



INSTITUTO
UNIVERSITÁRIO
DE LISBOA

Design and Integration of novel Transmission Techniques for Coverage, Power Consumption and Data Rate Improvements in Wireless Communication Networks

Vasco Rafael Jerónimo Velez

PhD in Information Science and Technology

Supervisors:

Doctor Nuno Manuel Branco Souto, Associate Professor (with Aggregation),
Iscte - Instituto Universitário de Lisboa

Doctor Pedro Joaquim Amaro Sebastião, Assistant Professor,
Iscte - Instituto Universitário de Lisboa

September 2023

Department of Information Science and Technology

Design and Integration of novel Transmission Techniques for Coverage, Power Consumption and Data Rate Improvements in Wireless Communication Networks

Vasco Rafael Jerónimo Velez

PhD in Information Science and Technology

Supervisors:

Doctor Nuno Manuel Branco Souto, Associate Professor (with Aggregation),
Iscte - Instituto Universitário de Lisboa

Doctor Pedro Joaquim Amaro Sebastião, Assistant Professor,
Iscte - Instituto Universitário de Lisboa

September 2023

Department of Information Science and Technology

Design and Integration of novel Transmission Techniques for Coverage, Power Consumption and Data Rate Improvements in Wireless Communication Networks

Vasco Rafael Jerónimo Velez

PhD in Information Science and Technology

Jury:

Doctor Américo Manuel Carapeto Correia, Full Professor, ISCTE - Instituto Universitário de Lisboa (President)

Doctor Rodolfo Alexandre Duarte Oliveira, Associate Professor (with Aggregation), Faculdade de Ciências e Tecnologias da Universidade de Lisboa

Doctor Pedro Manuel de Almeida Carvalho Vieira, Assistant Professor, Instituto Superior de Engenharia de Lisboa - ISEL

Doctor João Carlos Marques Silva, Assistant Professor, ISCTE - Instituto Universitário de Lisboa

Doctor Nuno Manuel Branco Souto, Associate Professor (with Aggregation), ISCTE - Instituto Universitário de Lisboa

September 2023

Acknowledgments

I am grateful to Professors Nuno Souto and Pedro Sebastião, my supervisors, for challenging me and providing guidance throughout my doctoral studies. Their knowledge and invaluable support enabled me to successfully complete my doctorate.

I would like to express my gratitude towards my mother and sister, who have consistently provided their unwavering support throughout this process.

Additionally, I would like to extend my thanks to Professor Américo Correia for his availability, patience, and invaluable expertise, which played a crucial role in the successful completion of this project.

Finally, I would like to acknowledge my colleagues and dear friend João Pavia for making this journey both enjoyable and fulfilling.

I would also like to express my gratitude to Instituto de Telecomunicações (IT), Iscte-IUL branch, for providing the necessary equipment for this production.

I extend a heartfelt "Thank you" to all those mentioned and appreciate your concern for me.

Agradecimentos

Agradeço aos Professores Nuno Souto e Pedro Sebastião, meus orientadores, por me terem desafiado e orientado ao longo do meu doutoramento. O seu conhecimento e apoio inestimável permitiram-me concluir com sucesso o meu doutoramento.

Gostaria de exprimir a minha gratidão à minha mãe e à minha irmã, que sempre me apoiaram inabalavelmente ao longo deste processo.

Gostaria ainda de agradecer ao Professor Américo Correia pela sua disponibilidade, paciência e conhecimentos inestimáveis, que desempenharam um papel crucial na conclusão bem sucedida deste projeto.

Por fim, gostaria de agradecer aos meus colegas e ao meu grande amigo João Pavia por terem tornado este percurso tão agradável e gratificante.

Gostaria também de expressar a minha gratidão ao Instituto de Telecomunicações (IT), pólo do Iscte-IUL, por ter disponibilizado o equipamento necessário para esta produção.

A todas as pessoas mencionadas o meu sincero "Obrigado" e agradeço a vossa preocupação para comigo.

Abstract

Since the number of active electronic devices connected to the network is increasing every year, finding energy-efficient and high-capacity solutions has become vital. With the constant evolution of wireless network technologies, regular updates have become a standard practice. It is crucial to employ relevant methods and technologies, even if they are still in the developmental or emerging stage in scientific community or the market.

The purpose of this study is to suggest effective solutions to enhance spectral and energy efficiency, with a focus on large scale Multiple Input Multiple Output (MIMO) approaches. With this aim, a scheme was proposed which enhances performance and reduces energy consumption by using Index Modulations (IM), which convey additional information bits mapped onto the indexes of the active resources. A system-level and physical study has demonstrated that the approach is capable of surpassing previous schemes with higher data transfer rates. Additionally, this thesis considered the use of a new type of reconfigurable surface which is capable of redirecting and reflecting electromagnetic waves, resulting in improved user reception. This innovative technology is known as Reconfigurable Intelligent Surfaces (RIS).

The main contributions of the thesis consist of proposing a new technique called Precoding-aided Transmitter side Generalized Space-Frequency Index Modulation (PT-GSFIM), that can select resources, specifically antennas and sub-carriers in an Multi User – Multiple Input Multiple Output (MU-MIMO) system. Additionally, the study includes an evaluation of Cloud Radio Access Network (C-RAN) deployment for fifth generation (5G) and Beyond 5G (B5G) radio networks based on PT-GSFIM designs for MIMO downlink transmissions; Finally, the research evaluates a RIS system's effectiveness in serving as a primary technology for future 6G networks operating within the Millimeter-Wave (mmWave) and Terahertz (THz) bands.

The initial contribution puts forward a precoding based approach combined with GSFIM transmissions. PT-GSFIM selects active antennas and subcarriers which carry amplitude and phase modulated symbols using additional information bits. This approach also uses Complex Rotation Matrices (CRMs) as a Signal Space Diversity (SSD) technique to take advantage of frequency diversity. The scheme was purpose-built for MU-MIMO systems, which uses a precoder to eliminate interference between users. Additionally, the study proposes three robust detection algorithms on receiver side, for the signals in this GSFIM schemes.

Continuing on the same topic, a system-level evaluation is proposed, using a 5G/B5G based C-RAN deployment, employing the GSFIM-based approach and the proposed detectors in a MU-MIMO system. Three scenarios were considered (two outdoor Urban Macro e Urban Micro and one Indoor) for 5G New Radio (NR) utilizing the frequency bands of 3.5 GHz and 28 GHz. The throughput values demonstrate superiority over traditional MU-MIMO schemes and generalized spatial modulations (GSM).

In the last contribution, a RIS-based large-scale MIMO system was presented and assessed as a potential solution for future post-5G or sixth Generation (6G) wireless communications. Specifically, an iterative algorithm was implemented, which successfully integrates precoder with individual phase tuning of the RIS panel elements in a MIMO-OFDM link that operates at mmWave and THz frequencies. A system evaluation was conducted in two scenarios, one based on Urban Micro Truncated (UMT) and the other on Indoor Open Office (IOO). The results illustrate enhancements in throughput values and coverage area.

These proposals are relevant to the problems mentioned above, providing potentially effective solutions in terms of spectral efficiency (SE) and energy efficiency (EE).

Keywords: MIMO, PT-GSFIM, RIS, 5G, 6G, SE e EE

Resumo

Uma vez que o número de dispositivos eletrônicos ativos aumenta todos os anos, tornou-se vital encontrar soluções eficientes em termos energéticos e de elevada capacidade. Com a constante evolução das tecnologias de redes sem fios, as atualizações regulares tornaram-se uma prática corrente. É crucial empregar métodos e tecnologias relevantes, mesmo que ainda estejam em fase de desenvolvimento ou emergentes na comunidade científica ou no mercado.

O objetivo deste estudo é sugerir soluções eficazes para melhorar a eficiência espectral e energética, centrando-se em abordagens baseadas em MIMO (Multiple Input Multiple Output) em grande escala. Com este objetivo, foi proposto um esquema que melhora o desempenho e reduz o consumo de energia através da utilização de modulações de índice (IM), o que implica a inclusão de bits adicionais de dados para identificar recursos ativos. Um estudo físico e ao nível do sistema demonstrou que a abordagem é capaz de ultrapassar esquemas anteriores com taxas de transferência de dados mais elevadas. Além disso, esta tese considerou a utilização de um novo tipo de superfície reconfigurável que é capaz de redirecionar e refletir ondas eletromagnéticas, resultando numa melhor receção pelo utilizador. Esta tecnologia inovadora é conhecida como Superfícies Inteligentes Reconfiguráveis (RIS).

As principais contribuições da tese consistem em propor um esquema de Modulação por Índice de Espaço-Frequência Generalizado do lado do Transmissor Auxiliado por Pré-codificação (PT-GSFIM), que pode selecionar recursos, especificamente antenas e sub-portadoras em um sistema Multi User - Multiple Input Multiple Output (MU-MIMO). Além disso, o estudo inclui uma avaliação da implementação de redes de acesso via rádio na nuvem (C-RAN) para redes de rádio de quinta geração (5G) e para além do 5G (B5G) com base em conceções PT-GSFIM, para transmissões de downlink MIMO; por último, a investigação avalia a eficácia de um sistema RIS para servir como tecnologia principal para futuras redes 6G que operam nas bandas de ondas milimétricas (mmWave) e terahertz (THz).

A contribuição inicial apresenta uma abordagem baseada na pré-codificação combinada com transmissões GSFIM. Este esquema utiliza bits de informação especiais para selecionar antenas e sub-portadoras que transportam símbolos modulados em amplitude e fase. As matrizes de rotação complexas (CRM) são também utilizadas nesta abordagem, como técnicas de diversidade de espaço de sinal (SSD), a fim de tirar partido da diversidade de frequência. O esquema foi concebido para sistemas MU-MIMO, que utilizam um pré-codificador para eliminar a interferência entre utilizadores. Adicionalmente, o estudo propõe três algoritmos de deteção fiáveis para os sinais neste esquema GSFIM.

A segunda contribuição propõe uma avaliação ao nível do sistema utilizando uma implementação C-RAN baseada em 5G/B5G, empregando a abordagem assente em GSFIM e os detetores propostos num sistema MU-MIMO. Foram considerados três cenários (dois exteriores Urban Macro e Urban Micro e um Indoor) para 5G NR utilizando as bandas de 3,5 GHz e 28 GHz. Os valores de transferência de dados, demonstram superioridade em relação aos esquemas MU-MIMO tradicionais e às modulações espaciais generalizadas (GSM).

Na última contribuição, um sistema MIMO em larga escala baseado em RIS foi apresentado e avaliado quanto à sua eficácia em futuras comunicações sem fios pós-5G ou de sexta geração (6G). Especificamente, foi implementado um algoritmo iterativo, que integra com sucesso o pré-codificador com sintonização de fase individual dos elementos do painel RIS numa ligação MIMO-OFDM que opera em frequências mmWave e THz. Foi efetuada uma avaliação do sistema em dois cenários, um baseado em Urban Micro Truncated (UMT) e o outro em Indoor Open Office (IOO). Os resultados ilustram melhorias nos valores de débito e na área de cobertura.

Estas propostas são relevantes para os problemas acima mencionados, fornecendo soluções potencialmente eficazes em termos de eficiência espectral (SE) e eficiência energética (EE).

Palavras-chave: MIMO, PT-GSFIM, RIS, 5G, 6G, SE, EE.

Contents

Acknowledgments	i
Agradecimientos.....	iii
Abstract	v
Resumo	vii
Contents	ix
List of Figures.....	x
List of Tables	xiii
List of Acronyms	xiv
List of Symbols.....	xvii
Chapter 1 - Introduction.....	1
1.1. Motivation and Scope	1
1.2. Objectives and Goals.....	5
1.3. Research Contributions.....	7
1.4. Thesis Overview	9
1.4.1. Index Modulations.....	9
1.4.2. Millimeter-Waves and Terahertz Communications and Reconfigurable Intelligent Surfaces.....	13
Chapter 2 - A Generalized Space-Frequency Index Modulation Scheme for Downlink MIMO Transmissions With Improved Diversity	19
Chapter 3 - System-Level Assessment of a C-RAN based on Generalized Space–Frequency Index Modulation for 5G New Radio and Beyond.....	35
Chapter 4 - Performance Assessment of a RIS-Empowered Post-5G/6G Network Operating at the mmWave/THz Bands	63
Chapter 5 - Conclusion and Future Work	79
5.1. Main Conclusions	79
5.2. Future Work	81
References	83
Appendix A	87
Appendix B	109
Appendix C	141

List of Figures

Figure 1.1 - Summary of the goals and research.	6
Figure 1.2 - Typical GSM scheme of transmitter side [29].	9
Figure 1.3 - Transmitter and receiver scheme of GSM-MU-MIMO [28]	10
Figure 1.4 - Comparison of BER performance versus SNR between GSM-MU-MIMO and MU-MIMO [28]	10
Figure 1.5 - Example of GSFIM scheme using space and frequency simultaneously [30].....	11
Figure 1.6 - Working principle of C-RAN [36].....	12
Figure 1.7 - Design of a Hybrid precoding, both transmitter and receiver sides [37]	13
Figure 1.8 - Example of and hybrid precoding design, intended to a THz system [37].....	14
Figure 1.9 - Throughput vs number of users when using hybrid precoding systems into a C-RAN [38].	15
Figure 1.10 - Example of a RIS panel, with several tunable elements [39]	16
Figure 1.11 - RIS example in a room, showing an increase on receiver signal [23]	16
Figure 2.1 - System Model: PT-GSFIM transmitter (a) and receiver (b)..	23
Figure 2.2 - Example of the structure of each PT-GSFIM symbol in the space-frequency domain with $N_f=4, N_{af}=3, N_s=4, N_a=2$	24
Figure 2.3 - Comparison of CRM rotation angle ϕ of PT-GSFIM in a MU scenario with $N_u=4, N_{tx}=8N_u, N_f=4, N_{af}=3, N_s=8, N_a=1, N_{rx}=4$, QPSK (SNR=0dB).....	30
Figure 2.4 - BER performance of PT-GSFIM in a MU scenario with $N_u=4, N_{tx}=5N_u, N_f=4, N_{af}=3, N_s=5, N_a=2, N_{rx}=5$, QPSK (5.75 bpcu per user)..	30
Figure 2.5 - Complexity of the different GSFIM detector algorithms for the MU scenario with $N_u=4, N_{tx}=5N_u, N_f=4, N_{af}=3, N_s=5, N_a=2, N_{rx}=5$, QPSK (5.75 bpcu per user)... ..	31
Figure 2.6 - BER performance of PT-GSFIM in a MU scenario with $N_u=4, N_{tx}=5N_u, N_f=4, N_{af}=3, N_s=5, N_a=2, N_{rx}=5$, 64-QAM (11.75 bpcu per user).....	31
Figure 2.7 - BER performance of PT-GSFIM in a MU scenario with $N_{tx}=5N_u, N_f=4, N_{af}=3, N_s=5, N_a=2, N_{rx}=4$, QPSK (5.75 bpcu per user).....	31
Figure 2.8 - BER performance of PT-GSFIM in a MU scenario with $N_u=4, N_{tx}=8N_u, N_s=8, N_a=1, N_{rx}=4$, QPSK (3 bpcu per user).....	31
Figure 2.9 - BER performance of PT-GSFIM, PT-SDIM, PTSFIM and conventional BD MU-MIMO ($N_u=8$).....	32
Figure 3.1 - Transmitter and Receiver structure [25].....	40
Figure 3.2 - Urban micro scenario layout: Macro layer: ISD = 200 m = 3R; Micro layer: 3 micro TRPs per macro TRPs.....	42
Figure 3.3 - Indoor Office scenario layout.....	43
Figure 3.4 - Comparison of BER performances of PT-GSFIM, GSM and conventional BD MU-MIMO, with $N_u = 8$, and SE = 12 bpcu.....	48

Figure 3.5 - BLER vs (E_s/N_0) dB curves, for PT-GSFIM in a MU scenario with $N_{rx} = 5$, 64QAM (11.75 bpcu and 12.5 bpcu), 16QAM (11.9375 bpcu) and 4QAM (5.75 and 6.5 bpcu).....	49
Figure 3.6 - Throughput vs number of users for different scenarios with $N_{tx} = 5N_u$, $N_f = 4$, $N_{af} = 3$, $N_s = 5$, $N_a = 2$, $N_{rx} = 5$, 64-QAM (11.75 bpcu per user) and 4QAM (5.75 bpcu per user), C-RAN cluster size 1.....	50
Figure 3.7 - Throughput vs number of users for different scenarios with $N_{tx} = 5N_u$, $N_f = 4$, $N_{af} = 3$, $N_s = 5$, $N_a = 2$, $N_{rx} = 5$, 64-QAM (11.75 bpcu per user) and 4QAM (5.75 bpcu per user), C-RAN cluster size 3.....	51
Figure 3.8 - Throughput performance (bps) vs CDF of Throughput for UMa scenario with $N_{tx} = 5N_u$, $N_f = 4$, $N_{af} = 3$, $N_s = 5$, $N_a = 2$, $N_{rx} = 5$, 64QAM (11.75 bpcu per user) and 4QAM (5.75 bpcu per user).....	52
Figure 3.9 - CDF of TRP Throughput vs TRP throughput (bps) for UMa scenario with $N_{tx} = 5N_u$, $N_f = 4$, $N_{af} = 3$, $N_s = 5$, $N_a = 2$, $N_{rx} = 5$, 64-QAM (11.75 bpcu per user) and 4QAM (5.75 bpcu per user).....	53
Figure 3.10 - Throughput vs number of active TRP antennas, for UMa scenario, $N_{rx} = 5$, for two C-RAN cluster sizes, modulations and number of antennas per user, namely, with $N_{tx}/N_u = 5$, 64QAM(5), 4QAM(5), with $N_{tx}/N_u = 8$, 64QAM(8), 4QAM (8) and with $N_{tx}/N_u = 12$, 16QAM(12).....	54
Figure 3.11 - Throughput vs number of active TRP antennas, for UMi scenario, $N_{rx} = 5$, for two C-RAN cluster sizes, modulations and number of antennas per user, namely, with $N_{tx}/N_u = 5$, 64QAM(5), 4QAM(5), with $N_{tx}/N_u = 8$, 64QAM(8), 4QAM (8) and with $N_{tx}/N_u = 12$, 16QAM(12).....	55
Figure 3.12 - Throughput vs number of active TRP antennas, for InD scenario, $N_{rx} = 5$, for two C-RAN cluster sizes, modulations and number of antennas per user, namely, with $N_{tx}/N_u = 5$, 64QAM(5), 4QAM(5), with $N_{tx}/N_u = 8$, 64QAM(8), 4QAM (8) and with $N_{tx}/N_u = 12$, 16QAM(12).....	56
Figure 3.13 - Throughput vs number of active TRP antennas, RAN cluster size 3, for different systems and scenarios, with $N_{tx}/N_u = 12$, $N_{sc} = 256$, $N_a = 2$, $N_{rx} = 5$, 12bpcu, GSM and MU-MIMO with 64QAM and GSFIM with 16QAM $N_f = 16$, $N_{af} = 13$	57
Figure 3.14 - Coverage vs percentage of TRP transmitted power, for UMa scenario, with $N_{tx}/N_u = 8$, $N_{sc} = 256$, $N_{rx} = 5$, 64QAM with 12.5 bpcu and 4QAM with 6.5 bpcu.....	58
Figure 4.1 - Example of a MIMO-OFDM communication system aided by RIS, consisting of BS/AP, UE, and RIS panel with N_{RIS} reflecting elements.....	66
Figure 4.2 - IOO environment where black dots are represented as AP, and blue dots RIS panels. Users are shown in red dots that are dispersed in random positions across the floor. IOO scenario without RIS (left) and with RIS assistance (right).....	69
Figure 4.3 - UMT scenario, without RIS (top image), with RIS placed at the same distance from the nearest BS (middle image) and with RIS placed at different distances of BS (bottom one).....	70

Figure 4.4 - Example of RIS Panel with 3 sub-panels used in all simulated environments (left), and a hexagonal grid where 4 BS are located in the center (right).....	70
Figure 4.5 - Achievable rate versus iterations number obtained with the proposed AM-SVD-APG for a scenario with $f=28\text{GHz}$, $N_{tx}=64$, $N_s=3$, $N_{rx}=16$, and $N_c=1$	71
Figure 4.6 - BER Comparison versus horizontal distance for a scenario with $f=28\text{GHz}$, $N_{tx}=64$, $N_s=3$, $N_{rx}=16$, $N_{ris}=144$ or 576 elements with $N_c=1$, where RIS panel is placed at a fixed position of (20m, 8m) and the user position changes between the receiver (RX) and the transmitter (TX).....	72
Figure 4.7 - BER versus transmitted power when the receiver is fixed at (55 m, 0 m), with a configuration of $f=28\text{GHz}$, $N_{tx}=64$, $N_s=3$, $N_{rx}=16$ and $N_c=132$, but with $N_{ris}=144$ elements (top) or 576 (bottom).....	72
Figure 4.8 - Throughput achievable with RISs in the UMT scenario, when $N_s=1$, $N_{tx}=64$, $N_{rx}=16$ and $N_c=132$	73
Figure 4.9 - Coverage versus transmitted power with RISs in the UMT scenario, when $N_s=1$, $N_{tx}=64$, $N_{rx}=16$ and $N_c=132$	73
Figure 4.10 - Throughput, demonstrating how RIS can affect the rate, as a function of the number of served UEs, when $N_{tx}=64$, $N_s=3$, $N_{rx}=16$. RIS are located at the same distance from nearest BS.....	74
Figure 4.11 - Throughput vs number of users when $N_s=3$, $N_{rx}=16$, $N_c=132$. RIS are located at the same distance from the nearest BS.....	74
Figure 4.12 - Comparison of different achievable throughput using $N_s=3$, $N_{rx}=16$, $N_c=132$ in the UMT scenario shown on bottom image of figure 3.....	75
Figure 4.13 - Coverage of throughput using fixed parameters as $N_s=3$, $N_{rx}=16$, $N_c=132$, relative to the realistic UMT scenario presented in previous Figure 12.....	75
Figure 4.14 - Throughput, demonstrating how RIS can affect the rate, in the IOO environment using $N_{tx}=64$, $N_s=2$, $N_{rx}=16$	76

List of Tables

Table 1.1 - Table of publications	7
Table 2.1 - Algorithm 1: OB-MMSE-based GSFIM detector	26
Table 2.2 - Algorithm 2: multiblock sMMP detector	27
Table 2.3 - Algorithm 3: admm-based gsfim detector	27
Table 2.4 - Number of real flops for the different gsfim detectors.....	29
Table 3.1 - Algorithm based in ADMM used for GSFIM detection [25]	47
Table 3.2 - SE in bps/Hz of different scenarios and cluster size 1 and 3.....	57
Table 4.1 - Algorithm 1: AM-SVD-APG	68
Table 4.2 - Parameters of simulated scenarios	70

List of Acronyms

3GPP – 3rd Generation Partnership Project

5G – 5th Generation of Mobile Network

5G NR – 5G New Radio

6G – 6th Generation of Mobile Network

ADMM – Alternating Direction Method of Multipliers

AM – Alternating Maximization

ANACOM – Autoridade Nacional de Comunicações

AoSA – Arrays of SubArrays

APG – Accelerated Proximal Gradient

AP – Access Point

AR – Augmented Reality

B5G – Beyond 5G

BBU – Base Band Unit

BER – Bit Error Rate

BS – Base Station

CA – Co-Author

CDF – Cumulative Distribution Function

CP – Cyclic Prefix

C-RAN – Cloud-Radio Access Network

CRM – Complex Rotation Matrices

DAoSA – Dynamic Arrays of SubArrays

DPS – Double Phase Shifters

EE – Energy Efficiency

FA – First Author

FDD – Frequency Division Duplex

FFT – Fast Fourier Transform

Gbps – Gigabits per second

GHz – Gigahertz

GSFIM – Generalized Spatial-Frequency IM

GSM – Generalized Spatial Modulation

IFFT – Inverse Fast Fourier Transform

IM – Index Modulations
 InD – Indoor
 IOO – Indoor Open Offices
 IoT – Internet of Things
 LOS – Line of Sight
 MIMO – Multiple-Input-Multiple-Output
 MLD – Maximum Likelihood Detector
 mmWave – millimeter-Wave
 M-QAM – Multilevel Quadrature Amplitude modulation
 MU – Multi-User
 MUI – Multi-User Interference
 NLOS– Non Line Of Sight
 OB-MMSE – Ordered Block Minimum Mean-Squared Error
 OCRM – Orthonormal Complex Matrices
 OFDM – Orthogonal Frequency Division Multiplexing
 PT-GSFIM – Precoding-aided Transmitter side Generalized Space-Frequency Index Modulation
 RF – Radio Frequency Chain
 RIS – Reconfigurable Intelligent Surfaces
 RRH – Remote Radio Head
 SE – Spectral Efficiency
 SINR – Signal to Interference plus Noise Ratio
 SM – Spatial Modulation
 sMMP – Multipath Matching Pursuit with slicing
 SNR – Signal to Noise Ratio
 SPS – Single Phase Shifters
 SSD – Signal Space Diversity
 SU – Single-User
 SVD – Singular Value Decomposition
 Tbps – Terabits per second
 TDD – Time Division Duplex
 THz – Terahertz
 TRPs – Transmission and Reception Points

TTI – Transmission Time Interval

UMa – Urban Macro

UMi – Urban Micro

UM-MIMO – Ultra Massive MIMO

UMT – Urban Micro Truncated

WHO – World Health Organization

List of Symbols

A – Area of elements of RIS

$\mathbf{a}_s(\phi_l^{S \rightarrow R}, \theta_l^{S \rightarrow R})$ – Denote the RIS and transmitter array response vectors at the (azimuth, elevation) angles of $(\phi_l^{R \rightarrow S}, \theta_l^{R \rightarrow S})$

$\mathbf{a}_r(\phi_l^{R \leftarrow S}, \theta_l^{R \leftarrow S})$ – Steering vectors of RIS and transmitter array response vectors at the (azimuth, elevation) angles of $(\phi_l^{R \leftarrow S}, \theta_l^{R \leftarrow S})$

$\mathbf{a}_D(\phi_l^{D \leftarrow S}, \theta_l^{D \leftarrow S})$ – Steering vectors of receiver

β – Path loss

k – Subcarrier

\mathbf{x} – Transmitted signal

λ – Wavelength

G – Antenna gains

\mathbf{W} – Combiner matrix

\mathcal{A}_0 – Complex value of symbols

\mathcal{O} – Complexity order

∇ – Complex-valued gradient

τ – Delay of path

ρ – Denotes the power per stream and per subcarrier

\mathbf{U} – Left singular vectors

\mathbf{n}, N_0 – Noise level

ϕ – Phase shifts

P_n – Power of noise

\mathbf{F} – Precoder matrix

\mathbf{y} – Received signal

\mathbb{S}_0 – Set of possible sorts of \mathbf{s}

\mathbf{s} – Symbol

\mathbf{H} - Channel Matrix

$\mathbf{H}[k]^{S,R}$ – Channel matrix between BS/AP and RIS panel

$\mathbf{H}[k]^{R,D}$ – Channel matrix between RIS and user

$\mathbf{H}[k]^{s,D}$ – Channel matrix between BS/AP and user

$\mathbf{\Lambda}$ – Diagonal matrix

E_s – Energy of symbol

D_F – Fraunhofer distance

\mathbf{I} – Identity matrix

\mathbf{A} – Matrix of OCRM

Π – Projection over a set

\mathbf{V} – Right singular vectors

Δf – Subcarrier spacing

μs – Microseconds

D – Distance/Spacing

f – Frequency

f_c – Carrier frequency

L – Length of path

m^{th} – Phase shifts of RIS element

N_a – Number of active antennas

N_{af} – Number of active subcarriers

N_f, N_c – Number of subcarriers

N_{ris} – Number of RIS elements

N_{rx} – Number of reception antennas

N_{RF}^{tx} – Number of RF chains at transmitter

N_s – Number of symbols

N_{tx}, N_t – Number of transmitting antennas

N_u – Number of users

P_{user} – Power of allocated to user

u – user

P – Penalties parameter

φ – Rotation angle

Chapter 1 - Introduction

The first chapter introduces the main goals of research and the work developed including all necessary contributions to this research. It also introduces the motivation and goals and the main overview of this thesis research.

1.1. Motivation and Scope

With the advancement of technology, communications are forced to evolve to maintain the quality criteria demanded by modern society. We are currently facing challenges that require digital communications to be reliable and efficient in order to be able to interconnect all types of devices required by human beings on planet Earth. Since the first generations of mobile communications, we have gone from communications to serving people to people, to serving all kinds of existing connections of devices, whether these are objects, machines, smart connections, and more. These types of connections improve the way communications interact within society and how people and cities live. These types of digital tools are the pillars of the modern world we live in today. Seeing that services and applications become increasingly digital, such as 3D technology or mixed reality for example, they will consume large amounts of data during the communication process [1]. Since energy is not unlimited, it is vital that we make sure that future devices and their communications consume less energy making them more efficient.

While the number of electronic devices connected to the network increases year by year, most of them are not the traditional mobile devices, such as smartphones. As it is reported, future networks need to be prepared for the emergence of newer types of communication and connections or various types of devices, such as: smart watches, sensors, implants, etc [2]. As the world's mobile data traffic increases, wireless network upgrades are mandatory with each new generation. Over several decades, we have experienced the evolution of mobile communications over five generations, through new technologies, standards, marketing, services and others [1]. Managing this growth while improving the mobile user experience requires continued network evolution, especially in topics of efficiency and resource allocation. Since technology is also integrated into any type of device these days, both industry and the scientific community are striving to make systems more efficient with each generation iteration providing improvements over its predecessor [3].

According to the report in [4], show that the growth of the mobile market with the adoption of the newly implemented mobile digital transmission technologies, specifically the fifth generation (5G), is expected to reach a staggering 472 Exabytes of traffic generated per month by 2028. This type of mobile traffic generated, contains an interesting fact, the traffic does not depend only on the traditional cell phone market, but also on other services and digital applications that require a mobile connection, such as gaming services, streaming, IoT (Internet of Things) sensors, localization, communication, etc [2]. It is now possible to use 5G mobile communications to display statistical and augmented reality (AR) data

on multiple user terminals at a Formula 1 race in real time. This type of application allows the user to experience the driver's view during the race [5]. Such experiences require technology that is not yet ready to be scaled up to the macroscopic level across all types of scenarios and environments on the rest of the world. During the *Internationale Automobil-Ausstellung* Mobility 2023 conference [6], several companies showed great interest in mobility associated with digital services for cars, as well as the various autonomous driving systems, or even self-driving cars. As mentioned before, these vehicles require a lot of connectivity, since there are a large number of systems that require iteration between the servers and the vehicle [7] [8].

Although studies indicate that traffic generated, whether by voice or other traditional mobile services, is indeed decreasing, on contrast, internet data consumption on mobile devices has increased dramatically [9]. This type of behavior may be associated with the pandemic SARS-CoV-2 declared by the World Health Organization (WHO) on March 11, 2020, which forced the transformation of services and applications to a digital format [10]. As can be seen from the study carried out in [11], there was indeed a period during the lockdown when mobile traffic did decrease, but after this period, the volume of devices has increased to values pre-pandemic and still growing. Even though people used their mobile devices, most of them used the fixed/domestic wireless networks, WI-FI for example [11]. Much of the traffic generated belongs to the video streaming category, whether it is video calls, streaming, etc, which in telecommunications implies the transfer of large volumes of data rate. *Autoridade Nacional de Comunicações* (ANACOM), the operator responsible for managing mobile traffic in Portugal, reported there was an increase of more than 50% over the previous year, relating only to the consumption of mobile Internet and its services. [3]

Despite the efforts that academia and industry made with 5G, there are already efforts to modernize it, either to a newer sub-versions or bands (5G New radio - 5G NR), or to a new standalone generation, the sixth generation (6G) [12]. Although 6G is expected to be ready by 2030, we are entering into the beginning of the era of the Industry 4.0 society and smart connections [13] [14].

The latest generation already being commercialized - 5G, shows good benefits for the standards of today. Given this situation, the networks are already operating in frequency bands above 7-15 Gigahertz (GHz) such as the millimeter-wave (mmWave) band with frequencies up to 300GHz. This increase allows the spectrum to become less dense and allows to accommodate more users or more bandwidth [15]. While the 5G network can already achieve reliable low-delay communications, it is expected that this parameter will be further reduced in future 6G networks. This will allow cloud servers or computing servers to almost process the data they receive in a small fraction of time [13]. It is rumored that 6G will provide connections with a higher level of location accuracy than the previous generation and data rates on the order of terabits, reducing latency and power consumption. [13].

One critical aspect is the importance of the energy efficiency, since mobile devices are subject to high energy consumption when active. To have a long lifetime, it is necessary to achieve low power consumption during communications. Many of these devices are not always physically connected to the

power grid, so they rely on cloud-based computing services to complete their tasks. To enhance the energy efficiency (EE), researchers proposed solution to optimize the resources already available in Multiple-Input-Multiple-Output (MIMO) antennas technology in the different domains such as frequency, space, channel or even time [13]. As time goes by, spectral resources start to become scarce and, as a result, it becomes necessary to use more antennas, in other words more energy. One of the solutions found by the scientific community is Index Modulations (IM), which attempts to dynamically arrange which resources should be assigned, allocated, or activated. IM are techniques that aim to improve spectral efficiency (SE) and energy consumption and are one of the promising techniques for the new generation of 6G. With this optimization, it becomes possible to improve the properties of communications in terms of EE and SE, while maintaining reduced complexity [16]. IM have attracted the attention of the research community in the last decade as a very promising solution, since they consist of indexing special symbols in the modulations used to make communication from larger constellations more efficient, thus improving SE. One of the great capabilities of IM is that it can reduce or even eliminate the interference that occurs between users when there are a high number of transmissions occurring at the same time. [15]. Since communication systems may contain multiple transmitting and receiving antennas, i.e. MIMO schemes, it is necessary to come up with schemes or modifications for cases where, instead of activating all the antennas for a single receiver, a combination of active antennas can be used for each receiver or some method that manages to disperse the data correctly [13]. Instead of classic phase and amplitude methods, it is possible to use the resources of larger constellations in certain periods, times or even domains, thus making communications more reliable and efficient, or even reducing electromagnetic interference. By using the various domains together, whether they are spatial, frequency, time or even channel, it is possible to form a kind of pattern which is only used to activate, for example, the transmit antennas of a MIMO system, or even the Orthogonal Frequency Division Multiplexing (OFDM) sub-carriers or slot times [15].

Even though 5G already achieves good performances using MIMO links and communication between devices, it still cannot currently provide the communication rates necessary for certain applications, such as edge computing, or the reliability that future services might need [17]. One other problem that 5G suffers, is also the lack of coverage when using the higher bands. According to some studies, 6G mobile network is expected to be able to accommodate a high density of devices, around 500 billion, and with a scalability 1000 times greater than 5G [18]. This will require the use of a volume of new frequencies that have not yet been explored, offering transfer values in the order of hundreds of gigabits (Gbps) or even terabits (Tbps) per second [19]. While working in these high frequency bands can bring advantages, there are physical limitations. Terahertz (THz) and sub-THz communications are a strong bet in this field, since this type of communication, despite being micro-scale (only used for small distance communications), offers good performance capacity in open spaces, i.e., when there are almost no obstacles and line of sight (LOS) communication is possible. Faced with this situation, this type of communication suffers from high attenuation over long distances which often results in a lower

signal to noise ratio (SNR) value [20]. However, in cases of Non LOS (NLOS), which is when receivers do not have a direct line with the transmitter, normally they can be impaired [13]. However, academic activity and researchers have already proposed solutions to this type of problem, in which it is expected that in future generations it will be possible to physically implement this solution. One solution that seems to be very promising is the use of special panels that can partially reflect part of the incident electromagnetic wave, taking advantage of some of the dissipated energy, resulting in a small signal boost at the receiver.

Reconfigurable Intelligent Surfaces (RIS) are a new type of hardware panel with a special surface material that allow the modification and manipulation of certain parameters of the electromagnetic wave by tuning its phases [21]. This type of passive element panel is made up of a predetermined number of small, low-cost antennas, which are able to modify the properties of the electromagnetic wave, creating a funnel or a radius of greater concentration for certain areas. RIS has been designed as a key technology to assist or aid wireless communications by improving the coverage of Base Stations (BSs) or Cells, or even by enhancing channel conditions. Although it is still very early days, it is expected that good performance will be achieved when these two technologies are more mature [22]. Normally these types of RIS surfaces are built as flat surfaces, depending of the favorable electromagnetic wave propagation conditions. Technologies such as mmWaves and THz are able to work in conjunction with RIS, providing some significant gains. They are very useful in cases of NLOS, where there are obstacles capable of blocking wireless communications [23].

The type of material/surfaces which RIS are made, can be placed on multiple types of surfaces, such as building roofs, facades, furniture, clothing, etc. Typically, are easy to install, as they are passive elements, and can be placed on multiple structures, or even on other types of surfaces [24]. One of the greatest benefits of RIS, is to offer SE improvements, as they are surfaces that can be reconfigured dynamically. When it comes to energy consumption, they are known as environmentally friendly, because they are considered to be passive systems, since they are able to use part of the propagation energy of the incoming waves [25]. Most of the times RIS are compatible with most wireless network technologies. With this type of technology, it will be possible to shape and adapt environments and sites in order to obtain the best possible coverage, without affecting power consumption [26].

1.2. Objectives and Goals

After the brief introduction about the topics related to the work that will be presented in this thesis, it is clear what problems or challenges wireless communications are currently facing. Working in higher frequency bands is not always the best solution for the different foreseen scenarios. In some cases it is necessary to optimize the resources, techniques and methods that already exist or are emerging, and demonstrate their performance and robustness. To this end, solutions such as IM are one of the most promising techniques for demonstrating the possible optimizations that are needed in mobile communications, whether as SE or EE. To continue along the same topic of efficiency, it is important to explore the concept of RIS, that is still emerging in the scientific community. RIS can also make it possible to improve SE and EE when used on higher frequency bands, such as mmWaves and sub-THz. Both types of techniques can be integrated into 5G/B5G systems or on future new generations such as 6G.

The main objective of this thesis is to propose a design and integration of spectrum and energy efficient techniques into wireless communication networks. Considering the aforementioned objective, the work consisted of three phases, each with specific sub-objectives, as follows:

- **Scheme using IM based on two domains used simultaneously (space and frequency) and targeted at multi-user (MU) downlink communications** – a new scheme was proposed that works on two IM domains (both in the spatial and frequency domain), which was given the name Precoding-aided Transmitter side Generalized Space-Frequency Index Modulation (PT-GSFIM), and manages to remove interference between users in MU-MIMO transmissions. The approach incorporates a precoder at the transmitter side to assist in the transmission of PT-GSFIM. Complex Rotation Matrices (CRM) has been used as a technique for Signal Space Diversity (SSD) to exploit frequency diversity, which, when combined with IM schemes corresponds to a novel approach. Three specific algorithms have also been developed to detect Generalized Spatial-Frequency IM (GSFIM) symbols. Since the precoder removes all the multi-user interference (MUI), the receiver only has to implement single-user (SU) detection. The performance, effectiveness and SE of the proposed PT-GSFIM approach was demonstrated and compared against other existing methods;
- **Integration of the PT-GSFIM into C-RAN deployments, followed by thorough system-level evaluation considering B5G networks** – the use of the PT-GSFIM algorithm in a MU system in a Cloud-Radio Access Network (C-RAN) was studied and evaluated, with the purpose of assessing its performance, effectiveness and SE in downlink transmissions in realistic Beyond 5G (B5G) scenarios. In order to analyze the potential benefits of the proposed PT-GSFIM-based C-RAN, a detailed system-level assessment was presented for three different three-dimensional scenarios (one indoor and two outdoor), taken from the standardized 5G New Radio (5G NR), using two different numerologies and frequency ranges;

- **Optimization and evaluation of reconfigurable intelligent surfaces based technologies, also named as RIS, integrated into post-5G/6G networks operating at higher frequency bands (mmWave/sub-THz)** – In this topic, an iterative algorithm has been proposed to achieve the joint design of the access point precoder and the phase-shifts of the RIS elements, considering large scale MIMO OFDM links. The study includes the system-level analysis of a RIS-aided network deployment in post-5G/6G. The network operates in two different frequency bands, mmWave and sub-THz, considering both near-field and far-field propagation models. Two distinct environments, Indoor Open Office (IOO) and Urban Micro Truncated (UMT), were examined. The findings demonstrate that the adoption of the proposed RIS-based approach can effectively increase the throughput and coverage area.

In the following sub-section 1.3, we will describe all the works published during this doctoral process. Sub-section 1.4 will then give an overview of each published article, highlighting the objectives and results found during the research. Figure 1 contains a summary of all the research carried along this thesis.

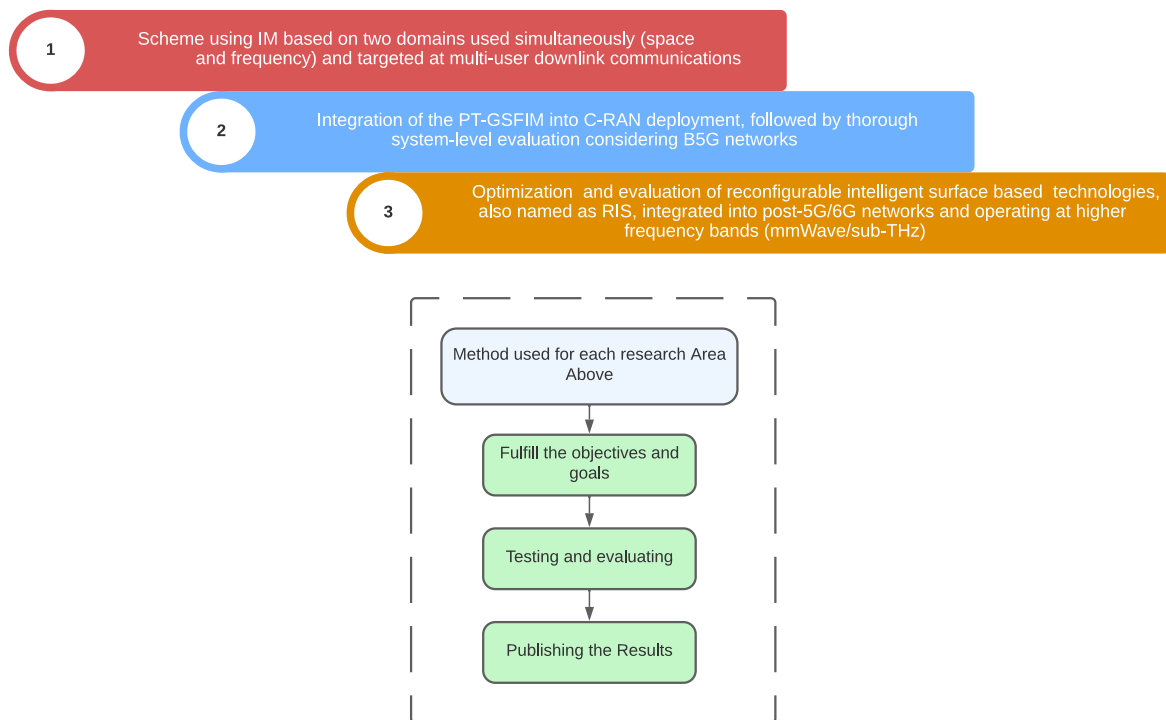


Figure 1 - Summary of the goals and research.

1.3. Research Contributions

The research carried out over the last few years has allowed some questions to be raised and answered. This work has resulted in the publication of some scientific articles on the topics mentioned above, resulting in a new scheme based on generalized index modulations in space and frequency, with the goal of improving SE and performance in terms of Bit Error Rate (BER). Proceeding on the same topic, we evolved to use the same approach but considering a system level design to prove that is possible to integrate this method on a MU-MIMO system based on standards for 5G NR systems and scenarios. In the last article, we explored the possibility of optimization and integration of combined RIS-assisted MIMO schemes operating in sub-THz bands into future post-5G/6G systems. This shown that is possible to improve the overall throughput and coverage area of receivers when we have cases NLOS with the BS.

Although this thesis focuses on the three articles presented in the chapters that follow, it was also necessary to develop and collaborate on other work within the same themes. This additional work was important to achieve the goals of the thesis resulting not only on the three main publications presented in this dissertation but also allowed the contribution to other three papers. All the published articles that resulted from this work are summarized in the table below, where they are marked as first author (FA), and as co-author (CA). There was a total of 6 publications, 3 FA and 3 CA.

