UNIVERSIDADE FEDERAL DE SANTA CATARINA

DEPARTAMENTO DE ENGENHARIA ELÉTRICA

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Bruno Fontana da Silva

Strategies for Multistage Design of Sparse Code Multiple Access Systems

Florianópolis, 09 de Dezembro de 2019.

BRUNO FONTANA DA SILVA

STRATEGIES FOR MULTISTAGE DESIGN OF SPARSE CODE MULTIPLE ACCESS SYSTEMS

Tese submetida à Universidade Federal de Santa Catarina como parte dos requisitos para a obtenção do grau de Doutor em Engenharia Elétrica.

Orientador: Bartolomeu Ferreira Uchôa Filho

Co-orientador: Danilo Silva

FLORIANÓPOLIS

2019

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Silva, Bruno Fontana da
Strategies for Multistage Design of Sparse Code Multiple Access Systems / Bruno Fontana da Silva ; orientador, Bartolomeu Ferreira Uchôa Filho, coorientador, Danilo Silva, 2019.
146 p.

Tese (doutorado) - Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia Elétrica, Florianópolis, 2019.

Inclui referências.

1. Engenharia Elétrica. 2. SCMA. 3. Multiple Access. 4. NOMA. 5. Message-Passing Algorithm. I. Uchôa Filho, Bartolomeu Ferreira. II. Silva, Danilo. III. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Engenharia Elétrica. IV. Título.

Bruno Fontana da Silva

STRATEGIES FOR MULTISTAGE DESIGN OF SPARSE CODE MULTIPLE ACCESS SYSTEMS

O presente trabalho em nível de doutorado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof. Cristiano Magalhães Panazio, Ph.D.

Universidade de São Paulo

Prof.ª Mylène Pischella, Ph.D.

Conservatoire National des Arts et Métiers

Prof. Rodrigo Caiado de Lamare, Dr.

Pontifícia Universidade Católica do Rio de Janeiro

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de Doutor em Engenharia Elétrica.

Bartolomeu Ferreira Uchoa Filho:47636211491
Digitally signed by Bartolomeu Ferreira Uchoa Filho:47636211491
Date: 2019.12.11 17:18:18 -03'00'

Bartolomeu Ferreira Uchôa Filho, Ph.D.

Coordenador do Programa de Pós-Graduação em Engenharia Elétrica

Universidade Federal de Santa Catarina

Bartolomeu Ferreira Uchoa Filho:47636211491
Digitally signed by Bartolomeu Ferreira Uchoa Filho:47636211491
Date: 2019.12.11 17:17:48 -03'00'

Orientador: Prof. Bartolomeu Ferreira Uchôa Filho, Ph.D.

Universidade Federal de Santa Catarina

Coorientador: Prof. Danilo Silva, Ph.D.

Universidade Federal de Santa Catarina

Florianópolis, 09 de Dezembro de 2019.

Para a minha avó,

Maria Regina Fontana do Amaral,

pelo tempo de ausência e pela distância

que este trabalho nos impôs.

Agradecimentos

Gostaria de registrar formalmente minha gratidão a todas as pessoas as quais contribuíram, direta ou indiretamente, para que este trabalho fosse concluído. Em particular,

aos meus pais, Tônio e Gilsonea, e a minha avó Maria Regina, pelos ensinamentos básicos da vida e pelo amor incondicional, essenciais para que eu siga dando os meus passos independentemente das dificuldades existentes no caminho;

à Verônica, pelo amor, companheirismo, e por trazer mais cores aos momentos da minha vida;

ao meu orientador Bartolomeu Ferreira Uchôa Filho, por sempre acreditar no meu trabalho, desde o início da minha vida acadêmica na pós-graduação, sempre me motivando a seguir em frente e me tornar um profissional cada vez melhor;

ao meu coorientador Danilo Silva, pelo apoio constante, pela paciência sempre renovável, e também pelas contribuições fundamentais para o progresso e melhoria das minhas pesquisas;

aos professores Didier Le Ruyet e Myléne Pischella, do *Conservatoire national des arts et métiers* (CNAM), e ao professor Cesar Azurdia-Meza, da *Universidad de Chile*, pelo acolhimento durante os períodos de estudos nos intercâmbios, pela atenção dispendida e por todas as colaborações para o meu trabalho e para minha formação como pesquisador;

aos colegas de laboratório do GPqCom, Juliana Inácio, Paulo Branco, Ana Scharf, José Clair e João Vinholi, e do CNAM, Luan Chen, Wafa Njima, Iandra Galdini e Dayse Correia, pela amizade, pelos cafés e conversas e pelas colaborações inerentes àqueles que trilham juntos os caminhos da academia;

ao amigo e colega do GPqCom, Ricardo Bohaczuk, em particular, que além de grande companheiro de bancada e dos cafezinhos, também prestou assistência técnica incondicional com o nosso servidor, ajuda sem a qual esse trabalho jamais seria possível;

aos grandes amigos Daniel Bertineti, Yuri Valadão, Vinícius Ludwig, Rodrigo Tiedt, Gustavo Baierle, Otávio Schmengler, Jean Diniz, Julie de Oliveira, Felipe Loose, João Otávio Cadó, Ptolomeu Palma, André Wyzykowski, Maria Cláudia e Alexandre Moreira, os quais acompanharam essa jornada em intensidades e proximidades diferentes, mas que possuem em comum um laço de amizade, confiança e respeito que eu espero jamais perder;

aos meus familiares, principalmente tios(as), primos(as), e em particular ao Diego Fontana, pelo apoio incondicinal de uma vida inteira, e pelos valiosos exemplos de vida;

aos antigos mestres, Renato Machado e Alexandre Campos, por plantarem as primeiras

sementes do meu interesse pela pesquisa científica, e por todas as bases e conselhos valiosos que deram início a essa longa jornada;

ao Instituto Federal Sul-rio-grandense, em particular aos colegas e alunos do Campus Sapiranga, pelo apoio e pela oportunidade de concluir este trabalho com a dedicação necessária;

à Universidade Federal de Santa Catarina e ao seu Programa de Pós-graduação em Engenharia Elétrica, pelo suporte à pesquisa científica nesta área e pela excelente qualidade exercida por todos os servidores e docentes, fundamentais para que eu pudesse desenvolver a minha formação;

ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pelos financiamentos diretos (141161/2016-7 e 204676/2018-5) e indiretos que apoiaram a realização deste trabalho.

“L’homme ne poursuit que des chimères”.

(Pierre-Simon Laplace)

Resumo da Tese apresentada à UFSC como parte dos requisitos
necessários para obtenção do grau de Doutor em Engenharia Elétrica

STRATEGIES FOR MULTISTAGE DESIGN OF SPARSE CODE MULTIPLE ACCESS SYSTEMS

Bruno Fontana da Silva

09 de Dezembro de 2019

Orientador: Bartolomeu Ferreira Uchôa Filho

Co-orientador: Danilo Silva

Área de concentração: Comunicações e Processamento de Sinais

Palavras-chave: SCMA, NOMA, acesso múltiplo, modulação multidimensional, detecção multiusuário

Número de páginas: 144

Múltiplo Acesso por Códigos Esparsos (SCMA) é um esquema de múltiplo acesso não-ortogonal (NOMA) por divisão de código. Neste esquema, J usuários compartilham K recursos ortogonais, sendo que o número de usuários é tipicamente maior do que o número de recursos, ou seja, $J > K$. Esse cenário é denominado sobrecarregado, de forma que é inevitável que ocorram colisões entre os usuários em alguns recursos. No SCMA, cada usuário espalha sua informação sobre os K recursos usando uma assinatura esparsa. Dessa maneira, o número de colisões em cada recurso é limitado a um número controlado e menor que o número total de usuários. A esparsidade das assinaturas permite o uso de um detector multiusuário baseado no algoritmo *message-passing* (MPA).

Nesta tese, são apresentadas estratégias multi-estágio para o projeto de sistemas SCMA *uplink*. Um estado da arte das pesquisas relacionadas é apresentado, segmentando a área em três categorias de trabalhos: projeto de *codebooks* SCMA, algoritmos de detecção multiusuário e métodos de alocação de recursos.

Uma das contribuições deste trabalho é um método de projeto de *codebooks* SCMA estruturados pelo espaço gerado por códigos de bloco. Com o método proposto, é possível projetar modulações M-árias N-dimensionais com diversidade de espaço de sinais $N - 1$ usando alfabetos de tamanho relativamente pequeno $q \ll M$. Os *codebooks* projetados apresentam bom desempenho de taxa de erro de bit e mantêm uma ordem de complexidade do MPA aceitável.

Uma segunda contribuição é um detector multiusuário baseado no MPA, utilizando seleção dinâmica de arestas e aproximação Gaussiana. Trata-se de uma modificação de um detector de baixa complexidade já conhecido na literatura, incluindo um limiar de seleção de arestas para ajustar o compromisso entre desempenho e ordem de complexidade de maneira paramétrica. Soluções analíticas para a ordem de complexidade média são derivadas para esse algoritmo. Os resultados mostram que é possível obter um desempenho de taxa-de-erro tão satisfatório quanto a solução original com uma ordem de complexidade média significativamente menor.

Por fim, a tese apresenta também uma análise sobre o projeto da alocação de recursos e, consequentemente, das assinaturas de espalhamento. Essa análise considera a capacidade do canal SCMA como figura de mérito. Em particular, a capacidade simétrica e o desempenho de sistemas com taxa simétrica são discutidos pra projetos com assinaturas regulares e irregulares. Neste contexto, considera-se um cenário celular no qual os canais possuem desvanecimentos de larga e de pequena escala.

Resumo Expandido

Introdução

Sistemas de múltiplo acesso não-ortogonal (NOMA) estão sendo considerados para atender aos requisitos do próximo padrão de telefonia celular (5G) [1, 2]. Dentre as categorias de NOMA, destacam-se NOMA por alocação de potência (PD) e por divisão de código (CD). Particularmente para CD-NOMA, um esquema que ganhou muita atenção da indústria e da comunidade científica foi o múltiplo acesso por códigos esparsos (SCMA) [3].

SCMA considera a transmissão simultânea de diversos usuários compartilhando um único meio. Em um sistema sobrecarregado (mais usuários do que recursos compartilhados), é inevitável que ocorram colisões entre os usuários. Para que seja possível separar a informação de cada usuário no receptor, utilizando um detector multiusuário (MUD), os esquemas SCMA utilizam assinaturas esparsas de espalhamento. Devido à esparsidade, o grau de colisão em cada recurso é controlado a um número muito menor que o total de usuários do sistema. Assim, é possível utilizar o algoritmo *message-passing* (MPA) como MUD [4], o que torna a complexidade de detecção relativamente baixa em relação à detecção ótima.

Cada usuário de um esquema SCMA espalha sua informação em apenas alguns recursos do total disponível. A informação é tratada como um símbolo (ou palavra-código) de uma modulação multidimensional (MDM), tipicamente denominada de *codebook* SCMA. O projeto de codebooks SCMA é discutido em diversos trabalhos [5–7]. Um codebook bem dimensionado possui propriedades que resultam em um bom desempenho em termos de taxa de erro. Projeto de *codebooks* SCMA é uma linha de pesquisa que ganhou destaque dentro da literatura.

Outro aspecto importante para o SCMA é o algoritmo para o MUD. Embora o MPA tenha uma complexidade relativamente baixa em sistemas de pequenas dimensões e com modulações pequenas, sua ordem de complexidade é exponencial no grau de colisão de cada recurso e polinomial no alfabeto das projeções complexas das MDMs. Portanto, qualquer aumento de eficiência espectral, seja pela sobrecarga (aumentando o grau de colisão) ou pela modulação (aumentando a taxa de transmissão) causa um aumento de complexidade significativo para o MPA. Alguns algoritmos MUD foram propostos para simplificar a detecção no SCMA, inclusive muitas variações do MPA [8–11].

Por fim, a própria alocação de recursos do MPA é uma etapa de projeto relevante, que afeta a capacidade do canal de múltiplo acesso [12–14]. A alocação de recurso pode ser interpretada

como o projeto de um grafo bipartido, semelhante ao projeto de códigos de matrizes de paridade de baixa densidade (LDPC) [15, 16]. Além disso, em um ambiente celular, o desvanecimento de larga escala e a alocação de potência necessitam de atenção durante o projeto, para que o sistema seja otimizado em relação à sua eficiência espectral e energética [17].

Objetivos

Este trabalho busca contribuir com métodos de projeto flexíveis e generalizados de sistemas SCMA *uplink*. Busca-se viabilizar o projeto de sistemas de grandes dimensões, com alta eficiência espectral e boa confiabilidade na comunicação.

Metodologia

O trabalho desenvolvido tem início com uma vasta análise do estado da arte das pesquisas relacionadas aos sistemas SCMA. Nesse estudo, identificaram-se três linhas principais de abordagem para o projeto desses sistemas: projeto de *codebooks* SCMA, algoritmos MUD de baixa-complexidade e projeto da alocação de recursos do *uplink* SCMA. A partir dessa análise, o trabalho buscou trazer contribuições em cada uma dessas linhas, ao mesmo tempo em que tenta trazer uma visão integrada mostrando as relações entre todos os parâmetros envolvidos nos diferentes estágios de projeto.

As contribuições propostas nessa tese consideram a análise do sistema SCMA sob diferentes aspectos: eficiência espectral global, desempenho de taxa de erro do enlace de comunicação, ordem de complexidade computacional do algoritmo receptor, capacidade simétrica do canal e complexidade de projeto do sistema.

Além disso, simulações numéricas considerando a de enlace são realizadas para SCMA codificado e não-codificado, sob canais com desvanecimento de pequena escala e também em ambientes celulares com perda de percurso (desvanecimento de larga-escala). Dessa forma, é possível analisar o desempenho do SCMA *uplink* em diferentes cenários de comunicação.

Outras figuras de mérito e formas de análise do sistema são apresentadas sob contexto ao longo do trabalho. Todas as simulações numéricas são baseadas no modelo matemático de comunicação digital em banda-base do SCMA *uplink*, em que todos os nós envolvidos (transmissores e receptor) possuem somente uma antena. Os recursos ortogonais podem ser interpretados como subportadoras de um sistema de múltiplo acesso por divisão ortogonal de frequência (OFDMA).

Resultados e Discussão

Na linha de projeto de *codebooks* SCMA, foi realizada a análise de desempenho das MDMs existentes na literatura e um estudo crítico das metodologias de projeto desses *codebooks*. Foi identificada uma semelhança entre alguns parâmetros e figuras de mérito dos *codebooks* SCMA com códigos corretores de erro (códigos de bloco). A partir dessa comparação, realizamos uma proposta de projeto de codebooks SCMA baseados em códigos de bloco lineares. O método proposto simplifica o projeto dos codebooks SCMA em relação aos métodos existentes da literatura e também garante propriedades desejáveis, como uma alta diversidade de espaço de sinais e códigos sobre alfabetos pequenos (auxiliando no controle da complexidade do MPA).

Em relação aos algoritmos para MUD, é realizada uma modificação de um algoritmo MPA com seleção de arestas existente. Essa modificação considera uma seleção dinâmica de arestas baseada em um limiar arbitrário. Ao definir esse limiar relativo aos ganhos instântâneos do canal de desvanecimento, a proposta mostra que é possível obter uma ordem de complexidade média menor no receptor, sem perda significativa do desempenho em relação à taxa de erro. A proposta viabiliza a operação de sistemas com alta eficiência espectral.

Por fim, na linha de projeto da alocação de recursos do *uplink*, realiza-se uma análise da capacidade simétrica do canal SCMA em ambientes celulares com desvanecimentos de larga e de pequena escala. Nesse contexto, comparamos os tipos de alocação de recurso denominados regular e irregular, no que diz respeito ao grau de espalhamento dos usuários. Através de simulações numéricas, mostra-se como realizar o projeto desses sistemas e compara-se o desempenho de ambos em termos de taxa de erro de bit.

Considerações Finais

Esta tese contribuiu com novas técnicas de projeto de sistemas SCMA *uplink*.

A técnica de projeto de *codebooks* SCMA baseados em códigos lineares permite o projeto de modulações multidimensionais a partir de diversos códigos algébricos clássicos da literatura, além de viabilizar esquemas de alta eficiência espectral sem comprometer a ordem de complexidade computacional do receptor MPA.

O algoritmo MUD proposto, utilizando seleção de arestas baseada em um limiar, também permite que esquemas com alta sobrecarga e alta eficiência espectral reduzam arbitrariamente a complexidade de detecção. Ao escolher apropriadamente o limiar, é possível garantir a confiabilidade do sistema mesmo com códigos corretores de erro de comprimento de bloco relativamente pequeno.

Por fim, a análise de capacidade simétrica dos sistemas SCMA em ambientes celulares mostra que é possível projetar assinaturas de espalhamento irregulares para os usuários, sem comprometer a confiabilidade média do sistema e regulando arbitrariamente o balanço de desempenho entre esses usuários.

Ressalta-se que as análises permitiram observar muitas potencialidades do sistema SCMA e que, apesar do mérito das contribuições, ainda existem desafios abertos para o projeto e otimização do sistema em cenários mais complexos.

Palavras-chave: SCMA, NOMA, modulações multidimensionais, algoritmo *message-passing*.

Abstract of Thesis presented to UFSC as a partial fulfillment
of the requirements for the degree of Doctor in Electrical Engineering

STRATEGIES FOR MULTISTAGE DESIGN OF SPARSE CODE MULTIPLE ACCESS SYSTEMS

Bruno Fontana da Silva

December 9th, 2019

Advisor: Bartolomeu Ferreira Uchôa Filho
Co-advisor: Danilo Silva
Area of concentration: Communications and Signal Processing
Keywords: SCMA, NOMA, Multiple Access, Multidimensional Modulation, Multiuser Detection
Number of pages: 144

Sparse Code Multiple Access (SCMA) is a code-domain non-orthogonal multiple access (NOMA). In this scheme, J users share K orthogonal resources, in a scenario where $J > K$. This scenario is said to be overloaded, since orthogonality is not possible and users will collide with each other in the resources. In SCMA, each user spreads its information using a sparse signature. Hence, the number of collisions in each resource is limited to a number significantly smaller than the number of users. The sparsity of the signatures allows for the use of a multiuser detector based on the message-passing algorithm (MPA).

In this thesis, multi-stage design strategies for uplink SCMA are discussed. A state-of-the-art review of related researches is provided, where SCMA is divided into three categories: SCMA codebook design methods, multiuser detection algorithms and resource allocation methods.

One contribution of this work is a design method of SCMA codebooks using the underlying structure of vector-subspaces generated by linear block codes. Using the proposed method, it is possible to design M-ary N-dimensional modulations with signal-space diversity $N - 1$ using alphabets of relatively small size $q \ll M$. The designed codebooks have a good error-rate performance while keeping a feasible detection complexity order with the MPA.

A second contribution is a threshold-based edge selection MPA. This is a modified receiver from an edge-selection MPA known in the literature, where the messages of each edge are calculated treating weak users as noise, according to a given threshold. A closed-form expression

for the average complexity order of the algorithm is also derived. Results show that it is possible to adjust the threshold in order to obtain a good trade-off between error-rate and complexity order, outperforming the original method.

Finally, a discussion on the resource allocation design for SCMA is presented. The analysis considers both complexity and information theory figures of merit. In particular, we discuss the achievable symmetric capacity and the error-rate performance of symmetric rate SCMA systems with regular and irregular spreading degrees. The scenario considers a wide-area network with both large-scale and small-scale fading channels.

List of Figures

List of Tables

Acronyms

5G	next generation of wireless communications systems
ALOHA	Additive Links On-line Hawaii Area
APSK	Amplitude-Phase Shift Keying
AWGN	Additive White Gaussian Noise
BER	bit error rate
BSA	Binary Switching Algorithm
CD	Code Domain
CDF	cumulative distribution function
CDMA	Code Division Multiple Access
CSI	channel state information
CSMA	Carrier Sense Multiple Access
ESGA	Edge-Selection with Gaussian Approximation
EXIT	Extrinsic Information Transfer
FDMA	Frequency Division Multiple Access
FEC	Forward Error Correction
FN	Function Node
GA	Gaussian Approximation
GLCP	Grouped Linear Constellation-Precoding
GRS	Generalized Reed-Solomon
HC	Hamming Code
LDPC	Low Density Parity Check
LDS	Low Density Signatures
MA	Multiple Access
MAC	Medium Access Control
MAP	maximum *a posteriori*
MC	Mother Constellation
MDM	Multidimentional Modulation
MDS	Maximum Distance Separable
MED	minimum Euclidean-distance
MIMO	Multiple-Input Multiple-Output (antennas)

MMF	max min-fairness
MPA	Message-Passing Algorithm
MPD	minimum product-distance
MUD	Multiuser Detection
NOMA	Non-Orthogonal Multiple Access
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
OMA	Orthogonal Multiple Access
OSI	Open Systems Interconnection
PAPR	Peak-to-Average Power Ratio
PD	Power Domain
PDF	probability density function
PEG	Progressive Edge-Growth
PHY	Physical Layer
PMF	probability mass function
QAM	Quadrature-Amplitude Modulation
QoS	Quality of Service
RA	Resource Allocation
RS	Reed-Solomon
SCMA	Sparse Code Multiple Access
SDMA	Space Division Multiple Access
SER	symbol error rate
SIC	Successive Interference Cancellation
SNR	Signal-to-Noise-Ratio
SS	Spread Spectrum
SSD	Signal-Space Diversity
TB	threshold-based
TDMA	Time Division Multiple Access
VN	Variable Node
WAN	Wide Area Network
WSC	Weighted Sum-Capacity

List of Symbols

Symbol	Description	Domain
J	number of users in the uplink	$\mathbb{Z}^+$
K	number of shared orthogonal resources	$\mathbb{Z}^+$
β	load factor, in SCMA $\beta = J/K > 1$	$\mathbb{R}^+$
M	number of modulation symbols	$\mathbb{Z}^+$
N	number of non-null chips in the spreading signatures	$\mathbb{Z}^+$
q	number of unique complex projections in each complex-dimension of the modulation	$\mathbb{Z}^+$
C	multidimensional constellation, M-ary, N complex-dimensions	$C \subset \mathbb{C}^N$
$\mathbf{c}_j$	symbol transmitted by user j	C
$\mathcal{X}_j$	sparse constellation of user j	$\mathcal{X}_j \subset \mathbb{C}^K$
$\mathbf{x}_j$	sparse symbol transmitted by user j	$\mathcal{X}_j$
$\mathbf{F}$	resource allocation matrix	$\mathbb{F}_2^{K\times J}$
$\mathbf{f}_j$	spreading signature of user j, i.e. the j-th column of $\mathbf{F}$	$\mathbb{F}_2^K$
$\mathbf{V}_j$	spreading operator of user j, such that $\mathbf{x}_j = \mathbf{V}_j\mathbf{c}_j$	$\mathbb{F}_2^{K\times N}$
d_k	number of user colliding in the resource k (irregular degrees)	$\mathbb{Z}^+$
d_f	number of user colliding in each resource (regular degree)	$\mathbb{Z}^+$
ϕ_k	set of users that collide in the resource k	-
$\phi_{k\setminus j}$	set of users that collide in the resource k, except user j	-
φ_j	set of resources where user j spreads its information	-
$\mathbf{h}_j$	fading coefficients of user j	$\mathbb{C}^K$
$\mathbf{z}$	additive noise at the receiver	$\mathbb{C}^K$
T_m	number of iterations of the message-passing algorithm	$\mathbb{Z}^+$
SE	system's overall spectral efficiency	$\mathbb{R}^+$

Contents

1 Introduction

Non-Orthogonal Multiple Access (NOMA) is a scheme in which multiple users of a communication system share the available resources of the medium in a non-orthogonal manner. Traditional Orthogonal Multiple Access (OMA) schemes, based on channelization, assign resources to users guaranteeing collision avoidance. This ensures a simple Multiuser Detection (MUD) at the receivers, due to orthogonality between signals of different users. On the other hand, in NOMA, the transmitted signals of different users will experience collisions (superposition on the same orthogonal resources).

The main motivation to break the collision avoidance paradigm is to increase the spectral efficiency of the system. In OMA, the number of supported users is limited to the number of available resources. If a Multiple Access (MA) scheme has more users than available orthogonal resources, the system is said to be *overloaded* and orthogonality is not possible anymore. Thus, NOMA is necessary to enable overloaded scenarios, increasing spectral efficiency and also the number of users that can be supported simultaneously [1, 2].

Among several NOMA schemes, SCMA has raised attention of the research and industry communities as a potential candidate for the next generation of wireless communications systems (5G). SCMA was first proposed in [3]. However, it is a generalization of another well-known scheme named Low Density Signatures (LDS) [18].

In both SCMA and LDS, Resource Allocation (RA) is a key step to enable reliable communication and feasible receiver complexity. Collisions are designed so that a limited number of users (ideally, much lower than the total number of users) will collide with each other in each resource. This is achieved by designing the users' signatures in a sparse manner. To do so, each user is limited to spread its information signal over a limited number of resources, typically much lower than the total number of available ones.

The work in [3] established SCMA transmitters as Multidimentional Modulation (MDM)

schemes, where each user maps a sequence of bits directly to a *codeword* from a multidimensional finite set of complex codewords, which is denoted as an SCMA *codebook*[1]. These codewords are further spread over the shared resources using the corresponding sparse signature designed for each user.

Another concern of NOMA is how to separate the superposed signals from different users at the receiver. In [3], the SCMA receiver builds on top of the joint-MUD proposed for LDS in [18]: the Message-Passing Algorithm (MPA) is applied as a (relatively) low-complexity alternative to the maximum *a posteriori* (MAP) detector. Because of the sparse RA, it is convenient to represent an SCMA scheme as a bipartite graph of user nodes and resource nodes. Thus, applying the MPA over this graph to detect the transmitted symbols can significantly save computation and achieve near-optimum performance.

In this work, we study SCMA on the uplink context, i.e. multiple users transmitting to a common destination (base-station). Similarly to the original work in [3], we deal with a multistage design approach for SCMA. From transmitter to receiver, this work analyses the SCMA performance and describes its design steps. Some contributions are proposed on the topics of codebook design, resource allocation and multiuser detection algorithms.

1.1 A Brief Overview of Medium Access Protocols

Before heading to the main topic, a general view of the field is discussed in this section.

1.1.1 OMA

Since the introduction of the Open Systems Interconnection (OSI) model, in the 1970s, industry and academia have designed communication systems that allow for the coexistence of different generations of devices and technologies (from 2G to 5G). Although the OSI model is a reference for the design of compatible communication networks and protocols, the TCP/IP model was developed earlier and is still applied in the industry, despite the fact that much has changed since its first version [19].

Although the models do not match perfectly, both OSI and the TCP/IP model have some equivalent layers. The Physical Layer (PHY) is the one at the bottom of the OSI stack, preceded by the data-link layer. In the TCP/IP, the bottom of the stack is an equivalent layer that groups functions of the PHY and data-link layers from the OSI model.

[1] In this work, the expressions “SCMA codebook”, “codebook”, “multidimensional constellation” and “modulation” may be used interchangeably.

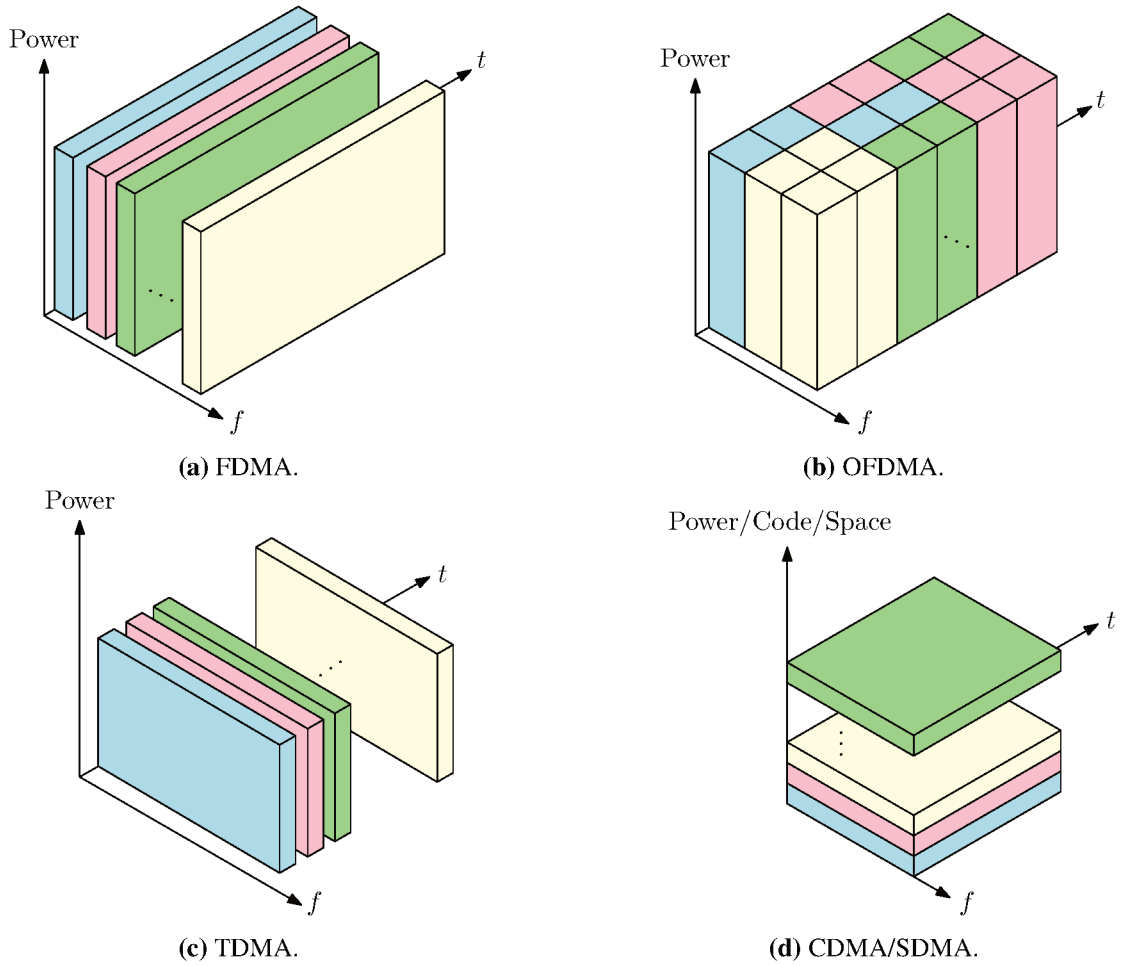
The PHY deals with the transmission of bits between two nodes over a physical medium. Its interface with the data-link layer is associated with protocols of Medium Access Control (MAC). The goal of these protocols is to give medium access permission to the multiple users that share the physical resources. In [19], the MAC protocols are separated into three categories: random access protocols, controlled-access protocols and channelization protocols.

Under random access protocols, the most popular are Carrier Sense Multiple Access (CSMA) and Additive Links On-line Hawaii Area (ALOHA)[2]. Random access implies that there is no hierarchy among users. In ALOHA, there is great potential of collisions in the receiver during the MA. CSMA was developed to reduce the chances of collision by "listening" to the medium before the transmission. In both cases, the receiver cannot properly decode the signals when there is a collision during the MA.

Regarding controlled-access protocols, the nodes jointly decide which users will be allowed to transmit. There are controls based on reservation, transmission queues, pooling methods and token-passing methods. In all cases, there is also a concern to avoid collision between users in the medium.

Finally, the channelization methods consider sharing the bandwidth of the link in different domains (time, frequency, code, space, etc.), between all users. Some common channelization protocols are Frequency Division Multiple Access (FDMA), Time Division Multiple Access (TDMA) and Code Division Multiple Access (CDMA). Figure 1.1 shows an illustration of the conventional OMA channelization protocols, including the variation of Space Division Multiple Access (SDMA).

[2]This is the original meaning of the acronym, although nowadays it has gained a broader meaning.



(a) FDMA. **(b)** OFDMA. **(c)** TDMA. **(d)** CDMA/SDMA.

Figure 1.1: Illustration of conventional OMA. Each color represents the signal block of a different user.

Also, the following box summarizes the main aspects of these protocols.

Conventional OMA

Frequency Domain in FDMA, channels are divided in disjoint bands on the frequency domain. Transmission over time is continuous and the multiple users are orthogonaly separated only in frequency domain. FDMA has been largely applied in analog cellular systems, such as AMPS and ETACS [20]. In digital communications and modern cellular networks, Orthogonal Frequency Division Multiple Access (OFDMA) systems play the equivalent MA role, using the Orthogonal Frequency Division Multiplexing (OFDM) subcarriers in frequency domain. Practical implementation issues of FDMA are related to carrier frequency offset (between the users themselves and between users and the destination), and out-of-band emission between different orthogonal bands or subcarriers.

Time Domain TDMA divides the time-domain in reserved time-slots, orthogonal be-

tween each other. These slots are distributed between multiple users. The distribution can be cyclic, and each cycle of user's time-slots is viewed as a frame. By analyzing a periodic frame with slots from multiple users, it will be noticed that they share the same frequency bandwidth in the medium. A difficulty of TDMA implementation is time-synchronism between the multiple users and the common destination. According to [20], TDMA protocol has been used in digital GSM, PDC, IS-54 and IS-136 standards.

Code Division in CDMA, multiple users transmit simultaneously in a continuous manner, occupying the full spectrum available in the shared medium. In order to make each user's signal distinguishable, the transmitters use a technique named Spread Spectrum (SS). In SS, the information symbol of the user is multiplied by a spreading sequence that uses more resources (time-slots or subcarriers) than the single one which the symbol would need in a common transmission. These sequences are similar to channel codes, and also have a corresponding rate with respect to RA. In the CDMA literature, these resources are also denoted "chips". It is possible to allocate different spreading sequences to each users. For example, by assigning orthogonal spreading sequences (orthogonal codes) to each user, it is possible to separate their messages at the receiver by using a correlation between the received signal and the known sequences. As described by [20], CDMA has been used in the digital telecommunication standards IS-95, W-CDMA and CDMA2000. For the uplink of IS-95, it is used a combination of orthogonal and non-orthogonal codes.

1.1.2 NOMA

In OMA, each resource (time-slot, subcarrier) is assigned to a single user, in order to avoid collision. Even in CDMA, where users collide in both time and frequency domains, the number of orthogonal spreading sequences that can be designed is limited to the number of available resources in the shared medium. This imposes a limitation to OMA: a multiple access cannot serve, simultaneously, more users than the number of available resources.

The interest in NOMA is to increase the capacity of multiple access schemes, allowing reliable communication in *overloaded* scenarios, i.e. more users than available resources of

the shared medium. Figure 1.2 shows an illustration of what future communication systems could look like when NOMA is considered. By using optimized methods of resource allocation, spectral efficiency could be improved. More users can be served simultaneously by dealing with collisions that can be resolved by modern MUD algorithms.

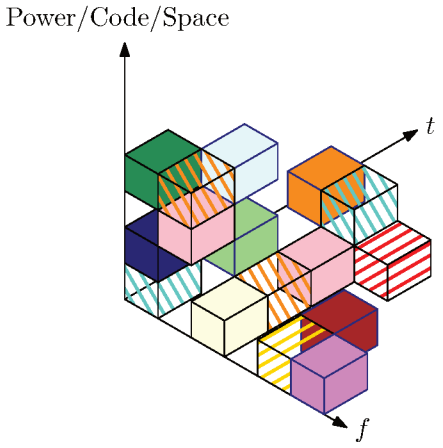


Figure 1.2: Illustration of NOMA. Each block filled with a solid color represents the signal block of a different user. Tilling patterns represent resources where two or more users have transmitted simultaneously, colliding in the same chip.

During the last decade, research on NOMA schemes has increased significantly, mostly motivated by the rush towards 5G standardization and requirements. Several schemes have been proposed in the literature [3, 18, 21]. According to [22], NOMA schemes can be classified into two main categories.

The first category considers multiplexing symbols based on Power Domain (PD), thus denoted PD-NOMA. In order to resolve the collisions and decode the symbols, PD-NOMA has proposed MUD algorithms based on Successive Interference Cancellation (SIC) [23].

Similarly to orthogonal CDMA, the second category of NOMA is based on Code Domain (CD), thus denoted CD-NOMA. To allow for overloaded systems, the authors in [4] have proposed designing sparse spreading sequences that would result in a low number of collisions in each chip. This scheme is the LDS multiple access. Later, in [3], LDS has been generalized to SCMA, where the spreading sequences are viewed as a support for the transmission of multidimensional symbols from an SCMA codebook. Thus, LDS has become a particular case of SCMA, corresponding to repetition codes.

Figure 1.3 shows the main categories of channelization MA with respect to its domains and orthogonality principles. As highlighted in the Figure, the present work concentrates on SCMA schemes.

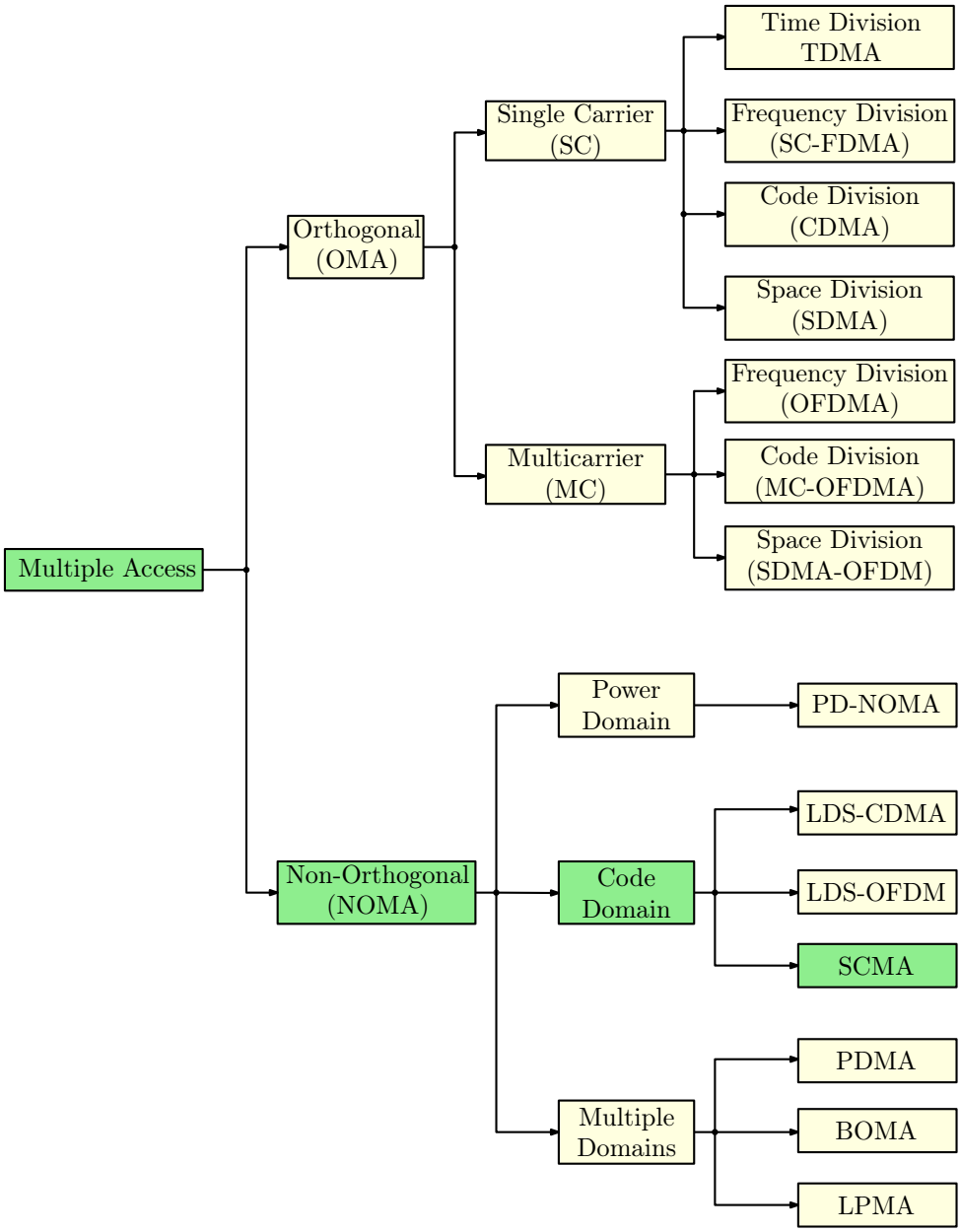


Figure 1.3: OMA and NOMA categories of multiple access. The NOMA schemes that are shown and were not mentioned in the text are: building block sparse-constellation based orthogonal multiple access (BOMA) and lattice-partition multiple access (LPMA). Figure adapted from [2].

1.2 Related Works on SCMA

Hundreds of papers related to SCMA have been published since the original work of [3]. From the point of view in the present work, a few research topics have been established as relevant problems for SCMA schemes:

(a) codebook design: the design methods to obtain good multidimensional constellations for the SCMA users;

(b) multiuser detection: the low-complexity receiver algorithms to separate the users' signals,

looking for more efficient algorithms with lower complexity than the original MPA, although trying to keep good error-rate performance;

(c) resource allocation: investigation of the design of the spreading sequences of each user, as well as the power allocation between SCMA users.

Although the research topics can be viewed by its main concern, it is difficult to decouple their design dependence in the full SCMA system. For example, large dimensional codebooks need more resources, influencing the resource allocation and causing more collisions per resource, which will increase the detection complexity.

Because of this design complexity, involving many parameters and codependent blocks, most of the literature has adopted a multistage design approach. This means that some assumptions are made in some stages (e.g., the resource allocation or the size of the systems) in order to optimize or improve the block of interest (e.g., the codebook with the best performance for a given system size).

