eq.(5.9).

5.4 Numerical results

We evaluate the proposed AE-based approaches by them comparing with three different state-of-the-art approaches, including the Lloyd algorithm [28], the greedy algorithm [26] and $G_{4,2}$ constellation [30]. We assume the number of transmit antennas $N_t = 2$, number of received antennas $N_r = 2$, coherence interval $T = 4$ and constellation size $M = 256$. Two types of transmission channel, i.i.d. Rayleigh and correlated fading channels Eq.(2.20) are considered. It is worth stressing that the proposed AEs are optimization methods for constellation design and are only involved in the training stage. After the offline-training, the learned constellations are implemented into look-up tables at the transmitter.

Firstly, we study the symbol error rate (SER) performance of AE-FC. Table 5.1 provides structure and details of it. In AE-FC, the rectified linear units (ReLU) activation function is applied at the intermediate layers and Xavier method [115] is applied for the weights initialization. We implement the orthonormalization with method 1 due to the back-propagation constraint of the SVD operation in Pytorch. In this simulation, ML detection Eq.(2.21) is adopted at the receiver.

Fig. 5.3 compares the SER performance of the constellations trained at different SNRs by the pro-

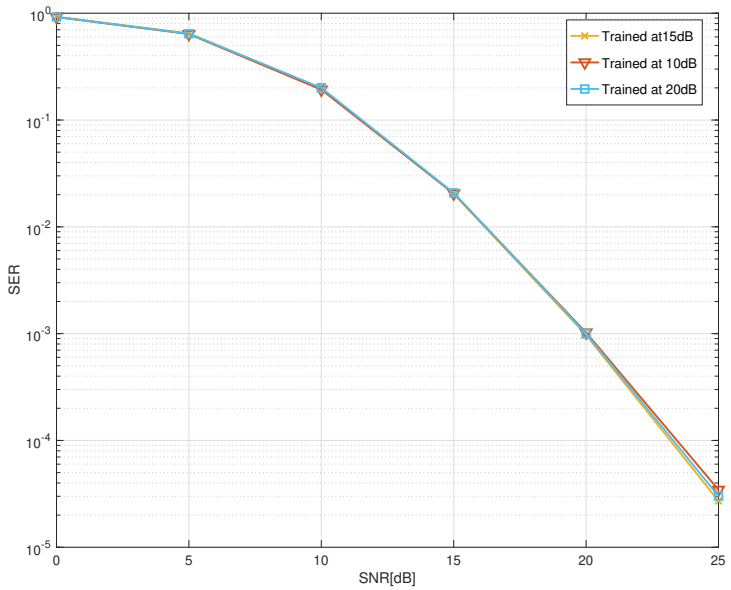


FIGURE 5.3 – SER performance of the proposed AE-FC with different training SNRs in i.i.d. Rayleigh channel.

posed AE-FC. It is obvious that the training SNR has negligible impact on the resulting constellation performance. This implies that the proposed AE-FC can render the constellations good performance over a wide range of SNR by training at only single SNR value. In the following, we only consider the codebook trained at 15dB.

Fig. 5.4 compares the SER performance of AE-FC and state-of-the-art in i.i.d. Rayleigh channel. The constellation proposed AE has performance gains over other constellations when SNR is greater than 10dB. At SER $= 10^{-3}$, it outperforms the greedy approach and $G_{4,2}$ constellations by 0.7dB, Lloyd approach by 1.1 dB and the non-Grassmannian constellation by 1.4dB. When generating Grassmannian constellations, the AE-CF approach considers the cross entropy metric between the input and the output, rather than chordal Frobenius distance used by the conventional approaches. Therefore, intuitively, the newly proposed AE-FC approach is more efficient to design Grassmannian constellations regarding lower SER.

Fig. 5.5 illustrates the SER performance comparison of AE-FC and state-of-the-art in correlated fading channel with $r = 0.9$. The resulting constellation of the proposed AE-FC outperforms the other state-of-the-art constellations by about 0.9dB at $SNR = 20$dB over correlated fading channel.

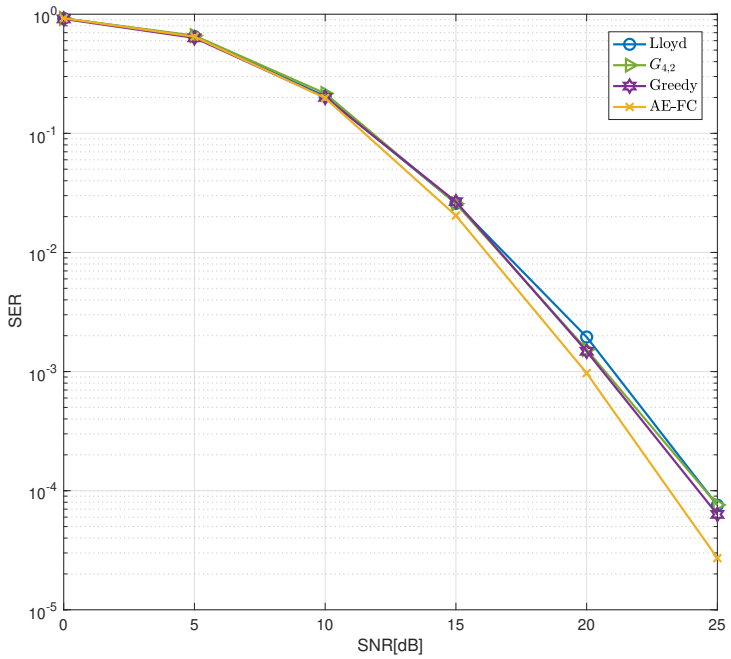


FIGURE 5.4 – SER performance of the proposed AE-FC in i.i.d. Rayleigh channel.

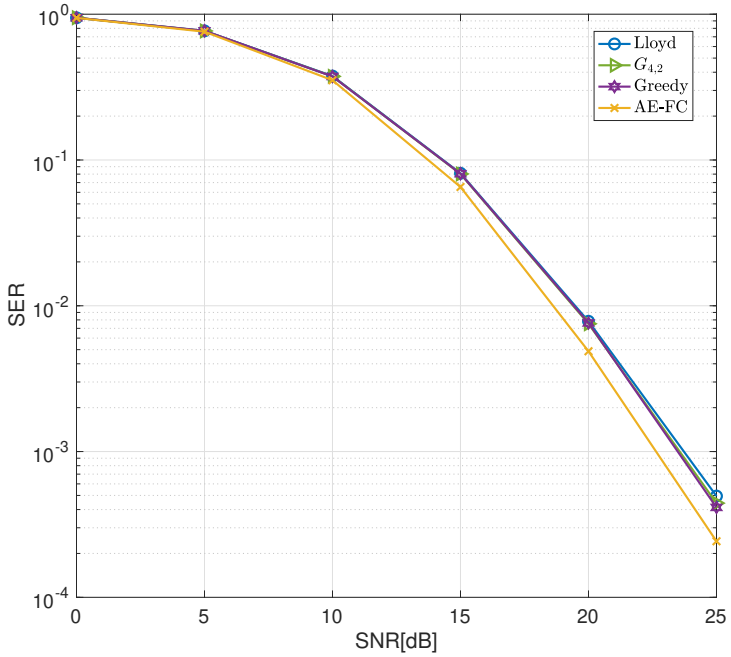


FIGURE 5.5 – SER performance of the proposed AE-FC in correlated fading channel with $r = 0.9$.

Secondly, we investigate the SER performance of the proposed AE-GLRT. Before diving into the performance of the resulting constellations, we first evaluate the CNN-GLRT detector. Fig. 5.6 compares the SER performance of CNN-GLRT and conventional GLRT detector. The $G_{4,2}$ constellation

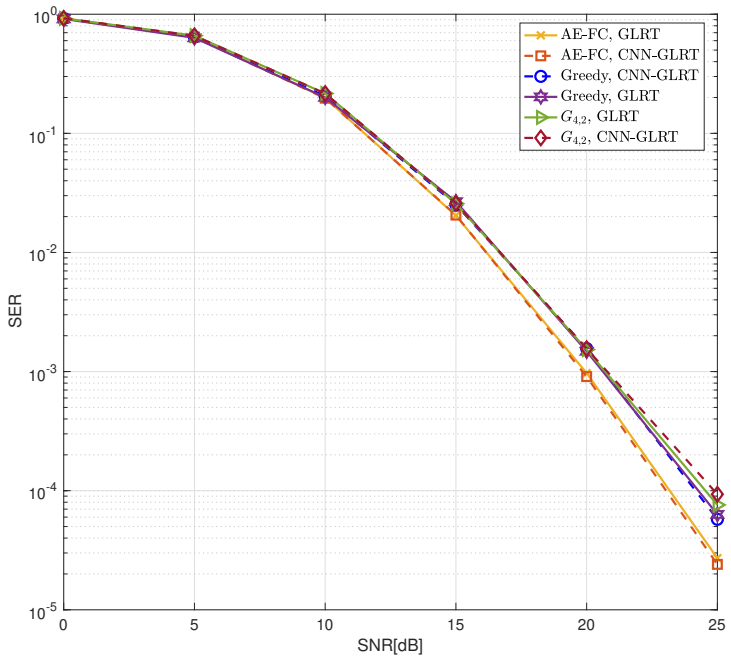


FIGURE 5.6 – SER performance of CNN-GLRT detector.

TABLE 5.2 – Structure and training hyper-parameters of the proposed AE-GLRT.

Name	AE-GLRT
N	2000
Nb of epochs	20
Nb of batches	20
Learning rate	0.0003
Optimizer	Adam
SNR	10dB
$\mathbf{O}_T$	$[256, 5120, 5120, 16]$
Channel	i.i.d. Rayleigh

[30] as well as the constellations designed by greedy algorithm [26] and the proposed AE-FC are simulated in the comparative SER performance of the two detectors. As shown in Fig. 5.6, CNN-GLRT has the same detection effectiveness as the conventional GLRT.

Table 5.2 lists the structure and all the hyper-parameters in the training of the proposed AE-GLRT. There are two hidden layers of the encoder of the proposed AE-GLRT. Fig. 5.7 illustrates the SER performance of the resulting codebook of AE-GLRT in i.i.d. Rayleigh channel. As expected, applying the fully-connected receiver of AE-FC results in poor SER performance. When simulating with the CNN-GLRT detector, the codebooks generated by the two proposed AE have close SER performance and both outperform the codebooks by the state-of-the-art approaches.