Table 1 - Table of Publications

Date	Publications
September 2020	Precoded generalized spatial modulation for downlink MIMO transmissions in beyond 5G networks (CA)
August of 2021	A generalized space-frequency index modulation scheme for downlink MIMO transmissions with improved diversity (FA)
September 2021	Low Complexity Hybrid Precoding Designs for Multiuser mmWave/THz Ultra Massive MIMO Systems (CA)
February 2022	System-level assessment of a C-RAN based on generalized space–frequency index modulation for 5G new radio and beyond (FA)
March 2022	System-level assessment of low complexity hybrid precoding designs for massive MIMO downlink transmissions in beyond 5G networks (CA)
May 2023	Performance Assessment of a RIS-empowered post-5G/6G network operating at the mmWave/THz bands (FA)

After presenting the summarized version of the contributed research in the above table, the details of each contribution are presented below, starting with the chronological order of publication, with peer review and including its Scimago/Scopus Journal Ranking:

- J. P. Pavia, **V. Velez**, B. Brogueira, N. Souto, and A. Correia, "Precoded Generalized Spatial Modulation for Downlink MIMO Transmissions in Beyond 5G Networks," *Applied Sciences*, vol. 10, no. 18, p. 6617, Sep. 2020, doi: 10.3390/app10186617.
❖ Journal: Applied Sciences from MDPI; ISSN: 20763417; H-Index:101; Quartile Q2.
- **V. R. J. Velez**, J. P. C. B. B. Pavia, N. M. B. Souto, P. J. A. Sebastião and A. M. C. Correia, "A Generalized Space-Frequency Index Modulation Scheme for Downlink MIMO Transmissions With Improved Diversity," in *IEEE Access*, vol. 9, pp. 118996-119009, 2021, doi: 10.1109/ACCESS.2021.3106547.
❖ Journal: IEEE Access from IEEE Inc.; ISSN: 21693536; H-Index:204; Quartile Q1;
- J. P. Pavia **V. Velez**, R. Ferreira, N. Souto, M. Ribeiro, J. Silva and R. Dinis, "Low Complexity Hybrid Precoding Designs for Multiuser mmWave/THz Ultra Massive MIMO Systems," *Sensors*, vol. 21, no. 18, p. 6054, Sep. 2021, doi: 10.3390/s21186054.
❖ Journal: Sensors from MDPI; ISSN: 14243210, 14248220; H-Index: 219; Quartile Q1.
- **V. Velez**, João Pedro Pavia, Catarina Rita, Carolina Gonçalves, Nuno Souto, Pedro Sebastião and Américo Correia, "System-Level Assessment of a C-RAN based on Generalized Space–Frequency Index Modulation for 5G New Radio and Beyond," *Applied Sciences*, vol. 12, no. 3, p. 1592, Feb. 2022, doi: 10.3390/app12031592.
❖ Journal: Applied Sciences from MDPI; ISSN: 20763417; H-Index:101; Quartile Q2.
- J. P. Pavia, **V. Velez**, N. Souto, M. Ribeiro, P. Sebastião, and A. Correia, "System-Level Assessment of Low Complexity Hybrid Precoding Designs for Massive MIMO Downlink Transmissions in Beyond 5G Networks," *Applied Sciences*, vol. 12, no. 6, p. 2812, Mar. 2022, doi: 10.3390/app12062812.
❖ Journal: Applied Sciences from MDPI; ISSN: 20763417; H-Index:101; Quartile Q2.
- **V. R. J. Velez**, J. P. C. B. B. Pavia, N. M. B. Souto, P. J. A. Sebastião and A. M. C. Correia, "Performance Assessment of a RIS-Empowered Post-5G/6G Network Operating at the mmWave/THz Bands," in *IEEE Access*, vol. 11, pp. 49625-49638, 2023, doi: 10.1109/ACCESS.2023.3277388.
❖ Journal: IEEE Access from IEEE Inc.; ISSN: 21693536; H-Index:204; Quartile Q1;

1.4. Thesis Overview

The work produced within this thesis is focused on two main research areas, that were previously mentioned. The study in each area was directed related to other co-authored papers, even if the main chapters of this thesis are based on publications as FA. As a result, we give a broad outline of all the work developed during the PhD in this sub-section, emphasizing and connecting all the contributions mentioned in the preceding paragraph.

1.4.1. Index Modulations

In an effort to make wireless communications more efficient, and achieve lower latencies and higher data rates, the scientific community has turned to innovative physical layer solutions to improve their SE and EE. One of the main solutions available for mobile communications that has been attracting researchers corresponds to IM schemes. In addition to obtaining good results in studies already carried out, they have advantages over other methods, such as that their hardware simplicity, which can be relatively low. IM is a digital modulation technique that allows *special* bits to be indexed within the structure of a transmission packet according to the active resources (whether they are antennas, carriers, radio frequency chains, etc), thus allowing additional information to be sent. There are several types of IM, the most common being Spatial Modulation (SM) and its variants [27].

The first contribution in this topic was through a CA study on the application of IM in spatial modulations, more specifically using the Generalized Spatial modulation (GSM) version [28]. In this study, a Multilevel Quadrature Amplitude modulation (M-QAM) constellation of up to 1024 levels was used, where a BS with N_t antennas transmits a number of symbols (N_s), to a user receiver with a number of receiver antennas (N_r). Traditional M-QAM models do not explore the spatial domain. When considering the incorporation of the IM, the index bits select the transmit antennas through a switcher, as shown on Figure 2.

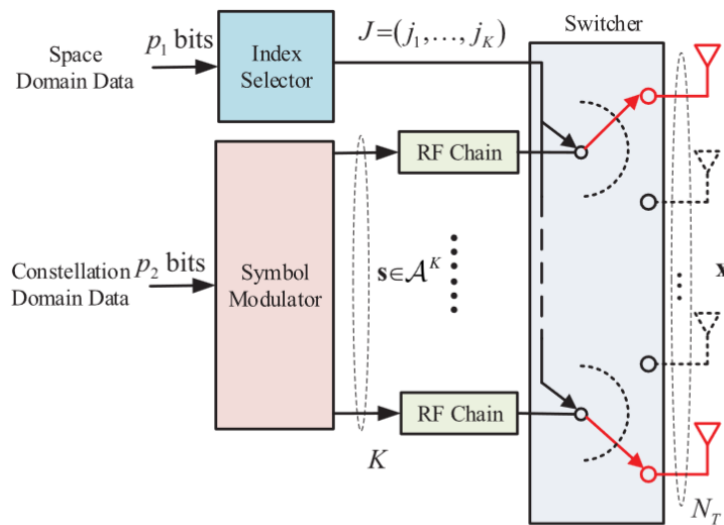


Figure 2 - Typical GSM scheme of transmitter side [29].

To better explore this capacity on IM scheme, in our paper, we used a GSM mapper to select the correct antenna on the spatial domain. Figure 3 shows the proposed scheme based on GSM used with the incorporated precoder. The singular value decomposition (SVD) based precoder combined with GSM, and with the assistance of an alternating direction method of multipliers (ADMM) algorithm at the receivers, was able to achieve better SE and EE than a conventional MIMO system. The proposed detector for the receivers also offers less complexity compared to its conventional counterpart. The proposed scheme was named GSM MU-MIMO. In order to eliminate inter-user interference and to convert the MU transmission into numerous independent SU links, a block diagonal precoder is implemented at the BS.

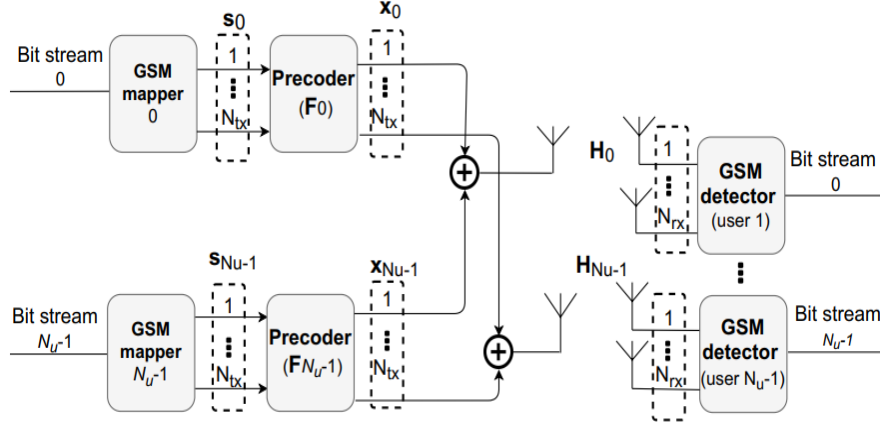


Figure 3 - Transmitter and receiver scheme of GSM-MU-MIMO [28].

In Figure 4, it is possible to verify that the algorithm provides a promising alternative, that can improve the system SE when compared to the use of higher-level modulations in conventional MU-MIMO. To compare the three algorithms, the same binary rate, measured in bits per second (bps), was maintained for each set of comparisons.

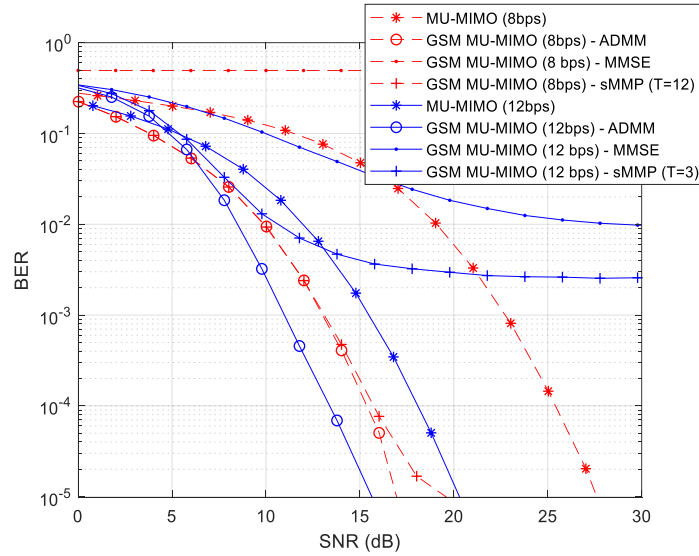


Figure 4 - Comparison of BER performance versus SNR between GSM-MU-MIMO and MU-MIMO [28].

Following on from the previous paper [28], we developed the IM-based approach resulting in the first FA paper [27]. This article is described in Chapter 2, where all the information is detailed. Although there have been studies on SU and MU in SM/GSM transmissions and IM-OFDM, one of the contributions was to propose an approach based on a SVD precoder combined with various IM domains, namely both the spatial and frequency domains. The proposed scheme, extends the initial approach presented in [30]. Our scheme, which we refer to as PT-GSFIM, combines the index bits to select the resources simultaneously, whether they are spatial (the antennas) or frequency (the OFDM sub-carriers). This ability to select resources allows us to bring improvements that IM offers. We have a number of transmitting antennas (N_{tx}) that represent the total number of available ones to use, a number of total available subcarriers (N_f), a number of active sub-carriers (N_{af}), and a number of active antennas (N_a). Each of these parameters can be changed to obtain the maximum performance. Figure 5 shows an example of how this type of scheme works. Each block corresponds to a selected sub-carrier with a selected antenna. To select the antennas which corresponds to the spatial slots available, we use the N_a parameter. The sub-carriers of OFDM are represented as N_{af} parameter. Together, they form a structure that contains the transmitting symbols represented as number of symbols (N_s). Then these symbols encode spatial bits and conventional modulated symbols and serve as the basis for constructing the complete PT-GSFIM symbol.

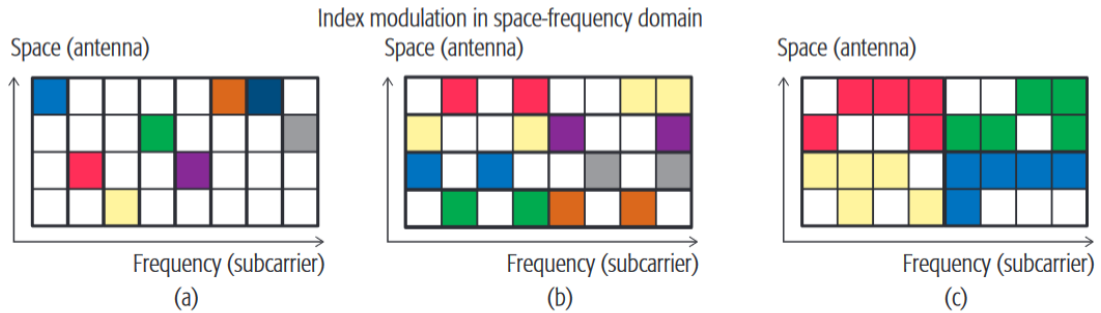


Figure 5 - Example of GSFIM scheme using space and frequency simultaneously [30].

In addition to the combination of precoder and IM, CRMs were also used to assist implementation as SSD on the transmitter side. SSDs, allows to make use of the frequency diversity. This technique can produce the potential gains of PT-GSFIM scheme in a context of downlink in a cellular transmission with MIMO schemes. The precoder scheme on the transmitter side allowed us to eliminate possible interference between the different users' signals, thus providing three new types of detectors. These new types of detectors allow the complexity to be reduced, while still enabling good reception performance. The three detector algorithms that have been used, were derived from existing methods and were modified in order to accommodate the PT-GSFIM signals. The algorithms correspond to Ordered Block Minimum Mean-Squared Error (OB-MMSE), Multipath Matching Pursuit with slicing (sMMP) and ADMM [31] [32] [33]. According to our analysis, the proposed algorithms offered less complexity than

a traditional transmission using a Maximum Likelihood Detector (MLD), improving BER performance, i.e., there is lower BER, thus improving its SE compared to MU-MIMO and GSM schemes. Also in our results, it was possible to use a lower constellation scheme in a downlink system, which allowed it to outperform other algorithms using higher constellations. Chapter 2, all the technical details of the study involved with the PT-GSFIM algorithm are explained in detail.

Still on the subject of IM, we continued the work carried out in the first article on PT-GSFIM, and considered the evolution and integration of the proposed approach into a C-RAN system based on 5G NR standards, described here [34]. The core idea behind C-RAN is to partition the BS functions by separating the main BS into the Base Band Unit (BBU) and Remote Radio Head (RRH), and then conglomerating the BBUs from numerous sites into a sole geographical point such as a cloud data center, applying cloud computing and virtualization approaches [35]. Another advantage is that it permits us to manage Radio Frequency (RF) chain resources remotely, without having to physically go to all the sites, allowing the management of carrier aggregation, for example. Figure 6 shows an example of a C-RAN.

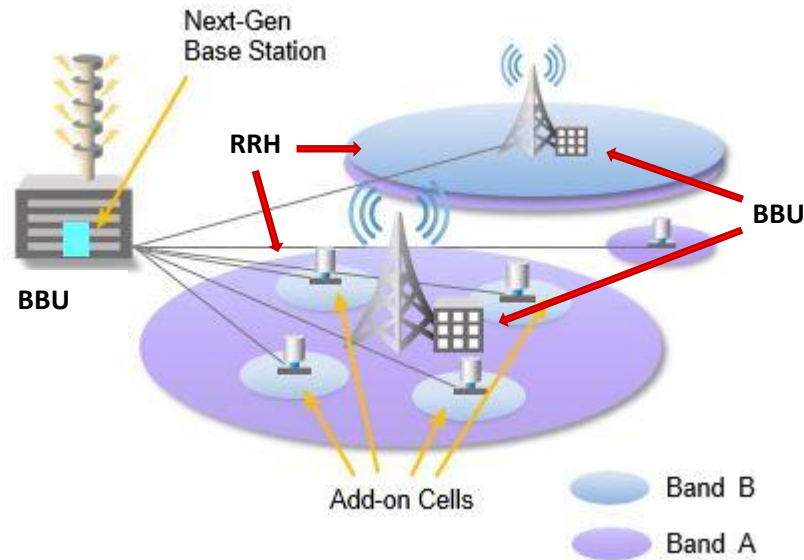


Figure 6 - Working principle of C-RAN [36].

In the system level evaluation considered in this work, we used a number of BSs and Access Points (APs) (which simulate BSs but they are small - micro cells) in order to try and recreate a real mobile device situation. We also considered three distinct scenarios using two numerologies of 5G NR, as stated, in order to test the overall SE at a system level. In this type of scenario, PT-GSFIM was studied as a downlink alternative to conventional MU-MIMO, with different M-QAM constellations adopted for comparison purposes. As with the PT-GSFIM approach presented in the previous work, an SVD precoder combined with CRM is used in the C-RAN integration in order to remove interference between users connected to the same BS/APs while providing additional diversity. Two of the three scenarios used are based on Urban Micro (street canyon) (UMi) – that focus on micro cells, and Urban Macro (UMa) – that focus on large cells, both of which were used with LOS and NLOS beams. The third scenario evaluated was an Indoor (InD), that represents the different types of offices where the receivers

are placed. On both urban and indoor scenarios, a frequency band of 30 GHz, and a 70 GHz was used (optional one along with the primary band of 30 GHz, but only on indoor scenario). All the specifications used are described in [34]. The assessment was based on link-level simulations, which provide input to the system-level simulator, which then simulates multiple deployments with multiple APs and users operating simultaneously. As expected, the PT-GSFIM approach outperformed the traditional GSM and MU-MIMO schemes. Chapter 3 presents the published article in which all the work summarized above is detailed.

1.4.2. Millimeter-Waves and Terahertz Communications and Reconfigurable Intelligent Surfaces

Millimeter-wave and terahertz technologies have gained popularity in the wireless research community due to their potential for providing large and underutilized bandwidths that can enable ultra-high-speed connections, for example the desired rates of 100 Gbps. However, working in the THz band presents challenges such as reflection and scattering losses along the transmission path, high dependency on distance. The high signal attenuation present at these frequencies necessitates the use of very large antenna arrays to achieve low complexity and power consumption, alongside of Hybrid analog/digital schemes/designs to obtain optimal performance. A hybrid design algorithm is suitable for both mmWave and THz MIMO systems. As for, we proposed in [37], a design that can accommodate various analog and digital architectures, such as phase shifters, switches, inverters, antenna selection, and more. This makes it applicable to support multi-user ultra-massive MIMO (UM-MIMO) in severely hardware-constrained THz systems.

Figure 7 shows an example of hybrid precoding scheme. Both the precoder and combiner consist of discrete digital and analog processing units, as the analog precoder is placed after of transmitting antennas e before of the FFT blocks on receiver.

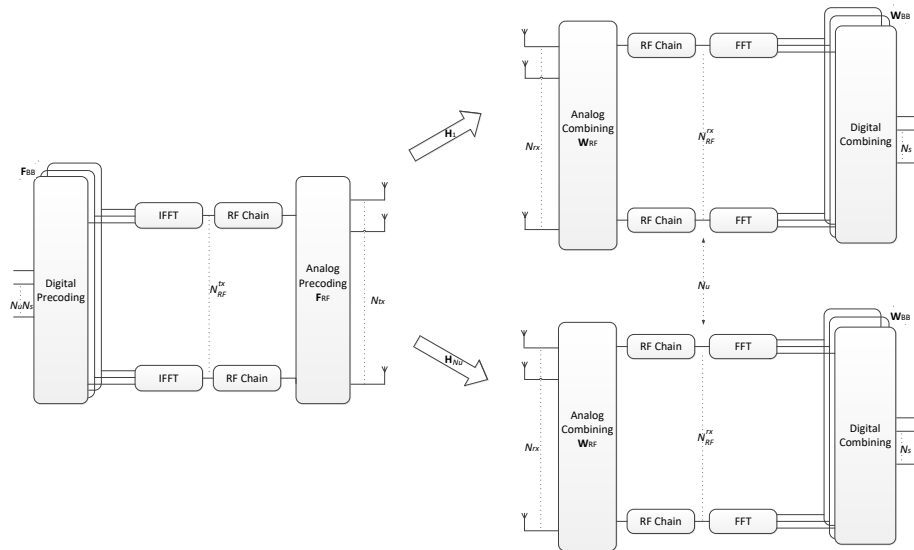


Figure 7 – Design of a Hybrid precoding, both transmitter and receiver sides [37].

Figure 8 shows the different analog architectures that were considered for the design of the proposed hybrid precoding. The different architectures can encompass various structures, including fully connected, arrays of subarrays (AoSA), and dynamic arrays of subarrays (DAoSA), which are suitable for THz communications.

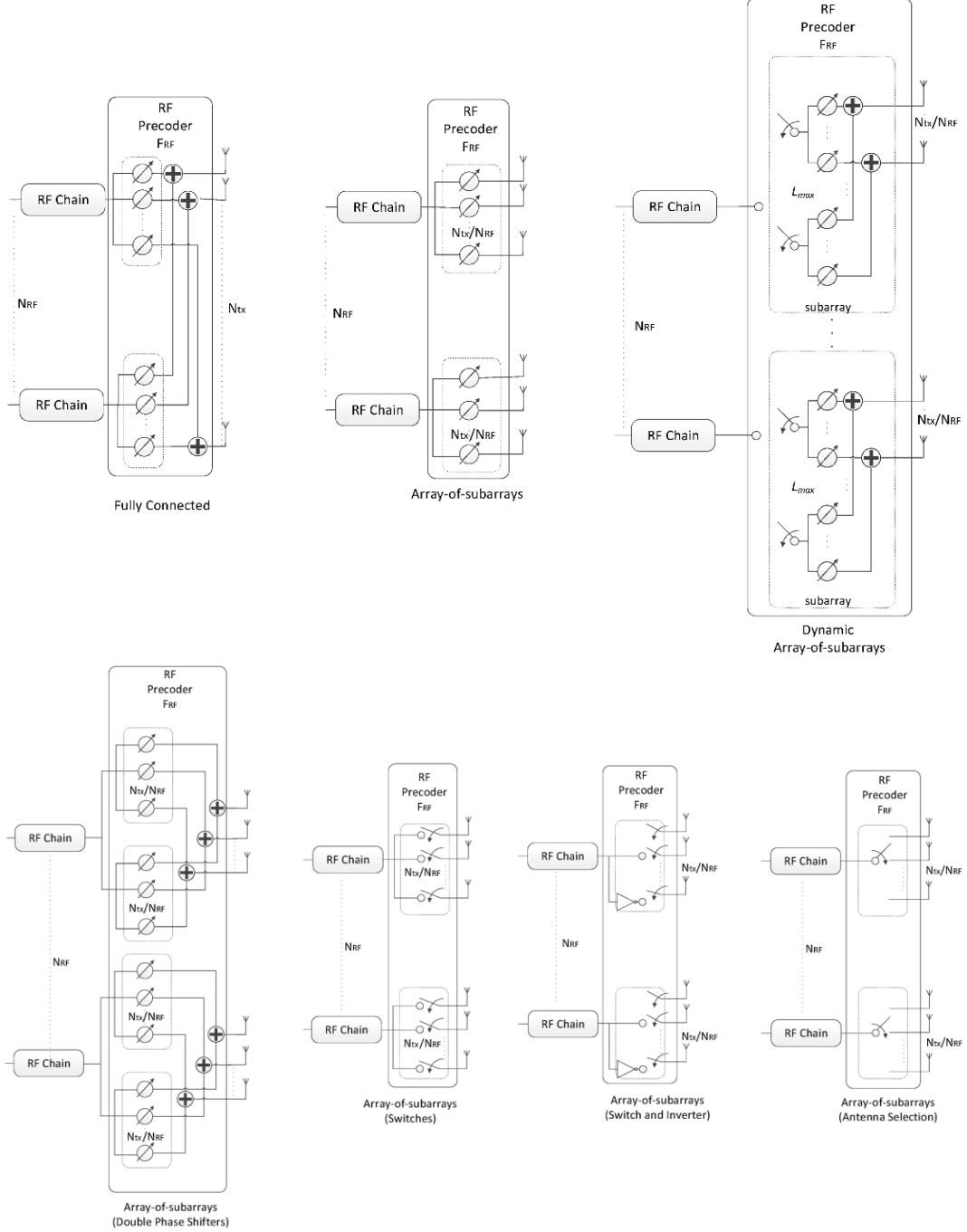


Figure 8 - Example of and hybrid precoding design, intended to a THz system [37].

Hybrid schemes benefit from lower energy consumption in analog schemes and reduced complexity in digital schemes, making energy consumption a primary target. Such schemes allow us to enhance system SE, while reducing the overall energy consumption.

Continuing the previous work on hybrid designs, a system-level evaluation was conducted on a C-RAN [38]. In this case we proposed an evaluation of a C-RAN designed for beyond-5G systems, integrating this hybrid schemes, with specifically focus on APs that operate in the mmWave/THz bands, supporting MU-MIMO transmission with massive/ultra-massive antenna arrays combined with low-complexity hybrid precoding architectures.

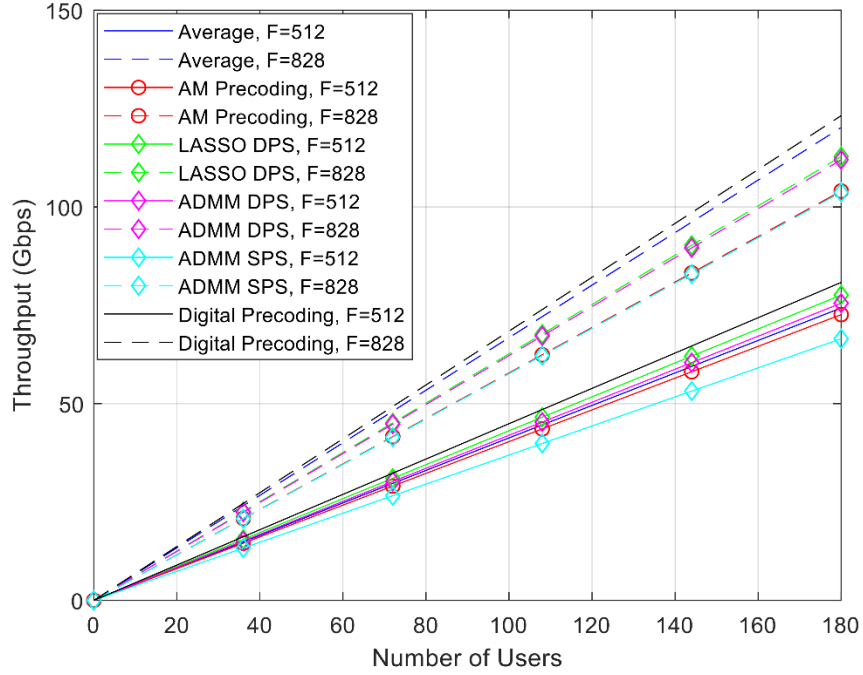


Figure 9 - Throughput vs number of users when using hybrid precoding systems into a C-RAN [38].

In figure 9, we observe that digital precoding, offers better throughput in contrast to the analog precoding, which benefits the power consumption. In this evaluation, we shown a throughput vs number of users, using $N_{tx}=100$ and $N_{rx}=4$, $N_{RF}^{tx}=12$, $N_s=2$. Parameter F in Figure 9 corresponds to the number of sub-carriers used in each configuration. Schemes based on double phase shifters (DPS) present a better performance when compared to the ones based on single phase shifters (SPS). In general, hybrid design algorithms provide a performance level similar to fully digital precoders while presenting less complexity for implementation purposes.

According to many studies [39], RISs are the next key technology for the development of mobile networks capable of meeting EE requirements. As mentioned above, this type of surface uses passive elements to *re-shape* the direction of the electromagnetic wave during its propagation. Figure 10 shows an example of a RIS panel comprising several tunable elements, where each element corresponds to a phase shifter. We can see an example in Figure 10 (a) where a symbol is transmitted and part of the reflection of the incident wave is reflected back to the receiving antenna (b) and (c).

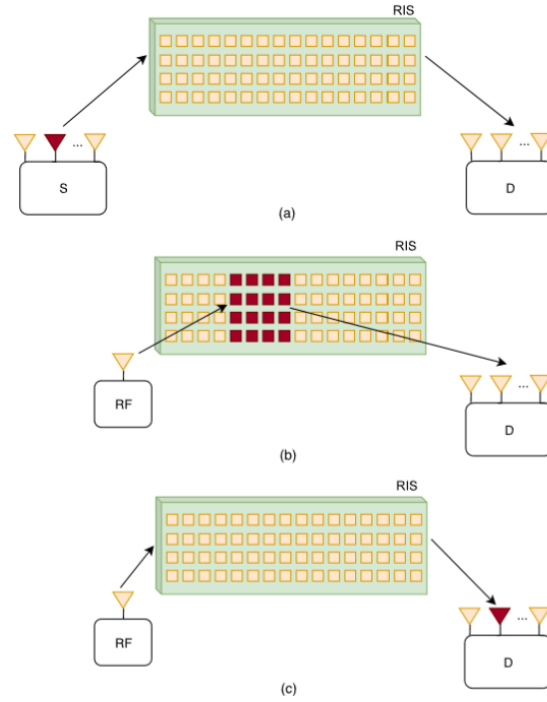


Figure 10 - Example of a RIS panel, with several tunable elements [40].

Figure 11, shows an example of a RIS deployment, assisting an AP with the objective to increase the received signal, on a NLOS case.

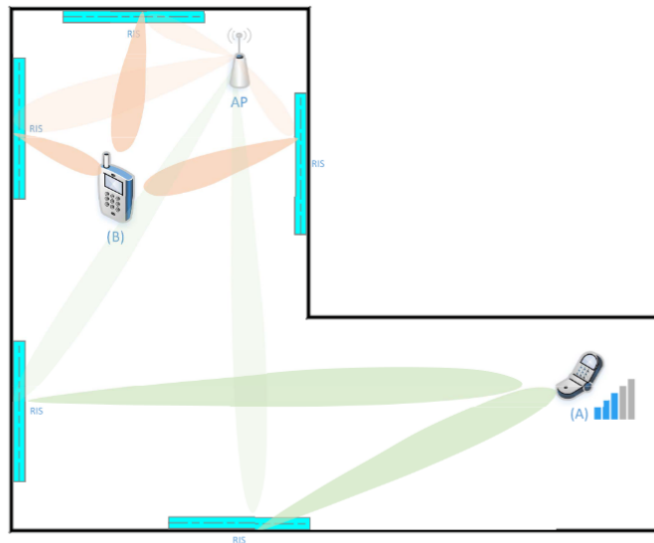


Figure 11 - RIS example in a room, showing an increase on receiver signal [23].

In this subject on RIS [41], we have considered combining with massive and ultra-massive arrays of antennas considering the integration into future mmWave/THz communication systems. The proposed approach was studied both at the physical and at the system level. The scenarios used are based in the UMT, where multiple BSs transmit the signals to multiples users/receivers, and in the IOO environment, which represents the casual offices around and inside of the cities. The UMT, represents dense urban scenarios (e.g. large cities). The IOO scenario contains some obstacles that could deflect or impact the propagation of original signal. We considered two scenarios from 5G NR Standards, one in

the mmWave band (28 GHz), and another one operating in sub-THz frequencies (100 GHz). This requires us to adopt two different propagation models. The aim was to demonstrate, in a traditional MIMO environment, how RIS can assist this type of communication, which is prone to high attenuation, and improve the coverage of the same system.

To accomplish this, we designed a new joint precoder and RIS optimization algorithm for OFDM transmissions with multiple sub-carriers and large-scale MIMO. The algorithm presented is called Alternating Maximization-Singular Value Decomposition-Accelerated Proximal Gradient (AM-SVG-APG) and consists of the combined use of alternating maximization (AM) and the Singular Value Decomposition (SVD) methods, with the accelerated proximal gradient (APG) used to perform the individual optimization of the RIS elements. The aforementioned algorithm performs the computational calculation to optimize the best phase shifts for each element of the RIS panel. Each RIS panel is composed of a number of (N_{ris}) RIS elements. The number of N_{ris} elements has influence on average BER and throughput results. The placement of RIS panels between the transmitter and receiver also impacts the performance, even if only slightly. Link-level simulation results have demonstrated that the proposed approach is highly effective in achieving the full potential gains of RIS-assisted communications. The results actually show two interesting situations: the transmitted power values show that it is possible to slightly improve performance in terms of BER, depending on the position of the RIS panels (remember that these elements are passive and do not amplify the signal); and that the throughput values in relation to the distribution of the service by users can offer improvements in signal coverage of up to 53% and up to 58% of throughput. This demonstrates the feasibility of achieving high transmission rates using RIS technology in both indoor and outdoor environments at mmWave and sub-THz bands.

All of the aforementioned details are provided in Chapter 4, which contains the entirety of the article, from its analysis and synthesis to its results.

Chapter 2 - A Generalized Space-Frequency Index Modulation Scheme for Downlink MIMO Transmissions With Improved Diversity

As mentioned above, multidimensional IM schemes are an alternative physical layer approach that can provide wireless communications with improved tradeoff between SE, power consumption and implementation complexity. As a possible solution, instead of adding more antennas and resources, IM could change the paradigm of how the various resources are activated through the transmission of implicit information *masked* inside these wireless resources.

This chapter presents the contribution of the proposed PT-GSFIM scheme, that combines part of the information bits into the respective active antennas and subcarriers, which also carry modulated symbols. As will be explained below, in this contribution, more importance has been given to MU cases, as there have not yet been many studies on this subject and adopting this type of approach. To improve even further the performance without causing more complexity on the transmitter and receiver, SSD techniques were considered, through the adoption of CRM, which increases the diversity for a downlink transmission in a MIMO system. One of the main benefits of the adopted precoder is the capability to remove MUI from the received signals. In order for the receivers to be able to directly decode the PT-GSFIM signals, three detection algorithms have been proposed: the OB-MMSE, sMMP, and ADMM. The performance results were obtained in terms of BER for different algorithms. We also tested, how different number of users (N_u) can impact the performance in a MU scenario, while keeping the same SE and same constellation.

Below is the published and accepted paper containing all the information on the scheme mentioned, including the method used to analyze it and the algorithms used for comparison purposes.

Paper details:

- Title: A Generalized Space-Frequency Index Modulation Scheme for Downlink MIMO Transmissions with Improved Diversity
- Date of Publication: 20 of august of 2021
- Journal: IEEE Access
- Scimago/Scopus Journal Ranking: Quartile 1
- Publisher: IEEE inc.

A Generalized Space-Frequency Index Modulation Scheme for Downlink MIMO Transmissions with Improved Diversity

Vasco Velez^{1,2}, João Pedro Pavia^{1,2}, Student Member, IEEE, Nuno Souto^{1,2}, Senior Member, IEEE, Pedro Sebastião^{1,2}, Member, IEEE, and Américo Correia^{1,2}, Senior Member, IEEE

¹Department of Information Science and Technology, ISCTE-Instituto Universitário de Lisboa, 1649-026 Lisboa, Portugal

²Instituto de Telecomunicações, 1049 - 001 Lisboa, Portugal

Corresponding author: Vasco Velez (e-mail: vasco_vez@iscte-iul.pt).

“This work was partially supported by the FCT - Fundação para a Ciência e Tecnologia under the grant 2020.05621.BD. The authors also acknowledge the funding provided by FCT/MCTES through national funds and when applicable co-funded EU funds under the project UIDB/50008/2020.”


ABSTRACT Multidimensional Index Modulations (IM) are a novel alternative to conventional modulations which can bring considerable benefits for future wireless networks. Within this scope, in this paper we present a new scheme, named as Precoding-aided Transmitter side Generalized Space-Frequency Index Modulation (PT-GSFIM), where part of the information bits select the active antennas and subcarriers which then carry amplitude and phase modulated symbols. The proposed scheme is designed for multiuser multiple-input multiple-output (MU-MIMO) scenarios and incorporates a precoder which removes multiuser interference (MUI) at the receivers. Furthermore, the proposed PT-GSFIM also integrates signal space diversity (SSD) techniques for tackling the typical poor performance of uncoded orthogonal frequency division multiplexing (OFDM) based schemes. By combining complex rotation matrices (CRM) and subcarrier-level interleaving, PT-GSFIM can exploit the inherent diversity in frequency selective channels and improve the performance without additional power or bandwidth. To support reliable detection of the multidimensional PT-GSFIM we also propose three different detection algorithms which can provide different tradeoffs between performance and complexity. Simulation results shows that proposed PT-GSFIM scheme, can provide significant gains over conventional MU-MIMO and GSM schemes.

INDEX TERMS Index Modulation (IM), Precoding-aided Transmitter side Generalized Space-Frequency Index Modulation (PT-GSFIM), Multiple-Input Multiple-Output (MIMO), Orthogonal Frequency Division Multiplexing (OFDM), Multiuser-MIMO (MU-MIMO).

I. INTRODUCTION

Due to the increased number of devices connected to the internet and to the amount of mobile traffic services growing up every day, new and more

efficient solutions are needed to improve current wireless networks. Therefore, newer solutions are necessary to offer higher mobility communications with a higher spectral and energy efficiency (EE). This implies much higher transmission rates, with substantial lower latencies, enabling new types of

The associate editor coordinating the review of this manuscript and approving it for publication was Chen Chen .

applications such as: extended reality, internet of everything, autonomous connected vehicles and so on. Amongst the potential physical layer solutions for future wireless systems, index modulation (IM) schemes have been captivating a lot of research interest in the wireless community. Instead of using conventional multiple-input multiple-output (MIMO) with a large number of radio frequency (RF) chains and scaled power consumption, IM has attracted researchers as an alternative because of its interesting tradeoffs between number of data streams, performance and EE [1], [2]. These characteristics benefit both the receivers and transmitters enabling reduced complexity implementations [3], [4]. All these advantages make IM schemes potential candidates for future of wireless communications, namely for 5G and 6G networks [5].

IM techniques convey additional information bits contained implicitly in the index of the selected element (in the spatial, frequency, temporal/time or channel domain), such as the active antenna subset, as is done in generalized spatial modulations (GSM) [6] or the active subcarrier subset, as is done in IM aided orthogonal frequency division multiplexing (IM-OFDM) [7].

Spatial modulation (SM) is an IM scheme that is a simplification of MIMO with lower implementation and computational complexity, and it only requires a single RF chain to convey additional information [8]. In SM, only one antenna is active at any given time which often transmits a conventional amplitude and phase modulated (APM) symbol, such as phase shift keying (PSK) or quadrature amplitude modulation (QAM). Alternatively, information can be sent solely through the index of the active transmit antenna, as is done in space shift keying (SSK) modulation [9]. Comparatively, SSK can be seen as a low complexity implementation of SM since it does not require inter-antenna synchronization (IAS) and the complexity of the detection is smaller. Still, the spectral efficiency (SE) is also smaller than SM. GSM generalizes the concept of SM to multiple active antennas during transmission. Instead of activating a single transmit antenna, several antennas are activated simultaneously in order to transmit multiple M -ary modulated symbols [10], [11]. The additional information is conveyed implicitly through the respective indexes of the antenna-activation pattern. Even though SM schemes present some advantages over conventional MIMO, such as mitigating inter-channel interference (ICI) and better performance in terms of bit-error rate (BER) and EE, they also bring some drawbacks. In fact, the use of only one active antenna (or a reduced number in the case of GSM) limits the SE for the same number of available transmit antennas, also sacrificing some diversity gain [2], [11], [13].

Channel state information (CSI) is required in many MIMO schemes. It can be exploited at the transmitted

side in order to reduce the complexity of receivers, and is referred to as CSIT, or at the receiver side, being referred to as CSIR, which is the most commonly used in SM/GSM schemes [6]. The use of CSIT combined with precoding can enable the use of SM, denoted as precoded SM (PSM), when there are strict complexity constraints at the receiver. In this case additional information bits are mapped to the index of the targeted receiver antenna during each transmission [12]. Generalized Precoding aided SM (GPSM) is an extension of PSM similar to GSM, which permits multiple receive antennas to be activated during each timeslot/transmission [13]. Quadrature SM (QSM) is another type of extension of SM, which divides the transmission into independent real and imaginary chunks that can be transmitted by different antennas simultaneously, with additional information bits being used for selecting the active antennas [14]. QSM is able to increase the SE of the system when compared to SM. Generalized Precoding aided QSM (GPQSM) generalizes the QSM concept to the receiver side, improving the SE over GPSM [15].

IM-OFDM, uses a similar approach to SM/GSM techniques but relies on the indices of existing subcarriers of a conventional OFDM transmission [16], [17]. In this case, only a subset of the subcarriers are activated depending on the index bits [18] and only these subcarriers convey M -ary signal modulated symbols [19], [20]. When index bits are triggered, they indicate which subcarriers inside a block of an OFDM frame should be activated during communication [4], [6]. In [21], the authors proposed a precoding technique, using IM on MIMO-OFDM systems, which tries to decrease the receiver complexity at the downlink. This is done by applying precoding to each subcarrier frequency so that the OFDM frames are received orthogonally without interference. A different approach is Layered Orthogonal Frequency Division Multiplexing with IM (L-OFDM-IM) which was proposed to increase SE in OFDM-IM scheme [22]. It splits the subcarriers into n layers, with the index bits selecting active subcarriers that can overlap but are distinguishable through different carefully designed constellations. Recently, a lot of attention has been directed to schemes that involve the use of multiple signal constellations aided by IM [23] – [26]. The objective is to bring some improvements in SE and diversity to OFDM-IM schemes. In [23], it was introduced an OFDM-IM scheme named dual mode IM aided OFDM (DM-OFDM), that divides the subcarriers into two major groups, which are modulated with different constellations “modes”, that correspond to two index subsets. Unlike an OFDM-IM system, where a subcarrier is switched on or off according to the index bits, DM-OFDM modulates all subcarriers which improves the SE [23]. In [24], generalized (G) DM-OFDM extends the previous idea to multiple constellations that are changeable according to the

information bits of each OFDM subblock, with a minor loss in performance. Additionally, the authors in [25] proposed a novel multiple-mode (MM) OFDM-IM scheme where all subcarriers are activated with different types of PSK/QAM signal constellations. In order to improve the number of transmission patterns, the full permutations of the distinguishable modes are also used to convey additional information. This approach was generalized in [26] (G-MM-OFDM-IM) by allowing different subcarriers to carry signal constellations of different sizes while conveying the same number of IM bits. These schemes were shown to improve the SE over OFDM-IM and DM-OFDM at the cost of additional complexity.

Even though in this paper we focus on the use of IM in the space and frequency domain, there are other approaches proposed in the literature. For example, IM can also be applied in the time domain (TD), by using time slots which are activated according to their indices, and can also be combined with space-time block codes (STBC) [27]. In [28], the authors presented a new type of IM-based uplink called IM-multiple access (IMMA), that uses a non-orthogonal multiple access (NOMA) method, operating with time slots similarly to transmission division multiple access (TDMA). In IMMA, each user can select his own time slot independently without the need of scheduling, but also, time slots can be shared with multiple users. Another different IM approach corresponds to channel domain IM (CD-IM), also known as media-based modulation (MBM), where information bits are transmitted via different channel realizations generated through available RF mirrors [29].

Regarding the research work that has been done in SM/GSM, besides single-user (SU) transmission, several multiuser (MU) schemes have been proposed for downlink and uplink scenarios. While the uplink can be addressed as a direct extension of SU GSM [30], the downlink requires a different approach. In the downlink, the users will typically have a limited number of antennas which can make them prone to MU interference (MUI). In this case one can resort to PSM to exploit the independence of multiple antennas and achieve higher SE and EE [31]. Resorting to precoding, the IM approach can be implemented at the transmitter side [34] or at the receiver side [31]. Receive SM (R-SM), consists in using some of information bits for selecting the receiver antennas, which will in fact receive the intended signals [31] – [33]. A specific precoder or special preprocessing stages must be adopted to achieve the desired antenna activation pattern at the receiver. One drawback of this approach concerns the fact that the receiver antennas must remain active all the time, even though only a portion of them receive symbols [32]. In [34], the authors adopted a different form of precoded space domain IM, which tries to improve SE by using a virtual GSM applied at the transmitter side through the use of precoding. In this

case, virtual active antennas transmit M -ary modulated symbols using a precoder in order to mitigate MUI and allow SU GSM detection.

Besides the different versions that IM can assume, which depend on the specific dimension that is exploited, several authors have recently started to propose multidimensional schemes that are based on combinations of multiple one-dimensional IM [4]. For example, in generalized space-frequency IM (GSFIM) [35], both SM and OFDM-IM schemes are combined to make use of the frequency domain and spatial domain simultaneously. In this case, index bits are used for selecting the active antennas and subcarriers at each transmission interval. The authors in [35] explored this approach showing good results both in achievable rates and BER when compared with conventional schemes. However, only SU scenarios were considered. One problem with OFDM, and consequently with OFDM-IM and GSFIM is that it does not exploit the frequency diversity which can degrade the performance when some subcarriers experience deep fades. In this case, low-rate channel codes can be adopted but this reduces the system's SE. As an alternative, in [36] the authors proposed the use of subcarrier interleaving for improving the performance of OFDM-IM.

Motivated by the work above, in this paper we assume an OFDM MU downlink transmission as the baseline and propose a precoding-aided, transmitter side multidimensional IM scheme which we refer to as PT-GSFIM. In PT-GSFIM, part of the information bits are used to select resources in the spatial and frequency domain whereas the remainder of the bits APM symbols. To help circumvent the poor performance problem of uncoded GSFIM we adopt the concept of signal space diversity (SSD), which was originally presented in [37], into the design of the signal. The use of SSD enables us to associate the bits over several subcarriers and benefit from the diversity effects inherent to a frequency selective channel, while keeping the mapping/de-mapping process simple. In this case, we integrate SSD into PT-GSFIM using complex rotation matrices (CRM), which were proposed in [38] to achieve time-diversity within the context of downlink transmission in a cellular system with multiple transmit antennas. At the transmitter side we consider that a precoder is used for removing MUI. Regarding the receiver, we propose several types of detectors with different tradeoffs in terms of complexity and performance. The main contributions of this paper can be summarized as follows:

- A frequency and space domain IM scheme that we refer to as PT-GSFIM, is designed for the downlink of MU-MIMO systems. To simplify the implementation, PT-GSFIM independently encodes part of the information bits onto spatial indexes (selecting active virtual antennas) and part onto frequency indexes (active subcarriers).

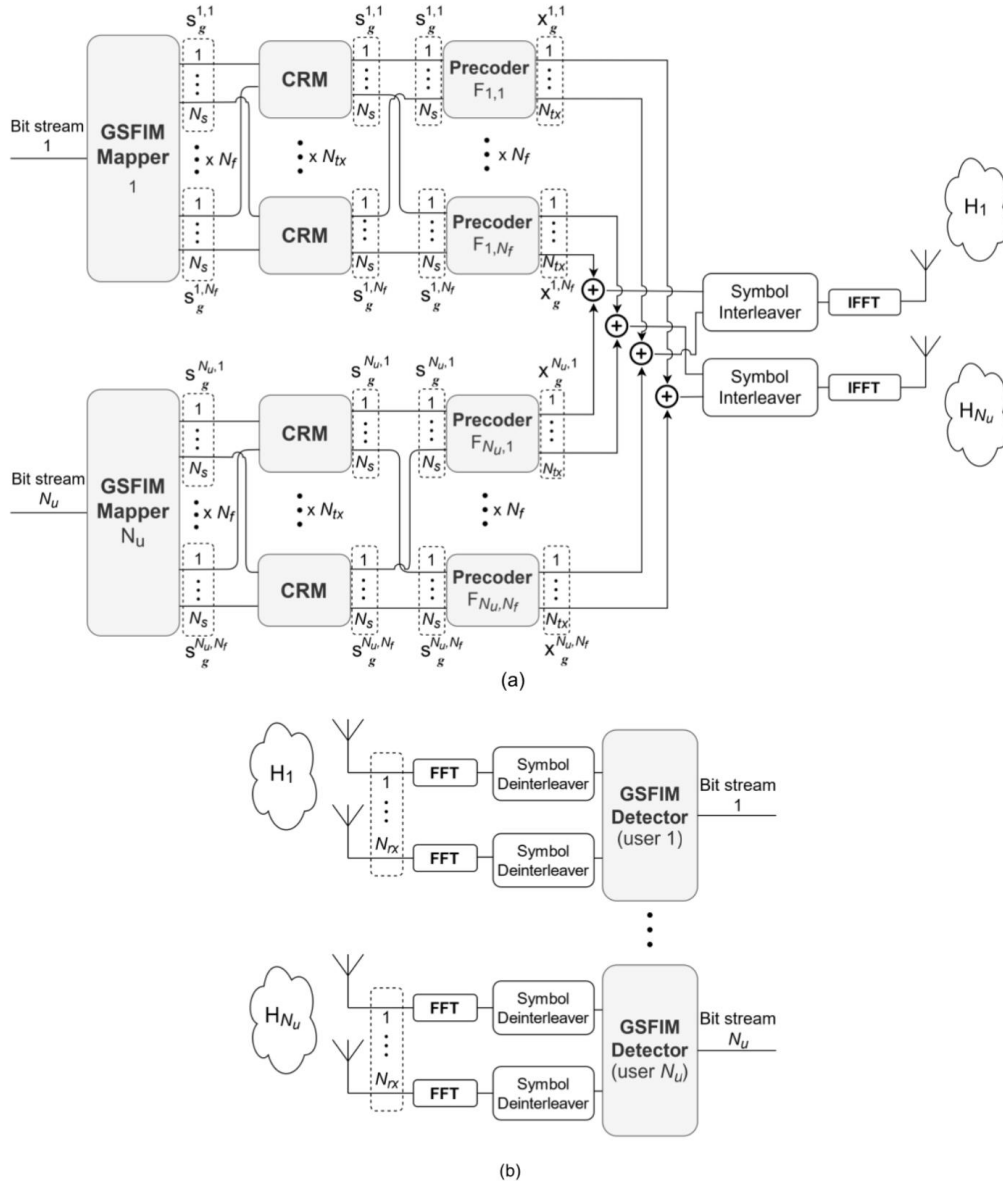


FIGURE 1. System Model: PT-GSFIM transmitter (a) and receiver (b).