For the codebook design procedure, the approach taken in works such as [5, 7, 24] is to assume that the resource allocation is regular (see Chapter 2) and given as an input for the design.

On the other hand, resource allocation analysis of SCMA is mostly concerned with the achievable rate limits of the multiple access scheme. Thus, works such as [25, 26] assume the use of given codebooks and/or resource allocation setups, normally for low sized schemes with four resources and six users, focusing on power allocation optimizations that typically maximize the sum-throughput of the system.

Regarding the proposed MUDs for SCMA, several approaches have been explored. In order to have fast convergence and reduce the number of operations, modified message update equations, with lower complexity throughout the MPA iterations, are proposed in [27–32]. Other approaches, such as [11, 33, 34], have tried to obtain linear complexity order combining powerful error-correcting codes with a simplified receiver algorithm. Most of these works have also evaluated the performance of the proposed MUD with some benchmark system (most of the times, four resources and six users) and with a fixed codebook of low number of dimensions (typically two complex dimensions and four codewords per codebook).

In Table 1.1, we summarize some of the relevant works published on these three subjects related to SCMA system design.

Table 1.1: Summary of representative research papers on SCMA, divided into three fields: codebook design, multiuser detection and resource allocation.

Scheme	References	Features
Codebook Design		
LDS-QAM	[4, 35]	repetition codes
Tx-QAM	[3, 5, 36]	lattice-rotation of MDMs, shuffling QAM projections
Low Number of Complex Projections (LNCP)	[5, 7, 8]	low-number of complex projections per complex-dimension (alphabet size smaller than codebook size)
Capacity-based	[6, 24]	optimization of constellation constrained capacity for the one-dimensional multiple access channel
Rotation-based	[37–39]	unitary rotation matrices to optimize product-distance or signal-space diversity
Star-QAM	[40, 41]	closed-form design of Star-QAM constellations to maximize minimum euclidean-distance
Spherical Codes	[42]	structure codebook-design based on spherical-codes
Genetic Algorithms	[43]	heuristic parameter optimization of structured MDMs
Autoencoders	[34, 44]	codebooks designed by unsupervised learning
Multiuser Detection Algorithms		
Original MPA	[3, 18]	based on the marginalization of MAP probabilities
q-MPA	[8]	exploits reduced alphabet sizes of each codebook
Fast-convergence MPA	[27–32]	early-stop of the detection of symbols with reliable beliefs
Edge-Selection MPA	[9, 10]	treats weak users as noise in the calculation of the MPA messages
Gaussian Approximation MPA	[9–11, 33]	simplifications of the MPA messages by Gaussian approximation on the interference
Turbo BICM-ID	[7, 45]	bit-interleaved coded modulation SCMA with iterative detection and decoding
Sphere and List Decoding	[46–53]	tree-search for the closest-vector decoder over structured codebooks
MIMO-SCMA	[11, 54–56]	algorithms for systems with multiple antennas
Neural Networks	[34, 44]	deep neural network multiuser detection (supervised learning)
Resource Allocation		
Achievable Limits	[12–14, 17, 57–59]	closed-form expressions to evaluate the channel capacity
Optimized Power Allocation	[25, 26, 60–62]	sum-rate maximization and power allocation
Subcarrier Assignment	[63, 64]	algorithms to assign the resources and to improve error-rate performance
LDPC-based Designs	[15, 16]	building the graph to maximize the girth

1.3 Goals and Contributions

The main goal of this thesis is to contribute with a generalized and flexible design method of uplink SCMA systems. The method must enable the design of large sized and large dimensional schemes, improving both the spectral efficiency and the performance of SCMA schemes. Moreover, the design approach must also take into account the feasibility of the multiuser detection at the receiver.

In order to propose practical design methods, this work also takes a multistage design approach, dealing with codebook design, multiuser detection and research allocation in a separate way. However, our proposed contributions try to make the design as flexible as possible and also evaluate the schemes in an integrated manner, checking for its robustness to scalability and overload.

More specifically, this thesis presents the following contributions:

(i) A novel codebook design method is presented. The SCMA codebooks are based on the underlying structure of error-correcting block codes in order to ensure a desired minimum signal-space diversity with a small-sized alphabet. In particular, when Maximum Distance Separable (MDS) codes are used, it is possible to achieve near-maximum minimum signal-space diversity[3]. A simple design procedure, of linear complexity, enables obtaining new families of codebooks using off-the-shelf linear error correcting codes. Several design examples are provided and it is shown that the proposed codebooks outperform other known design methods with respect to error-rate metrics and design complexity order.

(ii) An improved MPA-based MUD is proposed. The algorithm builds on top of the works from [9, 10], which in turn are based on edge-selection and Gaussian approximation techniques. Using a modified criterion to the edge-selection, with the aid of an arbitrary threshold, the proposed method can achieve as good or even better error-rate performance with lower complexity. Since the proposed method has a random complexity order, it is shown in an analytical analysis that the expected value of the complexity order can be adjusted to a desired value within a bounded region, allowing for a flexible trade-off between complexity and performance. Results are shown in a coded SCMA scenario for different systems and codebook sizes.

[3] Actually, with this approach, it becomes clear that SCMA codebooks only maximize the minimum signal-space diversity if they are based on an alphabet size at least as big as the codebook size.

(iii) Finally, we consider methods for the resource allocation stage of SCMA in a wide-area network scenario. Based on channel capacity limits from information theory, a rate analysis of SCMA is used to help in the design of codebook assignment and subcarrier assignment methods. This approach enables the design of heterogeneous setups, where each user may be able to transmit spreading over a different number of resources. The irregular resource assignments take into account the resulting complexity order at the receiver. Design examples are provided for symmetric rate SCMA systems.

1.4 Document Organization

In Chapter 2, an in-depth survey of SCMA is presented. State-of-the-art literature is reviewed, presenting comprehensive examples of all the features present in each stage of SCMA, as well as defining important mathematical notation and modeling assumptions for the remaining of the thesis.

In Chapter 3, the proposed codebook design method is presented. We show the general aspects of a multistage design focused on the codebook construction according to desired input parameters. Several examples are shown, and complementary details are illustrated in the Appendices.

In Chapter 4, we present the proposed multiuser detection based on an improvement from methods of the literature. This chapter describes the original MUDs that motivated the work and the details of the proposed modifications. A rigorous analysis of the complexity order is discussed. Simulation results of the proposed algorithm are done, evaluating its performance in several scenarios.

In Chapter 5, we conduct an analysis of the resource allocation design of SCMA. Figures of merit from information theory are used to evaluate the capacity of NOMA systems. We specifically evaluate the symmetric capacity of SCMA systems. With these analysis, we propose codebook assignment and subcarrier assignment algorithms for regular and irregular SCMA. We evaluate the achievable rates, complexity order of the receiver and the error-rate performance.

Finally, Chapter 6 summarizes the conclusions of the present work and enumerates a few possibilities of future works on the field. The remaining of the document includes also the literature references and complementary materials on the Appendices.

• • •

2

SCMA: State of The Art

In an uplink of an SCMA scheme, J users transmit their information to a common destination, e.g. the base-station, by spreading their signals over K orthogonal resources of the shared physical medium. Generally, throughout this work, all the considered scenarios are *overloaded*, which means that the reader must always assume $J > K$.

Wireless communications are considered for the implementation of SCMA. Unless otherwise mentioned, each of the transmitters and the receiver have only *one antenna*. Also, this work assumes channel models where the amplitude of the received signal is subject to *small-scale fading*, which follows a *Rayleigh distribution*.

In this chapter, it is assumed that no error correcting codes are applied, and the schemes are denoted as *uncoded*. Also, initially, *large-scale path-loss* is not considered. Coded scenarios and large-scale path-loss will be discussed in Chapters 4 and 5, respectively.

2.1 General SCMA Transmitter

Consider the transmitter $\mathcal{T}_j$ of user j. A binary message $\mathbf{b}_j \in \mathbb{F}_2^{k_j}$ will be transmitted using a codeword from an SCMA codebook $\mathcal{X}_j \subset \mathbb{C}^K$. The one-to-one mapping function $\mu_j^{-1} : \mathcal{X}_j \rightarrow \mathbb{F}_2^{k_j}$ denotes the binary labeling associated with each codeword.

The codebook size is $\left|\mathcal{X}_j\right| = M_j$, such that $M_j = 2^{k_j}$. User j is allowed to spread its signal only over $N_j < K$ of the shared resources. This means that the codewords $\mathbf{x}_j \in \mathcal{X}_j$ have only N_j non-null entries, while $K - N_j$ entries will be null.

An intermediate codebook $C_j \subset \mathbb{C}^{N_j}$, an M_j-ary N_j-dimensional modulation, represents the non-sparse MDM associated with user j. In order to spread the symbols $\mathbf{c}_j \in C_j$, a binary matrix $\mathbf{V}_j \in \mathbb{F}_2^{K \times N_j}$ applies a linear spreading operation.

Figure 2.1 illustrates the SCMA encoder procedure for user j. This encoding is done for

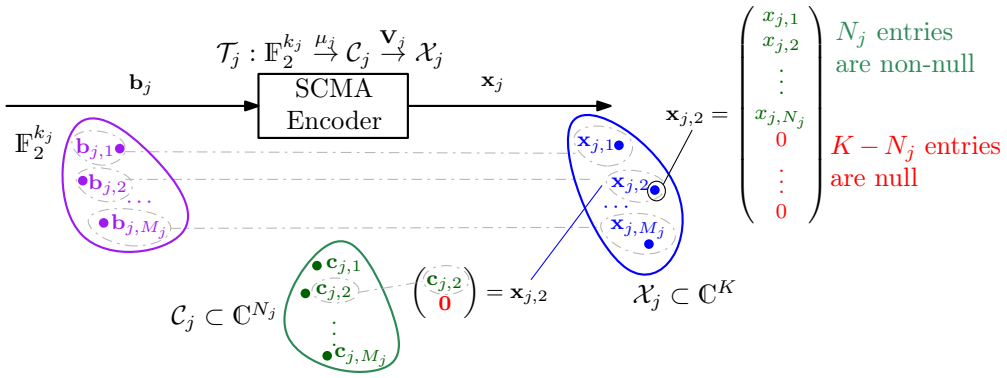


Figure 2.1: Illustration of the SCMA encoder $\mathcal{T}_j$ of user j.

all user $j = 1, \ldots, J$, so that they can transmit their symbols that will superpose at the receiver. Note that $\mathcal{T}_j$ is function of K, M_j, μ_j, N_j, $\mathcal{C}_j$ and $\mathbf{V}_j$. The sparse codebook $\mathcal{X}_j$ is completely defined by $\mathcal{C}_j$ and $\mathbf{V}_j$.

Each user transmits over a channel whose fading coefficients are given by a vector $\mathbf{h}_j \in \mathbb{C}^K$ with i.i.d. random variables. Each entry is distributed as $h_{j,k} \sim \mathcal{CN}(0,1)$, such that $|h_{j,k}|$ is a Rayleigh random variable. At the receiver, an Additive White Gaussian Noise (AWGN) is present, represented by the random vector $\mathbf{z}$ whose entries are i.i.d. and given as $z_k \sim \mathcal{CN}(0, \sigma^2)$.

The received signal $\mathbf{y} \in \mathbb{C}^K$ of an SCMA scheme is given by the following model:

$$\mathbf{y} = \sum_{j=1}^{J} \operatorname{diag}(\mathbf{h}_j)\, \mathbf{x}_j + \mathbf{z}. \tag{2.1}$$

This equation can also be represented in a different notation:

$$\mathbf{y} = \sum_{j=1}^{J} \operatorname{diag}(\mathbf{h}_j)\, \mathbf{V}_j \mathbf{c}_j + \mathbf{z} \tag{2.2a}$$

$$= \sum_{j=1}^{J} \mathbf{H}_j \mathbf{c}_j + \mathbf{z} \tag{2.2b}$$

where $\mathbf{H}_j = \operatorname{diag}(\mathbf{h}_j)\, \mathbf{V}_j \in \mathbb{C}^{K \times N_j}$ is the sparse effective channel of user j.

For convenience, we assume that the transmit power is normalized, such that $E\left[\|\mathbf{c}_j\|^2\right] = P = 1 \ \forall j$. Thus, we define the Signal-to-Noise-Ratio (SNR) as $\text{SNR} = \dfrac{P}{\sigma^2}$ and the bit energy per noise power of user j as $E_{b,j}/\sigma^2 = \text{SNR}/k_j$ (per channel use).

Note that, in the model of Equations (2.2), the spreading matrices $\mathbf{V}_j$ create a sparse resource allocation for each user. Moreover, since these matrices are fixed, independent of the selected codewords, there is a fixed assignment of resources for each user during the transmissions. This

assignment, or resource allocation, can also be represented in a graph structure, as illustrated in Figure 2.2.

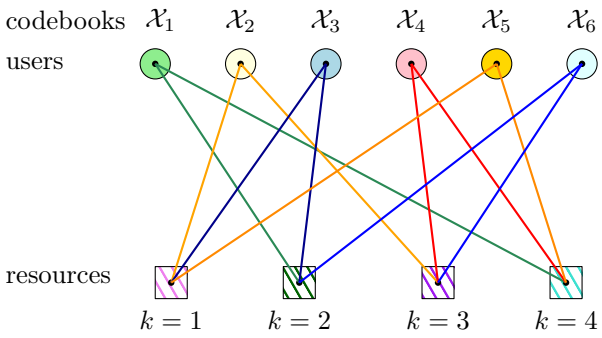


Figure 2.2: Graph representation of the SCMA resource allocation. This scheme has $J = 6$ users sharing $K = 4$ resources. Each user spreads its codewords over 2 resources and there are exactly 3 users colliding in each resource.

Thus, $\mathbf{V}_j$ can also be represented with a binary vector $\mathbf{f}_j$, whose entries $f_{j,k}$ are 1 if resource k is assigned to user j, otherwise $f_{j,k} = 0$. Concatenating the vectors $\mathbf{f}_j$ into a matrix $\mathbf{F} \in \mathbb{F}_2^{K \times J}$, the graph structure of SCMA can be represented by a sparse matrix, similar to Low Density Parity Check (LDPC) codes. The example of Figure 2.2 is a benchmark matrix from [3], which we denoted as:

$$\mathbf{F}_{4\times 6} = \begin{bmatrix} 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}. \tag{2.3}$$

As an example, for user $j = 2$, we have $\mathbf{f}_2 = \begin{bmatrix} 1 & 0 & 1 & 0 \end{bmatrix}^T$ and $\mathbf{V}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}^T$.

Using matrix $\mathbf{F}$, we can easily describe the graph structure of SCMA. Define the sets $\varphi_j = \{k : f_{j,k} = 1\}$, of the resources where user j spreads its codeword, and $\phi_k = \{j : f_{j,k} = 1\}$, of the users colliding in resource k. Thus, the relationships $N_j = |\varphi_j|$ and $d_k = |\phi_k|$ are also named spreading degree of user j and collision degree of resource k, respectively.

From LDS to SCMA, this graph structure of the resource allocation has motivated the use of the MPA as a MUD. In this detection algorithm, a user is referred as a Variable Node (VN) and a resource as a Function Node (FN). More details of the MPA are discussed in Section 2.5.

2.2 Multistage Design Methods

As discussed in Section 2.1, an SCMA scheme can be represented with parameters J, K, $\mathbf{F}$, and $\mathcal{T}_j\left(M_j, \mu_j, N_j, C_j\right)\ \forall j$. Considering some figure of merit $\mathfrak{F}$ to be minimized in the SCMA

system, an optimized design would consider the following problem:

$$\begin{aligned} \underset{J,K,\mathbf{F},\mathcal{T}_j\left(M_j,\mu_j,N_j,C_j\right)\forall j}{\text{minimize}} \quad & \mathfrak{F}\left(J,K,\mathbf{F},\mathcal{T}_j\left(M_j,\mu_j,N_j,C_j\right)\forall j\right) \\ \text{subject to} \quad & \mathfrak{C} \end{aligned} \tag{2.4}$$

where $\mathfrak{C}$ represents a list of constraints, including such as $J > K$, $N_j \leq K\ \forall j$, $P \leq 1$, etc.

Note that the resource allocation matrix $\mathbf{F}$ is dependent on several parameters (e.g. J, K, N_j), causing some dependence with the design of the codebooks. Also, to the best of our knowledge, there is no practical convex function to be used as figure of merit in this optimization problem. Typically, error-rate metrics could be considered, however we are unaware of the existence of practical analytical forms of this metrics for SCMA.

Hence, Problem (2.4) is not practical to be approached unless a few assumptions can be made, such as defining arbitrary parameters, regular structure among users, decoupling dependence of design stages, and so on. This multistage method, given some input parameters, was the approach suggested in [3] and followed in many other works [7, 24].

In order to simplify the SCMA design procedure, we will consider the following assumptions:

i. J and K are arbitrarily chosen to attend some desired load factor $\beta = J/K > 1$;

ii. all codebooks C_j have size $M_j = M$ ($k_j = k$ bits/codeword, $\frac{E_{b,j}}{\sigma^2} = \frac{E_b}{\sigma^2}\ \forall j$);

iii. $\mathbf{F}$ follows a regular design, such that columns have weight $N_j = N\ \forall j$ and rows have collision degree $d_k = d_f\ \forall k$.

Note that the assumption of regular column degree also implies that all codebooks C_j have N complex dimensions.

Also, regarding the codebooks, some works [3, 6, 38, 39] have proposed designing J codebooks C_j based on a Mother Constellation (MC) C combined with rotation operators. As explained in [35], this idea of rotations is highly motivated by AWGN channels without fading or downlink SCMA. When random channels with fading are considered in the uplink, random phase will affect each user, and the receiver will not be able to undo this effect for each user independently.

Therefore, in the present work, since our scope is uplink SCMA over Rayleigh channels, we consider another assumption to simplify the design procedure:

iv. all users apply the same codebook $C_j = C$, with the same binary labeling $\mu_j = \mu$.

Finally, for the given desired parameter tuple (J, K, M, N), the resulting design problem is described as:

$$\begin{aligned} \underset{C,\mu,\mathbf{F}}{\text{minimize}} \quad & \mathfrak{F}(C, \mu, \mathbf{F}) \\ \text{subject to} \quad & \mathfrak{C}. \end{aligned} \tag{2.5}$$

In many works [5, 7, 24], Problem (2.5) is further simplified by decoupling the codebook design, related to (μ, C), from the resource allocation design, associated with $\mathbf{F}$. Thus, for codebook design, it is assumed that $\mathbf{F}$ is an independent design, and only the knowledge of M and N are taken as inputs. On the other hand, resource allocation works such as [25, 26] assume the use of optimal codebooks on top of the optimized design of $\mathbf{F}$.

2.3 Codebook Design Methods

2.3.1 LDS Codebooks

Originally, LDS codebooks were designed by spreading a symbol x, taken from a Quadrature-Amplitude Modulation (QAM) set of size M ($\mathcal{Q}_M$), over some sparse signature $\mathbf{f}$. Then, the transmitted codeword would be $x\mathbf{f}$. This is similar to a repetition code.

Considering the RA matrix $\mathbf{F}_{4\times 6}$, an LDS-QAM codebook with parameters $(M, N) = (4, 2)$ can be described[1] as:

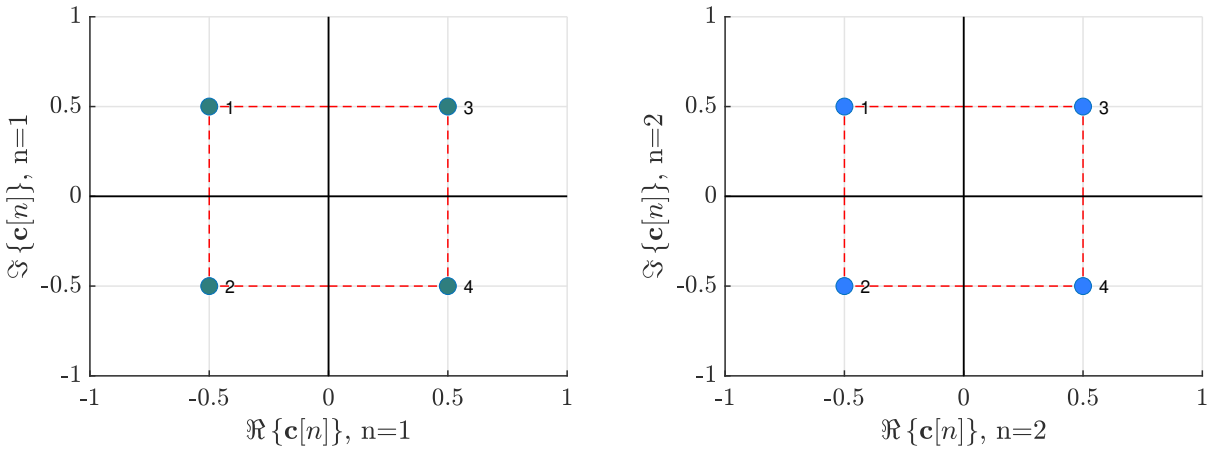
$$C = \frac{1}{\sqrt{E}} \begin{bmatrix} 0 & 0 & 0 & 0 \\ +1+i & -1+i & -1-i & +1-i \\ 0 & 0 & 0 & 0 \\ +1+i & -1+i & -1-i & +1-i \end{bmatrix}. \tag{2.6}$$

where E is a normalizing factor to ensure that $P = 1$. Figure 2.3 illustrates the complex projections of this LDS codebook in each complex dimension.

2.3.2 Tx-QAM Codebooks

The first SCMA codebooks were proposed in [3, 5]. The examples are given for the benchmark RA matrix $\mathbf{F}_{4\times 6}$, with $N = 2$ complex-dimensions per codeword. The authors suggest using a mother constellation C, based on a single-user MDM with good properties, such as large minimum Euclidean-distance (MED) and large minimum product-distance (MPD).

[1] We slightly abuse the notation representing the set of codewords C as a matrix, where each column corresponds to a codeword. This notation is convenient to describe SCMA codebooks and discuss its properties.



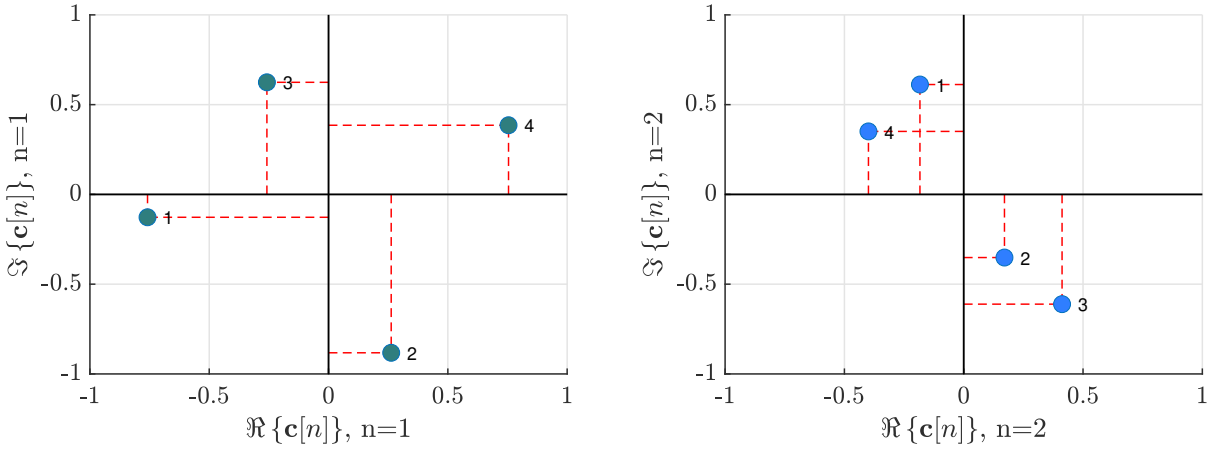
(a) $n = 1$. **(b)** $n = 2$.

Figure 2.3: LDS-QAM codebook, with $(M,N) = (4,2)$ and labeled with Gray mapping according to the first dimension.

Constellation operators [3] (e.g., rotation matrices) may be applied to the MC in order to differentiate the codebooks between users.

The methods proposed in [3, 5] suggest composing the entries of each complex-dimension of the codebook from mixed real/imaginary parts of QAM symbols. Permutations between the complex-dimensions will lead to the composition of the codewords.

In Figure 2.4, a codebook with parameters $(M,N) = (4,2)$ is illustrated. This codebook is given as an example in [36]. The first and second complex dimensions correspond to the plots 2.4a and 2.4b, respectively. The numbers refer to the binary labeling in natural order (represented by the codeword indexes $1,\ldots,M$). Assuming normalized power $P = 1$ for both codebooks, T4-QAM of Figure 2.4 has a shaping gain of 1,25 (dB) with respect to the LDS-QAM codebook of Figure 2.3 [36], because of its increased MED.



(a) $n = 1$. **(b)** $n = 2$.

Figure 2.4: Codebook T4-QAM from [36]. Labeled projections $\Re + i\Im$ in each dimension.

SCMA codebooks from [3, 5] are denoted Tx-QAM, where x is replaced by *the number of unique complex projections in each complex-dimension* of the codebook. Throughout the present work, we may refer to this definition simply as the "projections" or the "alphabet" of the complex-dimensions.

In [8], the authors give more details on the design of codebooks with a low number of projections. They also proposed a low-complexity receiver that exploits this type of codebook design. In Section 2.5, more details are reviewed regarding this receiver.

2.3.3 $A_{M,q}$ APSK Codebooks

In [7], a multistage method is proposed to build the codebooks. For a desired codebook size M and number of complex-dimensions N, the authors propose building a MC on top of an one-dimensional Amplitude-Phase Shift Keying (APSK) alphabet $\mathcal{A}_q$. Moreover, the APSK alphabet may be chosen with $q \leq M$ elements, in order to enable codebooks with low number of projection.

To obtain the N-dimensional M-ary codebook, the following steps are proposed in [7]:

i. Choose an arbitrary M-lengthed permutation, denoted π_1, of the elements in $\mathcal{A}_q$. There will be $M - q$ non-unique elements. This permutation, $\pi_1\left(A_q\right)$, will correspond to the M projections of the first complex-dimension of C, labeled $1, \ldots, M$.

ii. Find the best set of $N-1$ M-lengthed permutations of $\mathcal{A}_q$, $\pi_2, \ldots, \pi_N$, that will compose remaining entries of the codewords.

By "best set of permutations", the authors in [7] mean the set which maximizes an arbitrary figure of merit related to the codebook C. They propose optimizing the cut-off rate, a metric described in [7, 37]. Also, because the complexity order of searching for the best permutations is large, $O\left(M!^{N-1}\right)$, the authors simplify the procedure to a sub-optimal approach, building one-dimension at a time. This reduces the complexity order of the design to $O\left((N-1)\,M!\right)$.

A last step discussed in [7] is to find the best binary labeling μ for the designed codebook. Based on a labeling figure of merit from [65], the solution proposed is to apply the Binary Switching Algorithm (BSA), iteratively, saving a few number of the best generated labelings. Later, these labelings can be chosen with an Extrinsic Information Transfer (EXIT) chart analysis for coded SCMA [7].

The codebooks designed in [7] are denoted $A_{M,q}$ codebooks. An illustration of the design method described above is shown in Figure 2.5. In Figure 2.6, the example $A_{8,8}$ from the

original work is illustrated. This codebook has $N = 2$ complex-dimensions and shows the best performance when compared to the other examples $A_{8,q}$ where $q < 8$.

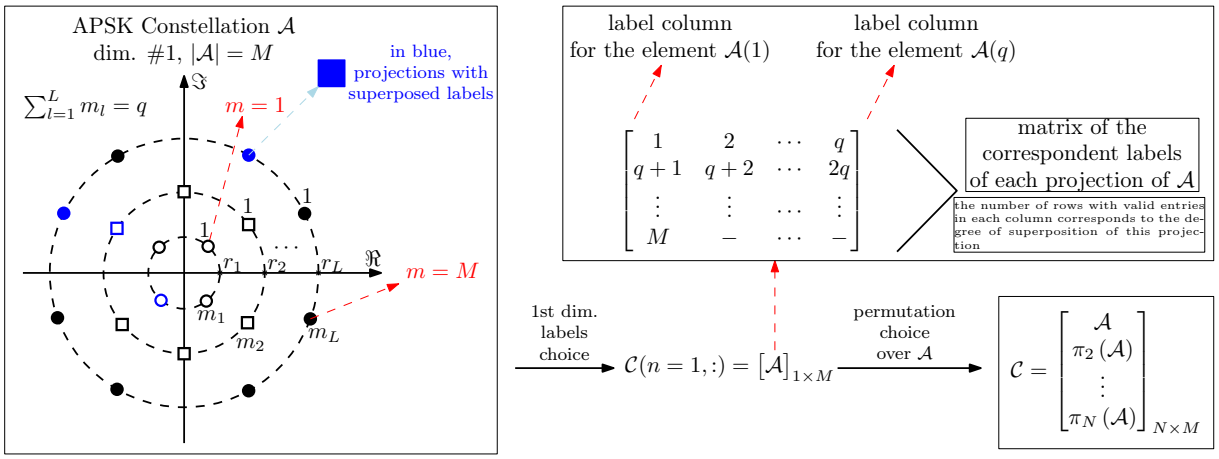
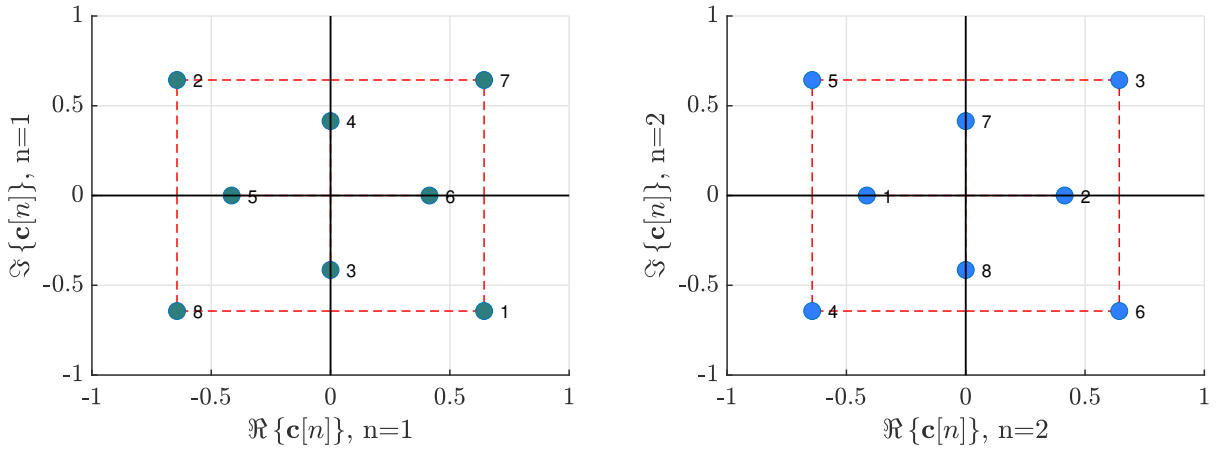


Figure 2.5: Illustration of the design method of $A_{M,q}$ SCMA codebooks. The APSK constelation is built with L energy rings, each one with a certain number m_l of PSK projections. The matrix structure with q columns of labels is just figurative. See more details of the design method in [7].



(a) $n = 1$. **(b)** $n = 2$.

Figure 2.6: Codebook $C = A_{8,8}$ from [7]. Labeled projections $\mathfrak{R} + i\mathfrak{I}$ in each dimension.

2.3.4 Other Design Methods

In [35], a survey on different codebook design methods is presented. The authors illustrate several codebooks, comparing their figures of merit and evaluating its performances over different types of channels.

Another codebook design method that deserves attention is from the works [6, 24]. The authors consider an information theoretical figure of merit, based on insights of the AWGN multiple access channel (collisions over one complex-dimension). Optimal phase-rotations can be designed for d_f users colliding in a single-resource. Then, the one-dimensional codebook is

expanded to N dimensions searching for permutation patterns, similar to [7].

In [66, 67], real-valued MDMs are designed for a single-user scenario. These MDMs can achieve maximum signal-space diversity, i.e. a diversity equal to the number of dimensions of the codebook. This is achieved by combining lattice constellations with Grouped Linear Constellation-Precoding (GLCP), which is a generalization of the rotation matrices for MDMs originally proposed in [68].

More works on SCMA codebook design methods are listed in Table 1.1.

2.3.5 Other Design Criterion

From the discussion in the previous sections, it can be observed that many design methods build the SCMA codebooks on top of a MC, which is designed based on single-user figures of merit for MDMs.

Joint design of J codebooks seems to be a difficult problem, and this approach is not well exploited in the current literature. Some works such as [69–71] deal with the joint stacked vector of J multidimensional symbols to derive upper bounds on the average pairwise error probability. However, evaluation of this metric may have large computational cost, since there are $M^J\left(M^J-1\right)$ error events. Nevertheless, a few insights are found in [69], which may serve as a guideline to the design of MCs.

Another issue is that, although some codebook design methods may be generalized, most of them show evaluation results only for $N = 2$ complex-dimensions and $M \leq 16$, using the benchmark $\mathbf{F}_{4\times6}$. Although this choice may be for comparison reasons, there is also a problem of scalability. Large-sized codebooks, with respect to M and N, may have poor performance with the increase of this parameters, due to energy limitations $P \leq 1$. Moreover, increasing M and N may also result in huge decoding complexity, when considering the MPA.

In the present work, we approach the SCMA codebook design problem aiming to solve these scalability and complexity issues. A proposed design method is discussed in Chapter 3.

2.3.6 Performance Comparison

In Figures 2.7 and 2.8, we compare the symbol error rate (SER) and bit error rate (BER), respectively, of some state-of-the-art SCMA codebooks described in Section 2.3. Monte Carlo simulations were done over the Rayleigh fading channel model. The resource allocation matrix is the benchmark $\mathbf{F}_{4\times6}$.

It can be noticed that all curves have a slope with diversity 2, related to the signal-space

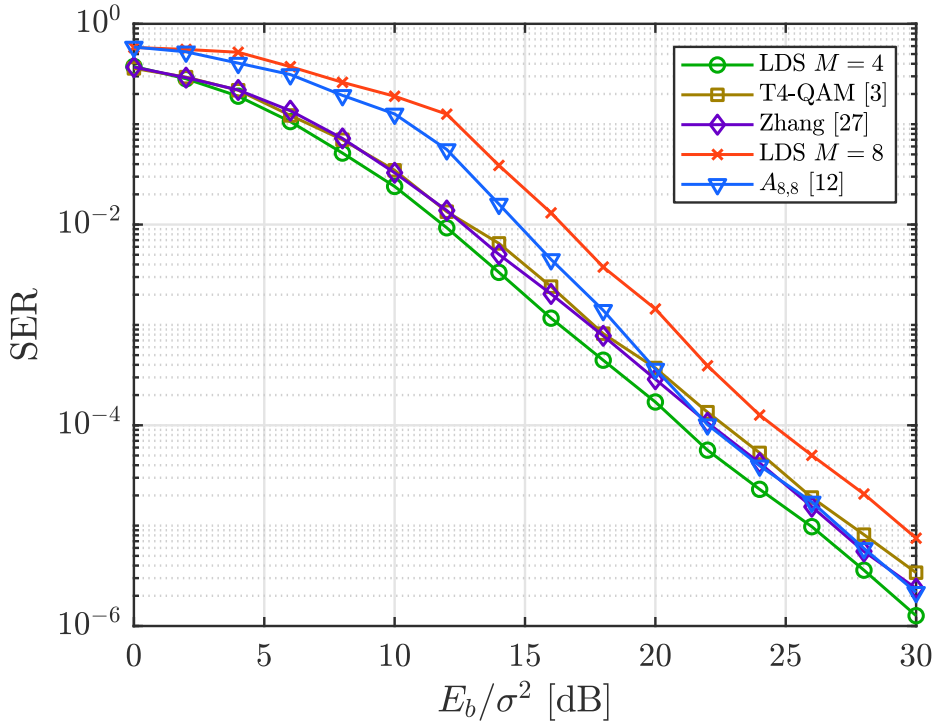


Figure 2.7: Symbol error rate of some codebooks described in Section 2.3. The resource allocation matrix is the benchmark $\mathbf{F}_{4\times6}$.

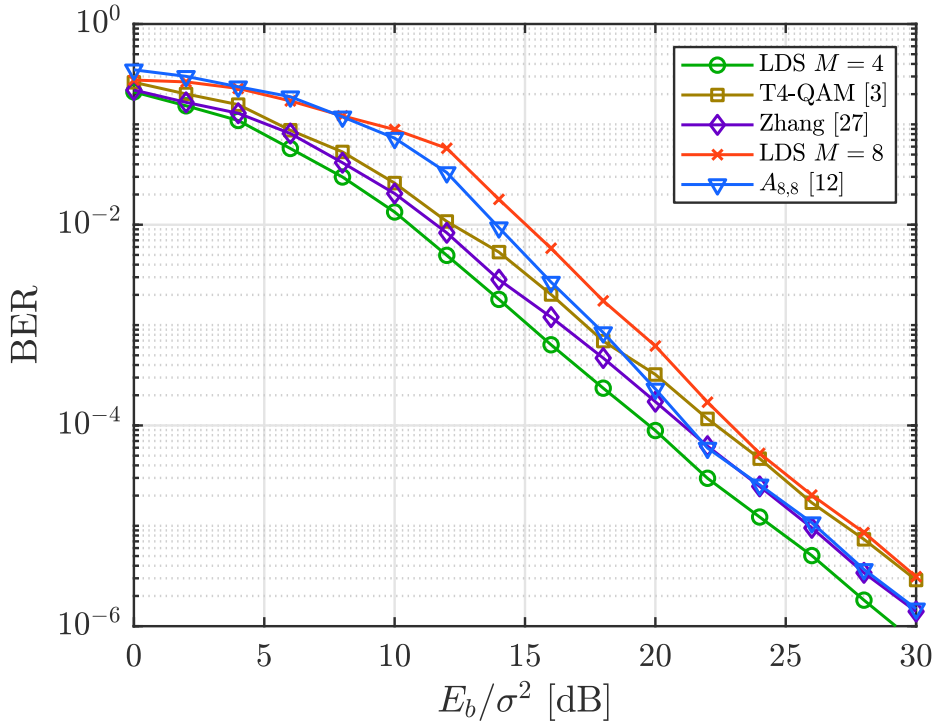


Figure 2.8: Bit error rate of some codebooks described in Section 2.3. The resource allocation matrix is the benchmark $\mathbf{F}_{4\times6}$.

diversity of these codebooks, which is maximum, since $N = 2$ for all codebooks.

Among the codebooks with $M = 4$ (including T4-QAM and Zhang), LDS-QAM codebook

seems to have the best performance. On the other hand, with $M = 8$, the codebook $A_{8,8}$ [7] shows a better coding gain in comparison to LDS designs.

The improvement from the SER to the BER is related to the binary labeling μ of each codebook. Since the labeling of T4-QAM was not provided in [36], there may be room for improvement. Also, for the codebook from [6], we replicated the one described for user $j = 1$, ignoring the user-specific rotations.

2.4 Resource Allocation Design

Resource allocation amounts to determining the spreading signature of each user, i.e. in which of the shared resources the users can transmit their symbols. This design is equivalent to determining the binary signature matrix $\mathbf{F}$.

2.4.1 Aiming for Achievable Limits

The works in [17, 57, 58] investigate achievable limits of the sum-capacity in NOMA systems under different scenarios, considering the channel model and channel state information (CSI). In the works [25, 26, 60–62], RA optimization aiming to maximize the achievable sum-rate is discussed.

LDS Sum-Capacity for the AWGN Channel from [13]

In [12–14, 59], the authors derive closed-form expressions for the sum-capacity of the LDS transmission (sparse CDMA) over AWGN channels (possibly with phase-rotations).

This derivations rely on the following hypotheses:

a. $\mathbf{x}_j = x_j \mathbf{f}_j$ (repetition codes)

b. $\mathbf{x} = \begin{bmatrix} x_1 & \cdots & x_J \end{bmatrix}^T \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_K)$ (joint symbol vector is Gaussian distributed).

In [12, 13], the focus is on semi-regular degree distribution (N, d_f) of $\mathbf{F}$, where $d_f = \beta N$. The authors describe the sum-capacity based on the characterization of the squared singular values of $\frac{1}{\sqrt{N}}\mathbf{F}\mathbf{F}^H$. From [13]:

$$C_{\text{sum}}^{(\text{LDS})} \triangleq \frac{1}{K} I(\mathbf{x}; \mathbf{y}|\mathbf{F}) \tag{2.7a}$$

$$= \frac{1}{K} E\left\{ \log_2 \det\left[\mathbf{I}_K + \frac{\text{SNR}}{N}\mathbf{F}\mathbf{F}^H \right] \right\} \tag{2.7b}$$

$$\underset{K\to\infty}{=} (\beta-1)\log_2\left[1+c_3\mathrm{SNR}-\frac{1}{4}\mathcal{F}\left(c_2\mathrm{SNR},\frac{c_3}{c_2}\right)\right] + c_1\log_2\left[1+(c_2+c_3)\,\mathrm{SNR}-\frac{1}{4}\mathcal{F}\left(c_2\mathrm{SNR},\frac{c_3}{c_2}\right)\right] - c_1\log_2\left[\frac{\left(1+d_f\mathrm{SNR}\right)^2}{\mathcal{G}\left(c_2\mathrm{SNR},\frac{d_f}{c_2},\frac{c_3}{c_2}\right)}\right] \tag{2.7c}$$

where $d_f = \beta N$, $c_1 = \left(\frac{\beta(N-1)+1}{2}\right)$, $c_2 = \left(\beta-\frac{1}{N}\right)$, $c_3 = \left(1-\frac{1}{N}\right)$ and

$$\mathcal{F}(x,z) \triangleq \left[\sqrt{x\left(1+\sqrt{z}\right)^2+1}-\sqrt{x\left(1-\sqrt{z}\right)^2+1}\right]^2 \tag{2.8}$$

$$\mathcal{G}(x,y,z) \triangleq \left[\frac{\sqrt{\left(y-\left(1-\sqrt{z}\right)^2\right)\left(x\left(1+\sqrt{z}\right)^2+1\right)}-\sqrt{\left(y-\left(1+\sqrt{z}\right)^2\right)\left(x\left(1-\sqrt{z}\right)^2+1\right)}}{\sqrt{y-\left(1-\sqrt{z}\right)^2}-\sqrt{y-\left(1+\sqrt{z}\right)^2}}\right]^2. \tag{2.9}$$

$x, y, z \in \mathbb{R}^+$ and $y \geq \left(1+\sqrt{z}\right)^2$.