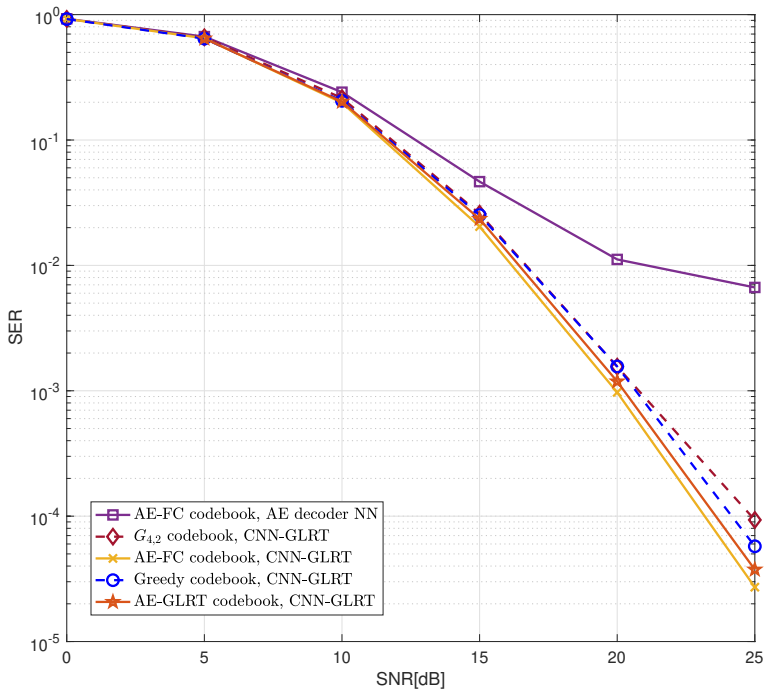


FIGURE 5.7 – SER performance of the proposed AE-GLRT in i.i.d. Rayleigh channel.

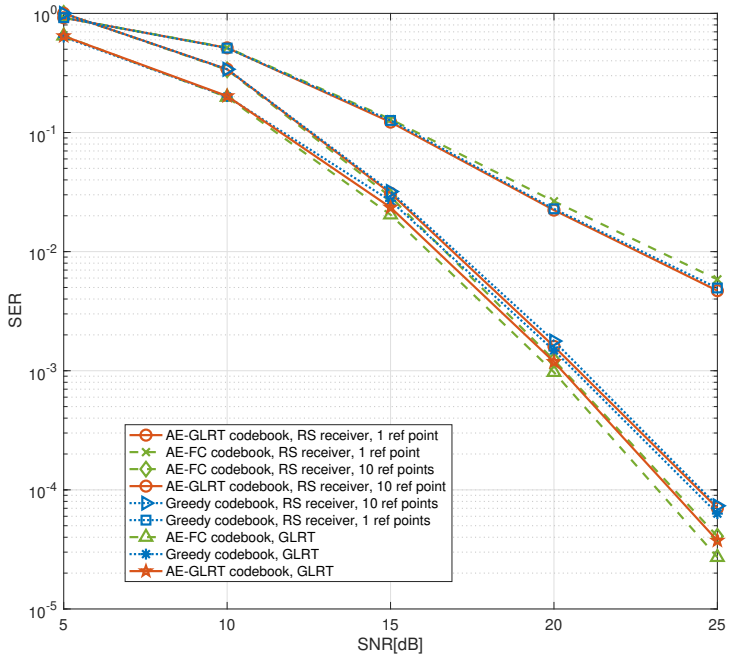


FIGURE 5.8 – SER performance of the reduced-search detector in i.i.d. Rayleigh channel.

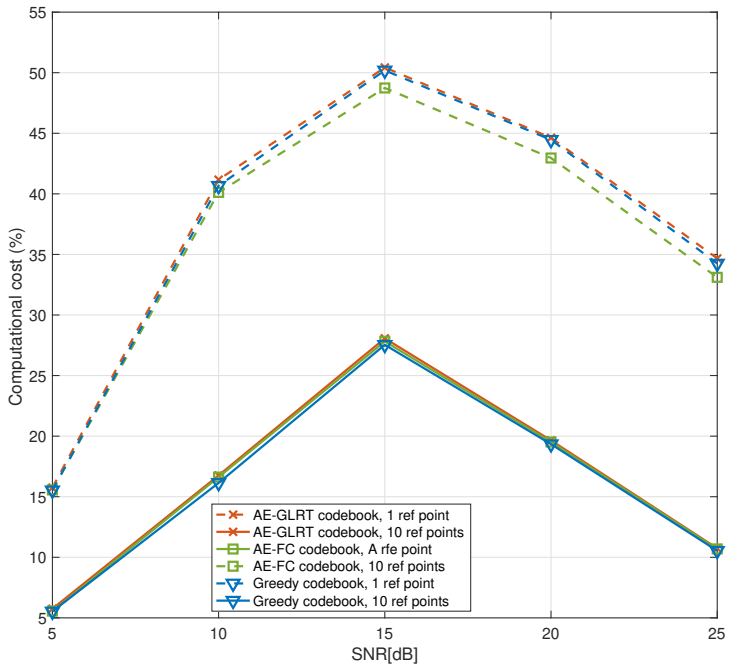
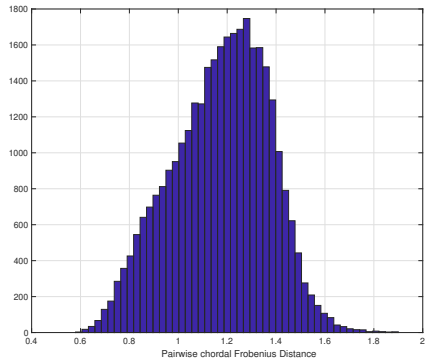


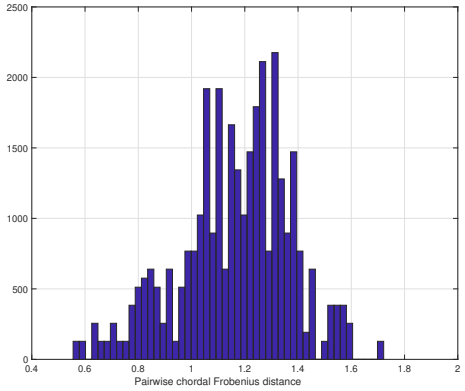
FIGURE 5.9 – Computational cost of the reduced-search detector.

Besides, we simulate the SER performance with the reduced-search (RS) detector proposed in [26]. Fig. 5.8 and Fig. 5.9 illustrates the SER performance and computational cost of different codebooks when simulating with the RS detector. Please note that the computational cost is a ratio of the number of constellation points considered in the RS detector and the codebook size M. The resulting codebook of the proposed AE-GLRT has the best SER performance and lowest computational cost compared to other codebooks when the RS detector has only one reference point.

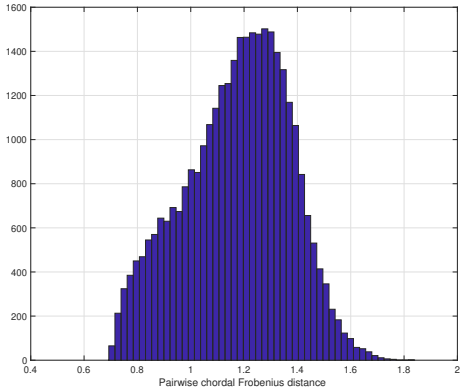
Fig. 5.10 shows the distribution and mean of the pairwise chordal Frobenius distance between the Grassmannian constellation points. It is notable that the distribution of constellation $G_{4,2}$ is different from the others since it is algebraically designed for the system with $T = 4$ and $N_t = 2$ while the others are generic optimization methods for NC MIMO systems. The resulting codebooks of the two proposed AEs resemble each other. It is noticeable that the greedy codebook has considerable large amount of low-value pairwise distances. The proposed AEs shows negligible mean distance gap compared to the Lloyd and greedy approaches.



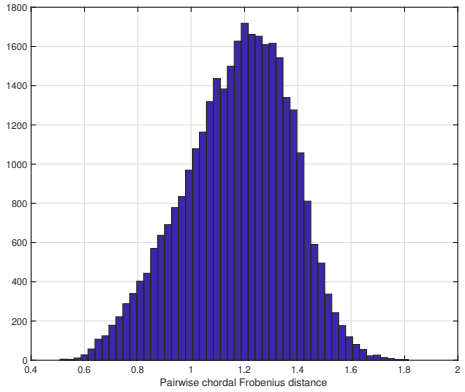
(a) Constellation by *Lloyd* ($d_{mean} = 1.1731$).



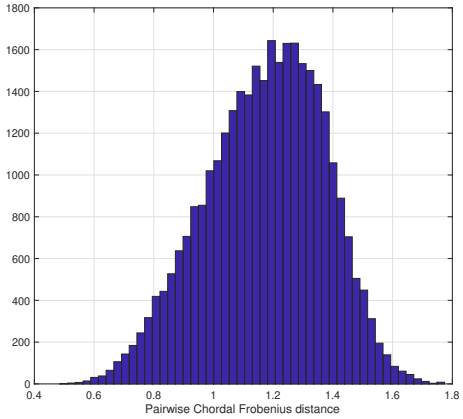
(b) Constellation $G_{4,2}$ [30] ($d_{mean} = 1.1634$).



(c) Constellation by greedy approach [26] ($d_{mean} = 1.1730$).



(d) Constellation by the proposed AE-FC ($d_{mean} = 1.1727$).



(e) Constellation by the proposed AE-GLRT ($d_{mean} = 1.1729$).

FIGURE 5.10 – Distribution of pairwise chordal Frobenius distance between constellation points.

5.5 Conclusion

In this chapter, we propose two novel AE-based schemes to design Grassmannian constellations for NC MIMO systems. To satisfy the requirement of Grassmannian constellations, we introduce an orthonormalization process which facilitates the back propagation for the training of the proposed AEs. Besides, we have proposed to implement the GLRT detection using a CNN. The proposed schemes are compared with various sate-of-the-art in the performance evaluation. The constellations learned by the novel approaches significantly outperforms the conventionally designed constellations in moderate and high SNR regime over i.i.d. Rayleigh and correlated fading channel. It might be suggested that the proposed AEs are promising optimization methods of Grassmannian constellation design for all SNRs in diverse channel conditions.

Chapitre 6

DL-based approach for SU-MDC design

Contenu

6.1 Introduction

In this chapter, we focus on MDC design for SU multidimensional transmission which is motivated by the potential OFDM-based schemes explained in section 2.3. For the details of the transmission model, please refer to section 2.3.1. We propose a novel DNN structure which is dedicated to MDC design over Rayleigh fading channels. Without applying an AE scheme, the proposed scheme exploits loss function in order to generate codebooks with good performance. Therefore, we propose two loss functions based on expertise in communication systems for the proposed approach. Contrary to the classical loss function which calculates the difference between the input and output messages, the proposed loss functions act on learned codebooks directly by maximizing essential codebook FoMs or decreasing theoretical symbol error probability (SEP). The proposed DNN structure is significantly simpler than an AE structure and can be regarded as a standalone optimization method. Furthermore, the novel approach is more advantageous than the data-driven AE techniques when a large amount of training data is not necessary. The proposed approach is evaluated by comparing with conventional state-of-the-art approaches and AE techniques.

6.2 Important FoMs in MDC design

In this section, we highlight the FoMs that are mostly considered in the design of MDCs. Since the average symbol power of a codebook significantly changes the FoMs, all the following expressions are based on the power constraint in Eq.(2.28) for fair comparison.

MED

$$d_E(\mathcal{C}) = \min_{m,m'} \left\{ \left\| \mathbf{x}_m - \mathbf{x}_{m'} \right\|, \quad 1 \leq m < m' \leq M \right\} \tag{6.1}$$

where $\mathbf{x}_m$ and $\mathbf{x}_{m'}$ are m-th and m'-th element of codebook $\mathcal{C}$ respectively.