- The remaining information bits are mapped onto M -ary modulated symbols. To handle MU downlink scenarios, PT-GSFIM incorporates a precoder applied to each subcarrier but to the whole set of transmit antennas (active and inactive) which keeps the local structure of a conventional transmitted GSM signal intact. This allows the removal of MUI and transforms a MU communication into multiple independent SU communications.
- To avoid the problem of having active subcarriers experiencing deep fades at the same time that some inactive subcarriers experience stronger channels, the proposed PT-GSFIM integrates SSD, implemented using CRM matrices applied over each sub-

block. Combined with subcarrier level interleaving, this approach can exploit the inherent diversity in frequency selective channels whilst keeping the structure of the PT-GSFIM signal intact from the point of view of the receivers. Improved performance can thus be obtained without any additional power or bandwidth.

To accomplish reliable detection of the multidimensional PT-GSFIM signal at the receiver's side, three different algorithms are presented. The proposed approaches are generalized versions of three known GSM algorithms that were modified in order to cope with the special structure of PT-GSFIM signals with CRM, which must consider the joint combination of active antennas and

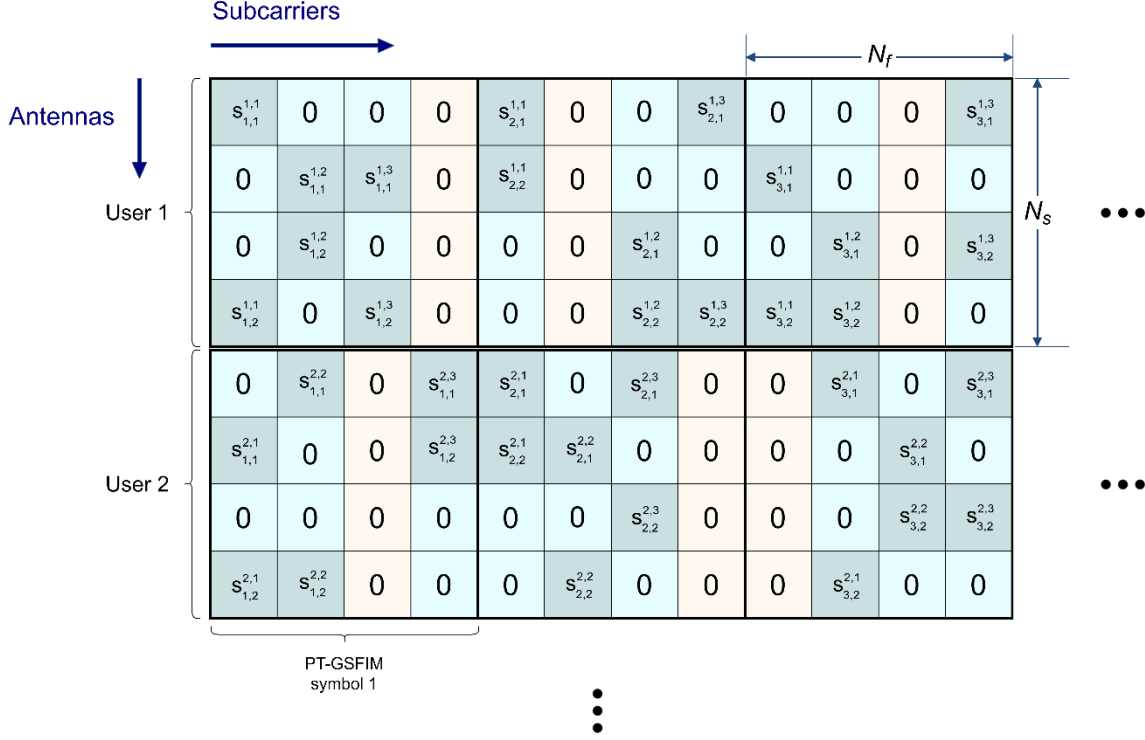


FIGURE 2. Example of the structure of each PT-GSFIM symbol in the space-frequency domain with $N_f=4$, $N_a=3$, $N_s=4$, $N_b=2$.

subcarriers. It is shown that the three derived algorithms provide different tradeoffs between performance and complexity.

The paper is organized as follows: Section II presents the system model which is followed by the PT-GSFIM transmitter structure formulation in section III. The three different algorithms proposed for signal detection are described in section IV alongside their respective complexities. Performance results are then presented in section V. Finally, the conclusions are outlined in section VI.

Notation: Matrices and vectors are denoted by uppercase and lowercase boldface letters, respectively. $(\cdot)^T$ and $(\cdot)^H$ denote the transpose and conjugate transpose of a matrix/vector, $\|\cdot\|_p$ is the ℓ_p -norm of a vector, $\|\cdot\|_0$ is its cardinality, $\text{supp}(\mathbf{x})$ returns the support of \mathbf{x} , $\lfloor \cdot \rfloor$ is the floor function and \mathbf{I}_n is the $n \times n$ identity matrix.

II. SYSTEM MODEL

Let us consider the downlink of an OFDM-based MU-MIMO system where a BS equipped with N_{tx} antennas transmits to N_u users, each with N_{rx} antennas, as shown in Fig. 1.

We consider also the adoption of an OFDM scheme, where part of the information is mapped onto conventional M -sized APM symbols and the other part is encoded both on spatial and frequency indexes. While the encoding onto the spatial indexes follows the same GSM approach from [39], the mapping onto the frequency indexes assumes the grouping of the

subcarriers into N_f -sized subblocks inside which only N_{af} subcarriers are active for each user. We assume a simple mapping of the bits to the transmitted block matrix, where the encoding onto spatial indexes is accomplished independently from the encoding to the frequency indexes, using simple look-up tables (LUTs): one for the spatial resources and one for the frequency resources in each subblock. While this independent mapping sacrifices some SE (the number of available space-frequency blocks that can be indexed is reduced) it makes the mapping/de-mapping process easier as well as the detection, as we will show further ahead. It is important to note also that the approach based on a pair of LUTs is adequate when the number of active antenna combinations and active subcarrier combinations is small, as is the case of the scenarios considered in this paper. If the number of combinations is large, then a combination strategy can be employed as described in [40] and [20]. Fig. 2 illustrates an example of the structure adopted for each PT-GSFIM symbol for the case of frequency domain sub-blocks with $N_f=4$ subcarriers each, where $N_{af}=3$ of them are active subcarriers, thus encoding frequency domain index bits. These active subcarriers convey GSM symbols ([39]) which encode both spatial bits and conventional modulated symbols, and are thus used as the base element for constructing the whole PT-GSFIM symbol. In this example each user's GSM symbol has a size of $N_s=4$ with $N_a=2$ active (nonzero) positions.

According to the figure, the g^{th} PT-GSFIM symbol, $\mathbf{s}_g^u \in \mathbb{C}^{N_s N_f \times 1}$, for the u^{th} user can be written as

$\mathbf{s}_g^u = \left[\left(\mathbf{s}_g^{u,1} \right)^T \dots \mathbf{0} \dots \left(\mathbf{s}_g^{u,N_{af}} \right)^T \right]^T$, which corresponds to the concatenation of $(N_f - N_{af})$ length- N_s vectors of zeros with N_{af} GSM symbol vectors $\mathbf{s}_g^{u,i} \in \mathbb{C}^{N_s \times 1}$ that are defined as $\mathbf{s}_g^{u,i} = \left[0 \quad s_{g,1}^{u,i} \quad 0 \quad \dots \quad 0 \quad s_{g,N_a}^{u,i} \quad 0 \right]^T$, where $i = 1, \dots, N_{af}$, $u = 1, \dots, N_u$ and $g = 1, \dots, N_{GSFIM}$. The symbols $s_{g,j}^{u,i}$, $j = 1, \dots, N_a$ are selected from an APM constellation. Assuming a MU downlink scenario where precoding is applied, we can write the transmitted signal as

$$\mathbf{x}_g = \sum_{u=1}^{N_u} \mathbf{F}_u \mathbf{s}_g^u = \mathbf{F} \mathbf{s}_g, \quad (1)$$

where $\mathbf{s}_g = \left[\left(\mathbf{s}_g^1 \right)^T \dots \left(\mathbf{s}_g^{N_u} \right)^T \right]^T$. Matrix $\mathbf{F} = \left[\mathbf{F}_1 \dots \mathbf{F}_{N_u} \right]$

with $\mathbf{F}_u \in \mathbb{C}^{N_{tx} N_f \times N_s N_f}$, denotes the precoder matrix. After the precoders (and interleavers, which are explained further ahead), we consider that the frequency domain symbols are grouped into N -sized blocks and converted to the time domain by the Inverse Fast Fourier Transform (IFFT) blocks. A cyclic prefix (CP) with length larger than the delay spread of the channel is then added to each block before transmission. Note that each of the described PT-GSFIM symbols carries a total of bits per user given by

$$N_{bits} = N_{af} \left(\left\lfloor \log_2 \left(\frac{N_s}{N_a} \right) \right\rfloor + N_a \log_2 M \right) + \left\lfloor \log_2 \left(\frac{N_f}{N_{af}} \right) \right\rfloor. \quad (2)$$

To illustrate, if 64-QAM is employed in the case of Fig. 2, a total of 47 bits are mapped to a symbol, i.e., an average of 11.75 bits per subcarrier are being transmitted to each user.

III. TRANSMITTER STRUCTURE

A. PRECODER DESIGN

Assuming that the CP has been dropped and the time domain samples have been converted to the frequency domain through an N -point Fast Fourier Transform (FFT), the g^{th} PT-GSFIM symbol received by the u^{th} user can be written as

$$\mathbf{y}_g^u = \mathbf{H}_g^u \mathbf{x}_g + \mathbf{n}_g^u = \mathbf{H}_g^u \mathbf{F}_u \mathbf{s}_g^u + \mathbf{H}_g^u \sum_{\substack{j=1 \\ j \neq u}}^{N_u} \mathbf{F}_j \mathbf{s}_g^j + \mathbf{n}_g^u, \quad (3)$$

where $\mathbf{y}_g^u \in \mathbb{C}^{N_{rx} N_f \times 1}$, $\mathbf{n}_g^u \in \mathbb{C}^{N_{rx} N_f \times 1}$ represents noise and

$\mathbf{H}_g^u \in \mathbb{C}^{N_{rx} N_f \times N_{tx} N_f}$ is the frequency domain channel matrix between the BS and user u . Since we are considering an OFDM system, the channel matrix

follows a block diagonal structure, i.e.,

$$\mathbf{H}_g^u = \text{blkdiag} \left\{ \mathbf{H}_g^{u,1}, \dots, \mathbf{H}_g^{u,f}, \dots, \mathbf{H}_g^{u,N_f} \right\} \quad (4)$$

with $\mathbf{H}_g^{u,f} \in \mathbb{C}^{N_{rx} \times N_{tx}}$, $f = 1, \dots, N_f$. In the right side of expression (3), the second term represents interference between users. Each of the precoder matrices can be designed in order to eliminate all this MUI at all the receivers following the block diagonalization approach described in [43]. However, since many positions of the PT-GSFIM symbol, \mathbf{s}_g^u , are empty, with part of the information encoded in the equivalent channel impulse responses, we do not apply any of the power loading optimization approaches from [43] (even though power control can still be applied between users). Due to the block diagonal structure of \mathbf{H}_g^u , the signals transmitted on each subcarrier only generate interference on that subcarrier. This means that the precoder matrices \mathbf{F}_u can be designed independently for each subcarrier and follow a block diagonal structure, namely

$$\mathbf{F}_u = \text{blkdiag} \left\{ \mathbf{F}_{u,1}, \dots, \mathbf{F}_{u,N_f} \right\}, \quad (5)$$

with $\mathbf{F}_{u,f} \in \mathbb{C}^{N_{tx} \times N_s}$, $f = 1, \dots, N_f$. To cancel all MUI, each of the component precoder matrices $\mathbf{F}_{u,f}$ is designed so as to enforce $\mathbf{H}_g^{v,f} \mathbf{F}_{u,f} = \mathbf{0}$ for all $v \neq u$. Let us define the matrix that concatenates all the channel matrices between the BS and all users except user k , in subcarrier f as

$$\tilde{\mathbf{H}}_g^{u,f} = \left[\left(\mathbf{H}_g^{1,f} \right)^T \dots \left(\mathbf{H}_g^{u-1,f} \right)^T \left(\mathbf{H}_g^{u+1,f} \right)^T \dots \left(\mathbf{H}_g^{N_u,f} \right)^T \right]^T. \quad (6)$$

This matrix models the propagation of the signal targeted to user u , when it arrives at all the other receivers where it will constitute interference, therefore it can be referred to as the interference channel matrix. To avoid the generated interference one can design the respective precoder matrix $\mathbf{F}_{u,f}$

using an orthonormal basis of the null space of $\tilde{\mathbf{H}}_g^{u,f}$. First, we compute the respective singular value decomposition (SVD) given by

$$\tilde{\mathbf{H}}_g^{u,f} = \tilde{\mathbf{U}}_{u,f} \tilde{\mathbf{\Lambda}}_{u,f} \left[\tilde{\mathbf{V}}_{u,f}^{(1)} \quad \tilde{\mathbf{V}}_{u,f}^{(0)} \right]^H, \quad (7)$$

where $\tilde{\mathbf{U}}_{u,f}$ is the matrix with the left-singular vectors, $\tilde{\mathbf{\Lambda}}_{u,f}$ is a rectangular diagonal matrix with the decreasing nonzero singular values, $\tilde{\mathbf{V}}_{u,f}^{(1)}$ contains the right singular vectors corresponding to the nonzero singular values and $\tilde{\mathbf{V}}_{u,f}^{(0)}$ contains the remainder right singular vectors. Since the columns of $\tilde{\mathbf{V}}_{u,f}^{(0)}$ span the null space of $\tilde{\mathbf{H}}_g^{u,f}$, to guarantee zero MUI each precoder matrix can be set as $\mathbf{F}_{u,f} = \tilde{\mathbf{V}}_{u,f}^{(0)} \left[\cdot, 1 : N_s \right]$. The resulting signal at each receiver then reduces to

$$\mathbf{y}_g^u = \hat{\mathbf{H}}_g^u \mathbf{s}_g^u + \mathbf{n}_g^u, \quad (8)$$

where $\hat{\mathbf{H}}_g^u = \mathbf{H}_g^u \mathbf{F}_u$ is the equivalent SU channel and no MUI exists. Bearing in mind that both \mathbf{H}_g^u and \mathbf{F}_u have a block diagonal structure, $\hat{\mathbf{H}}_g^u$ will also be block diagonal. According to the presented system model, only a few positions of \mathbf{s}_g^u will tend to be nonzero. Therefore, \mathbf{s}_g^u is a sparse vector and this sparsity can be exploited for reducing the transmitted power. To accomplish this, we can apply the same approach that we discussed in [34] for virtual GSM transmissions, where an alternative signal is generated which minimizes the transmitted power while ensuring that the signals arriving at the receivers are identical to the original ones.

Although in this paper we are assuming fully-digital precoders, it is possible to reduce the implementation complexity of the proposed solution by adopting an hybrid precoder design where the signal processing is split into two separate parts: a reduced digital one and an analog part which is typically supported on analog phase shifters. The hybrid design can be simply achieved through direct approximation of the fully-digital precoder matrices using the product of smaller digital precoder matrices and an analog precoder matrix (the same for all subcarriers), as described in [41] and [42]. As also explained in [42], an additional cancellation step should be included in this approximation-based approach in order to remove residual inter-user interference.

B. COMPLEX ROTATION MATRICES

According to the signal model (8) where $\hat{\mathbf{H}}_g^u$ is block diagonal, each individual GSM symbol vector composing a PT-GSFIM is subject to the effect of the channel on a single subcarrier. In order to exploit the inherent diversity of frequency selective channels that are typical in mobile propagation environments, we can apply SSD techniques. SSD was originally proposed in [37] and can be used in an OFDM system so as to improve its performance without requiring additional power or bandwidth. A simple way to implement SSD is to resort to CRM so as to associate each GSM symbol that makes up the main PT-GSFIM symbol to different subcarriers instead of only one. The process of applying CRM consists in working with a rotated super-symbol. In this case we apply the spreading over the frequency direction only, i.e., all the different spatial components are subject to the same rotation. The rotated length- $N_s N_f$ super-symbol for each user u can be expressed as

$$\tilde{\mathbf{s}}_g^u = (\mathbf{A}_{N_f} \otimes \mathbf{I}_{N_s}) \cdot \mathbf{s}_g^u, \quad (9)$$

and the transmitted signal becomes (1)

TABLE 1. Algorithm 1: OB-MMSE-based GSFIM detector

```

1: Input:  $\mathbf{y}_g^u, \tilde{\mathbf{H}}_g^u, \mathbf{I}, V_{th} = 2N_{rx}\sigma^2$ .
2:  $z_j = \frac{1}{(\tilde{\mathbf{H}}_g^u[:,j])^H \tilde{\mathbf{H}}_g^u[:,j]} (\tilde{\mathbf{H}}_g^u[:,j])^H \mathbf{y}_g^u, j = 1, \dots, N_f N_s$ 
3:  $\mathbf{w} = [w_1, w_2, \dots, w_{N_{comb}}]^T$ ,
    $w_i = \sum_{k \in I_i} z_k^2, I_i \in \mathbf{I}, i = 1, 2, \dots, N_{comb}$ 
4:  $[k_1, k_2, \dots, k_{N_{comb}}] = \arg \text{sort}(\mathbf{w})$ .
5:  $j = 1$ .
6: While  $j \leq N_{comb}$  do
7:  $\hat{\mathbf{s}}_g^u[\bar{I}_j] \leftarrow 0$ 
    $\hat{\mathbf{s}}_g^u[I_j] \leftarrow \Pi_{A_{N_{of} N_u}} \left( \left( (\tilde{\mathbf{H}}_g^u[:, I_j])^H \tilde{\mathbf{H}}_g^u[:, I_j] + 2\sigma^2 \mathbf{I}_{N_{of} N_s} \right)^{-1} \right.$ 
    $\left. \times (\tilde{\mathbf{H}}_g^u[:, I_j])^H \mathbf{y}_g^u \right)$ 
    $d_j = \|\mathbf{y}_g^u - \tilde{\mathbf{H}}_g^u[:, I_j] \cdot \hat{\mathbf{s}}_g^u\|_F^2$ .
8: If  $d_j < V_{th}$  then
9:  $\hat{\mathbf{s}}_g^u[\bar{I}_j] \leftarrow 0, \hat{\mathbf{s}}_g^u[I_j] \leftarrow \hat{\mathbf{s}}_g^u[I_j]$ , break
10: else
11:  $j = j + 1$ ;
12: end if
13: end while
14: If  $j > N_{comb}$  then
15:  $m = \arg \min_{j \in \{1, \dots, N_{comb}\}} d_j$ ,
    $\hat{\mathbf{s}}_g^u[\bar{I}_m] \leftarrow 0, \hat{\mathbf{s}}_g^u[I_m] \leftarrow \hat{\mathbf{s}}_g^u[I_m]$ .
16: end if
17: Output:  $\hat{\mathbf{s}}_g^u$ .

```

$$\mathbf{x}_g = \sum_{u=1}^{N_u} \mathbf{F}_u (\mathbf{A}_{N_f} \otimes \mathbf{I}_{N_s}) \mathbf{s}_g^u = \mathbf{F} (\mathbf{I}_{N_u} \otimes (\mathbf{A}_{N_f} \otimes \mathbf{I}_{N_s})) \mathbf{s}_g \quad (10)$$

Matrix $\mathbf{A}_{N_f} \in \mathbb{C}^{N_f \times N_f}$ can be selected from the family of orthonormal complex matrices (OCRM), which are defined as

$$\mathbf{A}_{M_{CRM}} = \begin{cases} \begin{bmatrix} e^{j\varphi} & je^{-j\varphi} \\ je^{j\varphi} & e^{-j\varphi} \end{bmatrix} / |\mathbf{A}_2|^{1/2}, M_{CRM} = 2 \\ |\mathbf{A}_2| = \det(\mathbf{A}_2) = 2 \\ \begin{bmatrix} \mathbf{A}_{M_{CRM}/2} & \mathbf{A}_{M_{CRM}/2} \\ \mathbf{A}_{M_{CRM}/2} & -\mathbf{A}_{M_{CRM}/2} \end{bmatrix} / |\mathbf{A}_{M_{CRM}}|^{1/M_{CRM}}, M_{CRM} > 2 \end{cases} \quad (11)$$

where $M_{CRM} = 2^n$ ($n \geq 1$), $|\mathbf{A}_{M_{CRM}}| = \det(\mathbf{A}_{M_{CRM}})$ and φ being the rotation angle [38]. To better reap the benefits of frequency diversity, a symbol interleaver (subcarrier-wise) should be employed so that the effective group of subcarriers allocated to each PT-GSFIM symbol is spread far apart within the overall bandwidth.

TABLE 2. Algorithm 2: multiblock sMMP detector

```

1: Input:  $\mathbf{y}_g^u, \tilde{\mathbf{H}}_g^u, L$ ;
2: Initial:  $k = 0, \mathbf{r}_1^{(0)} = \mathbf{y}_g^u, T^{(0)} = \emptyset$ ;
3: While  $k < N_a N_{af}$  do
4:  $k = k + 1, q = 0, T^{(k)} = \emptyset$ ;
5: For  $i = 1$  to  $|T^{(k-1)}|$  do
     $[t_1, t_2, \dots, t_L] = \text{argsort} \left( \left\| \left( \tilde{\mathbf{H}}_g^u \right)^H \mathbf{r}_i^{(k-1)} \right\| \right)$ . (sort  $L$  best
    indices)
6: For  $j = 1$  to  $L$  do
7:  $t_{imp} = t_i^{(k-1)} \cup \{t_j\}$ ;
8: If  $(t_{imp} \notin T^{(k)})$  and (number of indices in  $t_{imp}$  inside
    subcarrier block of  $t_j$  is  $\leq N_a$ ) and (number of
    subcarrier blocks with indices in  $t_{imp}$  is  $\leq N_{af}$ ) then
9:  $q = q + 1$ ;
10:  $t_q^{(k)} = t_{imp}$ ;
11:  $T^{(k)} = T^{(k)} \cup \{t_q^{(k)}\}$ ;
12:  $(\tilde{\mathbf{s}}_g^u)^{(k),(q)} \leftarrow 0$ 
     $(\tilde{\mathbf{s}}_g^u[t_q^{(k)}])^{(k),(q)} \leftarrow \prod_{A_q^{(k)}} \left( \left( \left( \tilde{\mathbf{H}}_g^u[:, t_q^{(k)}] \right)^H \tilde{\mathbf{H}}_g^u[:, t_q^{(k)}] \right)^{-1} \right.$ 
     $\left. \times \left( \tilde{\mathbf{H}}_g^u[:, t_q^{(k)}] \right)^H \mathbf{y}_g^u \right)$ 
13:  $\mathbf{r}_q^{(k)} = \mathbf{y}_g^u - \tilde{\mathbf{H}}_g^u[:, t_q^{(k)}] (\tilde{\mathbf{s}}_g^u)^{(k),(q)}$ .
14: end if
15: end for
16: end for
17: end while
18:  $q^* = \arg \min_q \left\| \mathbf{r}_q^{(N_a N_{af})} \right\|_2$ ;
19: Output:  $\hat{\mathbf{s}}_g^u = (\tilde{\mathbf{s}}_g^u)^{(N_a N_{af})}(q^*)$ .

```

IV. SIGNAL DETECTION

A. PROPOSED ALGORITHMS

According to the transmitted structure described in the previous section, where a precoder removes all the MUI, the receiver only needs to implement SU detection. If no CRM is employed, the block diagonal structure of $\hat{\mathbf{H}}_g^u$ in (8) allows us to simplify the detection process. In this case the techniques proposed for GSM, such as the ordered block minimum mean-squared error (OB-MMSE) detector from [44], the multipath matching pursuit with slicing (sMMP) from [45] or the alternating direction method of multipliers (ADMM) based detector from [39], can be directly applied to PT-GSFIM by simply preceding those algorithms by an active subcarrier detection step. This initial detection can be accomplished by working with matrix $\mathbf{Y} = \text{vec}_{N_{af} \times N_f}^{-1}(\mathbf{y}_g^u)$, where $\text{vec}_{m \times n}^{-1}(\cdot)$ denotes the inverse of the vectorization operator

TABLE 3. Algorithm 3: ADMM-based GSFIM detector

```

1: Input:  $\mathbf{x}^0, \mathbf{r}^0, \mathbf{z}^0, \mathbf{u}^0, \mathbf{v}^0, \mathbf{w}^0, \tilde{\mathbf{H}}_g^u, \mathbf{y}_g^u, \rho_x, \rho_r, \rho_z, Q$ 
2:  $f_{best} = \infty$ .
3:  $\Phi \leftarrow \left( \left( \tilde{\mathbf{H}}_g^u \right)^H \tilde{\mathbf{H}}_g^u + (\rho_x + \rho_r + \rho_z) \mathbf{I}_{N_f N_s} \right)^{-1}$ .
4: for  $t = 0, 1, \dots, Q-1$  do
5:  $\mathbf{s}^{(t+1)} \leftarrow \Phi \left( \left( \tilde{\mathbf{H}}_g^u \right)^H \mathbf{y}_g^u + \rho_x (\mathbf{x}^{(t)} - \mathbf{u}^{(t)}) + \rho_r (\mathbf{r}^{(t)} - \mathbf{v}^{(t)}) \right.$ 
     $\left. + \rho_z (\mathbf{z}^{(t)} - \mathbf{w}^{(t)}) \right)$ .
6:  $(\mathbf{x}^{(t+1)}) \leftarrow \Pi_{S_0} \left( (\mathbf{s}^{(t+1)}) + (\mathbf{u}^{(t+1)}) \right)$ .
7:  $\mathbf{r}^{(t+1)} \leftarrow \Pi_J \left( (\mathbf{s}^{(t+1)}) + \mathbf{v}^{(t)} \right)$ .
8:  $\mathbf{z}^{(t+1)} \leftarrow \Pi_{A_0 N_{af}} \left( (\mathbf{s}^{(t+1)}) + \mathbf{w}^{(t)} \right)$ .
9:  $I \leftarrow \text{supp}(\mathbf{x}^{(t+1)}) \cap \text{supp}(\mathbf{r}^{(t+1)})$ .
10: If  $t = Q-1$  then
11:  $\tilde{\mathbf{s}}_g^u[\bar{I}] \leftarrow 0$ 
     $\tilde{\mathbf{s}}_g^u[I] \leftarrow \prod_{A_q^{N_{af} N_a}} \left( \left( \left( \tilde{\mathbf{H}}_g^u[:, I] \right)^H \tilde{\mathbf{H}}_g^u[:, I] + 2\sigma^2 \mathbf{I}_{N_{af} N_s} \right)^{-1} \right.$ 
     $\left. \times \left( \tilde{\mathbf{H}}_g^u[:, I] \right)^H \mathbf{y}_g^u \right)$ 
    (polishing)
12: else
13:  $\tilde{\mathbf{s}}_g^u[\bar{I}] \leftarrow 0, \tilde{\mathbf{s}}_g^u[I] \leftarrow \mathbf{z}^{(t+1)}[I]$ .
14: end if
15: If  $f(\tilde{\mathbf{s}}_g^u) < f_{best}$  then
16:  $\hat{\mathbf{s}}_g^u[\bar{I}] \leftarrow 0, \hat{\mathbf{s}}_g^u[I] \leftarrow \tilde{\mathbf{s}}_g^u[I]$ .
17:  $f_{best} = f(\tilde{\mathbf{s}}_g^u)$ .
18: end if
19:  $\mathbf{u}^{(t+1)} \leftarrow \mathbf{u}^{(t)} + \mathbf{s}^{(t+1)} - \mathbf{x}^{(t+1)}$ .
20:  $\mathbf{v}^{(t+1)} \leftarrow \mathbf{v}^{(t)} + \mathbf{s}^{(t+1)} - \mathbf{r}^{(t+1)}$ .
21:  $\mathbf{w}^{(t+1)} \leftarrow \mathbf{w}^{(t)} + \mathbf{s}^{(t+1)} - \mathbf{z}^{(t+1)}$ .
22: end for
23: Output:  $\hat{\mathbf{s}}_g^u$ .

```

$\text{vec}(\cdot)$, which is defined as $\text{vec}_{m \times n}^{-1}: \mathbb{C}^{mn \times 1} \rightarrow \mathbb{C}^{m \times n}$ such that $\text{vec}_{m \times n}^{-1}(\text{vec}(\mathbf{X})) = \mathbf{X}$ for any $\mathbf{X} \in \mathbb{C}^{m \times n}$. The indexes of the N_{af} columns of \mathbf{Y} with larger Euclidean norm and which also match a valid active subcarrier combination are selected as those which will be processed. Each of these columns, together with the corresponding channel matrix $\tilde{\mathbf{H}}_g^{u,f}$ are then used as inputs for the GSM detection algorithms mentioned previously, which will estimate the active antennas and APM symbols for that individual subcarrier. In the case of CRM being employed, a different approach must be adopted since each GSM symbol is spread over several subcarriers. In this case the received signal can be written as

$$\begin{aligned} \mathbf{y}_g^u &= \hat{\mathbf{H}}_g^u \tilde{\mathbf{s}}_g^u + \mathbf{n}_g^u \\ &= \tilde{\mathbf{H}}_g^u \mathbf{s}_g^u + \mathbf{n}_g^u, \end{aligned} \quad (12)$$

where $\tilde{\mathbf{H}}_g^u = \hat{\mathbf{H}}_g^u (\mathbf{A}_{N_f} \otimes \mathbf{I}_{N_s})$ is the overall equivalent

channel “seen” by user u for the g^{th} PT-GSFIM symbol. In the following we propose three different algorithms that can be adopted for the detector. The first one is a generalization of the OB-MMSE from [44] and is shown in Algorithm 1. The main differences in the presented algorithm reside in step 2 and 3 which have to be computed over both the spatial and frequency dimensions and require working with the equivalent channel matrix for the whole PT-GSFIM symbol $\tilde{\mathbf{H}}_g^u$. Furthermore, instead of working with the possible antenna combinations, the extended algorithm considers the possible joint combination of active antennas and subcarriers, and sorts these $N_{\text{comb}} = 2^{\left\lfloor \log_2 \left(\binom{N_s}{N_a} \right) + \left\lfloor \log_2 \left(\binom{N_f}{N_{af}} \right) \right\rfloor \right\rfloor}$ combinations according to the measured reliability (step 4). In this case we use \mathbb{I} to denote the set of possible supports of \mathbf{s}_g^u , i.e., $\mathbb{I} = \{I_1, \dots, I_{N_{\text{comb}}}\}$ with $I_i = \text{supp}(\mathbf{s}_g^u)$ and $i = 1, 2, \dots, N_{\text{comb}}$. Each of these possible antennas and subcarrier combinations is processed using a block MMSE detector applied with a fixed support (step 7). This step involves an element-wise projection over set \mathcal{A}_0 , denoted as $\Pi_{\mathcal{A}_0^{N_s N_f}}(\cdot)$.

The algorithm runs until the threshold condition is satisfied (step 8), otherwise it will end only after all the combinations are processed in which case the best estimate will be selected.

The second approach is based on the sMMP from [45] which we extend for a multiblock structured signal such as GSFIM. The resulting detector is shown in Algorithm 2. It follows a greedy strategy similar to the orthogonal matching pursuit (OMP) algorithm [46], but instead of performing a tree search along a single path, it uses parallel search to build a list of candidates and selects the best estimate at the end. The algorithm comprises a total of $N_a N_{af}$ iterations. Defining L as the number of child candidates, in each iteration, the algorithm expands each of the previous iteration candidates into L newer candidates with an additional component that is selected so as to maximize the correlation with the previous residual. While the total number of candidates can keep increasing along the iterations, some of them may overlap and can be removed. In order to adapt the algorithm to GSFIM signals, it has to work jointly with the spatial and frequency dimensions and with the equivalent channel matrix for the whole PT-GSFIM symbol. Furthermore, step 8 has to be modified so that a new candidate is only added to the list if the selected component does not exceed a total of N_a nonzero components in that subcarrier, nor if it results in a total of subcarriers with nonzero components exceeding N_{af} .

The third detector is based on the application of ADMM, following an approach that is an extension of the one adopted in [39] for GSM. First, we formulate the maximum likelihood detector (MLD) detection problem as

$$\min_{\mathbf{s}_g^u} f(\mathbf{s}) \triangleq \|\mathbf{y}_g^u - \tilde{\mathbf{H}}_g^u \mathbf{s}_g^u\|_2^2 \quad (13)$$

$$\text{subject to } \mathbf{s}_g^u \in \mathcal{A}_0^{N_s N_f} \quad (14)$$

$$\text{supp}(\mathbf{s}_g^{u,i}) \in \mathbb{S}_0, \quad i = 1, \dots, N_f \quad (15)$$

$$\text{supp}(\mathbf{s}_g^u) \in \mathbb{I}, \quad (16)$$

where \mathcal{A}_0 represents the complex valued APM constellation set, including symbol 0. \mathbb{S}_0 represents the set of possible supports of $\mathbf{s}_g^{u,i}$ according to possible GSM symbols i.e., it is the set with valid (virtual) active antenna combinations including the null support (all antennas inactive). \mathbb{I} denotes the set of possible supports of $\mathbf{s}_g^{u,i}$ according to the valid active subcarrier combinations. It is important to highlight that since this problem formulation is different from the original one presented in [39], mostly due to constraints (15) and (16), the whole algorithm has to be derived again, as will be done next. Introducing auxiliary variables, \mathbf{x} , \mathbf{r} , \mathbf{z} , we can integrate constraints (13)-(14) into the objective function and at the same time make it separable by rewriting the MLD problem as

$$\min_{\mathbf{s}_g^u} f(\mathbf{s}_g^u) \triangleq \|\mathbf{y}_g^u - \tilde{\mathbf{H}}_g^u \mathbf{s}_g^u\|_2^2 + \sum_{i=1}^{N_f} I_{\mathbb{S}_0}(\mathbf{x}^i) + I_{\mathbb{I}}(\mathbf{r}) + I_{\mathcal{A}_0^{N_s N_f}}(\mathbf{z}) \quad (17)$$

$$\text{subject to } \mathbf{s}_g^{u,i} = \mathbf{x}^i, \quad i = 1, \dots, N_f \quad (18)$$

$$\mathbf{s}_g^u = \mathbf{r} \quad (19)$$

$$\mathbf{s}_g^u = \mathbf{z}. \quad (20)$$

Working with \mathbf{s} instead of \mathbf{s}_g^u in the expressions, the augmented Lagrangian function (ALF) for this problem is given by

$$\begin{aligned} L_{\rho_x, \rho_r, \rho_z}(\mathbf{s}, \mathbf{x}, \mathbf{r}, \mathbf{z}, \mathbf{u}, \mathbf{v}, \mathbf{w}) = & \|\mathbf{y}_g^u - \tilde{\mathbf{H}}_g^u \mathbf{s}\|_2^2 + \sum_{i=1}^{N_f} I_{\mathbb{S}_0}(\mathbf{x}^i) + I_{\mathbb{I}}(\mathbf{r}) \\ & + I_{\mathcal{A}_0^{N_s N_f}}(\mathbf{z}) + \rho_x \left(\|\mathbf{s} + \mathbf{u} - \mathbf{x}\|_2^2 - \|\mathbf{u}\|_2^2 \right) + \rho_r \left(\|\mathbf{s} + \mathbf{v} - \mathbf{r}\|_2^2 - \|\mathbf{v}\|_2^2 \right) \\ & + \rho_z \left(\|\mathbf{s} + \mathbf{w} - \mathbf{z}\|_2^2 - \|\mathbf{w}\|_2^2 \right) \end{aligned} \quad (21)$$

where ρ_x, ρ_r, ρ_z are penalty parameters and $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{C}^{N_s N_f \times 1}$ are scaled dual variables. We then apply the dual ascent method to the dual problem which requires us to iteratively accomplish the independent minimization of the augmented Lagrangian over \mathbf{s} , \mathbf{x} , \mathbf{r} and \mathbf{z} combined with the dual variables update through gradient ascent. This results in the following sequence of steps:

• *Step 1: Minimization of the ALF over \mathbf{s} .* This step consists in solving

$$\begin{aligned} \mathbf{s}^{(t+1)} = \min_{\mathbf{s}} \{ & \|\mathbf{y}_g^u - \tilde{\mathbf{H}}_g^u \mathbf{s}\|_2^2 + \rho_x \left(\|\mathbf{s} + \mathbf{u} - \mathbf{x}\|_2^2 - \|\mathbf{u}\|_2^2 \right) \\ & + \rho_r \left(\|\mathbf{s} + \mathbf{v} - \mathbf{r}\|_2^2 - \|\mathbf{v}\|_2^2 \right) + \rho_z \left(\|\mathbf{s} + \mathbf{w} - \mathbf{z}\|_2^2 - \|\mathbf{w}\|_2^2 \right) \}. \end{aligned} \quad (22)$$

A closed-form solution to this problem can be obtained from $\nabla_{\mathbf{s}} L_{\rho_x, \rho_r, \rho_z}(\mathbf{s}, \mathbf{x}, \mathbf{r}, \mathbf{z}, \mathbf{u}, \mathbf{v}, \mathbf{w}) = 0$ which, leads to

$$\mathbf{s}^{(t+1)} = \Phi \left(\left(\tilde{\mathbf{H}}_g^u \right)^H \mathbf{y}_g^u + \rho_x \left(\mathbf{x}^{(t)} - \mathbf{u}^{(t)} \right) + \rho_r \left(\mathbf{r}^{(t)} - \mathbf{v}^{(t)} \right) + \rho_z \left(\mathbf{z}^{(t)} - \mathbf{w}^{(t)} \right) \right) \quad (23)$$

with

$$\Phi = \left(\left(\tilde{\mathbf{H}}_g^u \right)^H \tilde{\mathbf{H}}_g^u + (\rho_x + \rho_r + \rho_z) \mathbf{I}_{N_f N_s} \right)^{-1}. \quad (24)$$

• *Step 2: Minimization of the ALF over \mathbf{x} .* This step solves

$$\mathbf{x}^{(t+1)} = \min_{\mathbf{x}} \left\{ \sum_{i=1}^{N_f} I_{\mathbb{S}_0}(\mathbf{x}^i) + \rho_x \left(\|\mathbf{s} + \mathbf{u} - \mathbf{x}\|_2^2 - \|\mathbf{u}\|_2^2 \right) \right\}. \quad (25)$$

The minimization involves an indicator function applied to each individual subcarrier and can be obtained as

$$\left(\mathbf{x}^i \right)^{(t+1)} = \Pi_{\mathbb{S}_0} \left(\left(\mathbf{s}^i \right)^{(t+1)} + \left(\mathbf{u}^i \right)^{(t+1)} \right), \quad i = 1, \dots, N_f, \quad (26)$$

i.e., as the projection over set \mathbb{S}_0 , $\Pi_{\mathbb{S}_0}(\cdot)$. This projection can be performed by selecting the N_a largest magnitude elements whose indices also match a valid active antenna combination.

• *Step 3: Minimization of the ALF over \mathbf{r} .* The third step solves

$$\mathbf{r}^{(t+1)} = \min_{\mathbf{r}} \left\{ I_{\mathbb{J}}(\mathbf{r}) + \rho_r \left(\|\mathbf{s} + \mathbf{v} - \mathbf{r}\|_2^2 - \|\mathbf{v}\|_2^2 \right) \right\}. \quad (27)$$

which also involves an indicator function and whose solution corresponds to the projection over set \mathbb{J} , namely

$$\mathbf{r}^{(t+1)} = \Pi_{\mathbb{J}} \left(\mathbf{s}^{(t+1)} + \mathbf{v}^{(t)} \right). \quad (28)$$

This projection can be accomplished by selecting the N_{af} columns of matrix $\mathbf{R} = \text{vec}_{N_s \times N_f}^{-1} \left(\mathbf{s}^{(t+1)} + \mathbf{v}^{(t)} \right)$ with larger Euclidean norm and which also match a valid active subcarrier combination, while nulling all the others. The result is then simply set as $\mathbf{r} = \text{vec}(\mathbf{R})$.

• *Step 4: Minimization of the ALF over \mathbf{z} .* The fourth step solves

$$\mathbf{z}^{(t+1)} = \min_{\mathbf{z}} \left\{ I_{\mathcal{A}_0^{N_s N_f}}(\mathbf{z}) + \rho_z \left(\|\mathbf{s} + \mathbf{w} - \mathbf{z}\|_2^2 - \|\mathbf{w}\|_2^2 \right) \right\}. \quad (29)$$

whose solution can also be obtained as a projection, but in this case over set $\mathcal{A}_0^{N_s N_f}$

$$\mathbf{z}^{(t+1)} = \Pi_{\mathcal{A}_0^{N_s N_f}} \left(\mathbf{s}^{(t+1)} + \mathbf{w}^{(t)} \right). \quad (30)$$

It can be implemented elementwise using simple rounding to the closest element in \mathcal{A}_0 .

• *Step 5: Dual variable update.* The last main step updates the dual variables according to

$$\mathbf{u}^{(t+1)} = \mathbf{u}^{(t)} + \mathbf{s}^{(t+1)} - \mathbf{x}^{(t+1)}, \quad (31)$$

TABLE 4. Number of real flops for the different GSFIM detectors

Detector	Complexity
MLD	$(8N_{rx}N_fN_aN_{af} + 4N_fN_{rx} - 1)N_{comb}M^{N_aN_{af}}$
GSFIM-OB-MMSE	$12N_{rx}N_fN_aN_{af} + 2N_sN_f$ $+ N_{comb} \left(N_aN_{af} - 1 + 4(N_aN_{af})^3 + 12(N_aN_{af})^2 N_{rx}N_f + \right.$ $\left. + 7(N_aN_{af})^2 + 14N_aN_{af}N_{rx}N_f + 4N_{rx}N_f - 1 \right)$
GSFIM-sMMP ($L=1$)	$(8N_sN_f^2N_{rx} + N_sN_f)N_aN_{af} + (5N_{rx}N_f - 2)$ $+ \sum_{k=1}^{N_s} (4k^3 + k^2(4N_{rx}N_f + 15) + k(20N_{rx}N_f - 5))$
GSFIM-sMMP ($L>1$)	$(8N_sN_f^2N_{rx} + N_sN_f) \left((1 - L^{N_aN_{af}}) / (1 - L) + (5N_{rx}N_f - 2) L^{N_aN_{af}} \right)$ $+ \sum_{k=1}^{N_s} (4k^3 + k^2(4N_{rx}N_f + 15) + k(20N_{rx}N_f - 5)) L^k$
GSFIM-ADMM	$4(N_sN_f)^3 + (N_sN_f)^2(4N_{rx}N_f + 7) + N_sN_f(12N_{rx}N_f - 1)$ $+ Q(8(N_sN_f)^2 + 59N_sN_f - 6N_f) + 4(N_aN_{af})^3$ $+ (N_aN_{af})^2(4N_{rx}N_f + 15) + N_aN_{af}(12N_{rx}N_f - 3)$

$$\mathbf{v}^{(t+1)} = \mathbf{v}^{(t)} + \mathbf{s}^{(t+1)} - \mathbf{r}^{(t+1)}, \quad (32)$$

$$\mathbf{w}^{(t+1)} = \mathbf{w}^{(t)} + \mathbf{s}^{(t+1)} - \mathbf{z}^{(t+1)}. \quad (33)$$

The overall sequence of steps is summarized in Algorithm 3, where \bar{I} denotes the complement of the support set I (i.e., $\bar{I} = \{1, \dots, N_s N_f\} \setminus I$) and Q is the maximum number of iterations. It is important to note that the projections of both step 6 and step 7 can be simplified to simple cardinality-based projections if set \mathbb{S}_0 or set \mathbb{J} are very large, i.e., there is a big number of valid active antenna combinations and/or valid active subcarrier combinations. In this case, projection $\Pi_{\mathbb{S}_0}(\cdot)$ is obtained by zeroing the $N_s - N_a$ smallest magnitude elements whereas $\Pi_{\mathbb{J}}(\cdot)$ is computed by zeroing the $N_f - N_{af}$ columns of \mathbf{R} with smaller Euclidean norm. Algorithm 3 includes a final polishing procedure (step 11) where the projected MMSE estimate is computed for the reduced problem with the support fixed. Regarding the initialization of the algorithm, the warm and random start procedures described in [39] can be directly extended. Note that compared with [39], the proposed ADMM algorithm works with the joint spatial and frequency dimensions and besides several differences in previously existing steps, it also includes two main additional steps, namely: an additional projection (step 7) and an additional dual variable update (step 20).