The result in [13] is also dependent on the corresponding graph of $\mathbf{F}$ being locally tree-like and convergent (in the large-system limit, as $K \to \infty$) to a bipartite Galton Watson tree (BGWT).

Results for $N_j = N = 1\ \forall j$, with irregular collision degrees, are available in [14, 59].

Remark: *as observed in [57], when considering repetition codes, the covariance matrix of symbols $\mathbf{f}_j x_j$ is not diagonal, because of the correlation among dimensions. On the other hand, for the joint LDS vector $\mathbf{x}$, it is natural to consider that $E\left[x_j x_i\right] = 0$ for different users $i \neq j$.*

Average SCMA Sum-Capacity for the Rayleigh Channel [17]

In the related work [17], a closed-form expression for the average sum-capacity of uplink SCMA is derived, assuming that the users are uniformly randomly deployed in a hexagon cell area. The authors assume regular distribution of $N_j = N\ \forall j$, with uniform power allocation, Rayleigh small-scale fading channels and path-loss coefficients $\alpha_j^{-1} = D_j^{0.5\rho}$, where ρ is the path-loss exponent and D_j is the distance between user j and the base-station.

The result in [17] is an analytical expression that mainly depends on the distances $D_{\min} = \min_j D_j$, $D_{\max} = \max_j D_j$ and the distribution of collision degrees d_k for $k = 1, \ldots, K$.

However, as described by the authors, the expression for the generalized scenario provides little insight. Hence, the authors simplify the result to the particular case where the users are in

the same radius, i.e. $D_{\min} = D_{\max}$ (i.e. the degraded Rayleigh channel). In this scenario, we adapt the notation from [17] to rewrite its result here:

$$C_{\text{sum}}^{(\text{SCMA})} = \frac{e^{\frac{R^\rho}{P}}}{\log(2)} \sum_{k=1}^{K} \sum_{i=1}^{d_k} \sum_{j=0}^{i-1} \binom{d_k}{i} \binom{i-1}{j} (-1)^{i-1-j} E_{d_k-j}\left[\frac{R^\rho}{P}\right], \tag{2.10}$$

where $E_n[x]$ is the exponential integral [72].

2.4.2 LDPC-inspired Designs

In [15, 16], the authors consider design methods for high-dimensional RA matrices in SCMA based on known constructions of LDPC matrices, independently of the CSI.

In [73], we consider the design of $\mathbf{F}$ based on Progressive Edge-Growth (PEG) construction [74]. More details are shown in Chapter 3.

2.4.3 Subcarrier Assignment Algorithms

Suboptimal strategies are proposed in [25, 26, 60–62], usually applying subcarrier assignment algorithms to find suboptimal results of the sum-rate, and sometimes followed by power allocation optimization [25, 26, 61].

In [64], the authors proposed order-based subcarrier assignment algorithms, using three different approaches: fixed user order, opportunistic allocation and proportional fairness.

2.5 Multiuser Detection

Due to SCMA's sparse feature, the joint multiuser detection can be performed using the message-passing algorithm, achieving near-optimal performance in comparison to maximum *a posteriori* joint detection.

The motivation for the MPA is the complexity order of the joint detection. If maximum likelihood is used, testing for all possible results of superposed symbols at the receiver, the complexity order of this joint detection is $O\left(M^J\right)$. With the use of the MPA, associated with the design of the sparse signatures, a complexity order of $O\left(M^{d_f}\right)$ can be obtained for the joint detection [3, 18]. Although both algorithms have exponential complexity order, using MPA can significantly reduce the computational cost at the receiver, as long as $d_f \ll J$.

To reduce the detection complexity even more, several low-complexity receivers based on modified MPA have been proposed.

In [75], the log-MPA is used to avoid computing large sums of exponentials.

In order to have fast convergence and reduce the number of operations, modified message-update equations, with lower complexity in the MPA iterations, are proposed in [27–32]. However, in all of these works, the first iteration still has exponential complexity order $O\left(M^{d_f}\right)$.

A SCMA codebook design method, based on low number of complex projections per complex dimension, was first proposed in [5] and further detailed in [8]. This method imposes a limitation $q < M$ to the number of complex projections in each complex dimension of the codebooks. With a modification of the MPA (which we denote here as q-MPA), the MUD complexity order reduces to $O\left(q^{d_f}\right)$.

Due to this exponential complexity, MPA solutions in [3, 5, 8, 18, 27–32, 75] are intractable for large overloads and rates. In [9, 10], an MPA based on Edge-Selection with Gaussian Approximation (ESGA) on interference has been proposed for coded SCMA. In ESGA-MPA, the complexity order is $O\left(M^{d}\right)$, and the exponent d can be arbitrarily limited within $1 \leq d \leq d_f$. Performance degradation is expected for the lowest complexity $d = 1$, which means full Gaussian Approximation (GA) on the interference. However, it is not clear if the scheme works combined with codebooks of large-valued M or N, or even combined with q-MPA.

Other existing MUDs for SCMA are listed in Table 1.1.

2.5.1 Original Message-passing Algorithm for SCMA

The MPA for SCMA approximates the optimum detection given by the MAP Equations. In Appendix A, we show the derivation of these Equations for SCMA schemes.

The MPA receivers are based on the bipartite graph structure of the resource allocation matrix $\mathbf{F}$. The users are considered as VNs and the resources as FNs. Recall the sets φ_j and ϕ_k defined in Section 2.1. For the regular designs assumed in Section 2.2, we have that $|\phi_k| = d_f$ $\forall k$ and $\left|\varphi_j\right| = N$ $\forall j$.

From the model of Equation (2.1), each FN has an observed sample y_k and can calculate the following metric:

$$f\left(y_k, \mathbf{X}_{[k]} | \mathbf{H}_{[k]}\right) = -\frac{1}{2\sigma^2}\left| y_k - h_{j,k}x_{j,k} - \sum_{i \in \phi_{k\backslash j}} h_{i,k}x_{i,k} \right|^2, \tag{2.11}$$

where $\phi_{k\backslash j} = \phi_k \setminus \{j\}$, and $\mathbf{X}_{[k]}$ and $\mathbf{H}_{[k]}$ represent the concatenation of complex-projections and channel coefficients, respectively, from users in ϕ_k.

The messages exchanged between the nodes, for $1 \leq t < T_m$ iterations of the MPA, are given

as:

$$I_{k\to j}^{(t)}\left(\mathbf{x}_j\right) = \sum_{\mathbf{x}_{\phi_{k\backslash j}(1)}} \cdots \sum_{\mathbf{x}_{\phi_{k\backslash j}\left(d_f-1\right)}} \left[\exp\left(f\left(y_k, \mathbf{X}_{[k]}|\mathbf{H}_{[k]}\right) + \sum_{i\in\phi_{k\backslash j}} \log I_{i\to k}^{(t-1)}\left(\mathbf{x}_i\right)\right)\right] \tag{2.12a}$$

$$I_{j\to k}^{(t)}\left(\mathbf{x}_j\right) = \prod_{k'\in\varphi_{j\backslash k}} I_{k'\to j}^{(t)}\left(\mathbf{x}_j\right), \tag{2.12b}$$

where $\varphi_{j\backslash k} = \varphi_j \setminus \{k\}$, $\phi_{k\backslash j}(i)$ is the index of the i-th user in $\phi_{k\backslash j}$ and the initial extrinsic message is given as $I_{j\to k}^{(0)}\left(\mathbf{x}_j\right) = p\left(\mathbf{x}_j\right)$.

Observe that

a. the messages exchanged in the MPA are M-dimensional vectors, whose entries correspond to probabilities regarding the discrete random variable $\mathbf{x}_j$;

b. before sending a message from node a to node b, normalization should be done with respect to the probability mass function (PMF) of $\mathbf{x}_j$, such that $\sum_{\mathbf{x}_j} I_{a\to b}^{(t)}\left(\mathbf{x}_j\right) = 1$;

c. in the last iteration ($t = T_m$), all FNs are taken into account, i.e. $I_j^{(T_m)}\left(\mathbf{x}_j\right) = \prod_{k\in\varphi_j} I_{k\to j}^{(T_m)}\left(\mathbf{x}_j\right)$;

d. in coded SCMA, when applying outer loops in iterative detection and decoding, $p\left(\mathbf{x}_j\right) = 1/M$ for the first loop; for the next loops, the new *a priori* information $p\left(\mathbf{x}_j\right)$ can be obtained using the extrinsic information from the output of an error-correcting decoder.

2.5.2 MPA Complexity Order

When applying the MPA in [18] for SCMA [3], the complexity order O_{MPA} of exchanging messages in the graph can be written as

$$O_{\text{MPA}} = T_o T_m \left(C_{\text{FN}}^{\text{MPA}} + C_{\text{VN}}\right), \tag{2.13}$$

where $C_{\text{VN}} = MJ(N-1)$ is the variable node complexity order (independent of the MUD operations at the FN) and $C_{\text{FN}}^{\text{MPA}} = Kd_f M^{d_f}$ is the function node complexity order of the original MPA for SCMA.

Note that the MPA has exponential complexity order with respect to the collision degrees d_f. This leads to a trade-off between the load factor $\beta = J/K$ and computational cost of the joint detection.

In [7, 8], the authors assume the use of codebooks with low number of projections to mitigate the effect of exponential complexity. Assuming the use of this type of codebooks, with

a limitation $q_j \leq q$ on the number of complex projections per complex dimension, and the use of the corresponding q-MPA in [8], we could limit (2.13) as

$$O_{\text{MPA}} \leq T_o T_m \left[C_{\text{FN}}^{q\text{-MPA}} + C_{\text{VN}} \right], \tag{2.14}$$

where $C_{\text{FN}}^{q\text{-MPA}} = K d_f q^{d_f}$.

2.5.3 Log-Domain MPA and max log-MPA

Computing the messages of the MPA in the log-domain was first proposed in [75].

Taking the logarithm of $I_{k\to j}^{(t)}\left(\mathbf{x}_j\right)$ in (2.12) leads to the use of the binary operation $\overset{\star}{\max}$, defined as:

$$\overset{\star}{\max}(a,b) = \max(a,b) + \log\left[1 + \exp\left(-|a-b|\right)\right], \tag{2.15}$$

which, alternatively, can be also written as:

$$m = \max_i x_i \tag{2.16a}$$

$$\log \sum_{i=1}^{n} \exp(x_i) = m + \log\left(\sum_{i=1}^{n} \exp\left(x_i - m\right) \right). \tag{2.16b}$$

The operation $\overset{\star}{\max}$ may also be known as the Jacobian Logarithm. This operation avoids computing large sums of exponentials, preventing numerical precision issues. Another benefit of Log-MPA is avoiding multiplications in the VNs, computing only the summation of the messages sent from the FNs.

The resulting MPA Equations are rewritten as:

$$I_{k\to j}^{(t)}\left(\mathbf{x}_j\right) = \overset{\star}{\max}_{\mathbf{x}_{\phi_{k\setminus j}(1)} \cdots \mathbf{x}_{\phi_{k\setminus j}\left(d_f-1\right)}} \left[f\left(y_k, \mathbf{X}_{[k]} | \mathbf{H}_{[k]}\right) + \sum_{i \in \phi_{k\setminus j}} \log I_{i\to k}^{(t-1)}\left(\mathbf{x}_i\right) \right] \tag{2.17a}$$

$$I_{j\to k}^{(t)}\left(\mathbf{x}_j\right) = \log\left(p\left(\mathbf{x}_j\right)\right) + \sum_{k' \in \varphi_{j\setminus k}} I_{k'\to j}^{(t)}\left(\mathbf{x}_j\right), \tag{2.17b}$$

A further simplification can be used, which is denoted max log-MPA. This algorithm simply replaces $\overset{\star}{\max}$ with a regular max operation, removing the regularization component of Equations (2.15) and (2.16b). Although this reduces the computational cost, it also causes some degradation in the error-rate performance.

2.5.4 MPA for Codebooks with Low Number of Projections

Initially proposed in [5] and further detailed in [8], codebooks with a low-number $q < M$ of complex projections in each complex-dimension can help to reduce the receiver complexity.

Because the alphabet of each user is reduced, there is a smaller number of possible combinations in each resource. Then the receiver complexity will be given by Equation (2.14). This is a polynomial complexity reduction, compared to (2.13). Unfortunately, it is still exponential in the collision degree d_f.

In order to exploit the reduced-size alphabet, the update equations of the MPA messages must be modified as described in [8]. A brief explanation of how this is done is given in the following box. We define this variation as q-MPA.

Message updates in q-MPA

For simplicity, drop the user index and consider a codebook $\mathcal{X}$ such that the projections x_k over FN k assume values from the alphabet $\mathcal{A}_k$, whose size is $|\mathcal{A}_k| = q$.

Suppose that two different symbols from the codebook, $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$: $\mathbf{x} \neq \mathbf{x}'$, have the same projection over FN k, i.e. $x_k = x'_k = x$.

When marginalizing over $\mathcal{X}$, both symbols $\mathbf{x}$ and $\mathbf{x}'$ will produce the same value $y_k - h_{j,k}x$ in $f\left(y_k, \mathbf{X}_{[k]}|\mathbf{H}_{[k]}\right)$ of Equation (2.12a). Thus, it can be concluded that

$$I_{k\to j}(\mathbf{x}) = I_{k\to j}(\mathbf{x}').$$

This means that we can obtain the marginalization by *sweeping only over the q unique projections* $x \in \mathcal{A}_k$, as long as we modify $I_{k\to j}(\mathbf{x})$ with $I_{k\to j}(x)$ $\forall \mathbf{x} : x_k = x$.

• • •

3

SCMA Codebook Design based on Block Codes

The discussion in Chapter 2 reveals that designing SCMA codebooks is mostly done based on single-user MDMs with good properties. Moreover, it is challenging to scale an SCMA system because of the complexity issues of the MPA detection.

In this section, we propose a novel codebook design method. We consider a multistage approach, similar to [7]. Also, the design is independent of $\mathbf{F}$, although the spreading degrees of the codebook and of the resource allocation matrix must match. The main motivation is to have a good trade-off with respect to:

i. good MDM properties;

ii. low-complexity codebook design method;

iii. low-complexity detection.

In [7], the authors used the following design order: (1) choose the alphabet of projections, with $q \leq M$ (2) extend to N-dimensions searching for N integer permutations of length M; and (3) choose a good binary labeling.

The good features of the method in [7] are the relatively simple design steps and the use of $q \leq M$ number of projections, which in turn allow for codebooks with good Peak-to-Average Power Ratio (PAPR) and for a relatively low-complexity receiver if q-MPA [8] is used (as long as $q \ll M$).

The main problem of this method relies on the complexity to find good permutation sets. When M and N slightly scale, the optimal approach becomes intractable. Even the suboptimal way of building the dimensions one at a time has great complexity. Also, when evaluating the codebooks in [7], the authors found strong performance degradation when $q < M$, which poses as a drawback of the method.

3.1 Hamming Distance and Signal-Space Diversity

In order to keep the good features of the method in [7], we also consider designing SCMA codebooks over alphabets of size $q < M$ in our proposed method. This property will keep the good PAPR of the codebooks and also will enable the low-complexity q-MPA MUD.

It turns out that the degradation of performance in the codebooks $A_{8,q}$ (with $q < M$) from [7] is more related to the benchmark resource allocation matrix $\mathbf{F}_{4\times 6}$. This benchmark matrix forces $N = 2$, and thus the maximum signal-space diversity is limited by $N = 2$ complex-dimensions.

Signal-Space Diversity (SSD) is a concept introduced in [68] that exploits the design of the projections of the multidimensional modulations. The definition of SSD is related to the Hamming distance $d_H(\mathbf{c}, \mathbf{c}')$ of two codewords $\mathbf{c}, \mathbf{c}' \in C$. For complex vectors, this means the number of different complex entries between two vectors compared dimension-wise.

Differently from single-user systems, in the uplink of multiple-users, it is not possible to equalize the phase of the different users simultaneously. Thus, the maximum achievable SSD in SCMA is not $2N$, as it is in [67, 68], but only N, due to the correlation between real and imaginary parts.

From the theory of Forward Error Correction (FEC) codes, the minimum Hamming distance $d_{\min}$ of a codebook with codeword length N is limited by an upper bound known as the Singleton bound [76]:

$$M \leq q^{N-d_{\min}+1} \tag{3.1}$$

which can be rearranged as

$$d_{\min} \leq 1 + N - b \tag{3.2}$$

given that $M = q^b$ for any integer value of $b > 0$.

The Singleton bound is not necessarily true for SCMA, since the codebooks are not linearly generated over the same field of the message space[1]. Also, the mapping from the messages to the codewords is not typically linear[2] in SCMA. Another difference is that, in linear FEC codes, it is necessary that the alphabet size q be a power of prime. Table 3.1 helps to summarize differences in the analogy of block codes and SCMA codebooks.

On the other hand, it is evident that SCMA codebooks will only have the largest minimum

[1] Actually, in SCMA, the message space is not defined as a field, although it could be represented by an extension field for codebooks with $M = 2^k$.

[2] However, a few works [49, 51] consider binary expansion as a way to generate linear mappings with some kind of complex-valued generator matrix.

Table 3.1: Parallel comparison between FEC codes and SCMA Codebooks.

	Message Space	**Codebook Space**
Block Codes	Message Alphabet: $\mathcal{M} = \{0, \ldots, q-1\}$ Message Vectors: $\mathbf{m} = \begin{bmatrix} m_1 & \cdots & m_b \end{bmatrix}^T, m_i \in \mathcal{M}$ Message Space Size: $M = q^b$	Codewords Alphabet: $\mathcal{A} = \{0, \ldots, q-1\}$ Codeword Vectors: $\mathbf{c} = \begin{bmatrix} c_1 & \cdots & c_N \end{bmatrix}^T, c_i \in \mathcal{A}$ Number of Codewords: M
SCMA	Message Alphabet: $\mathcal{M} = \{0, \ldots, M-1\}$ Scalar Messages: $m \in \mathcal{M}$ Message Space Size: M	Codewords Alphabet: $\mathcal{A} = \{c_1, \ldots, c_q\} \quad c_i \in \mathbb{C}$ Codeword Vectors: $\mathbf{c} = \begin{bmatrix} c_1 & \cdots & c_N \end{bmatrix}^T, c_i \in \mathcal{A}$ Number of Codewords: M

Hamming distance N *if the codewords have* M *unique projections in each complex-dimension*, that is, if the projections of the codewords are designed such that $c_n \neq c'_n$ for $n = 1, \ldots, N$ and for all pairs $(\mathbf{c}, \mathbf{c}')$ such that $\mathbf{c} \neq \mathbf{c}'$. A very simple design that always achieves this property is the LDS codebooks.

Otherwise, for any SCMA codebook with low number of complex projections per complex dimension $q < M$, the Singleton bound for the minimum Hamming distance will hold true. Thus, for any SCMA codebook with $q < M$, the largest minimum Hamming distance is bounded by $d_{\min} \leq 1 + N - b$, with $b > 1$.

In fact, considering the theory of FEC codes, codebooks that achieve the Singleton bound are classified as MDS codes. When $b = 2$, the MDS codes achieve $d_{\min} = N - 1$.

Based on the previous discussion, we conclude that an M-ary SCMA codebook $C \subset \mathbb{C}^N$, with each complex-dimension designed with complex-projections over a q-ary complex-valued alphabet $\mathcal{A}$, constrained with $q < M$, will have its minimum signal-space diversity L upper-bounded by the following relationship:

$$L \leq N - 1. \tag{3.3}$$

3.2 Underlying Vector-Space Structure

The analogy between FEC codes and SCMA goes beyond the Hamming distance and the SSD of MDMs. The M-length integer permutations $\pi_1, \ldots, \pi_N$ of $\mathcal{A}_q$, used to build the SCMA codebook in the design method [7], can be seen as a list of M N-dimensional vectors whose

entries are taken over the alphabet $\mathbb{Z}_q$.

Consider then a $(N, b, d)_q$ block code, where d is the minimum Hamming distance of the code and q is the alphabet length. Indeed, looking the other way around, any block code can be seen as a stack of N M-length permutations over $\mathbb{Z}_q$ (or $\mathbb{F}_q$, if q is a power of prime). The idea is illustrated in the following example.

Example 1

Consider a binary ($q = 2$) parity check code of length $N = 3$, for messages of $b = 2$ bits (i.e., 1 parity check bit), whose $M = q^b = 4$ codewords are listed in the following matrix structure:

$$C = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}. \tag{3.4}$$

The minimum Hamming distance of this code is $d_{\min} = 2$, which also matches the Singleton bound ($N - b + 1 = 2$). Thus, this is also an MDS code.

If the elements of the field $\mathbb{F}_2$ are replaced with the alphabet $\mathcal{A}_2 = \{x_1, x_2 : x_i \in \mathbb{C}\}$, the resulting structure

$$C = \begin{bmatrix} x_1 & x_1 & x_2 & x_2 \\ x_1 & x_2 & x_1 & x_2 \\ x_1 & x_2 & x_2 & x_1 \end{bmatrix} \tag{3.5}$$

could represent an SCMA Codebook of $M = 4$ codewords, $N = 3$ complex-dimension, signal-space diversity $L = 2$, and with a low number of complex projections $q = 2$.

The integer permutations underlying the structure of this codebook are represented by the matrix $\mathbf{\Pi}^q_{N,M}$:

$$\mathbf{\Pi}^2_{3,4} = \begin{bmatrix} \pi_1(\mathbb{Z}_2 + 1) \\ \pi_2(\mathbb{Z}_2 + 1) \\ \pi_3(\mathbb{Z}_2 + 1) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 1 & 2 & 1 & 2 \\ 1 & 2 & 2 & 1 \end{bmatrix}. \tag{3.6}$$

This is just another way to represent the block code in Equation (3.4), where, for simplicity, we show the permutation entries in the range $[1, q]$ $(\mathbb{Z}_q + 1)$, instead of $[0, q-1]$ $(\mathbb{Z}_q)$.

From the previous example, we note that is also possible to extract a block-code from a known SCMA codebook. The next example illustrates this.

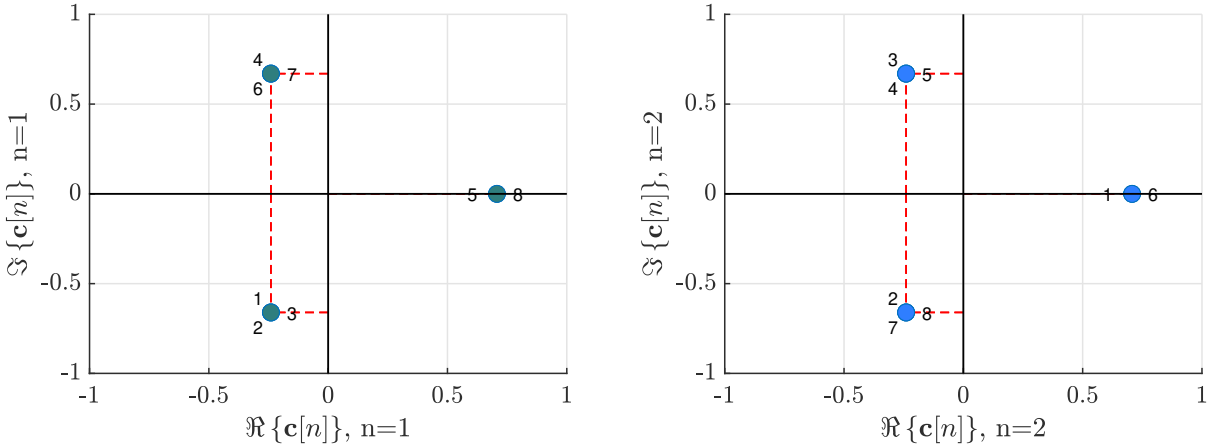
Example 2

Consider the SCMA codebook $A_{8,3}$ from [7]. The design is shown in Figure 3.1. This codebook has the following underlying structure of integer permutations:

$$\mathbf{\Pi}_{2,8}^{3} = \begin{bmatrix} 3 & 3 & 3 & 2 & 1 & 2 & 2 & 1 \\ 1 & 3 & 2 & 2 & 2 & 1 & 3 & 3 \end{bmatrix}. \tag{3.7}$$

This would not be a good block code, since its minimum Hamming distance is $d = 1$. Thus, as a FEC, it would not be able to correct any errors. Actually, this is not even a good SCMA codebook, mainly because of its SSD $L = 1$. In [7], it shows the worst performance compared to the other $A_{8,q}$ codebooks that have $q > 3$. Thus, the performance drawback does not help its low-complexity reduction at the receiver, which is due to the small value of q.

The bottleneck of $A_{8,3}$ is its length, $N = 2$. As a FEC, larger block-length is necessary in order to increase the minimum Hamming distance. The same is valid for increased SSD in SCMA codebooks.



(a) $n = 1$. **(b)** $n = 2$.

Figure 3.1: Codebook $A_{8,3}$ from [7].

We can summarize the above discussion stating that an M-ary N-dimensional SCMA codebook C, with each complex-dimension designed with complex-projections over a q-ary complex-valued alphabet $\mathcal{A}_q$, for any q power of a prime number, has an underlying vector-subset of the vector-space $\mathbb{F}_q^N$, represented by a matrix $\mathbf{\Pi}_{N,M}^{q} \in \mathbb{F}_q^{N \times M}$ whose columns are the elements of the subset.

In the present work, we will concentrate in q being power of prime, since this enables the

use of many types of algebraic block codes based on finite fields $\mathbb{F}_q$.

3.3 Proposed SCMA Codebook Design Method

In this section, we apply the idea of underlying vector-space structure to design SCMA codebooks over small alphabets, but with good MDM properties. Moreover, the proposed method eliminates the exhaustive search of permutations in [7]. The main insight is changing the order of the design steps in [7]: the permutations are generated first, using algebraic operations.

Most of the noteworthy SCMA codebooks rely on a MDM with good properties for single-user scenarios [5, 7]. From the many figures of merit to evaluate MDMs, the SSD L plays an important role. The slope of the asymptotic average pairwise error probability increases with L. Thus, it is expected that MDMs with large SSD L will have better performance, at least in large SNR regimes.

Herein, the SSD is associated with the minimum Hamming distance of a block code. Motivated by this figure of merit, we propose to design the SCMA codebook using the underlying structure of well-known error-correcting block codes. In particular, if MDS codes are used, the SCMA codebooks could achieve a SSD $L = N - b + 1$. Since $b > 1$, the propose method can design SCMA codebooks with SSD at most $L = N - 1$.

Finally, we will focus on structures of linear codes, since it easier to represent a linear code by just specifying its generator matrix and also to span the vector-subspace (codewords) with linear operations.

3.3.1 Step 1: Choosing the Block-Code

Consider a q-ary finite field $\mathbb{F}_q = \{\alpha_1, \alpha_2, \ldots, \alpha_q\}$, where q is a power of prime. A linear block code is the result of a one-to-one mapping between each vector $\mathbf{u}$ of the message-space $\mathbb{F}_q^b$ and the elements $\mathbf{v}$ of the vector-subspace $\mathcal{V} \subset \mathbb{F}_q^N$ generated by a generator matrix $\mathbf{G} \in \mathbb{F}_q^{b \times N}$. Both the message-space and the generated vector-subspace have size q^b.

We propose that an N-dimensional M-ary SCMA codebook C be structured using the vector-subspace generated by $\mathbf{G}$ by associating the field elements $\alpha_1, \ldots, \alpha_q$ with the set $\mathcal{A}_q = \{x_1, \ldots, x_q\}$, where $x_i \in \mathbb{C}$ for $i = 1, \ldots, q$. The elements of $\mathcal{A}_q$ must be unique values and the mapping $\mathbb{F}_q \to \mathcal{A}_q$ must be one-to-one.

If the generator matrix of an MDS linear code is chosen, the resulting minimum Hamming distance of the linear code will be $d_{\text{mds}} = N - b + 1$. This result translates to a SSD $L = d_{\text{mds}}$

in the SCMA codebook C. Since $b = 1$ usually corresponds to repetition codes, we consider SCMA codebooks based on MDS block codes with $b \geq 2$.

In summary,

a. choose b and q, considering that $|C| = M \leq q^b$;
b. choose a generator matrix $\mathbf{G}_{b\times N}$ of an MDS code,
c. use $\mathbf{G}$ to generate a $(N, b, d_{\text{mds}})_q$ MDS code,
d. in each of the codewords, replace the field elements $\alpha_1, \ldots, \alpha_q$ by the elements $x_1, \ldots, x_q$, respectively.

Design Example 1

In this example, we use a Generalized Reed-Solomon (GRS) construction in order to achieve length $N = q$.

Consider a $(N, b) = (4, 2)$ code, so that $d_{\text{mds}} = N - 1 = 3$ and we can use alphabet size $q = 4$. This defines the maximum size of the codebook, $M \leq 16$, and the SSD $L = 3$. We fix $M = 16$ and use all the codewords in order to have a 16-ary MDM.

The $b \times N$ generator matrix of this GRS, in systematic form, is given by

$$\mathbf{G}_{2\times 4} = \begin{bmatrix} 1 & 0 & \alpha & \alpha^2 \\ 0 & 1 & \alpha^2 & \alpha \end{bmatrix}, \tag{3.8}$$

which spans 16 vectors in $\mathbb{F}_4^4$. We replace the field elements $0, 1, \alpha, \alpha^2$ by the respective complex elements of $\mathcal{A}_q = \{x_1, x_2, x_3, x_4\}$ in each dimension.

At this point, C is structured as:

$$C = \begin{bmatrix} x_1 & x_3 & x_4 & x_2 & x_1 & x_3 & x_4 & x_2 & x_1 & x_3 & x_4 & x_2 & x_1 & x_3 & x_4 & x_2 \\ x_1 & x_1 & x_1 & x_1 & x_3 & x_3 & x_3 & x_3 & x_4 & x_4 & x_4 & x_4 & x_2 & x_2 & x_2 & x_2 \\ x_1 & x_4 & x_2 & x_3 & x_2 & x_3 & x_1 & x_4 & x_3 & x_2 & x_4 & x_1 & x_4 & x_1 & x_3 & x_2 \\ x_1 & x_2 & x_3 & x_4 & x_4 & x_3 & x_2 & x_1 & x_2 & x_1 & x_4 & x_3 & x_3 & x_4 & x_1 & x_2 \end{bmatrix}. \tag{3.9}$$

Thus, the structure of this GRS code resulted in an SCMA codebook with parameters $(M, N, q, L) = (16, 4, 4, 3)$.

Design Example 2

Consider the codebook design for $M = 8$ codewords and spreading degree $N = 4$. We set $q = 3$ and $b = 2$.

Using a Hamming Code (HC) construction for $(N, b, d)_q = (4, 2, 3)_3$, we obtain an MDS code with $d_{\text{mds}} = 3$. The $b \times N$ generator matrix in systematic form, given by

$$\mathbf{G}_{2\times 4} = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 2 \end{bmatrix}. \tag{3.10}$$

$\mathbf{G}_{2\times 4}$ is able to span a 9-ary vector-subspace of $\mathbb{F}_3$, which is actually an MDS code. We denote this SCMA codebook design as the HC construction.

In order to obtain a 8-ary codebook we must expurgate one codeword. This procedure will be done using figures of merit detailed in the following Section 3.3.2.

3.3.2 Step 2: Designing the Complex Projections $\mathcal{A}_q$

Once the underlying vector-subspace structure is generated, it is necessary to choose the q complex-values of the codebook projections in $\mathcal{A}_q$. In [7], $\mathcal{A}_q$ is arbitrarily chosen as an APSK constellation as the first step of codebook design.

Since we have a fixed codebook structure, we relax the choice of $\mathcal{A}_q$ to any set of complex-values. It is possible to optimize this choice maximizing some MDM figure of merit. The previously designed SSD L is not affected by this step, as long as the q projections have no multiplicity, i.e. $x_i \neq x_j$ for all $i \neq j$.

Also, since the MPA receiver calculates the messages independently in each resource, the codebooks could also have a different set $\mathcal{A}_q$ in each of the N dimensions, i.e., $\mathcal{A}_q^{(n)}$ for $n = 1, \ldots, N$. The choice of the N sets $\mathcal{A}_q^{(n)}$ can optimize a codebook figure of merit of interest.

For example, the mean cutoff rate approximation $\Psi(C)$ suggested in [7, 37] can be used in the search for the qN projections. Thus, the problem could be described as:

$$\begin{aligned} &\underset{\mathcal{A}_q^{(1)},\ldots,\mathcal{A}_q^{(N)}}{\text{maximize}} && \Psi(C) \\ &\text{subject to} && E\left[\|\mathbf{c}\|^2\right] = 1 \end{aligned} \tag{3.11}$$

where

$$\Psi(C) = \log_2 M - \log\left[1 + \frac{1}{M}\sum_{\mathbf{c}\in C}\sum_{\substack{\mathbf{c}'\in C\\ \mathbf{c}'\neq\mathbf{c}}}\prod_{n=1}^{N}\left(1 + \frac{\left|c_n - c'_n\right|^2}{4N_0}\right)^{-1}\right]. \tag{3.12}$$

Other figures of merit to be considered are the minimum (squared) Euclidean distance $\min_{i,j} d_E^2\left(\mathbf{c}_i, \mathbf{c}_j\right)$, the minimum L-product distance $\min_{i,j} d_{p,L}\left(\mathbf{c}_i, \mathbf{c}_j\right)$ and the PAPR.

During simulation analysis, we observed that optimizations over $\Psi(C)$ do not improve significantly the performance of the system when compared with arbitrary choices for $\mathcal{A}_q$ such as APSK projections, especially for low q such as 3 or 4.

Thus, for codebooks with small values of q, we simplify $\mathcal{A}_q^{(n)} = \mathcal{A}_q\ \forall n$ by choosing arbitrary q complex projections among known constellations such as QAM and APSK.

On the other hand, for cases where q^b is greater than the desired M and once $\mathcal{A}_q$ is fixed, other optimizations may be considered, even for arbitrary $\mathcal{A}_q$. In this case, it is possible to expurgate $q^b - M$ vectors in order to optimize some figure of merit dependent on C, e.g. maximizing $\Psi(C)$.

For the HC in the second design example of Section 3.3.1, we chose arbitrary APSK projections for $\mathcal{A}_3$ and we optimized the MDM over $\Psi(C)$ by removing the HC codeword produced by the message vector $\mathbf{u} = (2, 2)$.

Design Example 3

Consider the following design of an SCMA codebook: $M = 64$ codewords, $N = 4$ complex-dimensions. By using a GRS code, we can pick $q = 8$ and $b = 2$, such that the minimum SSD will be $L = N - b + 1 = 3$. The $(4, 2)$ block code is designed over $\mathbb{F}_{2^3}$ with the generator matrix:

$$\mathbf{G}_{2\times 4} = \begin{bmatrix} 1 & 0 & \alpha+1 & \alpha^2+1 \\ 0 & 1 & \alpha & \alpha^2 \end{bmatrix}. \tag{3.13}$$

To design the alphabets $\mathcal{A}_q$, we consider random searches (2500 iterations) for q amplitudes $c_q \sim \mathcal{N}(0, 2)$ and q angles $\theta_q \sim \mathcal{U}(-\pi, \pi)$.

Assuming a maximization problem, the objective function can be d_E^2, $d_{p,L}$ or $\Psi(C)$ (at 10 dB of E_b/N_0).

Considering them independently, for three given search realizations, we were able to find the following tuple of metrics:

a. $\left(d_E^2, d_{p,L}, \Psi(C)\right) = (0.3613, 0.0419, 5.1381)$, when maximizing d_E^2,

b. $\left(d_E^2, d_{p,L}, \Psi(C)\right) = (0.4230, 0.0514, 5.2063)$, when maximizing $d_{p,L}$,

c. $\left(d_E^2, d_{p,L}, \Psi(C)\right) = (0.3366, 0.0549, 5.2157)$, when maximizing $\Psi(C)$.

Consider an objective function $f = w_1 \times d_E^2 + w_2 \times d_{p,L} + w_3 \times \Psi(C)$. Optimization can be performed maximizing the linear combination of the metrics. The set of weights w_1, w_2 and w_3 can also be tuned. When using $w_1 = w_3 = 1$ and $w_2 = 10$, we found $\left(d_E^2, d_{p,L}, \Psi(C)\right) = (0.4645, 0.0605, 5.1816)$.

However, we remark that results may vary in each realization, since random search is not an efficient optimization method. Proper heuristic algorithms with more structure could improve the results.

Comparing with an arbitrary choice of an APSK alphabet, using projections parametrized with $\mathbf{m} = [4, 4]$ (see [7]), we obtain $\left(d_E^2, d_{p,L}, \Psi(C)\right) = (0.5770, 0.0830, 5.2064)$.

Although the random optimization showed that large values of $\Psi(\mathcal{X})$ can be found, the arbitrary solution with two-radius APSK seems to provide a good trade-off, also with the largest euclidean-distance and product-distance outputs.

3.3.3 Step 3: Designing the Binary Labeling

Once the codebook C is fully designed from Steps 1 and 2, the SER of the system is defined for a given MUD. Yet one can perform optimization by choosing a binary labeling function μ among the $M!$ possible solutions in order to improve the resulting BER.

For classical one-dimensional (complex) rectangular QAM constellations, Gray-code labeling provides a reduction factor of $1/k$ from the SER to the BER curve. For general MDMs, the BSA has been proposed in [77] in order to find a good labeling by optimizing a cost function based on the Euclidean distance between codeword pairs.

In our method, we also consider the BSA approach. Thus, the last step of our design method is the same last step of the method in [7].

3.3.4 Complexity Reduction in the Codebook Design Stage

To extend the design from 1 to N dimensions, Step 2 of the codebook design method in [7] proposes exhaustive search of integer permutations, which has complexity $O\left((M!)^{N-1}\right)$, intractable even for small values of M and N such as $M = 8, N = 3$. Even the suboptimal approach in [7] has complexity $O\left((N-1) \times M!\right)$ which is also intractable for high order constellations.

In [24], the codebook design involves evaluating a cost function with complexity $O\left(M^{d_f}\right)$ (same complexity as the original MPA receiver), which clearly is not a practical design as M and d_f increase; additionally, in a second step of their method, they also face the permutation search problem of [7].

In the Step 1 of our proposed method, we obtained a set of integer permutations in a straightforward manner, using $\mathbf{G}_{b\times N}^{T}$ and linear operations to span a vector space. The complexity order of this procedure is $O\left(M\right)$. Hence, our proposed method avoids large complexity during the codebook design stage and enables a straightforward design of high-order constellations, with the extra benefit of ensuring arbitrary SSD up to the MDS bound. If desired, one may also use non-MDS codes, which limits the SSD to $L < N - b + 1$, but may relax some design constraints.

3.3.5 On the Complexity Order of the q-MPA for High-Dimensional SCMA Codebooks

Regardless of the codebook design method, when using the MPA as MUD, the receiver complexity is still exponential, as discussed in Section 2.5.

However, using $q < M$ complex projections per dimension in the codebook can significantly reduce the complexity of the MPA receiver [8]. In this sense, by using small alphabet size $q < M$ in our proposed Step 1, we maintain the benefits of codebooks proposed in [8].

Even assuming the use of codebooks with low number of projections, with a small value for q, and the use of the corresponding q-MPA, there is an issue with the exponent d_f in the complexity order of Equation (2.14).

It can be easily shown that $d_f = \beta N$, where $\beta = J/K > 1$ represents the SCMA load factor. This implies that increasing the load, or the spreading degree N, will dramatically increase the MPA complexity order. Thus, we face a trade-off. Assuming $q < M$ implies that the minimum signal-space diversity of $\mathcal{C}$ will be lower than N, which can cause a degradation in the error rate performance [73]. To balance this effect, codebooks with larger N could be designed, but

this will cause an increase in the average and maximum collision degrees. The only remaining degree of freedom would be the load factor, which could be reduced at the cost of reducing the spectral efficiency.

We can rewrite Equation (2.14) to emphasize this trade-off, replacing d_f with βN:

$$O_{\text{MPA}} = T_o T_m \left[JN \left(q^{\beta N} + M \right) - JM \right]. \tag{3.14}$$

Thus, the effect of the collision degree d_f still poses a problem for increasing values of load factor and spreading degree, and consequently, for the use of codebooks with large SSD.

Regardless of this problem, when considering the use of q-MPA as MUD, it is desirable to use q as small as possible. Nonetheless, there are other constraints related to block codes that may limit the choice of q during the codebook design procedure, depending on (N, b, d). We discuss it in the following section.

3.3.6 The Main Conjecture of Linear Codes

Note that, in MDS codes, given the message-length b and the alphabet size q (i.e., codebook size $M = q^b$), the minimum distance of the code d_{mds} is proportional to the block length N of the code [78].

Let $N_{\max}(b, q)$ be the maximum length of an MDS block-code, with fixed b and q. According to [78], there is a conjecture denoted as *"the main conjecture" on linear MDS codes*, originally presented in [79], which disposes that:

$$N_{\max}(b, q) = \begin{cases} q + 2 & \text{if } q = 2^t \text{ and } b \in \{3, q - 1\} \\ q + 1 & \text{otherwise} \end{cases}. \tag{3.15}$$

Table 3.2 shows many results, reviewed in [78], for the upper bound on the block-code length under many conditions of the parameters b and q. The last column of the table also shows the MDS minimum distance in each case.