MPD

$$d_P(\mathcal{C}) = \min_{m,m'} \left\{ \prod_{n=1}^{2N} \|x_{m,n} - x_{m',n}\|, \quad 1 \leq m < m' \leq M \right\} \tag{6.2}$$

where $x_{m,n}$ and $x_{m',n}$ are the n-th entry of vector $\mathbf{x}_m$ and $\mathbf{x}_{m'}$ respectively. In the computation of the MPD, we only consider the parameter n satisfying $\|x_{m,n} - x_{m',n}\| > \delta$. δ is the threshold that distinguishes different constellation projections in each real dimension. It is set as $\delta = 0.001$ in this thesis.

SSD

$$d_S(\mathcal{C}) = \min_{m,m'} \left\{ d_H(\mathbf{x}_m, \mathbf{x}_{m'}), \quad 1 \leq m < m' \leq M \right\} \tag{6.3}$$

where $d_H(\mathbf{x}_m, \mathbf{x}_{m'})$ represents the Hamming distance considering the threshold $\delta = 0.001$.

6.3 Proposed DL-based approach for MDC design

The proposed DL-based approach is elaborated in this section. The overall DNN structure will first be introduced, then two types of loss function devised for the DNN structure will be explained in detail.

6.3.1 The proposed DNN structure

The proposed approach consists in a fully-connected DNN with multiple layers. Its structure is denoted as $\mathbf{O} = [O_0, \cdots, O_L, O_{L+1}]$, where L is the number of hidden layers and O_l ($l = 0, 1, \cdots, L+1$) represents the neuron number at the l-th layer. Fig. 6.1 illustrates the scheme of the proposed approach, where function $f(\cdot)$ denotes the proposed DNN structure. Input of $f(\cdot)$ are the one-hot vectors $\mathbf{r}_m \in \mathbb{R}^{M\times 1}$ ($m \in \{1, \cdots, M\}$), in which only one element, indicating the transmitted message, is one and the others are zero. Hence, there is $(r_m)_k = \begin{cases} 1, & m = k \\ 0, & m \neq k \end{cases}$, where $(r_m)_k$ is the k-th element of $\mathbf{r}_m$. The corresponding output of DNN $f(\cdot)$ with input $\mathbf{r}_m$ is codeword $\mathbf{x}_m \in \mathbb{R}^{2N\times 1}$, formulated as

$$\mathbf{x}_m = f(\mathbf{r}_m) \tag{6.4}$$

The codebook learned by the proposed approach comprises all the codeword $\mathbf{x}_m$ ($m \in \{1, \cdots, M\}$), expressed as

$$\mathcal{C} = f(\mathbf{R}) \tag{6.5}$$

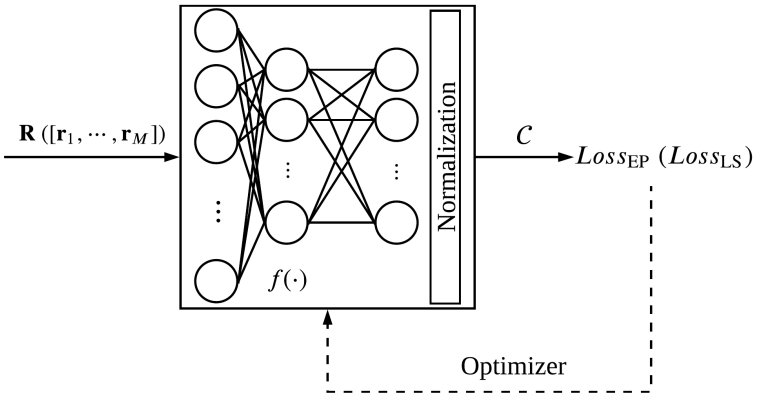


FIGURE 6.1 – Training scheme of the proposed DNN.

where matrix $\mathbf{R} = [\mathbf{r}_1, \cdots, \mathbf{r}_M]$ concatenates all the input one-hot vectors. The neuron number of the input and output layer of DNN $f(\cdot)$ are determined by the sizes of the one-hot vectors and output codewords, which are $O_0 = M$ and $O_{L+1} = 2N$, respectively. In order to ensure the power constraint of the codebook in Eq.(2.28), a normalization layer is implemented as the final process inside $f(\cdot)$, which is given as

$$f(\mathbf{r}_m) = \frac{\tilde{f}(\mathbf{r}_m)}{\sqrt{\frac{\|\tilde{f}(\mathbf{r}_m)\|_2^2}{N}}} \tag{6.6}$$

where $\tilde{f}(\mathbf{r}_m)$ represents the output of multiple dense layers before the normalization process. Different from the AE techniques composed of an encoder and decoder neural network and aiming at recovering the input message of encoder neural network at the output of decoder neural network, the proposed approach significantly relies on its loss function to learn codebooks with good performance as illustrated in Fig. 6.1. In the next subsections, we will introduce the two loss functions compatible with the above DNN structure, based respectively on classical FoMs and error probability.

6.3.2 Loss function based on MED and MPD

As introduced previously that FoMs are important metrics to evaluate MDCs, they can be applied in the design as well. Consequently, the first loss function for the proposed approach aims at maximizing MED and MPD at the same time to resolve the MDC optimization problem. It is denoted as $Loss_{\mathrm{EP}}$. Accordingly, $f_{\mathrm{EP}}(\cdot)$ refers to the proposed DNN using the loss function $Loss_{\mathrm{EP}}$ for training and its resulting codebook is denoted as $\mathcal{C}_{\mathrm{EP}}$. Considering the characteristics of the gradient descent, the loss

function is defined as the negative of the sum of MED and MPD. $Loss_{\mathrm{EP}}$ is formulated as

$$Loss_{\mathrm{EP}} = -\Big(d_E(\mathcal{C}_{\mathrm{EP}}) + \alpha d_P(\mathcal{C}_{\mathrm{EP}}) \Big) \tag{6.7}$$

where α is a coefficient chosen to coordinate the value scale of MED and MPD to facilitate the training. α is always positive and can be optimized separately as a hyper-parameter.

6.3.3 Loss function based on theoretical SEP

Since the objective of the proposed approach is to design MDCs with good performance, for example low error probabilities, it is intuitive to exploit theoretical SEP in the loss function. Because the value of SEP is rather small, much lower than 1, it could probably become very trivial during training, which possibly make the DNN converge toward a local optimum. As the logarithm function is monotonically increasing and can map SEP to a negative real number, it is applied in the loss function in order to enable a more efficient training of the proposed approach. Consequently, the second type of loss function is denoted as $Loss_{\mathrm{LS}}$ and its resulting codebook is denoted as $\mathcal{C}_{\mathrm{LS}}$. $Loss_{\mathrm{LS}}$ is written as

$$Loss_{\mathrm{LS}} = \log\left(\frac{1}{M} \sum_{m=1}^{M} \sum_{\substack{m_1=1 \\ m_1 \neq m}}^{M} P\big(f_{\mathrm{LS}}(\mathbf{r}_m) \to f_{\mathrm{LS}}(\mathbf{r}_{m_1})\big) \right) \tag{6.8}$$

where $f_{\mathrm{LS}}(\mathbf{r}_m)$ and $f_{\mathrm{LS}}(\mathbf{r}_{m_1})$ represent the m-th and m_1-th codeword learned by the proposed DNN with the loss function $Loss_{\mathrm{LS}}$ respectively. $P\big(f_{\mathrm{LS}}(\mathbf{r}_m) \to f_{\mathrm{LS}}(\mathbf{r}_{m_1})\big)$ is the pairwise error probability (PEP) between the two codewords. Assuming that the real dimension-wise distance between any two codewords is unique, the PEP between codeword $\mathbf{a}$ and $\mathbf{b}$ can be approximated as [41]

$$P(\mathbf{a} \to \mathbf{b}) = \sum_{n=1}^{2N} \left(\frac{1-\mu_n}{2} \right) \prod_{\substack{n'=1 \\ n' \neq n}}^{2N} \frac{\delta_n^2}{\delta_n^2 - \delta_{n'}^2} \tag{6.9}$$

where δ_n is the n-th element of vector $\boldsymbol{\delta} = |\mathbf{a} - \mathbf{b}|$ defined as absolute distance vector between $\mathbf{a}$ and $\mathbf{b}$, and $\mu_n = \sqrt{\frac{\delta_n^2}{4N_0 + \delta_n^2}}$.

6.4 Numerical results

In this section, performance of the proposed approach is investigated by comparing it with various state-of-the-art. Both MDCs designed by AE techniques, including model-aware [43] and model-free

TABLE 6.1 – Training hyper-parameters the proposed DNN-based approach.

	f_{EP}		f_{LS}	
	$(16, 4)$	$(256, 8)$	$(16, 4)$	$(256, 8)$
Batch size	M			
Batch number	1			
Epoch number	5000	10000	5000	1000
Learning rate	$1 * 10^{-4}$	$1 * 10^{-5}$	$3 * 10^{-4}$	$1 * 10^{-5}$
Training SNR	10dB		20dB	

TABLE 6.2 – Training hyper-parameters of AEs.

	Model-aware		Model-free	
	$(16, 4)$	$(256, 8)$	$(16, 4)$	$(256, 8)$
Batch size	4000	8000	4000	8000
Batch number	100	200	100	200
Epoch number	50	50	50	50
Learning rate	$3 * 10^{-5}$	$1 * 10^{-5}$	$3 * 10^{-4}$	$3 * 10^{-5}$
Training SNR	10dB			

[42], and conventional approaches, lower projection based MDC design [67] and GLCP with rotation [35] are studied. For the lower projection based MDC design, (M, M_q) denotes M-ary codebook with M_q projected points over every (real) dimension. The transmission system sending one out of M codewords through $2N$ real channel uses is represented as $(M, 2N)$. Two sets of transmission system are considered in the comparison, namely, $(16, 4)$ and $(256, 8)$. The proposed DNN structures are $\mathbf{O} = [M, 32, 32, 2N]$ and $\mathbf{O} = [M, 512, 512, 512, 2N]$ for the system $(16, 4)$ and $(256, 8)$, respectively. Table 6.1 shows the training hyper-parameters of the novel approach. For f_{EP}, the coefficient α equals to 10 in the system $(16, 4)$ and 10000 in the system $(256, 8)$. Identical activation function ReLU is adopted for all the hidden layers. Xavier method [115] is employed for the DNN weight initialization and Adam optimizer is applied during the training. $\mathbf{O}_T$ is the encoder DNN structure of baseline AEs [43, 42] which is same as $\mathbf{O}$. Model-aware and model-free AEs have the same decoder neural network structure which is $\mathbf{O}_R = [2N, 48, 48, M]$ for the system $(16, 4)$ and $\mathbf{O}_R = [2N, 96, 96, 96, K]$ for the system $(256, 8)$. Training hyper-parameters for the AEs are listed in Table 6.2. By comparing Table 6.1 and Table 6.2, it is obvious that the AEs need hundreds of thousands of training data, while the proposed DNN only requires one training batch with size M due to the only input $\mathbf{R}$. Therefore, compared to the data-driven AEs, the novel approach is more effective when labeled data are scarce in the system. The novel approach and AEs are regarded as optimization methods and only involved

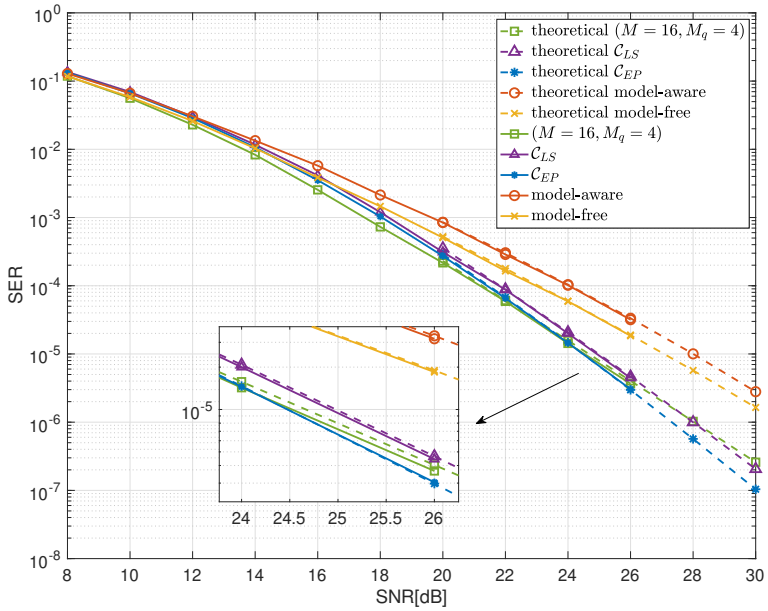


FIGURE 6.2 – SER of different codebooks in system $(16, 4)$.

in the off-line training. Thus, after training, all the learned codebooks are made into look-up tables at the transmitter and the ML algorithm Eq.(2.30) is applied at the receiver as it can exploit all the benefits of the SSD.