B. COMPUTATIONAL COMPLEXITIES

In Table 4, we present the respective worst-case complexities in number of real floating-point operations (flops) of the three proposed algorithms alongside the maximum likelihood detector. In the case of GSFIM-OB-MMSE, N_{comb} corresponds to the number of possible supports of \mathbf{s}_g^u and is given by

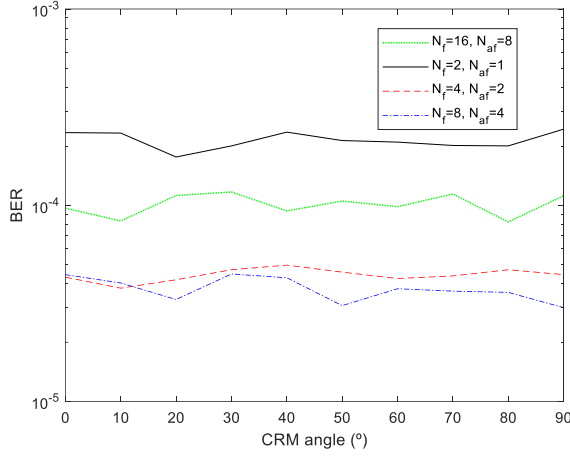


FIGURE 3. Comparison of CRM rotation angle ϕ of PT-GSFIM in a MU scenario with $N_u=4$, $N_{tx}=8N_u$, $N_f=4$, $N_{af}=3$, $N_s=8$, $N_a=1$, $N_{rx}=4$, QPSK (SNR=0dB).

$$N_{comb} = 2^{N_{af} \left[\log_2 \left(\frac{N_s}{N_a} \right) \right] + \left[\log_2 \left(\frac{N_f}{N_{af}} \right) \right]}. \quad (34)$$

In the presented complexities it is assumed that no earlier termination of the GSFIM-ADMM algorithm occurs. Furthermore, to arrive at the expressions we considered that each complex-valued sum, multiplication and squared absolute value computation counts as 2, 6 and 3 real flops. We also took into account that some matrix operations do not need to be repeated after the first iteration (ex: matrix multiplication $(\tilde{\mathbf{H}}_g^u)^H \mathbf{y}_g^u$ in step 5 of GSFIM-ADMM algorithm). In section V we use the presented complexity expressions to compare the three algorithms in different scenarios.

V. PERFORMANCE RESULTS

In this section, the performance of the proposed PT-GSFIM scheme with the different detection algorithms are assessed and compared against other MU-MIMO systems. Monte Carlo simulations were run according to the system model presented previously. The adopted channel model was the Extended Typical Urban model (ETU) [47] (similar conclusions could be reached for other severely time-dispersive channels). It is assumed that all the channel coefficients are independently drawn according to a complex Gaussian distribution $\mathcal{CN}(0,1)$, and that all users experience the same path-loss. The symbols $s_{g,f}^{u,i}$ transmitted in the active positions (subcarrier and antenna) are randomly selected from an M -QAM constellation with equal probabilities and with $E \left[\left| s_{g,f}^{u,i} \right|^2 \right] = 1$. To understand the impact of the CRM rotation angle ϕ on the performance of PT-GSFIM, Fig. 3 shows the BER results obtained for different sizes of N_f . We assume a scenario with $N_u=4$, $N_{tx}=8N_u$, $N_s=8$, $N_a=1$, $N_{rx}=4$, SNR=0 dB (per user), QPSK and GSFIM-ADMM detector. It can be observed that in general the results do not exhibit significant sensitivity to the

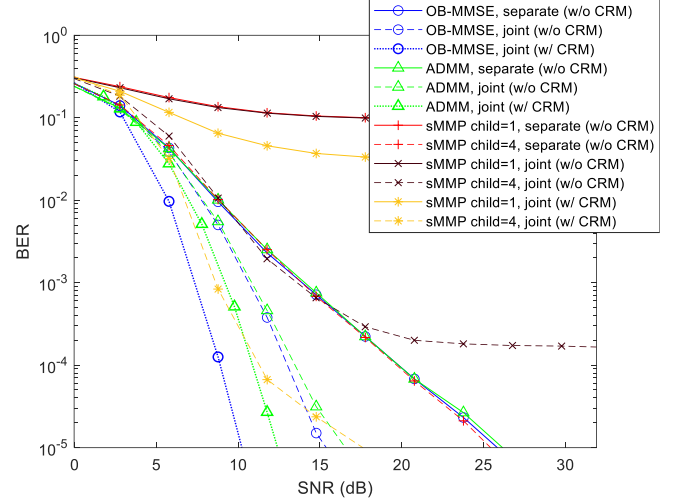


FIGURE 4. BER performance of PT-GSFIM in a MU scenario with $N_u=4$, $N_{tx}=5N_u$, $N_f=4$, $N_{af}=3$, $N_s=5$, $N_a=2$, $N_{rx}=5$, QPSK (5.75 bpcu per user).

rotation angle (variations are due to simulation accuracy). What is relevant is the spreading of the PT-GSFIM symbols over several subcarriers through the use of CRM combined with the interleaver in order to obtain additional diversity. It is important to note that even though this figure can show better BERs for smaller N_f , larger N_f tend to achieve better results at higher signal-to-noise ratios (SNRs) as we will show further ahead. Next, we evaluate the performance of the different receivers and the impact of the use of CRM in PT-GSFIM, with a rotation angle ϕ of 30° and reminding that the CRM is applied to each set of N_f subcarriers. We assume a base scenario with $N_u=4$, $N_{tx}=5N_u$, $N_f=4$, $N_{af}=3$, $N_s=5$, $N_a=2$ and $N_{rx}=5$, which corresponds to a SE of 5.75 bits per channel use (bpcu) and per user. In Fig. 4, three different BER curves are shown for each of the proposed detectors. The “separate” curve corresponds to the direct application of the GSM detector preceded by an active subcarrier detection step. In this case no CRM is applied. The curves where “joint” appears in the legend consider the joint detection of the GSFIM symbols using Algorithm 1-3. For this case, curves with and without CRM are included. It can be observed that the joint detection clearly provides substantial gains over the “separate” approach. The OB-MMSE algorithm achieved the best results followed by ADMM (with 10% of penalty). As for sMMP, it can be seen that it clearly underperforms. It is important to note that sMMP relies on consecutively finding, through correlation, the columns of the channel matrix that are more closely related the residuals of the constructed candidates (step 5 of algorithm 2). Since the cross-correlation between the different columns of the channel matrix has no guarantee of being close to 0 for the scenario of Fig. 4 (channel matrix is random and squared), this will tend to result in errors even if no noise exists, causing irreducible BER floors. Increasing the number of child nodes, L , combined with CRM can improve the results, lowering the error floor. In fact, in general, the use of

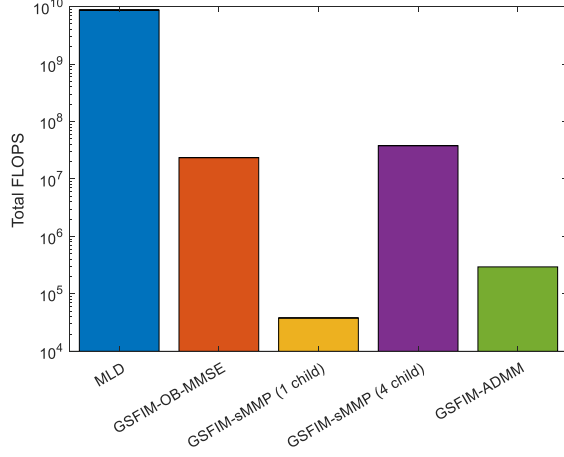


FIGURE 5. Complexity of the different GSFIM detector algorithms for the MU scenario with $N_u=4$, $N_{tx}=5N_u$, $N_r=4$, $N_{af}=3$, $N_s=5$, $N_a=2$, $N_{rx}=5$, QPSK (5.75 bpcu per user).

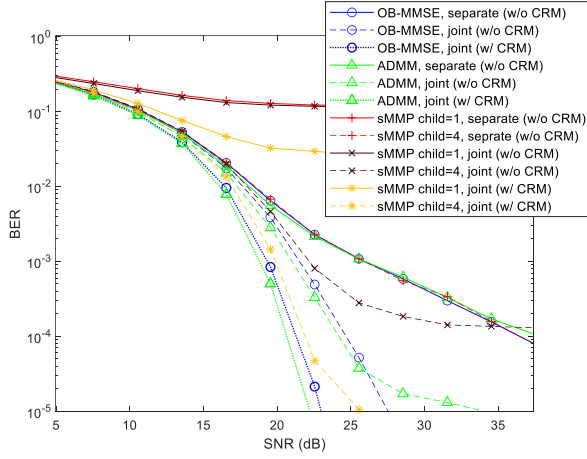


FIGURE 6. BER performance of PT-GSFIM in a MU scenario with $N_u=4$, $N_{tx}=5N_u$, $N_r=4$, $N_{af}=3$, $N_s=5$, $N_a=2$, $N_{rx}=5$, 64-QAM (11.75 bpcu per user).

CRM has a large positive impact on the performance of the studied receivers. For example, when using OB-MMSE, CRM provides a gain of 5 dB at a BER of 10^{-5} . It is important to note that even though OB-MMSE with CRM gives the best results with QPSK, it requires a higher computational complexity than the other approaches, as N_f increases. This can be observed in Fig. 5, where we plot the complexity in flops of the different algorithms, using the expressions provided in section IV.B. As reference we also show the complexity of the MLD whose total number of flops is several orders of magnitude higher than the other alternatives, making it an unpractical solution. We can also see that the lowest complexity approach is the GSFIM-sMMP with 1 child node. However, this case revealed a poor performance in Fig. 4. Increasing the number of child nodes improves the BER results of GSFIM-sMMP but also increases its complexity, making GSFIM-ADMM the approach with the best tradeoff in terms of performance and complexity.

Based on the previous scenario, the signal constellation was increased from QPSK to 64-QAM.

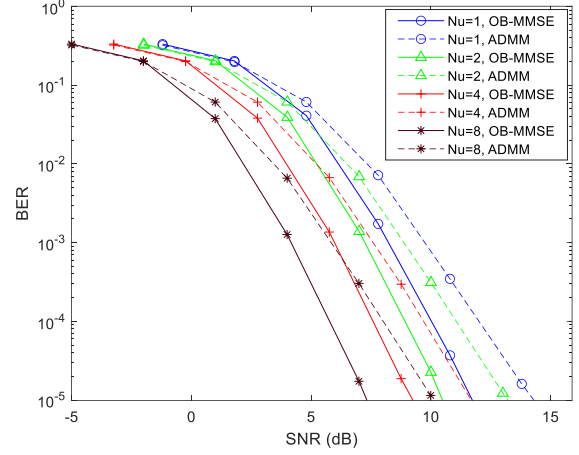


FIGURE 7. BER performance of PT-GSFIM in a MU scenario with $N_{tx}=5N_u$, $N_r=4$, $N_{af}=3$, $N_s=5$, $N_a=2$, $N_{rx}=4$, QPSK (5.75 bpcu per user).

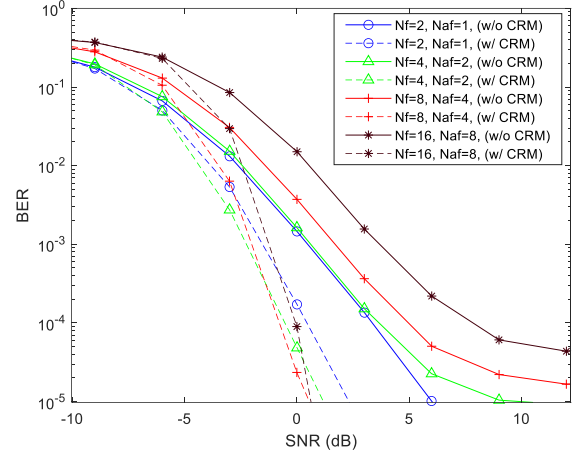


FIGURE 8. BER performance of PT-GSFIM in a MU scenario with $N_u=4$, $N_{tx}=8N_u$, $N_s=8$, $N_a=1$, $N_{rx}=4$, QPSK (3 bpcu per user).

The results are shown in Fig. 6. Most of the relative behavior of the curves observed in Fig. 4 repeats itself in this graph. However, in this case, ADMM achieves the best results followed by OB-MMSE. As verified previously, the use of CRM clearly improves the results.

Next, we study the influence of the number users on the performance of the receivers, considering basically the same scenario of Fig. 4, with a QPSK signal constellation and the use of CRM. Fig. 7 shows that independently of the chosen algorithm, the results improve by increasing the number of users. The improvement is achieved due to the additional transmit antennas (since $N_{tx}=5N_u$) combined with the power minimization step applied at the transmitter (detailed in [34]). It is important to note also that due to the use of block diagonalization based precoders, which remove all MUI and transform the signal arriving at each user to an equivalent SU channel (as shown in eq. (8)), there is no loss of diversity in this scenario. In fact, since we always use $N_{tx}=5N_u$ and $N_s=5$ when we increase the number of users, the equivalent received SU model

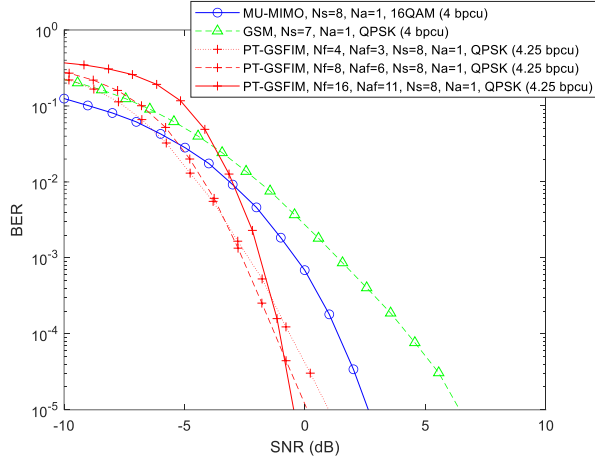


FIGURE 9. BER performance of PT-GSFIM, PT-SDIM, PTSFIM and conventional BD MU-MIMO ($N_u=8$).

(eq. (8)) always maintains the same dimensions.

In the next scenario, which is shown in Fig. 8, we evaluate the impact of the CRM size, defined according to different values of N_f and N_{af} . Due to the very high computational complexity of OB-MMSE for large CRMs, we consider the adoption of the ADMM algorithm only. The system is operating with $N_u=4$, $N_{tx}=8N_u$, $N_s=8$, $N_a=1$, $N_{rx}=4$, and QPSK. Note that the SE of the different curves is not always exactly the same, but it stays very close to 3 bpcu per user (varies between 3 and 3.3). As reference, curves without CRM are also included. It can be observed that without CRM, using larger subcarrier subblocks (N_f) tends to degrade the performance due to the more challenging detection task making the curves exhibit irreducible BER floors. In fact, it is important to note this figure corresponds to a difficult underdetermined scenario, where the number of observations is smaller than the number of unknowns. With CRM, the behavior is reversed. Even though, for low SNRs a larger N_f achieves worse BERs, at higher SNRs the curves fall with a steeper drop and end up achieving better results. Comparing the curves of PT-GSFIM with CRM against the curves without CRM for the same dimensions in term of N_f and N_{af} , we can observe gains between 3 dB (smallest N_f) and 8 dB (largest N_f) for a BER of 10^{-4} , due to the additional diversity.

In the last set of results shown in Fig. 9, we present a comparison between multiple scenarios using a conventional Block Diagonalization (BD) MU-MIMO scheme from [43], GSM MU-MIMO from [34], and the proposed PT-GSFIM. Regarding PT-GSFIM, we present three different configurations, while keeping the same SE. For all the curves we consider $N_u=8$, with the SE being close to 4 bpcu per user. ADMM based detection was adopted for all the schemes, i.e., for BD MU-MIMO, GSM MU-MIMO and PT-GSFIM. It can be observed that the proposed PT-GSFIM schemes can outperform both GSM and conventional MU-MIMO

and the gains are higher when the symbols are spread over a larger number of subcarriers (N_f).

VII. CONCLUSION

In this paper we described a MU-MIMO system where a base station transmits precoded space-frequency domain IM symbols. The proposed PT-GSFIM scheme adopts SSD techniques which allows it to benefit from the diversity effects inherent to a frequency selective channel. To support the detection of PT-GSFIM signals at the receiver, three different algorithms were presented. As expected, it was observed that by increasing the size of the subcarrier subblocks while keeping the same SE, can result in improved performance when SSD is employed, since the symbols are spread over a larger number of subcarriers, benefiting from the frequency diversity effect. Performance results also demonstrated that the proposed scheme can outperform both GSM and conventional MU-MIMO.

ACKNOWLEDGMENT

This work was partially supported by the FCT - Fundação para a Ciência e Tecnologia under the grant 2020.05621.BD. The authors also acknowledge the funding provided by FCT/MCTES through national funds and when applicable co-funded EU funds under the project UIDB/50008/2020.

REFERENCES

- [1] E. Basar, "Index modulation techniques for 5G wireless networks," *IEEE Communications Magazine*, vol. 54, no. 7, pp. 168-175, Jul. 2016.
- [2] X. Cheng, M. Zhang, M. Wen, and L. Yang, "Index Modulation for 5G: Striving to Do More with Less," *IEEE Wireless Communications*, vol. 25, no. 2, pp. 126-132, Apr. 2018.
- [3] E. Basar, M. Wen, R. Mesleh, M. Di Renzo, Y. Xiao, and H. Haas, "Index Modulation Techniques for Next-Generation Wireless Networks," *IEEE Access*, vol. 5, pp. 16693-16746, 2017.
- [4] S. Doğan Tusha, A. Tusha, E. Basar, and H. Arslan, "Multidimensional Index Modulation for 5G and Beyond Wireless Networks," *Proceedings of the IEEE*, vol. 109, no. 2, pp. 170-199, Feb. 2021.
- [5] Q. Li, M. Wen, B. Clerckx, S. Mumtaz, A. Al-Dulaimi, and R. Q. Hu, "Subcarrier Index Modulation for Future Wireless Networks: Principles, Applications, and Challenges," *IEEE Wireless Communications*, vol. 27, no. 3, pp. 118-125, Jun. 2020.
- [6] J. Wang, S. Jia and J. Song, "Generalised Spatial Modulation System with Multiple Active Transmit Antennas and Low Complexity Detection Scheme," *IEEE Transactions Wireless Communications*, vol. 11, no. 4, pp. 1605-1615, Apr. 2012.
- [7] M. Wen, B. Ye, E. Basar, Q. Li, and F. Ji, "Enhanced Orthogonal Frequency Division Multiplexing With Index Modulation," *IEEE Transactions on Wireless Communications*, vol. 16, no. 7, pp. 4786-4801, Jul. 2017.
- [8] M. Di Renzo, H. Haas, A. Ghayeb, S. Sugiura, and L. Hanzo, "Spatial Modulation for Generalized MIMO: Challenges, Opportunities, and Implementation," *Proceedings of the IEEE*, vol. 102, no. 1, pp. 56-103, Jan. 2014.
- [9] J. Jeganathan, A. Ghayeb, L. Szczecinski, and A. Ceron, "Space Shift Keying Modulation for MIMO Channels,"

- IEEE Transactions on Wireless Communications*, vol. 8, no. 7, pp. 3692-3703, Jul. 2009.
- [10] T. L. Narasimhan, P. Raviteja, and A. Chockalingam, "Generalized Spatial Modulation in Large-Scale Multiuser MIMO Systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 7, pp. 3764-3779, Jul. 2015.
 - [11] A. Younis, N. Serafimovski, R. Mesleh, and H. Haas, "Generalised spatial modulation," in *2010 Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA, 2010, pp. 1498-1502.
 - [12] L. Yang, "Transmitter Preprocessing Aided Spatial Modulation for Multiple-Input Multiple-Output Systems," in *2011 IEEE 73rd Vehicular Technology Conference (VTC Spring)*, 2011, pp. 1-5.
 - [13] R. Zhang, L. Yang, and L. Hanzo, "Generalised Pre-Coding Aided Spatial Modulation," *IEEE Transactions on Wireless Communications*, vol. 12, no. 11, pp. 5434-5443, Nov. 2013.
 - [14] R. Mesleh, S. S. Ikki, and H. M. Aggoune, "Quadrature spatial modulation," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 6, pp. 2738-2742, Jun. 2015.
 - [15] J. Li, M. Wen, X. Cheng, Y. Yan, S. Song, and M. H. Lee, "Generalized Precoding-Aided Quadrature Spatial Modulation," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 2, pp. 1881-1886, Feb. 2017.
 - [16] S. Sugiura, T. Ishihara, and M. Nakao, "State-of-the-Art Design of Index Modulation in the Space, Time, and Frequency Domains: Benefits and Fundamental Limitations," *IEEE Access*, vol. 5, pp. 21774-21790, Oct. 2017.
 - [17] R. Abu-alhiga and H. Haas, "Subcarrier-index modulation OFDM," in *2009 IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications*, Tokyo, Japan, 2009, pp. 177-181.
 - [18] E. Basar, "On Multiple-Input Multiple-Output OFDM with Index Modulation for Next Generation Wireless Networks," *IEEE Transactions on Signal Processing*, vol. 64, no. 15, pp. 3868-3878, Aug. 2016.
 - [19] J. Li, M. Wen, X. Jiang, and W. Duan, "Space-Time Multiple-Mode Orthogonal Frequency Division Multiplexing With Index Modulation," *IEEE Access*, vol. 5, pp. 23212-23222, 2017.
 - [20] E. Basar, Ü. Aygölü, E. Panayircı, and H. V. Poor, "Orthogonal Frequency Division Multiplexing With Index Modulation," *IEEE Transactions on Signal Processing*, vol. 61, no. 22, pp. 5536-5549, Nov. 2013.
 - [21] S. Gao, M. Zhang and X. Cheng, "Precoded Index Modulation for Multi-Input Multi-Output OFDM," *IEEE Transactions on Wireless Communications*, vol. 17, no. 1, pp. 17-28, Jan. 2018.
 - [22] J. Li, S. Dang, M. Wen, X. Jiang, Y. Peng, and H. Hai, "Layered Orthogonal Frequency Division Multiplexing With Index Modulation," *IEEE Systems Journal*, vol. 13, no. 4, pp. 3793-3802, Dec. 2019.
 - [23] T. Mao, Z. Wang, Q. Wang, S. Chen, and L. Hanzo, "Dual-Mode Index Modulation Aided OFDM," *IEEE Access*, vol. 5, pp. 50-60, 2017.
 - [24] T. Mao, Q. Wang, and Z. Wang, "Generalized Dual-Mode Index Modulation Aided OFDM," *IEEE Communications Letters*, vol. 21, no. 4, pp. 761-764, Apr. 2017.
 - [25] M. Wen, E. Basar, Q. Li, B. Zheng, and M. Zhang, "Multiple-Mode Orthogonal Frequency Division Multiplexing With Index Modulation," *IEEE Transactions on Communications*, vol. 65, no. 9, pp. 3892-3906, Sep. 2017.
 - [26] M. Wen, Q. Li, E. Basar, and W. Zhang, "Generalized Multiple-Mode OFDM With Index Modulation," *IEEE Transactions on Wireless Communications*, vol. 17, no. 10, pp. 6531-6543, Oct. 2018.
 - [27] T. Mao, Q. Wang, Z. Wang, and S. Chen, "Novel Index Modulation Techniques: A Survey," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 315-348, 1st Quart. 2019.
 - [28] S. Althunibat, R. Mesleh, and T. F. Rahman, "A Novel Uplink Multiple Access Technique Based on Index-Modulation Concept," *IEEE Transactions on Communications*, vol. 67, no. 7, pp. 4848-4855, Jul. 2019.
 - [29] E. Basar, "Media-Based Modulation for Future Wireless Systems: A Tutorial," *IEEE Wireless Communications*, vol. 26, no. 5, pp. 160-166, Oct. 2019.
 - [30] N. Souto and A. Correia, "Frequency Domain Equalization for Single and Multiuser Generalized Spatial Modulation Systems in Time Dispersive Channels," *IEEE Wireless Communications Letters*, vol. 9, no. 3, pp. 316-320, Mar. 2020.
 - [31] P. Yang, Y. Xiao, M. Xiao, and Z. Ma, "NOMA-Aided Precoded Spatial Modulation for Downlink MIMO Transmissions," *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 3, pp. 729-738, Jun. 2019.
 - [32] A. Raafat, A. Agustin, and J. Vidal, "Downlink Multi-User Massive MIMO Transmission Using Receive Spatial Modulation," *IEEE Transactions on Wireless Communications*, vol. 19, no. 10, pp. 6871-6883, Oct. 2020.
 - [33] M. Maleki, K. Mohamed-Pour, and M. Soltanalian, "Receive Spatial Modulation in Correlated Massive MIMO With Partial CSI," *IEEE Transactions on Signal Processing*, vol. 67, no. 5, pp. 1237-1250, Mar. 2019.
 - [34] N. Souto and A. Correia, "A Precoding Aided Space Domain Index Modulation Scheme for Downlink Multiuser MIMO Systems" *IEEE Transactions On vehicular Technology*, vol. 69, no. 10, pp. 12333 - 12337, Oct. 2020.
 - [35] T. Datta, H. S. Eshwariaiah, and A. Chockalingam, "Generalized Space-and-Frequency Index Modulation," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 7, pp. 4911-4924, Jul. 2016.
 - [36] Y. Xiao, S. Wang, L. Dan, X. Lei, P. Yang, and W. Xiang, "OFDM With Interleaved Subcarrier-Index Modulation," *IEEE Communications Letters*, vol. 18, no. 8, pp. 1447-1450, Aug. 2014.
 - [37] J. Boutros and E. Viterbo, "Signal Space Diversity: A Power- and Bandwidth-Efficient Diversity Technique for the Rayleigh Fading Channel," *IEEE Transactions on Information Theory*, vol. 44, no. 4, pp. 1453-1467, Jul. 1998.
 - [38] A. Correia, "Optimized Complex Constellations for Transmitter Diversity," *Wireless Personal Communications Journal*, vol. 20, no. 3, pp. 267-284, Mar. 2002.
 - [39] H. Lopes and N. Souto, "Iterative Signal Detection for Large-Scale GSM-MIMO Systems," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 8, pp. 7734-7738, Aug. 2018.
 - [40] T. L. Narasimhan and A. Chockalingam, "On the Capacity and Performance of Generalized Spatial Modulation," *IEEE Communications Letters*, vol. 20, no. 2, pp. 252-255, Feb. 2016.
 - [41] N. Souto, J. Silva, J. Pavia, and M. Ribeiro, "An alternating direction algorithm for hybrid precoding and combining in millimeter wave MIMO systems," *Physical Communication*, vol. 34, pp. 165-173, 2019.
 - [42] X. Yu, J. Zhang, and K. B. Letaief, "Doubling Phase Shifters for Efficient Hybrid Precoder Design in Millimeter-Wave Communication Systems," *Journal of Communications and Information Networks*, vol. 4, no. 2, pp. 51-67, Jun. 2019.
 - [43] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing Methods for Downlink Spatial Multiplexing in Multiuser MIMO Channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 461-471, Feb. 2004.
 - [44] Y. Xiao, Z. Yang, L. Dan, P. Yang, L. Yin, and W. Xiang, "Low-Complexity Signal Detection for Generalized Spatial Modulation," *IEEE Communications Letters*, vol. 18, no. 3, pp. 403-406, Mar. 2014.
 - [45] B. Shim, S. Kwon, and B. Song, "Sparse Detection With Integer Constraint Using Multipath Matching Pursuit," *IEEE Communications Letters*, vol. 18, no. 10, pp. 1851-1854, Oct. 2014.
 - [46] J. A. Tropp and A. C. Gilbert, "Signal Recovery From Random Measurements Via Orthogonal Matching Pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655-4666, Dec. 2007.

- [47] "Evolved Universal Terrestrial Radio Access (E-UTRA); Base Station (BS) radio transmission and reception, " *3GPP TS 36.104 v15.3.0*, June 2018.



VASCO R. J. VELEZ received the B.Sc. and M.Sc. degrees in Telecommunications and Computer Engineering from the ISCTE-IUL, in 2016 and 2018, respectively. In 2017, he joined at IT (Instituto de Telecomunicações) as part of his M.Sc. thesis, he worked in field of spoofing of GNSS signals. He is currently pursuing

the Ph.D degree in Information Science and Technology for effective RF techniques against unauthorized UAVs for his thesis. His research interests include the development of new technologies for communications systems, wireless networks, signal processing for telecommunications, MIMO schemes, MIMO Radar, source localization, GNSS signals and unmanned aerial vehicles.



JOÃO PEDRO C. B. B. PAVIA graduated in Telecommunications and Computer Engineering from the ISCTE-IUL, in 2018. At the same year, he joined at Instituto de Telecomunicações (IT) as a graduate research assistant. He worked for an international project funded by FCT (Fundação para a Ciência e Tecnologia) and TUBITAK (The Scientific and Technological Research Council of

Turkey) with aim of understanding the interaction between glow discharge detectors and THz waves in order to develop a new generation of imaging systems for safety applications. João Pedro is currently working on the design of wireless communications schemes for the mmWave and THz communications for his PhD thesis. His research interests include the development of new technologies for the Terahertz-band, wireless networks, signal processing for communications, MIMO schemes, electromagnetics, big data and machine learning. He is a student member of the IEEE and he served as chair of his student branch from 2015 to 2018. He was also involved as local organizer in several conferences promoted by ISCTE-IUL and by IT. Recently, he served as reviewer for the 2021 IEEE 94th Vehicular Technology Conference: VTC2021-Fall.



NUNO M. B. SOUTO graduated in aerospace engineering-avionics branch, in 2000 in Instituto Superior Técnico, Lisbon, Portugal and received his Ph.D. in 2006. From November 2000 to January 2002, he worked as a researcher in the field of automatic speech recognition for Instituto de Engenharia e Sistemas de Computadores, Lisbon, Portugal. He joined the ISCTE-University Institute of

Lisbon, as an assistant professor in 2006. He is a researcher at IT (Instituto de Telecomunicações), Portugal, since 2002 and has been involved in several international research projects and many national projects. His research interests include wireless networks, signal processing for communications, OFDM, single carrier transmission with frequency domain equalization, channel estimation, synchronization, MIMO schemes, wireless sensor networks and unmanned aerial vehicles. He is a member of the IEEE Signal Processing Society.



PEDRO J. A. SEBASTIÃO received the Ph.D. degree in electrical and computer engineering from IST. He is currently a Professor with the ISCTE-IUL's Information Science and Technology Department. He is the Board Director of the AUDAX-ISCTE-Entrepreneurship and Innovation Center, ISCTE, responsible for the LABS LISBOA incubator and a

Researcher at the Institute of Telecommunications. He has oriented several master's dissertations and Ph.D. theses. He is the author or coauthor of more than two hundred scientific articles and he has been responsible for several national and international Research and Development projects. He has been an Expert and Evaluator of more than one hundred national and international Civil and Defense Research and Development projects. It has several scientific, engineering, and pedagogical awards. Also, he has organized or co-organized more than fifty national and international scientific conferences. He planned and developed several postgraduate courses in technologies and management, entrepreneurship and innovation, and transfer of technology and innovation. He has supported several projects involving technology transfer and creation of start-ups and spin offs of value to society and market. He has developed his professional activity in the National Defense Industries, initially in the Office of Studies and later as the Board Director of the Quality Department of the Production of New Products and Technologies. He was also responsible for systems of communications technology in the Nokia-Siemens business area. His main research interests include monitoring, control and communications of drones, unnamed vehicles, planning tools, stochastic process (modeling and efficient simulations), the Internet of Things, and efficient communication systems.



AMÉRICO M. C. CORREIA (Senior Member, IEEE) received the B.Sc. degree in electrical engineering from the University of Angola, in 1983, and the M.Sc. and Ph.D. degrees in telecommunications from the Instituto Superior Técnico (IST), Lisbon, Portugal, in 1990 and 1994, respectively. From 1991 to 1999, he was with IST as an Assistant

Professor. From August 2000 to November 2007, he was an Associate Professor with ISCTE. Since December 2007, he has been a Full Professor with ISCTE, Lisbon University Institute, Portugal. He has been a Researcher with the IT (Instituto de Telecomunicações), Portugal, since 1992. He was a Post-Graduated Researcher with RWTH, Aachen, Germany, from October 1994 to September 1995. He visited the Nokia Research Center from September to December 1998 as a Visiting Scientist. From September 2000 to August 2001, he joined Ericsson Eurolab Netherlands. His main research topics include, CDMA, OFDMA, NOMA, MIMO, radio resource management, broadcasting technology, and enhanced multimedia broadcast/multicast services. He is a member of Communications Society of IEEE and has been a member of "Ordem dos Engenheiros" since 1987.

Chapter 3 - System-Level Assessment of a C-RAN based on Generalized Space–Frequency Index Modulation for 5G New Radio and Beyond

Following on the work started in the previous paper about IM schemes, we proceeded to the study and system level evaluation of the integration the PT-GSFIM approach into C-RAN. In this case we considered a C-RAN deployment based on the standards of 5G and B5G systems, and we evaluated the performance of the adoption of PT-GSFIM links in MU-MIMO scenarios. Three different scenarios from the standards of 5G NR of 3GPP alliance were considered, along with two different frequencies. For UMa and UMi scenarios, we used two options of carrier frequencies: one with 3.5 GHz and another one with 28 GHz. For the indoor office scenario, we used an mmWave band operating at 30 GHz or 70 GHz. We assumed three types of modulations, (from 4QAM up to 64QAM). Both scenarios have pedestrians moving at 3 km/h, with LOS or NLOS paths. In this contribution, instead of comparing simply the BER performance, we used the average throughput instead. We keep the same SE in every case to a better understanding and comparison.

From our results we observed that the proposed PT-GSFIM based C-RAN obtains better results compared to a traditional MU-MIMO approach. In fact, we could also observe that it was possible to use a lower modulation constellation and obtain the same SE and better performance than the standard MU-MIMO. This shows that IM can potentially offer better SE and EE in mobile communications, whether in urban or indoor scenarios. All the information mentioned, with more details, is presented below.

Paper details:

- Title: System-Level Assessment of a C-RAN based on Generalized Space–Frequency Index Modulation for 5G New Radio and Beyond
- Date of Publication: 2 February 2022
- Journal: Applied Sciences
- Scimago/Scopus Journal Ranking: Quartile 2
- Publisher: MDPI.

Article

System-Level Assessment of a C-RAN based on Generalized Space-Frequency Index Modulation for 5G New Radio and Beyond

Vasco Velez ^{1,2,*}, João Pedro Pavia ^{1,2}, Catarina Rita ¹, Carolina Gonçalves ¹, Nuno Souto ^{1,2}, Pedro Sebastião ^{1,2} and Américo Correia ^{1,2}

¹ Department of Information Science and Technology, ISCTE-Instituto Universitário de Lisboa, 1649-026 Lisbon, Portugal; Joao_Pedro_Pavia@iscte-iul.pt (J.P.P.); Catarina_Rita@iscte-iul.pt (C.R.); Carolina_Loureiro@iscte-iul.pt (C.G.); Nuno.Souto@iscte-iul.pt (N.S.); Pedro.Sebastiao@iscte-iul.pt (P.S.); Americo.Correia@iscte-iul.pt (A.C.)

² Instituto de Telecomunicações, 1049-001 Lisbon, Portugal

* Correspondence: Vasco_Velez@iscte-iul.pt

Citation: Velez, V.; Pavia, J.P.; Rita, C.; Gonçalves, C.; Souto, N.; Sebastião, P.; Correia, A. System-Level Assessment of a C-RAN based on Generalized Space-Frequency Index Modulation for 5G New Radio and Beyond. *Appl. Sci.* **2022**, *12*, 1592. <https://doi.org/10.3390/app12031592>

Academic Editor:
Mario Marques Da Silva

Received: 12 January 2022
Accepted: 31 January 2022
Published: 2 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Index Modulations (IM) have been attracting considerable research efforts in recent years as it is considered a promising technology which can enhance the spectral and energy efficiency and help cope with the rising demand of mobile traffic in future wireless networks. In this paper, we propose a cloud radio access network (C-RAN) suitable for fifth generation (5G) and beyond systems, where the base stations (BSs) and access points (APs) transmit multidimensional IM symbols, which we refer to as Precoding-aided Transmitter side Generalized Space-Frequency IM (PT-GSFIM). The adopted PT-GSFIM approach is an alternative multiuser multiple-input multiple-output (MU-MIMO) scheme that avoids multiuser interference (MUI) while exploiting the inherent diversity in frequency selective channels. To validate the potential gains of the proposed PT-GSFIM based C-RAN, a thorough system level assessment is presented for three different three-dimensional scenarios taken from standardized 5G New Radio (5G NR), using two different numerologies and frequency ranges. Throughput performance results indicate that the 28 GHz band in spite of its higher bandwidth and higher achieved throughput presents lower spectral efficiency (SE). The 3.5 GHz band having lower bandwidth and lower achieved throughput attains higher SE. Overall, the results indicate that a C-RAN based on the proposed PT-GSFIM scheme clearly outperforms both generalized spatial modulation (GSM) and conventional MU-MIMO, exploiting its additional inherent frequency diversity.

Keywords: Index Modulation; Precoding; PT-GSFIM; Multi-user MIMO; System level simulation; 5G and beyond 5G

1. Introduction

As technology advances in wireless communications, more innovative solutions have been appearing. These solutions should be able to handle the increase of the number of devices connected to the network

and also the amount of mobile traffic services that are growing day by day. Newer generations are expected to offer better system robustness, higher mobility, spectral and energy efficiency than previous ones.

Current fifth generation (5G) technology is capable, efficient and flexible, but can be improved further. 5G works in bands below 7 GHz, which corresponds to the Frequency Range 1 (FR1) of 5G New Radio (5G NR) [1], and above 24 GHz, corresponding to Frequency Range 2 (FR2) of 5G NR, and can accommodate much more users than the last generation counterpart. It uses the massive multiple input multiple output (MIMO) concept which allows to achieve 100x higher efficiency without requiring more base stations (BSs). Massive MIMO and non-orthogonal multiple access (NOMA) are widely used in wide band networks and can accommodate a higher number of users and streams but fail to achieve a higher spectral efficiency (SE) that is required to meet the challenges and requirements of beyond 5G (B5G) and future generations (6G) [2,3]. Regarding recent research work that has been done in the physical layer, one of the potential candidates to improve future and current networks is the use of index modulation (IM) schemes [4,5].

IM has surged in recent years as a promising technique to convey additional information bits, named as index bits, that are distributed through a certain order, and select certain resource elements such as antennas, subcarriers, slots and/or channels. This capability of activating certain resources of communications, through the indices of the building blocks, can result in improvements on energy efficiency (EE) and SE, and also reduces the complexity of the receivers [3,6]. IM can be applied together with current technologies like massive MIMO and can be used in high mobility scenarios, like mobile communications [5]. Spatial modulation (SM) and orthogonal frequency division multiplexing with IM (OFDM-IM) schemes are well-known examples of IM schemes [4,7].

SM was first introduced in order to simplify the MIMO schemes to a lower computational complexity and implementation, as it only requires a single RF chain to convey the index bits [8]. This type of scheme can be combined with conventional modulation techniques such as amplitude and Phase Modulation (APM), Phase Shift Keying (PSK) or Quadrature Amplitude Modulation (QAM). Although SM only uses one active transmitter, in [9] the concept of SM was generalized to enable multiple active antennas simultaneously, which was named as Generalized SM (GSM). IM schemes like SM and GSM, are usually more suitable in large scale communications, allowing superior EE and better BER results than conventional MIMO [5, 9, 10]. The addition of precoding techniques allows the implementation of SM relying on the activation of antennas at the receiver instead of the transmitter, enabling lower implementation complexity at the receiver [10].

OFDM-IM uses an alternative approach as it explores IM applied at the frequency domain instead of the spatial domain, relying on the indices of subcarriers of conventional OFDM transmission [11], [12]. Like SM/GSM, in OFDM-IM only a portion of subcarriers are activated according to their indices, and only those convey M -ary modulated signals [13]. Various authors have modified and improved the basic OFDM-IM scheme. An example is using precoding techniques to improve further the communications in downlink between BSs and receivers. This results in reduced complexity at the receiver side, which tends to lower the interference [14]. The use of interleavers on OFDM subcarriers is another possible approach for achieving considerable improvements. In

[15], the authors compared the performance of OFDM-IM with a subcarrier-level block interleaver (OFDM-ISIM) against traditional OFDM and OFDM-IM. It was shown that the former was capable of achieving better BER results. Layered OFDM-IM (L-OFDM-IM) is another example of an IM scheme which divides the OFDM subcarriers into multiple layers [16]. Dual Mode OFDM-IM (DM-OFDM-IM), Multi-Mode OFDM-IM (MM-OFDM-IM), and their generalized versions, are other good examples of IM based OFDM schemes. Both of these schemes use groups of designed constellations, which can be selected according to their indices and desired communication [17-20]. All of these mentioned schemes tried to improve SE over OFDM-IM.

It is important to note that IM can also be applied to time and channel domains. In [21], the authors presented a scheme where IM is applied in the time domain. It uses the concept of time slots, similar to Time Division Multiple Access (TDMA), which is named as IM Multiple Access (IMMA). On the other hand, the IM approach in [22] uses the channel approach to convey additional information through channel realizations. This is also known as media-based modulation (MBM).

Many recent works on IM are focused on multiuser (MU) communications, especially in downlink situations. The receivers usually have a limited number of antennas, which are prone to multiuser interference (MUI). On the other hand, the uplink has less complexity, and can be addressed like an extension of GSM [23]. Some authors have proposed versions of IM based on multidimensional schemes, that are built with a combination of spatial and frequency domains [7]. In [24], the authors have merged both SM and OFDM-IM schemes simultaneously to produce a scheme named as Generalized Space-Frequency IM (GSFIM). In this new technique, the index bits, can select which antenna and subcarrier should be activated at each transmission. Although only single user (SU) cases were explored, the authors showed good results in BER performance and transmission rates, when compared with other schemes. More recently, in [25] the authors extended the same concept, proposing multidimensional IM schemes that work in MU scenarios. Both frequency and spatial domains are exploited for downlink MU communications with assistance of Signal Space Diversity (SSD), showing promising results. In [26], a scheme that combines space-frequency code and MIMO OFDM-IM was proposed. It was shown that the presented scheme was capable of increasing the transmit diversity order.

5G and beyond networks using radio access virtualization strategies and advanced computational platforms will exploit network densification. The virtual cell concept removes the traditional cell boundary for the device and provides a consequent reduction of the detrimental "cell-edge experience" by the terminal. Traditionally, devices are associated with a cell and, as a consequence, the link performance may degrade as a terminal moves away from the cell center. In a virtualized cloud radio access network (C-RAN), the network determines which BSs or access points (APs) are to be associated with each terminal. The cell moves with the terminal in order to provide a cell-center experience throughout the entire network. Each terminal is served by its preferred set of APs. The actual serving set for a terminal may contain one or multiple BSs and the terminal's data is partially or fully available at some or a small set of potential serving BS. The BS controller (Central Processor) will accommodate each terminal with its preferred set and transmission mode

at every communication instance while considering load and channel state information (CSI) knowledge associated with the BS [27].

5G NR system [28- 30] uses scalable OFDM numerology introducing specific subcarrier spacings (Δf), Transmission Time Interval (TTI), Cyclic Prefix (CP), and the number of slots. Higher numerology indices correspond to larger subcarrier spacings (SCSs), ranging from 15 kHz up to 480 kHz according to the equation: $\Delta f = 15 \text{ kHz} \times 2^n$. The numerology index n depends on various factors (i.e., service requirements, deployment type, carrier frequency, etc.). The introduction of wider SCS is essential for mitigating inter-carrier interference (ICI) and phase noise at millimeter wave (mmWave) frequencies. As SCS widens, the TTI assumes smaller values ranging from 1 ms to 31.25 μs . 5G NR was designed to lower interference and increase EE by reducing always-on transmissions, which is a crucial aspect to extend the lifetime of IoT devices. 5G NR ensures forward compatibility as it is prepared for its future 6G evolution in use cases and technologies. The introduction of mini-slots made 5G NR capable to guarantee low-latency requirements for ultra-reliable and low latency communications (URLLC). At mmWave frequencies, high capacity and extreme data rates are possible, even though higher frequencies introduce limitations in coverage due to increased signal attenuation [31]. 5G NR specifications, are projected to accommodate dense urban scenarios, but with the introduction of mmWave, it can be more likely to have some part of the signal blocked by obstacles or severely affected by distance, which causes a substantial decrease of the signal strength, becoming hard to compensate even with advanced signal processing techniques [32, 33].

Motivated by the work above, in this paper we study the adoption of Precoding-aided Transmitter side Generalized Space-Frequency IM (PT-GSFIM) as an alternative multi-user MIMO (MU-MIMO) scheme for the downlink in 5G and beyond systems. Besides exploiting both frequency and space domain resources through IM, different M -QAM constellation orders can be supported so as to accommodate flexible tradeoffs between SE, performance and complexity. In this paper we extend our previous work in [25], where we presented the design of a precoder and different detection algorithms for PT-GSFIM. The present scheme uses the same precoder to remove the MUI between BS and receivers, breaking the MU communication into equivalent small SU links and achieving a lower complexity at the receiver side. Furthermore, we adopt subcarrier-level interleaving combined with SSD techniques as it allows the proposed approach to exploit the inherent diversity in frequency selective channels and achieve improved performance without requiring additional power or bandwidth. In this paper, we focus on the system level assessment of a C-RAN based on the PT-GSFIM scheme. The system level evaluation is performed for numerologies one and two from 5G NR in three 3D scenarios with different parameters, such as, modulation, number of transmitted antennas per user, bandwidth and frequency carrier, and it is benchmarked against two alternative MU-MIMO schemes. System level simulations were performed based on PT-GSFIM link level results between the BSs and the multiple terminals, where it is considered that the alternating direction method of multipliers (ADMM) detection algorithm from [25] is applied on the receiver side. System level evaluation demonstrates that C-RAN deployments in several indoor and outdoor scenarios, including urban or mixed, can achieve significant performance and coverage improvements over typical cellular networks.