In [80], a generator matrix is shown for the construction of linear MDS codes based on Reed-Solomon (RS) construction, for the parameters $N = q + 1$, with q power of prime. The $b \times N$ generator matrix $\mathbf{G}$ is given as:

Table 3.2: Maximum block-code lengths $N_{\max}(b,q)$ reviewed in [78] (table adapted from the original work).

Condition	$N_{\max}(b,q)$	d_{mds}
[79] $\begin{cases} q = 2^t \text{ and } [b = 3 \text{ or } b = q-1] \\ \text{otherwise} \end{cases}$	$\begin{cases} = q+2 \\ = q+1 \end{cases}$	$\begin{cases} = q-b+3 \\ = q-b+2 \end{cases}$
$\forall\, b,q$	$> b+1$	≥ 2
$b \geq q$	$= b+1$	$= 2$
$b \geq 2$	$\leq q+b-1$	$\leq q$
p prime, $p < b$, $p \mid (q-1)$	$\leq q+b-2$	$\leq q-1$
$b \geq 4$, $q > 2$, $36 \mid q$; q odd	$\leq q+b-2$	$\leq q-1$
$b \geq 4$, $q > 2$, $36 \mid q$; q even	$\leq q+b-3$	$\leq q-2$
$b \geq 3$, $q > 2$, $4 \mid q$	$\leq q+b-2$	$\leq q-1$
$b \leq q$ e $q = p^h$, for p prime	$\leq q+b+1-\min(b,p)$	$\leq q+2-\min(b,p)$
$q = p^h$, for p prime and $p < b < q$	$\leq q+b-p$	$\leq q-p+1$
$q = p^h$, for p prime and $2 < q-p+1 < b < q-2$	$\leq q+1$	$\leq q-b+2$
$q = p^h$, for p prime and $h > 1$ e $b \leq 2p-2$	$\leq q+1$	$\leq q-b+2$
$q = p^h$, for p prime and $h > 1$ e $q-2p+4 \leq b \leq q$	$\leq q+1$	$\leq q-b+2$

$$\mathbf{G}_{b\times N} = \begin{bmatrix} 1 & 1 & \cdots & 1 & 0 \\ g_1 & g_2 & \cdots & g_q & \vdots \\ \vdots & \vdots & & \vdots & 0 \\ g_1^{b-1} & g_2^{b-1} & \cdots & g_q^{b-1} & 1 \end{bmatrix}, \tag{3.16}$$

where $g_i \in \mathbb{F}_q$, for $i = 1,\ldots,q$, are the finite field elements of the alphabet $\mathbb{F}_q$. When using this construction in an SCMA codebook, we will refer to it as the SB-RS codebooks. A design example is shown below.

Design Example 4

A SCMA codebook will be designed using the SB-RS construction of Equation (3.16). The goal is to design an 8-ary codebook ($M = 8$) with SSD $L = 4$.

To achieve the largest SSD with the smallest N, we need to set b as low as possible, thus $b = 2$. This results in $N = L + b - 1 = 5$ complex-dimensions.

Then, we know that $M \leq q^b$, so $q \geq \sqrt{8} \approx 2.8$. The choice $q = 3$ would be interesting to ensure low-complexity for the q-MPA. However, from Table 3.2, there's only one condition (the main conjecture) that would achieve $L = 5$, and this condition demands that $q = 2t$ (even q). Thus, we need larger q. If we choose $q = 4$, we can use the SB-RS construction, resulting in $N = q + 1 = 5$ complex-dimensions and SSD $L = N - 1 = q = 4$.

Then, the SB-RS 2×5 generator matrix is:

$$\mathbf{G}_{2\times 5} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 \\ 0 & \alpha & \alpha^2 & 1 & 1 \end{bmatrix}. \tag{3.17}$$

Note that this matrix is able to span 16 codewords, since the message-space size is $q^b = 4^2 = 16$. In order to obtain an 8-ary codebook, we need to expurgate half of these codewords.

Since q is small, we choose APSK projections for $\mathcal{A}_q$ and exhaustively search for the best 8 out of 16 codewords that maximize the cutoff rate $\Psi(C)$. This results in the following codebook:

$$C = \begin{bmatrix} x_1 & x_3 & x_4 & x_2 & x_1 & x_3 & x_4 & x_2 \\ x_1 & x_1 & x_1 & x_1 & x_3 & x_3 & x_3 & x_3 \\ x_1 & x_2 & x_3 & x_4 & x_4 & x_3 & x_2 & x_1 \\ x_1 & x_4 & x_2 & x_3 & x_2 & x_3 & x_1 & x_4 \\ x_1 & x_2 & x_3 & x_4 & x_2 & x_1 & x_4 & x_3 \end{bmatrix} \tag{3.18}$$

where $\mathcal{A}_q = \{x_1, x_2, x_3, x_4\}$ is designed with 1 APSK ring specified by the vector $\mathbf{m} = \begin{bmatrix} 4 \end{bmatrix}$ (see details of the multiple-rings APSK design in [7]).

3.4 Designing F for High-Dimensional Codebooks

Most papers on SCMA [3, 7, 24] are restricted to the analysis of a low dimension 4×6 matrix and codebooks with $N = 2$ dimensions. However, as pointed out in [15, 16], the performance of the MPA receiver can be improved when using matrices with larger girth (minimum length of a cycle in a Tanner graph), which is not possible in very dense (low dimensional) matrices.

Moreover, to the best of our knowledge, no other work has considered feasible codebooks with $N > 2$, limiting the SSD L of the MDM to be $1 \leq L \leq 2$. Since we want to improve the MPA performance, with high-dimensional codebooks, we consider matrices with $N > 2$ and we design $\mathbf{F}$ with larger dimensions in order to find larger girth and less dense matrices.

We choose to use PEG [74] algorithm to find a $K \times J$ binary matrix $\mathbf{F}$ with regular column-degree N. The impact of these design guidelines will be evaluated in Section 3.5. A design

example can be seen in Figure 3.2 for input parameters $(J, K, N, d_f) = (20, 16, 4, 5)$.

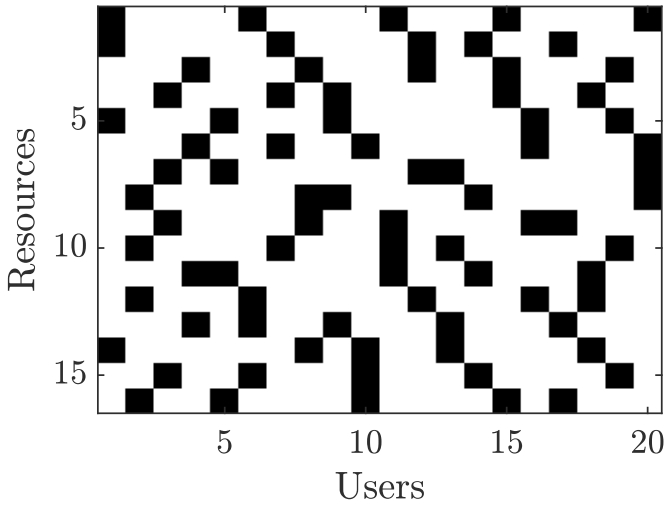


Figure 3.2: Design example of the resource allocation matrix $\mathbf{F}$ for $J = 20$ users, $K = 16$ resources and $N = 4$. Black squares represent entries $F_{k,j} = 1$, corresponding to resource k allocated to user j. White spaces represent null entries, meaning that the user does not spread over this resource. It has girth 6 (480 cycles) and density 25%.

3.5 Performance Evaluation

In this section, we evaluate codebooks designed with the proposed method. We ran numerical simulations of the link level communication in the baseband model, evaluating the bit-error rate as the key performance indicator.

Monte Carlo simulations are done over the i.i.d. Raleigh channel. The system is uncoded. In all simulations, the MUD is the q-log-MPA.

We consider several setups of parameters for the codebook design.

3.5.1 Codebooks with $N = 4$

In this section, we consider the design of codebooks with $N = 4$ complex dimensions. We reuse the designed codebooks from Design Examples 1 and 2, based on the GRS and HC constructions, respectively.

To the best of our knowledge, there are no published SCMA schemes with $N = 4$ and tractable decoding complexity.

Since the codebooks in [7] have outperformed the ones in [5], and since we consider that the design complexity of [24] is not practical, even for moderate M and d_f, we compare our proposed MDMs with 2 different codebooks that we have designed following the method from [7]. We have considered APSK based $A_{M,q}$ for $(M, q) = \{(8, 3), (16, 4)\}$. Since exhaustive search of the permutations is not tractable, we considered the suboptimal approach proposed for the

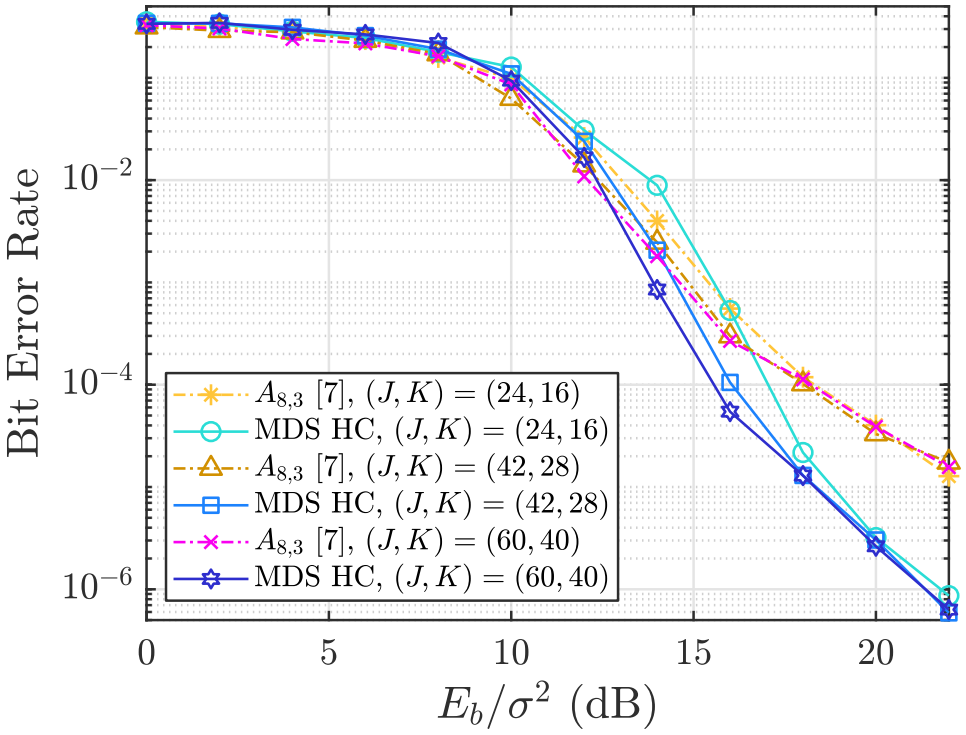


Figure 3.3: MPA performance for different density levels of the resource allocation matrices $\mathbf{F}$. The load is fixed to $J/K = 150\%$ ($d_f = 6$), and the codebook structure is the HC-based one, with $(N,k,M,q) = (4,2,8,3)$. $\mathbf{F}_{24\times16}$ and $\mathbf{F}_{42\times28}$ have girth 4 (31 and 6 cycles, respectively) and $\mathbf{F}_{60\times40}$ has girth 6 (667 cycles).

step 2 in [7] and we searched the permutations randomly over $T = 10^5$ iterations per dimension (complexity $O((N-1)T)$).

Both the proposed and benchmark codebooks consider APSK projections which are distributed over the rings according to vectors $\mathbf{m} = [3]$ and $\mathbf{m} = [4]$ for $q = 3$ and $q = 4$, respectively (see more details in [7]). In Table 3.3, we provide values of some figures of merit of these codebooks.

In Figure 3.3, the results for the HC codebook are shown. Note that the larger diversity of our designs outperforms the $A_{M,q}$ codebooks in the large-SNR regime. Also, it can be observed that the large-sized matrices improve the performance of the system, specially for the MDS codebooks.

Similar results are found for the 16-ary codebooks. The MDS codebooks have the best performance in the high-SNR regime, where the larger signal-space diversity dominates the slope of the BER. The moderate-SNR region has an improvement when the size of the matrix $\mathbf{F}$ is larger, since it allows for a better design with the PEG algorithm.

The spectral efficiency of these schemes is given by the following relationship:

$$\mathrm{SE} = \beta \times \log_2 M. \tag{3.19}$$

Table 3.3: Figures of merit comparison for the M-ary 4-dim. codebooks.

Design Methods	(M,q,N)	$\min_{i,j} d_E^2(\mathbf{c}_i,\mathbf{c}_j)$	$\min_{i,j} d_{p,L}(\mathbf{c}_i,\mathbf{c}_j)$	L	$\Psi(C)$ @ 8 dB
Proposed Codebooks	$(8,3,4)$	2.25	0.6495	3	2.2864
	$(16,4,4)$	2.00	0.5000	3	3.0167
Method from [7]	$(8,3,4)$	1.50	0.7500	2	2.2337
	$(16,4,4)$	1.00	0.5000	2	2.9548

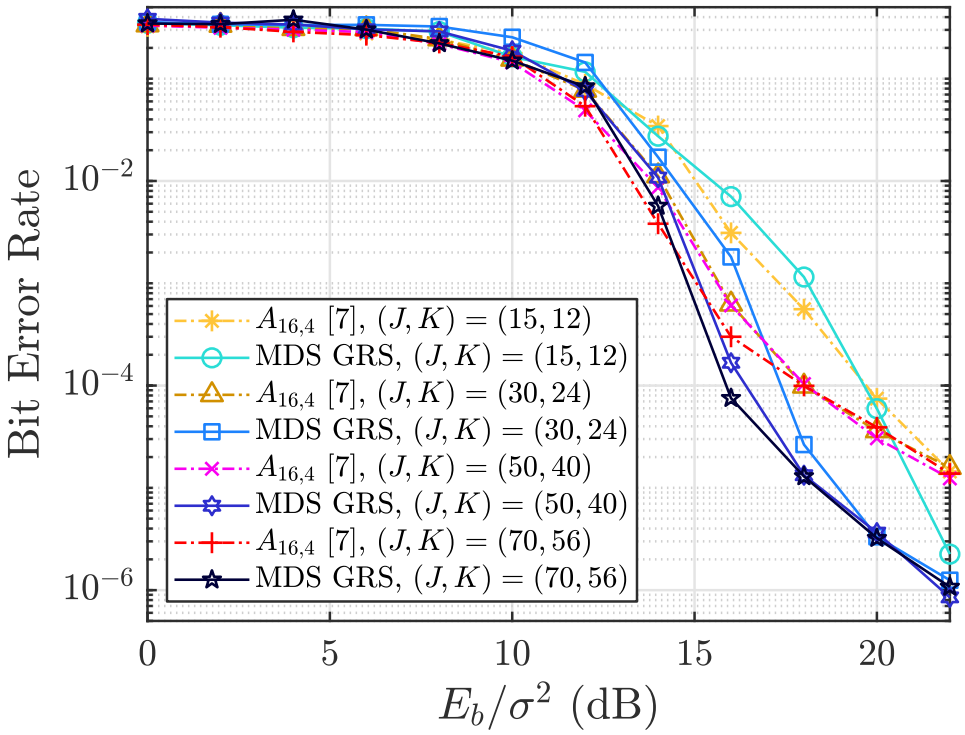


Figure 3.4: MPA performance for different density levels of the resource allocation matrices $\mathbf{F}$. The load is fixed to $J/K = 125\%$ ($d_f = 5$), and the codebook structure is the GRS-based one, with $(N,k,M,q) = (4,2,16,4)$. $\mathbf{F}_{15\times12}$ and $\mathbf{F}_{30\times24}$ have girth 4 (25 and 2 cycles, respectively), and $\mathbf{F}_{50\times40}$ and $\mathbf{F}_{70\times56}$ have girth 6 (320 and 268 cycles, respectively).

In the first scenario, SE $= 1.5 \times 3 = 4.5$ bits/resource. In the second scenario, SE $= 1.25 \times 4 = 5$ bits/resource. Thus, although the overload is smaller in the second scenario, we were able to increase the spectral efficiency with a high-order modulation, without compromising significantly the BER nor the receiver complexity.

3.5.2 Effect of increasing N

In this Section, we consider 8-ary codebooks ($M = 8$) and several values of spreading degree: $N \in \{3,4,5,6\}$. The design for $N = 4$ was already given by the HC codebook. For the other designs, we consider the following approaches.

For $N = 3$, we build codebook $(3,2)$-GRS over $\mathbb{F}_3$. The code's generator matrix is given as

$$\mathbf{G}_{2\times 3} = \begin{bmatrix} \mathbf{I}_2 & \begin{bmatrix} 2 & 2 \end{bmatrix}^T \end{bmatrix}, \tag{3.20}$$

where $\mathbf{I}_n$ is the $n \times n$ identity matrix. Since $\mathbf{G}_{2\times 3}$ is able to span 9 codewords, we follow the expurgation procedure described in Section 3.3.2, with the criteria of maximizing the cutoff rate $\Psi(C)$. The solution leads to the expurgation of the codeword $\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T$. The mapping of $\mathbb{F}_3$ to $\mathcal{A}_3$ is straightforward, $i \rightarrow x_i$ for $i = 0, 1, 2$. The elements of $\mathcal{A}_3$ are taken arbitrarily from APSK projections, which are distributed over the rings according to the parametric vector $\mathbf{m} = [3]$ (see [7]).

Codebooks for $N = 5$ and $N = 6$ are based on RS codes over $\mathbb{F}_2^2$ ($q = 4$) and $\mathbb{F}_5$ ($q = 5$), respectively. The generator matrices are constructed following the method in [80], thus we denote them $(5,2)$-SB-RS and $(6,2)$-SB-RS. The matrices are given as

$$\mathbf{G}_{2\times 5} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 \\ 0 & \alpha & \alpha^2 & 1 & 1 \end{bmatrix}, \tag{3.21}$$

and

$$\mathbf{G}_{2\times 6} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 3 & 4 & 1 \end{bmatrix}. \tag{3.22}$$

Similarly to $(3,2)$-GRS, $\mathcal{A}_q$ is defined arbitrarily for codebook $(5,2)$-SB-RS, with APSK projections using $\mathbf{m} = [4]$.

The complex projections of codebook $(6,2)$-SB-RS were chosen following the general optimization described in Section 3.3.2. The figure of merit was $\Psi(\mathcal{X})$. The resulting projections are

$$\begin{aligned} \mathcal{A}_5 = \{&0.3767e^{-i0.82\pi}, 0.3793e^{-i0.53\pi}, 0.3882e^{i0.45\pi}, \\ &0.4204e^{i0.57\pi}, 0.4713e^{-i0.04\pi}\}. \end{aligned}$$

Since $(5,2)$-SB-RS spans 16 codewords and $(6,2)$-SB-RS spans 25 codewords, expurgation with respect to the maximization of $\Psi(\mathcal{X})$ is also employed to obtain 8-ary codebooks. Because there are many expurgated codewords to be listed, we leave this procedure to be carried on by the reader.

Table 3.4 summarizes the figures of merit for the designed codebooks. In Figure 3.5,

Table 3.4: Designed 8-ary codebooks and its parameters.

Codebook ID	N	q	L	$\min_{i,j} d_E^2(\mathbf{c}_i, \mathbf{c}_j)$	$\min_{i,j} d_{p,L}(\mathbf{c}_i, \mathbf{c}_j)$
$(3,2)$-GRS	3	3	2	2.0000	1.0000
$(4,2)$-HC	4	3	3	2.2500	0.6495
$(5,2)$-SB-RS	5	4	4	2.0000	0.2263
$(6,2)$-SB-RS	6	5	5	1.7761	0.0238

we compare the performance of these codebooks, alongside with $A_{8,8}$ from [7], which has $N = 2$ complex-dimensions. The scenario considers $J = 20$ users sharing $K = 15$ resources ($\beta = 133\%$) in an i.i.d. Rayleigh channel. We apply MPA with $T_m = 5$ iterations. In this setup, the codebooks' length $N = 2,3,4,5$ and 6 correspond to average collision degrees $\overline{d_f} = 2.67, 4, 5.33, 6.67$ and 8, respectively. The non-integer values of $\overline{d_f}$ assume slightly irregular degrees (a few FNs have degree $d_k = \lceil \overline{d_f} \rfloor$).

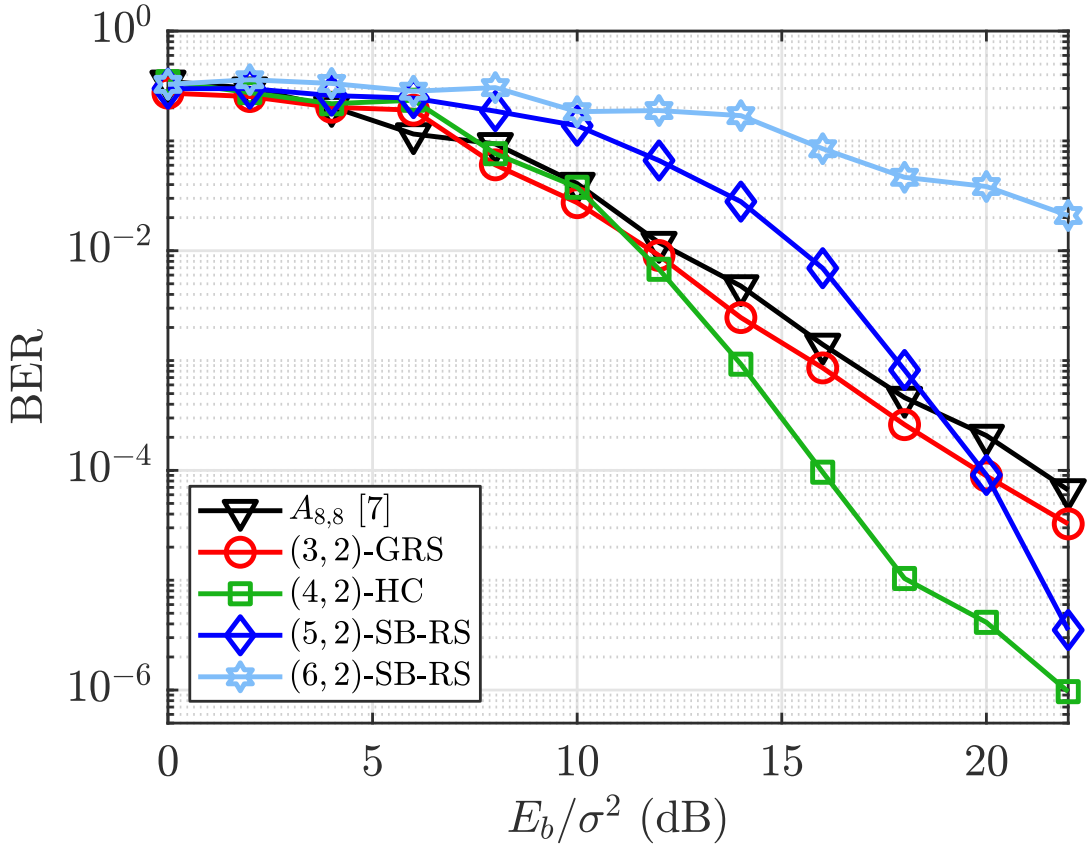


Figure 3.5: BER performance for codebooks of size $M = 8$ with several codeword lengths N.

The results show that $(3,2)$-GRS has a better performance than $A_{8,8}$, because of the shaping gain of using a smaller alphabet size q, even though the former has greater length than the latter. The codebook with the best performance in this scenario is $(4,2)$-HC, since it combines a small alphabet with a larger SSD.

The codebooks with $N = 5$ and $N = 6$ have been designed to have larger SSD. However, it can be noticed that the effect of the SSD only shows up in the large E_b/σ^2 regime. Moreover, the codebook $(6, 2)$-SB-RS increases the collision degree to a very large value ($d_f = 8$), which degrades the performance in the simulated range of SNR.

3.5.3 Codebooks with large M

In [7], a 64-ary codebook with 4 complex dimensions is illustrated, with an alphabet of $q = 16$ complex projections. This codebook is named $A_{64,16}$. However, the authors do not show simulation results for this codebook.

In this Section, we consider the codebook of Design Example 4, with $M = 64$ codewords over $N = 4$ complex-dimensions, using an alphabet of $q = 8$ complex projections in each dimension.

We compare our design with the codebook $A_{64,16}$. It is clear that even with q-MPA, the complexity order of $A_{64,16}$ is $O\left(16^{d_f}\right)$, while our proposed MDS codebook has complexity order $O\left(8^{d_f}\right)$. Also, the codebook $A_{64,16}$ has as signal-space diversity $L = 2$, while our MDS design ensured $L = 3$.

Since the complexity is high even for moderate collision degree, we evaluate this codebooks in a scenario of low level overload, $\beta = 125\%$, such that $d_f = 5$. Note that the complexity for $A_{64,16}$ is still very high, $O\left(T_m J N 10^6\right)$.

For this simulation, we have designed a resource allocation matrix for $J = 30$ users sharing $K = 24$ resources. Note that the spectral efficiency is high, $\text{SE} = \beta \log_2 M = 1.25 \times 6 = 7.5$ bits/resource. Figure 3.6 shows the BER performance of these codebooks for the i.i.d. Rayleigh channel, considering $T_m = 5$ iterations of the MPA.

Results show that $A_{64,16}$ has a poor performance in the low to moderate SNR region. The proposed MDS codebook also begins its slope in a moderate SNR point, but it has a very good BER close to 10^{-6} at $E_b/\sigma^2 = 28$ dB. This gain is a combined effect of a larger SSD ($L = 3$) and smaller alphabet size ($q = 8$).

The proposed outperforms $A_{64,16}$ not only regarding the BER, but also with respect to the complexity order of the MPA. For the proposed scenario, the GRS codebook has complexity order of $O\left(T_m J N q^{d_f}\right) \approx O\left(1.96608 \times 10^7\right)$ (related only to the FN operations, since the VNs do not burden significantly the receiver complexity). The codebook $A_{64,16}$ has a complexity order of approximately $O\left(6.29146 \times 10^8\right)$ (32 times larger).

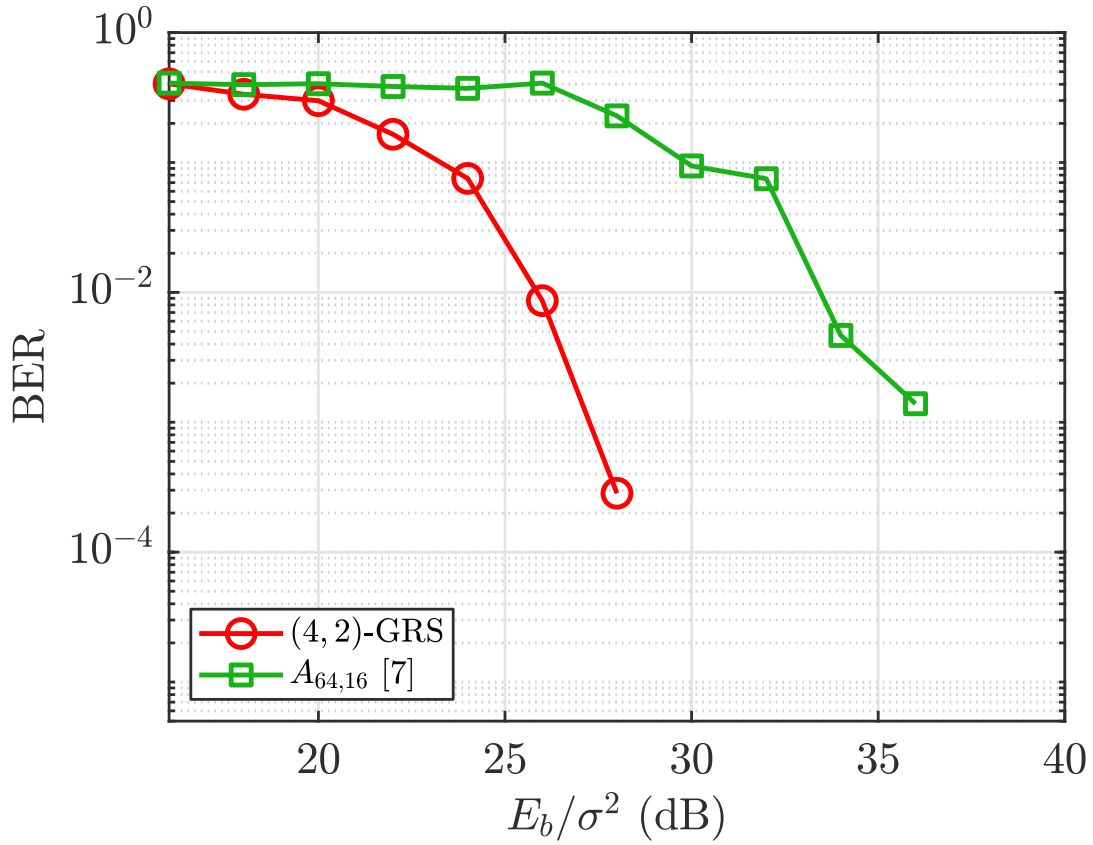


Figure 3.6: BER performance for two codebooks of size $M = 64$ in a scenario of load factor $\beta = 125\%$.

• • •

4

Threshold-based Edge-Selection MPA for SCMA Receivers

This chapter presents an improved version of the ESGA-MPA from [9, 10]. The modification to ESGA considers a threshold-based selection (instead of fixed) of the strongest neighboring edges. Thus, we dynamically apply partial GA on interference, according to the channel gains realizations in each resource.

The probability of a function node evaluating $D + 1$ neighboring users in each message calculation is also derived for the i.i.d. Rayleigh fading channel. From this derivation, we find an analytical expression for the complexity order of our proposed algorithm.

In addition, we also combine our MUD with codebooks of low number of projections, in order to enable q-MPA. Results show that, with a proper design, it is possible to arbitrarily reduce the receiver complexity while maintaining a good performance trade-off.

4.1 Revisiting the System Model

In this Chapter, we consider bit-interleaved coded modulation schemes with iterative detection and decoding. Without loss of generality, user j encodes its data $\mathbf{m}_j \in \mathbb{F}_2^k$ into a codeword $\mathbf{d}_j \in \mathbb{F}_2^n$ using a FEC of rate $R = k/n$. Neglecting the use of bit interleavers (π), each codeword is divided into N_s packets of b bits each, such that $n = N_s \times b$.

Each packet of b bits is mapped to an N-dimensional SCMA codeword $\mathbf{c}_j$ of an M-ary MDM $C \subset \mathbb{C}^N$, where $M = 2^b$, which is the SCMA "mother-codebook". The spectral efficiency is rewritten as:

$$\text{SE} = R \times \beta \times \log_2 M. \tag{4.1}$$

The remaining of the model is the same as in Chapter 2. At the receiver, a MUD employs T_m iterations for each of the N_s channel uses. The soft information of all the n bits of each user is then forwarded to J parallel FEC decoders. Outer loops with T_o iterations can be employed

in the turbo receiver configuration, as illustrated in Figure 4.1.

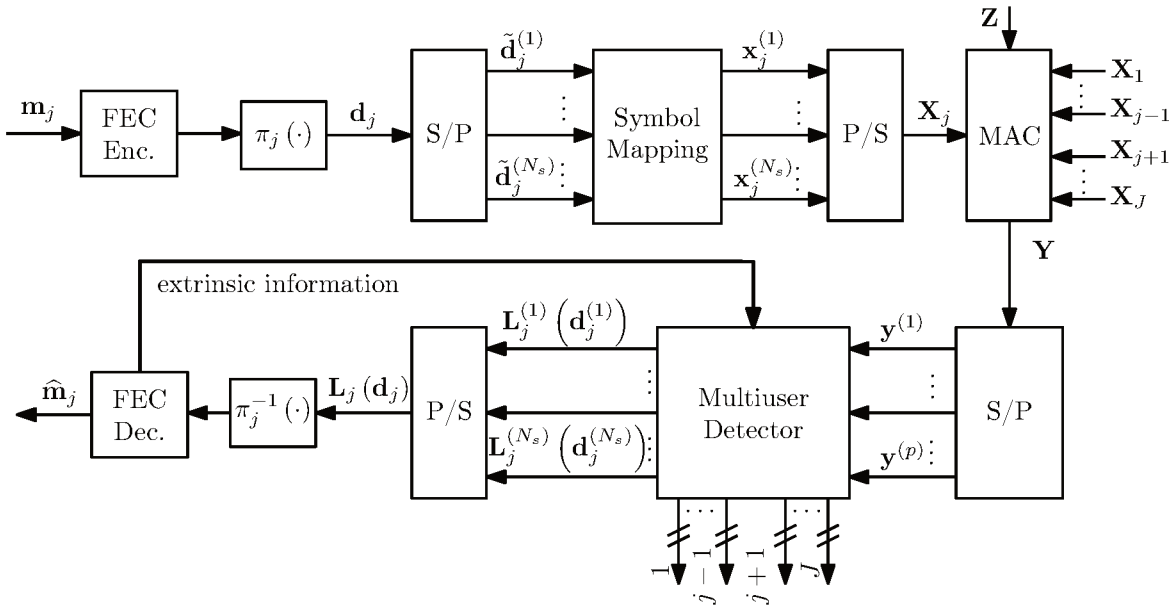


Figure 4.1: Block diagram of the SCMA system model considering bit-interleaved coded modulation with iterative detection and decoding.

4.2 ESGA-MPA in [9, 10]

Several other works have proposed receiver modifications in order to tackle the complexity issues of the original MPA. The edge selection approach in [9, 10] can arbitrarily reduce the exponent in the complexity order, with the use of partial GA on the interference of Equation (2.11). However, when the complexity order tends to linear with respect to d_f (by cutting too many edges), performance degradation is shown to be significant, to an extent that even powerful FECs are rendered useless.

From Equation (2.1), we can write the sample y_k, received in FN k, as

$$y_k = h_{j,k}x_{j,k} - \sum_{l\in\omega_{j,k}} h_{l,k}x_{l,k} - \underbrace{\sum_{i\in\overline{\omega_{j,k}}} h_{i,k}x_{i,k} + z_k}_{z_{j,k}}, \tag{4.2}$$

where $\omega_{j,k}$ and $\overline{\omega_{j,k}}$ are disjoint sets of "strong" and "weak" neighboring users (of user j), respectively. Also, the union of the two sets builds $\phi_{k\setminus j}$, i.e. $\phi_{k\setminus j} = \omega_{j,k} \cup \overline{\omega_{j,k}}$.

In [9, 10], the set $\omega_{j,k}$ of the "strongest" neighbors of user j is built taking the d_s strongest out of the $d_f - 1$ users in $\phi_{k\setminus j}$, with respect to their channel gains $\left|h_{j,k}\right|^2$. The remaining $d_f - 1 - d_s$ users compose $\overline{\omega_{j,k}}$, and are treated as interference along with AWGN.

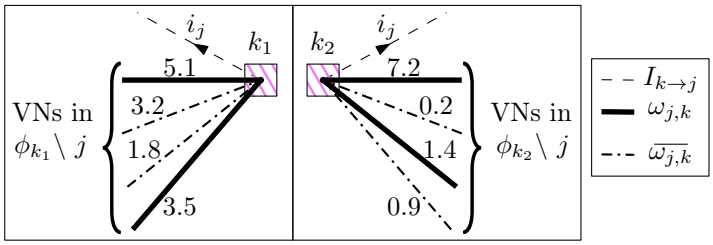


Figure 4.2: Example of edge selection procedure of ESGA in two FNs (k_1 and k_2) with $d_f = 5$ and selecting the $d_s = 2$ strongest neighboring users. The numbers over the neighboring edges illustrate the channel gains. The weakest edges will be treated as Gaussian noise.

Gaussian approximation on the interference is proposed by the authors. The mean $\mu_{j,k} = E\left[z_{j,k}\right]$ and variance $\sigma_{j,k} = \mathrm{Var}\left[z_{j,k}\right]$ of $z_{j,k}$ must be obtained from the superposition of the corresponding random variables in Equation (4.2). Below, we replicate the equations of $\mu_{j,k}$ and $\sigma_{j,k}$ [9]:

$$\begin{aligned}
\mu_{j,k}^{(t)} &= \sum_{i \in \overline{\omega_{j,k}}} h_{i,k} \overline{x_{i,k}}(t) \\
&= \sum_{i \in \overline{\omega_{j,k}}} h_{i,k} \left[\sum_{\mathbf{x}_i \in \mathcal{X}} I_{i \to k}^{(t)}\left(\mathbf{x}_i\right) x_{i,k} \right] && \text{(4.3a)} \\
\sigma_{j,k}^{(t)} &= \sum_{i \in \overline{\omega_{j,k}}} \left|h_{i,k}\right|^2 \mathrm{var}\left(x_{i,k}(t)\right) + \sigma^2 \\
&= \sum_{i \in \overline{\omega_{j,k}}} \left|h_{i,k}\right|^2 \left[\sum_{\mathbf{x}_i \in \mathcal{X}} \left(I_{i \to k}^{(t)}\left(\mathbf{x}_i\right) \left|x_{i,k}\right|^2 \right) - \sum_{\mathbf{x}_i \in \mathcal{X}} \left| I_{i \to k}^{(t)}\left(\mathbf{x}_i\right) x_{i,k} \right|^2 \right] + \sigma^2. && \text{(4.3b)}
\end{aligned}$$

Note that the index t indicates the current MPA iteration. The authors in [31] suggested updating $\mu_{j,k}$ and $\sigma_{j,k}$ at every iteration of the MPA, improving the performance of the GA.

Finally, the MPA is applied to calculate the messages from FN k to VN j using the effective noise z'_k and marginalizing the message calculation of Equation (2.12a) only with respect to the strong neighbors in $\omega_{j,k}$. The modified Equation (2.12a) from [31] is replicated here:

$$I_{k \to j}^{(t)}\left(\mathbf{x}_j\right) = \sum_{\mathbf{x}_{\omega_{j,k}(1)}} \cdots \sum_{\mathbf{x}_{\omega_{j,k}(d_s)}} \left[\exp\left(-\frac{1}{2{\sigma_{j,k}^{(t)}}^2} \left| y_{j,k} - \sum_{i \in \omega_{j,k}} h_{i,k} x_{i,k} - \mu_{j,k}^{(t-1)} \right|^2 + \sum_{i \in \omega_{j,k}} \log I_{i \to k}^{(t-1)}\left(\mathbf{x}_i\right) \right) \right] \tag{4.4}$$

where $y_{j,k} = y_k - h_{j,k} x_{j,k}$.

In Figure 4.2, an illustration of the edge selection procedure of ESGA is shown for two FNs, assuming $d_f = 5$ and $d_s = 2$.

4.3 ESGA-MPA Complexity Order

Since the work in [9, 10] does not assume q-MPA, the complexity order of the ESGA-MPA is based on the modified Equation (2.13), given as

$$\mathcal{O}_{\text{ESGA}} = N_s T_o T_m \left(C_{\text{FN}}^{\text{ESGA}} + C_{\text{VN}} \right), \tag{4.5}$$

where

$$C_{\text{FN}}^{\text{ESGA}} = K d_f M \left[M^{d_s} + \underbrace{d_f - 1 - d_s}_{\text{GA cost}} \right]. \tag{4.6}$$

As it can be seen, $0 \leq d_s \leq d_f - 1$, where $d_s = 0$ means complete GA on the collision and $d_s = d_f - 1$ corresponds to the original MPA.

The ESGA-MPA proposed in [9, 10] has the major advantage of arbitrarily controlling the MUD complexity without giving up on the MPA detector structure, since the MPA is supposed to approximate the maximum *a posteriori* detector.

A main disadvantage in the ESGA-MPA of [9, 10] is that small values of d_s can cause strong degradation in error-rate performance when the channels gains are very similar, i.e. all the neighboring users are "strong enough" to influence the message of FN k to user j (see the example of FN k_1 in Figure 4.2). On the other hand, excessive computation is incurred for large values of d_s when considering the influence of "weak" neighboring channels (relatively small channel gain, as in FN k_2 of Figure 4.2).

Thus, the discussion suggests that the value of d_s could be chosen adaptively, according to the channel gains. Under this motivation, in the following section we propose a modification of the ESGA-MPA.

4.4 Proposed Threshold-Based ESGA-MPA

In order to tackle the disadvantage of the ESGA-MPA regarding the choice of the strong neighbors, we propose a threshold-based (TB) version of the ESGA-MPA, which we denote as TB-ESGA-MPA.

4.4.1 Strong and weak edges definition based on a threshold

The proposed TB-ESGA-MPA can be viewed as the application of the ESGA-MPA using a different definition for the strong and weak neighboring edges of user j in FN k. Moreover, the

concept of the threshold-based selection is applied independently in each FN, which is different from the original works proposed in [9, 10], where a fixed number of edges are cut in each FN.

For ease of notation, we define the channel fading norm $\left|h_{j,k}\right|$ of user j in the resource k as $\sqrt{a_{j,k}}$. Among the neighbors of user j colliding in FN k, i.e. $\forall i \in \phi_{k\setminus j}$, we define the strongest neighbor channel gain as $h^\star = \max_i a_{i,k}$ and its respective user as $i^\star = \arg\max_i a_{i,k}$.

Without loss of generality, assume that there will be at least 2 users colliding in FN k, i.e. $d_f \geq 2$, and that the users $i \in \phi_{k\setminus j}$ are indexed by $1, 2, \ldots, d_f - 1$.