6.4.1 SER performance

Firstly, SER performance is studied under the assumption of perfect CSI knowledge at the receiver. In this case, ML algorithm in Eq.(2.30) can be rewritten as $\hat{\mathbf{x}} = \arg\max_{\mathbf{x}} p(\mathbf{y}|\mathbf{x}, \tilde{\mathbf{h}})$.

Fig. 6.2 compares the SER performance of different MDCs in the system $(16, 4)$. For high SNR values, theoretical SEP instead of simulation SER is considered. Codebook $\mathcal{C}_{\text{EP}}$ and $\mathcal{C}_{\text{LS}}$ both outperform the codebooks learned by the AEs when SNR $\geqslant$ 7dB. They approximately have 2dB and 2.5dB performance gain over the codebook learned by the model-free and model-aware AE at SER $= 10^{-5}$, respectively. Codebook $\mathcal{C}_{\text{LS}}$ has negligible SER degradation compared to codebook $(M = 16, M_q = 4)$ [67] at high SNRs. Moreover, codebook $\mathcal{C}_{\text{EP}}$ outperforms codebook $(M = 16, M_q = 4)$ when SNR $>$ 24dB.

Fig. 6.3 shows the SER performance of different codebooks in the system $(256, 8)$. The SER performance of codebook $\mathcal{C}_{\text{EP}}$ and $\mathcal{C}_{\text{LS}}$ is nearly the same. They outperform the resulting codebook of the model-free AE 1.6dB and that of model-aware AE 2dB, when SNR is equal to 20dB. In high SNR

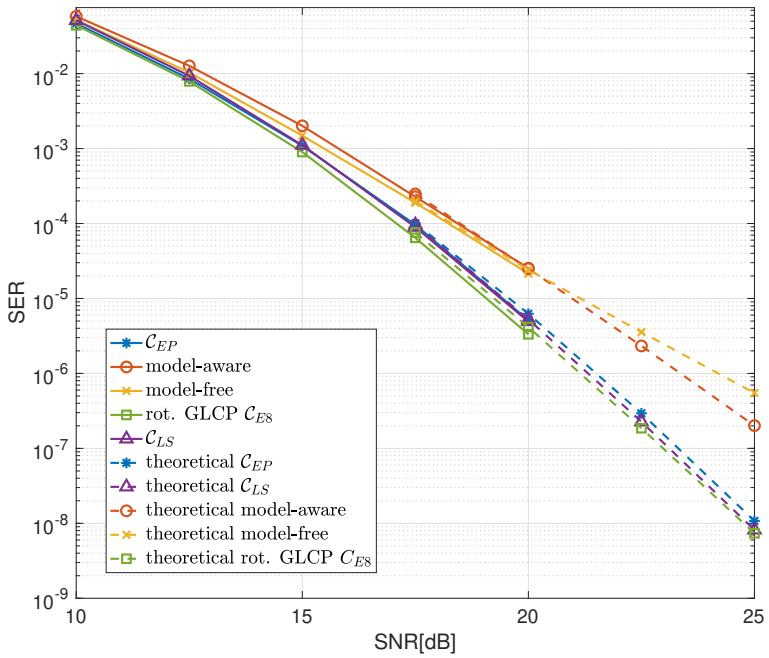


FIGURE 6.3 – SER of different codebooks in system $(256, 8)$.

region ($20\text{dB} \leqslant \text{SNR} \leqslant 25\text{dB}$), performance of the codebooks learned by the novel approach is very close to that of the codebook rot. GLCP $\mathcal{C}_{E8}$ [35]. Considering that codebook rot. GLCP $\mathcal{C}_{E8}$ shows prominent SER performance in the performance evaluation of paper[35], the proposed scheme is an efficient approach.

Secondly, in order to evaluate the robustness of various MDCs, SER performance is studied in CSI uncertainty scenario. In this scenario, the receiver has access to an estimate of $\tilde{\mathbf{h}}$ with elements corrupted by i.i.d. additive Gaussian noise expressed by $\mathcal{N}_c(0, \sigma_h^2)$. We consider $\sigma_h^2 = 0.05$ in the simulation. Fig. 6.4 illustrates the SER performance of MDCs with different CSI conditions in the system $(256, 8)$. Because of the CSI uncertainty, SER performance of codebook $\mathcal{C}_{\text{EP}}$, $\mathcal{C}_{\text{LS}}$ and rot. GLCP $\mathcal{C}_{E8}$ deteriorates about 0.5dB. While SER performance of the codebooks learned by the AEs is degraded by 0.75dB in CSI uncertainty scenario. It implies that the MDCs obtained using the proposed approach have the same robustness level compared to the ones obtained by the conventional approaches.

6.4.2 FoMs of MDCs

The essential FoMs of different MDCs in the two simulation systems are listed in Table 6.3. For the system $(16, 4)$, the resulting MCDs of the AEs have very low MPD values, thus their SER performance

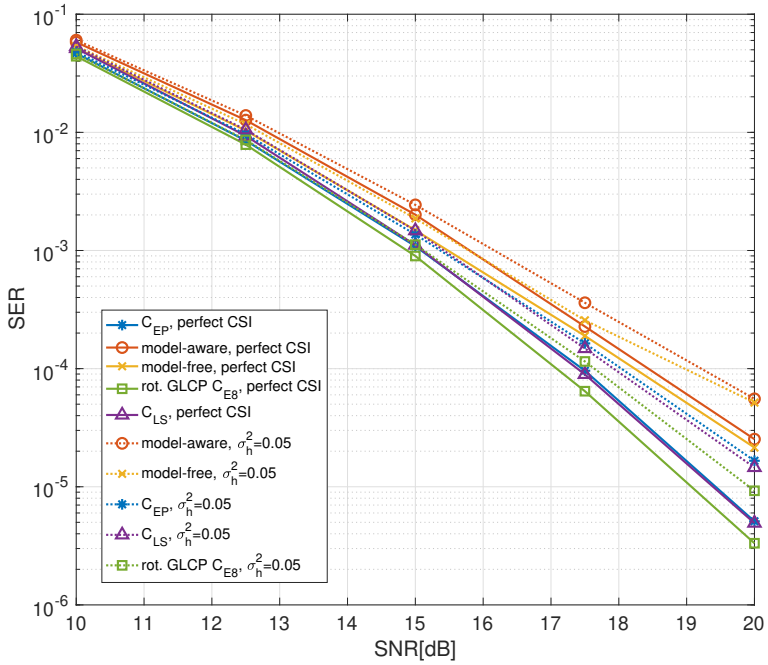


FIGURE 6.4 – SER comparison of different CSI scenarios in system $(256, 8)$.

TABLE 6.3 – FoM of codebooks in different systems.

System $(16, 4)$				System $(256, 8)$			
MDC	MED	MPD	SSD	MDC	MED	MPD	SSD
f_{EP}	1.248	0.016	4	f_{EP}	1.660	5.386×10^{-8}	6
f_{LS}	0.946	0.003	4	f_{LS}	1.025	1.442×10^{-8}	7
Model-aware	1.328	5.879×10^{-4}	4	Model-aware	1.055	1.339×10^{-9}	6
Model-free	1.421	5.549×10^{-5}	4	Model-free	1.359	5.873×10^{-9}	7
$(M = 16, M_q = 4)$	1.261	0.158	3	rot. GLCP $\mathcal{C}_{E8}$	1.952	3.308×10^{-5}	7

in higher SNR regime is much worse than that of other MDCs. Although codebook $(M = 16, M_q = 4)$ [67] has the highest MED and MPD, codebook $\mathcal{C}_{\text{EP}}$ still outperforms it at high SNR values. The reason is that codebook $\mathcal{C}_{\text{EP}}$ has high SSD than codebook $(M = 16, M_q = 4)$. It verifies the importance of exploiting SSD in MDC design. For the system $(256, 8)$, MPD of the codebooks $\mathcal{C}_{\text{EP}}$ and $\mathcal{C}_{\text{LS}}$ are significantly higher than that of codebooks learned by the AEs. The reason might be that the training of the AEs does not consider the specific requirements of MDC design over Rayleigh fading channels. As codebook $\mathcal{C}_{\text{LS}}$ has higher SSD than codebook $\mathcal{C}_{\text{EP}}$, the former has lower SER, especially at high SNR values. SSD of codebook rot. GLCP $\mathcal{C}_{E8}$ is less than 8 (full diversity) because of the threshold used in Eq.(6.3) $(\delta = 0.001)$.

6.5 Conclusion

In this chapter, a novel approach, consisting of a fairly simpler DNN structure compared to AE techniques, is proposed for MDC design over Rayleigh fading channels. Two loss functions are devised for the proposed approach. For performance evaluation, the proposed approach is compared with data-driven AEs and conventional approaches. The novel approach with both loss functions can significantly outperform AEs in terms of SER. In the simulation, it shows better or very similar performance than conventional approaches depending on the system settings or the adopted loss function. Furthermore, the FoM investigation demonstrates that the codebooks learned by the novel approach have much better MPD than those learned by AEs. Last but not least, the proposed approach is not constrained by the training dataset.

Chapitre 7

Conclusion and perspectives

Contenu

7.1 Conclusion

This thesis covers the study of three potential enabling techniques for the 6G, which are SCMA, NC MIMO and SU multidimensional transmission. Specifically, we have worked on the MUD for different SCMA systems and have explored DL-based techniques of the scheme design for NC MIMO and SU multidimensional transmission. A comprehensive introduction regarding the three techniques of technical backgrounds, system model and state-of-the-art is presented in chapter 2.