The paper is organized as follows: Section 2 presents the model for the PT-GSFIM system, Section 3 presents the transmitter and receiver structure as well as the different 5G NR scenarios that are considered in the evaluation. Section 4 presents and discusses the link and system level simulations results, whereas the conclusions are outlined in Section 5.

Notation: Matrices and vectors are denoted by uppercase and lowercase boldface letters, respectively. $(\cdot)^T$ and $(\cdot)^H$ denote the transpose and conjugate transpose of a matrix/vector, $\|\cdot\|_p$ is the ℓ_p -norm of a vector, $\|\cdot\|_0$ is its cardinality, $\text{supp}(\mathbf{x})$ returns the support of \mathbf{x} , $\lfloor \cdot \rfloor$ is the floor function and \mathbf{I}_n is the $n \times n$ identity matrix.

2. System Model

We consider the downlink of an OFDM based MU-MIMO system, where a BS equipped with N_{tx} antennas, transmits to N_u users. Each user is assumed to have N_{rx} antennas each, wherein that transmission is composed by grouping the subcarriers into N_f -sized subblocks, with only N_{af} of them active at any given moment, as seen in Figure 1.

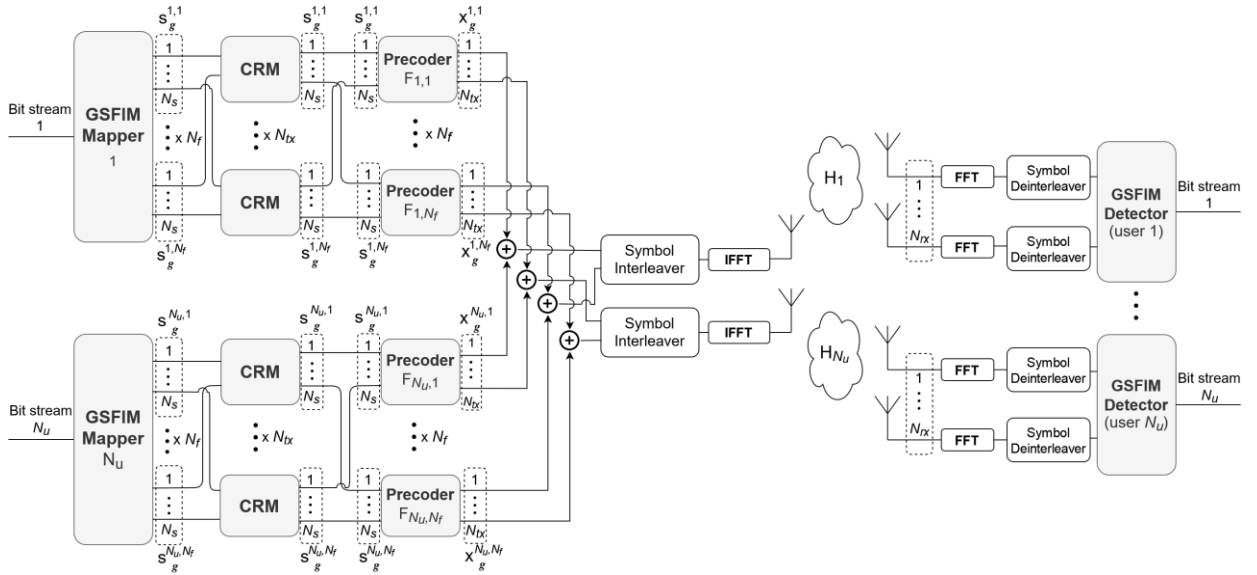


Figure 1. Transmitter and Receiver structure [25].

As an example, if we have a scenario with $N_f=8$ and $N_{af}=6$, then only 6 out of each group of 8 subcarriers are selected, and only the selected ones are activated to convey modulated symbols through their respective indices. This pattern of activation of subcarriers is also combined with the pattern of activation of the spatial domain which means that not all antennas will be transmitting in an active subcarrier. In fact, denoting N_s as the total number of streams assigned to the antennas on each subcarrier, then only N_a of the positions will contain modulated symbols which will correspond to active antennas.

2.1. Precoding-aided Transmitter side Generalized Space-Frequency Index Modulation (PT-GSFIM)

We adopt an OFDM based MU-MIMO as basis, where part of the information is mapped to conventional M -sized APM symbols and the remaining is encoded according to spatial and frequency indices. We assume a transmission block matrix where the bits are mapped in a simple

way, with the encoding of the spatial indices executed independently from the frequency indices. In this case simple look-up-tables (LUTs) are inside each subblock: one for the spatial resources and another for the frequency resources. Although some SE is sacrificed when the mapping is done independently (the number of available space-frequency blocks that are indices, normally are reduced), it simplifies the detection process as well the mapping/de-mapping process, as we will show further below. It is important to highlight that the approach based on a pair of LUTs is only adequate when a small number of both active antenna combinations and active subcarrier combinations exists. When this number is large, a different strategy can be employed, as described in [11, 34].

Defining a PT-GSFIM symbol as $\mathbf{s}_g^u \in \mathbb{C}^{N_s N_f \times 1}$, for the u^{th} user, it can be written as $\mathbf{s}_g^u = \left[\left(\mathbf{s}_g^{u,1} \right)^T \cdots \mathbf{0} \cdots \left(\mathbf{s}_g^{u,N_{af}} \right)^T \right]^T$, which corresponds to the concatenation of $(N_f - N_{af})$ length- N_s vectors of zeros with N_{af} GSM symbol vectors $\mathbf{s}_g^{u,i} \in \mathbb{C}^{N_s \times 1}$. These symbols are defined as $\mathbf{s}_g^{u,i} = \left[0 \quad s_{g,1}^{u,i} \quad 0 \quad \cdots \quad 0 \quad s_{g,N_a}^{u,i} \quad 0 \right]^T$, where $i = \{1, \dots, N_{af}\}$, $u = \{1, \dots, N_u\}$ and $g = \{1, \dots, N_{GSFIM}\}$ (N_{GSFIM} is the number of PT-GSFIM symbols inside an OFDM block). Each of the N_{GSFIM} symbols is composed of all available antenna and subcarrier positions, where each is activated and deactivated according to their indices. The symbols $s_{g,j}^{u,i}$, where $j = \{1, \dots, N_a\}$, are selected from an APM constellation. Since we are considering a MU downlink situation, where precoding is applied, the expression for the transmitted signal can be written as

$$\mathbf{x}_g = \sum_{u=1}^{N_u} \mathbf{F}_u \mathbf{s}_g^u = \mathbf{F} \mathbf{s}_g, \quad (1)$$

where $\mathbf{s}_g = \left[\left(\mathbf{s}_g^1 \right)^T \cdots \left(\mathbf{s}_g^{N_u} \right)^T \right]^T$ and $\mathbf{F} = \left[\mathbf{F}_1 \cdots \mathbf{F}_{N_u} \right]$, with $\mathbf{F}_u \in \mathbb{C}^{N_{tx} N_f \times N_s N_f}$, represents the precoder matrix. After the precoders and interleavers (which are explained further ahead), the frequency domain symbols are concatenated into N -sized blocks and converted to time domain through an Inverse Fast Fourier Transform (IFFT). Before the transmission occurs, a CP with length greater than the delay spread of the channel is added to each block. It is important to observe that each PT-GSFIM symbol can carry a total of bits per user calculated through

$$N_{bits} = N_{af} \left(\left\lfloor \log_2 \left(\frac{N_s}{N_a} \right) \right\rfloor + N_a \log_2 M \right) + \left\lfloor \log_2 \left(\frac{N_f}{N_{af}} \right) \right\rfloor, \quad (2)$$

where N_s is the number of available spatial positions on each subcarrier which is constrained to $N_s \leq N_{tx}/N_{u_f}$ and M is the size of the constellation adopted for the conventional modulated symbols.

2.2. 5G New Radio 3D Scenarios

There are several 5G test scenarios [47]. The three most important ones are described here and later evaluated using the multidimensional PT-GSFIM signals [25].

2.2.1 Urban macro

The urban macro (UMa) homogeneous deployment scenario focuses on large cells and continuous coverage. This scenario will be interference-limited, using macro transmission and reception points (TRPs, radio BSs above rooftop level). The inter site distance (ISD) in this scenario is 500 m, with the BS antenna height being 25 m. The carrier frequency is 3.5 GHz or 28 GHz. The bandwidth for 3.5 GHz and 28 GHz is up to 100 MHz and 1 GHz, respectively. Full buffer model is assumed. A total of 20 user equipments (UEs) are distributed per sector of each cell, with 80% users being indoor with 3 km/h velocity and the remaining 20% being in cars with 30 km/h velocity. All parameters used correspond to the Table 7.2-1 in [36].

2.2.2 Urban micro – Street canyon

The urban micro (UMi) deployment scenario focuses on urban micro cells and high user densities and traffic loads in city centers and dense urban areas. This scenario will be interference-limited, using micro TRPs with micro cells. The ISD for the micro cells is 200 m. There are 3 micro cells per macro cell, with the micro-BS antenna height being 10 m. The carrier frequency and bandwidths for micro cell are the same as for urban macro cell. Full buffer traffic model is assumed. A total of 10 UEs are distributed per micro sector with 80% users being indoor with a moving speed of 3 km/h and the remaining 20% being in cars with a velocity of 30 km/h. All parameters used correspond to the Table 7.2-1 in [36]. The layout of this scenario is illustrated on Figure 2.

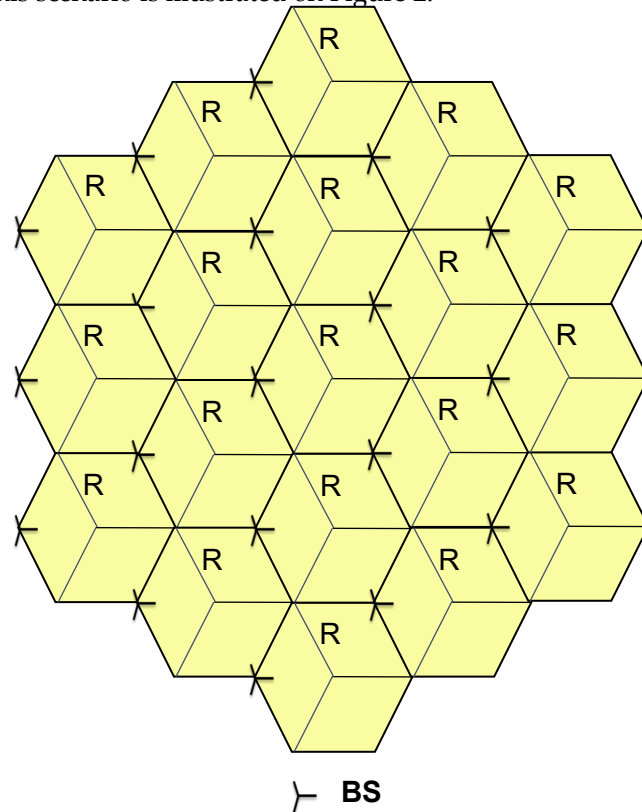


Figure 2. Urban micro scenario layout: Macro layer: ISD = 200 m = 3R; Micro layer: 3 micro TRPs per macro TRPs.

2.2.3 Indoor hotspot

The indoor hotspot deployment scenario focuses on small coverage per cell and high user throughput or user density in buildings. This scenario represents indoor offices (InDs) with a total area of $120\text{ m} \times 50\text{ m}$. There are 12 small cells which are deployed with an ISD of 20 m . In this case the BS antenna height is 3 m . The carrier frequency options are 30 and 70 GHz (mmWaves). The bandwidth for both 30 and 70 GHz carriers is up to 1 GHz . A total of 10 users per cell are distributed uniformly and all users are indoors with 3 km/h velocity. Full buffer model is assumed. The layout of this scenario is illustrated on Figure 3. All parameters used correspond to the Table 7.2-2 in [36].

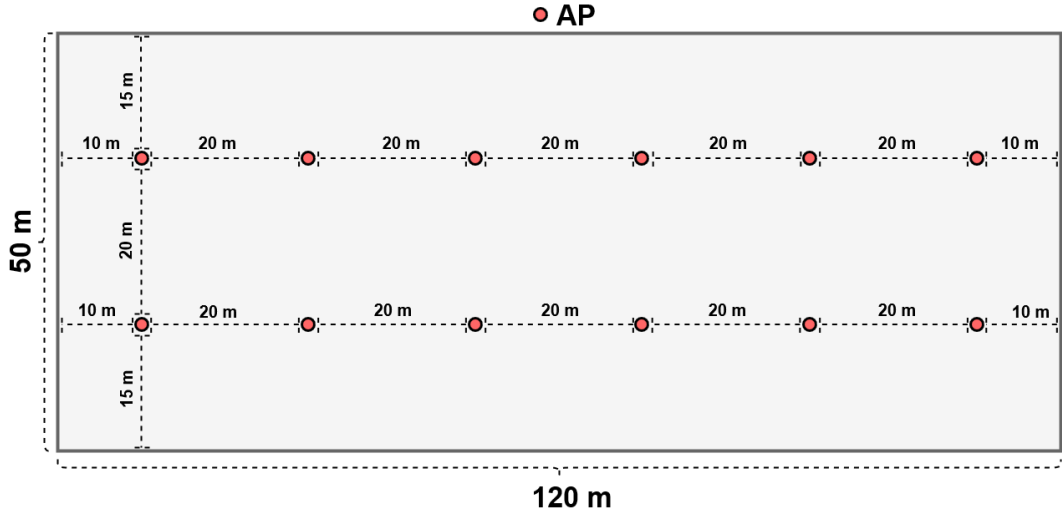


Figure 3. Indoor Office scenario layout.

Our 3D simulation channel model considers the above mentioned indoor and outdoor wireless propagation environments in terms of physical aspects of mmWave frequencies while numerous 5G NR channel models issues are adopted [35, 36]. For any considered operating frequency and scenario, the number of clusters, number of subarrays (scatterers) per cluster, and the positions of the clusters can be determined by the detailed steps and procedures of [35, 36]. According to the 5G 3GPP 3D channel models, the number of clusters and scatterers are determined using the Poisson and uniform distributions with specific parameters. Since, 5G NR extends the operating frequency range by comprising the frequency bands below 7 GHz (i.e., FR1) and in the range from 24.25 GHz to 52.6 GHz (i.e., FR2), different antenna solutions and techniques need to be employed depending on the utilized spectrum. For lower frequencies, up to a moderate number of antennas can be activated (i.e., 64). In higher frequency bands, the transmission is characterized by a considerable signal attenuation that limits the network coverage. To overcome this limitation, one of the key features is the adoption of a large number of multi-antenna elements having a given aperture to increase the transmission/reception capability of MU-MIMO and beamforming. Since, managing transmissions in higher frequency bands is complicated, beam management is necessary to establish the correspondence between the directions of the transmitter and the receiver-side beams by identifying the most suitable beam pair for both downlink and uplink.

3. Transmitter and Receiver Structure

In this section we describe the transmitter and receiver structures adopted. Section 3.2. addresses the use of complex rotation matrices (CRM) so as to integrate SSD into PT-GSFIM symbol design. The last subsection, the adopted receiver algorithm.

3.1. Precoder Design

Assuming the time domain samples have been converted to the frequency domain through an N -point Fast Fourier Transform (FFT) and the CP has been discarded, the g^{th} PT-GSFIM symbol received by the user can be written as

$$\mathbf{y}_g^u = \mathbf{H}_g^u \mathbf{x}_g^u + \mathbf{n}_g^u = \mathbf{H}_g^u \mathbf{F}_u \mathbf{s}_g^u + \mathbf{H}_g^u \sum_{\substack{j=1 \\ j \neq u}}^{N_u} \mathbf{F}_j \mathbf{s}_g^j + \mathbf{n}_g^u, \quad (3)$$

where $\mathbf{y}_g^u \in \mathbb{C}^{N_{rx} N_f \times 1}$, $\mathbf{n}_g^u \in \mathbb{C}^{N_{rx} N_f \times 1}$ represents the noise and $\mathbf{H}_g^u \in \mathbb{C}^{N_{rx} N_f \times N_{tx} N_f}$ is the frequency domain channel matrix between the BS and the user u . Since, we are considering an OFDM scheme, the channel matrix follows a block diagonal structure, and can be written as

$$\mathbf{H}_g^u = \text{blkdiag} \left\{ \mathbf{H}_g^{u,1}, \dots, \mathbf{H}_g^{u,f}, \dots, \mathbf{H}_g^{u,N_f} \right\}, \quad (4)$$

with $\mathbf{H}_g^{u,f} \in \mathbb{C}^{N_{rx} \times N_{tx}}$, $f = 1, \dots, N_f$. As seen in the right side of (3), the second term represents the interference between users. Using the approach of block diagonalization (BD) described in [37] for the design of the precoder matrices \mathbf{F}_u , it is possible to eliminate all MUI from all the receivers. It is important to note that the signals conveyed on each subcarrier only generate interference on that subcarrier due to the block diagonal structure of \mathbf{H}_g^u . This means that the precoder matrices \mathbf{F}_u , for each subcarrier can be built independently and follow a block diagonal structure, i.e.

$$\mathbf{F}_u = \text{blkdiag} \left\{ \mathbf{F}_{u,1}, \dots, \mathbf{F}_{u,N_f} \right\}, \quad (5)$$

with $\mathbf{F}_{u,f} \in \mathbb{C}^{N_{tx} \times N_s}$, $f = 1, \dots, N_f$. Each of the component precoder matrices $\mathbf{F}_{u,f}$ is generated so as to enforce $\mathbf{H}_g^{v,f} \mathbf{F}_{u,f} = \mathbf{0}$ for all $v \neq u$, cancelling all MUI. Let us define the matrix that concatenates all the channel matrices between the base station and all users, except user u , for the subcarrier f as

$$\tilde{\mathbf{H}}_g^{u,f} = \left[\left(\mathbf{H}_g^{1,f} \right)^T \dots \left(\mathbf{H}_g^{u-1,f} \right)^T \left(\mathbf{H}_g^{u+1,f} \right)^T \dots \left(\mathbf{H}_g^{N_u,f} \right)^T \right]^T. \quad (6)$$

This matrix models the transmission of the signal targeted at the u^{th} user when it reaches all of the other receivers, which causes interference, thus corresponding to the interference channel matrix. The appropriate precoder matrix $\mathbf{F}_{u,f}$, can be designed using an orthonormal basis of the null space of $\tilde{\mathbf{H}}_g^{u,f}$. Starting from the singular value decomposition (SVD) of $\tilde{\mathbf{H}}_g^{u,f}$, which is given by

$$\tilde{\mathbf{H}}_g^{u,f} = \tilde{\mathbf{U}}_{u,f} \tilde{\mathbf{\Lambda}}_{u,f} \left[\tilde{\mathbf{V}}_{u,f}^{(1)} \tilde{\mathbf{V}}_{u,f}^{(0)} \right]^H, \quad (7)$$

where $\tilde{\mathbf{U}}_{u,f}$ is the matrix with the left-singular vectors, $\tilde{\mathbf{\Lambda}}_{u,f}$ is a rectangular diagonal matrix with the decreasing nonzero singular values, $\tilde{\mathbf{V}}_{u,f}^{(1)}$ contains the right singular vectors corresponding to the nonzero singular values and $\tilde{\mathbf{V}}_{u,f}^{(0)}$ contains the remainder right singular vectors. To guarantee zero MUI in each precoder matrix, each precoder can be set as

$\mathbf{F}_{u,f} = \tilde{\mathbf{V}}_{u,f}^{(0)}[:, 1:N_s]$. In this case, at each receiver, the resulting signal is reduced to

$$\mathbf{y}_g^u = \hat{\mathbf{H}}_g^u \mathbf{s}_g^u + \mathbf{n}_g^u, \quad (8)$$

where $\hat{\mathbf{H}}_g^u = \mathbf{H}_g^u \mathbf{F}_u$ is the equivalent SU channel. It can be observed that there is no presence of MUI in the resulting signal. Keeping in mind that both \mathbf{H}_g^u and \mathbf{F}_u have a block diagonal structure, $\hat{\mathbf{H}}_g^u$ will be block diagonal as well. Only a few positions of \mathbf{s}_g^u will tend to be nonzero, according to the stated system model. As a result, \mathbf{s}_g^u is a sparse vector, and its sparsity can be used to lower the transmitted power. To do this, we can use the same strategy that we considered in [38] for virtual GSM transmissions, in which an alternative signal is generated to reduce the transmitted power while assuring that the signals arriving at receivers are similar to the original ones. Although, we assume fully-digital precoders in this paper, it is possible to simplify the implementation complexity of the proposed solution by using a hybrid precoder design comprising a reduced digital part and an analog part supported for example on analog phase shifters (and which are the same for all subcarriers). The hybrid design can be easily accomplished following the approximation-based approach described in [39].

3.2. Complex Rotation Matrices

A GSM symbol vector forming a PT-GSFIM is subject to the influence of the channel on a single subcarrier since, according to the signal model of equation (8), $\hat{\mathbf{H}}_g^u$ is block diagonal. However, we can incorporate SSD techniques into the design of PT-GSFIM in order to take advantage of the inherent diversity of frequency selective channels commonly seen in mobile propagation situations. SSD was first proposed in [40] and can be used to increase the performance without requiring more power or bandwidth. In the proposed approach we apply SSD by spreading each GSM symbol that makes up the primary PT-GSFIM symbol over different subcarriers through the use of CRM. Working with a rotating super-symbol is part of the CRM application procedure. In this situation, we only apply the spreading in the frequency direction, which means that all of the spatial components are rotated simultaneously. For each user u , the rotated length- $N_s N_f$ super-symbol can be expressed as

$$\tilde{\mathbf{s}}_g^u = (\mathbf{A}_{N_f} \otimes \mathbf{I}_{N_s}) \cdot \mathbf{s}_g^u, \quad (9)$$

and the transmitted signal becomes (1)

$$\mathbf{x}_g = \sum_{u=1}^{N_u} \mathbf{F}_u (\mathbf{A}_{N_f} \otimes \mathbf{I}_{N_s}) \mathbf{s}_g^u = \mathbf{F} (\mathbf{I}_{N_u} \otimes (\mathbf{A}_{N_f} \otimes \mathbf{I}_{N_s})) \mathbf{s}_g. \quad (10)$$

Matrix $\mathbf{A}_{N_f} \in \mathbb{C}^{N_f \times N_f}$ can be chosen from the orthonormal complex matrices (OCRM) family, which is defined as

$$\mathbf{A}_{M_{CRM}} = \begin{cases} \begin{bmatrix} e^{j\varphi} & je^{-j\varphi} \\ je^{j\varphi} & e^{-j\varphi} \end{bmatrix} / |\mathbf{A}_2|^{1/2}, M_{CRM} = 2 \\ |\mathbf{A}_2| = \det(\mathbf{A}_2) = 2 \\ \begin{bmatrix} \mathbf{A}_{M_{CRM}/2} & \mathbf{A}_{M_{CRM}/2} \\ \mathbf{A}_{M_{CRM}/2} & -\mathbf{A}_{M_{CRM}/2} \end{bmatrix} / |\mathbf{A}_{M_{CRM}}|^{1/M_{CRM}}, M_{CRM} > 2 \end{cases}, \quad (11)$$

where $M_{CRM} = 2^n$, $|\mathbf{A}_{M_{CRM}}| = \det(\mathbf{A}_{M_{CRM}})$ and φ is the rotation angle [41]. To get the most out of frequency diversity, a subcarrier-wise interleaver is applied so that the effective group of subcarriers assigned to each PT-GSFIM symbol is spread out across the entire bandwidth. This interleaver is similar to the subcarrier-level block interleaver in [15]. However, in the case of PT-GSFIM its effect is leveraged since with the use of CRM, each individual GSM symbol is spread over the different subcarriers allocated to the corresponding PT-GSFIM subblock.

3.3. Receiver Design

In the previous section, the transmitter side was described, where the use of precoder removes all MUI from the receiver side. This means that only SU detection needs to be implemented at the receivers.

Due to the use of CRM, each GSM symbol comprising a PT-GSFIM block is dispersed across numerous subcarriers. As a consequence, the received signal can be written as

$$\begin{aligned} \mathbf{y}_g^u &= \hat{\mathbf{H}}_g^u \tilde{\mathbf{s}}_g^u + \mathbf{n}_g^u \\ &= \tilde{\mathbf{H}}_g^u \mathbf{s}_g^u + \mathbf{n}_g^u, \end{aligned} \quad (12)$$

where $\tilde{\mathbf{H}}_g^u = \hat{\mathbf{H}}_g^u (\mathbf{A}_{N_f} \otimes \mathbf{I}_{N_s})$ is the overall equivalent channel that the user u “perceives” for the g^{th} PT-GSFIM symbol. To design the detector, first, we formulate the maximum likelihood detection (MLD) problem as

$$\min_{\mathbf{s}_g^u} f(\mathbf{s}) \triangleq \|\mathbf{y}_g^u - \tilde{\mathbf{H}}_g^u \mathbf{s}_g^u\|_2^2, \quad (13)$$

$$\text{subject to } \mathbf{s}_g^u \in \mathcal{A}_0^{N_s N_f}, \quad (14)$$

$$\text{supp}(\mathbf{s}_g^{u,i}) \in \mathbb{S}_0, \quad i = 1, \dots, N_f, \quad (15)$$

$$\text{supp}(\mathbf{s}_g^u) \in \mathbb{J}, \quad (16)$$

where \mathcal{A}_0 represents the set with the complex valued symbols from the adopted APM constellation, including symbol 0. \mathbb{S}_0 denotes the set of valid supports of $\mathbf{s}_g^{u,i}$, according to the possible GSM symbols. \mathbb{J} represents the set of valid supports of $\mathbf{s}_g^{u,i}$ according to all possible active subcarrier combinations. Applying the alternating direction method of multipliers (ADMM), following the approach described in in [42], results in the algorithm of Table 1. In the algorithm, \bar{I} designates the complement of the support set I and Q is the number maximum of iterations.

In line 6, $\Pi_{\mathbb{S}_0}(\cdot)$ denotes the projection over \mathbb{S}_0 which can be completed by choosing the N_a largest magnitude elements, and whose indices also correspond to a valid active antenna combination. If set \mathbb{S}_0 is very large, the projection can be simplified to a basic cardinality-based one. In this situation the projection can be obtained by turning to zeros the $N_s - N_a$ smallest magnitude elements. Step 7 requires computing the projection over set \mathbb{J} , $\Pi_{\mathbb{J}}(\cdot)$, which may be achieved by nulling all, except the N_{af} columns of the matrix $\mathbf{R} = \text{vec}_{N_s \times N_f}^{-1}(\mathbf{s}^{(t+1)} + \mathbf{v}^{(t)})$ which have a higher Euclidean norm and which also match a valid active subcarrier combination. If set \mathbb{J} is very large, the projection can also be simplified to a basic cardinality-based projection. In this case we simply need to zero the $N_f - N_a$ columns of \mathbf{R} with smaller Euclidean norm. Step 8 involves a projection over set $\mathcal{A}_0^{N_s N_f}$, $\Pi_{\mathcal{A}_0^{N_s N_f}}(\cdot)$ which can be computed

elementwise through simple rounding to the nearest element in \mathcal{A}_0 . The projection in step 11 can be performed in a similar way.

Table 1. Algorithm based in ADMM used for GSFIM detection [25].

1:	Input: $\mathbf{x}^0, \mathbf{r}^0, \mathbf{z}^0, \mathbf{u}^0, \mathbf{v}^0, \mathbf{w}^0, \tilde{\mathbf{H}}_g^u, \mathbf{y}_g^u, \rho_x, \rho_r, \rho_z, Q$
2:	$f_{best} = \infty$
3:	$\Phi \leftarrow \left(\left(\tilde{\mathbf{H}}_g^u \right)^H \tilde{\mathbf{H}}_g^u + (\rho_x + \rho_r + \rho_z) \mathbf{I}_{N_f N_s} \right)^{-1}$
4:	for $t = 0, 1, \dots, Q-1$ do
5:	$\mathbf{s}^{(t+1)} \leftarrow \Phi \left(\left(\tilde{\mathbf{H}}_g^u \right)^H \mathbf{y}_g^u + \rho_x (\mathbf{x}^{(t)} - \mathbf{u}^{(t)}) + \rho_r (\mathbf{r}^{(t)} - \mathbf{v}^{(t)}) + \rho_z (\mathbf{z}^{(t)} - \mathbf{w}^{(t)}) \right)$
6:	$(\mathbf{x}^i)^{(t+1)} \leftarrow \Pi_{\mathcal{S}_0} \left((\mathbf{s}^i)^{(t+1)} + (\mathbf{u}^i)^{(t+1)} \right)$
7:	$\mathbf{r}^{(t+1)} \leftarrow \Pi_{\mathbb{J}} \left(\mathbf{s}^{(t+1)} + \mathbf{v}^{(t)} \right)$
8:	$\mathbf{z}^{(t+1)} \leftarrow \Pi_{\mathcal{A}_0^{N_s N_f}} \left(\mathbf{s}^{(t+1)} + \mathbf{w}^{(t)} \right)$
9:	$I \leftarrow \text{supp}(\mathbf{x}^{(t+1)}) \cap \text{supp}(\mathbf{r}^{(t+1)})$
10:	If $t = Q-1$ then
11:	$\tilde{\mathbf{s}}_g^u[\bar{I}] \leftarrow 0$ $\tilde{\mathbf{s}}_g^u[I] \leftarrow \Pi_{\mathcal{A}^{N_{of} N_a}} \left(\left(\left(\tilde{\mathbf{H}}_g^u[:, I] \right)^H \tilde{\mathbf{H}}_g^u[:, I] + 2\sigma^2 \mathbf{I}_{N_{of} N_s} \right)^{-1} \times \left(\tilde{\mathbf{H}}_g^u[:, I] \right)^H \mathbf{y}_g^u \right)$
12:	else
13:	$\tilde{\mathbf{s}}_g^u[\bar{I}] \leftarrow 0, \tilde{\mathbf{s}}_g^u[I] \leftarrow \mathbf{z}^{(t+1)}[I]$
14:	end if
15:	If $f(\tilde{\mathbf{s}}_g^u) < f_{best}$ then
16:	$\hat{\mathbf{s}}_g^u[\bar{I}] \leftarrow 0, \hat{\mathbf{s}}_g^u[I] \leftarrow \tilde{\mathbf{s}}_g^u[I]$
17:	$f_{best} = f(\tilde{\mathbf{s}}_g^u)$
18:	end if
19:	$\mathbf{u}^{(t+1)} \leftarrow \mathbf{u}^{(t)} + \mathbf{s}^{(t+1)} - \mathbf{x}^{(t+1)}$
20:	$\mathbf{v}^{(t+1)} \leftarrow \mathbf{v}^{(t)} + \mathbf{s}^{(t+1)} - \mathbf{r}^{(t+1)}$
21:	$\mathbf{w}^{(t+1)} \leftarrow \mathbf{w}^{(t)} + \mathbf{s}^{(t+1)} - \mathbf{z}^{(t+1)}$
22:	end for
23:	Output: $\hat{\mathbf{s}}_g^u$

4. Numerical Results

In this section we present numerical simulations, both link and system level, for the PT-GSFIM based downlink scheme integrated into a 5G NR system. Results are presented in bit error rate (BER) and measure the performance of the signal across the entire communication chain, from transmitter to receiver. Link performance results, namely, block error rate (BLER) results are used as input by the system level simulator. We consider a virtualized C-RAN, where the network determines which BSs or APs are to be associated with each terminal. The total number of BSs or APs depends on the scenario and are indicated on Tables 1 and 2. Each BS or AP consists of three TRPs, each one equipped with N_{tx} antennas, while

users have N_{rx} antennas. The BS or AP array configuration corresponds to 3 uniform linear arrays, each with 256 antennas, where the separation between antennas of the array is half wavelength [43]. The system level block diagram can be found in references [35, 36, 45, 45].

4.1. Link level simulations

In this section, the performance of the proposed PT-GSFIM scheme is assessed and compared against other MU-MIMO systems. Monte Carlo simulations were run according to the system model presented previously.

Figure 4 shows a comparison of block diagonalization (BD) based MU-MIMO scheme from [37], GSM MU-MIMO from [38], and the PT-GSFIM from [25]. We assume a scenario with $N_u = 8$, $N_a = 2$ and $N_{rx} = 5$ for all schemes, while keeping the same SE, close to 12 bits per channel use (bpcu). SE is computed through Equation (2), divided by N_f . We also employ a total of $N_{tx} = 12N_u$ antennas for all 3 schemes, with PT-GSFIM and BD MU-MIMO using the ADMM algorithm for the detection, and GSM using the OB-MMSE algorithm from [44]. The CRM adopted in the PT-GSFIM scheme, considers a rotation angle φ of 30° . Both GSM and PT-GSFIM are using an 16QAM constellation whereas 64QAM is used in BD MU-MIMO. To ensure a fair comparing, all schemes have a SE close to 12 bpcu per user. It can be observed that the PT-GSFIM scheme outperforms both GSM and traditional MU-MIMO, and the gains are greater when the symbols are distributed over a larger number of subcarriers (N_f). At the end of the paper, we will compare these three MU-MIMO schemes based on system level simulations.

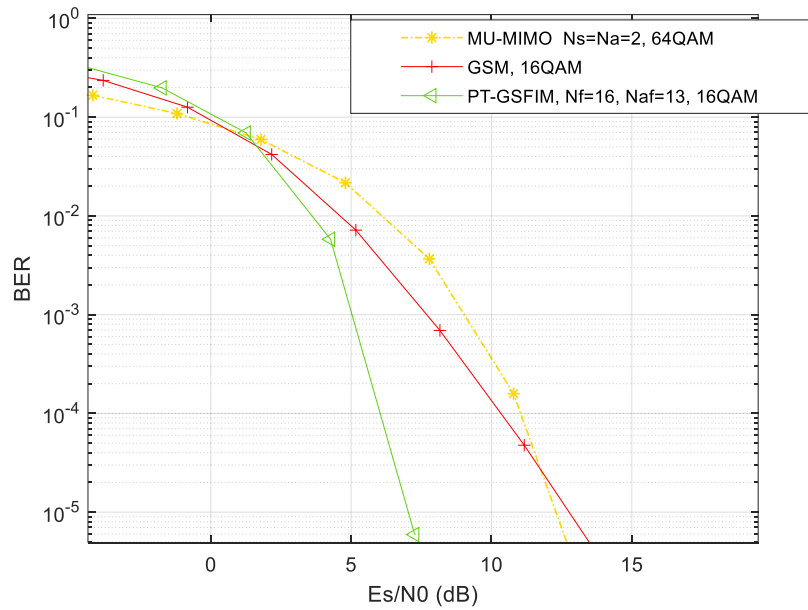


Figure 4. Comparison of BER performances of PT-GSFIM, GSM and conventional BD MU-MIMO, with $N_u = 8$, and SE = 12 bpcu.

A second set of simulations were performed in order to analyze the BLER performance versus the energy per symbol to noise power spectral density (E_s/N_0) in dB for the proposed PT-GSFIM system. For these BLER simulations, a minimum of 25000 blocks were transmitted for computing each BLER result.

In Figure 5, BLER is presented versus (E_s/N_0) in dB, for three cases of M -QAM modulations namely, $M \in \{4, 16, 64\}$. We have different SEs for 64QAM, namely 12.5 bpcu and 11.75 bpcu, for 16QAM the SE is 11.9375 bpcu and for 4QAM (QPSK) we have 6.5 bpcu and 5.75 bpcu. The difference between them is the number of transmitting antennas per user which were 8 antennas per user for 64QAM (12.5 bpcu) and 4QAM (6.5 bpcu), 5 antennas per user are used for 64QAM (11.75 bpcu) and 4QAM (5.75 bpcu), whereas 12 antennas per user are used for 16QAM (11.9375 bpcu). As expected, higher values of M require higher values of E_s/N_0 to reach the reference BLER = 10^{-1} . Higher number of antennas per user increases the bpcu and decreases the E_s/N_0 required to reach the reference BLER.

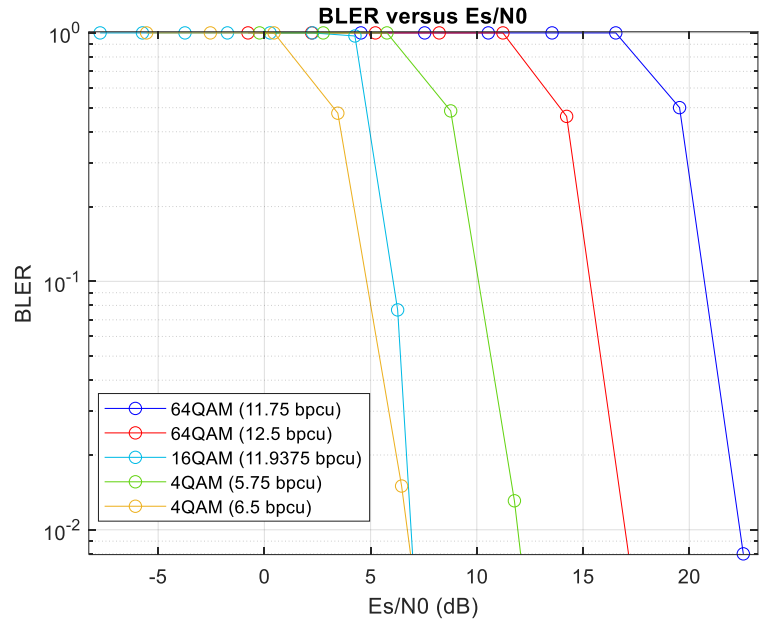


Figure 5. BLER vs (E_s/N_0) dB curves, for PT-GSFIM in a MU scenario with $N_{rx} = 5$, 64QAM (11.75 bpcu and 12.5 bpcu), 16QAM (11.9375 bpcu) and 4QAM (5.75 and 6.5 bpcu).

4.2. System level Simulations

Using the PT-GSFIM link level results, in this subsection we describe several system level simulations that were performed for a C-RAN incorporating PT-GSFIM, considering different scenarios and 5G NR numerologies. The signal-to-noise ratio (SNR) in dB considered in the system level simulations is obtained from $SNR = (E_s/N_0) + 10 \log(R_s/B)$ dB, where R_s is the total transmitted symbol rate per antenna and user, B is the total bandwidth (we considered 20 MHz at 3.5 GHz and 50 MHz at 28 GHz), and E_s/N_0 is the ratio of symbol energy to noise spectral density in dB. Values of E_s/N_0 are obtained from the link level BLER results. 5G NR frame structure has both Frequency Division Duplex (FDD) used in the paired spectrum, while Time Division Duplex (TDD) is used for the unpaired spectrum. We chose TDD in this work and 5G NR numerology 1 for UMi and UMa scenarios at 3.5 GHz, and numerology 2 for InD scenario at 28 GHz [43]. For UMi and UMa scenarios the $B = 20$ MHz with normal CP, where the subcarrier

spacing is 30 kHz and 28 OFDM symbols are transmitted in every subframe of 1 ms. For InD scenario the $B = 50$ MHz, the subcarrier spacing is 60 kHz and 28 OFDM symbols are transmitted in every subframe of 0.5 ms. Other simulator parameters are as follow: maximum TRP transmit power of 49 dBm for UMa, 44 dBm for UMi, or 30 dBm for InD, receiver spectral noise power density -174 dBm/Hz, CP overhead 5%, pilots/TRP = 50 and uniform linear arrays with uni-polarized antennas. Each user feedbacks all CSI and signal-to-interference-plus-noise ratio (SINR) to TRPs during uplink time slots. Equal duration of downlink and uplink time slots is assumed. Users are placed uniformly around TRPs and random waypoint mobility around TRPs is employed. We must consider that the full buffer is used. The static clustering technique partitions the network into three adjacent TRPs sets where each user is served by at least one TRP, while the others perform inter-user interference. When C-RAN cluster size is one it means that there is no C-RAN BS clustering, this is the traditional cellular system. The UMi street canyon scenario is a high dense urban one where 80% of users are inside buildings and 20% of users are in the streets (such as UMa). However, the inter-site distance is 200 m which allows to decrease the total transmitted power from 49 dBm (UMa) to 44 dBm. Scenarios UMa and UMi street canyon can be directly compared because they use both numerology 1 and the same carrier frequency (3.5 GHz) and bandwidth $B = 20$ MHz. However, the scenario InD cannot be directly compared with the others because it uses numerology 2 with the double of transmitted symbols per 1 ms, having a different bandwidth, $B = 50$ MHz, and a carrier frequency of 28 GHz.

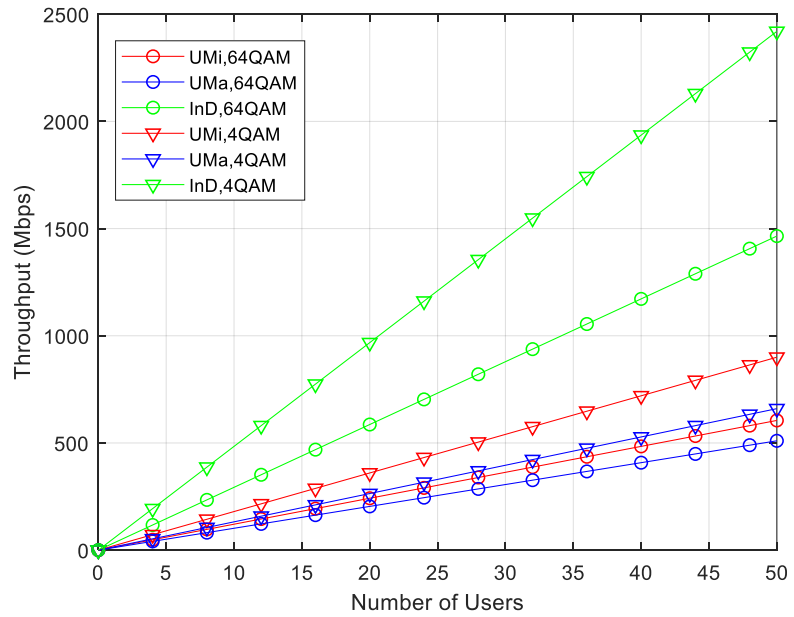


Figure 6. Throughput vs number of users for different scenarios with $N_{tx} = 5N_u$, $N_f = 4$, $N_{af} = 3$, $N_s = 5$, $N_a = 2$, $N_{rx} = 5$, 64-QAM (11.75 bpcu per user) and 4QAM (5.75 bpcu per user), C-RAN cluster size 1.

Based on the parameters of Figure 6, and noting that we considered $N_u = 50$, 5 transmit antennas per user $N_{tx} = 5N_u$, and $N_{rx} = 5$ receive antennas per user, there are a total of 250 active antennas at each sector (one TRP per sector). 64QAM requires a higher SNR compared to 4QAM. InD is expected to offer the highest throughput due to its smaller size. This

occurs independently of the modulation. UMi scenario assures the next best throughput performance whereas UMa is the scenario with the lowest throughput. It is observed that the highest throughput performance is for InD with 4QAM modulated packets. This is due to InD scenario having twice the bit rate compared to the other scenarios. Also, InD has strong inter-interference and high propagation losses, however, most of 4QAM packets are received correctly due to its required SINR being small. By contrast, most of 64QAM packets are received with error because of its required SINR being high, resulting in lower average throughput. In other words, 64QAM packets are only correctly received close to the BS, while 4QAM packets are correctly received almost everywhere. Scenario UMi has also strong inter-interference between sites and there is additional outdoor to indoor propagation loss for pedestrians that are inside the buildings. The throughput performance of 64QAM packets is lower than 4QAM packets in all scenarios. The explanation for this is the same given for InD scenario. Scenario UMa has also strong inter-interference between sites, and due to the highest ISD, users at the cell borders suffer the smallest SINR. It has also additional outdoor to indoor propagation loss for pedestrians that are inside the buildings. As a consequence, the throughput performance of UMa is the worst of the three scenarios. In UMa, the throughput performance of 4QAM packets is slightly higher than 64QAM packets.

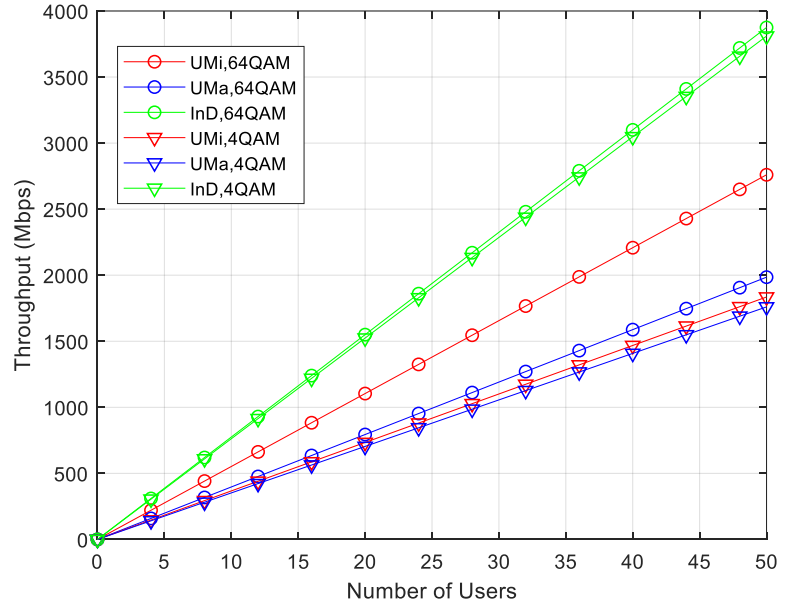


Figure 7. Throughput vs number of users for different scenarios with $N_{tx} = 5N_u$, $N_f = 4$, $N_{af} = 3$, $N_s = 5$, $N_a = 2$, $N_{rx} = 5$, 64-QAM (11.75 bpcu per user) and 4QAM (5.75 bpcu per user), C-RAN cluster size 3.