The set $\Phi_{j,k}$, formed by the users with channel gains "as strong as the strongest neighbor" in FN k, is defined as

$$\Phi_{j,k} = \left\{ i : a_{i,k} \geq r h^\star,\ i \in \phi_{k\setminus j},\ i \neq i^\star \right\}, \tag{4.7}$$

for some threshold $rh^\star$, with $0 \leq r \leq 1$.

Finally, we propose the set of the "strongest" neighbors of user j in FN k to be

$$\omega_{j,k} = \Phi_{j,k} \cup \left\{ i^\star \right\} \tag{4.8}$$

and, consequently,

$$\overline{\omega_{j,k}} = \phi_{k\setminus j} - \omega_{j,k}. \tag{4.9}$$

Using these sets, obtained for $k = 1, \ldots, K$ and $j = 1, \ldots, J$, the remaining steps of ESGA-MPA [9, 10] can then be applied to MUD. Similar to the original receiver, once a new realization of the channel gains $h_{j,k}$ is estimated, these sets must be updated.

In summary, the TB-ESGA-MPA can be described in four steps:

i. define a threshold r (a percentage of the value of the strongest neighboring channel gain);

ii. for user j at FN k, first locate the strongest neighboring channel gain $h^\star$;

iii. include all the other neighbors that satisfy the threshold $rh^\star$ in the "strong edges" group;

iv. apply the ESGA-MPA message update rule[1] of Equation (4.4) [9, 10], $\forall \mathbf{x}_j \in \mathcal{X}$, and repeat the procedure for all the remaining (j, k).

Algorithm 1 shows the pseudo-code for the application of TB-ESGA in the FN operations of the MPA during iteration t.

[1] If q-MPA is considered, the application is $\forall x_{j,k} \in \mathcal{A}_q$.

Algorithm 1 Pseudo-code of the application of TB-ESGA in the FN operations in iteration t.

procedure TB-ESGA-FNs(r, C, ϕ_k $\forall k$, φ_j $\forall j$, $\mathbf{h}_j$ $\forall j$)
  for $k = 1, \ldots, K$ **do** ▷ For every FN
    for $j \in \phi_k$ **do** ▷ For every VN connected to FN k
      $h^\star \leftarrow \max_{i \in \phi_{k\setminus j}} |h_{i,k}|^2$ ▷ Select the strongest channel gain colliding with user j in res. k
      $i^\star \leftarrow \arg\max_{i \in \phi_{k\setminus j}} |h_{i,k}|^2$ ▷ Index of the user with channel gain $h^\star$
      $\omega_{j,k} \leftarrow \phi_{k\setminus j}$ and $\overline{\omega_{j,k}} \leftarrow \emptyset$ ▷ Initialize strong and weak users' sets
      for $i \in (\phi_{k\setminus j} - \{i^\star\})$ **do** ▷ For every user colliding with user j in res. k, except $i^\star$
        if $a_{i,k} < r h^\star$ **then** ▷ Check if the user fails the threshold (if it is "*too weak*")
          $\omega_{j,k} \leftarrow \omega_{j,k} - \{i\}$ ▷ Remove user from the set of strong users
          $\overline{\omega_{j,k}} \leftarrow \overline{\omega_{j,k}} \bigcup \{i\}$ ▷ Add user in the set of weak users
        end if
      end for
      $d_s(j,k) \leftarrow |\omega_{j,k}|$ ▷ Effective number of strong users
      Calculate $\mu_{j,k}^{(t)}$ and $\sigma_{j,k}^{(t)}$ using (4.3a) and (4.3b), respectively ▷ Effective noise's mean and var.
      Calculate the messages from FN k to VN j using (4.4) ▷ Complexity order $O\left(q^{d_s(j,k)}\right)$
    end for
  end for
end procedure

Figure 4.3 illustrates the edge selection procedure of TB-ESGA for two FNs with $d_f = 5$, using $r = 0.5$ for the threshold. Compared to the example of Figure 4.2, which has the same FNs and associated neighbors channel gains, it can be seen that the proposed TB-ESGA dynamically evaluates when to consider each of the edges. In FN k_1 of Figure 4.3, due to the application of the threshold, a third edge, that was arbitrarily ignored in ESGA-MPA of Figure 4.2, is now included. However, for FN k_2 of Figure 4.3, the proposed edge-selection avoids excessive computation on relatively small channel gains.

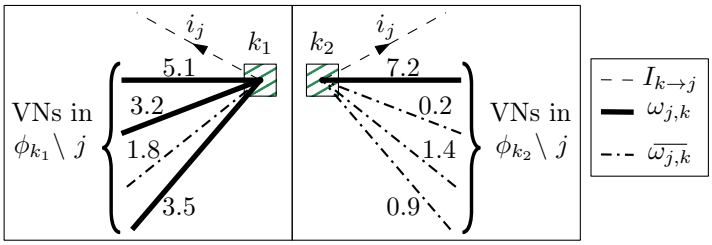


Figure 4.3: Example of edge selection procedure of TB-ESGA in two FNs (k_1 and k_2) with $d_f = 5$ with a threshold $r = 0.5$. In FN k_1, only one edge will be treated as noise. In FN k_2, except for the strongest edge, all the other users will be treated as noise.

Note that, even with the ESGA or TB-ESGA, the output of the MPA can be used to generate soft-bit information and to calculate the log-likelihood ratios of the transmitted bits. It is expected that edge-selection and Gaussian approximation will cause some degradation in the error-rate performance, although they reduce the order of the exponential complexity in the FN

operations. To compensate this degradation, improvements that have linear complexity can be done in the decoder side, such as optimizing the number of inner and outer loops of iterative detection and decoding or finding optimal scaling values for the extrinsic log-likelihood ratios.

4.4.2 Average Complexity Order of the Proposed MUD

Since the channel gains are random, a consequence of the proposed TB-ESGA-MPA is that the composition of $\omega_{j,k}$ also becomes random, as well as its cardinality.

Let $d_s(j,k) \triangleq \left|\omega_{j,k}\right|$, i.e. the cardinality of the set of strong neighbors. Under Rayleigh block fading channel model, we derive the probability of $d_s(j,k)$ being equal to $D+1$, for $0 \leq D \leq d_f - 2$. With this probability, we are also able to derive an expression for the average complexity order of the proposed TB-ESGA-MPA.

For the i.i.d. Rayleigh channel, the channel gains $a_{j,k}$ are distributed according to

$$f(x) = e^{-x} \tag{4.10a}$$

$$F(x) = 1 - e^{-x}, \tag{4.10b}$$

where $f(x)$ is the probability density function (PDF) of $a_{j,k}$ and $F(x)$ is the cumulative distribution function (CDF) of $a_{j,k}$.

Note that $d_s(j,k) \geq 1$, since $i^\star$ will always belong to $\omega_{j,k}$. Hence, the randomness is associated with the set $\Phi_{j,k}$. For some threshold rx, the probability that $\left|\Phi_{j,k}\right|$ is equal to D, for $0 \leq D \leq d_f - 2$, is given by

$$\begin{aligned} &P\left[\left|\Phi_{j,k}\right| = D\right] = \\ &\binom{d_f-2}{D} P\left[a_1, \ldots, a_D \geq rx\right] P\left[a_{D+1}, \ldots, a_{d_f-1} < rx\right] &\text{(4.11a)} \\ &= \binom{d_f-2}{D} \left[1 - T(rx)\right]^D \left[T(rx)\right]^{d_f-2-D} &\text{(4.11b)} \end{aligned}$$

where $T(rx) = F(rx)/F(x)$.

Let $p_\Phi(d_s(j,k))$ be the PMF of the integer random variable $d_s(j,k)$, such that $p_\Phi(D+1) = P\left[d_s(j,k) = D+1\right] = P\left[\left|\Phi_{j,k}\right| = D\right]$.

Given r and x, we have

$$\overline{d_s}(r,x) = E_\Phi\left[d_s(j,k)\right] = \sum_{d=1}^{d_f-1} d p_\Phi(d) \tag{4.12a}$$

$$= \sum_{d=0}^{d_f-2} (d+1) \binom{d_f-2}{d} [1-T(rx)]^d [T(rx)]^{d_f-d-2}, \tag{4.12b}$$

where we denote $\overline{d_s}(r,x)$ as the expected size of the strong neighbors sets for a given threshold rx.

It is also important to evaluate the expectation over the exponential term of the ESGA complexity order, which is a function of $d_s(j,k)$:

$$\overline{q^{d_s(j,k)}}(r,x) = E_\Phi \left[q^{d_s(j,k)} \right] = \sum_{d=1}^{d_f-1} q^d p_\Phi(d) \tag{4.13a}$$

$$= \sum_{d=0}^{d_f-2} \left(q^{d+1} \right) \binom{d_f-2}{d} [1-T(rx)]^d [T(rx)]^{d_f-d-2}, \tag{4.13b}$$

Thus, similar to Equation (4.5), we can write the complexity order of the proposed method, for fixed x and r, as:

$$O_{\mathrm{TB}}(r,x) = N_s T_o T_m \left[C_{\mathrm{FN}}^{\mathrm{TB}}(r,x) + C_{\mathrm{VN}} \right], \tag{4.14}$$

where

$$C_{\mathrm{FN}}^{\mathrm{TB}}(r,x) = K d_f q \left[\overline{q^{d_s(j,k)}}(r,x) + d_f - 1 - \overline{d_s}(r,x) \right] \tag{4.15}$$

relates to the FN complexity order.

We can now estimate the average complexity of the TB-ESGA-MPA, as a function of the parameter r, by taking the expectation of $O_{\mathrm{TB}}\left(\overline{d_s}(r,x)\right)$ over x:

$$\begin{aligned} \overline{O_{\mathrm{TB}}}(r) &= E_x \left[O_{\mathrm{TB}}(r,x) \right] \\ &= N_s T_o T_m \left[\overline{C_{\mathrm{FN}}}(r) + C_{\mathrm{VN}} \right], \end{aligned} \tag{4.16}$$

where

$$\begin{aligned} \overline{C_{\mathrm{FN}}}(r) &= E_x \left[C_{\mathrm{FN}}(r,x) \right] \\ &= K d_f q \left[\overline{q^{d_s}}(r) + d_f - 1 - \overline{d_s}(r) \right] \end{aligned} \tag{4.17}$$

and

$$\overline{q^{d_s}}(r) = E_x \left[\overline{q^{d_s(j,k)}}(r,x) \right]$$

$$= \int_0^\infty \overline{q^{d_s(j,k)}}(r,x)\, f_x(x)\, dx, \tag{4.18}$$

$$\overline{d}_s(r) = E_x\left[\overline{d}_s(r,x)\right]$$

$$= \int_0^\infty \overline{d}_s(r,x)\, f_x(x)\, dx. \tag{4.19}$$

Closed-form expressions for $\overline{q^{d_s}}(r)$ and $\overline{d}_s(r)$ are difficult to obtain, but the integrals in (4.18) and (4.19) can be evaluated numerically, regardless of the distribution of x. If $x = h^\star$, then x is a first order statistics [81], and its PDF and CDF are given as

$$f_x(x) = (d_f - 1)\,[F(x)]^{d_f-2}\, f(x) \tag{4.20a}$$

$$F_{(d_f-1)}(x) = P\left[h^\star \le \sqrt{x}\right]$$

$$= [F(x)]^{d_f-1} \tag{4.20b}$$

More specifically, for the i.i.d. Rayleigh channel:

$$f_x(x) = (d_f - 1)e^{(1-d_f)x}\,(e^x - 1)^{d_f-2}. \tag{4.21}$$

4.4.3 TB-ESGA Design Guidelines

In this section, we provide some guidelines for the system design of TB-ESGA. We suggest the following Algorithm 2 as an empirical method for choosing the value of the threshold parameter r. To the best of our knowledge, there is no analytical expression for the frame- or bit-error rate of SCMA that considers the effects of the MPA receiver. Thus, so far, Monte Carlo simulations are typically necessary to evaluate the final performance of an SCMA system.

Algorithm 2 Iterative procedure to choose the threshold of TB-ESGA.

procedure THRESHOLD SELECTION(E_b/σ^2, FER$^\circ$,δ_r)
  $r \leftarrow 1 + \delta_r$
  while FER > FER$^\circ$ **do**
    $r \leftarrow r - \delta_r$
    Run a FER link level simulation of TB-ESGA at the operation point of E_b/σ^2
    if r is zero **and** FER > FER$^\circ$ **then**
      Declare failure. E_b/σ^2 possibly too low.
    end if
  end while
  return r
end procedure

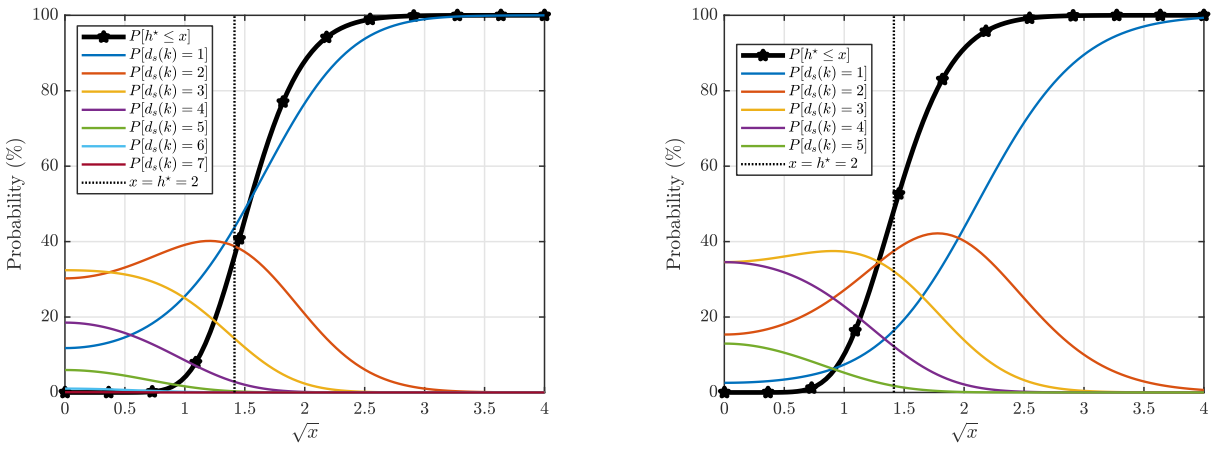
The algorithm starts with $r = 1$, which has the lowest complexity. Then, for a desired operation point of E_b/σ^2, the frame error rate (FER) is obtained with a Monte Carlo simulation of the system model. If the current FER is lower or equal to the target performance value FER°, the algorithm stops and returns the value of r. Otherwise, it decrements r by a value of δ_r and then repeats the procedure. If the algorithm reaches $r = 0$ (original MPA) and the FER did not reach the target value, the algorithm declares failure. The failure could mean that the E_b/σ^2 is not sufficiently large to reach the desired FER. In this case, one can increase the operation point E_b/σ^2 or consider redesigning other system parameters, such as R, C (M, N) and β.

4.5 Numerical Results

4.5.1 Edge Selection Probability

In Figure 4.4a, we show the results of the probability in Equation (4.11b) for a FN having collision degree $d_f = 8$, using a threshold of $r = 70\%$. We can notice that $P[h^\star \leq 2] \approx 37\%$ and, for $x = h^\star = 2$, there is a chance of 44.1% that $d_s(j,k) = 1$, 39.1% that $d_s(j,k) = 2$ and 14.6% that $d_s(j,k) = 3$. That is, the probability of a high degree rapidly decreases. Note that the probability of an effective collision degree of $d_s(j,k) + 1 = 6$ is less than 0.5%. Thus, we have an opportunity to greatly reduce ESGA's complexity, as we quantify in the next section.

Figure 4.4b shows as a second example the probability in (4.11b) for a FN with collision degree $d_f = 6$, using a threshold of $r = 40\%$. The reduced pool size $d_f - 2$ also influences the CDF of $h^\star$. The relaxed threshold increases the chances of selecting edges and, for each value of D, it makes the probability curves $P[d_s(j,k) = D]$ "wider" than those in Figure 4.4a. For the same value of $x = 2$, we now have $P[d_s(j,k) = 2] \approx 38\% > P[d_s(j,k) = D]$ for all $D \neq 2$.



(a) $d_f = 8$, using a threshold of $r = 70\%$ of $h^\star$.

(b) $d_f = 6$, using a threshold of $r = 40\%$ of $h^\star$.

Figure 4.4: Results of Equation (4.11b) for $D = 0, \ldots, d_f - 2$.

4.5.2 Average Complexity Order of Proposed TB-ESGA-MPA

In Figure 4.5, we show the results of numerical evaluation of Equation (4.16) for Scenarios in Table 4.1. The predicted complexity order (solid lines) matches the one obtained with Monte Carlo simulations (markers). Recall that $r = 0$ and $r = 1$ correspond to the original MPA and to full GA (except for $i^\star$, meaning $d_s = 1$), respectively, yielding deterministic complexities.

Table 4.1: Simulation scenarios with fixed parameters.

Scenario ID	(J, K)	(M, q, N)	d_f	β	N_s	SE (bits/res)
(A)	$(20, 10)$	$(4, 4, 2)$	4	200%	64	2.0
(B)	$(20, 12)$	$(8, 3, 3)$	5	166%	43	2.5
(C)	$(18, 12)$	$(16, 4, 4)$	6	150%	32	3.0

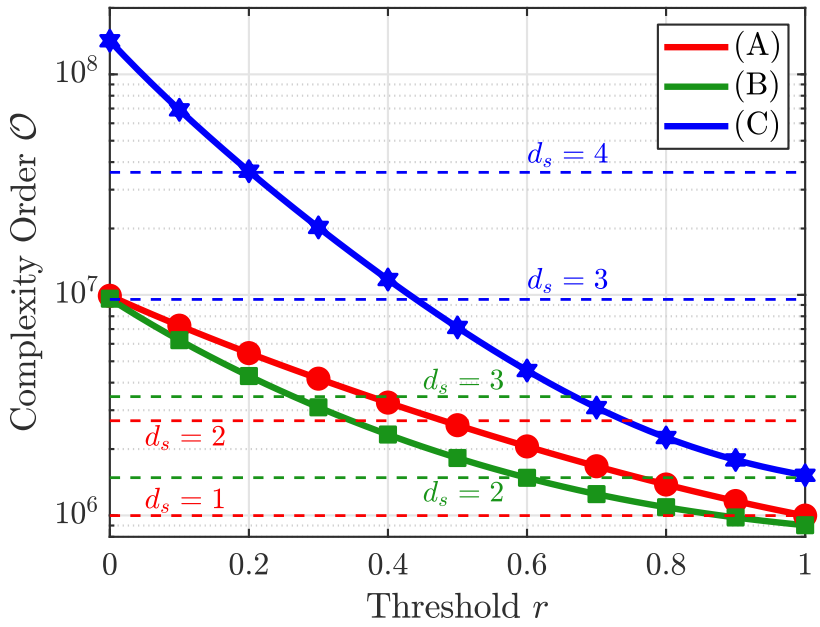


Figure 4.5: Curves of $\overline{O_{\text{TB}}}(r)$ from Equation (4.16) for Scenarios in Table 4.1. Markers correspond to complexities evaluated by Monte Carlo simulation. Dashed lines correspond to O_{ESGA} (considering LNCP) of ESGA-MPA with fixed d_s, as shown in the annotations, and $r = 0$ corresponds to the complexity order of the original MPA ($d_s = d_f - 1$).

We also added the complexity of ESGA for a few choices of $d_s < d_f - 1$. From the complexity order prediction above, the value of r can be chosen to reduce the average complexity with respect to ESGA. However, predicting the complexity order gives no guarantee of good error-rate performance, neither for ESGA nor for TB-ESGA. Thus, for each system and for each choice of parameters, it is important to evaluate the system error-rate performance.

4.5.3 Choosing the value of r for each scenario

Figure 4.6 shows the resulting FER curves (solid lines with markers) of simulation scenarios A, B and C for the full range of r. We considered $\delta_r = 0.1$ and an operation point of $E_b/\sigma^2 = 22$ dB. For design purposes, we also plot again the average complexity order curves of TB-ESGA (dashed lines).

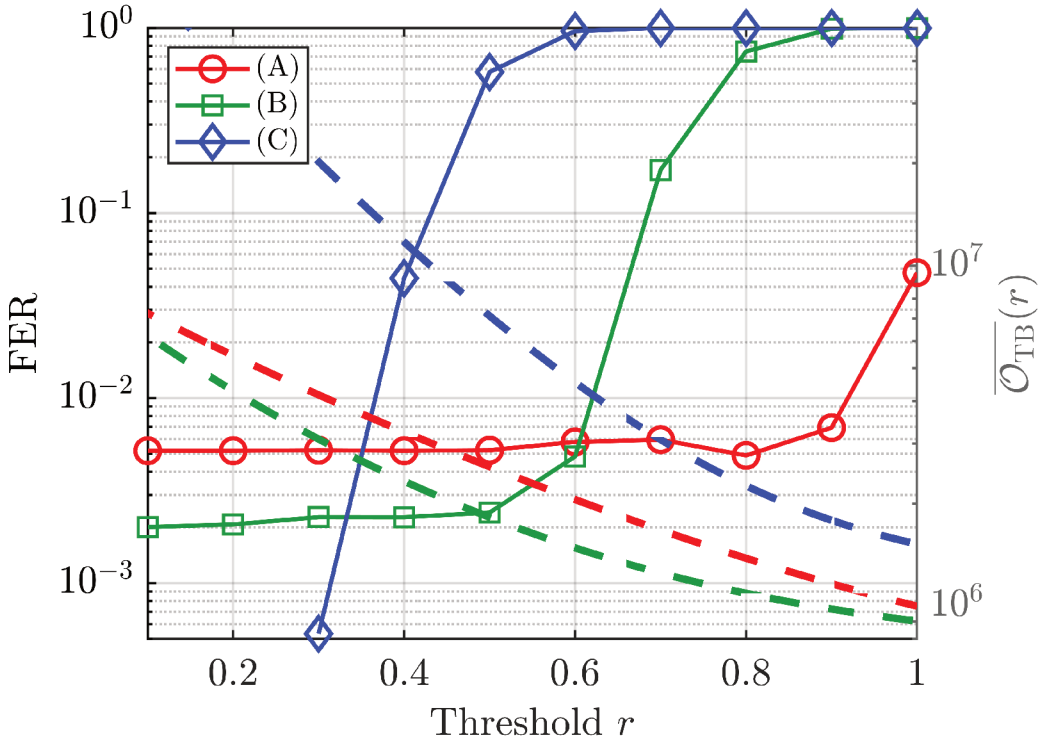


Figure 4.6: Resulting FER of each simulation Scenario according to the parameter r. E_b/σ^2 is fixed at 22 dB.

It can be noticed, in Figure 4.6, that some good trade-offs can be found for r. In each scenario, there is a plateau transition in the FER curves, where one can adjust r to achieve good error-rate performance at the lowest possible complexity. Reducing r too much seems to be inefficient after this transition, since the FER does not decrease significantly anymore. This differential value of FER could also be included as a stopping criteria in Algorithm 2, instead of using FER°.

For scenario A, the choice of $r = 0.8$ is exactly at the point of lower complexity where the FER curve is still at its minimum value. In scenario B, choosing $r = 0.5$ would be the optimal choice with respect to the error-rate performance, but we will use $r = 0.6$ to save some computation time since the FER is about the same value reached with $r = 0.8$ in scenario A. In scenario C, the complexity order of the scheme is relatively larger than the previous ones, thus we choose $r = 0.4$ in order to have the FER as small as possible without increasing the

complexity order of the receiver more than $O\left(10^7\right)$.

4.5.4 Error Rate Performance and Complexity Reduction

We have simulated link level SCMA under different scenarios (A, B and C), according to the parameters in Table 4.1. In all simulations, the FEC is a $(n,k,R) = \left(128,64,\frac{1}{2}\right)$ short block length LDPC, available in [82], using $T_f = 30$ belief propagation iterations. The BICM-ID SCMA iterations are fixed as $(T_o,T_m) = (3,5)$. LNCP codebooks were designed using the methods from [7, 73]. We provide some figures of merit of these codebooks, shown in Table 4.2.

Table 4.2: Designed SCMA codebooks and its parameters.

Design Method	**Codebook** (M,q,N)	MSSD	$d^2_{E,\min}$	$d_p^{(\min)}$
A_{44} from [7]	$(4,4,2)$	2	2	1.0
GRS based on [73]	$(8,3,3)$	2	2	1.0
GRS from [73]	$(16,4,4)$	3	2	0.5

In Table 4.3, we provide the complexity order $\overline{O}$ of each MUD in each scenario. The values of O_{unc} and O_{cod} is the complexity order of the original MPA for uncoded and coded SCMA, respectively, calculated according to Equation (2.14) using $(T_o,T_m) = (1,5)$ $(T_o,T_m) = (3,5)$, respectively.

Figures 4.7, 4.8 and 4.9 show the evaluation of the FER of scenarios (A), (B) and (C), respectively. Each frame has $64 \times J$ information bits (note that the J encoders are independent of each other).

In scenario (A), we observe in Fig. 4.7 that the FER difference between ESGA-MPA with $d_s(j,k) = 2\ \forall j,k$ and the proposed TB-ESGA-MPA is almost negligible. However, the average complexity of TB-ESGA-MPA is about half of the former, as shown in Table 4.1. If we reduce the complexity of ESGA using $d_s(j,k) = 1\ \forall j,k$, its complexity is 10% less with respect to the proposed TB-ESGA-MPA, however the consequence is a strong FER degradation, being the worse in performance.

Table 4.3: Average complexity order for the different scenarios, comparing the ESGA-MPA from [9, 10] with the proposed TB-ESGA-MPA.

ID	MUD	$\log_{10}\left(\bar{O}\right)$	$\frac{\bar{O}}{O_{\text{unc}}}$ (%)	$\frac{\bar{O}}{O_{\text{cod}}}$ (%)
(A)	ESGA $d_s(j,k) = 2\ \forall j,k$	6.43	81.40	27.13
	ESGA $d_s(j,k) = 1\ \forall j,k$	6.00	30.23	10.07
	TB-ESGA $d_s(j,k) \in [1,3]$, $r = 0.8$	6.14	41.82	13.92
(B)	ESGA $d_s(j,k) = 3\ \forall j,k$	6.53	107.92	35.97
	ESGA $d_s(j,k) = 2\ \forall j,k$	6.17	46.31	15.43
	TB-ESGA $d_s(j,k) \in [1,4]$, $r = 0.6$	6.17	46.24	15.45
(C)	ESGA $d_s(j,k) = 4\ \forall j,k$	7.56	75.95	25.31
	ESGA $d_s(j,k) = 3\ \forall j,k$	6.97	20.16	6.71
	TB-ESGA $d_s(j,k) \in [1,5]$, $r = 0.4$	7.07	24.79	8.27

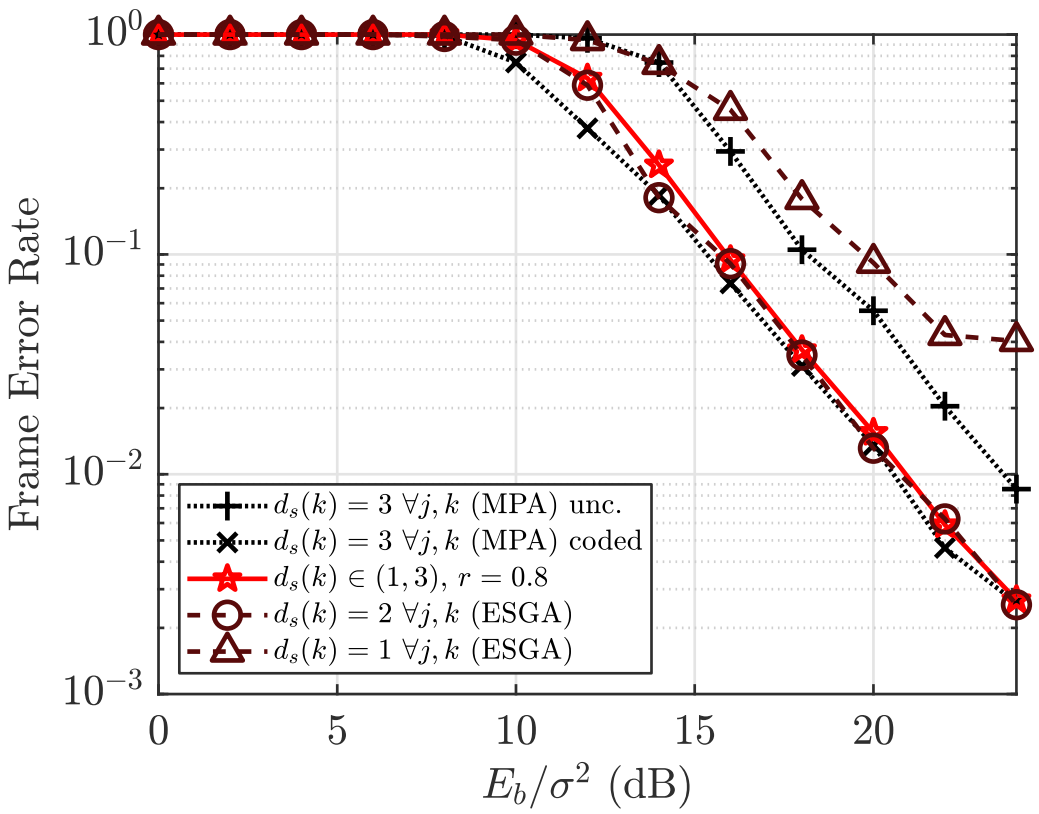


Figure 4.7: Frame error-rate of simulation scenario (A).

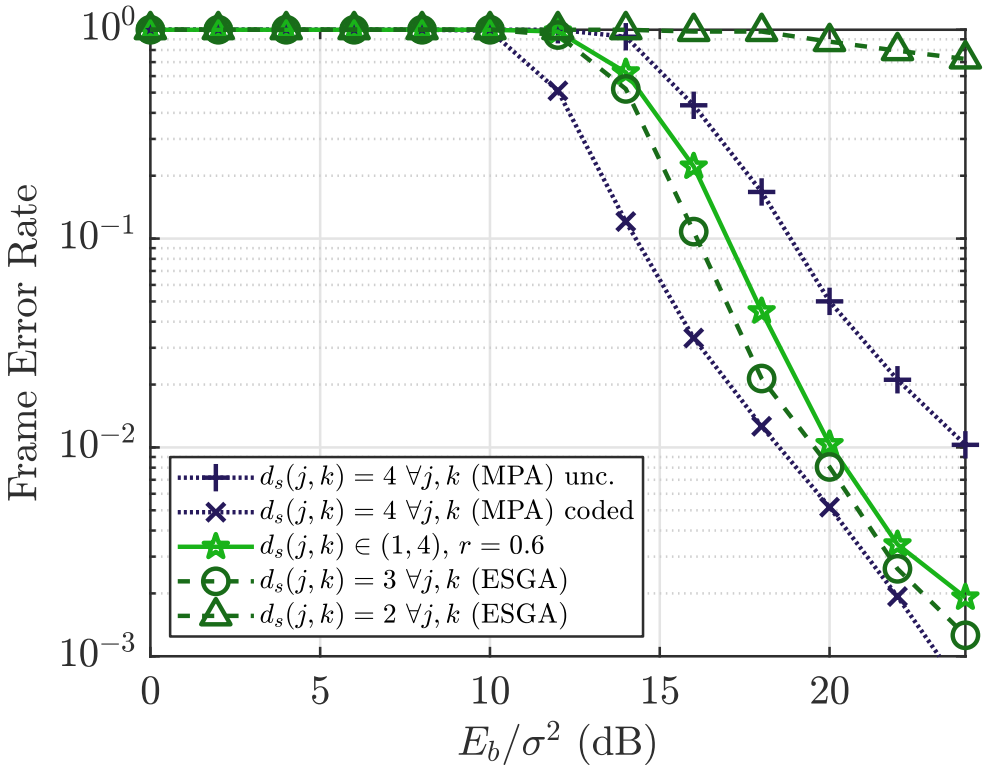


Figure 4.8: Frame error-rate of simulation scenario (B).

In scenario (B), we see in Fig. 4.8 that ESGA-MPA with $d_s\,(j,k) = 3\ \forall j,k$ has a better FER performance than TB-ESGA-MPA, with a difference of approximated 2 dB at 10^{-1}. However, the complexity order of TB-ESGA-MPA from Table 4.1 shows a reduction of about 60% with respect to fixed $d_s\,(j,k) = 3$. Reducing to $d_s\,(j,k) = 2\ \forall j,k$ in ESGA-MPA achieves about the same complexity order of TB-ESGA-MPA, but once again at the cost of severe performance degradation. The degradation of FER in ESGA-MPA with $d_s\,(j,k) = 2\ \forall j,k$ makes this strategy inadequate for application in this scenario.

In scenario (C), the LNCP codebook size is larger ($M = 16$) and the LNCP alphabet has size $q = 4$. Because of the larger number of dimensions of this codebook ($N = 4$), the overload caused an increased number of collisions in each FN ($d_f = 6$). The FER of ESGA-MPA in Figure 4.9 shows that this strategy does not work in this scenario. On the other hand, TB-ESGA-MPA with $r = 0.4$ is able to cope with this LNCP codebook and with the large number of collisions, having an acceptable FER performance with relatively low complexity in comparison with O_{ref}.

In summary, we observed in simulations that the threshold r must be tuned for each scenario. Increasing M or d_f (which is proportional to N and J/K) will cause significant performance degradation if r is not properly reduced. Hence, a trade-off between complexity, performance and spectral efficiency must be considered in the system design.

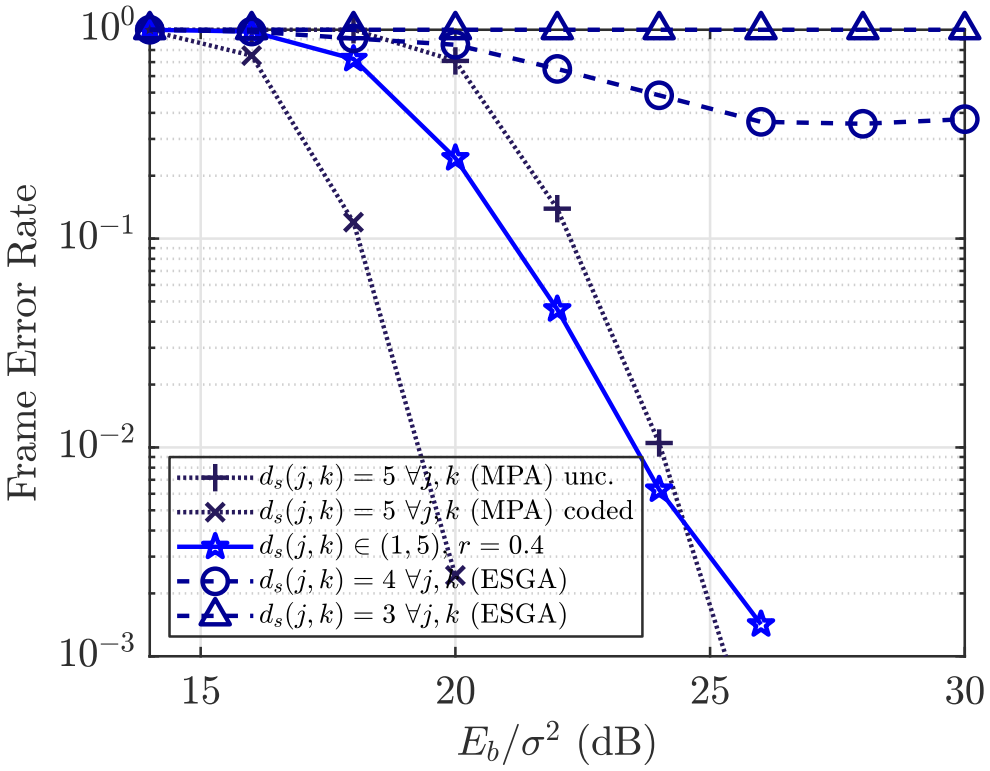


Figure 4.9: Frame error-rate of simulation scenario (C).

• • •

5

Resource Allocation Methods for Wide Area Network SCMA

In this chapter, a novel approach for the resource allocation design of SCMA is discussed. We consider now a system named Wide Area Network (WAN) SCMA (WAN-SCMA), which takes into account the large-scale path-losses of each user selected for the uplink.

It presents an analysis of some relevant figures of merit, which depend on the design of $\mathbf{F}$. Given these metrics, we evaluate the benefits of designing $\mathbf{F}$ with irregular spreading degrees. We propose design algorithms to find suboptimal solutions that can enhance the discussed metrics when comparing to regular designs. Related works have been discussed in Section 2.4.

5.1 Revisiting the System Model

For the following discussion, it is important to review some assumptions on the SCMA system model. In this chapter, we will consider that the J users are located in a WAN, each one in a different position inside the cell, at a certain distance D_j [km] to the base-station. Therefore, large-scale path-loss will be taken into account.

The large-scale path-loss coefficients are denoted α_j, and its model is typically related to the distances D_j (such that $\alpha_j \propto D_j$) and to other geographical aspects of the electromagnetic environment. The specific model for this path-loss will be defined later, for simulation purposes.

Regarding power allocation, the following assumptions are considered:

i. each user transmits with total power P_j,

ii. all users transmit with the same power, $P_j = P \; \forall j$,

iii. the power allocated in each resource is $p_{k,j}$, such that $P_j = \sum_{k \in \varphi_j} p_{k,j} = P$,

iv. uniform power allocation over the assigned resources, $p_{k,j} = p_j = P/N_j \ \forall(k, j)$,

v. the users' indexes are defined such that $\alpha_1 < \ldots < \alpha_J$ (or $D_1 < \ldots < D_J$).

vi. the codebooks C_j still have normalized power, but it will not be represented with P anymore, i.e. $E\left[\|\mathbf{c}_j\|^2\right] = 1 \neq P$.

Considering the path-loss in the cell, we redefine the effective SNR of user j as:

$$\gamma_j = \alpha_j^{-1}\text{SNR} = \frac{\alpha_j^{-1}P}{\sigma^2} \tag{5.1}$$

where σ^2 is the noise power in each resource, modeled as in (2.1). The bit energy per noise power of user j becomes $E_{b,j}/\sigma^2 = \gamma_j/k_j$ (per channel use). Note that assumption v. and Equation (5.1) imply that $\gamma_1 > \ldots > \gamma_J$.

Figure 5.1 shows an illustration of the system model of WAN-SCMA.

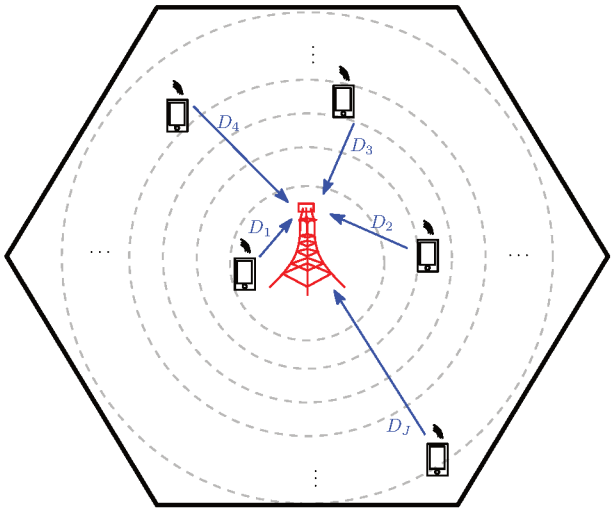


Figure 5.1: WAN-SCMA: illustration of uplink users located at different distances (radius) from the base-station.

Thus, we modify equation 2.1 to rewrite the received vector of uplink SCMA, which is now given as

$$\mathbf{y} = \sum_{j=1}^{J} P^{1/2}\alpha_j^{-1/2}\,\text{diag}\left(\mathbf{h}_j\right)\mathbf{x}_j + \mathbf{z} \tag{5.2a}$$

$$= \sum_{j=1}^{J} \mathbf{H}_j\mathbf{c}_j + \mathbf{z} \tag{5.2b}$$

where $\mathbf{H}_j = P^{1/2}\alpha_j^{-1/2}\,\text{diag}\left(\mathbf{h}_j\right)\mathbf{V}_j$ is the sparse effective channel of user j.

From Section 2.1, recall the definitions of the sets of VNs colliding in FN k, ϕ_k, and of FNs associated with user j, φ_j. In this chapter, we will consider the possibility that the spreading degrees N_j and the collision degrees d_k can be irregular. In other words, the Hamming weights of the rows and columns of $\mathbf{F}$ can be different.

Example 3

Consider the following irregular design for an SCMA system with $J = 6$ users sharing $K = 4$ resources:

$$\mathbf{F} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \tag{5.3}$$

In this design, the spreading degrees are $N_1 = N_4 = N_5 = 1$, $N_2 = N_3 = 4$, and $N_6 = 2$. The collision degrees are also slightly irregular: $d_1 = 4$, while $d_2 = d_3 = d_4 = 3$.

5.2 Figures of Merit

In this section, we review some figures of merit considered for the SCMA system.

From a practical point of view, we discuss the receiver complexity when employing the MPA MUD. On the other hand, resource allocation usually takes into consideration figures of merit from the channel rate region, in order to assign codebooks and resources that can enhance the system performance.

We briefly discuss the metrics associated with both of these issues.

5.2.1 MPA Complexity Order for Irregular SCMA

Section 2.5 and Chapter 4 provided some insight on the complexity order of the MPA for the case of regular resource allocation, i.e. regular designs of $\mathbf{F}$ ($d_k = d_f \ \forall k$ and $N_j = N \ \forall j$).

In summary, it was pointed out that the complexity order of the MAP MUD for SCMA is proportional to M^J, which is intractable as J increases, even for small M, while the FN complexity of the MPA is proportional to M^{d_f} [3], which is much lower than the optimum MAP when $d_f \ll J$.