First of all, we have focused on iterative MUD algorithms for SCMA systems, where two transmission scenarios are considered. In the first scenario, we have proposed a low-complexity iterative algorithm called FO-GAA for the simple coded SCMA transmission. It has the lowest complexity among several simplified algorithms and a very fast convergence, meanwhile its BER performance loss is acceptable. In the second scenario, a highly overload SCMA transmission is introduced where two groups of SCMA share the same REs. At the receiver, we have employed iterative MUD algorithms for SCMA and SIC simultaneously. In view of the trade-off between MUD performance and spectral efficiency, we have proposed an optimization method of LDPC channel code rates given convergence requirement by analyzing the EXIT chart of the transmission. The highly overload scheme shows substantial spectral efficiency improvement and good BER performance in comparison to the conventional SCMA scheme.

Secondly, we have reviewed the state-of-the-art SD-based detectors for SCMA systems. Among them, SD-SCMA and GSD-SCMA are the most promising as they are applicable to the generalized SCMA transmission. We have elaborated on these two state-of-the-art. As a contribution, we have introduced SQRD and Schnorr-Euchner enumeration to further enhance their decoding efficiency and accelerate the tree search speed. However, SD-SCMA requires a constraint on the codebook power and GSD-SCMA computational complexity grows dramatically with overloading factor and codebook size. We have demonstrated an efficient pruning methodology in an example of a simple overload transmission. To address the complexity issue of GSD-SCMA, we have proposed the SGSD-SCMA detector which employs the pruning methodology to GSD-SCMA. Accordingly, two pruning algorithms are proposed for the SGSD-SCMA.Besides, we have derived the theoretical error rates of the proposed pruning algorithms under an assumption of the average number of projection of the codebook in each real dimension. Regarding the simulation performance, we have shown that the propose SGSD-SCMA

is especially advantageous for codebooks with moderate overloading factor and large size.

Next, we have discussed the Grassmannian constellation construction for NC MIMO systems.Seeing the particularity of AEs, we have proposed AE-FC consisting of fully-connected DNN as the encoder and decoder as an approach to construct Grassmannian constellation. To guarantee the restriction of Grassmannian constellations and carry out the back propagation of the training at the same time, we have introduced an orthonormalization process in the encoder. On the other hand, we have pointed out the similarity between the calculation of the GLRT detector and 2D convolutional layer. We have implemented GLRT on a CNN which is named CNN-GLRT. By applying the CNN-GLRT as the decoder, we have proposed AE-GLRT, the second scheme to construct Grassmannian constellations. The performance evaluation illustrates that the proposed AE-based approaches outperforms the conventional approaches in terms of the resulting codebook SER performance.

Finally, we have studied the MDCs design in the SU multidimensional transmission. We have analyzed some significant MoFs for MDCs in Rayleigh fading channels. Contrary to other AE-based schemes, we have proposed a new simple DL structure for MDCs design and have devised two loss functions dedicated to the new scheme. The proposed scheme performs the training driven by the desired FoMs or performance. As a result, the proposed DL structure is not data-hungry even suitable for data scarcity. The resulting MDCs of the proposed scheme shows substantial outperformance over the MDCs designed by AEs or conventional schemes.

7.2 Perspectives and future works

As we are finalizing this thesis, there have been various novel transmission schemes and cutting-edge proposals for the future wireless networks. For the continuity of this thesis, we hereunder provide some possible extensions that have not been addressed in the thesis and perspectives that can be developed in the future

1 DL-aided SD-based detector for SCMA : In [83, 116], DL techniques are applied to predict the initial squared radius of SD which helps to accelerate the tree search and effectively decrease the computational complexity. As we have proposed a low-complexity SD-based detector for SCMA named SGSD-SCMA, it is interesting to extend the study to the DL-aided SD-based detectors for SCMA in order to further enhance the MUD performance and/or reduce computational

complexity.

2 Codebook design for SD-based SCMA detectors : There have been research on developing low-projection SCMA codebooks [63, 117] in order to decrease the computational complexity for the MPA detector. However, we have revealed that low-projection SCMA codebooks deteriorate the error rate of the proposed pruning algorithms as well as the SGSD-SCMA detector. This implies that the desirable SCMA codebooks for SD-based detectors may be very different from those for the MPA detector. It would be wise to investigate the particular FoMs and approaches of the codebooks design for SD-based SCMA detectors.

3 Grant-free SCMA (GF-SCMA) : In order to carter to sporadic mMTC and reduce signaling overhead, UL grant-free transmission scheme has been taken into account in 3GPP NR [118, 119]. Due to the non-orthogonality nature, SCMA has been considered in grant-free transmission scheme, called GF-SCMA. Because grant-free scheme is contention-based, joint MPA (JMPA) carrying out joint active user detection and data decoding is investigated [120, 121]. Authors in [122] further integrate channel estimation in the receiver. Still, there are challenges of GF-SCMA regarding the design of contention transmission units, collision detection and etc. For future work, we could start from studying the receiver of GF-SCMA and then explore the solutions for other challenges.

4 NC MIMO in practical wireless networks : In [123], the authors compare the performance of NC MIMO schemes and the coherent MIMO schemes over a temporally-correlated channel. The simulation results show that the NC MIMO with Grassmannian signaling equipped with large number of transmit antennas outperforms its coherent counterparts in scenarios with mobility at high SNR values. This suggests that NC MIMO is promising for vehicular communications [123]. Moreover, the NC MIMO for block-fading channels is anticipated to arise in future vehicle-to-vehicle and airborne communication networks in order to handle the high mobility and rapid channel and demand variations [124]. However, there are still many challenges of the NC MIMO for practical implementations : such as labeling, low-complexity detection and peak-to-average power ratio reduction. Thus, it is interesting to study the NC MIMO for future 6G scenarios in practice.

Bibliographie

[1] W. Saad, M. Bennis et M. Chen, "A vision of 6g wireless systems : Applications, trends, technologies, and open research problems," *IEEE Network*, vol. 34, n°. 3, p. 134–142, 2020.

[2] L. Bariah, L. Mohjazi, S. Muhaidat, P. C. Sofotasios, G. K. Kurt, H. Yanikomeroglu et O. A. Dobre, "A prospective look : Key enabling technologies, applications and open research topics in 6g networks," *IEEE Access*, vol. 8, p. 174 792–174 820, 2020.

[3] Y. Chen, A. Bayesteh, Y. Wu, B. Ren, S. Kang, S. Sun, Q. Xiong, C. Qian, B. Yu, Z. Ding, S. Wang, S. Han, X. Hou, H. Lin, R. Visoz et R. Razavi, "Toward the standardization of non-orthogonal multiple access for next generation wireless networks," *IEEE Communications Magazine*, vol. 56, n°. 3, p. 19–27, 2018.

[4] 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.

[5] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland et F. Tufvesson, "6g wireless systems : Vision, requirements, challenges, insights, and opportunities," *Proceedings of the IEEE*, vol. 109, n°. 7, p. 1166–1199, 2021.

[6] Lizhong Zheng et D. N. C. Tse, "Communication on the Grassmann manifold : a geometric approach to the noncoherent multiple-antenna channel," *IEEE Transactions on Information Theory*, vol. 48, n°. 2, p. 359–383, 2002.

[7] Z. Liu, Y. Xin et G. Giannakis, "Linear constellation precoding for ofdm with maximum multipath diversity and coding gains," *IEEE Transactions on Communications*, vol. 51, n°. 3, p. 416–427, 2003.

[8] O. Simeone, "A very brief introduction to machine learning with applications to communication systems," *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, n°. 4, p.

648–664, 2018.

[9] D. Gündüz, P. de Kerret, N. D. Sidiropoulos, D. Gesbert, C. R. Murthy et M. van der Schaar, "Machine learning in the air," *IEEE Journal on Selected Areas in Communications*, vol. 37, n°. 10, p. 2184–2199, 2019.

[10] T. Gruber, S. Cammerer, J. Hoydis et S. t. Brink, "On deep learning-based channel decoding," dans *2017 51st Annual Conference on Information Sciences and Systems (CISS)*, 2017, p. 1–6.

[11] F. Liang, C. Shen et F. Wu, "An iterative bp-cnn architecture for channel decoding," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, n°. 1, p. 144–159, 2018.

[12] N. Samuel, T. Diskin et A. Wiesel, "Learning to detect," *IEEE Transactions on Signal Processing*, vol. 67, n°. 10, p. 2554–2564, 2019.

[13] N. Shlezinger, R. Fu et Y. C. Eldar, "Deep soft interference cancellation for mimo detection," dans *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, p. 8881–8885.

[14] H. He, C.-K. Wen, S. Jin et G. Y. Li, "Deep learning-based channel estimation for beamspace mmwave massive mimo systems," *IEEE Wireless Communications Letters*, vol. 7, n°. 5, p. 852–855, 2018.

[15] C. She, C. Sun, Z. Gu, Y. Li, C. Yang, H. V. Poor et B. Vucetic, "A tutorial on ultrareliable and low-latency communications in 6g : Integrating domain knowledge into deep learning," *Proceedings of the IEEE*, vol. 109, n°. 3, p. 204–246, 2021.

[16] A. Jagannath, J. Jagannath et T. Melodia, "Redefining wireless communication for 6g : Signal processing meets deep learning with deep unfolding," *IEEE Transactions on Artificial Intelligence*, p. 1–1, 2021.

[17] X. Meng, Y. Wu, Y. Chen et M. Cheng, "Low complexity receiver for uplink scma system via expectation propagation," dans *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, 2017, p. 1–5.

[18] X. Fu, M. Pischella et D. Le Ruyet, "On gaussian approximation algorithms for scma," dans *2019 16th International Symposium on Wireless Communication Systems (ISWCS)*, 2019, p. 155–160.

[19] L. Li, J. Wen, X. Tang et C. Tellambura, "Modified sphere decoding for sparse code multiple access," *IEEE Communications Letters*, vol. 22, n°. 8, p. 1544–1547, 2018.

[20] G. Chen, J. Dai, K. Niu et C. Dong, “Optimal receiver design for scma system,” dans *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2017, p. 1–6.

[21] M. Vameghestahbanati, E. Bedeer, I. Marsland, R. H. Gohary et H. Yanikomeroglu, “Enabling sphere decoding for scma,” *IEEE Communications Letters*, vol. 21, n°. 12, p. 2750–2753, 2017.

[22] M. Vameghestahbanati, I. Marsland, R. H. Gohary et H. Yanikomeroglu, “A novel sd-based detection for generalized scma constellations,” *IEEE Transactions on Vehicular Technology*, vol. 68, n°. 10, p. 10 278–10 282, 2019.

[23] B. M. Hochwald et T. L. Marzetta, “Unitary space-time modulation for multiple-antenna communications in rayleigh flat fading,” *IEEE Transactions on Information Theory*, vol. 46, n°. 2, p. 543–564, 2000.

[24] T. L. Marzetta et B. M. Hochwald, “Capacity of a mobile multiple-antenna communication link in rayleigh flat fading,” *IEEE transactions on Information Theory*, vol. 45, n°. 1, p. 139–157, 1999.

[25] A. Barg et D. Nogin, “Bounds on packings of spheres in the grassmann manifold,” *IEEE Transactions on Information Theory*, vol. 48, n°. 9, p. 2450–2454, 2002.