In Figure 7 the RAN cluster size is 3, so the network is partitioned into three adjacent site sets and each user is served at the same time by three sites (generating much less inner-interference). It is also considered the same maximum $N_u = 50$ from previous scenario. From the comparison between Figures 7 and 6, it is clear that there is an improved throughput for clusters of 3 sites, due to much lower inter-interference between sites, specially for 64QAM in all scenarios. InD continues to offer the highest throughput due to its smaller size and double transmitted bit rate when

compared to the other scenarios. The average throughput gain due to cluster size 3 for UMi is 4.6 for 64QAM and 2.0 for 4QAM. The gain for UMa is 3.9 for 64QAM and 2.7 for 4QAM. The gain for InD is 2.6 for 64QAM and 1.6 for 4QAM. For 4QAM (5.75 bpcu), the increase of throughput between 3C and 1C is less obvious due to its lower SE allowing higher cell coverage. For 50 users in UMi with 3C, the total throughput of 64QAM is 2.760 Gbps, which corresponds to a sector SE of 138 bps/Hz that is equivalent to almost 2.8 bps/Hz per user. For 50 users in InD, the total throughput of 64QAM is 3.875 Gbps, which corresponds to a sector SE of 77.5 bps/Hz that is equivalent to 1.55 bps/Hz per user. This is in agreement with higher frequency bands (28 GHz) which imply higher bandwidths (50 MHz) but lower SE, while lower frequency bands (3.5 GHz) result in lower bandwidths (20 MHz) but higher SE.

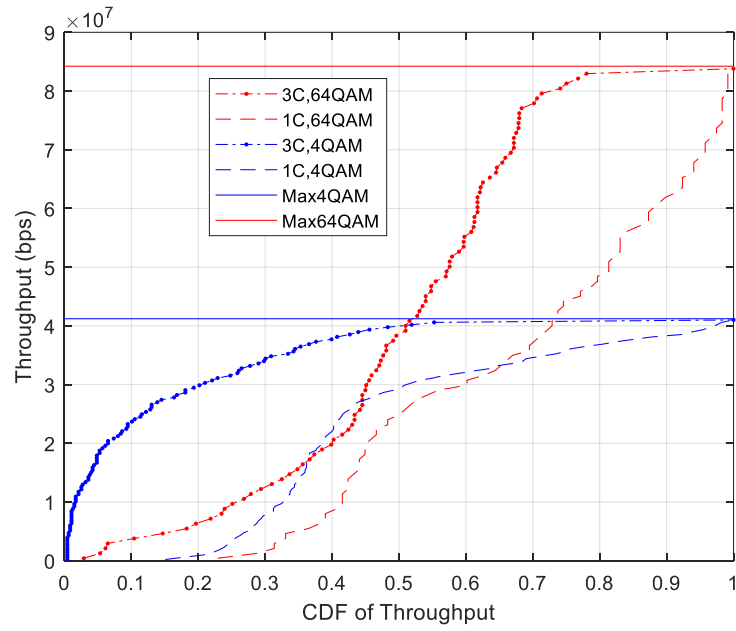


Figure 8. Throughput performance (bps) vs CDF of Throughput for UMa scenario with $N_{tx} = 5N_{tr}$, $N_f = 4$, $N_{af} = 3$, $N_s = 5$, $N_a = 2$, $N_{rx} = 5$, 64QAM (11.75 bpcu per user) and 4QAM (5.75 bpcu per user).

Figure 8 presents the cumulative distribution function (CDF) of each user served by one TRP of UMa scenario, based on the throughput results presented on Figure 6 and Figure 7. This CDF statistics is based on the throughput per user achieved as a function of the distance between terminals and serving TRPs. Each TRP has $N_{tx} = 250$ active antennas serving 50 users each with $N_{rx} = 5$ antennas. We consider the C-RAN UMa scenario with two different cluster sizes namely, a cluster of size one (1C)) and a cluster of size three (3C). As expected, only for 3C there is a percentage of users that achieve the maximum throughput. With 3C, not only users that are close to the TRPs are able to get the maximum throughput. About 20% of users for 64QAM and 50% of users for 4QAM are in this situation of achieving maximum throughput. For 4QAM with 1C, only users that are located quite close to the TRP antenna (less than 1%) are able to get maximum throughput. For 64QAM with 1C no user is able to get the maximum throughput due to high inter-interference. For 4QAM with 1C, only users that are located quite close to the TRP antenna (less than 1%) are able to get maximum throughput. For 64QAM with 1C

no user is able to get the maximum throughput due to high inter-interference. So, for 64QAM most of the packets are not received correctly while for 4QAM many of the packets are correctly received.

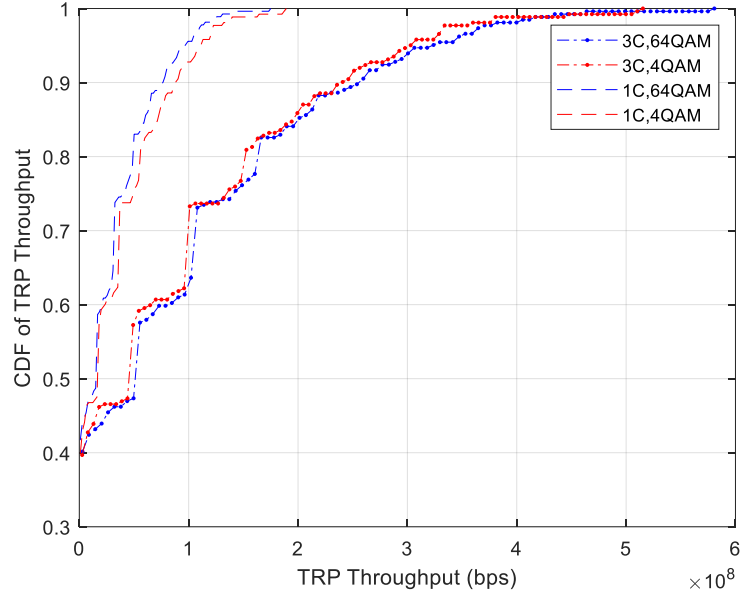


Figure 9. CDF of TRP Throughput vs TRP throughput (bps) for UMa scenario with $N_{tx} = 5N_u$, $N_f = 4$, $N_{af} = 3$, $N_s = 5$, $N_a = 2$, $N_{rx} = 5$, 64-QAM (11.75 bpcu per user) and 4QAM (5.75 bpcu per user).

Figure 9 presents the CDF of a TRP for UMa scenario serving 20 users, during the simulation run with a duration of 100 seconds. These statistics are based on the throughput performance of each one of all users measured at each TTI = 0.5 ms during the simulation run. When we compare the maximum throughput values of Figure 8 and 9, we can check an increase of 14 for 4QAM and 7 for 64QAM. These numbers refer to 3C RAN clustering confirming its lower inter-interference. We can also observe in Figure 9, the difference of three times more throughput achieved by RAN clustering 3C compared to the case 1C. This occurs independently of the modulation. For 3C, there is a slight throughput gain of 64QAM modulated packets compared to 4QAM packets. For 1C, there is a slight throughput gain of 4QAM packets compared to 64QAM. This corroborates the results of Figures 6 and Figure 7, for the value of $N_u = 20$.

In Figure 10, we only consider the UMa scenario, we compare the throughput for three cases of number of transmitting antennas per user N_{tx}/N_u , with RAN clusters size of 1 and 3. Based on the BLER results of Figure 5, the best BLER performance with highest bpcu per user belongs to $N_{tx}/N_u = 12$ followed by $N_{tx}/N_u = 8$ and $N_{tx}/N_u = 5$. However, higher number of transmitting antennas per user implies that lower number of users are served, because the total number of TRP antennas is fixed and equals to 256. The receiving antennas per user is kept fixed and equals to 5. It should be point out, that it is not clear what N_{tx}/N_u would provide higher throughput. From Figure 8, we conclude that the best N_{tx}/N_u ratio depends on the RAN cluster size and modulation. For cluster size 1 the highest throughput performance is for $N_{tx}/N_u = 12$, using 16QAM and the lowest throughput is for 64QAM with $N_{tx}/N_u = 5$.

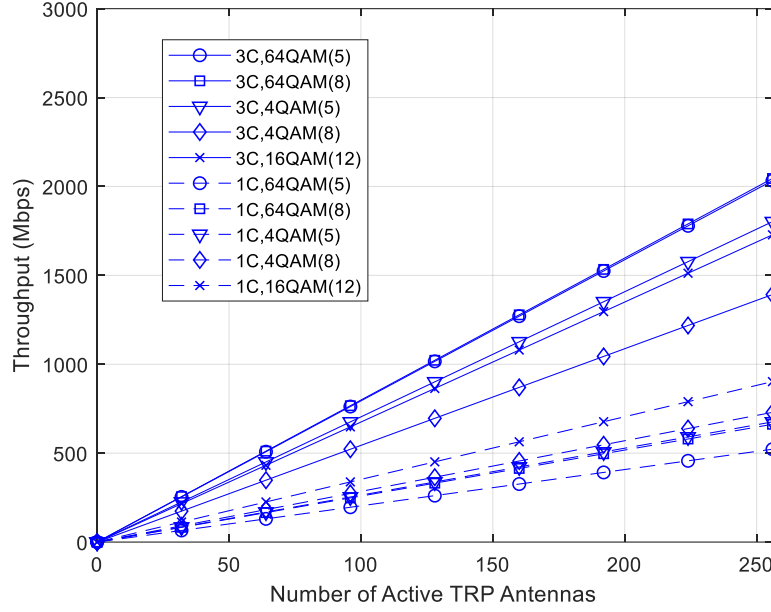


Figure 10. Throughput vs number of active TRP antennas, for UMa scenario, $N_{rx}=5$, for two C-RAN cluster sizes, modulations and number of antennas per user, namely, with $N_{tx}/N_u=5$, 64QAM(5), 4QAM(5), with $N_{tx}/N_u=8$, 64QAM(8), 4QAM(8) and with $N_{tx}/N_u=12$, 16QAM(12).

This result can be explained by the existence of high inter-interference in this cluster size, which affect mostly high order modulations with maximum number of users served at the same time. The case with $N_{tx}/N_u=12$ using a constellation of 16QAM has the minimum number of users served at the same time but has highest bpcu per user, resulting in the lowest interference, which is the best case. However, for cluster size 3 there is much less inter-interference between sites. In this case, the highest throughput performance is achieved for the two 64QAM cases namely, with $N_{tx}/N_u=8$ or $N_{tx}/N_u=5$. The lowest throughput is achieved for $N_{tx}/N_u=8$, using 4QAM. The latter is explained by the low order modulation and bpcu associated to a limited number of served users compared to the case of $N_{tx}/N_u=5$. With a small level of inter-interference of cluster size 3, high order modulations present the best throughput performance. From these results we conclude that there is an optimum N_{tx}/N_u to achieve the highest throughput, but it depends on the cluster size and modulation. Higher N_{tx}/N_u with 16QAM is recommended for C-RAN cluster size 1 with high level of inter-interference and lower number of users per site. Lower N_{tx}/N_u with 64QAM is recommended for RAN cluster size 3 with lower inter-interference but higher number of users per site.

Figure 11 corresponds to Figure 10, but in this case we consider the UMi scenario. When we compare both figures, it is obvious that the best throughput performance is achieved in the UMi scenario. This was expected due to lower ISD, in spite of a higher number of users per unit of area. The total number of TRP antennas is the same and equals to 256. The number of receiving antennas per user is kept fixed and equals to 5. It is interesting to find the N_{tx}/N_u that provides higher throughput.

From Figure 11, we can conclude that the best N_{tx}/N_u depends on the C-RAN cluster size and modulation. For cluster size 1 the highest throughput performance is obtained for $N_{tx}/N_u=12$, using 16QAM and the lowest throughput is achieved for 64QAM with $N_{tx}/N_u=5$. The high inter-

interference of this cluster size makes high order modulation with maximum number of users served at the same time, the worst case of achieved throughput. The case $N_{tx}/N_u = 12$, using 16QAM has the minimum number of users served simultaneously but with highest bpcu per user, and continues to present the lowest interference corresponding to the best case. Cluster size 3 has much less inter-interference between sites. In this case, the highest throughput performance is achieved for 64QAM, with $N_{tx}/N_u = 5$.

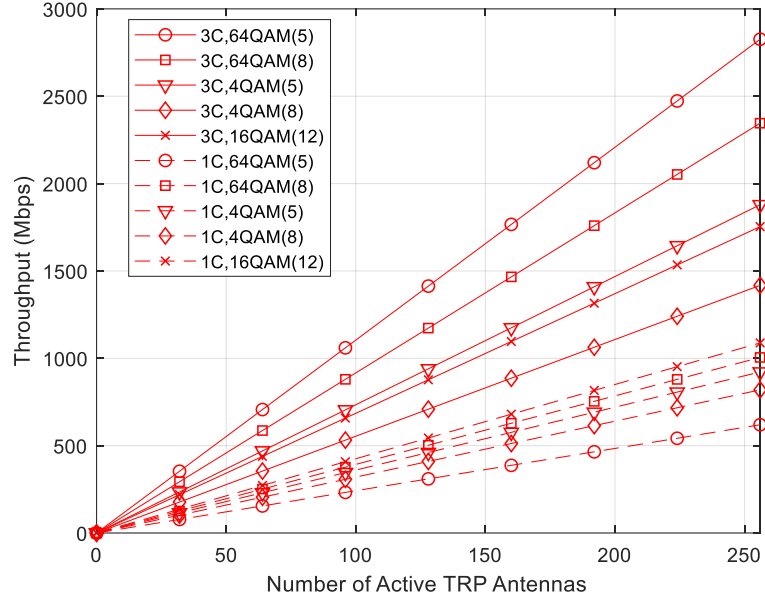


Figure 11. Throughput vs number of active TRP antennas, for UMi scenario, $N_{rx} = 5$, for two C-RAN cluster sizes, modulations and number of antennas per user, namely, with $N_{tx}/N_u = 5$, 64QAM(5), 4QAM(5), with $N_{tx}/N_u = 8$, 64QAM(8), 4QAM(8) and with $N_{tx}/N_u = 12$, 16QAM(12).

The lowest throughput is achieved for $N_{tx}/N_u = 8$, using 4QAM. The latter is explained by the low order modulation and bpcu associated to a reduced number of served users compared to the case of $N_{tx}/N_u = 5$. Cluster size 3 has a small level of inter-interference, which allows the case of 64QAM, with $N_{tx}/N_u = 5$ to achieve higher throughput than 64QAM with $N_{tx}/N_u = 8$, due to a higher number of users with almost the same bpcu. Considering what we have been observing throughout this study, we conclude that there is a value of N_{tx}/N_u that can achieve the highest throughput, but it depends on the cluster size and modulation. Higher N_{tx}/N_u with 16QAM is recommended for C-RAN cluster size 1 with high level of inter-interference and lower number of users per site. 64QAM with $N_{tx}/N_u = 5$ is recommended for C-RAN cluster size 3 with lower inter-interference but higher number of users per site.

Figure 12 corresponds to Figures 10 and 11, but in this case, we consider the InD scenario. When we compare the three figures, it is obvious that the best throughput performance is attained in the InD scenario. This was expected due to the two-fold increase in the transmitted bit rate and lower ISD, in spite of the highest number of users per unit of area. The total number of TRP antennas is the same and equals to 256. The number of receiving antennas per user is kept fixed and equals

to 5. In this case we also wanted to find the value of N_{tx}/N_u that could provide the highest throughput in this InD scenario at 28 GHz.

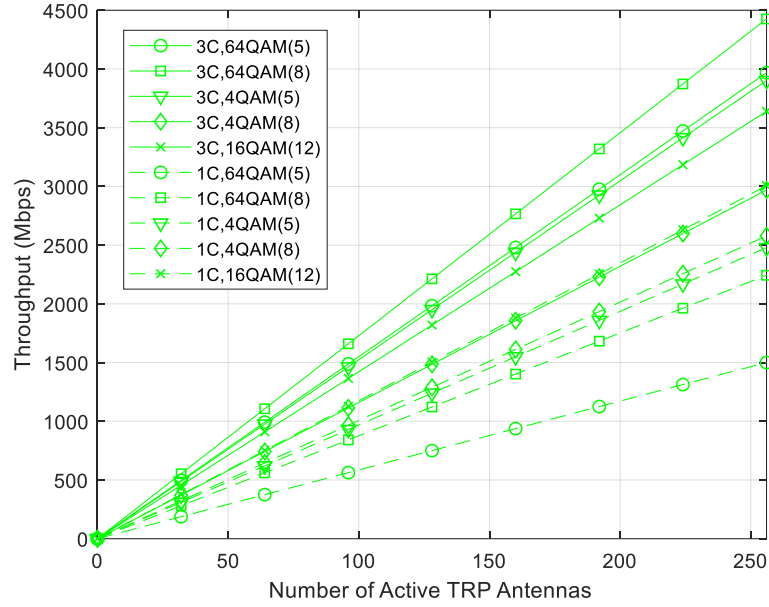


Figure 12. Throughput vs number of active TRP antennas, for InD scenario, $N_{rx} = 5$, for two C-RAN cluster sizes, modulations and number of antennas per user, namely, with $N_{tx}/N_u = 5$, 64QAM(5), 4QAM(5), with $N_{tx}/N_u = 8$, 64QAM(8), 4QAM(8) and with $N_{tx}/N_u = 12$, 16QAM(12).

From Figure 12, we conclude that the best N_{tx}/N_u depends on the C-RAN cluster size and modulation. For cluster size 1 the highest throughput performance continues to be of $N_{tx}/N_u = 12$, using 16QAM and the lowest throughput is obtained for 64QAM with $N_{tx}/N_u = 5$. The high inter-interference of cluster size 1 makes high order modulation with the maximum number of users served at the same time the worst case of achieved throughput. The case of 16QAM with $N_{tx}/N_u = 12$, in spite of having the minimum number of users served at the same time, has the highest bpcu per user and continues to present the lowest interference. For cluster size 3, there is much less inter-interference between sites, and the highest throughput performance is obtained for 64QAM, with $N_{tx}/N_u = 8$. The lowest throughput is achieved for 4QAM with $N_{tx}/N_u = 5$. The latter is explained by the low order modulation and lower bpcu associated to this case in spite of the maximum number of served users per site. Cluster size 3 has a small level of inter-interference and allows 64QAM with $N_{tx}/N_u = 8$ to achieve higher throughput than 64QAM with $N_{tx}/N_u = 5$, due to a higher bpcu despite a lower number of users per site. As before, we conclude that there is a value of N_{tx}/N_u that can achieve the highest throughput, but it depends on the cluster size and modulation. Higher N_{tx}/N_u with 16QAM is recommended for RAN cluster size 1 with high level of inter-interference and lower number of users per site. 64QAM with $N_{tx}/N_u = 8$ is recommended for RAN cluster size 3 with lower inter-interference but higher number of users per site.

Table 2 presents the SE of the three scenarios, namely UMa, UMi and InD, for C-RAN cluster size of 1 and 3, based the throughput results taken from Figures 10, 11 and 12. It is observed that cluster 1C, 16QAM(12) offers the highest SE for all scenarios. For cluster 3C, 64QAM(5) offers the

highest SE for scenario UMi whereas 64QAM(8) offers the highest SE for scenario InD. For UMa scenario both 64QAM(5) and 64QAM(8) offer the highest SE.

Table 2. SE in bps/Hz of different scenarios and cluster size 1 and 3.

Modulation(SE)	UMa,3C	UMi,3C	InD,3C	UMa,1C	UMi,1C	InD,1C
64QAM(5)	2.00	2.75	1.55	0.51	0.61	0.59
64QAM(8)	2.00	2.29	1.73	0.65	0.98	0.88
4QAM(5)	1.76	1.84	1.52	0.66	0.90	0.97
4QAM(8)	1.36	1.38	1.16	0.71	0.80	1.00
16QAM(12)	1.69	1.71	1.42	0.88	1.06	1.17

In Figure 13, we compare the throughput for three different systems, namely, the proposed PT-GSFIM, BD MU-MIMO scheme from [37] and GSM MU-MIMO from [38], where we keep constant $N_{tx}/N_u = 12$ and $N_{sc} = 256$, resulting in approximately 12 bpcu. We evaluate the throughput for

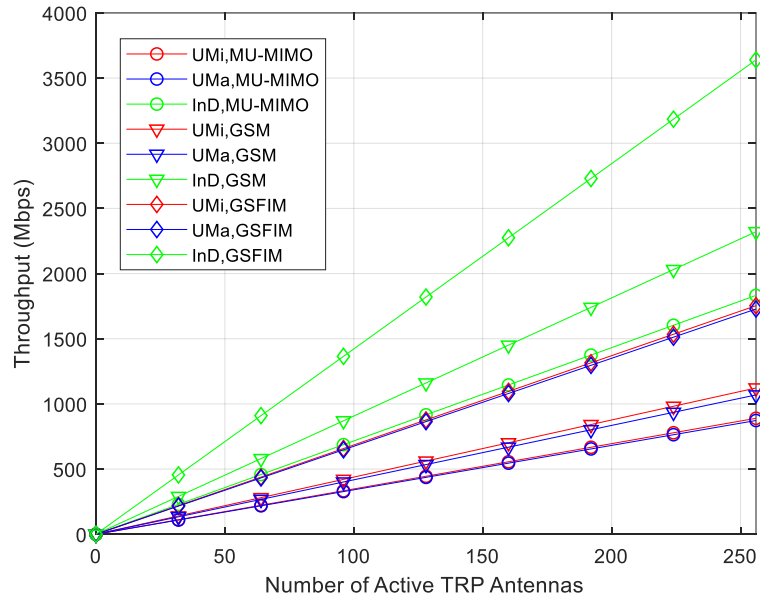


Figure 13. Throughput vs number of active TRP antennas, RAN cluster size 3, for different systems and scenarios, with $N_{tx}/N_u = 12$, $N_{sc} = 256$, $N_a = 2$, $N_{rx} = 5$, 12bpcu, GSM and MU-MIMO with 64QAM and GSFIM with 16QAM $N_f = 16$, $N_{af} = 13$.

three scenarios, namely, UMa, UMi and InD. We observe that independently of the system, the throughput performance of InD scenario is the best as expected. This is due to its double transmitted bit rate compared to the other scenarios, the combination of $N_{tx}/N_u = 12$, with cluster size 3, resulting in smaller number of users having lower inter-interference. Based on the results of Figure 13, the proposed PT-GSFIM scheme clearly outperforms both GSM and conventional MU-MIMO, taking advantage of the symbols being spread over a larger number of subcarriers, exactly, $N_f = 16$, $N_{af} = 13$, thus exploiting the frequency diversity.

Therefore, for C-RAN cluster size three, both indoor and outdoor scenarios indicate that the total throughput achieved by our scheme is

higher than conventional MU-MIMO and better system performance than the GSM counterpart.

In Figure 14, we compare the coverage of C-RAN clusters size 1 and 3, for 64QAM and 4QAM both with $N_{tx}/N_u = 8$. As expected, the coverage of modulation 4QAM is higher than the coverage of 64QAM independently of the cluster size. We also observe that for 100% of transmitted power the coverage gain of cluster size 3 (3C) compared to cluster size 1 (1C) is 1.9 for 4QAM and increases to 3.1 for 64QAM. Cluster size 3 increases especially the coverage of high order modulations such as 64QAM modulation.

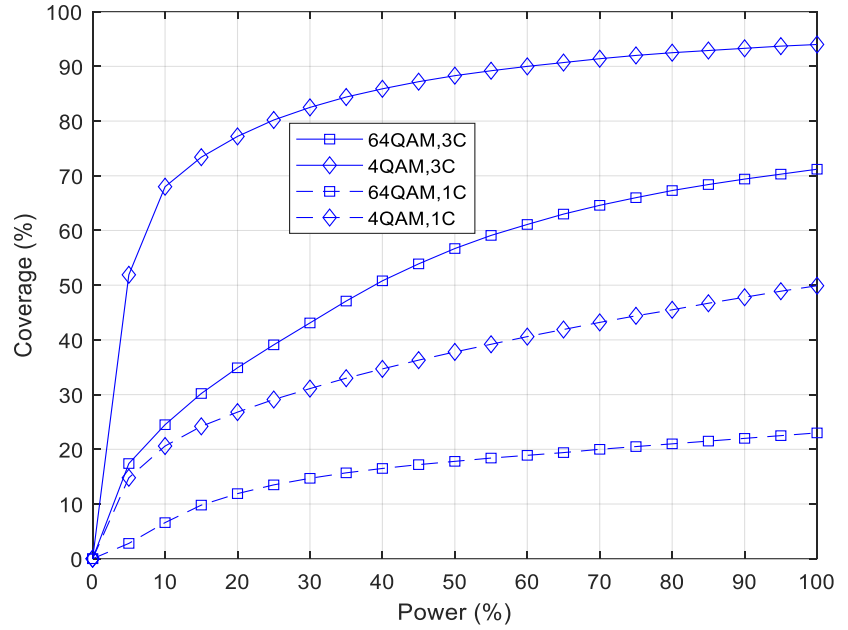


Figure 14. Coverage vs percentage of TRP transmitted power, for UMa scenario, with $N_{tx}/N_u = 8$, $N_{sc} = 256$, $N_{rx} = 5$, 64QAM with 12.5 bpcu and 4QAM with 6.5 bpcu.

5. Conclusions

In this paper we described a C-RAN aimed at 5G and beyond systems, where each BS transmits precoded space-frequency domain IM symbols. In the proposed approach, part of the information is conveyed implicitly on the indices of the active OFDM subcarriers and antennas, which also transmit normal APM signals carrying the remaining bits. The adopted PT-GSFIM scheme can benefit from the diversity effects inherent to frequency selective channels and avoid inter-user interference at each receiver, which enables the C-RAN to accommodate flexible tradeoffs between SE, performance and complexity. A thorough system level evaluation of the proposed PT-GSFIM based C-RAN was performed considering standardized 5G NR scenarios with two different cluster sizes. With cluster size three, the network is portioned into three adjacent TRPs sets where each user is served by three TRPs, while the others perform inter-user interference. For RAN cluster size is one, we have the traditional cellular system. C-RAN system level simulations considered three different three-dimensional scenarios using two different numerologies and frequency ranges. Namely, UMa and UMi scenarios using numerology one at 3.5 GHz and InD scenario using numerology two at 28 GHz. Throughput performance results indicate that, in spite of

its higher bandwidth and higher achieved throughput, the 28 GHz band presents lower SE. The 3.5GHz band having lower bandwidth and lower achieved throughput attains higher SE. There is an optimum number of transmitting antennas per user N_{tx}/N_u to achieve the highest throughput, but it depends on the cluster size of C-RAN and modulation. Higher number of N_{tx}/N_u with 16QAM is optimum for traditional cellular networks with high level of inter-site interference and lower number of users per site. Lower number of N_{tx}/N_u with 64QAM is optimum for C-RAN with lower inter-site interference but higher number of users per site. Overall, the system evaluation shows that the proposed PT-GSFIM scheme can outperform both GSM and conventional MU-MIMO, exploiting its additional inherent frequency diversity. As future work we intend to apply other alternative numerical and probabilistic methods which may be less time intensive in order to achieve accurate system level evaluation results.

Author Contributions: Conceptualization, V.V., J.P.P., N.S., and A.C.; methodology, V.V., J.P.P., N.S., and A.C.; software, V.V., J.P.P., C.R., C.G., N.S., P.S. and A.C.; validation, V.V., J.P.P., C.R., C.G., N.S., P.S. and A.C.; formal analysis, V.V., J.P.P., C.R., C.G., N.S., P.S. and A.C.; investigation, V.V., J.P.P., C.R., C.G., N.S., P.S. and A.C.; resources, A.C. and N.S.; data curation, V.V., J.P.P., C.R., C.G., N.S., P.S. and A.C.; writing—original draft preparation, V.V., J.P.P., N.S. and A.C.; writing—review and editing, V.V., J.P.P., C.R., C.G., N.S., P.S. and A.C.; visualization, V.V., J.P.P., C.R., C.G., N.S., P.S. and A.C.; supervision, A.C. and N.S.; project administration, A.C. and N.S.; funding acquisition, A.C. and N.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the FCT—Fundação para a Ciência e Tecnologia under the grant 2020.05621.BD. The authors also acknowledge the funding provided by FCT/MCTES through national funds and when applicable co-funded EU funds under the project UIDB/50008/2020.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. 3rd Generation Partnership Project (3GPP). TS 38.101 v14.1.1, 5GNR; User Equipment (UE) radio transmission and reception, Release 15, August 2017. Available online: www.3gpp.org/ftp/Specs/archive/38_series/38.101-1/38101-1-001.zip (accessed on 24 January 2022).
2. Marques da Silva, M.; Dinis, R. Power-Ordered NOMA with Massive MIMO for 5G Systems. *Appl. Sci.* **2021**, *11*, 3541. <https://doi.org/10.3390/app11083541>
3. Basar, E., Wen, M., Mesleh, R., Di Renzo, M., Xiao, Y. and Haas, H. Index Modulation Techniques for Next-Generation Wireless Networks. *IEEE Access.* **2017**, *5*, pp.16693-16746.
4. Basar, E. Index modulation techniques for 5G wireless networks. *IEEE Communications Magazine.* **2016**, *54*(7), pp.168-175.
5. Cheng, X., Zhang, M., Wen, M. and Yang, L. Index Modulation for 5G: Striving to Do More with Less. *IEEE Wireless Communications.* **2018**, *25*(2), pp.126-132.
6. Correia, A., Souto, N., Sebastiao, P., Gomez-Barquero, D. and Fuentes, M. Broadcasting Scalable Video with Generalized Spatial Modulation in Cellular Networks. *IEEE Access.* **2020**, *8*, pp.22136-22144.
7. Dogan Tusha, S., Tusha, A., Basar, E. and Arslan, H. Multidimensional Index Modulation for 5G and Beyond Wireless Networks. *Proceedings of the IEEE.* **2021**, *109*(2), pp.170-199.

8. Mesleh, R.; Haas, H.; Sinanovic, S.; Ahn, C.W.; Yun, S. Spatial Modulation. *IEEE Trans. Veh. Technol.*. **2008**, *57*, 2228–2241.
9. Younis, A.; Serafimovski, N.; Mesleh, R.; Haas, H. "Generalised spatial modulation." In Proceedings of the Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA 2010 (7-10 Nov), pp. 1498–1502.
10. Zhang, R., Yang, L. and Hanzo, L. Generalised Pre-Coding Aided Spatial Modulation. *IEEE Transactions on Wireless Communications*. **2013**, *12*(11), pp.5434–5443.
11. Basar, E., Aygolu, U., Panayirci, E. and Poor, H. Orthogonal Frequency Division Multiplexing with Index Modulation. *IEEE Transactions on Signal Processing*. **2013**, *61*(22), pp.5536–5549.
12. Basar, E. On Multiple-Input Multiple-Output OFDM with Index Modulation for Next Generation Wireless Networks. *IEEE Transactions on Signal Processing*. **2016**, *64*(15), pp.3868–3878.
13. Li, J., Wen, M., Jiang, X. and Duan, W. Space-Time Multiple-Mode Orthogonal Frequency Division Multiplexing with Index Modulation. *IEEE Access*. **2017**, *5*, pp.23212–23222.
14. Gao, S., Zhang, M. and Cheng, X. Precoded Index Modulation for Multi-Input Multi-Output OFDM. *IEEE Transactions on Wireless Communications*. **2018**, *17*(1), pp.17–28.
15. Xiao, Y.; Wang, S.; Dan, L.; Lei, X.; Yang, P.; Xiang, W. OFDM With Interleaved Subcarrier-Index Modulation. *IEEE Communications Letters* **2014**, *18*, 1447–1450.
16. Li, J., Dang, S., Wen, M., Jiang, X., Peng, Y. and Hai, H. Layered Orthogonal Frequency Division Multiplexing with Index Modulation. *IEEE Systems Journal*. **2019**, *13*(4), pp.3793–3802.
17. Mao, T., Wang, Z., Wang, Q., Chen, S. and Hanzo, L. Dual-Mode Index Modulation Aided OFDM. *IEEE Access*. **2017**, *5*, pp.50–60.
18. Mao, T., Wang, Q. and Wang, Z. Generalized Dual-Mode Index Modulation Aided OFDM. *IEEE Communications Letters*. **2017**, *21*(4), pp.761–764.
19. Wen, M., Basar, E., Li, Q., Zheng, B. and Zhang, M. Multiple-Mode Orthogonal Frequency Division Multiplexing with Index Modulation. *IEEE Transactions on Communications*. **2017**, *65*(9), pp.3892–3906.
20. Wen, M., Li, Q., Basar, E. and Zhang, W. Generalized Multiple-Mode OFDM With Index Modulation. *IEEE Transactions on Wireless Communications*. **2018**, *17*(10), pp.6531–6543.
21. Althunibat, S., Mesleh, R. and Rahman, T. A Novel Uplink Multiple Access Technique Based on Index-Modulation Concept. *IEEE Transactions on Communications*. **2019**, *67*(7), pp.4848–4855.
22. Basar, E. Media-Based Modulation for Future Wireless Systems: A Tutorial. *IEEE Wireless Communications*. **2019**, *26*(5), pp.160–166.
23. Souto, N. and Correia, A. Frequency Domain Equalization for Single and Multiuser Generalized Spatial Modulation Systems in Time Dispersive Channels. *IEEE Wireless Communications Letters*. **2020**, *9*(3), pp.316–320.
24. Datta, T., Eshwaraiah, H. and Chockalingam, A. Generalized Space-and-Frequency Index Modulation. *IEEE Transactions on Vehicular Technology*. **2016**, *65*(7), pp.4911–4924.
25. Velez, V., Pavia, J., Souto, N., Sebastiao, P. and Correia, A. A Generalized Space-Frequency Index Modulation Scheme for Downlink MIMO Transmissions with Improved Diversity. *IEEE Access*. **2021**, *9*, pp.118996–119009.
26. Wang, L.; Chen, Z.; Gong, Z.; Wu, M. Space-Frequency Coded Index Modulation with Linear-Complexity Maximum Likelihood Receiver in the MIMO-OFDM System. *IEEE Signal Processing Letters*. **2016**, *23*, 1439–1443.
27. Zaidi, S., Ben Smida, O., Affes, S., Vilaipornsawai, U., Zhang, L. and Zhu, P. User-Centric Base-Station Wireless Access Virtualization for Future 5G Networks. *IEEE Transactions on Communications*. **2019**, *67*(7), pp.5190–5202.
28. 3rd Generation Partnership Project (3GPP). TS 38.211 v15.2.0. 5G/NR Physical Channels and Modulation, (Release 15), 2018. Available online: www.3gpp.org/ftp/Specs/archive/38_series/38.211/38211-f30.zip. (accessed on 30 September 2021).
29. 3rd Generation Partnership Project (3GPP). TS 38.214; NR; Physical layer procedures for data; (Release 15), 2020. Available online: www.3gpp.org/ftp/Specs/archive/38_series/38.214/38214-fb0.zip. (accessed on 30 September 2021).
30. 3rd Generation Partnership Project (3GPP). TS 38.213; NR; Physical layer procedures for control; (Release 15), 2020. Available online: www.3gpp.org/ftp/Specs/archive/38_series/38.213/38213-fb0.zip. (accessed on 30 September 2021).
31. Zaidi, A., Baldemair, R., Tullberg, H., Bjorkegren, H., Sundstrom, L., Medbo, J., Kilinc, C. and Da Silva, I. Waveform and Numerology to Support 5G Services and Requirements. *IEEE Communications Magazine*. **2016**, *54*(11), pp.90–98.
32. Begishev, V., Samuylov, A., Moltchanov, D., Machnev, E., Koucheryavy, Y. and Samouylov, K. Connectivity Properties of Vehicles in Street Deployment of 3GPP NR Systems. **2018** *IEEE Globecom Workshops (GC Wkshps)*.

33. Gkonis, P., Trakadas, P. and Kaklamani, D. A Comprehensive Study on Simulation Techniques for 5G Networks: State of the Art Results, Analysis, and Future Challenges. *Electronics*. **2020**, 9(3), p.468.
34. Lakshmi Narasimhan, T. and Chockalingam, A. On the Capacity and Performance of Generalized Spatial Modulation. *IEEE Communications Letters*. **2016**, 20(2), pp.252-255.
35. 3rd Generation Partnership Project (3GPP). TR 36.873 Study on 3D channel model for LTE (Release 12), 2014 Available online: www.3gpp.org/ftp/Specs/archive/36_series/36.873/36873-c00.zip (accessed on 30 September 2021).
36. 3rd Generation Partnership Project (3GPP). TR 38.901 Study on channel model for frequencies from 0.5 to 100 GHz (Release 14), 2017. Available online: www.3gpp.org/ftp/Specs/archive/38_series/38.901/38901-e20.zip (accessed on 30 September 2021).
37. Spencer, Q., Swindlehurst, A. and Haardt, M. Zero-Forcing Methods for Downlink Spatial Multiplexing in Multiuser MIMO Channels. *IEEE Transactions on Signal Processing*. **2004**, 52(2), pp.461-471.
38. Souto, N. and Correia, A. A Precoding Aided Space Domain Index Modulation Scheme for Downlink Multiuser MIMO Systems. *IEEE Transactions on Vehicular Technology*. **2020**, 69(10), pp.12333-12337.
39. Pavia, J.; Velez, V.; Ferreira, R.; Souto, N.; Ribeiro, M.; Silva, J.; Dinis, R. Low Complexity Hybrid Precoding Designs for Multiuser mmWave/THz Ultra Massive MIMO Systems. *Sensors*. **2021**, 21, 6054.
40. Boutros, J.; Viterbo, E. Signal space diversity: a power- and bandwidth-efficient diversity technique for the Rayleigh fading channel. *IEEE Transactions on Information Theory*. **1998**, 44, 1453-1467.
41. Correia, A. Optimized Complex Constellations for Transmitter Diversity, *Wireless Personal Communications Journal*. **2002**, 20(3), pp. 267-284.
42. Lopes, H.; Souto, N. Iterative Signal Detection for Large-Scale GSM-MIMO Systems. *IEEE Transactions on Vehicular Technology* **2018**, 67, 7734-7738.
43. Van Chien, T.; Björnson, E. *Massive MIMO Communications*. In *5G Mobile Communications*; Xiang, W., Zheng, K., Shen, X., Eds.; Springer: Berlin, Germany, **2017**.
44. Xiao, Y.; Yang, Z.; Dan, L.; Yang, P.; Yin, L.; Xiang, W. Low-Complexity Signal Detection for Generalized Spatial Modulation. *IEEE Communications Letters* **2014**, 18, 403-406.
45. Correia, A.; Silva, M.M. *Link and system level simulation for MIMO*. In *MIMO Processing for 4G and Beyond: Fundamentals and Evolution*; Silva, M.M., Monteiro, F.A., Eds.; CRC Press: Boca Raton, FL, USA, **2014**.
46. Pavia, J.; Velez, V.; Brogueira, B.; Souto, N.; Correia, A. Precoded Generalized Spatial Modulation for Downlink MIMO Transmissions in Beyond 5G Networks. *Applied Sciences* **2020**, 10, 6617.
47. 3rd Generation Partnership Project (3GPP). TR 38.913 5G; Study on scenarios and requirements for next generation access technologies, version 16.0.0 Release 16, July 2020. Available online: www.3gpp.org/ftp/Specs/archive/38_series/38.913/38913-g00.zip (accessed on 18 November 2021).

Chapter 4 - Performance Assessment of a RIS-Empowered Post-5G/6G Network Operating at the mmWave/THz Bands

As stated in Section 1.4, a new technology that is emerging in scientific community and could potentially boost the performance for next generation of wireless communications like the 6G, is the use of RIS. Like before the concepts of EE also applies here. This new type of reflecting surfaces can help the transmitted electromagnetic wave to arrive at the target receivers through proper tuning. They are made of several passive elements, which do not need any source of electric power. RIS panels are composed of a certain number of elements that can be tuned individual to *re-shape* the incoming radio wave and direct it to a better or certain point. This can affect the coverage of an array of antennas, or even increase the capacity of system, for example.

In this third article, our primary focus was to demonstrate the benefits how RIS can impact massive/ultra-massive MIMO-OFDM systems within the context of a post-5G or 6G wireless system operating in the mmWave/THz band. To accomplish the joint design of the precoder and RIS phase shifts, we proposed an iterative algorithm based on the alternating maximization (AM) method combined with singular value decomposition (SVD) and water filling, and also with the accelerated proximal gradient (APG) method. The proposed approach considers multiple subcarrier transmission and was evaluated considering the integration into a post-5G/6G network deployment operating in two different bands, one in the mmWave (28 GHz) and the other one on a sub-THz band (100 GHz). As before, the scenarios used are based on urban and indoor scenarios, in which there are differences from the indoor and outdoor scenarios in the previous article, since there are now RIS panels scattered throughout the scenarios. From our results the average throughput does not seem to decrease significantly when the RIS panel is positioned at varying distances from the nearest BS/APs.

All details and information mentioned about RIS are now presented below, including the obtained results.

Paper Details:

- Title: Performance Assessment of a RIS-Empowered Post-5G/6G Network Operating at the mmWave/THz Bands
- Date of Publication: 17 May 2023
- Journal: IEEE Access
- Scimago/Scopus Journal Ranking: Quartile 1
- Publisher: IEEE inc.

Received 18 April 2023, accepted 13 May 2023, date of publication 17 May 2023, date of current version 25 May 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3277388



RESEARCH ARTICLE

Performance Assessment of a RIS-empowered post-5G/6G network operating at the mmWave/THz bands

Vasco Velez^{1,2}, João Pedro Pavia^{2,3}, Nuno Souto^{1,2}, Senior Member, IEEE, Pedro Sebastião^{1,2}, Member IEEE, and Américo Correia^{1,2}, Senior Member, IEEE

¹Departament of Information Science and Technology, ISCTE - Instituto Universitário de Lisboa, 1649-026 Lisboa, Portugal

²Instituto de Telecomunicações, 1049-001 Lisboa, Portugal

³COPELABS, Universidade Lusófona, Campo Grande 376, 1749-024 Lisboa, Portugal

Corresponding author: Vasco Velez (e-mail: Vasco_vez@iscte-iul.pt).

This work was funded by FCT/MCTES through national funds and when applicable co-funded by EU funds under the project UIDB/50008/2020.

ABSTRACT Reconfigurable Intelligent Surfaces (RISs) are considered to be a key enabling technology for 6G as they can potentially provide a boost in performance with a high energy efficiency. RISs rely on the use of arrays with a large number of low-cost quasi-passive reflecting elements which can be individually tuned in order to shape the radio wave propagation. This can effectively enable the implementation of smart radio environments, increasing the capacity and improving the coverage of the system. Since most RISs related studies focus on evaluating the gains of RIS based solutions in simplified communication scenarios, in this paper we investigate the potential benefits of RISs when integrated into future wireless networks within the context of post-5G/ 6G systems. With this aim, we present an iterative algorithm for accomplishing the joint design of the access point precoder and phase-shifts of the RIS elements considering a multi-stream multiple-input multiple output (MIMO) orthogonal frequency division multiplexing (OFDM) link. Based on this approach, we then present the system-level evaluation of a RIS-aided post-5G/6G network deployment operating in two different bands, mmWave and sub-THz, and which considers both near-field and far-field propagation models. The results obtained in two different environments namely, Indoor Open Office (IOO) and Urban Micro Truncated (UMT), show that the adoption of the proposed RIS-based approach can effectively improve the throughput and coverage area.

INDEX TERMS Reconfigurable Intelligent Surfaces, MIMO, System Level Evaluation, mmWave, TeraHertz

I. INTRODUCTION

In recent years, services based on mobile internet have grown exponentially in wireless cellular networks. The considerable increase in the number of connected devices requires the development of new strategies that are capable to improve the quality and reliability of mobile communications. This is an effort that has been enabling the research community to propose several novel solutions for future evolutions of

wireless networks [1]. Increasing spectral and energy efficiency is a principle that is transverse to each generation of mobile communications, whose basis started with multiplexing techniques (e.g. Orthogonal Frequency Division Multiplexing (OFDM) techniques), which were later on expanded with the spatial domain resorting to methods such as multiple input multiple output (MIMO) schemes. Recent improvements include index modulations which can explore the indexes of resource blocks to carry information, for example the spatial resources, frequency resources or even both, as in the case of

The associate editor coordinating the review of this manuscript and approving it for publication was Adao Silva.

generalized spatial-frequency index modulation (GSFIM) [2].

Recent studies have been documented in the literature which can potentially address some of the challenges that must be tackled in future sixth generation wireless networks (6G) [3]. One particular approach within the context of the physical layer encompasses reconfigurable intelligent surfaces (RIS) [4]. This new paradigm shift can potentially increase the quality of communication links in wireless communications [5]. The main idea underlying the use of RIS is the creation of smart environments which can improve wireless communications while also being able to target other type of applications such as location, sensing, and wireless power transfer [6]. Since networks are constantly developing towards a software-based solution, the notion of smart environments is not new.

RIS can be synthesized as an “intelligent” surface comprising a large set of periodic elements that can change the phase (and also the amplitude) of incident waves [7]. These structures are programmable, and they can be used to control electromagnetic wave propagation by generating constructive or destructive interference in the desired directions. Such feature allows us to improve the quality of the links between the transmitter and receiver. RIS may be attached to practically any surface, including walls, furniture, building panels, and clothes. Another advantage of this technology is that its operation is typically quasi-passive, making it a low complexity and energy efficient approach. Due to the low power consumption and the possibility to be embedded in surrounding objects, these surfaces can be seen as a cost-effective solution for future wireless networks [8]. One of the main objectives of using RIS aided systems is to extend the coverage of a wireless network. Considering this purpose, the authors in [9] analyzed the optimal location and orientation of a RIS in order to maximize the cell coverage when considering the communication from a BS to a single user. It was concluded that besides being located at a moderate distance from the BS, the RIS should be deployed vertically to the direction from the BS to the RIS. Although these types of surfaces can implement passive beamforming towards the receiver, they also work in full-duplex [4].

The large distance attenuation that takes place in the millimeter wave (mmWave) and Terahertz (THz) bands makes it difficult to achieve large coverage under the limitation of the maximum available base station (BS) transmit power. This phenomenon tends to worsen especially in systems designed for outdoor environments (even when considering highly directional antenna arrays). Another particular characteristic that is verified at mmWaves and THz bands is that the propagation channel tends to be spatially sparse, which results in a lower number of

propagation paths. Such challenges can be coped with the aid of RISs, since these devices can operate as a centralized beamformer that can increase the channel gains. Moreover, RIS can also create additional propagation paths around major obstacles and enable line-of-sight (LOS) links to distant receivers [10].