In general, the collision degrees d_k could be irregular, designing $\mathbf{F}$ with arbitrary size and degrees to optimize some figure of merit. Let $\overline{d_f} = \frac{1}{K}\sum_{k=1}^{K} d_k$ and $\overline{N} = \frac{1}{J}\sum_{j=1}^{J} N_j$ be the average

collision degree and the average spreading degree, respectively. Note that N_j and d_k are integer values $\forall j, k$.

The relationship of spreading degrees N_j and collision degrees d_k is tied up to the desired load factor of SCMA, and is given as $K\overline{d_f} = J\overline{N}$. Recall that we are only interested in overloaded scenarios, meaning that the load factor is always $\beta = J/K > 1$.

Hence, a regular matrix forces $d_f > N$. For irregular row weights only, $\overline{d_f} > N$, while irregular column weights only forces $d_f > \overline{N}$. In general, we can say that $\overline{d_f} = \beta \times \overline{N} > \overline{N}$.

Assuming the use of codebooks designed for q-MPA (see Section 2.5.4) and imposing a limitation $q_j \leq q \; \forall j$, we can say that, for T_m iterations of the MPA, the overall MUD complexity order is

$$O_{\text{MPA}} \leq T_m \times \left(\sum_{k=1}^{K} d_k q^{d_k} + MJ \left(\overline{N} - 1 \right) \right). \tag{5.4}$$

A very important aspect to be considered is the dependence of (5.4) on the average dimension $\overline{N}$ of the codebooks. Assuming the case of $d_k = d_f \; \forall k$ ($\mathbf{F}$ with regular row weight only), we can rewrite (5.4) as

$$O_{\text{MPA}} \leq T_m \times \left(J\overline{N} q^{\beta \times \overline{N}} + MJ \left(\overline{N} - 1 \right) \right) \tag{5.5}$$

and observe that the upper bound on O_{MPA} is exponential with $\overline{N}$ and with the load factor β.

For the cases where $\mathbf{F}$ is irregular in both N_j and d_k, we can argue that the complexity will be highly determined by both $\overline{N}$ and $\max_k d_k$. Controlling average d_k is achieved by fixing β and imposing a limitation on $\overline{N}$.

In order to control $\max_k d_k$ for a given distribution of N_j, we should design $\mathbf{F}$ similarly to a LDPC matrix with concentrated check-node degrees [83, 84], such that $d_k \approx \overline{d_f}$. This goal imposes a trade-off constraint in the subcarrier assignment step, which in practice limits the variance of the collision degree and suggests designs of $\mathbf{F}$ oriented to regular row weights.

5.2.2 Information Theory Metrics

NOMA Rate Region: MU-MIMO Analogous Model

The system model of SCMA is analogous to a Multiple-Input Multiple-Output (antennas) (MIMO) multiple-access channel. From [85], adapting notation, limits to the achievable rates of the MU MIMO channel are given[1] as:

[1]This capacity equation relies on the assumption that $\mathbf{x}_j \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{Q}_j\right)$ to maximize the entropy $h(\mathbf{y})$. Hence, $\mathbf{y} \sim \mathcal{CN}\left(\mathbf{0}, \sigma^2 \mathbf{I}_K + \mathbf{H}_j \mathbf{Q}_j \mathbf{H}_j^H\right)$.

$$R_j \leq \mathcal{I}_j = \log_2 \det \left[\mathbf{I}_K + \frac{1}{\sigma^2} \mathbf{H}_j \mathbf{Q}_j \mathbf{H}_j^H \right] \tag{5.6}$$

where $\mathbf{Q}_j = E\left[\mathbf{c}_j \mathbf{c}_j^H\right] = \frac{1}{M_j} \sum_{m=1}^{M_j} \mathbf{c}_j^{(m)} \mathbf{c}_j^{(m)H}$ is the covariance matrix of the codebook C_j.

The sum-capacity is similarly given as

$$\sum_{j=1}^{J} R_j \leq \mathcal{I}_{\text{sum}} = \log_2 \det \left[\mathbf{I}_K + \sum_{j=1}^{J} \frac{1}{\sigma^2} \mathbf{H}_j \mathbf{Q}_j \mathbf{H}_j^H \right]. \tag{5.7}$$

Finally, for a generalized subset of users, $\mathcal{J} \subseteq \{1, \ldots, J\}$ we have the rate bounds defined by the following equation:

$$\boxed{\sum_{j \in \mathcal{J}} R_j \leq \mathcal{I}_{\mathcal{J}} = \log_2 \det \left[\mathbf{I}_K + \sum_{j \in \mathcal{J}} \frac{1}{\sigma^2} \mathbf{H}_j \mathbf{Q}_j \mathbf{H}_j^H \right].} \tag{5.8}$$

Note that the rate region depends on $\mathbf{F}$, since the channel matrices $\mathbf{H}_j = P^{1/2} \alpha_j^{-1/2} \operatorname{diag}\left(\mathbf{h}_j\right) \mathbf{V}_j$ are the effective sparse channels of each user (see Eq. (2.2)). On the other hand, the covariance matrices $\mathbf{Q}_j$ depend only on the codebook design of the MDM $C_j\left(M_j, N_j\right)$.

Symmetric Capacity

Considering the choice of rates $R_1 = R_2 = \ldots = R_J = R$, named *symmetric rates*, the Equation (5.8) becomes

$$\sum_{j \in \mathcal{J}} R \leq \mathcal{I}_{\mathcal{J}} \tag{5.9a}$$

$$|\mathcal{J}|\, R \leq \mathcal{I}_{\mathcal{J}} \tag{5.9b}$$

$$R \leq \frac{\mathcal{I}_{\mathcal{J}}}{|\mathcal{J}|} \tag{5.9c}$$

Definition: Symmetric Capacity

The symmetric capacity is defined as

$$C_{\text{sym}} = \min_{\mathcal{J}} \frac{\mathcal{I}_{\mathcal{J}}}{|\mathcal{J}|}. \tag{5.10}$$

An achievable symmetric rate must be $R \leq C_{\text{sym}}$. The bottleneck of the symmetric capacity

depends on the set $\mathcal{J}_{\text{sym}} = \arg\min_{\mathcal{J}} \frac{\mathcal{I}_{\mathcal{J}}}{|\mathcal{J}|}$.

5.3 Optimization Problem for the Resource Allocation

The interest in the symmetric capacity is to simplify the design of SCMA, assuming that all users transmit at the same rate, with codebook sizes $M_1 = \ldots = M_J = M$. Maximizing this bottleneck would be the most fair approach in order to allow that all the J users can jointly transmit at its highest rate with error probability arbitrarily as low as possible.

On the other hand, the joint-detection of the MPA has exponential complexity and it might impose some constraints for the design of the SCMA graph, depending on the overload and spreading degrees.

Hence, given the small-scale fading coefficients $\mathbf{h}_j$ and the effective SNR γ_j of each users, we wish to design the resource allocation matrix $\mathbf{F}$ that maximizes the symmetric capacity, assuming some design constraints. The optimization problem is written below:

Symmetric Capacity Maximization

$$\max_{\mathbf{F}} \min_{\mathcal{J}} \quad \frac{\mathcal{I}_{\mathcal{J}}}{|\mathcal{J}|} \tag{5.11a}$$

$$\text{subject to} \quad N_{\min} \leq N_j \leq N_{\max} \; \forall j, \tag{5.11b}$$

$$\frac{1}{J}\sum_{j=1}^{J} N_j \leq \overline{N}_{\max}. \tag{5.11c}$$

Recall that N_j is associated with $\mathbf{F}$, since $N_j = \mathbf{f}_j^T \mathbf{f}_j$ is the spreading degree of user j (Hamming weight of column j).

Constraint (5.11b) limits the maximum and minimum number of resources that can be assigned for each user. The sum-constraint (5.11c) imposes a constraint on the maximum allowed average VN degree, meaning that we cannot increase all N_j boundlessly, which consequently limits the maximum average FN degree for a given load factor β. Jointly, these $J + 1$ constraints take into account the complexity issues of the figure of merit from Section 5.2.1.

With respect to the minimization in (5.11a), there are $2^J - 1$ sets $\mathcal{J}$ involved in the evaluation of the objective function itself. Even for a known $\mathbf{F}$, the complexity order of evaluating all sets to obtain (5.10) is $O\left(2^J\right)$, and it becomes computationally draining for large J. The drawbacks of this exhaustive search hinders dynamic assessment of C_{sym} and even its Monte Carlo simulations (ergodic values for some channel model, e.g. Rayleigh).

At the moment, we do not have a solution to this problem. Nevertheless, for evaluation purposes of all the constraints of the rate-region, it would be useful to find the set $\mathcal{J}$ that is the bottleneck without exhaustive search.

5.4 A Lower-Bound for the Symmetric Capacity

5.4.1 Assumptions on the Codebook Design

Note that, in the effective channel $\mathbf{H}_j$, the result of diag $(\mathbf{h}_j)\,\mathbf{V}_j$ is a diagonal matrix, since the k-th row of $\mathbf{V}_j$ is multiplied by the corresponding fading $h_{j,k}$, for $k = 1, \ldots, K$.

If we can rewrite the argument of the sum-capacities as diagonal matrices, it will simplify the capacity constraints.

This simplification is possible if we assume $\mathbf{Q}_j$ is diagonal[2], i.e. $\mathbf{Q}_j = \text{diag}\,(\mathbf{q}_j)$. Then,

$$\mathbf{H}_j\mathbf{Q}_j\mathbf{H}_j^H = \frac{1}{\sigma^2}\left(P^{1/2}\alpha_j^{-1/2}\,\text{diag}\,(\mathbf{h}_j)\,\mathbf{V}_j\right)\mathbf{Q}_j\left(P^{1/2}\alpha_j^{-1/2}\,\text{diag}\,(\mathbf{h}_j)\,\mathbf{V}_j\right)^H \tag{5.12a}$$

$$= \gamma_j\,\text{diag}\,(\tilde{\mathbf{q}}_j)\,\text{diag}\left(\mathbf{h}_j\mathbf{h}_j^H\right) \tag{5.12b}$$

where $\tilde{\mathbf{q}}_j = \mathbf{V}_j\mathbf{q}_j$ is a $K \times 1$ vector, corresponding to the sparse spread version of the $N \times 1$ vector $\mathbf{q}_j$. The steps of the derivation are straightforward and follow from the property that the product of two diagonal matrices is commutative.

Using (5.12b) in (5.9c):

$$R \leq \frac{1}{|\mathcal{J}|}\log_2\det\left[\mathbf{I}_K + \sum_{j\in\mathcal{J}}\gamma_j\,\text{diag}\,(\tilde{\mathbf{q}}_j)\,\text{diag}\left(\mathbf{h}_j\mathbf{h}_j^H\right)\right]. \tag{5.13}$$

5.4.2 The Equivalent K-Parallel MAC Channels

Let $\mathbf{M}_{\mathcal{J}} = \mathbf{I}_K + \sum_{j\in\mathcal{J}}\gamma_j\,\text{diag}\,(\tilde{\mathbf{q}}_j)\,\text{diag}\left(\mathbf{h}_j\mathbf{h}_j^H\right)$ be the notation for the resulting matrix in the argument of the determinant of (5.13). Note that $\mathbf{M}_{\mathcal{J}}$ is a diagonal matrix, whose k-th diagonal entry is given by

$$M_{\mathcal{J},(k,k)} = 1 + \sum_{j\in\mathcal{J}}\gamma_j\tilde{q}_{j,k}\left\|h_{j,k}\right\|^2.$$

Recall that resource k only contains users in ϕ_k due to the graph structure of $\mathbf{F}$. Thus, the summation may include this limitation to avoid the zero-valued entries of $\tilde{q}_{j,k}$.

[2]This constraint should be addressed in codebook design. It can be added as a constraint in the design steps of Chapter 3.

Assume[2] also that $q_{j,n} = \dfrac{1}{N_j}\ \forall n = 1,\ldots,N$. The sparse version $\tilde{\mathbf{q}}_j$ will also contain the same entries in its non-null positions. Thus, we can rewrite

$$\boxed{M_{\mathcal{J},(k,k)} = 1 + \sum_{j\in\mathcal{J}\cap\phi_k} \frac{\gamma_j}{N_j}\left\|h_{j,k}\right\|^2} \tag{5.14}$$

Finally, it is possible to write (5.13) as

$$\boxed{R \leq \frac{1}{|\mathcal{J}|}\log_2\left[\prod_{k=1}^{K} M_{\mathcal{J},(k,k)}\right]} \tag{5.15a}$$

$$= \frac{K}{|\mathcal{J}|}\log_2\left[\sqrt[K]{\prod_{k=1}^{K} M_{\mathcal{J},(k,k)}}\right] \tag{5.15b}$$

$$= \frac{K}{|\mathcal{J}|}\left[\frac{1}{K}\sum_{k=1}^{K}\tilde{C}_k\left(M_{\mathcal{J},(k,k)}\right)\right] \tag{5.15c}$$

where $\tilde{C}_k\left(M_{\mathcal{J},(k,k)}\right) = \log_2\left[M_{\mathcal{J},(k,k)}\right] = \log_2\left[1 + \sum_{j\in\mathcal{J}\cap\phi_k}\frac{\gamma_j}{N_j}\left\|h_{j,k}\right\|^2\right]$. Note that $G\left(M_{\mathcal{J},(1,1)},\ldots,M_{\mathcal{J},(K,K)}\right) = \sqrt[K]{\prod_{k=1}^{K} M_{\mathcal{J},(k,k)}}$ is the **geometric mean** of the diagonal elements of $\mathbf{M}_{\mathcal{J}}$, while $\frac{1}{K}\sum_{k=1}^{K}\tilde{C}_k(\cdots)$ is the **arithmetic mean** of the sum-capacities of the K equivalent parallel channels.

The *insight* is that the symmetric rate R must be less than the **smaller average sum-capacity among the K parallel MAC channels scaled by $K\,|\mathcal{J}|^{-1}$.** In each channel, *at most* d_f users collide ($\mathcal{J}\cap\phi_k$) and determine the sum-power of the capacity in this channel. Designing the graph sets φ_j, ϕ_k (i.e., $\mathbf{F}$) will influence the result.

Recall that among the sets $\mathcal{J}$, we have J sets of $|\mathcal{J}| = 1$ (constraint of the single-user capacities) and 1 set of $|\mathcal{J}| = J$ (sum-capacity constraint). These $J+1$ sets *may* be crucial constraints in the final result. We consider that evaluation of the $J+1$ typical constraints related to the J single-user capacity bounds and to the sum-capacity bound is not an issue and may serve as interesting figures of merit.

However, without exhaustive search, it is not simple to determine which of the $2^J - 1$ sets $\mathcal{J}$ will produce the minimum value to upper bound R. The challenge is to have a simpler way to evaluate the remaining combinations of $\binom{J}{i}$ when $i = 2,\ldots,J-1$.

5.4.3 Single-User Capacity Constraints

When $\mathcal{J} = \{j\}$, for $j = 1,\ldots,J$, we take (5.15b) and write:

$$C_1 = C_{\text{su}}^{\min} = \min_j K \log_2 \left[\sqrt[K]{\prod_{k=1}^{K} M_{\{j\},(k,k)}} \right] \tag{5.16}$$

to represent the minimum among the J single-user capacities.

5.4.4 Sum-Capacity Constraint

When $\mathcal{J} = \{1, \ldots, J\}$, we take (5.15b) and write:

$$C_J = C_{\text{sum}} = \frac{1}{\beta} \log_2 \left[\sqrt[K]{\prod_{k=1}^{K} M_{\{1,\ldots,J\},(k,k)}} \right] \tag{5.17}$$

to represent the sum-capacity of the system. Note that $\beta = J/K$ is the load factor.

5.4.5 The Remaining Subsets' Sum-Capacity Constraints

We have already evaluated $J+1$ out of $2^J - 1$ possible constraints. There are $2^J - J$ remaining ones, related to sum-constraints that take L out of J users at a time, for $2 \leq L \leq J-1$.

The worse case scenario for the capacity is to have the worse set of K-parallel MAC channels. When taking subsets of $L \geq 2$ users, the worst case scenario will happen if instead of transmitting orthogonaly, the L users collide with each other.

The maximum collision degree in each parallel channel must be considered, i.e. $d_f = \left\lceil \beta \overline{N} \right\rceil$. If the d_f strongest users of the subset collide among themselves, the average capacity of the K-parallel channels will be worst than if one of these users transmit alone or with less interference.

The following algorithm tries to identify the worst case for each possible value of $|\mathcal{J}|$, assuming that γ_j, N_j and $h_{j,k}$ is known $\forall j, k$. These values may serve as a lower-bound.

The output of the algorithm are $J-2$ lower bounds $C_2, \ldots, C_{J-1}$, where C_i is the lower bound of (5.10) for any subset $\mathcal{J}$ with cardinality $|\mathcal{J}| = i$.

The algorithm relies on the insight that

$$(1+P) \leq \left(1 + \frac{P}{N}\right)^N$$

for every pair of (P, N) s.t. $P \geq 0$ and $N \geq 1$. Note that in the l.h.s. we could have written $P = N \times \frac{P}{N}$.

Note also that this algorithm has complexity order $O(\alpha(J-1))$, where α should take into account the complexity from rows 4 to 9.

Algorithm 3 Irregular Design for Uplink SCMA

procedure LOWER BOUND CASE3($J, K, \mathbf{N}, \boldsymbol{\gamma}, \mathbf{H}$)
  for $L \in \{2, \ldots, J-1\}$ **do** ▷ $J-2$ values
    $p_{j,k} = \gamma_j \left\| h_{j,k} \right\|^2$ for $j = 1, \ldots, J$ and $k = 1, \ldots, K$ ▷ SNR of user j in subcarrier k
    $\mathcal{P} = \left\{ p_j : p_j = \min_k p_{j,k} \right\}$ ▷ mininum SNRs of each user among the K subcarriers
    $\mathcal{P}_{\mathcal{J}} = m(\mathcal{P}, L)$ ▷ $m(\mathcal{P}, L)$ selects the subset with the L smallest values of set $\mathcal{P}$
    Let $p_{j_1}, \ldots, p_{j_L}$ be the elements of $\mathcal{P}_{\mathcal{J}}$ in ascending order ▷ the L-weakest users, most faded and/or with largest path-losses
    Compose the vector $\mathbf{p} = \left[\underbrace{\frac{p_{j_L}}{N_{j_L}}, \ldots, \frac{p_{j_L}}{N_{j_L}}}_{N_{j_L}\times} \quad \cdots \quad \underbrace{\frac{p_{j_1}}{N_{j_1}}, \ldots, \frac{p_{j_1}}{N_{j_1}}}_{N_{j_1}\times} \right]$ ▷ uniformly distributed powers of each user $j_L, \ldots, j_1$
    Reshape $\mathbf{p}$ into the $K \times d_f$ matrix $\mathbf{P}$, row-wise (fill the missing entries with zeros) ▷ largest powers will collide the most allowed times
    $\mathbf{s} = \mathbf{P} \times \mathbf{1}_{d_f}$, ▷ Sum the elements of each row, where entries $s_1, \ldots, s_K$ are the sum-power of each parallel channel
    $C_L = \frac{K}{L} \log_2 \left(\sqrt[\kappa]{\prod_{k=1}^{K} [1 + s_k]} \right)$ ▷ lower-bound of symm. capacity for all subsets $\mathcal{J}$ of length $|\mathcal{J}| = L$
  end for
  return $C_2, \ldots, C_{J-1}$
end procedure

Compared to the complexity order $O\left(2^J - J\right)$ of searching all the subsets, it looks more efficient.

It should be possible to improve the algorithm, given $\mathbf{F}$, to ensure that the lower-bounds are closer to the real bottlenecks of each $|\mathcal{J}|$.

5.4.6 Lower-bound for the Symmetric Capacity

From that, we assume that the lower-bound of the symmetric capacity will be:

$$\boxed{C_{\text{sym}} \geq C_{\text{sym}}^{\text{LB}} = \min_i \left(C_1, C_2, \ldots, C_{J-1}, C_{\text{sum}} \right)} \tag{5.18}$$

where we recall that $C_1 = C_{\text{su}}^{\min}$ and $C_J = C_{\text{sum}}$.

Using this lower bound and choosing $R = 2^k \leq C_{\text{sym}}$, symmetric rate codebooks can be designed with $M = \left\lfloor \log_2 C_{\text{sym}}^{\text{LB}}, \right\rfloor$. Note that J MDMs $C_1, \ldots, C_J$ must be designed with the respective pairs of parameters (M, N_j).

5.5 Numerical Results of Symmetric Capacity

In this Section, we consider the AWGN channel, thus $h_{j,k} = 1 \; \forall j, k$. Given $\mathbf{N}$, we build $\mathbf{F}$ with PEG algorithm.

Note that γ_j will be distributed according to the distance distribution among users, which in

turn are the parameters that determine the path-loss. We assume that the distance range of users from base station D_j [km] is uniformly spaced in D_j^2.

The irregular distributions of N_j are obtained with heuristics of an AWGN free-interference OMA channel, described in Appendix C. In those equations, N_j is a function of γ_j. Thus, if the users are in the same radius, resulting in $\gamma_j = \gamma \ \forall j$, the distribution results in regular SCMA $N_j = N = \lfloor \overline{N}_{\max} \rceil$. Two design approaches are used: uniformly weighted sum-capacity (WSCU) and max min-fairness (MMF).

Table 5.1 defines the scenario considered for the results in this section.

Table 5.1: Fixed simulations parameters.

Parameter	**Value**	**Meaning**
P	20 dBm	Tx power
Δ_f	15 kHz	subcarrier bandwidth
N_0	-174 dBm/Hz	noise power spectral density
Path Loss Model	$128.1 + 37.6 \log_{10}(D[\text{km}])$;	LTE Urban
K	10	num. of shared resources
J	15	num. of users
β	150%	load factor of SCMA
$2^J - 1$	32767	total num. of subsets $\mathcal{J}$

In Table 5.2, we compare the results of C_{sym}, (5.10), and $C_{\text{sym}}^{\text{LB}}$, (5.18), assuming regular and irregular distributions of N_j subject to $\overline{N} = 3$.

Table 5.2: Comparing C_{sym} for regular vs. irregular schemes for several regimes of user positions, considering $[N_{\min}, N_{\max}] = [2, 6]$. The design of $\mathbf{F}$ has some influence on the exact values. When multiple variable node degree have the same weight, permutations of the columns could lead to different results, since users have different path-losses.

Distances	C_{sym}							
	Reg $N = 2$		Reg $N = 3$		MMF		WSCU	
	Eq. (5.10)	Eq. (5.18)	Eq. (5.10)	Eq. (5.18)	Eq. (5.10)	Eq. (5.18)	Eq. (5.10)	Eq. (5.18)
$\mathcal{U}(0.25^2, 0.50^2)$	9.5182	6.4278	9.5308	7.7015	9.5083	8.1420	9.5363	5.3693
$\mathcal{U}(0.25^2, 1.50^2)$	6.5000	3.4660	6.6023	4.1977	6.4898	5.0444	6.6953	3.0327
$\mathcal{U}(0.5^2, 2.50^2)$	4.2360	2.6214	4.3671	2.5709	4.2505	3.0396	4.3011	1.9394
$\mathcal{U}(0.75^2, 1.75^2)$	5.2331	3.8424	5.2849	3.6559	5.2549	4.2972	5.3238	2.6696

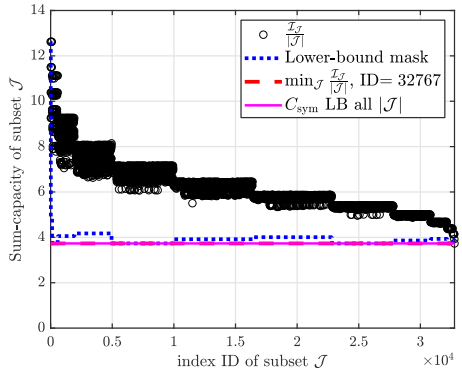
Evidence shows that the regular N_j or WSCU strategies have the largest symmetric capacity in most of the scenarios. The distance profile and the design of $\mathbf{F}$ will affect the ultimate result. In any case, there is little difference between C_{sym} of the different schemes in the scenarios of Table 5.2. This might suggest that PEG already does a good job in the design of $\mathbf{F}$ for a given distribution of $\mathbf{N}$.

With respect to the lower bound (5.18), it seems to have the largest values with the MMF strategy. However, this has no relevance, since it can be seen that the lower-bound order among

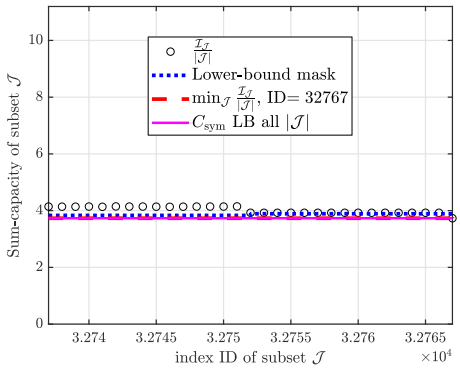
the schemes does not translate into the same performance order for the real C_{sym}. Thus, it is necessary to improve Algorithm 3 to find a tighter bound. The gap between C_{sym} and $C_{\text{sym}}^{\text{LB}}$ changes according to the distance profile of the scenarios.

If there is enough time for computation, it is preferable to use the effective symmetric capacity of Equation (5.10), since it allow for the design of schemes with larger symmetric rate. On the other hand, the application of the lower bound is suitable for designs with large J. This can be the result of large SCMA system, with respect to the parameters (J, K), or due to large load factor β (high overload regime).

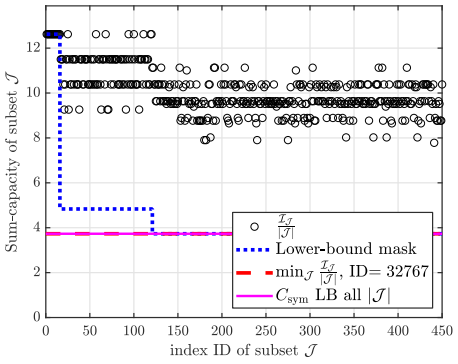
In Figure 5.2, we show another example to illustrate the results of the constraints of the symmetric rate R over all the users' subsets $\mathcal{J}$. The lower bound values $C_1, \ldots, C_J$ are also shown, denoted as the *lower-bound mask* for $\mathcal{I}_{\mathcal{J}}/|\mathcal{J}|$. A second example is illustrated in 5.3.



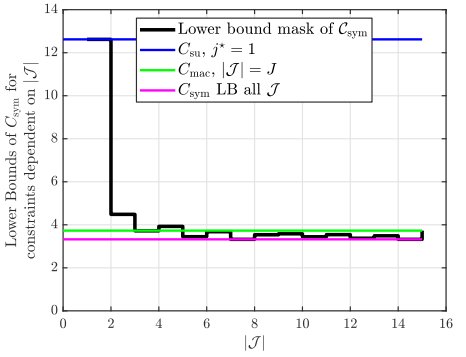
(a) Rate constraints for each set of $\mathcal{J}$.



(b) Detail where C_{sym} was found.



(c) Detail of the first $J \times 100$ subsets $\mathcal{J}$.

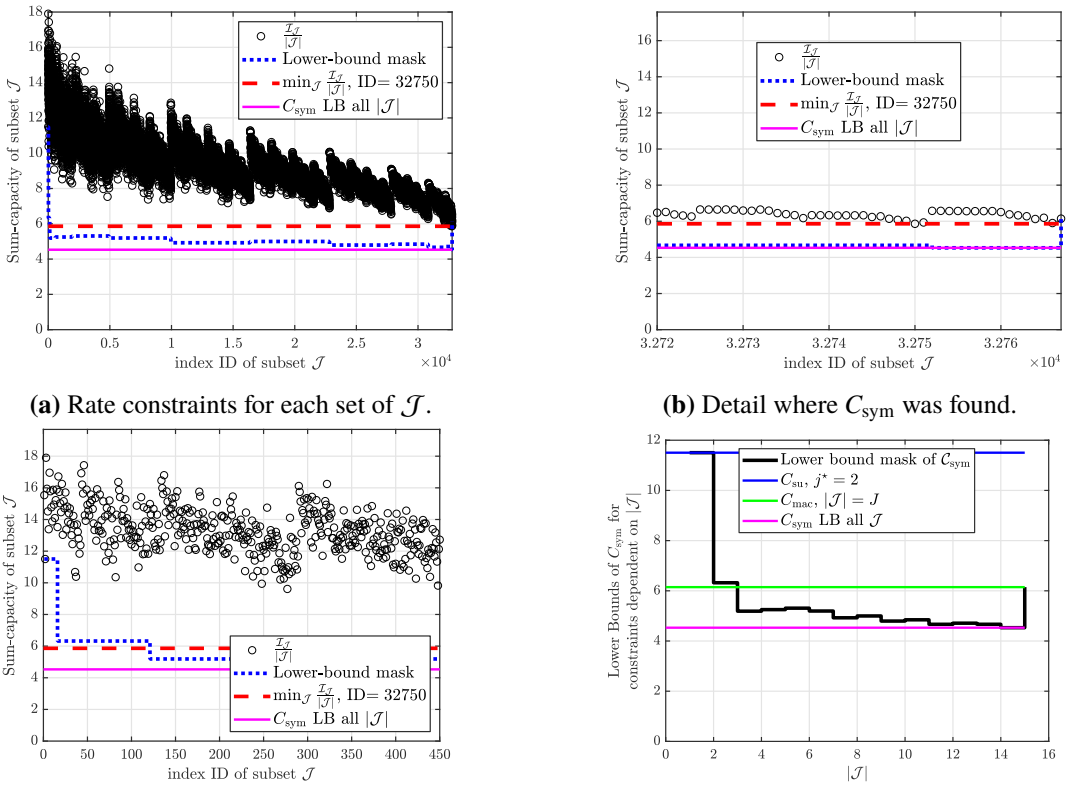


(d) Lower bound mask, without exhaustive search.

Figure 5.2: Example of $N_j = 4\ \forall j$. $D_j^2 \sim \mathcal{U}\left(D_{j,\min}^2, D_{j,\max}^2\right) = \mathcal{U}\left(1.75^2, 1.75^2\right)$ (moderate $\gamma_j = \gamma \approx 15$ dB $\forall j$, users are distant but in the same radius).

5.6 Bit Error Rate Performance SCMA

In this section, we consider link-level simulations of WAN-SCMA over Rayleigh fading channels. The simulation parameters are defined in Table 5.3. We evaluate the BER of systems



(a) Rate constraints for each set of $\mathcal{J}$.

(b) Detail where C_{sym} was found.

(c) Detail for the first $100 \times J$ subsets $\mathcal{J}$.

(d) Lower bound mask, without exhaustive search.

Figure 5.3: Example of MMF strategy ($\overline{N_j} = 3$), $N_j = [1,1,2,2,2,2,3,3,3,3,4,4,4,5,6]$. $D_j^2 \sim \mathcal{U}\left(D_{j,\min}^2, D_{j,\max}^2\right) = \mathcal{U}\left(0.25^2, 1.75^2\right)$. Power in the range of $(\gamma_J, \gamma_1) = (15.00, 46.77)$ dB.

with different strategies for the design of **N**.

Table 5.3: Simulations parameters.

Parameter	**Value**	**Meaning**
P	20 dBm	Tx power
Δ_f	15 kHz	subcarrier bandwidth
N_0	-174 dBm/Hz	noise power spectral density
Path Loss Model	$128.1 + 37.6 \log_{10}(D[\text{km}])$;	LTE Urban
D_j	$(0.75, 1.25)$ [km]	distance range of users from base station, D_j^2 uniformly spaced
$[N_{\min}, N_{\max}]$	$[3,5]$	min. and max. bounds for codebook dimensions
$\overline{N}_{\max}$	4.0	maximum average VN degree
K	16	num. of shared resources
J	24	num. of users
β	150.00%	load factor of SCMA
T_m	6	max. MPA iterations

Irregular designs use the heuristics from Appendix C. Also, given **N**, **F** is designed with PEG algorithm and the matrix is fixed during the Monte Carlo simulations of the channel.

In Table 5.4 we show the symmetric capacity (and its lower bound) calculated for the AWGN channel. It can be seen that the difference between schemes' capacities is not much. To comply with the AWGN capacity, the codebooks from Chapter 3, Section 3.5.2, are reused. Thus, each user will transmit with a rate of $R = \log_2 M = 3$ bits per channel use.

Table 5.4: Results of symmetric capacity over the AWGN channel. The parameters of the scenario are from Table 5.3.

N_j	C_{sym} Eq. (5.10)	$C_{\text{sym}}^{(\text{LB})}$ Eq. (5.18)
2	5.8520	4.9861
3	5.8629	4.6264
4	5.8653	4.4938
3 – 5 MMF	5.8601	4.9915
3 – 5 WSCU	5.8633	3.9267

Figure 5.3 compares the complexity order of the different schemes when using the MPA. The codebook with $N = 3$ is the one with the smaller complexity, due to its small alphabet size $q = 3$ and small spreading degree. Nevertheless, all strategies have the same order of complexity $O\left(10^5\right)$.

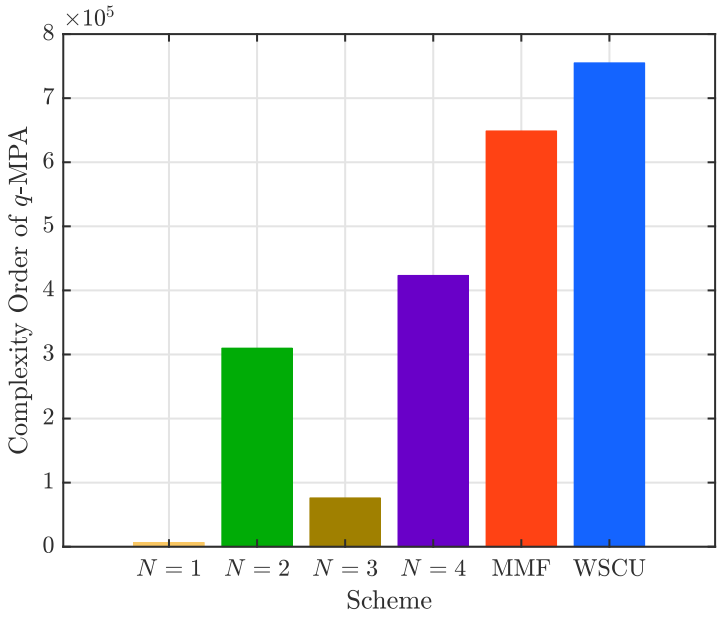


Figure 5.4: Average complexity order of each regular and irregular scheme.

Figure 5.5 shows the BER results of each scheme. It can be noticed that the schemes with smallest average BER are $N = 4$ and MMF. Regarding these schemes, although the average BER is approximately the same, the MMF resource allocation balances out the BER between users that are close to the base station and the ones that are far from it. The MMF BER has three

distinguishable regimes, related to the the spreading degrees $N = 3,4$ and 5.

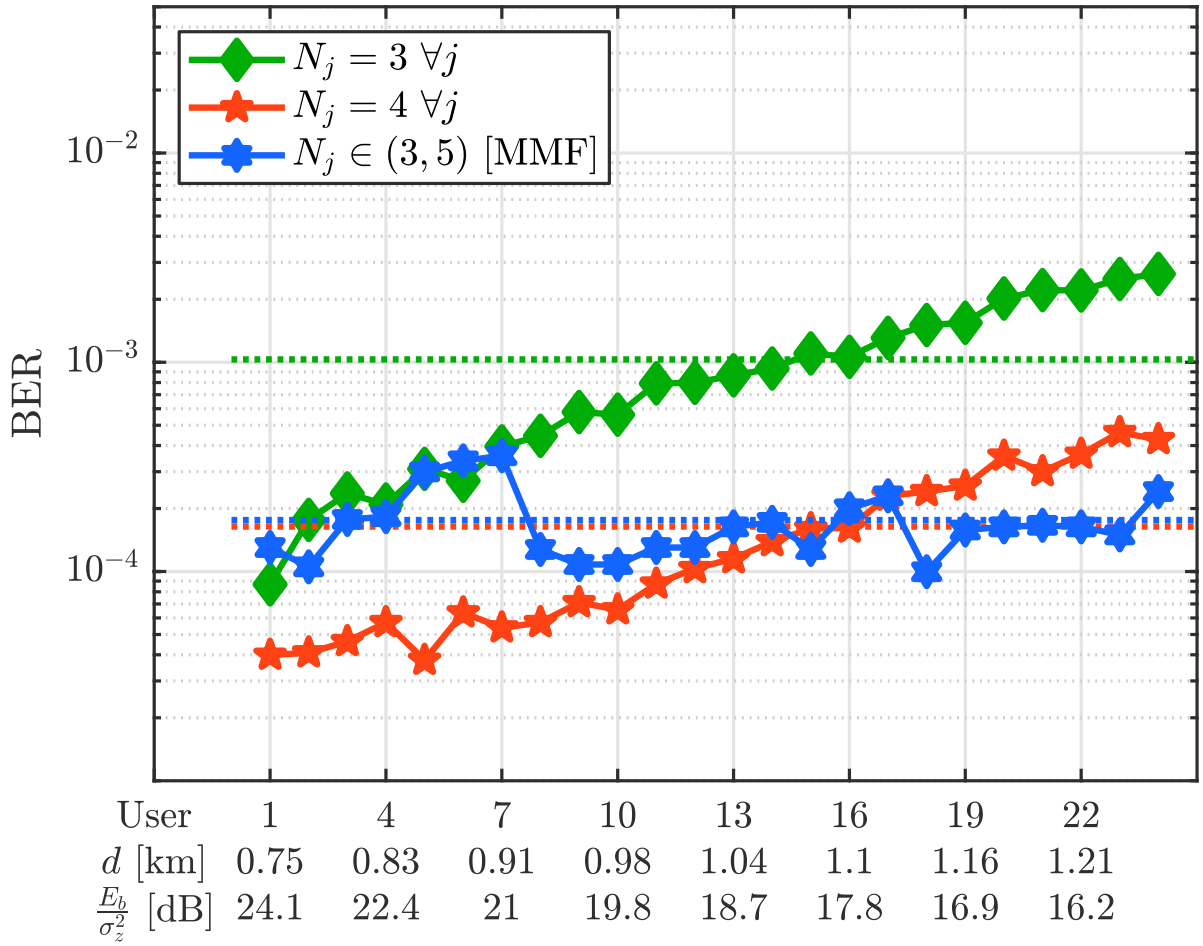


Figure 5.5: BER of irregular MMF ($N \in [3,5]$) vs. regular $N = \{3,4\}$, over the Rayleigh small-scale fading channels. Min. errors: 20 symbol errors per user. Max iterations: 2×10^6 ($P_s = 10^{-5}$).

• • •

6 Conclusion

In this thesis, we have presented an overview of the state-of-the-art research on SCMA. More importantly, we have approached SCMA system design in a multistage manner, proposing contributions in three directions: codebook design, low-complexity multiuser detection and resource allocation methods.

With respect to the proposed codebook design methods, we have shown that SCMA codebooks (or multidimensional modulations, in general) can be viewed as M-ary sets of N-dimensional complex-vectors which have an underlying subspace structure analogous to the generated subspace of a linear block code. Therefore, classic error-correcting codes can be used as a building block of SCMA codebooks, which inherit the minimum Hamming distance property from the block codes. The use of q-ary alphabets, such that $q \ll M$, enables the design of SCMA codebooks with signal-space diversity up to $N-1$, while keeping the complexity order of the FN operations of the MPA $O\left(q^{\beta N}\right)$ under control in comparison to M-ary alphabets.

On the contribution to low-complexity MUD algorithms, we have proposed an improved edge-selection MPA, based on works proposed in the literature. Our modification includes a threshold parameter to separate weak and strong users in the interference cancellation equations of the messages. By properly choosing the threshold, it is possible to have error-performance as good as the original algorithm but with an average complexity significantly smaller. Moreover, the scheme works with codebooks of large spectral-efficiency and small-size alphabets, while the original algorithm seems to fail in some of these scenarios.

Finally, we have discussed resource allocation methods with regular and irregular spreading degrees. Both the MPA complexity order and the channel capacity of the system are considered to formulate the optimization problems. However, a solution could not be derived because of the complexity involved with the objective function. Nevertheless, we have designed suboptimal systems and evaluated the symmetric capacity in order to simulate symmetric rate SCMA

systems in cellular scenarios with large- and small-scale fading channels. Results show that it is possible to work with codebooks of different dimensions and irregular resource allocation with similar performance as regular designs.

6.1 Future works

For the continuity of this research, the author would like to suggest the following directions:

1. further investigation on linear complexity multiuser detection for SCMA is necessary to make the system more appealing for future implementations with large-valued system parameters;

2. recent results using deep-learning aided receivers [34, 44] show promising results, although more investigation is needed to make it work with larger systems and under Rayleigh fading channels;

3. optimization of the resource allocation for SCMA with respect to the channel capacity seems to be a challenging problem; further investigation on how to maximize the symmetric capacity with feasible design algorithms seems to be an interesting direction, which could enable the design of improved and highly reliable SCMA communication systems;

4. several works have extended the SCMA model to MIMO systems; however, MIMO implies that the transmit symbols of each antenna from the same user may cause self-interference at the receiver antennas, increasing dramatically the MPA complexity as the parameters scale; it should be interesting to evaluate the proposed codebooks and TB-ESGA-MPA in MIMO scenarios to investigate the trade-offs involved.

5. finally, the SCMA codebooks based on MDS block codes have good properties and a simple design method; the bottleneck for application of these type of codebooks with large dimensions and large signal-space diversity is limited by the complexity order of the multiuser detection, since the FN operations of the MPA are exponential to βN; hence, it would be desirable to simplify the operations of the receiver considering the structure of these specific codebooks.