[26] R. H. Gohary et T. N. Davidson, “Noncoherent mimo communication : Grassmannian constellations and efficient detection,” *IEEE Transactions on Information Theory*, vol. 55, n°. 3, p. 1176–1205, 2009.

[27] D. Agrawal, T. Richardson et R. Urbanke, “Multiple-antenna signal constellations for fading channels,” *IEEE Transactions on Information Theory*, vol. 47, n°. 6, p. 2618–2626, 2001.

[28] Shengli Zhou et Baosheng Li, “Ber criterion and codebook construction for finite-rate precoded spatial multiplexing with linear receivers,” *IEEE Transactions on Signal Processing*, vol. 54, n°. 5, p. 1653–1665, 2006.

[29] Y. Wu, K. Ruotsalainen et M. Juntti, “Unitary space–time constellation design based on the chernoff bound of the pairwise error probability,” *IEEE transactions on information theory*, vol. 54, n°. 8, p. 3842–3850, 2008.

[30] I. Kammoun et J. . Belfiore, “A new family of grassmann space-time codes for non-coherent mimo systems,” *IEEE Communications Letters*, vol. 7, n°. 11, p. 528–530, 2003.

[31] J. Boutros et E. Viterbo, “Signal space diversity : a power- and bandwidth-efficient diversity technique for the rayleigh fading channel,” *IEEE Transactions on Information Theory*, vol. 44, n°. 4, p. 1453–1467, 1998.

[32] J. Boutros, E. Viterbo, C. Rastello et J. . Belfiore, “Good lattice constellations for both rayleigh fading and gaussian channels,” *IEEE Transactions on Information Theory*, vol. 42, n°. 2, p. 502–518, 1996.

[33] J. C. Inácio, B. F. Uchôa-Filho et D. Le Ruyet, “Grouped linear constellation-precoding for subcarrier index modulation ofdm,” dans *2017 International Symposium on Wireless Communication Systems (ISWCS)*, 2017, p. 308–313.

[34] J. C. Inácio, B. F. Uchôa-Filho et D. Le Ruyet, “Exploiting signal space diversity in ofdm with grouped subcarriers : Going beyond subcarrier index modulation,” *IEEE Wireless Communications Letters*, vol. 7, n°. 4, p. 650–653, 2018.

[35] J. C. Inácio, B. F. Uchôa-Filho, D. Le Ruyet et S. Montejo-Sánchez, “Full diversity multidimensional codebook design for fading channels : The combinatorial approach,” *IEEE Transactions on Communications*, vol. 68, n°. 7, p. 4104–4116, 2020.

[36] Y. Zhou, Q. Yu, W. Meng et C. Li, “Scma codebook design based on constellation rotation,” dans *2017 IEEE International Conference on Communications (ICC)*, May 2017, p. 1–6.

[37] S. ten Brink, “Convergence of iterative decoding,” *Electronics Letters*, vol. 35, n°. 10, p. 806–808, May 1999.

[38] 3GPP, “Multiplexing and channel coding,” n°. 38.212, 07 2018, version 15.2.0.

[39] H. Nikopour et H. Baligh, “Sparse code multiple access,” dans *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Sep. 2013, p. 332–336.

[40] M. Taherzadeh, H. Nikopour, A. Bayesteh et H. Baligh, “Scma codebook design,” dans *2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall)*, 2014, p. 1–5.

[41] J. Bao, Z. Ma, G. K. Karagiannidis, M. Xiao et Z. Zhu, “Joint multiuser detection of multidimensional constellations over fading channels,” *IEEE Transactions on Communications*, vol. 65, n°. 1, p. 161–172, 2017.

[42] F. A. Aoudia et J. Hoydis, "Model-free training of end-to-end communication systems," *IEEE Journal on Selected Areas in Communications*, vol. 37, n°. 11, p. 2503–2516, 2019.

[43] 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.

[44] M. Series, "Minimum requirements related to technical performance for imt-2020 radio interface (s)," *Report*, p. 2410–0, 2017.

[45] R. Li, "Network 2030 : Market drivers and prospects," dans *Proc. 1st International Telecommunication Union Workshop on Network*, vol. 2030, 2018.

[46] M. Giordani, M. Polese, M. Mezzavilla, S. Rangan et M. Zorzi, "Toward 6g networks : Use cases and technologies," *IEEE Communications Magazine*, vol. 58, n°. 3, p. 55–61, 2020.

[47] C. Han, Y. Wu, Z. Chen *et al.*, "Network 2030 a blueprint of technology, applications and market drivers towards the year 2030 and beyond," 2018.

[48] R. Hoshyar, F. P. Wathan et R. Tafazolli, "Novel low-density signature for synchronous cdma systems over awgn channel," *IEEE Transactions on Signal Processing*, vol. 56, n°. 4, p. 1616–1626, 2008.

[49] 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.

[50] C. Xu, N. Ishikawa, R. Rajashekar, S. Sugiura, R. G. Maunder, Z. Wang, L.-L. Yang et L. Hanzo, "Sixty years of coherent versus non-coherent tradeoffs and the road from 5g to wireless futures," *IEEE Access*, vol. 7, p. 178 246–178 299, 2019.

[51] B. Hochwald et T. Marzetta, "Unitary space-time modulation for multiple-antenna communications in rayleigh flat fading," *IEEE Transactions on Information Theory*, vol. 46, n°. 2, p. 543–564, 2000.

[52] B. Hochwald et W. Sweldens, "Differential unitary space-time modulation," *IEEE Transactions on Communications*, vol. 48, n°. 12, p. 2041–2052, 2000.

[53] G. Caire, G. Taricco et E. Biglieri, "Bit-interleaved coded modulation," *IEEE Transactions on Information Theory*, vol. 44, n°. 3, p. 927–946, 1998.

[54] C.-Y. Yang et M.-K. Ku, "Ldpc coded ofdm modulation for high spectral efficiency transmission," dans *2008 4th European Conference on Circuits and Systems for Communications*, 2008, p. 280–284.

[55] Y. Hori et H. Ochiai, "Performance analysis and interleaver structure optimization for short-frame bicm–ofdm systems," *IEEE Transactions on Wireless Communications*, vol. 15, n°. 1, p. 651–662, 2016.

[56] A. Yazar, S. Dogan-Tusha et H. Arslan, "6g vision : An ultra-flexible perspective," *ITU Journal on Future and Evolving Technologies*, vol. 1, n°. 9, p. 1–20, 2020.

[57] A. M. Jaradat, J. M. Hamamreh et H. Arslan, "Modulation options for ofdm-based waveforms : Classification, comparison, and future directions," *IEEE Access*, vol. 7, p. 17 263–17 278, 2019.

[58] E. Basar, U. Aygolu, E. Panayırcı et H. V. Poor, "Orthogonal frequency division multiplexing with index modulation," *IEEE Transactions on Signal Processing*, vol. 61, n°. 22, p. 5536–5549, 2013.

[59] R. Fan, Y. J. Yu et Y. L. Guan, "Generalization of orthogonal frequency division multiplexing with index modulation," *IEEE Transactions on Wireless Communications*, vol. 14, n°. 10, p. 5350–5359, 2015.

[60] C.-X. Wang, M. D. Renzo, S. Stanczak, S. Wang et E. G. Larsson, "Artificial intelligence enabled wireless networking for 5g and beyond : Recent advances and future challenges," *IEEE Wireless Communications*, vol. 27, n°. 1, p. 16–23, 2020.

[61] J. Jagannath, N. Polosky, A. Jagannath, F. Restuccia et T. Melodia, "Machine learning for wireless communications in the internet of things : A comprehensive survey," *Ad Hoc Networks*, vol. 93, p. 101913, 2019. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S1570870519300812

[62] M. Vameghestahbanati, I. D. Marsland, R. H. Gohary et H. Yanikomeroglu, "Multidimensional constellations for uplink scma systems—a comparative study," *IEEE Communications Surveys Tutorials*, vol. 21, n°. 3, p. 2169–2194, 2019.

[63] B. Fontana da Silva, D. Silva, B. F. Uchôa-Filho et D. Le Ruyet, "A multistage method for scma codebook design based on mds codes," *IEEE Wireless Communications Letters*, vol. 8, n°. 6, p. 1524–1527, Dec 2019.

[64] M. Gao, W. Ge, P. Zhang et Y. Zhang, "An efficient codebook design for uplink scma," *IEEE Access*, vol. 8, p. 211 665–211 675, 2020.

[65] L. Yu, P. Fan, D. Cai et Z. Ma, "Design and analysis of scma codebook based on star-qam signaling constellations," *IEEE Transactions on Vehicular Technology*, vol. 67, n°. 11, p. 10 543–10 553, 2018.

[66] K. Xiao, B. Xia, Z. Chen, B. Xiao, D. Chen et S. Ma, "On capacity-based codebook design and advanced decoding for sparse code multiple access systems," *IEEE Transactions on Wireless Communications*, vol. 17, n°. 6, p. 3834–3849, 2018.

[67] J. Bao, Z. Ma, M. Xiao, T. A. Tsiftsis et Z. Zhu, "Bit-interleaved coded scma with iterative multiuser detection : Multidimensional constellations design," *IEEE Transactions on Communications*, vol. 66, n°. 11, p. 5292–5304, 2018.

[68] C. Jiang et Y. Wang, "An uplink scma codebook design combining probabilistic shaping and geometric shaping," *IEEE Access*, vol. 8, p. 76 726–76 736, 2020.

[69] F. R. Kschischang, B. J. Frey, H.-A. Loeliger *et al.*, "Factor graphs and the sum-product algorithm," *IEEE Transactions on information theory*, vol. 47, n°. 2, p. 498–519, 2001.

[70] T. Minka *et al.*, "Divergence measures and message passing," Technical report, Microsoft Research, Rapport technique, 2005.

[71] W. B. Ameur, P. Mary, M. Dumay, J.-F. Hélard et J. Schwoerer, "Performance study of mpa, log-mpa and max-log-mpa for an uplink scma scenario," dans *2019 26th International Conference on Telecommunications (ICT)*, 2019, p. 411–416.

[72] B. F. d. Silva, D. Le Ruyet et B. F. Uchôa-Filho, "Threshold-based edge selection mpa for scma," *IEEE Transactions on Vehicular Technology*, vol. 69, n°. 3, p. 2957–2966, 2020.

[73] Y. Wang et L. Qiu, "Edge selection-based low complexity detection scheme for scma system," dans *2016 IEEE 84th Vehicular Technology Conference (VTC-Fall)*, 2016, p. 1–5.

[74] E. Viterbo et J. Boutros, "A universal lattice code decoder for fading channels," *IEEE Transactions on Information Theory*, vol. 45, n°. 5, p. 1639–1642, 1999.

[75] A. Chan et I. Lee, "A new reduced-complexity sphere decoder for multiple antenna systems," dans *2002 IEEE International Conference on Communications. Conference Proceedings. ICC 2002 (Cat. No.02CH37333)*, vol. 1, 2002, p. 460–464 vol.1.

[76] S. Yang et L. Hanzo, "Fifty years of mimo detection : The road to large-scale mimos," *IEEE Communications Surveys Tutorials*, vol. 17, n°. 4, p. 1941–1988, 2015.