In [11], the authors addressed an indoor scenario considering the use of RISs and the application of deep learning algorithms for maximizing the SNRs at the receivers. Their work focused on the design of a deep neural network that can provide the best received signal strength between users and transmitters, with the aid of RIS panels. In [12] the authors studied a system operating at the THz band using a RIS for an indoor and outdoor scenario that optimizes the phase shifts of the individual elements in order to assist an ultra-massive MIMO (UM-MIMO) communication link. Performance results showed that the approach was capable of effectively extending the communication range. The authors in [13], developed an algorithm to calculate the ideal phases for each RIS element in order to maximize the capacity of the transmission. In their study, a MIMO-OFDM link with frequency-selective fading channel and perfect channel state information was considered. Promising gains were shown on high and low SNRs values.

According to the literature, due to their energy efficiency, RIS can be a viable alternative to traditional amplifiers and relays when considering multi-user communication scenarios, as demonstrated by the authors in [14]. The authors in [15] described a system with a single access point (AP) that distributes packets to several users. The system is able to improve the performance of both orthogonal multiple access (OMA) and non-orthogonal multiple access (NOMA) with the aid of RIS. Although, they considered only a single antenna for both transmitter and receiver, it was proved that RIS can enhance both the capacity and rate of the system. Nevertheless, point-to-point communications in MIMO systems aided by RIS are still a challenge.

Although some results obtained with test bench prototypes exist in the literature [16], [17], they all comprise small-scale configurations. Furthermore, the use of RIS-assisted systems combined with MIMO configurations in Cloud Radio Access Network (C-RAN) requires more in-depth research into the optimization process and overall impact. The authors in [18] considered a RIS-aided cellular network and also presented an algorithm for joint optimization of the active beamforming at the BS and passive beamforming of the RIS. Their simulation results showed some performance improvements against other existing algorithms. In [19], the authors presented one of the first system-level studies of a RIS-aided network deployment, using two frequencies of fifth generation new radio (5G NR) namely, the C-band (3.5GHz) and mmWave band (28GHz). While

assuming a simplified operation in the far-field region with the RIS configured as anomalous reflectors, they demonstrated through a three-dimensional simulator how RISs can benefit a typical 5G urban network. Also, in [20], authors studied a system level design with an improved antenna model that analyses the pathloss, power and overall coverage between the transmitter and receiver with the aid of relayed RIS. Their results showed that the impact of the placement of RIS can affect the performance of the system, specially at edges of the sector cell. Despite the promising results of these initial studies, more research on the integration of RIS in future wireless networks is required before large-scale experimental deployments can become a reality. In fact, at the mm-Wave or THz frequencies, the distances between BSs will tend to be short enabling users to be connected to more than one BS simultaneously. Therefore, it is important to study whether RIS panels integrated into these Micro/Pico cells can effectively increase the coverage area as well as the throughput within the context of mmWave/THz networks in post-5G/6G RAN.

Motivated by the above work, in this paper we study the integration of RIS-assisted MIMO communications operating in mmWave /sub-THz bands into future post-5G/6G networks. Considering this objective, we first present an algorithm for joint precoding and RIS optimization which is then used as a basis for the system-level evaluation of a post-5G/6G RAN operating with multiple APs/BSs, multiple RIS panels and multiple users. Several numerical evaluations are presented and analyzed under different configurations, which demonstrate the effectiveness of the proposed optimization algorithm and the beneficial impact of the RISs. The main contributions of this paper can be summarized as follows:

- Targeting scenarios where the communication for different users is based on orthogonal multiple access, we propose a new algorithm that jointly computes the precoder and the phase shifts of RIS panels placed in between the receiver and the transmitter. Such strategy aims to maximize the achievable rate in a multicarrier point-to-point MIMO communication. The proposed algorithm uses the alternating maximization (AM) method to decouple the optimization variables and split the main problem into two simpler ones. The first subproblem is then solved using the singular value decomposition (SVD) combined with water filling whereas the second one is addressed with the accelerated proximal gradient (APG) approach. We refer to the resulting algorithm as AM-SVD-APG;
- We integrated the proposed RIS aided scheme into a post-5G/6G RAN based on MIMO-OFDM which operates at the mmWave and sub-THz bands and considered numerologies 3 and 5

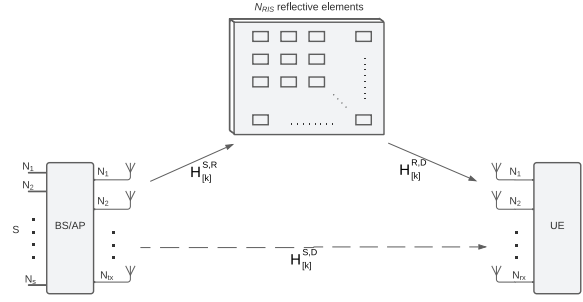


FIGURE 1. Example of a MIMO-OFDM communication system aided by RIS, consisting of BS/AP, UE, and RIS panel with N_{RIS} reflecting elements.

of 5G New Radio in order to perform a thorough system level assessment. For this assessment, both near-field and far-field propagation models were considered.

- A thorough system level evaluation was presented and analyzed for two main scenarios of 5G NR. The first scenario was set for mmWave (28GHz), and considered an Urban Micro Truncated (UMT) environment where several BSs transmit to multiple receivers with the aid of RIS panels placed around the cells. The second scenario was set for sub-THz (100GHz) and considered a deployment in an indoor environment surrounded by obstacles (Indoor Open Office - IOO). Results show that the throughput and coverage area of a post-5G/6G RAN can effectively improve, when integrating the proposed RIS-based approach.

The paper is organized as follows: Section II presents the system model for the RIS-assisted MIMO-OFDM system. Section III derives the joint precoding and RIS design algorithm followed by the description of the evaluated system-level scenarios in section IV. Performance results are then presented and analyzed in section V. Finally, the conclusions are outlined in section VI.

Notation: Bold lower and upper-case letters represent vectors and matrices, respectively. $\mathbb{C}^{a \times b}$ denotes the space of complex matrices of dimensions $a \times b$, $(\cdot)^H$ denotes the conjugate transpose of a matrix/vector, $\|\cdot\|_F$ is the Frobenius norm of a matrix, $\text{diag}(\mathbf{a})$ is a diagonal matrix with elements of \mathbf{a} on its diagonal, and \mathbf{I}_n is the $n \times n$ identity matrix.

II. SYSTEM MODEL

Targeting practical scenarios where the communication for different users is based on orthogonal multiple access, as for example in orthogonal frequency division multiple access (OFDMA), let us consider the downlink connection of a point-to-point MIMO communication system operating at the mmWave/THz band, as illustrated in Figure 1. In this figure, we observe an AP/BS with an

array of N_{tx} antennas transmitting to a user equipment (UE) with N_{rx} antennas, with the aid of a RIS panel with N_{RIS} elements. It is assumed that the system adopts OFDM to cope with frequency selective fading. In this study we essentially consider two scenarios. The first scenario is based on an outdoor environment in which a BS serves several users, with different carrier frequency and distance pathways than those discussed in the second scenario. The second scenario, on the other hand, is based on indoor environment, in which surrounding obstacles can easily block or obstruct the communication link. Such fact presents itself as critical especially in cases where direct LOS is used, or when the signal suffers attenuation losses. Both cases can be assisted with RIS panels with N_{RIS} reflective elements each with the purpose of improving communication links.

Every transmission comprises up to N_s simultaneous data streams per subcarrier k , which are represented as $\mathbf{s}[k] = [s_{1,k} \dots s_{N_s, N_c}]^T$, where $s_{i,k} \in \mathbb{C}$ corresponds to an amplitude and phase modulated symbol with $\mathbb{E}[\|\mathbf{s}[k]\|^2] = N_s$. Using a baseband representation, the signal arriving at a user at each subcarrier k , $\mathbf{R}[k] \in \mathbb{C}^{N_s \times 1}$, can be modeled after the combiner as

$$\mathbf{R}[k] = \sqrt{\rho} \mathbf{W}[k]^H \mathbf{H}[k] \mathbf{F}[k] \mathbf{s}[k] + \mathbf{W}[k]^H \mathbf{n}[k], \quad (1)$$

where $k = 1, \dots, N_c$, N_c is the total number of subcarriers allocated to the target user, ρ denotes the power per stream and per subcarrier, $\mathbf{W}[k] \in \mathbb{C}^{N_{rx} \times N_s}$ is the user combining matrix, $\mathbf{H}[k] \in \mathbb{C}^{N_{rx} \times N_{tx}}$ is the overall frequency domain channel matrix between the BS/AP and user, which includes the paths through the RIS panel, $\mathbf{F}[k] \in \mathbb{C}^{N_{tx} \times N_s}$ is the BS precoder matrix, $\mathbf{s}[k] \in \mathbb{C}^{N_s \times 1}$ is the vector of symbols and $\mathbf{n}[k] \in \mathbb{C}^{N_{rx} \times 1}$ is the noise vector whose elements follow an independent zero mean circularly symmetric Gaussian distribution with covariance $\sigma_n^2 \mathbf{I}_{N_{rx}}$. The total channel matrix $\mathbf{H}[k]$ can be represented as

$$\mathbf{H}[k] = \mathbf{H}[k]^{S,D} + \mathbf{H}[k]^{R,D} \mathbf{\Phi} \mathbf{H}[k]^{S,R}, \quad (2)$$

where, $\mathbf{H}[k]^{S,D} \in \mathbb{C}^{N_{rx} \times N_{tx}}$ is the channel between the BS/AP and the user, $\mathbf{H}[k]^{R,D} \in \mathbb{C}^{N_{rx} \times N_{RIS}}$ is the channel between the RIS panel and user, $\mathbf{\Phi} \in \mathbb{C}^{N_{RIS} \times N_{RIS}}$ is the matrix that models the effect of the RIS elements, having a diagonal structure with $\mathbf{\Phi} = \text{diag}(\boldsymbol{\varphi})$, where $\boldsymbol{\varphi} = [\varphi_1, \dots, \varphi_{N_{RIS}}]^T$ and φ_m denotes the phase shift of the m^{th} element of the RIS, and $\mathbf{H}[k]^{S,R} \in \mathbb{C}^{N_{RIS} \times N_{tx}}$ is the channel between the BS/AP and the RIS panel.

III. JOINT PRECODING AND RIS OPTIMIZATION

After introducing the system model in the previous section, we present the problem formulation which aims at maximizing the achievable rate in a MIMO-OFDM communication link aided by RIS panels. Considering a total of N_f subcarriers (with N_c subcarriers allocated to the target user), a cyclic prefix length of N_{CP} and assuming perfect channel knowledge, then the achievable rate in bits/s/Hz for the user is given by [13]

$$R = \frac{N_f}{(N_f + N_{CP}) N_c} \sum_{k=1}^{N_c} \log_2 \det \left(\mathbf{I}_{N_s} + \frac{\rho}{P_n} \mathbf{F}[k]^H \mathbf{H}[k]^H \times \right. \\ \left. \times \mathbf{H}[k] \mathbf{F}[k] \right). \quad (3)$$

Defining the power constraint at each subcarrier as $\|\mathbf{F}[k]\|_F^2 \leq N_s$ (which results in a total effective transmitted power of $P_{user} = \rho N_s N_c$ for that user) and assuming that the RIS panel only allows tuning the phase shifts of the individual reflecting elements, with this tuning being the same at all subcarriers, then we can formulate the problem as

$$\min_{\substack{\mathbf{F}[k] \in \mathbb{C}^{N_{tx} \times N_s} \\ \boldsymbol{\varphi} \in \mathbb{C}^{N_{RIS} \times 1}}} f(\mathbf{F}[k], \boldsymbol{\varphi}) = - \sum_{k=1}^{N_c} \ln \det \left(\mathbf{I}_{N_s} + \frac{\rho}{P_n} \mathbf{F}[k]^H \times \right. \\ \left. \times \mathbf{H}[k]^H \mathbf{H}[k] \mathbf{F}[k] \right) \\ \text{subject to } \|\mathbf{F}[k]\|_F^2 \leq N_s, \\ |\varphi_i| = a, \quad i = 1, \dots, N_{RIS}, \quad (4)$$

where $\mathbf{F}[k] \in \mathbb{C}^{N_{tx} \times N_s}$ is the BS/AP precoder matrix, P_n denotes the noise power in each sub-band ($P_n = \sigma_n^2$), and a is the amplitude of each reflection element in the RIS. It is important to note that the constants present in the achievable rate expression (3) were dropped in the definition of the objective function $f(\mathbf{F}[k], \boldsymbol{\varphi})$ and that the combiner is not part of the adopted optimization problem.

To solve problem (4), first we use the AM method, in order to decouple the optimization variables and simplify the problem. This approach allows problem (4) to be split into two subproblems which are simpler to address. The first subproblem is defined by fixing $\boldsymbol{\varphi}$ in (4) and then solving over $\mathbf{F}[k]$, which results in the following formulation

$$\min_{\mathbf{F}[k] \in \mathbb{C}^{N_{tx} \times N_s}} - \sum_{k=1}^{N_c} \ln \det \left(\mathbf{I}_{N_s} + \frac{\rho}{P_n} \mathbf{F}[k]^H \mathbf{H}[k]^H \mathbf{H}[k] \mathbf{F}[k] \right) \\ \text{s.t. } \|\mathbf{F}[k]\|_F^2 \leq N_s. \quad (5)$$

This problem can be solved using the SVD combined with a water filling algorithm [21][22]. The SVD of the channel matrix at subcarrier k can be written as

$$\mathbf{H}[k] = \mathbf{U}[k] \mathbf{\Lambda}[k] \mathbf{V}[k], \quad (6)$$

where $\mathbf{U}[k]$ is the matrix with left singular vectors, $\Lambda[k]$ is the diagonal matrix with the singular values in decreasing order, and $\mathbf{V}[k]$ is the matrix containing the right singular vectors. The precoder matrix is then set as $\mathbf{F}[k] = \mathbf{V}_{:,1:N_s}[k] \mathbf{D}[k]$, where $\mathbf{D}[k] = \text{diag}(d_1, \dots, d_{N_s})$ is obtained using a water-filling algorithm applied to the singular values of $\mathbf{H}[k]$, (that is, to the diagonal of $\Lambda[k]$) with a total allocated power of N_s .

The second subproblem corresponds to minimizing (4) over $\boldsymbol{\phi}$, with the precoder matrices $\mathbf{F}[k]$ ($k = 1, \dots, N_c$) fixed and can be rewritten as

$$\min_{\boldsymbol{\phi} \in \mathbb{C}^{N_{ris} \times 4}} - \sum_{k=1}^{N_c} \ln \det \left(\mathbf{I}_{N_s} + \frac{\rho}{P_n} \mathbf{F}[k]^H \mathbf{H}[k]^H \mathbf{H}[k] \mathbf{F}[k] \right) + I_{|\boldsymbol{\phi}|=a}(\boldsymbol{\phi}), \quad (7)$$

where $I_{|\boldsymbol{\phi}|=a}(\boldsymbol{\phi})$ is the set indicator function which returns zero if $|\phi_i| = a$ for any $i = 1, \dots, N_{ris}$, and returns $+\infty$ otherwise. The use of the indicator function allows us to incorporate the RIS phase shifts constraint defined in (3) directly into the objective function, thus resulting in an unconstrained optimization problem. To address the optimization problem in (7) which is nonconvex and difficult to solve due to the constant amplitude phase-shift constraint, we apply an APG method [23]. This method relies on iteratively applying a gradient-based step, followed by a proximal mapping. In order to improve the typical slow convergence of proximal gradient methods [24], the APG method computes the gradient step over an extrapolated point, which is a linear combination of the previous two estimates. Therefore, in the proposed approach we adopt the extrapolated variable defined as

$$\mathbf{y}^{(q)} = \boldsymbol{\phi}^{(q)} + \frac{q}{q+3} (\boldsymbol{\phi}^{(q)} - \boldsymbol{\phi}^{(q-1)}), \quad (8)$$

where q is the iteration number and $q/(q+3)$ corresponds to the extrapolation parameter (as suggested in [23]). The algorithm will thus consist of iteratively updating the estimate $\boldsymbol{\phi}^{(q+1)}$, by applying the proximal operator to a gradient step over $\mathbf{y}^{(q)}$, namely

$$\boldsymbol{\phi}^{(q+1)} = \text{prox}_{\alpha \mathbf{I}_{|\boldsymbol{\phi}|=a}}(\mathbf{y}^{(q)} - \alpha^{(q)} \nabla f(\mathbf{y}^{(q)})), \quad (9)$$

with $\alpha^{(q)}$ denoting the step size (can be found through backtracking line search [23]). The proximal operator for a function g is defined as $\text{prox}_g(\mathbf{z}) = \argmin_{\hat{\mathbf{x}}} g(\hat{\mathbf{x}}) + \frac{1}{2} \|\hat{\mathbf{x}} - \mathbf{z}\|^2$ which for (9), can be calculated as the projection over the set of vectors whose elements have modulus to equal a . Therefore, we can rewrite (9) as

TABLE 1. Algorithm 1: am-svd-apg

ALGORITHM 1: AM-SVD-APG	
1: Input: $\mathbf{R}[k]$, $\mathbf{H}[k]$, Q , $\boldsymbol{\phi}^{(0)}$	
2: For $q = 1, \dots, Q$ do	
3: $\mathbf{F}[k] = \mathbf{V}_{:,1:N_s}[k] \mathbf{D}[k]$, $k = 1, \dots, N_c$	
4: $\mathbf{y}^{(q)} = \boldsymbol{\phi}^{(q)} + \frac{q}{q+3} (\boldsymbol{\phi}^{(q)} - \boldsymbol{\phi}^{(q-1)})$	
5: $\boldsymbol{\phi}^{(q+1)} = (\mathbf{y}^{(q)} - \alpha \nabla f(\mathbf{y}^{(q)})) \oslash [\mathbf{y}^{(q)} - \alpha \nabla f(\mathbf{y}^{(q)})]$	
6: End for	
7: Output: $\mathbf{F}[k]$, $\boldsymbol{\phi}$, $k = 1, \dots, N_c$	

$$\boldsymbol{\phi}^{(q+1)} = (\mathbf{y}^{(q)} - \alpha^{(q)} \nabla f(\mathbf{y}^{(q)})) \oslash [\mathbf{y}^{(q)} - \alpha^{(q)} \nabla f(\mathbf{y}^{(q)})], \quad (10)$$

where \oslash denotes the Hadamard division. The gradient of $f(\mathbf{y})$ can be computed using (the details of the derivation are presented in Appendix A)

$$\begin{aligned} \nabla f(\mathbf{y}^{(q)}) = & -\frac{\rho}{P_n} \sum_{k=1}^{N_c} \text{diag} \left[\left(\mathbf{H}[k]^{R,D} \right)^H \mathbf{H}[k] \mathbf{F}[k] \times \right. \\ & \times \left(\mathbf{I}_{N_s} + \frac{\rho}{P_n} \mathbf{F}[k]^H \mathbf{H}[k]^H \mathbf{H}[k] \mathbf{F}[k] \right)^{-1} \\ & \left. \mathbf{F}[k]^H \left(\mathbf{H}[k]^{S,R} \right)^H \right]. \end{aligned} \quad (11)$$

Table 1, summarizes all steps of the proposed joint precoding and RIS optimization algorithm, which we refer to as AM-SVD-APG.

Regarding the complexity, the proposed algorithm of AM-SVD-APG mainly involves the SVD computations in step 3 as well as some matrix/vector multiplications and a small matrix inversion in step 5. It can be seen that this results in an overall computational complexity order of $\mathcal{O}(Q N_c (N_{tx} N_{rx} N_{ris} + N_{tx} N_{rx}^2 + N_{rx}^3))$. For comparison, the APG algorithm in [12] has a complexity order of $\mathcal{O}(N_c N_{tx} N_{rx}^2 + N_c Q (N_{tx} N_{rx} N_{ris} + N_{rx}^3 + N_{rx}^2 N_{ris} + N_s N_{tx} \times N_{rx}))$ and the projected gradient method (PGM) algorithm from [32] has a complexity of $\mathcal{O}(Q (N_{rx}^3 + N_{rx}^2 N_{tx} + N_{rx} N_{tx}^2 + N_{rx} N_{tx} N_{ris} + N_{tx}^3))$ (it is important to note that PGM does not consider multicarrier transmissions). While the complexity of AM-SVD-APG algorithm has a strong dependency on the receiver array size due to the SVD that is calculated in all iterations, in a typical downlink scenario the receiver array is much smaller than that of the transmitter array and RIS panel. Therefore, the complexity growth of AM-SVD-APG and APG is mainly caused by the product of these three dimensions weighted by the number of sub-carriers and iterations, namely $\mathcal{O}(Q N_c N_{tx} N_{rx} N_{ris})$. As for PGM, it also grows with the product of these dimensions as well as with $\mathcal{O}(Q N_{tx}^3)$.

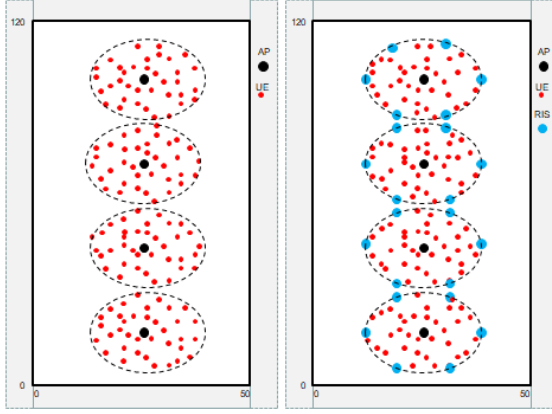


FIGURE 2. IOO environment where black dots are represented as AP, and blue dots RIS panels. Users are shown in red dots that are dispersed in random positions across the floor. IOO scenario without RIS (left) and with RIS assistance (right).

IV. SYSTEM-LEVEL CONFIGURATION AND SCENARIOS

After presenting the signal model for the individual links between BS and user as well as the proposed joint precoding and RIS design algorithm, in this section we describe the setup adopted for the post-5G/6G RAN system level assessment, detailing the target scenarios that were evaluated. The considered scenarios are based on two modified 3GPP environments [25]: the IOO environment, which corresponds to a traditional office space, and the UMT, which corresponds to a truncated Urban Micro Outdoor environment, that is based on urban micro dense areas.

The IOO scenario focuses on high user density inside buildings with offices that have indoor obstacles, such as walls and corridors. This scenario assumes that common offices have areas around $120 \times 50 \text{ m}^2$, as illustrated in Figure 2. In this case we considered sub-THz links with a carrier frequency of $f_c = 100 \text{ GHz}$, and the bandwidth is $B = 0.4 \text{ GHz}$. UEs are uniformly placed, with a minimum distance of 1 meter around APs and RIS. The red dots on Figure 2 show the UEs settled on the scenario. The UEs are randomly and uniformly placed in each scenario at a given distance and angle from the APs or from the associated RIS panel. Depending on the distance, there is a probability of having a LOS and NLOS links. In the case of the NLOS links, it was also considered fading due to multipath following a Rayleigh distribution with shadowing effect according to a Lognormal distribution. All parameters and equations are described in Tables 7.4.1-1 and 7.4.2-1 of [25].

A maximum of two UEs can be served simultaneously by each RIS panel, with the maximum distance between RIS and each UE being 7 m. The panels are represented as the bigger blue dots in Figure 2. On the APs side, they are represented as black dots, and they can support up to 14 UEs associated to each one. The maximum distance between AP and UEs is 17 m. For each AP placed, there are 6 nearest RIS panels coupled to that AP. The maximum distance between AP

and RIS panels, is 17 m. For comparison purposes, we also simulated the same scenario, but without any influence of RIS panels, as shown on the left side of Figure 2.

The UMT scenario illustrated in Figure 3, focuses on micro cells with high user densities. This represents cases of high traffic loads in city centers with dense indoor areas. According to the scenario specifications, it has a truncated area of $470 \times 400 \text{ m}^2$. In our study, the operating frequency was set to $f_c = 28 \text{ GHz}$ and the bandwidth to $B = 0.4 \text{ GHz}$. The Inter Site Distance (ISD) of UMT scenario is 200 m. UEs are also uniformly placed with a minimum distance of 1 meter around their respective BSs and RIS panels. The distribution of UEs across the scenario was based on the same spatial distribution of UEs described before for the previous scenario. The maximum distance between each UE and the associated BS is up to 89 m. Three different configurations were considered for the UMT scenario. In first one, each BS can accommodate up to 25 UEs, and does not have any aid from RISs, as shown on the top image of Figure 3. In the second configuration, almost all RIS panels were placed at 85 m from the nearest BS and the maximum distance between each UE and the corresponding RIS panel is 18 m. Up to two UEs can be linked to each corresponding RIS panel with a total of 20 RIS and 4 BS deployed. On average there are 5 RIS inside each BS coverage area, as represented in the middle image of Figure 3. In the third configuration the RIS panels are placed at 67 m (north and south sides), 58 m (left and right sides) or 88 m (diagonally) from the nearest BS, as displayed on bottom image of Figure 3. In this last deployment, there are 26 RIS panels and 4 BS which means that on average there are 6.5 RIS linked to each BS coverage area. This alternative setup manages to be more realistic, because RIS panels are placed at different distances. We must consider that it may not be always possible to place all RIS panels at exactly the same distance from the BSs, in a realistic deployment.

We consider that each RIS panel can be divided into small panels (called sub-panels), with N_{RIS} corresponding then to the number of elements in each sub-panel. This can be seen in Figure 4 where RISs are integrated into a hexagonal grid on the right size of same figure.

The total power transmitted by each AP at 100 GHz is set to 100mW (20dBm) for the IOO scenario, whereas 3.16W (35dBm) are transmitted from each BS at 28 GHz, in the case of the UMT scenario, see Table 7.8.1 of [25]. To serve the UEs, two types of links are considered namely a direct link between BS/AP and UE and an indirect link through the RIS. In the IOO case, due to the high attenuation at the sub-THz band for NLOS, it is assumed that there is only a direct link or only an indirect link, which is represented as 1C. For the UMT case there are double-links, where UEs are served

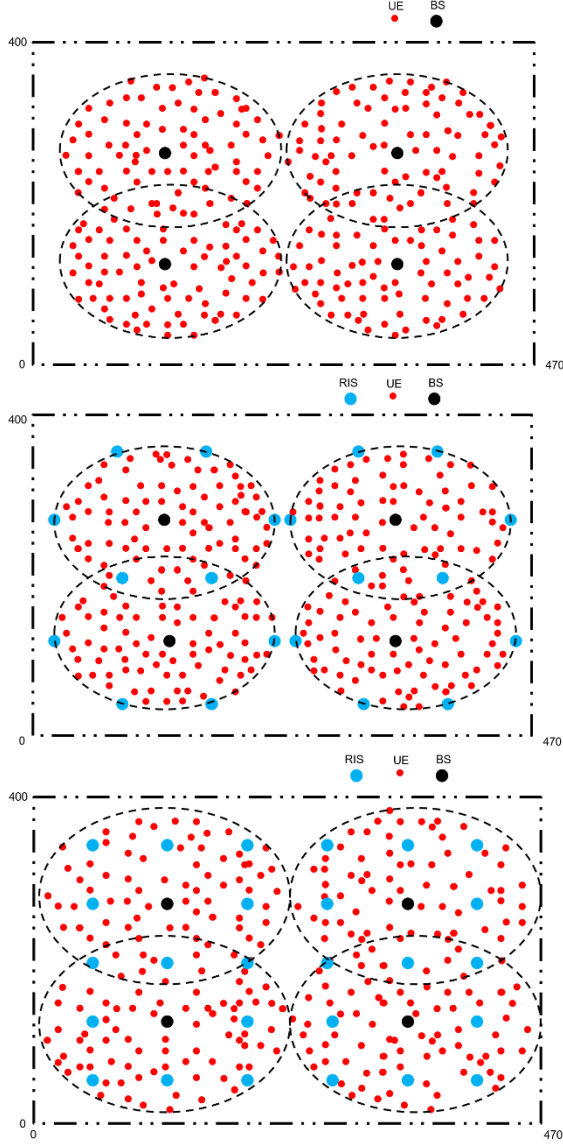


FIGURE 3. UMT scenario, without RIS (top image), with RIS placed at the same distance from the nearest BS (middle image) and with RIS placed at different distances of BS (bottom one).

simultaneously by a RIS and BS, which corresponds to 2C. The noise power is $N_0 = -88$ dBm for the bandwidth $B = 0.4$ GHz. The spacing between each element of the RIS panel is $d_{RIS} = \lambda/2 = 5.4$ mm (28GHz) and $d_{RIS} = 1.5$ mm (100GHz), resulting in an area of $A = 5.4^2 = 29.16$ mm² and $A = 2.25$ mm² per element. The gains of the individual antenna elements of the arrays are 0 dBi for both the transmitter and receiver. Table 2 summarizes the parameters used in our two scenarios.

V. NUMERICAL RESULTS

In this section, we start by presenting the performance of the proposed algorithm AM-SVD-APG using link-level simulations and then extend the evaluation to system level considering the integration into a post-5G/6G RAN. In our evaluation, we assume that the channels can have a LOS component and N_{ray} NLOS

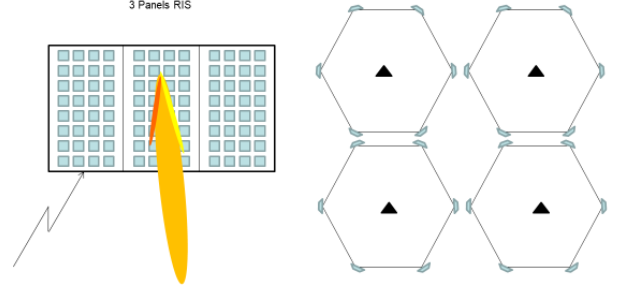


FIGURE 4. Example of RIS Panel with 3 sub-panels used in all simulated environments (left), and a hexagonal grid where 4 BS are located in the center (right).

TABLE 2. Parameters of simulated scenarios.

Parameters	IOO	UMT
Area (m ²)	120x50 m ²	470x400 m ²
Carrier Frequency	100 GHz	28 GHz
Cell Range w/o RIS	10 m	67 m
Cell Range w/ RIS	17 m	85 m
UE mobility	3 km/h	3 km/h
UE distribution (horizontal)	Uniform	Uniform
Maximum UEs attached/sector	Up to 14	Up to 25
Transmit Power	20 dBm	35 dBm
AP/BS Gains	0 dBi	0 dBi
Area of RIS elements	2.25 mm ²	29.46 mm ²

paths. For the case when the RIS is located far from the transmitter, i.e., when the distance between them, $d^{S,R}$, is larger than the Fraunhofer distance defined as $D_F \triangleq 2L_{array}^2/\lambda$ [26] (L_{array} is the largest dimension of the array), we assume a far-field propagation model with planar wavefronts and express the channel frequency response between the AP and RIS as

$$\mathbf{H}[k]^{S,R} = \sqrt{\beta_{LOS}^{S,R}} e^{-j2\pi \frac{d^{S,R}}{\lambda}} \mathbf{a}_R(\phi_0^{R \leftarrow S}, \theta_0^{R \leftarrow S}) \mathbf{a}_S^H(\phi_0^{S \rightarrow R}, \theta_0^{S \rightarrow R}) + \left(\sqrt{\frac{\beta_{NLOS}^{S,R}}{K_{Rice}}} \sum_{l=1}^{N_{ray}} \alpha_l^{S,R} \mathbf{a}_R(\phi_l^{R \leftarrow S}, \theta_l^{R \leftarrow S}) \times \mathbf{a}_S^H(\phi_l^{S \rightarrow R}, \theta_l^{S \rightarrow R}) e^{-j2\pi \tau_l f_k} \right), \quad (12)$$

where $f_k = f_c + \frac{B}{N_c} \left(k - 1 - \frac{N_c - 1}{2} \right)$, B is the bandwidth, f_c is the carrier frequency, τ_l is the delay of path l and K_{rice} specifies the ratio between the LOS and NLOS components. The coefficients $\alpha_l^{S,R}$ represent the complex gains of the NLOS rays, $\beta_{LOS}^{S,R}$ is the path loss of the LOS path and $\beta_{NLOS}^{S,R}$ denotes the path loss of the NLOS channel. This path loss can be approximated as [27]

$$\beta_{NLOS}^{S,R} = \frac{G_{tx} A_R}{4\pi (d_{S \leftrightarrow R})^2} e^{-k_{abs}(f) d_{S \leftrightarrow R}} \quad (13)$$

where G_{tx} is the transmit antenna gain, A_R is the RIS element area, $k_{abs}(f)$ is the molecular absorption coefficient at frequency f [28], and $d_{S \leftrightarrow R}$ is the

distance/length between the transmitter (source) and the RIS panel (RIS).

The vectors $\mathbf{a}_R(\phi_l^{R \leftarrow S}, \theta_l^{R \leftarrow S})$ and $\mathbf{a}_S(\phi_l^{S \rightarrow R}, \theta_l^{S \rightarrow R})$, denote the RIS and transmitter array response vectors at the (azimuth, elevation) angles of $(\phi_l^{R \leftarrow S}, \theta_l^{R \leftarrow S})$ and $(\phi_l^{S \rightarrow R}, \theta_l^{S \rightarrow R})$, respectively. If a uniform planar array (UPA) is adopted, then the steering vector for the transmitter can be written as [29]

$$\mathbf{a}_S(\phi_l^{S \rightarrow R}, \theta_l^{S \rightarrow R}) = \begin{bmatrix} 1, \dots, e^{j \frac{2\pi}{\lambda} d_{ant} (p \sin \theta_l^{S \rightarrow R} \sin \theta_l^{S \rightarrow R} + q \cos \theta_l^{S \rightarrow R})}, \dots, e^{j \frac{2\pi}{\lambda} d_{ant} ((\sqrt{N_{tx}}-1) \sin \theta_l^{S \rightarrow R} \sin \theta_l^{S \rightarrow R} + (\sqrt{N_{tx}}-1) \cos \theta_l^{S \rightarrow R})} \end{bmatrix}^T, \quad (14)$$

where $p, q = 0, \dots, \sqrt{N_{tx}} - 1$ are the indices of the respective antennas, λ is the wavelength and d_{ant} is the inter-element spacing. The steering vectors can also be described using a similar notation for both $\mathbf{a}_R(\phi_l^{R \leftarrow S}, \theta_l^{R \leftarrow S})$ and $\mathbf{a}_D(\phi_l^{D \leftarrow S}, \theta_l^{D \leftarrow S})$. In the case when the distance between the RIS and the transmitter is smaller than the Fraunhofer distance we assume a near-field propagation model with spherical wavefronts. In this case the expression in (12) must include the effect of the distances of the paths between each individual transmit antenna element and each RIS element [30]. Regarding the other channels, namely $\mathbf{H}[k]^{R,D}$ and $\mathbf{H}[k]^{S,D}$, we adopt the same model, and thus they can also be expressed similarly to (12) (for the far field case).

Regarding the link level evaluation, the results are presented in terms of bit error rate (BER) and measure the single-link performance between the transmitter and the user, considering the cases of no RIS and of RIS-aided communication. Whereas the precoder and RIS phases shifts are computed according to Algorithm I, for the combiner we apply a minimum mean squared error (MMSE) based equalizer computed using the equivalent channel as

$$\mathbf{W}^H[k] = \left((\mathbf{H}[k] \mathbf{F}[k])^H \mathbf{H}[k] \mathbf{F}[k] + \frac{\rho}{P_n} \mathbf{I}_{N_s} \right)^{-1} \times (\mathbf{H}[k] \mathbf{F}[k])^H. \quad (15)$$

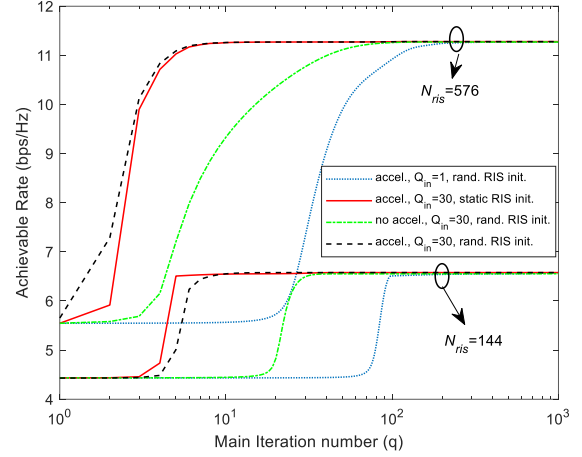


FIGURE 5. Achievable rate versus iterations number obtained with the proposed AM-SVD-APG for a scenario with $f=28\text{GHz}$, $N_{tx}=64$, $N_s=3$, $N_{rx}=16$, and $N_c=1$.

A. LINK LEVEL EVALUATION

The assessment of the RIS optimization algorithm was performed through Monte Carlo simulations for different types of configurations, in a RIS assisted OFDM-MIMO system that operates both in IOO and UMT scenarios. The coordinates of the AP/BS and User are (0 m, 0 m), and (1 m to d , 0 m), respectively. Each variation of parameter d along the simulations, corresponds to changing the distance between the receiver and the transmitter (which will also affect the distance between the UE and the RIS panel).

The subcarrier spacing (SCS), transmission time interval (TTI), cyclic prefix (CP), and the number of symbols per slot, are all defined by the proposed 5G NR scalable OFDM numerology [31]. Our UMT and IOO scenarios were simulated considering the numerologies 3 and 5, respectively. In each scenario, we considered antenna arrays with $N_{tx}=64$ up to 256 elements at the transmitter side. The receiver side has the same number of antennas namely, $N_{rx}=16$, in all simulations. Different numbers of transmitted symbols per subcarrier were considered in the simulations, ranging from $N_s=1$ up to $N_s=3$. The number of OFDM subcarriers (N_c) used in the evaluations is $N_c=60, 120, 132$ or 180 . All these numbers are multiples of 12, which corresponds to the number of subcarriers of a physical resource block in 5G NR [31].

To evaluate the convergence behavior of the proposed AM-SVD-APG algorithm as well as the effect of the initialization, Figure 5 present the achievable rate versus the iteration number, considering the UMT scenario with the following parameters: $N_{tx}=64$, $N_s=3$, $N_{rx}=16$ and $N_c=1$. The RIS panel is located at (50 m, 5

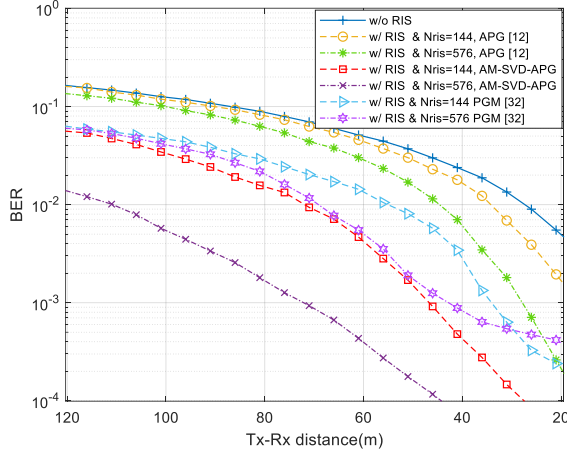


FIGURE 6. BER Comparison versus horizontal distance for a scenario with $f=28\text{GHz}$, $N_{tx}=64$, $N_s=3$, $N_{rx}=16$, $N_{ris}=144$ or 576 elements with $N_c=1$, where RIS panel is placed at a fixed position of (20m, 8m) and the user position changes between the receiver (RX) and the transmitter (TX).

m) and the user at (55 m, 0 m). In the figure we define Q_{in} as the number of times steps 4 and 5 of the algorithm are repeated inside each main iteration. The term ‘*accel*’ is used in the legend when we refer to the algorithm working with the extrapolated variable, as defined in (8). In the ‘*no accel*’ case, the algorithm does not work with the extrapolated variable when updating the phases of the RIS elements, in other words we set the extrapolation parameter to 0 and (8) reduces to $\mathbf{y}^{(q)} = \boldsymbol{\phi}^{(q)}$. We can observe that the accelerated algorithm takes less than 10 main iterations to converge if steps 4 and 5 of the algorithm are repeated several times inside each main iteration ($Q_{in}=30$). This is a consequence of the fact that steps 4 and 5 implement an inner iterative method for providing an estimate for the solution of subproblem (7). Therefore, repeating them several times potentially provides a better solution for the RIS phases before proceeding to the following main iteration and attempt to solve subproblem (5) in step 3 to obtain an updated precoder matrix. It is important to note however, that this fast convergence happens only when using the extrapolation step (accelerated version of AM-SVD-APG) since the non-accelerated algorithm with $Q_{in}=30$ requires at least 50-100 main iterations to converge. Furthermore, it can also be observed that different initializations for the RIS, such as the static RIS with all the elements set as $\phi_i=1$ (simple reflector), or a random initialization of the phases, do not seem to impact the final solution, resulting only in minor variations on the number of required iterations for the algorithm to converge.

Figure 6 shows the BER performance versus the distance for the individual link between the transmitter and a UE. For this comparison we use a power of $P_{user}=35\text{dBm}$ and we fixed the RIS panel at position (20m, 8m). By analyzing this figure, it is possible to observe curves for five different configurations. The

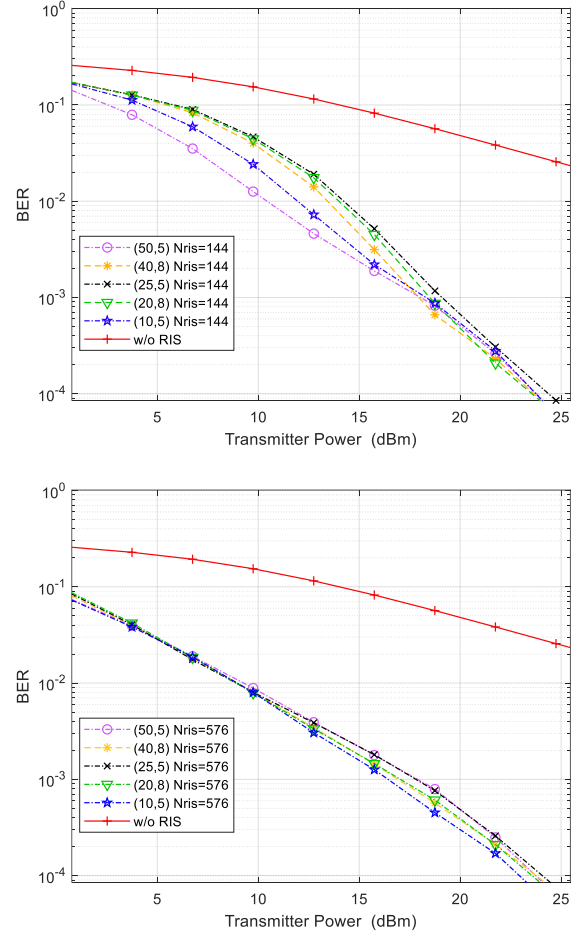


FIGURE 7. BER versus transmitted power when the receiver is fixed at (55 m, 0 m), with a configuration of $f=28\text{GHz}$, $N_{tx}=64$, $N_s=3$, $N_{rx}=16$ and $N_c=132$, but with $N_{ris}=144$ elements (top) or 576 (bottom).

parameters used to draw the comparison on UMT scenario are: $N_{tx}=64$, $N_s=3$, $N_{rx}=16$, $N_{ris}=144$ or 576 elements, and $N_c=1$. Curves with the proposed AM-SVD-APG algorithm are presented and assume the existence of direct and indirect links between transmitter and receiver. To compare with our algorithm, we include results obtained with the APG algorithm from [12], and with the well-known PGM algorithm from [32]. While the curves of the RIS aided communication links are always better than the conventional case without RIS, the best BER performance is achieved with AM-SVD-APG algorithm, with higher gains obtained when using a larger number of RIS elements.

By considering the same base configuration, we also present the BER versus transmitted power allocated to a user (i.e. P_{user}) in Figure 7 considering various positions of the RIS panel. In order to better understand the impact of the RIS, we divide these results into two figures, one with only 144, and the other with 576 RIS elements. The UE was placed at a fixed position of (55, 0) m. On the top image of Figure 7, we can observe the different distances used when a RIS panel with 144 elements is placed between the transmitter and the receiver (UE).

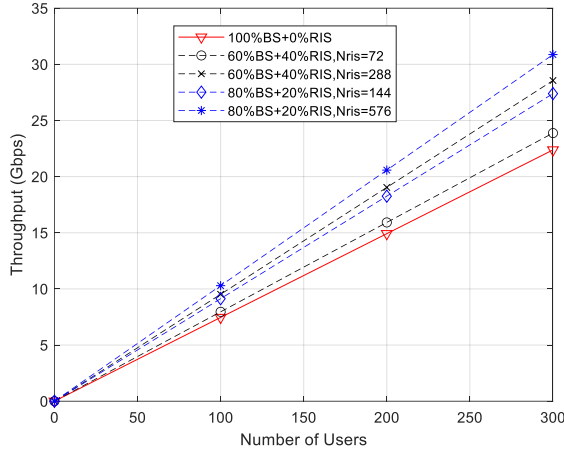


FIGURE 8. Throughput achievable with RISs in the UMT scenario, when $N_s=1$, $N_{tx}=64$, $N_{rx}=16$ and $N_c=132$.

The reference curve corresponds to a link without any aid from a RIS. It can be observed that the best curve obtained with $N_{ris}=144$ corresponds to the case where the RIS is placed at (50,5) m which represents a setup where the UE is closer to the AP/BS. Another curve showing good results correspond to the case of the RIS placed at (10,5) m, i.e., close to the transmitter. As expected, the RIS panel introduces a positive impact in the communication link by decreasing the BER at the UE for the same transmitted power. When the number of RIS elements increases to $N_{ris}=576$, the performance improves, while becoming less sensitive to the placement of the RIS.

B. SYSTEM LEVEL ANALYSIS

In the system level evaluation, we measured the overall throughput across the downlink considering a post-5G/6G RAN integrating the proposed RIS-aided scheme, as described in section IV. The main goal was to evaluate the achievable binary rate or throughput (in Gbps) versus the number of users served by both BS/AP and RIS.