6.2 Published Papers

During the development of this thesis, the following papers have been published:

1. B. Fontana da Silva, C. A. A. Meza, D. Silva and B. F. Uchôa-Filho, "**Exploiting spatial diversity in overloaded MIMO LDS-OFDM multiple access systems**," 2017 IEEE 9th Latin-American Conference on Communications (LATINCOM), Guatemala City, 2017, pp. 1-6. doi: 10.1109/LATINCOM.2017.8240149

2. B. F. da Silva, D. Silva, B. F. Uchôa-Filho, and D. Le Ruyet. "**A Multistage Method for SCMA Codebook Design Based on MDS Codes**". In: IEEE Wireless Communications Letters (2019). doi: 10.1109/LWC.2019.2925801.

3. Bruno Fontana da Silva, Bartolomeu F. Uchôa-Filho and Didier Le Ruyet, "**Receptor SCMA de Baixa Complexidade Baseado em um MPA com Seleção Dinâmica de Arestas e Aproximação Gaussiana**", XXXVII Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrT), Petrópolis (RJ), Brazil, September 2019.

4. B. F. da Silva, D. Silva, D. Le Ruyet and B. F. Uchôa-Filho. "**Threshold-Based Edge Selection MPA for SCMA**". In: IEEE Transactions on Vehicular Technology (2019). *Submitted on July 31st 2019. Current status: Major Review*.

Other papers published during the course of this Ph.D., unrelated to the main subject of the thesis and/or in collaboration (as co-author) with other researchers:

5. B. Fontana da Silva and Bartolomeu F. Uchôa-Filho, "**Detector de Baixa Complexidade para Modulação Espacial usando Equalização Linear com Forçagem a Inteiros**", XXXV Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrT), São Pedro (SP), Brazil, September 2017/9. doi: 10.14209/SBRT.2017.252.

6. Ana L. Scharf, Bartolomeu F. Uchôa-Filho, Bruno Fontana da Silva and Didier Le Ruyet "**Decodificação Iterativa em Árvore de Baixa Complexidade para SCMA**", XXXVII Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrT), Petrópolis (RJ), Brazil, September 2019.

7. Jose Clair Menezes Júnior, Bruno Fontana da Silva and Bartolomeu F. Uchôa-Filho, "**LDPC Code Design for Two-Layer Spatial Modulation**", Journal of Communication and Information Systems, VOL. 34, No.1, 2019. doi: 10.14209/jcis.2019.23

8. D. P. Bertineti, L. N. Canha, A. P. Medeiros, R.M. de Azevedo, and B. F. da Silva, "**Heuristic Scheduling Algorithm for Load Shift DSM Strategy in Smart Grids**

and IoT Scenarios", Innovative Smart Grid Technologies (ISGT) Latin America 2019, IEEE Power and Energy Society (PES), Gramado (RS), Brazil, September 2019. doi: 10.1109/ISGT-LA.2019.8895488.

6.3 Financial Support

This work has been funded by National Council for Scientific and Technological Development (CNPq) with scholarships 141161/2016-7 and 204676/2018-5. The author also acknowledges the support of *Instituto Federal Sul-rio-grandense* for the absence from work granted in "*Portaria* 1777/2018" of the Federal Public Service.

References

[1] 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, and R. Razavi. "Toward the Standardization of Non-Orthogonal Multiple Access for Next Generation Wireless Networks". In: *IEEE Communications Magazine* 56.3 (Mar. 2018), pp. 19–27. DOI: 10.1109/MCOM.2018.1700845.

[2] Y. Cai, Z. Qin, F. Cui, G. Y. Li, and J. A. McCann. "Modulation and Multiple Access for 5G Networks". In: *IEEE Communications Surveys Tutorials* 20.1 (2018), pp. 629–646. DOI: 10.1109/COMST.2017.2766698.

[3] H. Nikopour and H. Baligh. "Sparse code multiple access". In: *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. Sept. 2013, pp. 332–336. DOI: 10.1109/PIMRC.2013.6666156.

[4] R. Hoshyar, F. P. Wathan, and R. Tafazolli. "Novel Low-Density Signature for Synchronous CDMA Systems Over AWGN Channel". In: *IEEE Transactions on Signal Processing* 56.4 (Apr. 2008), pp. 1616–1626. DOI: 10.1109/TSP.2007.909320.

[5] M. Taherzadeh, H. Nikopour, A. Bayesteh, and H. Baligh. "SCMA Codebook Design". In: *2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall)*. Sept. 2014, pp. 1–5. DOI: 10.1109/VTCFall.2014.6966170.

[6] S. Zhang, K. Xiao, B. Xiao, Z. Chen, B. Xia, D. Chen, and S. Ma. "A capacity-based codebook design method for sparse code multiple access systems". In: *2016 8th International Conference on Wireless Communications Signal Processing (WCSP)*. Oct. 2016, pp. 1–5. DOI: 10.1109/WCSP.2016.7752620.

[7] J. Bao, Z. Ma, M. Xiao, T. A. Tsiftsis, and Z. Zhu. "Bit-Interleaved Coded SCMA With Iterative Multiuser Detection: Multidimensional Constellations Design". In: *IEEE Transactions on Communications* 66.11 (Nov. 2018), pp. 5292–5304. DOI: 10.1109/TCOMM.2017.2782325.

[8] A. Bayesteh, H. Nikopour, M. Taherzadeh, H. Baligh, and J. Ma. "Low Complexity Techniques for SCMA Detection". In: *2015 IEEE Globecom Workshops (GC Wkshps)*. San Diego, CA, Dec. 2015, pp. 1–6. DOI: 10.1109/GLOCOMW.2015.7414184.

[9] Y. Du, B. Dong, Z. Chen, J. Fang, P. Gao, and Z. Liu. “Low-Complexity Detector in Sparse Code Multiple Access Systems”. In: *IEEE Communications Letters* 20.9 (Sept. 2016), pp. 1812–1815. DOI: 10.1109/LCOMM.2016.2592912.

[10] Y. Wang and L. Qiu. “Edge Selection-Based Low Complexity Detection Scheme for SCMA System”. In: *2016 IEEE 84th Vehicular Technology Conference (VTC-Fall)*. Montreal, QC, Sept. 2016, pp. 1–5. DOI: 10.1109/VTCFall.2016.7880976.

[11] X. Meng, Y. Wu, Y. Chen, and M. Cheng. “Low Complexity Receiver for Uplink SCMA System via Expectation Propagation”. In: *2017 IEEE Wireless Communications and Networking Conference (WCNC)*. Mar. 2017, pp. 1–5. DOI: 10.1109/WCNC.2017.7925590.

[12] O. Shental, B. M. Zaidel, and S. S. Shitz. “Low-density code-domain NOMA: Better be regular”. In: *2017 IEEE International Symposium on Information Theory (ISIT)*. June 2017, pp. 2628–2632. DOI: 10.1109/ISIT.2017.8007005.

[13] B. M. Zaidel, O. Shental, and S. S. Shitz. “Sparse NOMA: A Closed-Form Characterization”. In: *2018 IEEE International Symposium on Information Theory (ISIT)*. June 2018, pp. 1106–1110. DOI: 10.1109/ISIT.2018.8437642.

[14] M. T. P. Le, G. C. Ferrante, G. Caso, L. De Nardis, and M. Di Benedetto. “On information-theoretic limits of code-domain NOMA for 5G”. In: *IET Communications* 12.15 (2018), pp. 1864–1871. DOI: 10.1049/iet-com.2018.5241.

[15] Q. He, B. Bai, D. Feng, H. Xu, and M. Zhu. “A Nonbinary LDPC-Coded SCMA System with Optimized Codebook Design”. In: *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*. Sept. 2017, pp. 1–6. DOI: 10.1109/VTCFall.2017.8288397.

[16] Li Li, Ping Zhi Fan, and Lajos Hanzo. “High dimensional codebook design for the SCMA down-link”. en. In: *IEEE Transactions on Vehicular Technology* (Aug. 2018), pp. 1–5.

[17] Z. Yang, X. Lei, Z. Ding, P. Fan, and G. K. Karagiannidis. “On the Uplink Sum Rate of SCMA System With Randomly Deployed Users”. In: *IEEE Wireless Communications Letters* 6.3 (June 2017), pp. 338–341. DOI: 10.1109/LWC.2017.2689766.

[18] R. Hoshyar, F. P. Wathan, and R. Tafazolli. “CTH06-4: Novel Low-Density Signature Structure for Synchronous DS-CDMA Systems”. In: *IEEE Globecom 2006*. Nov. 2006, pp. 1–5. DOI: 10.1109/GLOCOM.2006.67.

[19] Behrouz A. Forouzan. *Data Communications and Networking*. English. 5 edition. New York, NY: McGraw-Hill Education, Feb. 2012.

[20] Andrea Goldsmith. *Wireless Communications*. English. 1 edition. Cambridge ; New York: Cambridge University Press, Aug. 2005.

[21] S. Chen, B. Ren, Q. Gao, S. Kang, S. Sun, and K. Niu. “Pattern Division Multiple Access—A Novel Nonorthogonal Multiple Access for Fifth-Generation Radio Networks”. In: *IEEE Transactions on Vehicular Technology* 66.4 (Apr. 2017), pp. 3185–3196. DOI: 10.1109/TVT.2016.2596438.

[22] S. M. R. Islam, N. Avazov, O. A. Dobre, and K. S. Kwak. "Power-Domain Non-Orthogonal Multiple Access (NOMA) in 5G Systems: Potentials and Challenges". In: *IEEE Communications Surveys Tutorials* PP.99 (2016), pp. 1–1. DOI: 10.1109/COMST.2016.2621116.

[23] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi. "Non-Orthogonal Multiple Access (NOMA) for Cellular Future Radio Access". In: *Vehicular Technology Conference (VTC Spring), 2013 IEEE 77th*. June 2013, pp. 1–5. DOI: 10.1109/VTCSpring.2013.6692652.

[24] K. Xiao, B. Xia, Z. Chen, B. Xiao, D. Chen, and S. Ma. "On Capacity-Based Codebook Design and Advanced Decoding for Sparse Code Multiple Access Systems". In: *IEEE Transactions on Wireless Communications* 17.6 (June 2018), pp. 3834–3849. DOI: 10.1109/TWC.2018.2816929.

[25] Z. Li, W. Chen, F. Wei, F. Wang, X. Xu, and Y. Chen. "Joint codebook assignment and power allocation for SCMA based on capacity with Gaussian input". In: *2016 IEEE/CIC International Conference on Communications in China (ICCC)*. July 2016, pp. 1–6. DOI: 10.1109/ICCChina.2016.7636728.

[26] M. Moltafet, N. M. Yamchi, M. R. Javan, and P. Azmi. "Comparison Study Between PD-NOMA and SCMA". In: *IEEE Transactions on Vehicular Technology* 67.2 (Feb. 2018), pp. 1830–1834. DOI: 10.1109/TVT.2017.2759910.

[27] L. Yang, Y. Liu, and Y. Siu. "Low Complexity Message Passing Algorithm for SCMA System". In: *IEEE Communications Letters* 20.12 (Dec. 2016), pp. 2466–2469. DOI: 10.1109/LCOMM.2016.2609382.

[28] L. Wei, B. Huang, and J. Zheng. "Low-Complexity Detectors for Uplink SCMA: Symbol Flipping and Dynamic Partial Marginalization-Based MPA". In: *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*. Porto, June 2018, pp. 1–5. DOI: 10.1109/VTCSpring.2018.8417841.

[29] D. Wei, Y. Han, S. Zhang, and L. Liu. "Weighted message passing algorithm for SCMA". In: *2015 International Conference on Wireless Communications Signal Processing (WCSP)*. Nanjing, China, Oct. 2015, pp. 1–5. DOI: 10.1109/WCSP.2015.7341312.

[30] Y. Du, B. Dong, Z. Chen, J. Fang, and L. Yang. "Shuffled Multiuser Detection Schemes for Uplink Sparse Code Multiple Access Systems". In: *IEEE Communications Letters* 20.6 (June 2016), pp. 1231–1234. DOI: 10.1109/LCOMM.2016.2551742.

[31] Y. Du, B. Dong, Z. Chen, J. Fang, and X. Wang. "A Fast Convergence Multiuser Detection Scheme for Uplink SCMA Systems". In: *IEEE Wireless Communications Letters* 5.4 (Aug. 2016), pp. 388–391. DOI: 10.1109/LWC.2016.2565581.

[32] J. Dai, K. Niu, C. Dong, and J. Lin. "Improved Message Passing Algorithms for Sparse Code Multiple Access". In: *IEEE Transactions on Vehicular Technology* 66.11 (Nov. 2017), pp. 9986–9999. DOI: 10.1109/TVT.2017.2741525.

[33] Xiaotian Fu, Mylene Pischella, and Didier Le Ruyet. "On Gaussian Approximation Algorithms for SCMA". In: *2019 16th International Symposium on Wireless Communication Systems (ISWCS)*. ISSN: 2154-0225, 2154-0217. Aug. 2019, pp. 155–160. DOI: 10.1109/ISWCS.2019.8877279.

[34] Minhoe Kim, Nam-I Kim, Woongsup Lee, and Dong-Ho Cho. "Deep Learning-Aided SCMA". en. In: *IEEE Communications Letters* 22.4 (Apr. 2018), pp. 720–723. DOI: 10.1109/LCOMM.2018.2792019.

[35] M. Vameghestahbanati, I. D. Marsland, R. H. Gohary, and H. Yanikomeroglu. "Multidimensional Constellations for Uplink SCMA Systems—A Comparative Study". In: *IEEE Communications Surveys Tutorials* 21.3 (2019), pp. 2169–2194. DOI: 10.1109/COMST.2019.2910569.

[36] Mahmoud Taherzadehboroujeni, Hosein Nikopour, Alireza Bayesteh, and Mohammadhadi Baligh. "System and Method for Designing and Using Multidimensional Constellations". Pat. US20140369434 A1. U.S. Classification 375/261, 375/260; International Classification H04L5/00, H04B7/04; Cooperative Classification H04L25/03929, H04L27/3444, H03M13/616, H04B7/0482, H04B7/0617, H04L25/03942, H04L5/0016, H04B7/0456, H04L5/0007. Dec. 2014.

[37] J. Bao, Z. Ma, Z. Ding, G. K. Karagiannidis, and Z. Zhu. "On the Design of Multiuser Codebooks for Uplink SCMA Systems". In: *IEEE Communications Letters* 20.10 (Oct. 2016), pp. 1920–1923. DOI: 10.1109/LCOMM.2016.2596759.

[38] D. Cai, P. Fan, X. Lei, Y. Liu, and D. Chen. "Multi-Dimensional SCMA Codebook Design Based on Constellation Rotation and Interleaving". In: *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*. May 2016, pp. 1–5. DOI: 10.1109/VTCSpring.2016.7504356.

[39] Y. Zhou, Q. Yu, W. Meng, and C. Li. "SCMA codebook design based on constellation rotation". In: *2017 IEEE International Conference on Communications (ICC)*. May 2017, pp. 1–6. DOI: 10.1109/ICC.2017.7996395.

[40] L. Yu, P. Fan, D. Cai, and Z. Ma. "Design and Analysis of SCMA Codebook Based on Star-QAM Signaling Constellations". In: *IEEE Transactions on Vehicular Technology* 67.11 (Nov. 2018), pp. 10543–10553. DOI: 10.1109/TVT.2018.2865920.

[41] L. Yu, P. Fan, and Z. Han. "Maximizing Spectral Efficiency for SCMA Systems With Codebooks Based on Star-QAM Signaling Constellations". In: *IEEE Wireless Communications Letters* 8.4 (Aug. 2019), pp. 1163–1166. DOI: 10.1109/LWC.2019.2909913.

[42] J. Bao, Z. Ma, M. A. Mahamadu, Z. Zhu, and D. Chen. "Spherical Codes for SCMA Codebook". In: *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*. May 2016, pp. 1–5. DOI: 10.1109/VTCSpring.2016.7504489.

[43] V. P. Klimentyev and A. B. Sergienko. "SCMA Codebooks Optimization Based on Genetic Algorithm". In: *European Wireless 2017; 23th European Wireless Conference*. May 2017, pp. 1–6.

[44] Jinzhi Lin, Shengzhong Feng, Zhile Yang, Yun Zhang, and Yong Zhang. "A Novel Deep Neural Network Based Approach for Sparse Code Multiple Access". In: *arXiv:1906.03169 [cs, eess, math, stat]* (June 2019). arXiv: 1906.03169.

[45] J. Dai, K. Niu, C. Dong, and J. Lin. "Scaling Factor Aided Iterative Multiuser Receiver for Sparse Code Multiple Access". In: *IEEE Communications Letters* 21.8 (Aug. 2017), pp. 1723–1726. DOI: 10.1109/LCOMM.2017.2702636.

[46] F. Wei and W. Chen. "A Low Complexity SCMA Decoder Based on List Sphere Decoding". In: *2016 IEEE Global Communications Conference (GLOBECOM)*. Dec. 2016, pp. 1–6. DOI: 10.1109/GLOCOM.2016.7841513.

[47] Monirosharieh Vameghestahbanati, Ebrahim Bedeer, Ian Marsland, Ramy H. Gohary, and Halim Yanikomeroglu. "Enabling Sphere Decoding for SCMA". en. In: *IEEE Communications Letters* 21.12 (Dec. 2017), pp. 2750–2753. DOI: 10.1109/LCOMM.2017.2747550.

[48] F. Wei and W. Chen. "Low Complexity Iterative Receiver Design for Sparse Code Multiple Access". In: *IEEE Transactions on Communications* 65.2 (Feb. 2017), pp. 621–634. DOI: 10.1109/TCOMM.2016.2631468.

[49] G. Chen, J. Dai, K. Niu, and C. Dong. "Optimal receiver design for SCMA system". In: *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. Oct. 2017, pp. 1–6. DOI: 10.1109/PIMRC.2017.8292420.

[50] Lin Yang, Xinying Ma, and Yunming Siu. "Low Complexity MPA Detector Based on Sphere Decoding for SCMA". en. In: *IEEE Communications Letters* 21.8 (Aug. 2017), pp. 1855–1858. DOI: 10.1109/LCOMM.2017.2697425.

[51] L. Li, J. Wen, X. Tang, and C. Tellambura. "Modified Sphere Decoding for Sparse Code Multiple Access". In: *IEEE Communications Letters* 22.8 (Aug. 2018), pp. 1544–1547. DOI: 10.1109/LCOMM.2018.2848273.

[52] L. Karakchieva and P. Trifonov. "Joint list multistage decoding with sphere detection for polar coded SCMA systems". In: *SCC 2019; 12th International ITG Conference on Systems, Communications and Coding*. Feb. 2019, pp. 1–6. DOI: 10.30420/454862018.

[53] M. Vameghestahbanati, I. Marsland, R. Gohary, and H. Yanikomeroglu. "A Novel SD-based Detection for Generalized SCMA Constellations". In: *IEEE Transactions on Vehicular Technology* (2019), pp. 1–1. DOI: 10.1109/TVT.2019.2932907.

[54] Y. Du, B. Dong, Z. Chen, P. Gao, and J. Fang. "Joint Sparse Graph-Detector Design for Downlink MIMO-SCMA Systems". In: *IEEE Wireless Communications Letters* 6.1 (Feb. 2017), pp. 14–17. DOI: 10.1109/LWC.2016.2623785.

[55] J. Dai, G. Chen, K. Niu, and J. Lin. "Partially Active Message Passing Receiver for MIMO-SCMA Systems". In: *IEEE Wireless Communications Letters* 7.2 (Apr. 2018), pp. 222–225. DOI: 10.1109/LWC.2017.2767588.

[56] W. Yuan, N. Wu, Q. Guo, Y. Li, C. Xing, and J. Kuang. "Iterative Receivers for Downlink MIMO-SCMA: Message Passing and Distributed Cooperative Detection". In: *IEEE Transactions on Wireless Communications* 17.5 (May 2018), pp. 3444–3458. DOI: 10.1109/TWC.2018.2813378.

[57] S. Chen, K. Peng, Y. Zhang, and J. Song. "A Comparative Study on the Capacity of SCMA and LDSMA". In: *2018 14th International Wireless Communications Mobile Computing Conference (IWCMC)*. June 2018, pp. 1099–1103. DOI: 10.1109/IWCMC.2018.8450269.

[58] Zhiqiang Wei, Lei Yang, Derrick Wing Kwan Ng, Jinhong Yuan, and Lajos Hanzo. "On the Performance Gain of NOMA over OMA in Uplink Communication Systems". In: *arXiv:1903.01683 [cs, math]* (Mar. 2019). arXiv: 1903.01683.

[59] M. T. P. Le, G. C. Ferrante, T. Q. S. Quek, and M. Di Benedetto. "Fundamental Limits of Low-Density Spreading NOMA With Fading". In: *IEEE Transactions on Wireless Communications* 17.7 (July 2018), pp. 4648–4659. DOI: 10.1109/TWC.2018.2828853.

[60] G. Xiong and J. Sun. “An Optimal Resource Allocation Algorithm Based on Sum Rate Maximization for Uplink SCMA system”. In: *2018 IEEE 18th International Conference on Communication Technology (ICCT)*. Oct. 2018, pp. 805–810. DOI: 10.1109/ICCT.2018.8599973.

[61] Y. Wang, M. Zhao, D. Deng, S. Zhou, and W. Zhou. “Fractional Sparse Code Multiple Access and Its Optimization”. In: *IEEE Wireless Communications Letters* 7.6 (Dec. 2018), pp. 990–993. DOI: 10.1109/LWC.2018.2846571.

[62] N. M. Balasubramanya, S. Payami, and M. Sellathurai. “Uplink Resource Allocation for Shared LTE and SCMA IoT Systems”. In: *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*. June 2018, pp. 1–5. DOI: 10.1109/VTCSpring.2018.8417621.

[63] Jianjun Peng, Wei Chen, Bo Bai, Xin Guo, and Chen Sun. “Joint Optimization of Constellation With Mapping Matrix for SCMA Codebook Design”. In: *IEEE Signal Processing Letters* 24.3 (Mar. 2017), pp. 264–268. DOI: 10.1109/LSP.2017.2653845.

[64] M. Dabiri and H. Saeedi. “Dynamic SCMA Codebook Assignment Methods: A Comparative Study”. In: *IEEE Communications Letters* 22.2 (Feb. 2018), pp. 364–367. DOI: 10.1109/LCOMM.2017.2764469.

[65] G. Caire, G. Taricco, and E. Biglieri. “Bit-interleaved coded modulation”. In: *IEEE Transactions on Information Theory* 44.3 (May 1998), pp. 927–946. DOI: 10.1109/18.669123.

[66] Juliana Camilo Inácio. “Design and analysis of fully-diverse multidimensional constellations for Rayleigh fading channels”. pt. PhD thesis. Florianópolis, Brazil: Universidade Federal de Santa Catarina (UFSC), 2018.

[67] J. C. Inácio, B. F. Uchôa-Filho, and D. Le Ruyet. “Exploiting Signal Space Diversity in OFDM With Grouped Subcarriers: Going Beyond Subcarrier Index Modulation”. In: *IEEE Wireless Communications Letters* 7.4 (Aug. 2018), pp. 650–653. DOI: 10.1109/LWC.2018.2807446.

[68] J. Boutros and E. Viterbo. “Signal space diversity: a power- and bandwidth-efficient diversity technique for the Rayleigh fading channel”. In: *IEEE Transactions on Information Theory* 44.4 (July 1998), pp. 1453–1467. DOI: 10.1109/18.681321.

[69] J. Bao, Z. Ma, G. K. Karagiannidis, M. Xiao, and Z. Zhu. “Joint Multiuser Detection of Multidimensional Constellations Over Fading Channels”. In: *IEEE Transactions on Communications* 65.1 (Jan. 2017), pp. 161–172. DOI: 10.1109/TCOMM.2016.2622706.

[70] S. C. Lim, N. Kim, and H. Park. “Uplink SCMA System With Multiple Antennas”. In: *IEEE Transactions on Vehicular Technology* 66.8 (Aug. 2017), pp. 6982–6992. DOI: 10.1109/TVT.2017.2664894.

[71] C. Dong, X. Cai, K. Niu, and J. Lin. “An Efficient SCMA Codebook Design Based on 1-D Searching Algorithm”. In: *IEEE Communications Letters* 22.11 (Nov. 2018), pp. 2234–2237. DOI: 10.1109/LCOMM.2018.2868938.

[72] Daniel Zwillinger, ed. *Table of Integrals, Series, and Products*. English. 8 edition. Amsterdam ; Boston: Academic Press, Oct. 2014.

[73] B. F. da Silva, D. Silva, B. F. Uchôa-Filho, and D. Le Ruyet. "A Multistage Method for SCMA Codebook Design Based on MDS Codes". In: *IEEE Wireless Communications Letters* (2019). DOI: 10.1109/LWC.2019.2925801.

[74] Xiao-Yu Hu, E. Eleftheriou, and D. M. Arnold. "Regular and irregular progressive edge-growth tanner graphs". In: *IEEE Transactions on Information Theory* 51.1 (Jan. 2005), pp. 386–398. DOI: 10.1109/TIT.2004.839541.

[75] S. Zhang, X. Xu, L. Lu, Y. Wu, G. He, and Y. Chen. "Sparse code multiple access: An energy efficient uplink approach for 5G wireless systems". In: *2014 IEEE Global Communications Conference*. Austin, TX, Dec. 2014, pp. 4782–4787. DOI: 10.1109/GLOCOM.2014.7037563.

[76] Todd K. Moon. *Error correction coding: mathematical methods and algorithms*. Hoboken, N.J: Wiley-Interscience, 2005.

[77] F. Schreckenbach, N. Gortz, J. Hagenauer, and G. Bauch. "Optimization of symbol mappings for bit-interleaved coded Modulation with iterative decoding". In: *IEEE Communications Letters* 7.12 (Dec. 2003), pp. 593–595. DOI: 10.1109/LCOMM.2003.821325.

[78] Svenja Huntemann. "The upper bound of general Maximum Distance Separable codes". en. In: *Math 4200: Honours Project* (May 2012), p. 45.

[79] F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-Correcting Codes, Volume 16*. English. Amsterdam: North Holland Publishing Co., 1977.

[80] Simeon Ball. "The Maximum Distance Separable Codes Conjecture". In: *3rd International Castle Meeting on Coding Theory and Applications*. Cardona Castle, Barcelona - Spain, 2011.

[81] Herbert A. David and Haikady N. Nagaraja. *Order Statistics*. English. Ed. 3. Hoboken, N.J: Wiley-Interscience, 2003.

[82] Consultative Committee for Space Data Systems (CCSDS). *Short Block Length LDPC Codes for TC Synchronization and Channel Coding*. en. Tech. rep. CCSDS 231.1-O-1. 2015, p. 39.

[83] E. Paolini, M. Chiani, and M. P. C. Fossorier. "Degree Distribution Design for LDPC Codes: A Derivative Matching Approach". In: *IEEE Transactions on Communications* 59.11 (Nov. 2011), pp. 3007–3015. DOI: 10.1109/TCOMM.2011.081111.100399.

[84] H. Chen and Z. Cao. "A Modified PEG Algorithm for Construction of LDPC Codes with Strictly Concentrated Check-Node Degree Distributions". In: *2007 IEEE Wireless Communications and Networking Conference*. Mar. 2007, pp. 564–568. DOI: 10.1109/WCNC.2007.109.

[85] Bruno Clerckx and Claude Oestges. *MIMO wireless networks: channels, techniques and standards for multi-antenna, multi-user and multi-cell systems*. 2nd ed. Oxford ; Waltham , MA: Academic Press, 2013.

[86] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth. "On the LambertW function". en. In: *Advances in Computational Mathematics* 5.1 (Dec. 1996), pp. 329–359. DOI: 10.1007/BF02124750.

[87] Rajendra K. Jain, Dah-Ming W. Chiu, and William R. Hawe. *A Quantitative Measure of Fairness and Discrimination for Resource Allocation in Shared Computer Systems*. Tech. rep. DEC-TR-301. Digital Equipment Corporation, Sept. 1984.

[88] H. SHI, R. V. Prasad, E. Onur, and I. G. M. M. Niemegeers. "Fairness in Wireless Networks:Issues, Measures and Challenges". In: *IEEE Communications Surveys Tutorials* 16.1 (2014), pp. 5–24. DOI: 10.1109/SURV.2013.050113.00015.

[89] M. Pischella and J. Belfiore. "Weighted Sum Throughput Maximization in Multicell OFDMA Networks". In: *IEEE Transactions on Vehicular Technology* 59.2 (Feb. 2010), pp. 896–905. DOI: 10.1109/TVT.2009.2036268.

[90] K. Pourtahmasi Roshandeh, M. Ardakani, and C. Tellambura. "Exact Solutions for Certain Weighted Sum-Rate and Common-Rate Maximization Problems". In: *IEEE Communications Letters* 22.5 (May 2018), pp. 1026–1029. DOI: 10.1109/LCOMM.2018.2808182.

[91] Félix Brah, Abdellatif Zaidi, Jérôme Louveaux, and Luc Vandendorpe. "On the Lambert-W function for constrained resource allocation in cooperative networks". In: *EURASIP Journal on Wireless Communications and Networking* 2011.1 (June 2011), p. 19. DOI: 10.1186/1687-1499-2011-19.

[92] F. Brah and L. Vandendorpe. "Constrained resource allocation for OFDMA wireless mesh networks with limited feedback". In: *2010 Future Network Mobile Summit*. June 2010, pp. 1–8.

[93] C. S. Chabalala and F. Takawira. "Adaptive spectrum aggregation for opportunistic resource management in multichannel networks". In: *2017 IEEE AFRICON*. Sept. 2017, pp. 161–166. DOI: 10.1109/AFRCON.2017.8095474.

[94] C. S. Chabalala and F. Takawira. "Hybrid Channel Assembling and Power Allocation for Multichannel Spectrum Sharing Wireless Networks". In: *2017 IEEE Wireless Communications and Networking Conference (WCNC)*. Mar. 2017, pp. 1–6. DOI: 10.1109/WCNC.2017.7925891.

[95] S. Toumpis, R. Muller, and J. Sayir. "On the transport capacity of a multiple access Gaussian channel". In: *International Workshop on Wireless Ad-Hoc Networks, 2004.* May 2004, pp. 191–195. DOI: 10.1109/IWWAN.2004.1525569.

Appendices

A

MAP Equations for SCMA

Using Bayes' theorem, the joint probability of the transmited symbols $\mathbf{x}_1, \ldots, \mathbf{x}_J$, given the observed signal $\mathbf{y}$ at the receiver, is written as:

$$p(\mathbf{x}_1, \ldots, \mathbf{x}_J|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x}_1, \ldots, \mathbf{x}_J)\, p(\mathbf{x}_1, \ldots, \mathbf{x}_J) \tag{A.1a}$$

$$p(\mathbf{x}_1, \ldots, \mathbf{x}_J|\mathbf{y}) = p(\mathbf{y}|\mathbf{x}_1, \ldots, \mathbf{x}_J) \prod_{j=1}^{J} p(\mathbf{x}_j) \tag{A.1b}$$

where equality holds in the assumption that $\mathbf{x}_i$ is independent of $\mathbf{x}_j$ $\forall i \neq j$.

The probability of the observed vector $\mathbf{y}$, given knowledge of the transmitted signals, is written as

$$p(\mathbf{y}|\mathbf{x}_1, \ldots, \mathbf{x}_J) = p(\mathbf{z}) \tag{A.2a}$$

$$= \prod_{k=1}^{K} p(z_k|\mathbf{x}_1, \ldots, \mathbf{x}_J) \tag{A.2b}$$

where equality holds for independent noise entries in $\mathbf{z}$.

The noise entries z_k are assumed to be i.i.d. complex-valued random variables distributed as $\mathcal{CN}\left(0, \sigma^2\right)$. Thus, we write its PDF as

$$p(z_k|\mathbf{x}_1, \ldots, \mathbf{x}_J) = \frac{1}{\pi\sigma_z^2} \exp\left(-\frac{1}{2\sigma_z^2} \left| y_k - \sum_{j \in \phi_k} \mathbf{h}_j[k]\mathbf{x}_j[k] \right|^2 \right). \tag{A.3}$$

Using (A.3) in (A.2b), we obtain

$$p\left(\mathbf{y}|\mathbf{x}_1,\ldots,\mathbf{x}_J\right) = \prod_{k=1}^{K}\left[\frac{1}{\pi\sigma_z^2}\exp\left(-\frac{1}{2\sigma_z^2}\left|y_k - \sum_{j\in\phi_k}\mathbf{h}_j[k]\mathbf{x}_j[k]\right|^2\right)\right]. \tag{A.4}$$

Finally, using the *maximum a posteriori* rule, we can write the jointly-estimated multidimensional symbols $\mathbf{x}_1^\star,\ldots,\mathbf{x}_J^\star$ as

$$\mathbf{x}_1^\star,\ldots,\mathbf{x}_J^\star = \arg\max_{\mathbf{x}_1,\ldots,\mathbf{x}_J} p\left(\mathbf{x}_1,\ldots,\mathbf{x}_J|\mathbf{y}\right) \tag{A.5a}$$

$$\propto \arg\max_{\mathbf{x}_1,\ldots,\mathbf{x}_J} p\left(\mathbf{y}|\mathbf{x}_1,\ldots,\mathbf{x}_J\right)\prod_{j=1}^{J} p\left(\mathbf{x}_j\right). \tag{A.5b}$$

It is possible to marginalize the MAP rule for the specific estimated symbol $\mathbf{x}_j^\star$ using the following operations:

$$p\left(\mathbf{x}_j\right) \propto \mathop{\mathrm{E}}_{\substack{\mathbf{x}_1,\ldots,\mathbf{x}_{j-1},\\ \mathbf{x}_{j+1},\ldots,\mathbf{x}_J}}\left[p\left(\mathbf{y}|\mathbf{x}_1,\ldots,\mathbf{x}_J\right)p\left(\mathbf{x}_j\right)\right] \tag{A.6a}$$

$$= \sum_{\mathbf{x}_1}\cdots\sum_{\mathbf{x}_{j-1}}\sum_{\mathbf{x}_{j+1}}\cdots\sum_{\mathbf{x}_J}\left[p\left(\mathbf{y}|\mathbf{x}_1,\ldots,\mathbf{x}_J\right)\prod_{j''=1}^{J} p\left(\mathbf{x}_{j''}\right)\right], \tag{A.6b}$$

where the summation $\sum_{\mathbf{x}_i}$ is a slight abuse of notation meaning that we will average $\mathbf{x}_i$ over all the M_i elements of $\mathcal{X}_i$. For the sake of simplicity, we will also introduce the notation

$$\sum_{\mathbf{X}\setminus\mathbf{x}_j} \triangleq \sum_{\mathbf{x}_1}\cdots\sum_{\mathbf{x}_{j-1}}\sum_{\mathbf{x}_{j+1}}\cdots\sum_{\mathbf{x}_J} \tag{A.7}$$

to improve the presentation of the following derivations.

Thus, we can write $\mathbf{x}_j^\star$ as it follows:

$$\mathbf{x}_j^\star = \arg\max_{\mathbf{x}_j} p\left(\mathbf{x}_j\right) \tag{A.8a}$$

$$\propto \arg\max_{\mathbf{x}_j}\sum_{\mathbf{X}\setminus\mathbf{x}_j}\left[p\left(\mathbf{y}|\mathbf{x}_1,\ldots,\mathbf{x}_J\right)\prod_{j''=1}^{J} p\left(\mathbf{x}_{j''}\right)\right] \tag{A.8b}$$

$$= \arg\max_{\mathbf{x}_j}\left(\pi\sigma_z^2\right)^{-K}\sum_{\mathbf{X}\setminus\mathbf{x}_j}\left[\exp\left(-\frac{1}{2\sigma_z^2}\sum_{k=1}^{K}\left|y_k - \sum_{j\in\phi_k}\mathbf{h}_j[k]\mathbf{x}_j[k]\right|^2\right)\prod_{j''=1}^{J} p\left(\mathbf{x}_{j''}\right)\right] \tag{A.8c}$$

$$\propto \arg\max_{\mathbf{x}_j} \sum_{\mathbf{X}\setminus\mathbf{x}_j} \left[\exp\left(-\frac{1}{2\sigma_z^2} \sum_{k=1}^{K} \left| y_k - \sum_{j\in\phi_k} \mathbf{h}_j[k]\mathbf{x}_j[k] \right|^2 \right) \prod_{j''=1}^{J} p\left(\mathbf{x}_{j''}\right) \right] \tag{A.8d}$$

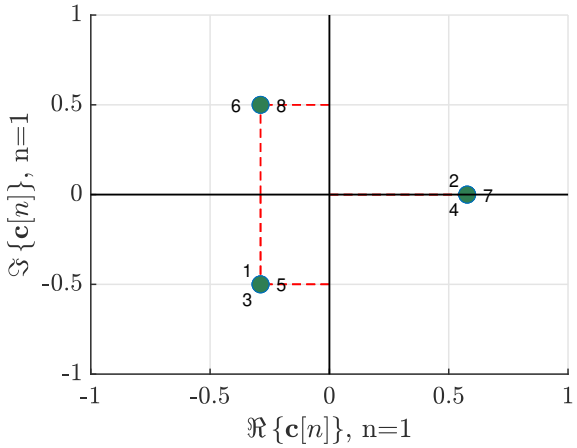
$$= \arg\max_{\mathbf{x}_j} \sum_{\mathbf{X}\setminus\mathbf{x}_j} \left[\exp\left(-\frac{1}{2\sigma_z^2} \sum_{k=1}^{K} \left| y_k - \sum_{j\in\phi_k} \mathbf{h}_j[k]\mathbf{x}_j[k] \right|^2 + \sum_{j''=1}^{J} \log p\left(\mathbf{x}_{j''}\right) \right) \right]. \tag{A.8e}$$

Equation (A.8e) is the basis for the message-passing algorithm equations. While the MAP rule for $\mathbf{x}_j^\star$ depends on the marginalization over $J-1$ interference-symbols, the MPA uses local estimations in each FN k by only marginalizing over $\phi_{k\setminus j}$, that has size $d_k - 1$ for some $d_k \ll J$.

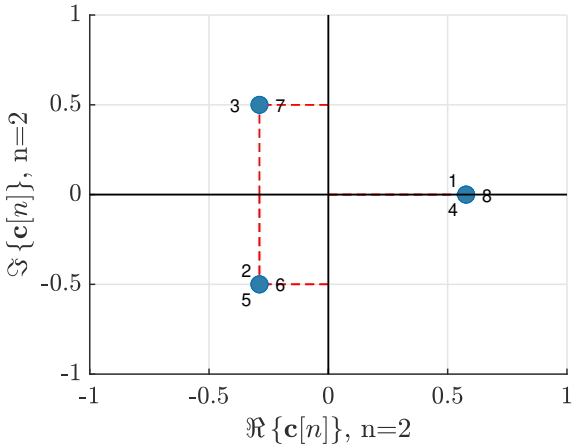
B

MDS Codebooks

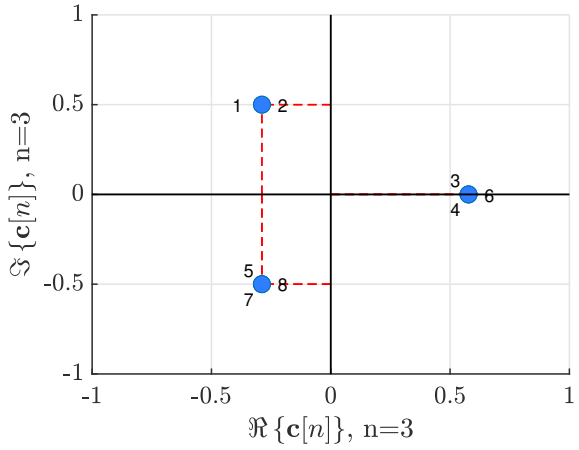
In this Appendix, we provide illustrations of the complex projections of the MDS codebooks in each of its complex-dimensions, along with the corresponding binary labeling (as an integer natural representation).



(a) $n = 1$.