[77] F. Wei et W. Chen, "Low complexity iterative receiver design for sparse code multiple access," *IEEE Transactions on Communications*, vol. 65, n°. 2, p. 621–634, 2017.

[78] B. Özbek et D. Le Ruyet, *Feedback strategies for wireless communication.* Springer, 2014.

[79] Z. Liu, Y. Xin et G. Giannakis, "Linear constellation-precoding for ofdm with maximum multipath diversity and coding gains," dans *Conference Record of Thirty-Fifth Asilomar Conference on Signals, Systems and Computers (Cat.No.01CH37256)*, vol. 2, 2001, p. 1445–1449 vol.2.

[80] M. Karlsson et E. Agrell, "Multidimensional modulation and coding in optical transport," *Journal of Lightwave Technology*, vol. 35, n°. 4, p. 876–884, 2017.

[81] D. Gómez-Barquero, C. Douillard, P. Moss et V. Mignone, "DVB-NGH : The next generation of digital broadcast services to handheld devices," *IEEE Transactions on Broadcasting*, vol. 60, n°. 2, p. 246–257, 2014.

[82] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga et A. Lerer, "Automatic differentiation in pytorch," 2017.

[83] M. Mohammadkarimi, M. Mehrabi, M. Ardakani et Y. Jing, "Deep learning-based sphere decoding," *IEEE Transactions on Wireless Communications*, vol. 18, n°. 9, p. 4368–4378, 2019.

[84] E. Nachmani, Y. Be'ery et D. Burshtein, "Learning to decode linear codes using deep learning," dans *2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2016, p. 341–346.

[85] C. Lu, W. Xu, H. Shen, H. Zhang et X. You, "An enhanced scma detector enabled by deep neural network," dans *2018 IEEE/CIC International Conference on Communications in China (ICCC)*, 2018, p. 835–839.

[86] T. J. O'Shea, K. Karra et T. C. Clancy, "Learning to communicate : Channel auto-encoders, domain specific regularizers, and attention," dans *2016 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, 2016, p. 223–228.

[87] M. Stark, F. Ait Aoudia et J. Hoydis, "Joint learning of geometric and probabilistic constellation shaping," dans *2019 IEEE Globecom Workshops (GC Wkshps)*, 2019, p. 1–6.

[88] B. Karanov, "End-to-end deep learning in optical fibre communication systems," Thèse de doctorat, UCL (University College London), 2020.

[89] F. Alberge, "Deep learning constellation design for the awgn channel with additive radar interference," *IEEE Transactions on Communications*, vol. 67, n°. 2, p. 1413–1423, 2019.

[90] Z.-R. Zhu, J. Zhang, R.-H. Chen et H.-Y. Yu, "Autoencoder-based transceiver design for owc systems in log-normal fading channel," *IEEE Photonics Journal*, vol. 11, n°. 5, p. 1–12, 2019.

[91] J. Tao, J. Xing, J. Chen, C. Zhang et S. Fu, "Deep neural hybrid beamforming for multi-user mmwave massive mimo system," dans *2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2019, p. 1–5.

[92] S. Dörner, S. Cammerer, J. Hoydis et S. t. Brink, "Deep learning based communication over the air," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, n°. 1, p. 132–143, 2018.

[93] S. Cammerer, F. A. Aoudia, S. Dörner, M. Stark, J. Hoydis et S. ten Brink, "Trainable communication systems : Concepts and prototype," *IEEE Transactions on Communications*, vol. 68, n°. 9, p. 5489–5503, 2020.

[94] V. Dumoulin et F. Visin, "A guide to convolution arithmetic for deep learning," *arXiv preprint arXiv :1603.07285*, 2016.

[95] Y. Wu, S. Zhang et Y. Chen, "Iterative multiuser receiver in sparse code multiple access systems," dans *2015 IEEE International Conference on Communications (ICC)*, June 2015, p. 2918–2923.

[96] B. Ghani, F. Launay, J. P. Cances, C. Perrine et Y. Pousset, "Iterative decoding for scma systems using log-mpa with feedback ldpc decoding," dans *Ubiquitous Networking*, O. Habachi, V. Meghdadi, E. Sabir et J.-P. Cances, édit. Cham : Springer International Publishing, 2020, p. 18–31.

[97] Y. Wang, L. Qiu et X. Li, "An efficient low complexity gaussian approximation-based scheme for scma detection," dans *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*. IEEE, 2017, p. 1–5.

[98] X. Meng, Y. Wu, Y. Chen et M. Cheng, "Low complexity receiver for uplink scma system via expectation propagation," dans *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, March 2017, p. 1–5.

[99] D. Koller et N. Friedman, *Probabilistic Graphical Models : Principles and Techniques - Adaptive Computation and Machine Learning*. The MIT Press, 2009.

[100] T. P. Minka, "Expectation propagation for approximate bayesian inference," *arXiv preprint arXiv :1301.2294*, 2013.

[101] ——, "A family of algorithms for approximate bayesian inference," Thèse de doctorat, Massachusetts Institute of Technology, 2001.

[102] S. Wu, L. Kuang, Z. Ni, J. Lu, D. Huang et Q. Guo, "Low-complexity iterative detection for large-scale multiuser mimo-ofdm systems using approximate message passing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, n°. 5, p. 902–915, Oct 2014.

[103] N. I. Kim et D. Cho, "Hybrid multiple access system based on non orthogonality and sparse code," dans *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, March 2017, p. 1–6.

[104] S. Sharma, K. Deka, V. Bhatia et A. Gupta, "Joint power-domain and scma-based noma system for downlink in 5g and beyond," *IEEE Communications Letters*, vol. 23, n°. 6, p. 971–974, June 2019.

[105] M. Moltafet, N. Mokari, M. R. Javan, H. Saeedi et H. Pishro-Nik, "A new multiple access technique for 5g : Power domain sparse code multiple access (psma)," *IEEE Access*, vol. 6, p. 747–759, 2018.

[106] M. El-Hajjar et L. Hanzo, "Exit charts for system design and analysis," *IEEE Communications Surveys Tutorials*, vol. 16, n°. 1, p. 127–153, First 2014.

[107] S. Zhang, K. Xiao, B. Xiao, Z. Chen, B. Xia, D. Chen et S. Ma, "A capacity-based codebook design method for sparse code multiple access systems," dans *2016 8th International Conference on Wireless Communications Signal Processing (WCSP)*, 2016, p. 1–5.

[108] D. J. MacKay, *Information theory, inference and learning algorithms*. Cambridge university press, 2003.

[109] T. Cui et C. Tellambura, "An efficient generalized sphere decoder for rank-deficient mimo systems," *IEEE Communications Letters*, vol. 9, n°. 5, p. 423–425, 2005.

[110] D. Wubben, R. Bohnke, J. Rinas, V. Kuhn et K.-D. Kammeyer, "Efficient algorithm for decoding layered space-time codes," *Electronics letters*, vol. 37, n°. 22, p. 1348–1350, 2001.

[111] D. Wubben, R. Bohnke, V. Kuhn et K.-D. Kammeyer, "Mmse extension of v-blast based on sorted qr decomposition," dans *2003 IEEE 58th Vehicular Technology Conference. VTC 2003-Fall (IEEE Cat. No.03CH37484)*, vol. 1, 2003, p. 508–512 Vol.1.

[112] B. Shim et I. Kang, "Sphere decoding with a probabilistic tree pruning," *IEEE Transactions on Signal Processing*, vol. 56, n°. 10, p. 4867–4878, 2008.

[113] Z. Li, W. Chen, F. Wei, F. Wang, X. Xu et Y. Chen, "Joint codebook assignment and power allocation for scma based on capacity with gaussian input," dans *2016 IEEE/CIC International Conference on Communications in China (ICCC)*, 2016, p. 1–6.

[114] T. Contributors, "Torch.nn.functional.conv2d," 2019. [En ligne]. Disponible : https://pytorch.org/docs/stable/generated/torch.nn.functional.conv2d.html#torch.nn.functional.conv2d

[115] S. K. Kumar, "On weight initialization in deep neural networks," *arXiv preprint arXiv :1704.08863*, 2017.

[116] D. Weon et K. Lee, "Learning-aided deep path prediction for sphere decoding in large mimo systems," *IEEE Access*, vol. 8, p. 70 870–70 877, 2020.

[117] A. Bayesteh, H. Nikopour, M. Taherzadeh, H. Baligh et J. Ma, "Low complexity techniques for scma detection," dans *2015 IEEE Globecom Workshops (GC Wkshps)*, 2015, p. 1–6.

[118] Lenovo, "Uplink Grant-free Access for 5G mMTC," *3GPP TSG RAN WG1 meeting, Lisbon Portugal*, p. document R1–1 609 398, Oct 2016.

[119] CATT, "Consideration on grant-free transmission," *3GPP TSG RAN WG1 meeting, Lisbon Portugal*, p. document R1–1 608 757, Oct 2016.

[120] F. Wang, Y. Zhang, H. Zhao, H. Huang et J. Li, "Active user detection of uplink grant-free scma in frequency selective channel," dans *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*, 2018, p. 1–6.

[121] A. Bayesteh, E. Yi, H. Nikopour et H. Baligh, "Blind detection of scma for uplink grant-free multiple-access," dans *2014 11th International Symposium on Wireless Communications Systems (ISWCS)*, 2014, p. 853–857.

[122] F. Wei, W. Chen, Y. Wu, J. Ma et T. A. Tsiftsis, "Message-passing receiver design for joint channel estimation and data decoding in uplink grant-free scma systems," *IEEE Transactions on Wireless Communications*, vol. 18, n°. 1, p. 167–181, 2019.

[123] J. Cabrejas, S. Roger, D. Calabuig, Y. M. M. Fouad, R. H. Gohary, J. F. Monserrat et H. Yanikomeroglu, "Non-coherent open-loop mimo communications over temporally-correlated channels," *IEEE Access*, vol. 4, p. 6161–6170, 2016.

[124] R. H. Gohary et H. Yanikomeroglu, "Noncoherent mimo signaling for block-fading channels : Approaches and challenges," *IEEE Vehicular Technology Magazine*, vol. 14, n°. 1, p. 80–88, 2019.

Annexe A

Sphere decoding in an example MIMO scenario

Consider a real-valued multiple antenna system with N transmit antennas and N receive antennas. The received signal is expressed as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{A.1}$$

where $\mathbf{x} \in \mathbb{R}^{N\times 1}$ is the transmitted signal vector, $\mathbf{H} \in \mathbb{R}^{N\times N}$ is the channel matrix and $\mathbf{n} \in \mathbb{R}^{N\times 1}$ is AWGN vector whose entries follow the distribution $\mathcal{N}(0, \sigma^2)$. The entries of $\mathbf{x}$ are drawn independently from the codebook $\mathcal{X} \subset \mathbb{R}^{M\times 1}$. Therefore, the ML decoding problem is expressed as

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathcal{X}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 \tag{A.2}$$

Instead of performing exhaustive search, SD tests the transmitted signal points that are within a sphere of radius d centered around the received signal $\mathbf{y}$ denoted as $S(\mathbf{y}, d)$. SD can be regarded as ML with conditions, which can be given as

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathcal{X}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 \tag{A.3}$$

$$\text{subject to } \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 \leqslant d^2 \tag{A.4}$$

In order to guarantee a structure easy for tree search, we first perform QR-decomposition of $\mathbf{H}$ as

$$\mathbf{H} = \mathbf{Q}\mathbf{R} \tag{A.5}$$

where $\mathbf{R}$ is an upper triangular matrix with positive diagonal elements and $\mathbf{Q}$ is an orthogonal matrix.

As a result, the SD is rewritten as

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathcal{X}} \|\mathbf{Q}^{\intercal}\mathbf{y} - \mathbf{R}\mathbf{x}\|^2 \tag{A.6}$$

$$= \arg\min_{\mathbf{x}\in\mathcal{X}} \|\mathbf{y}' - \mathbf{R}\mathbf{x}\|^2 \tag{A.7}$$

$$\text{subject to } \|\mathbf{y}' - \mathbf{R}\mathbf{x}\|^2 \leqslant d^2 \tag{A.8}$$

where $\mathbf{y}' = \mathbf{Q}^{\intercal}\mathbf{y}$. Due to the upper triangular matrix $\mathbf{R}$, $\mathbf{y}' - \mathbf{R}\mathbf{x}$ has a structure as

$$\begin{bmatrix} y'_1 \\ y'_2 \\ \vdots \\ y'_N \end{bmatrix} - \begin{bmatrix} r_{1,1} & r_{1,2} & \cdots, & r_{1,N} \\ 0 & r_{2,2} & \cdots, & r_{2,N} \\ \vdots & \ddots & \vdots & \\ 0 & 0 & \cdots & r_{N,N} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \tag{A.9}$$

which enables the tree search. Hence, the condition is Eq.(A.8) can be expressed as

$$d^2 \geq \sum_{n=1}^{N} \left(y'_n - \sum_{l=n}^{N} r_{n,l} x_l \right)^2 \tag{A.10}$$

$$= (y'_N - r_{N,N}x_N)^2 + \left(y'_{N-1} - \sum_{l=N-1}^{N} r_{l,l} x_l \right)^2 + \cdots$$

We denote the terms in the right-hand side as the branch metrics. Since $\mathbf{R}$ is in an upper triangular format, it is natural to start the search from the bottom layer and trace upward. The recursive relationship at the n-th ($1 \leqslant n \leqslant N$) layer is

$$b_n = \left(y'_n - \sum_{l=n}^{N} r_{n,l} x_l \right)^2 \tag{A.11}$$

$$p_n = p_{n+1} + b_n \tag{A.12}$$

where b_n is the n-th layer branch metric and p_n is the n-th layer path metric. At the beginning, $P_{N+1} = 0$ is initialized. At the N-th layer (the bottom layer), the necessary condition for $\mathbf{R}\mathbf{x}$ being within the new sphere, $S(\mathbf{y}', d)$, is

$$(y'_N - r_{N,N}x_N)^2 \leqslant d^2 \tag{A.13}$$

The corresponding range of x_N is

$$\frac{y'_i - d}{r_{N,N}} \leqslant x_N \leqslant \frac{y'_i + d}{r_{N,N}} \tag{A.14}$$

To apply this to all the layers, we define upper bound u_n and lower bound l_n of the x_n. Consequently, at the n-th layer, the range of x_n becomes

$$l_n \leqslant x_n \leqslant u_n \tag{A.15}$$

where

$$l_i = \frac{1}{r_{n,n}}\left(y'_n - \sum_{l=n+1}^{2Jd_v} r_{n,l}x_l - \sqrt{d^2 - p_{n+1}}\right) \tag{A.16}$$

$$u_i = \frac{1}{r_{n,n}}\left(y'_n - \sum_{l=n+1}^{2Jd_v} r_{n,l}x_l + \sqrt{d^2 - p_{n+1}}\right) \tag{A.17}$$

$$\tag{A.18}$$

We define vector $\mathbf{e}_n$ to store the values of x_n satisfying the condition in Eq.(A.15) at the n-th layer, formulated as

$$\mathbf{e}_n = \mathcal{F}_{\text{enum}}(l_n, u_n, \mathcal{X}) \tag{A.19}$$

where $\mathcal{F}_{\text{enum}}(l_n, u_n, \mathcal{X})$ denotes the function enumerating all elements in $\mathcal{X}$ between l_n and u_n. We define s_n as the length of vector $\mathbf{e}_n$, equivalent to the number of candidates at n-th layer. To accelerate the tree search, we adopt Schnorr-Euchner (SE) strategy for candidates ordering in each layer [112]. In Schnorr-Euchner enumeration, the candidates are examined and sorted based on their path metric values in each layer. Recall the path metric expression in Eq.(A.12), p_{n+1} is identical for all child nodes having the same parent and the ordering is based on the branch metric. Specifically, after determining the candidates according to Eq.(A.19), we first calculate the branch metric values as

$$b_n(i) = \left(y'_n - \sum_{l=n+1}^{N} r_{n,l}x_l - r_{n,n}e_n(i)\right)^2 \tag{A.20}$$

where $i \in \{1, 2, \cdots, s_n\}$ and $e_n(i)$ refers to the i-th element of $\mathbf{e}_n$. Then, we reorder the elements of $\mathbf{e}_n$ in ascending order based on the associated values of $b_n(i)$. α_n is defined to represent the index of the chosen candidate in the n-th layer. At the n-th layer, the path metric p_i is updated by

$$p_n = p_{n+1} + b_n(\alpha_n) \tag{A.21}$$

The details of SD for real-valued MIMO is illustrated in flowchart Fig. A.1.

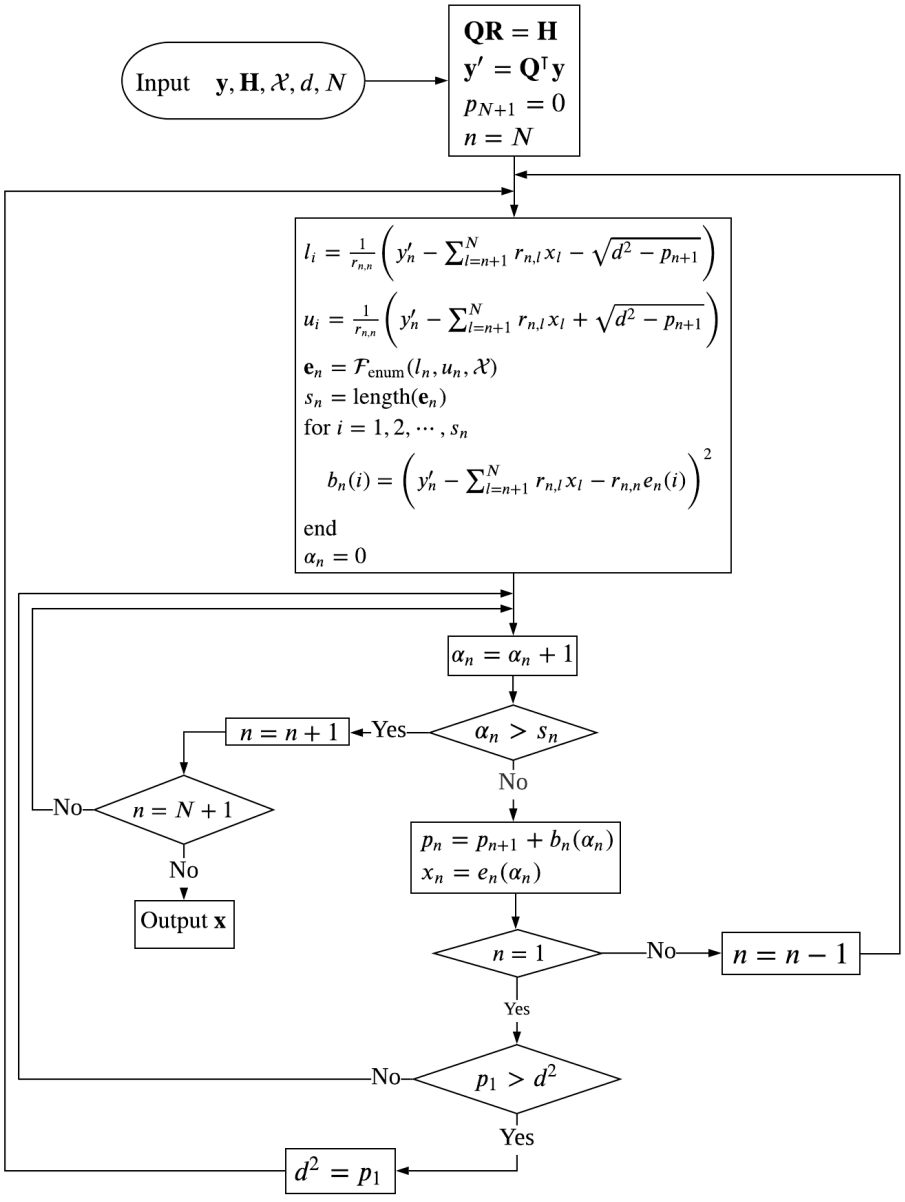


FIGURE A.1 – Flowchart of the example MIMO transmission.

le cnam

Xiaotian FU

Scheme Design and Advanced Receivers for SCMA and Non-coherent MIMO in Future Wireless Networks

HESAM UNIVERSITÉ

Résumé : Poussée par la demande de connectivité et le développement rapide du multimédia, la sixième génération (6G) de réseaux sans fil a récemment fait l'objet de discussions. Dans cette thèse, nous étudions trois schémas prometteurs pour la 6G, qui sont l'accès multiple par code creux (SCMA), la transmission MIMO (multiple-input multiple-output) non cohérente (NC) et la transmission multidimensionnelle à utilisateur unique (SU). Tout d'abord, nous travaillons sur les algorithmes itératifs de détection multi-utilisateurs (MUD) dans la transmission SCMA simple et fortement chargée. Par ailleurs, nous proposons deux algorithmes d'élagage et introduisons un détecteur SCMA basé sur le décodage par sphères à faible complexité. Ensuite, nous étudions la construction de constellations Grassmanniennes pour les systèmes MIMO NC. Deux auto-encodeurs sont proposés et évalués. Enfin, nous tirons parti des techniques DL et introduisons une nouvelle approche pour la conception de constellations multidimensionnelles dans la transmission multidimensionnelle SU.

Mots clés : SCMA, courbes EXIT, décodage par sphères, NC MIMO, constellation grassmannienne, dictionnaire multidimensionnel

Abstract : Driven by the demand of connectivity and fast development of multimedia, the sixth generation (6G) of wireless networks has recently been under discussion. In this thesis, we study three promising schemes for the 6G, which are sparse code multiple access (SCMA), non-coherent (NC) multiple-input multiple-output (MIMO) and single-user (SU) multidimensional transmission. Firstly, we work on the iterative multiuser detection (MUD) algorithms in both simple and highly overload SCMA transmission. Beside, we propose two pruning algorithms and introduce a low-complexity sphere decoding-based SCMA detector. Next, the Grassmannian constellation construction for NC MIMO systems is studied. Two autoencoder-based schemes are proposed and evaluated. Lastly, we take advantage of DL techniques and introduce a novel approach for designing multiudimensional constellations in the SU multidimensional transmission.

Keywords : SCMA, EXIT chart, sphere decoding, NC MIMO, Grassmannian constellation, multidimensional codebook