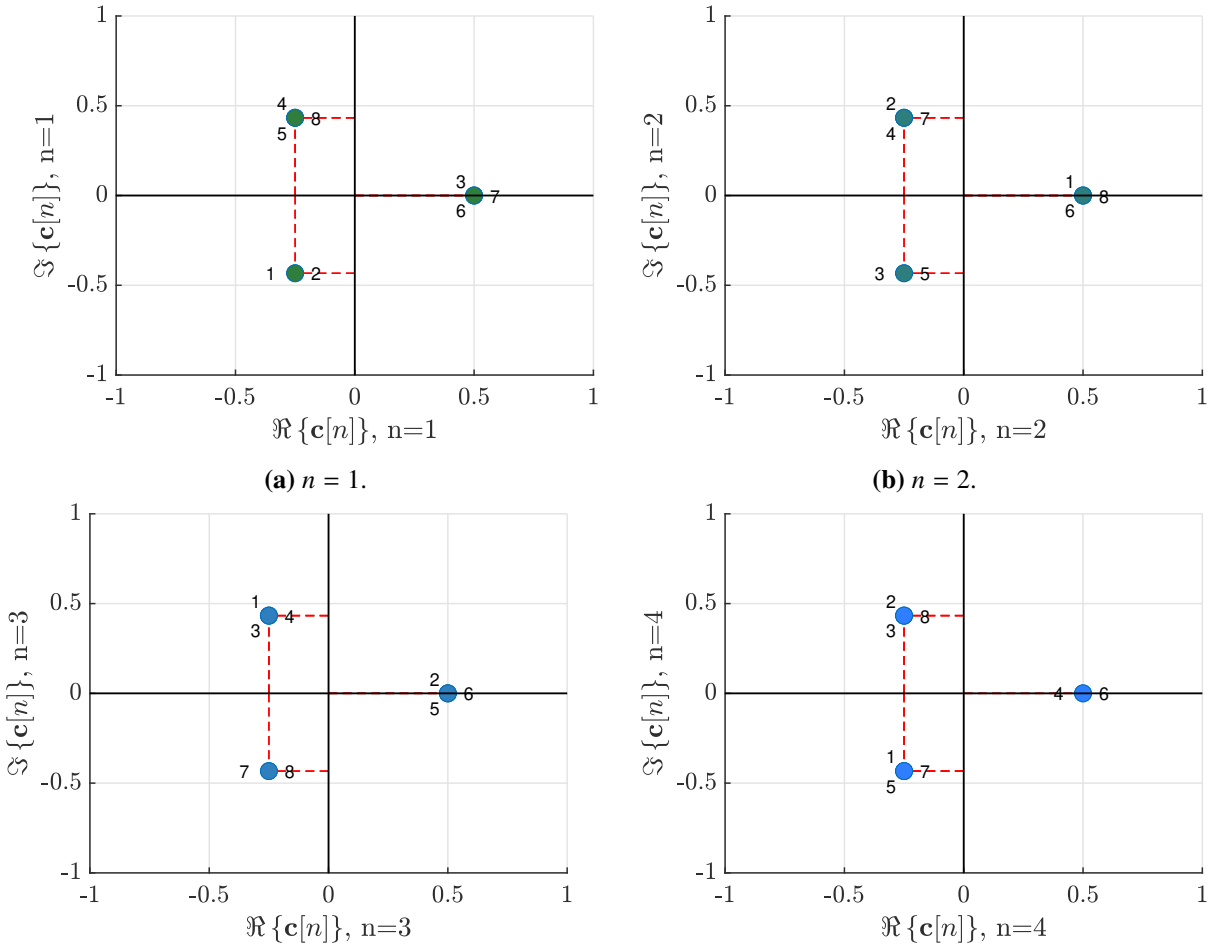


(b) $n = 2$.



(c) $n = 3$.

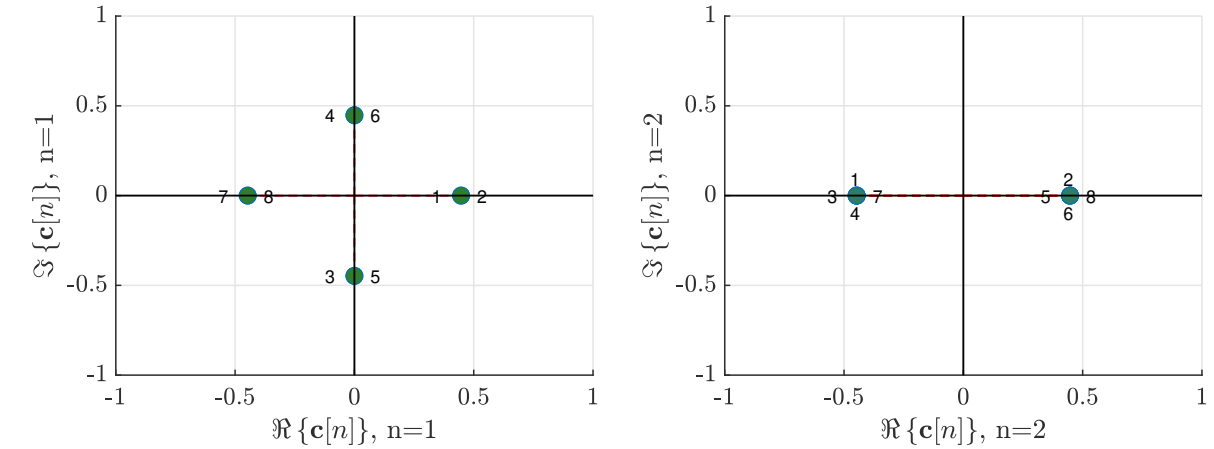
Figure B.1: Codebook $(M, N, q) = (8, 3, 3)$ based on a block code $(3, 2)$-GRS.



(a) $n = 1$. **(b)** $n = 2$.

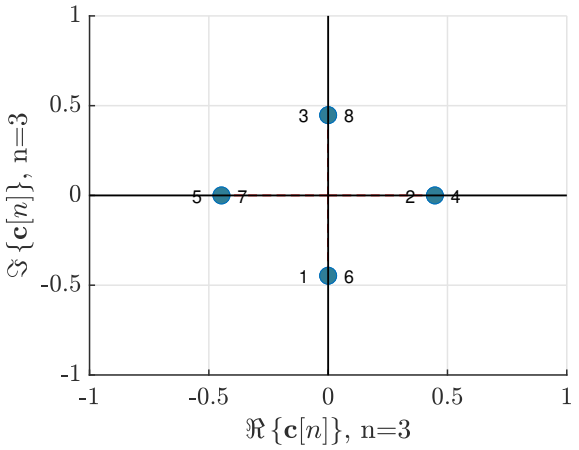
(c) $n = 3$. **(d)** $n = 4$.

Figure B.2: Codebook $(M, N, q) = (8, 4, 3)$ based on a block code $(4, 2)$-HC.

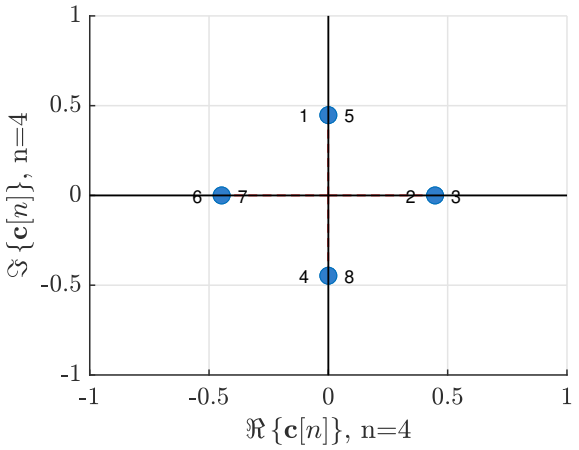


(a) $n = 1$.

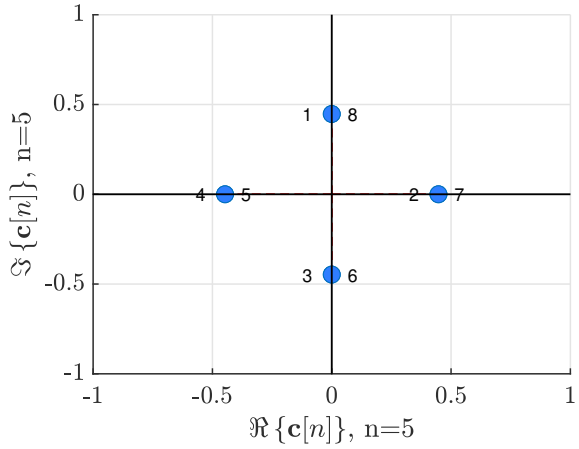
(b) $n = 2$.



(c) $n = 3$.

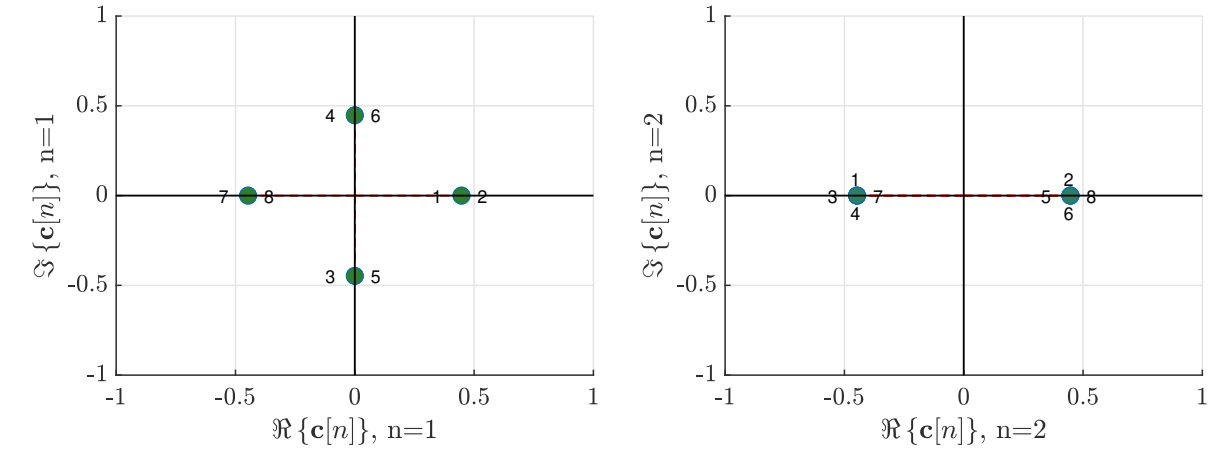


(d) $n = 4$.



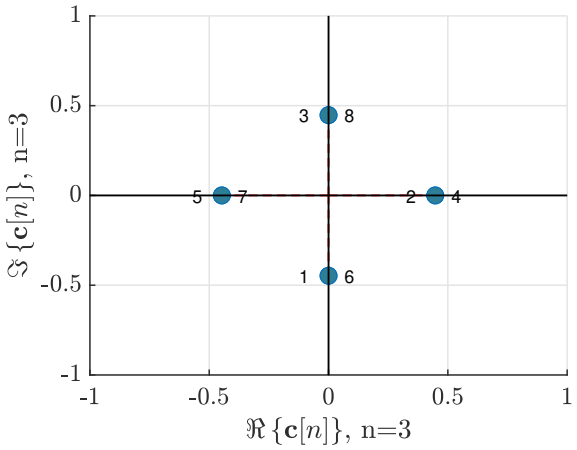
(e) $n = 5$.

Figure B.3: Codebook $(M, N, q) = (8, 5, 4)$ based on a block code $(5, 2)$-Ball.

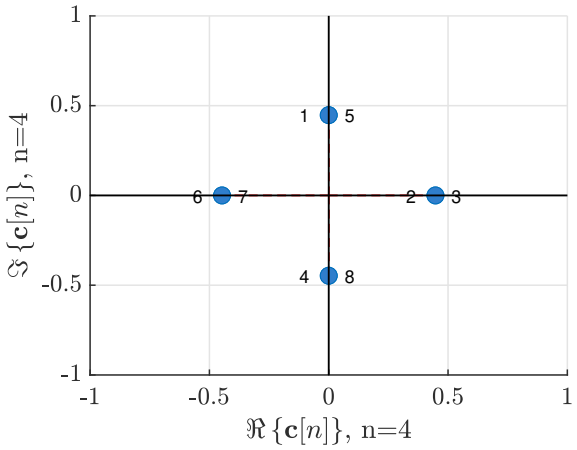


(a) $n = 1$.

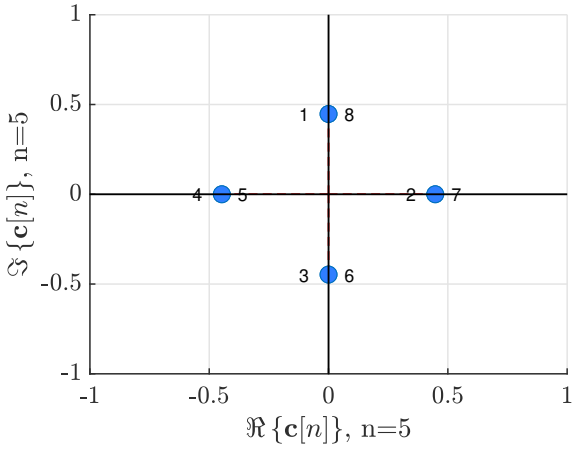
(b) $n = 2$.



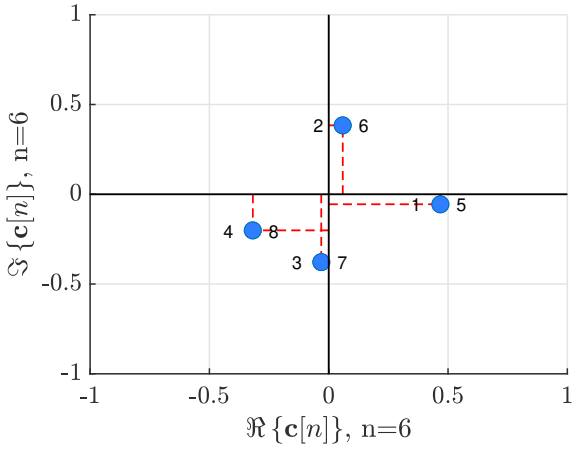
(c) $n = 3$.



(d) $n = 4$.

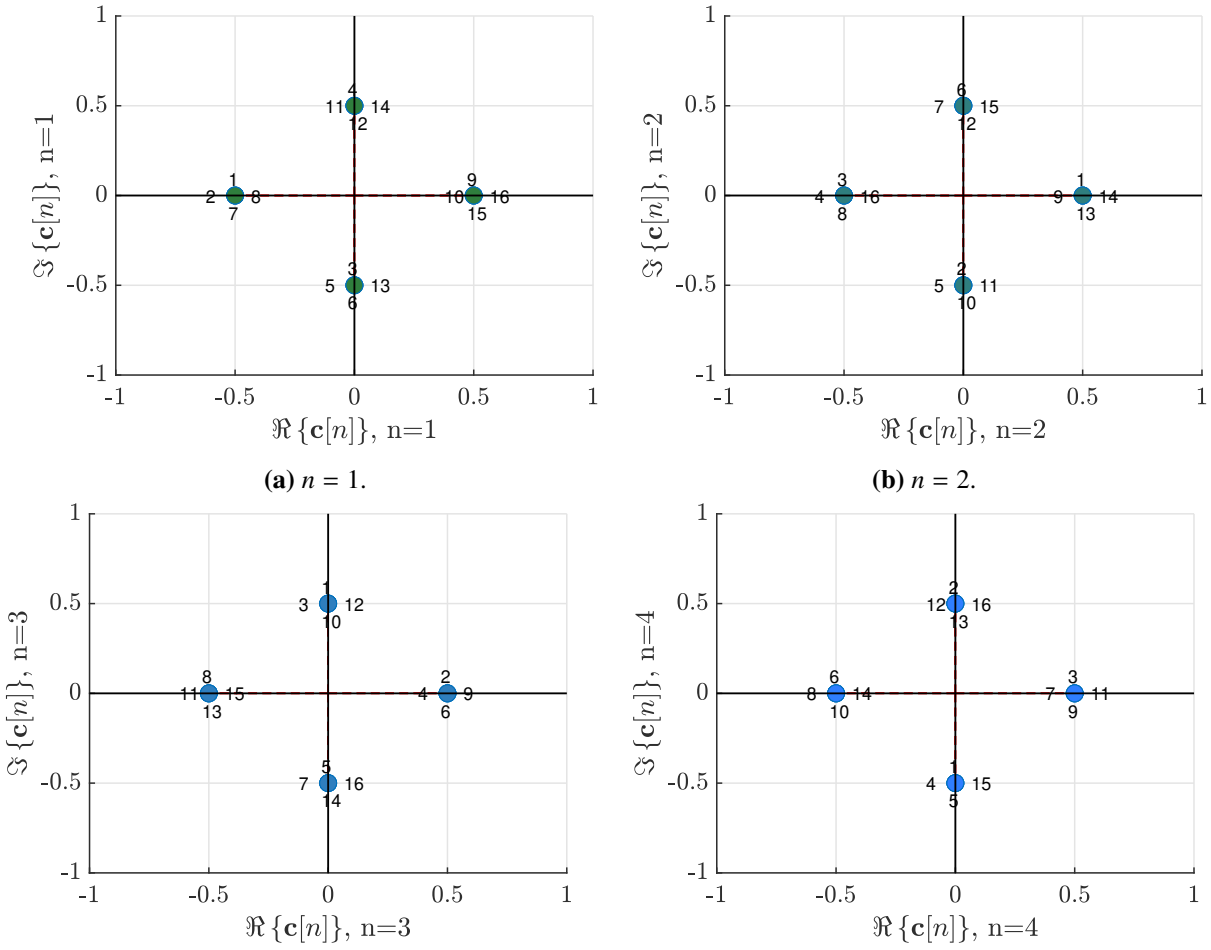


(e) $n = 5$.



(f) $n = 6$.

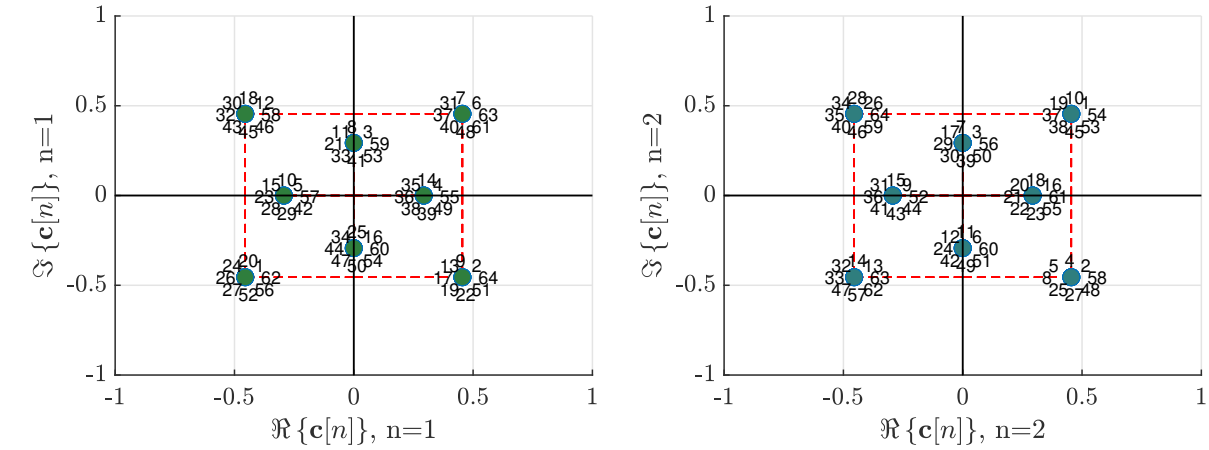
Figure B.4: Codebook $(M, N, q) = (8, 6, 5)$ based on a block code $(6, 2)$-Ball.



(a) $n = 1$. **(b)** $n = 2$.

(c) $n = 3$. **(d)** $n = 4$.

Figure B.5: Codebook $(M, N, q) = (16, 4, 4)$ based on a block code $(4, 2)$-GRS.



(a) $n = 1$. **(b)** $n = 2$.

(c) $n = 3$. **(d)** $n = 4$.

Figure B.6: Codebook $(M, N, q) = (64, 4, 8)$ based on a block code $(4, 2)$-GRS.

C

Resource Assignment Optimization for an OMA Free-Interference AWGN Channel

In this Appendix, we consider the problem of resource allocation in a AWGN channel OMA system, with constraints of limited number of the total number of resources.

We will solve two optimization problems for this system by relaxing the integer constraints and using convex optimization.

The solutions of these problems are taken as heuristics for the design of irregular SCMA in Chapter 5.

C.1 Equivalent OMA Free-Interference AWGN Channel

Consider that J users in a wide-area network have been select for an uplink OMA transmission in a AWGN OMA channel. Each user is located at a distance D_j of the base-station, and has a path-loss coefficient $\alpha_j \propto D_j^{-\rho}$ for some path-loss exponent $\rho \in \mathbb{R}$, $\rho > 1$, related to the large-scale path-loss model of this area.

The channel has $J \times \overline{N}_{\max}$ orthogonal resources to be assigned for the J selected users, such that each user will receive N_j resources. As a consequence, $\sum_{j=1}^{J} N_j \leq J\overline{N}_{\max}$ or equivalently $\overline{N} = \frac{1}{J}\sum_{j=1}^{J} N_j \leq \overline{N}_{\max}$.

The number of resource assigned to each user is also constrained to be within $N_{\min} \leq N_j \leq N_{\max}$. Remark that $N_j, N_{\min}$ and $N_{\max}$ must be integer values, although $\overline{N}_{\max}$ can be a continuous value as long as $J\overline{N}_{\max}$ is integer. Vector $\mathbf{N} \in \mathbb{N}^J$ concatenates the number of assigned resources for each user, i.e. the its entries are $N_1, \ldots, N_J$.

Each user has a total power of P that will be split uniformly over the N_j resources. In each of the N_j assigned resources, the users will transmit without interference, but the orthogonal channels have additive noise with power σ^2.

Thus, the SNR of each user is written as

$$\gamma_j = \frac{\alpha_j^{-1} P}{\sigma^2}. \tag{C.1}$$

Since user j is transmitting without interference over N_j orthogonal parallel AWGN channels, using a uniform power allocation strategy, its total channel capacity is written as:

$$C_j = N_j \log_2 \left(1 + \frac{\gamma_j}{N_j}\right). \tag{C.2}$$

C.2 MMF Resource Assignment

From a practical point of view, it is important that the system provides a good Quality of Service (QoS) for all users in the cell. In this sense, the minimum throughput among the users sharing the medium can be an important metric. We denote it as the max min-fairness (MMF) of the system.

C.2.1 The MMF Optimization Problem

We wish to design $\mathbf{N}$ to maximize $\min_j C_j$. The MMF problem is described as follows in the box below.

MMF Optimization Problem

$$\max_{\mathbf{N}} \min_{j} \quad C_j \tag{C.3a}$$

$$\text{subject to} \quad N_{\min} \leq N_j \leq N_{\max} \; \forall j, \tag{C.3b}$$

$$\frac{1}{J} \sum_{j=1}^{J} N_j \leq \overline{N}_{\max}. \tag{C.3c}$$

Constraint (C.3b) limits the maximum and minimum number of resources that can be assigned for each user.

The sum-constraint (C.3c) imposes a constraint on the maximum allowed average number of resources per users, meaning that we cannot increase all N_j boundlessly because of the finite-sized system limits.

Since these degrees N_j are integer numbers, problem C.3 is an integer programming problem,

of difficult tractability.

As a consequence of the constraints in (C.3), there will be limitations for the maximum achievable target data-rate, $\min_j C_j$. The value $\min_j C_j$ will be chosen as the largest value within a range $(C_{\mathrm{LB}}, C_{\mathrm{UB}})$ that can achieve equality in (C.3c).

Given some bottleneck value of C_j in this interval, the MMF is achieved if all users equalize to rate $C_j = r \in (C_{\mathrm{LB}}, C_{\mathrm{UB}})$. If we want all users to have the same rate $C_j = r \ \forall j$, we need to solve the following problem for each user:

$$r - N_j \log_2 \left(1 + \frac{\gamma_j}{N_j}\right) = 0 \tag{C.4}$$

This problem is equivalent to finding the intersection point of two functions dependent on N_j, i.e. $\frac{r}{N_j} = \log_2 \left(1 + \frac{\gamma_j}{N_j}\right)$.

C.2.2 Solution of the MMF Problem

To approach it, we relax the integer constraint and assume continuous variables $\mathbf{N} \in \mathbb{R}^J$. Thus, for each user, the problem has only one real solution, which is the root of Equation (C.4).

The details of the derivation are shown in Section C.2.3. Since $\mathbf{N}$ must have integer entries, we consider rounding and clipping, resulting in the following equation for N_j:

$$N_j = \left\lfloor -\frac{\gamma_j}{1 + \frac{\gamma_j}{c} W_{-1}\left(-\frac{c}{\gamma_j} e^{-\frac{c}{\gamma_j}}\right)} \right\rceil_{N_{\min}}^{N_{\max}} \tag{C.5}$$

where $W_{-1}(x)$ is the $k = -1$ branch of the Lambert W function [86], $c = \log(2) r$ and $\lfloor x \rceil_a^b = \min(b, \max(a, x))$ is a clipping function.

Solution (C.5) has the constraint $\frac{\gamma_j}{C_j} > \log(2)$, which relates to the *Shannon limit* of -1.59 dB of bit energy to noise power spectral density ratio. Thus, this ratio's constraint must be considered when using $W_{-1}(x)$ in (C.5).

We also consider the Jain's Fairness Index (FI) [87, 88] metric to evaluate the solution. Jain's FI is given as

$$\mathrm{FI}(x) = \frac{\bar{x}^2}{x_{\mathrm{rms}}^2}, \tag{C.6}$$

where $\bar{x}$ is the mean value of x and x_{rms} is the root mean square value of x.

In terms of the rate upper bounds from (C.2), when $x = C_j$, clearly the value of FI equals 1 for the case of $C_j = r \ \forall j$. When relaxing the integer constraint on $\mathbf{N} \in \mathbb{N}^J$ to continuous

values $\mathbf{N} \in \mathbb{R}^J$, without clipping (for positive N_j), it is possible to achieve the metric FI= 1 using Equation (C.5). However, integer N_j and clipping values of $N_{\min}$ and $N_{\max}$ may limit FI to be lower than 1. Nonetheless, even simple rounding can achieve values of FI very close to 1.

C.2.3 Proof of MMF Problem's Solution

In Equation (C.4), drop the user index j, for notation brevity, and let $x = N$ and $w = 1 + \frac{\gamma}{x}$. Then, we have

$$c = x \log\left(1 + \frac{\gamma}{x}\right) \tag{C.7a}$$

$$c = \frac{\gamma}{w-1} \log(w) \tag{C.7b}$$

$$\left(-\frac{c}{\gamma} w\right) e^{\left(-\frac{c}{\gamma} w\right)} = -\frac{c}{\gamma} e^{\left(-\frac{c}{\gamma}\right)} \tag{C.7c}$$

$$x = -\frac{\gamma}{1 + \frac{\gamma}{c} W\left(-\frac{c}{\gamma} e^{\left(-\frac{c}{\gamma}\right)}\right)} \tag{C.7d}$$

where step (C.7c) is solved applying Lambert W function $W(x)$ in both sides of the equality. Evaluating domain and image of $W(x)$ in (C.7d) we can conclude both that $k = -1$ branch of W_k is the one which simultaneously provides a valid solution for $x > 0$ and the root of (C.4), and also that $\frac{\gamma}{c} > 1$, which relates to the Shannon limit on E_b/N_0.

C.3 Weighted Sum-Capacity Resource Assignment

In multiple-access channels, the sum-capacity $C_{\text{sum}} = \sum_{j=1}^{J} C_j$ is used as a figure of merit for the total throughput that the channel can support. Although sum-capacity maximization is a common goal in RA problems, the Weighted Sum-Capacity (WSC) $C_{\text{W-sum}}$ maximization can also be considered as a figure of merit to deal with QoS or queue trade-offs with performance [89], among other applications [90].

The expression of $C_{\text{W-sum}}$ is given as

$$C_{\text{W-sum}} = \sum_{j=1}^{J} a_j C_j \tag{C.8}$$

where a_j is the weight of user j. Without loss of generality we assume that $\sum_{j=1}^{J} a_j = J$. Note that, in this OMA context, the sum capacity is a special case of the weighted-sum capacity, when

$a_1 = \ldots = a_J = 1$. We shall name the latter as the "uniform" weights case (WSCU).

It is important to mention that properly choosing the weights a_j helps to achieve different goals. For example, if $a_j = 1$ for all users, we have a solution which will benefit the users with larger SNR. On the other hand, $a_j \propto D_j^{\omega}$ (for some $\omega > 0$) leads to a fairness-based approach which prioritizes service to the farthest users. Any other criterion can be chosen for a_j, according to desired purposes.

C.3.1 The WSC Optimization Problem

The WSC optimization problem is written as follows in the box below.

Weighted Sum-Capacity Maximization

$$\max_{\mathbf{N}} \quad C_{\text{W-sum}} \tag{C.9a}$$

$$\text{subject to} \quad N_{\min} \leq N_j \leq N_{\max} \ \forall j, \tag{C.9b}$$

$$\frac{1}{J}\sum_{j=1}^{J} N_j \leq \overline{N}_{\max}. \tag{C.9c}$$

subject to the same constraints of Problem (C.3).

Since $\mathbf{N} \in \mathbb{N}^J$, (C.9) is an integer programming problem, with linear constraints. Hence, we relax again the integer constraint and assume continuous variables $\mathbf{N} \in \mathbb{R}^J$.

Under the relaxed conditions, it can be shown that the objective function $C_{\text{W-sum}}$ in Problem (C.9) is convex. As the constraints in (C.9) are linear, (C.9) becomes a convex optimization problem that has only one global optimal solution, which can be given by the Karush-Kuhn-Tucker (KKT) conditions.

The details of the derivation are shown in Section C.3.2. Since $\mathbf{N}$ must have integer entries, we consider again rounding and clipping, resulting in the final solution, given as:

$$N_j = \left\lfloor -\frac{\gamma_j}{1 + \frac{1}{W_0\left(-e^{-\left(\frac{\mu}{a_j}+1\right)}\right)}} \right\rfloor_{N_{\min}}^{N_{\max}} \tag{C.10}$$

where $W_0(x)$ is the $k = 0$ branch of the Lambert W function [86]. The constant μ in Equation

(C.10) is chosen in order to achieve equality in (C.9c). To choose μ, we use a root-finding numerical solution, e.g. the bisection method.

C.3.2 Solution of the max. WSC Problem

In this Appendix, we provide the derivation of Equation (C.10). This derivation is very similar to the work [91]. We follow from the relaxed problem in Section C.3.1, where the optimization variables were assumed to be continuous.

For simplicity, assume $\mathbf{N} = \mathbf{x}$ and define $\mathbf{x} = \begin{bmatrix} x_1 & \cdots & x_J \end{bmatrix}$, $f(\mathbf{x}) = -R_{j,\text{sum}}/\log(2)$ and the constraint inequalities on $\mathbf{x}$ as $g_j(\mathbf{x})$. From the derivatives of $f(\mathbf{x})$, one can show that it is a convex function.

From the problem described in Eq. (C.9), we write the Lagrangian function as:

$$\begin{aligned}\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = \underbrace{-\sum_{j=1}^{J} a_j x_j \log\left(1 + \frac{\gamma_j}{x_j}\right)}_{f(\mathbf{x})} + \sum_{j=1}^{J} \lambda_{1,j} \underbrace{\left(x_j - N_{\max}\right)}_{g_{1,j}(\mathbf{x}) \leq 0} \\ + \sum_{j=1}^{J} \lambda_{2,j} \underbrace{\left(-x_j + N_{\min}\right)}_{g_{2,j}(\mathbf{x}) \leq 0} + \lambda_3 \underbrace{\left(\sum_{j=1}^{J} x_j - J\overline{N}_{\max}\right)}_{g_3(\mathbf{x}) \leq 0}\end{aligned} \tag{C.11}$$

where $g_{1,j}(\mathbf{x}) = x_j - N_{\max}$, $g_{2,j}(\mathbf{x}) = -x_j + N_{\min}$ and $g_3(\mathbf{x}) = \sum_{j=1}^{J} x_j - J\overline{N}_{\max}$ are taken from the maximum, minimum and maximum average user REs constraints, respectively.

Solving for $\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = 0$ results in Equation (C.10) with the aid of Lambert W function $W_0(x)$ (an example of the use of $W(x)$ is given in Section C.2.3).

From the KKT conditions $\lambda_{1,j}, \lambda_{2,j}, \lambda_3 > 0 \ \forall j$, we conclude that $\mu = \lambda_3 + \lambda_{2,j} - \lambda_{1,j} > 0$. Other bounds for μ,

$$\text{LB} = \max_j a_j \left(\frac{1}{1 + \frac{\gamma_j}{N_{\max}}} - \left[1 + \log\left(\frac{1}{1 + \frac{\gamma_j}{N_{\max}}} \right) \right] \right), \tag{C.12a}$$

$$\text{UB}_1 = \min_j a_j \left(\frac{1}{1 + \frac{\gamma_j}{N_{\min}}} - \left[1 + \log\left(\frac{1}{1 + \frac{\gamma_j}{N_{\min}}} \right) \right] \right), \tag{C.12b}$$

come from the solution of KKT conditions $g_{i,j}(\mathbf{x}) \leq 0 \ \forall j$ and $i = 1, 2$.

For $g_3(\mathbf{x})$ we must consider the slackness property on variables $\lambda_{1,j}$ and $\lambda_{2,j}$, leading to

$\mu = \lambda_3$. To deal with clipping limits, the lower bound (C.12a) and the upper bound (C.12b) can be relaxed choosing $N_{\max} = \infty$ and $N_{\min} = 0$, respectively. Therefore, the interval for λ_3 will be $(0, \infty)$.

At last, λ_3 is found within this interval using some root-finding algorithm for KKT condition $\lambda_3 g_3\left(\mathbf{x}^\star\right) = 0$ and $\lambda_3 > 0$, where $\mathbf{x}^\star$ is calculated using Equation (C.10) and clipped at the desired limits of $[N_{\min}, N_{\max}]$.

For the particular case where $a_j = a \ \forall j$, it is possible to restrict λ_3 with an upper-bound given by:

$$\mathrm{UB}_2 = a\left(\frac{1}{1+\frac{\sum_{j=1}^{J}\gamma_j}{J\bar{N}_{\max}}} - \left\lceil 1+\log\left(\frac{1}{1+\frac{\sum_{j=1}^{J}\gamma_j}{J\bar{N}_{\max}}}\right)\right\rceil\right). \tag{C.13}$$

C.4 Numerical Results

Some numerical results of the strategies proposed in Sections C.2 and C.3 are presented here. Table C.1 shows the setup of simulation parameters.

Table C.1: Fixed simulations parameters.

Parameter	Value	Meaning
P	20 dBm	Tx power
Δ_f	15 kHz	subcarrier bandwidth
N_0	-174 dBm/Hz	noise power spectral density
Path Loss Model	$128.1 + 37.6\log_{10}(D[\mathrm{km}])$;	LTE Urban
$[N_{\min}, N_{\max}]$	$[1, 8]$	min. and max. bounds for codebook dimensions
$\bar{N}_{\max}$	4.0	maximum average VN degree
K	32	num. of shared resources
J	48	num. of users
O	150%	load factor of SCMA
D_j	$(0.25, 1.75)$ [km]	distance range of users from base station, D_j^2 uniformly spaced

The results obtained for the AWGN capacities of each strategy are shown in Table C.2. The yellow highlight emphasize the best results of each column considering regular N_j, while blue highlight emphasize the best result of irregular strategy for each column. The acronyms WSCF (fair-based) and WSCO (opportunistic) are used for the WSC strategies with weights defined as $a_j \propto D_j^{3.76}$ and $a_j \propto D_j^{-3.76}$, respectively.

It can be observed that irregular assignment outperforms regular assignment in each strategy (e.g., MMF outperforms in $\min_j C_j$ and FI, while WSCU outperforms in sum-throughput and

Table C.2: AWGN capacity analysis of each strategy.

N_j	$\sum a_j C_j$	$\sum C_j$	$\min_j C_j$	$\max_j C_j$	$\mathrm{FI}(C_j)$
1	F: 284.39 O: 668.62	362.61	5.03	15.54	0.91
2	F: 475.40 O: 1241.31	630.73	8.14	29.08	0.89
3	F: 632.69 O: 1777.85	864.11	10.59	41.86	0.87
4	F: 768.90 O: 2290.94	1075.39	12.62	54.16	0.85
5	F: 889.99 O: 2786.60	1270.57	14.36	66.09	0.84
1 – 8 MMF	422.10	880.26	13.63	23.04	0.99
1 – 8 WSCF	926.95	751.89	7.37	22.89	0.89
1 – 8 WSCU	1283.47	1283.47	5.03	100.31	0.54
1 – 8 WSCO	4176.79	1273.66	5.03	100.31	0.51

$\max_j C_j$). Note that the only regular assignment that outperforms the irregular ones is $N_j = 5$ $\forall j$, which needs 48 extra resources in comparison to MMF or WSCU.

In this simulations the distributions of N_j for MMF is

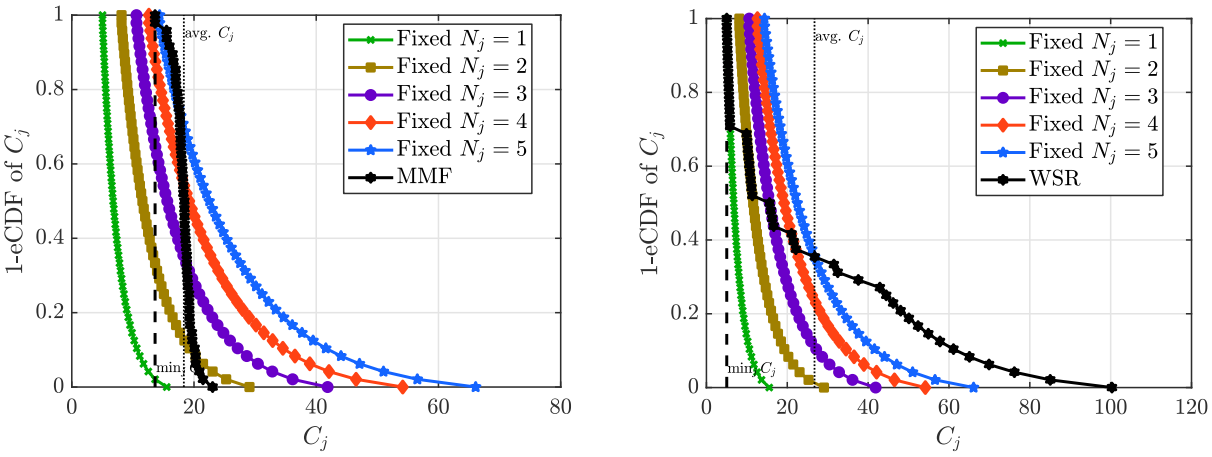
- $N_j = 1$: users 1, 2
- $N_j = 2$: users 3, 4, 5, 6, 7, 8, 9, 10, 11, 12
- $N_j = 3$: users 13, 14, 15, 16, 17, 18, 19, 20, 21, 22
- $N_j = 4$: users 23, 24, 25, 26, 27, 28, 29, 30, 31
- $N_j = 5$: users 32, 33, 34, 35, 36, 37
- $N_j = 6$: users 38, 39, 40, 41, 42
- $N_j = 7$: users 43, 44, 45, 46
- $N_j = 8$: users 47, 48

and for WSCU

- $N_j = 1$: users 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48,
- $N_j = 2$: users 26, 27, 28, 29, 30, 31, 32, 33, 34,
- $N_j = 3$: users 22, 23, 24, 25,

- $N_j = 4$: users 19, 20, 21,
- $N_j = 5$: users 18,
- $N_j = 6$: users 16, 17,
- $N_j = 7$: users 15,
- $N_j = 8$: users 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14,

Figures C.1a and C.1b illustrate the empirical probability that C_j is greater than a certain value, based on the distribution of C_j obtained with regular and irregular designs. It can be noted that MMF provides more fairness due to the vertical slope of the curve, where the minimum capacity (at the top of the Figure) outperforms the eCDF of regular designs.



(a) MMF strategy.

(b) WSCU strategy.

Figure C.1: Curves of 1−eCDF of C_j.

C.5 On the Use of the Lambert W Function

The Lambert W function has been used in other RA problems. The definition of the Lambert W function $W(x)$ represents the solutions y of the equation

$$ye^y = x, \tag{C.14}$$

where x can be any complex number. Thus, $y = f^{-1}(ye^y) = W(ye^y) = W(x)$.

If x is a real number, the domain of the Lambert W function is limited to $\left(-e^{-1}, \infty\right)$. Its image is divided in two branches: $k = -1$ for $x \in \left(-e^{-1}, 0\right)$ and $k = 0$ for $x \geq 0$. The illustration

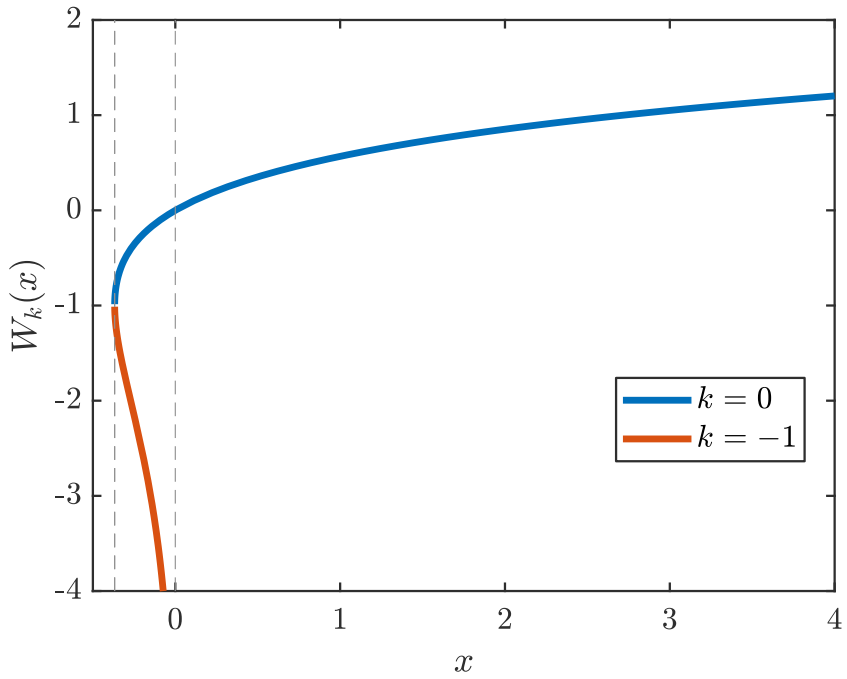


Figure C.2: Lambert W function domain and image: two main branches.

of the function's domain and image is shown in Figure C.2. Note that, when $x = -e^{-1}$, the function is still defined where the two branches meet, and its value is $y = -1$.

In the works [91, 92], we can find similar solutions using it in order to solve constrained RA problems in OFDMA mesh networks. In the works [93, 94], the authors employ the function while solving power allocation and channel selection problems, with the goal of node capacity maximization, in the context of cognitive radio systems.

In another application, the work of [95] has applied the Lambert W function in the context of the transport capacity, a weighted function of the channel capacity.