To evaluate how the RIS operates in our system level scenarios, we compared a standard communication without any effect of RISs against a deployment containing RIS panels. The former consists only of direct links between the transmitters and the receivers. The other cases consist of a combination of direct link connections and RIS-aided connections. They are represented as a percentage of users that are receiving signals from the BS plus the percentage of users with RIS connections, namely as %BS + %RIS. For example, 100%BS+0%RIS, means that all users are attached to the BS, whereas 60%BS+40%RIS, represents 60% of spread out users linked to a BS only and the remaining 40% served by RIS panels also. When we consider the UMT scenario, the case 100%BSs+0%RIS assumes that the UEs are uniformly distributed within a radius of 85 m (see Figure 3 top). On the other hand, when

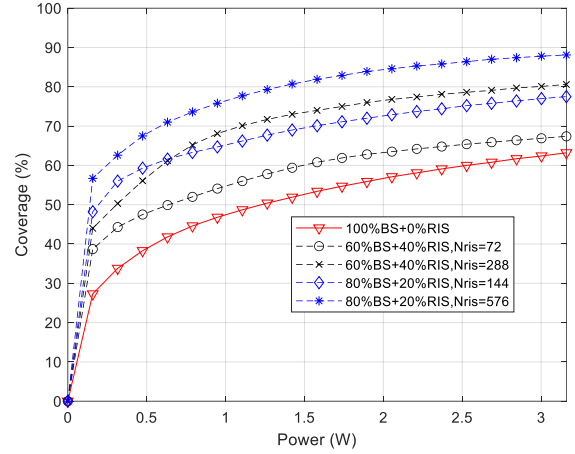


FIGURE 9. Coverage versus transmitted power with RISs in the UMT scenario, when $N_s=1$, $N_{tx}=64$, $N_{rx}=16$ and $N_c=132$.

simulating cases with BSs and RIS operating simultaneously, users that are connected to BSs will be uniformly distributed within a circle with a radius of 67 m. Moreover, the users attached to RIS panels will be uniformly distributed inside the ring defined between previous circle and a circle with a radius of 85 m (see Figure 3 middle). This means that users served by RISs are distributed across the exterior ring (RIS area) which has an area representing 38% of the total area defined by the larger 85 m radius circle.

In Figure 8, it is shown the throughput performance versus number of users in the UMT scenario. The throughput curves of Figure 8 were obtained with a maximum transmitted power of 3.16 W. It is worth mentioning that we kept most of the settings adopted in Figures 6-8. We can observe that in Figure 8, we also provide a curve representing the case without RIS. In these results we used $N_s=1$, $N_c=132$ and a RIS panel with $N_{RIS}=144$ or 576 elements. It is important to note that 4 different RIS sizes appear since each RIS panel is divided into sub-panels when serving more than 1 user. Therefore, the cases of $N_{RIS}=72$ or 288 correspond to cases of 2 users being associated to a panel of 144 and 576, respectively. When the UEs are served with a direct link only (without any RIS), we have a total average throughput of approximately 22.4 Gbps for 300 users (represented as a red line). If RIS panels are added to the C-RAN, the throughput can be increased between 7% to 28% when considering a distribution of UEs of 60%BS+40%RIS, represented by black lines. Note that in this case, N_{RIS} elements correspond to half size, since two UEs have been assigned per RIS. For the other case of 80%BS+20%RIS (blue lines), the increase in throughput is between, 22% and 38%. As we observed before, we can conclude that by increasing the number of RIS elements, the communication link can be boosted. In fact, the case with $N_{RIS}=576$ elements achieve the best performance across the UEs.

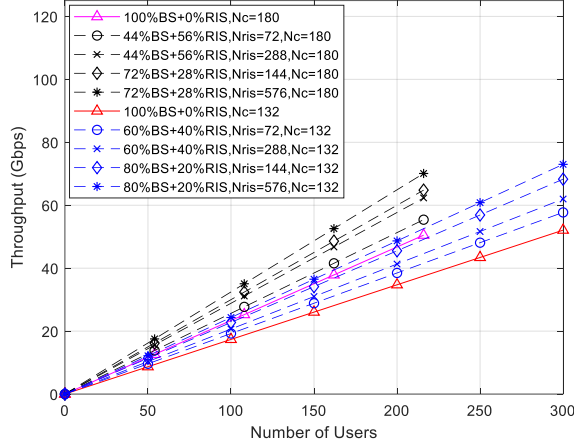


FIGURE 10. Throughput, demonstrating how RIS can affect the rate, as a function of the number of served UEs, when $N_{tx}=64$, $N_s=3$, $N_{rx}=16$. RIS are located at the same distance from nearest BS.

Figure 9 presents the coverage versus transmitted power equivalent for the same conditions of Figure 8. The comparison between Figure 9 and Figure 8 shows that there is a direct correspondence between the throughput performance and the associated coverage. For the maximum transmitted power of 3.16 W, the smallest coverage is achieved by the case of 100%BS+0%RIS with 63% whereas the highest coverage is 88% which is obtained by the case 80%BS+20%RIS with $N_{RIS}=576$. This represents a coverage gain of 38%. In order to provide a clear example of how RISs can impact the performance of the network, we also tested different types of configurations in the UMT scenario. For this purpose, we adopted a higher number of N_c combined with more spatial streams and larger transmit arrays.

Figure 10 presents the throughput results with this different arrangement considering a configuration of $N_{tx}=256$, $N_s=3$, $N_{rx}=16$, $N_{ris}=144$ or 576, and $N_c=132$ or 180. When we increase the number of subcarriers per user (N_c) from 132 to 180, we need to reduce the total number of users, from 300 to 216, in order to keep constant, the total number of subcarriers with data $N_f=3300$. It is observed that the case with distribution of users 80%BS+20%RIS can achieve a gain of 40% when using $N_c=132$, and the case 72%BS+28%RIS achieves a gain 39% when using $N_c=180$. It is important to mention that these two cases ($N_c=132$ and 180) with 100%BS+0%RIS, even with a higher number of UEs connected to BSs, cannot achieve the same throughput as the cases with the aid of RIS. This behavior can be explained because UEs were spread out using a different spatial distribution from the distribution of UEs that are served only by BSs. For a fair comparison, there should be 38% of UEs attached to RIS and the remaining 62% of users attached to BSs. The curves with 60%BS+40%BS or 44%BS+56%RIS have two UEs attached to each RIS panel. They correspond

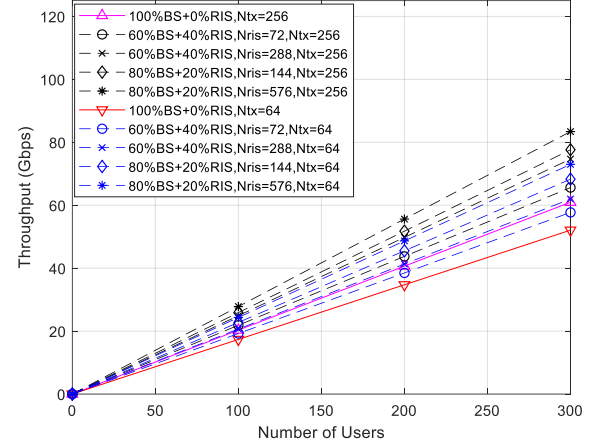


FIGURE 11. Throughput vs number of users when $N_s=3$, $N_{rx}=16$, $N_c=132$. RIS are located at the same distance from the nearest BS.

to half size RIS, namely, $N_{RIS}=288$ and 72. They have a worse spatial distribution of users when compared to the case of all UEs attached to a BS. That is the reason why the performance results of $N_{RIS}=144$ are better than with $N_{RIS}=288$. Considering $N_{RIS}=288$ as reference, the throughput gain for $N_c=132$ is 19% and for $N_c=180$ is 23%. In the case of $N_{RIS}=144$, the throughput gain is 31% for $N_c=132$ and 28% for $N_c=180$. With $N_{RIS}=72$, the throughput gain is 11% for $N_c=132$ and for $N_c=180$ the gain is 10%.

From this point on, we will compare how the transmitter antenna array can affect the overall throughput of the system. In Figure 11, we evaluate the throughput using $N_c=132$ and $N_{tx}=64$ or 256 antennas but keeping the other parameters the same as in the previous case. As it was seen before, the best throughput can be achieved with the highest number of RIS elements, namely with 576 RIS elements. It can be observed that with 80%BS+20%RIS, $N_{RIS}=576$ and $N_{tx}=256$ it is possible to achieve up to 37% better performance than the standard link (100%BS+0%RIS with $N_{tx}=256$). The best case for $N_{tx}=64$, corresponds to 80%BS+20%RIS with $N_{RIS}=576$, where we observe an increase of 40% face to the 100%BS+0%RIS. The worst case for $N_{tx}=256$, corresponds to 60%BS+40%RIS with $N_{RIS}=72$, where it can be seen an increase of 11% compared to the 100%BS+0%RIS case.

Based on these results, we can expect that the proposed RIS based approach can be effective in MIMO schemes with large antenna arrays and also with RIS panels with a large number of elements. To assess this behavior, we tested the same UMT scenario but with RIS panels located at different distance from the nearest BS, as illustrated in Figure 3 (bottom image), as it is a more realistic approach to place the RIS panels.

Figure 12 shows the simulated throughput when considering the bottom scenario of Figure 3, which represents a more realistic/typical deployment. The same parameters configuration adopted in Figure 11 was used in this case, which allows us to verify that

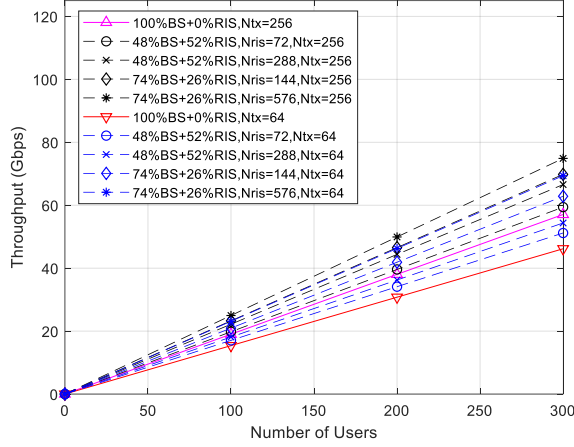


FIGURE 12. Comparison of different achievable throughput using $N_s=3$, $N_r=16$, $N_c=132$ in the UMT scenario shown on bottom image of figure 3.

similar results were obtained. The maximum transmitted power was set as 3.16 W. There is a different spatial distribution of UEs when compared to the previous two figures. In the cases of 74%BS+26%RIS, there is a single UE attached to each RIS, and two for the cases of 48%BS+52%RIS. The latter spatial distribution of UEs is worse than the distribution of UEs for 100%BS+0%RIS, while the opposite occurs for the former one. The spatial distributions that have a better uniformization are close to the cases with spatial distributions of 80%BS+20%RIS and 60%BS+40%RIS respectively. This last one is the best that has a uniform spatial distribution of UEs. The best results are achieved when using $N_{tx}=256$ antennas, a distribution of 74%BS+26%RIS and $N_{RIS}=576$ elements. In such case, we can observe an increase of 58% over the case 100%BS+0%RIS. The worst case happens when we have $N_{tx}=64$, which corresponds to 48%BS+56%RIS, considering $N_{RIS}=72$. In such case, we observed an increase of only 4% compared to the 100%BS+0%RIS results.

Figure 13 presents the coverage versus the total power transmitted by each BS for the same setup of Figure 12. When we compare Figure 13 with Figure 12, we observe that there is a direct correspondence between the throughput performance and the associated coverage. For the maximum transmitted power of 3.16 W with $N_{tx}=64$, the smallest coverage is achieved in the case 100%BS+0%RIS with 43.6% whereas the highest coverage is obtained in the case of 74%BS+26%RIS and $N_{RIS}=576$ with 66.8%. This is a coverage gain of 52%. Moreover, for $N_{tx}=256$, the smallest coverage is achieved for 100%BS+0%RIS resulting in 53.9%, whereas the highest coverage is 70.7% which is obtained for 74%BS+26%RIS with $N_{RIS}=576$. This is a coverage gain of 31%. When comparing the coverage performance of Figures 9, which considers $N_s=1$, against Figure 13, which considers $N_s=3$, we observe

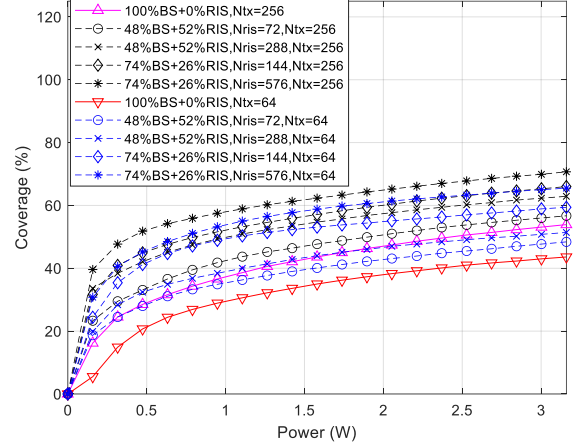


FIGURE 13. Coverage of throughput using fixed parameters as $N_s=3$, $N_r=16$, $N_c=132$, relative to the realistic UMT scenario presented in previous Figure 12.

a coverage loss of 31% as the number of spatial streams per user increases.

From this point on we will consider the IOO environment, in which all RIS are placed at a distance of 10 m from the nearest AP. To simulate the IOO scenario only with APs, we distributed uniformly the UEs within a radius of 17 m. When simulating the system with APs and RIS, the UEs that are connected to APs are uniformly distributed within a radius of 10 m, while the users attached to RIS are uniformly distributed within a radius of 7 m (see Figure 2). The area ratio of a circle with radius 10 m over another with radius 17 m is about 35%. This means that the users served by RIS panels should be 65% of the total on average. Figure 14 shows the throughput that can be achieved with $N_c=60$, and $N_c=120$. This scenario assumes a different carrier frequency than the one considered in the UMT scenario. The system operates at 100 GHz which, due to the shorter wavelength, allows us to work with more elements at the RIS, i.e., $N_{RIS}=256$ up to 2048. It is clear that the best throughput that can be achieved in this scenario corresponds to the cases with a higher number of RIS elements, namely $N_{RIS}=2048$. We can also observe that the instance where a RIS-aided UE is attached to a dedicated RIS, with a distribution of 57%AP + 43%RIS, $N_c=60$ and $N_{RIS}=2048$, has a gain of 29% when compared to the standard link (100%AP+0%RIS). The case labeled as 14%AP + 84%RIS with $N_c=120$, $N_{RIS}=2048$, has a gain of 43% when compared to the reference 100%AP+0%RIS.

Considering the cases with $N_c=120$, we reduce the users that are attached to the AP to keep a single user per RIS. Furthermore, as we double the number of subcarriers per user, N_c , we need to reduce to half the total number of users, namely from 168 to 84, in order to keep constant, the total number of subcarriers allocated with data symbols ($N_f=840$). It can be observed that the case with 14%AP + 86%RIS and N_c

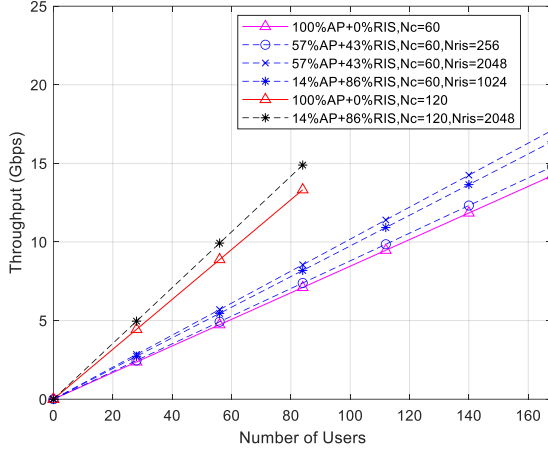


FIGURE 14. Throughput, demonstrating how RIS can affect the rate, in the IOO environment using $N_{tx}=64$, $N_s=2$, $N_{rx}=16$.

$=60$ can achieve higher throughputs than $N_c = 120$, for the maximum number of users served simultaneously. This is explained by the worse spatial distribution of UEs with $N_c = 120$ when compared to the distribution with $N_c = 60$. It is also important to point out that the adoption of RISs allow us to achieve throughput improvements of 24% for 14%AP+86%RIS with $N_{RIS}=1024$ and a gain of 13% when we only use 256 reflective elements (case 57%AP+43%RIS with $N_{RIS}=256$).

VI. CONCLUSION

Due to its potential performance and implementation gains, RISs are considered a key technology for future wireless networks, in particular for post-5G and 6G systems. Within this scope, in this paper we reported system-level assessments of RIS-aided post-5G/6G RAN deployments operating at mmWave (28 GHz) and sub-THz (100 GHz) bands. Considering scenarios where the communication for different users is based on orthogonal multiple access, we first proposed an iterative algorithm for accomplishing joint precoding and RIS optimization in multicarrier point-to-point MIMO communications. Link level simulation results showed that the proposed approach can be more effective than other existing approaches for harvesting the potential gains of RIS-aided communications. This algorithm was then integrated into the system level evaluation of a post 5G/6G RAN operating with multiple BSs, RIS panels and users. The assessment was performed over two different scenarios, IOO and UMT, and considered both near-field and far-field propagation models.

Our numerical results showed that it is possible to provide large transmissions rates with the aid of RIS both in outdoor and indoor environments at mmWave and sub-THz bands. In fact, at 28 GHz the results showed improvements in the overall throughput, with gains of up to 58% over a deployment without any RIS, and coverage gains of up to 53% over the standard

communication without any aid. It was thus observed that it is possible to extend the coverage from cell/BS, and also achieves better SNR at the UEs, increasing the overall performance of the system. By increasing the number of elements in each RIS panel, it is possible to improve the results even further. Furthermore, to obtain the best throughput gains, RIS panels should be placed near the edge of the cells. However, the throughput performance seems to not degrade substantially when RIS are placed at different distances of the nearest BS/AP. As future work, we intend to extend the joint active and passive beamforming algorithm as well as the system level evaluation to multiuser downlink/uplink and multi-cell MIMO scenarios, incorporating realistic low-resolution RIS phases-shifts and imperfect channel knowledge.

VII. APPENDIX A

A. DERIVATION OF $\nabla_{\Phi^*} f(\Phi)$

To derive the complex-valued gradient $\nabla_{\Phi^*} f(\Phi)$ we adopt the procedure described in [33]. First, we write the complex differential of $f(\mathbf{F}[k], \Phi)$ with respect to Φ^* as

$$df = - \sum_{k=1}^{N_c} \text{Tr} \left\{ \frac{\rho}{P_n} \left(\mathbf{I}_{N_s} + \frac{\rho}{P_n} \mathbf{F}[k]^H \mathbf{H}[k]^H \mathbf{H}[k] \mathbf{F}[k] \right)^{-1} \times \right. \\ \left. \times \left(\mathbf{F}[k]^H \mathbf{H}[k]^H d\mathbf{H}[k] \mathbf{F}[k] + \mathbf{F}[k]^H d\mathbf{H}[k]^H \mathbf{H}[k] \mathbf{F}[k] \right) \right\} \quad (16)$$

From the total channel matrix expression (2) we can also directly obtain the following complex differential

$$d\mathbf{H}[k] = \mathbf{H}[k]^{R,D} \text{diag}(d\Phi) \mathbf{H}[k]^{S,R} \quad (17)$$

Inserting this expression into (1) and using the following relation that is simple to verify for a generic matrix \mathbf{Z}

$$\text{Tr}\{\mathbf{Z} \text{diag}(d\Phi)\} = \text{diag}(\mathbf{Z})^T d\Phi, \quad (18)$$

we can rewrite (15) as

$$df = - \frac{\rho}{P_n} \sum_{k=1}^{N_c} \text{diag} \left[\mathbf{H}^{S,R}[k] \mathbf{F}[k] \left(\mathbf{I}_{N_s} + \frac{\rho}{P_n} \mathbf{F}[k]^H \mathbf{H}[k]^H \times \right. \right. \\ \left. \left. \times \mathbf{H}[k] \mathbf{F}[k] \right)^{-1} \mathbf{F}[k]^H \mathbf{H}[k]^H \mathbf{H}^{R,D}[k] \right]^T d\Phi_i - \frac{\rho}{P_n} \sum_{k=1}^{N_c} \times \\ \times \text{diag} \left[\mathbf{H}^{R,D}[k]^H \mathbf{H}[k] \mathbf{F}[k] \left(\mathbf{I}_{N_s} + \frac{\rho}{P_n} \mathbf{F}[k]^H \mathbf{H}[k]^H \mathbf{H}[k] \times \right. \right. \\ \left. \left. \times \mathbf{F}[k] \right)^{-1} \mathbf{F}[k]^H \mathbf{H}^{S,D}[k]^H \right]^T d\Phi_i^* \quad (19)$$

Then using table 3.2 from [33] results in following gradient expression

$$\nabla_{\Phi^*} f(\Phi) = -\frac{\rho}{P_n} \sum_{k=1}^{N_s} \text{diag} \left[\left(\mathbf{H}[k]^{R,D} \right)^H \mathbf{H}[k] \mathbf{F}[k] \times \right. \\ \left. \times \left(\mathbf{I}_{N_s} + \frac{\rho}{P_n} \mathbf{F}[k]^H \mathbf{H}[k]^H \mathbf{H}[k] \mathbf{F}[k] \right)^{-1} \right. \\ \left. \mathbf{F}[k]^H \left(\mathbf{H}[k]^{S,R} \right)^H \right] \quad (20)$$

which corresponds to the equation (11).

ACKNOWLEDGMENT

This work was funded by FCT/MCTES through national funds and when applicable co-funded by EU funds under the project UIDB/50008/2020.

REFERENCES

- [1] Chafii, M., Bariah, L., Muhaidat, S.H., & Debbah, M., "Twelve scientific challenges for 6g: Rethinking the foundations of communications theory," in *arXiv preprint arXiv:2207.01843*, 2022, doi: doi.org/10.48550/arXiv.2207.01843.
- [2] V. R. J. Velez, J. P. C. B. B. Pavia, N. M. B. Souto, P. J. A. Sebastião and A. M. C. Correia, "A Generalized Space-Frequency Index Modulation Scheme for Downlink MIMO Transmissions With Improved Diversity," in *IEEE Access*, vol. 9, pp. 118996-119009, 2021, doi: 10.1109/ACCESS.2021.3106547.
- [3] Basharat, S., Khan, M., Iqbal, M., Hashmi, U.S., Zaidi, S.A.R., Robertson, I., "Exploring reconfigurable intelligent surfaces for 6G: State-of-the-art and the road ahead," in *IET Commun.* 16, 1458–1474, 2022, doi: doi.org/10.1049/cmu2.12364.
- [4] E. Basar, "Transmission Through Large Intelligent Surfaces: A New Frontier in Wireless Communications," 2019 European Conference on Networks and Communications (EuCNC), 2019, pp. 112-117, doi: 10.1109/EuCNC.2019.8801961.
- [5] E. Björnson, Ö. Özdogan and E. G. Larsson, "Reconfigurable Intelligent Surfaces: Three Myths and Two Critical Questions," in *IEEE Communications Magazine*, vol. 58, no. 12, pp. 90-96, December 2020, doi: 10.1109/MCOM.001.2000407.
- [6] E. Björnson, H. Wymeersch, B. Matthiesen, P. Popovski, L. Sanguinetti and E. de Carvalho, "Reconfigurable Intelligent Surfaces: A signal processing perspective with wireless applications," in *IEEE Signal Processing Magazine*, vol. 39, no. 2, pp. 135-158, March 2022, doi: 10.1109/MSP.2021.3130549.
- [7] Yang F, Pitchappa P, Wang N., "Terahertz Reconfigurable Intelligent Surfaces (RISs) for 6G Communication Links," in *Micromachines (Basel)*, 13(2), 285, Feb 2022, doi: 10.3390/mi13020285. PMID: 35208409; PMCID: PMC8879315.
- [8] H. Zhang, B. Di, L. Song and Z. Han, "Reconfigurable Intelligent Surfaces Assisted Communications With Limited Phase Shifts: How Many Phase Shifts Are Enough?," in *IEEE Transactions on Vehicular Technology*, vol. 69, no. 4, pp. 4498-4502, April 2020, doi: 10.1109/TVT.2020.2973073.
- [9] S. Zeng, H. Zhang, B. Di, Z. Han and L. Song, "Reconfigurable Intelligent Surface (RIS) Assisted Wireless Coverage Extension: RIS Orientation and Location Optimization," in *IEEE Communications Letters*, vol. 25, no. 1, pp. 269-273, Jan. 2021, doi: 10.1109/LCOMM.2020.3025345.
- [10] M. A. ElMossallamy, H. Zhang, L. Song, K. G. Seddik, Z. Han and G. Y. Li, "Reconfigurable Intelligent Surfaces for Wireless Communications: Principles, Challenges, and Opportunities," in *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 3, pp. 990-1002, Sept. 2020, doi: 10.1109/TCCN.2020.2992604.
- [11] C. Huang, G. C. Alexandropoulos, C. Yuen and M. Debbah, "Indoor Signal Focusing with Deep Learning Designed Reconfigurable Intelligent Surfaces," 2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), 2019, pp. 1-5, doi: 10.1109/SPAWC.2019.8815412.
- [12] J. Praia, J. P. Pavia, N. Souto, and M. Ribeiro, "Phase Shift Optimization Algorithm for Achievable Rate Maximization in Reconfigurable Intelligent Surface-Assisted THz Communications," *Electronics*, vol. 11, no. 1, p. 18, Dec. 2021, doi: 10.3390/electronics11010018.
- [13] S. Zhang and R. Zhang, "Capacity Characterization for Intelligent Reflecting Surface Aided MIMO Communication," in *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1823-1838, Aug. 2020, doi: 10.1109/JSAC.2020.3000814.
- [14] C. Huang, G. C. Alexandropoulos, A. Zappone, M. Debbah and C. Yuen, "Energy Efficient Multi-User MISO Communication Using Low Resolution Large Intelligent Surfaces," 2018 IEEE Globecom Workshops (GC Wkshps), 2018, pp. 1-6, doi: 10.1109/GLOCOMW.2018.8644519.
- [15] X. Mu, Y. Liu, L. Guo, J. Lin and N. Al-Dhahir, "Capacity and Optimal Resource Allocation for IRS-Assisted Multi-User Communication Systems," in *IEEE Transactions on Communications*, vol. 69, no. 6, pp. 3771-3786, June 2021, doi: 10.1109/TCOMM.2021.3062651.
- [16] J. P. Pavia, Velez, V., Ferreira, R., Souto, N., Ribeiro, M., Silva, J., Dinis, R., "Low Complexity Hybrid Precoding Designs for Multiuser mmWave/THz Ultra Massive MIMO Systems," *Sensors*, vol. 21, no. 18, p. 6054, Sep. 2021, doi: 10.3390/s21186054.
- [17] X. Pei et al., "RIS-Aided Wireless Communications: Prototyping, Adaptive Beamforming, and Indoor/Outdoor Field Trials," in *IEEE Transactions on Communications*, vol. 69, no. 12, pp. 8627-8640, Dec. 2021, doi: 10.1109/TCOMM.2021.3116151.
- [18] S. Liu, P. Ni, R. Liu, Y. Liu, M. Li and Q. Liu, "BS-RIS-User Association and Beamforming Designs for RIS-aided Cellular Networks," 2021 IEEE/CIC International Conference on Communications in China (ICCC), 2021, pp. 563-568, doi: 10.1109/ICCC52777.2021.9580193.
- [19] B. Sihlbom, M. I. Poulakis and M. D. Renzo, "Reconfigurable Intelligent Surfaces: Performance Assessment Through a System-Level Simulator," in *IEEE Wireless Communications*, doi: 10.1109/MWC.015.2100668.
- [20] Gu, Q., Wu, D., Su, X., Wang, H., Cui, J., & Yuan, Y., "System-level Simulation of Reconfigurable Intelligent Surface assisted Wireless Communications System," in *arXiv preprint arXiv:2206.14777*, 2022, doi: doi.org/10.48550/arXiv.2206.14777.
- [21] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585-595, 1999.
- [22] A. Goldsmith, S. A. Jafar, N. Jindal and S. Vishwanath, "Capacity limits of MIMO channels," in *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 684-702, June 2003, doi: 10.1109/JSAC.2003.810294.
- [23] N. Parikh and S. P. Boyd, "Proximal algorithms," *Found. Trends Optim.*, vol. 1, no. 3, pp. 123-231, 2014.
- [24] A. Beck, *First-Order Methods in Optimization* (MOS-SIAM Series on Optimization). Philadelphia, PA, USA: SIAM, 2017.
- [25] 3GPP TR 38.901 V.16.1.0, "5G - Study on channel model for frequencies from 0.5 to 100 GHz", Nov. 2020.
- [26] F. Bohagen, P. Orten and G. E. Oien, "On spherical vs. plane wave modeling of line-of-sight MIMO channels," in *IEEE Transactions on Communications*, vol. 57, no. 3, pp. 841-849, March 2009, doi: 10.1109/TCOMM.2009.03.070062.
- [27] K. Dovelos, S. D. Assimonis, H. Quoc Ngo, B. Bellalta and M. Matthaiou, "Intelligent Reflecting Surfaces at Terahertz Bands: Channel Modeling and Analysis," 2021 IEEE International Conference on Communications Workshops (ICC Workshops), Montreal, QC, Canada, 2021, pp. 1-6, doi: 10.1109/ICCWshops50388.2021.9473890.
- [28] J. M. Jornet and I. F. Akyildiz, "Channel Modeling and Capacity Analysis for Electromagnetic Wireless Nanonetworks in the Terahertz Band," in *IEEE Transactions on Wireless Communications*, vol. 10, no. 10, pp. 3211-3221, October 2011, doi: 10.1109/TWC.2011.081011.100545.
- [29] C. Balanis, *Antenna Theory*. Wiley, 1997.

- [30] J. Huang et al., "Reconfigurable Intelligent Surfaces: Channel Characterization and Modeling," in *Proceedings of the IEEE*, 2022, doi: 10.1109/JPROC.2022.3186087.
- [31] 3GPP TS 38.201 "NR; Physical layer; General description", Release 17, 2021.
- [32] N. S. Perović, L. -N. Tran, M. Di Renzo and M. F. Flanagan, "Achievable Rate Optimization for MIMO Systems With Reconfigurable Intelligent Surfaces," in *IEEE Transactions on Wireless Communications*, vol. 20, no. 6, pp. 3865-3882, June 2021, doi: 10.1109/TWC.2021.3054121.
- [33] A. Hjörungnes, *Complex-Valued Matrix Derivatives With Applications in Signal Processing and Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2011.



VASCO R. J. VELEZ received the B.Sc. and M.Sc. degrees in Telecommunications and Computer Engineering from the ISCTE-IUL, in 2016 and 2018, respectively. In 2017, he joined at IT (Instituto de Telecomunicações) as part of his M.Sc. thesis, he worked in field of spoofing of GNSS signals. He is currently pursuing the Ph.D degree in Information

Science and Technology for effective RF techniques against unauthorized UAVs for his thesis. His research interests include the development of new technologies for communications systems, wireless networks, signal processing for telecommunications, MIMO schemes, MIMO Radar, source localization, GNSS signals and unmanned aerial vehicles.



JOÃO PEDRO C. B. B. PAVIA received his PhD in Information Science and Technology from the ISCTE-University Institute of Lisbon, in 2022. At his early stage research career, he worked for an international project funded by FCT and TUBITAK aiming to understand the interaction between glow discharge detectors and THz waves in order to develop a new

generation of imaging systems for security applications. He was also involved as local organizer in several conferences and symposiums promoted by ISCTE and by Instituto de Telecomunicações (IT). He is currently an assistant professor at Universidade Lusófona and a researcher at COPELABS and at IT. His research interests include the development of new technologies for the Terahertz band, wireless communication networks, signal processing, MIMO schemes, metamaterials, big data, machine learning and cybersecurity. João Pedro serves as reviewer for journals of Multidisciplinary Digital Publishing Institute (MDPI) and he was also a reviewer for the IEEE Vehicular Technology Conference: VTC2021-Fall, VTC2022 Spring and VTC2023-Spring.



NUNO M. B. SOUTO graduated in aerospace engineering-avionics branch, in 2000 in Instituto Superior Técnico, Lisbon, Portugal and received his Ph.D. in 2006. From November 2000 to January 2002, he worked as a researcher in the field of automatic speech recognition for Instituto de Engenharia e Sistemas de Computadores, Lisbon, Portugal. He

joined the ISCTE-University Institute of Lisbon, as an assistant professor in 2006. He is a researcher at IT (Instituto de Telecomunicações), Portugal, since 2002 and has been involved in several international research projects and many national projects.

His research interests include wireless networks, signal processing for communications, OFDM, single carrier transmission with frequency domain equalization, channel estimation, synchronization, MIMO schemes, wireless sensor networks and unmanned aerial vehicles. He is a member of the IEEE Signal Processing Society.



PEDRO J. A. SEBASTIÃO received the Ph.D. degree in electrical and computer engineering from IST. He is currently a Professor with the ISCTE-IUL's Information Science and Technology Department. He is the Board Director of the AUDAX-ISCTE-Entrepreneurship and Innovation Center, ISCTE, responsible for the LABS LISBOA incubator and a

Researcher at the Institute of Telecommunications. He has oriented several master's dissertations and Ph.D. theses. He is the author or coauthor of more than two hundred scientific articles and he has been responsible for several national and international Research and Development projects. He has been an Expert and Evaluator of more than one hundred national and international Civil and Defense Research and Development projects. It has several scientific, engineering, and pedagogical awards. Also, he has organized or co-organized more than fifty national and international scientific conferences. He planned and developed several postgraduate courses in technologies and management, entrepreneurship and innovation, and transfer of technology and innovation. He has supported several projects involving technology transfer and creation of start-ups and spin offs of value to society and market. He has developed his professional activity in the National Defense Industries, initially in the Office of Studies and later as the Board Director of the Quality Department of the Production of New Products and Technologies. He was also responsible for systems of communications technology in the Nokia-Siemens business area. His main research interests include monitoring, control and communications of drones, unnamed vehicles, planning tools, stochastic process (modeling and efficient simulations), the Internet of Things, and efficient communication systems.



AMÉRICO M. C. CORREIA (Senior Member, IEEE) received the B.Sc. degree in electrical engineering from the University of Angola, in 1983, and the M.Sc. and Ph.D. degrees in telecommunications from the Instituto Superior Técnico (IST), Lisbon, Portugal, in 1990 and 1994, respectively. From 1991 to 1999, he was with IST as an Assistant

Professor. From August 2000 to November 2007, he was an Associate Professor with ISCTE. Since December 2007, he has been a Full Professor with ISCTE, Lisbon University Institute, Portugal. He has been a Researcher with the IT (Instituto de Telecomunicações), Portugal, since 1992. He was a Post-Graduated Researcher with RWTH, Aachen, Germany, from October 1994 to September 1995. He visited the Nokia Research Center from September to December 1998 as a Visiting Scientist. From September 2000 to August 2001, he joined Ericsson Eurolab Netherlands. His main research topics include, CDMA, OFDMA, NOMA, MIMO, radio resource management, broadcasting technology, and enhanced multimedia broadcast/multicast services. He is a member of Communications Society of IEEE and has been a member of "Ordem dos Engenheiros" since 1987.

Chapter 5 - Conclusion and Future Work

In this final chapter, it is presented a brief summarization with the final thoughts about the presented work, and the future work which could be continued in the near future. It is divided into two sub-sections: the first one contains a brief summary of the work done, indicating the most important conclusions of the work, and the second one, presents a list of potential research topics that can be explored in a near future.

5.1. Main Conclusions

The aim of this thesis was to propose and evaluate new techniques that exploit both spatial and frequency domains in order to enhance coverage, reducing power consumption, and increasing data rates for future wireless communication networks. With this goal in mind, three different objectives were addressed: IM schemes that utilize both space and frequency domains to target MU downlink communications; integration of the IM schemes into C-RAN deployments in the context of B5G networks; optimization and assessment of RIS assisted techniques for future networks operating in the mmWave/sub-THz range.

As a result of this work, several contributions have been made in this area, all within the scope of MIMO communications: two in index modulations schemes, and one in reconfigurable intelligent surface technologies.

In the first contribution concerning IM algorithms, an approach was taken at the physical level, in a MU-MIMO system, where a BS transmits a precoded signal in the frequency and spatial domains simultaneously. The proposed scheme, called PT-GSFIM, combines previous versions of GSM and OFDM-IM into a multidimensional modulation which is further processed through CRM for improved diversity, and showed good performance results when compared to conventional MU-MIMO and GSM models. In addition, three different detection algorithms were proposed to enable the detection of the PT-GSFIM signal when used in an MUI environment. These detectors also demonstrated a low complexity, in the order of a few tens of magnitude lower than an MLD. As predicted, it was observed that increasing the size of subcarrier subblocks while maintaining the same SE can enhance performance when using SSD. This is because symbols are spread over a greater number of subcarriers, thus taking advantage of the frequency diversity effect. Performance results also demonstrated that the proposed scheme can outperform both GSM and traditional MU-MIMO systems. The data indicates a significant improvement in overall system performance. The initial findings suggest that the proposed scheme may prove to be a promising solution for future wireless communication networks. Still on the subject of IM, it was shown that the same system could be integrated in a C-RAN deployment based on B5G networks. Therefore, a C-RAN was described, in which the aforementioned scheme (PT-GSFIM) was adopted as an alternative to conventional MU-MIMO. A thorough system-level evaluation of the deployed C-RAN was performed in indoor and outdoor scenarios, in which multiple BS and APs transmitted to several

users. Two types of 5G NR frequencies were used according to the 3GPP model channels, 3.5 GHz to the UMa and UMi scenarios and 28 GHz for InD scenario. To differentiate the types of cluster sizes (clusters are the sets of transmission and reception points - TRPs) that make up the zones in which the antennas are directed), two types were used: 1 or 3. It should be noted that when cluster 3 (C3) is used, some MUI is to be expected, since the users are being served by the 3 TRPs located within the BS. Given this situation, the PT-GSFIM scheme managed to outperform the traditional MIMO and conventional GSM scheme in terms of throughput, as in the previous article. However, an interesting fact should be noted: although the values obtained with the 28 GHz frequency bands offer greater bandwidth, their SE seems low, while in the opposite case in the 3.5 GHz band, the contrary is true. In view of this situation, there is an optimum number of transmission antennas to place when using PT-GSFIM, although this depends on the C-RAN cluster used and the modulation to be used.

As far as RIS is concerned, it has an enormous potential in terms of performance and implementation gains and is therefore already being touted as a key technology for the next generations of wireless networks. So, on the subject of RIS, the last contribution focused on a RIS-assisted wireless transmission system which was developed and integrated into a B5G/6G network deployment, operating in two different frequency bands: one at 28 GHz, in the mmWave band, and the other at 100 GHz, already in sub-THz spectrum. It is well known that working in higher frequency bands requires coping with greater attenuations with distance. That said, the proposed scheme for assisting large scale multi-carrier MIMO-type point-to-point communications through the adoption of several RIS panels is accompanied by a precoder which are jointly optimized through the presented AM-SVD-APG algorithm. Tests and evaluations were carried out at link level and system level to show the possible gains offered by RIS. We state that higher number of elements in each panel and their position in relation to the transmitter have a strong influence on the overall system performance. According to the results, the best position is at the border of the cells. Even so, by placing the RIS panel elsewhere, better results can be obtained than without any assistance of RIS, since its performance is not completely degraded.

Regarding the scenarios tested, UMT and IOO, the numerical results show that it was possible to obtain gains in the order of over 50%, both in terms of throughput and coverage compared to a conventional situation. With this coverage increase, it is possible to extend the coverage radius of the BS or cell and improve the SNR values in the receivers, thus improving the overall performance of the entire system. Another important point is that the greater the number of elements in each panel, the better are the results.

5.2. Future Work

This work has presented several areas which can be further explored in future research. Based on the findings outlined in this thesis, further enhancements can be implemented to augment the value of this work. Some potential areas of improvement may include:

- Develop a solution considering the issue of channel estimation for the PT-GSFIM scheme – Channel estimation can still pose a challenge for this algorithm due to the added complexity of working with numerous carriers. However, incorporating error variance into the detection algorithm design can improve the robustness of the scheme;
- Expanding the RIS-aided scheme to work with multiple panels simultaneously for multiple users – Although our previous work used RIS panels that only assumed that each user had their own sub-panel, we can now expand and envision the use of multiple RIS panels for multiple users simultaneously within the same region;
- Expanding on the PT-GSFIM scheme, it would be valuable to develop a form to incorporate the adoption of RIS schemes. For this to work, modifications to the joint optimization algorithm for the precoder and phase-shifts of each RIS panel are needed. Furthermore, the receiver algorithms would also need to be modified;
- Extend the study of RIS to include STAR RIS panels, which not only reflect signals but also allow them to pass through – STAR RIS is a new concept that involves the surface's ability to reflect and transmit the incoming wireless signals, making them semi-transparent for full-space control of signal propagation in communication schemes. The optimization of the algorithm proposed in [41], can also be extended in order to accommodate this type of STAR-RIS panel.

References

- [1] Z. Wang, Y. Du, K. Wei, K. Han, X. Xu, G. Wei, W. Tong, P. Zhu, J. Ma, J. Wa, G. Wang, X. Yan, J. Xiang, H. Huang, R. Li, X. Wang, Y. Wang, S. Sun, S. Suo, Q. Gao and X. Su, "Vision, application scenarios, and key technology trends for 6G mobile communications," *Sci. China Inf. Sci.*, vol. 65, no. 151301, March 2022.
- [2] N. Raman, "Mobile data traffic: how big can disruptive growth be?," 20 12 2022. [Online]. Available: <https://www.linkedin.com/pulse/mobile-data-traffic-how-big-can-disruptive-growth-narayan-raman/>. [Accessed 27 08 2023].
- [3] "Mobile Internet traffic increased by 56.5% compared to the same quarter of 2022," ANACOM, 22 06 2023. [Online]. Available: <https://www.anacom.pt/render.jsp?contentId=1746694>. [Accessed 18 07 2023].
- [4] Ericsson, "Ericsson Mobility Report," Ericsson, Stockholm, Sweden, June, 2023.
- [5] C. Nelson, "Formula 1 Augmented Reality: The Race to the Metaverse is On," LinkedIn, 24 10 2022. [Online]. Available: <https://www.linkedin.com/pulse/formula-1-augmented-reality-race-metaverse-craig-nelson?trk=pulse-article>. [Accessed 25 07 2023].
- [6] "'People will love it': The future market in autonomous driving," IAA-Mobility, 05-10 09 2023. [Online]. Available: <https://www.iaa-mobility.com/en/visionary-club/season-3/the-future-market-in-autonomous-driving>. [Accessed 20 09 2023].
- [7] "Mobileye at IAA Mobility 2023," Mobileye, 2023. [Online]. Available: <https://www.mobileye.com/iaa-2023/>. [Accessed 11 09 2023].
- [8] "Mobileye: The Road to Fully Autonomous Vehicles," IAA mobility, 2023. [Online]. Available: <https://www.iaa-mobility.com/en/newsroom/news/autonomous-driving/the-road-to-fully-autonomous-vehicles>. [Accessed 11 09 2023].
- [9] A. Feldmann, O. Gasser, F. Lichtblau, E. Pujol, I. Poesse, C. Dietzel, D. Wagner, M. Wichtlhuber, J. Tapiador, N. V.-. Rodriguez, O. Hohlfeld and G. Smaragdakis, "A Year in Lockdown: How the Waves of COVID-19 Impact Internet Traffic," vol. 64, no. 7, pp. 101-108, July 2021.
- [10] J. Whalley, V. Stocker and W. Lehr, "COVID-19 and the Internet: Lessons Learned," in *Beyond the Pandemic? Exploring the Impact of COVID-19 on Telecommunications and the Internet*, Emerald Publishing Limited, Bingley, 2023, pp. 17-69.
- [11] A. Feldmann, O. Gasser, F. Lichtblau, E. Pujol, I. Poesse, C. Dietzel, D. Wagner, M. Wichtlhuber, J. Tapiador, N. V.-. Rodriguez, O. Hohlfeld and G. Smaragdakis, "Implications of the COVID-19 Pandemic on the Internet Traffic," vol. 2, pp. 1-5, 2021.
- [12] A. Conway, "The definitive guide to 5G: Everything you need to know," 30 12 2022. [Online]. Available: <https://www.xda-developers.com/5g/>. [Accessed 12 09 2023].
- [13] H. Viswanathan and P. E. Mogensen, "Communications in the 6G Era," *IEEE Access*, vol. 8, pp. 57063-57074, 2020.
- [14] "Samsung's 6G White Paper Lays Out the Company's Vision for the Next Generation of Communications Technology," Samsung, 14 07 2020. [Online]. Available: <https://news.samsung.com/global/samsungs-6g-white-paper-lays-out-the-companys-vision-for-the-next-generation-of-communications-technology>. [Accessed 01 09 2024].
- [15] S. Dang, R. D. Marco, M. Wen, M. Chafii, Y. Ko, B. F. Uchôa-Filho and A. Younis, "Editorial: Index Modulation for 6G Communications," *Frontiers in Communications and Networks*, vol. 2, no. 794311, 2021.
- [16] J. Li, S. Dang, M. Wen, Q. Li, Y. Chen, Y. Huang and W. Shang, "Index Modulation Multiple Access for 6G Communications: Principles, Applications, and Challenges," *IEEE Network*, vol. 37, no. 1, pp. 52-60, February 2023.

- [17] S. Elmeadawy and R. . M. Shubair, "6G Wireless Communications: Future Technologies and Research Challenges," in *International Conference on Electrical and Computing Technologies and Applications (ICECTA)*, Ras Al Khaimah, United Arab Emirates, 2019.
- [18] S. Alraih, I. Shayea, M. Behjati, R. Nordin, N. F. Abdullah, A. Abu-Samah and D. Nandi, "Revolution or Evolution? Technical Requirements and Considerations towards 6G Mobile Communications," *Sensors*, vol. 22, no. 3, p. 762, 2022.
- [19] S. Mumtaz, C. Jiang, A. Tölli, A. Al-Dulaimi, M. M. Butt, H. M. Asif and M. I. Ashraf, "Guest Editorial: 6G: The Paradigm for Future Wireless Communications," *IEEE Wireless*, vol. 29, no. 1, pp. 14-15, 2022.
- [20] Y. Corre, G. Gougeon, J.-B. Doré, S. Bicaïs, B. Miscopein, E. Faussurier, M. Saad, J. Palicot and F. Bader, "Sub-THz Spectrum as Enabler for 6G Wireless communications up to 1 Tbits/s," in *6G Wireless Summit*, Levi Lapland, Finland, 2019.
- [21] E. Basar, M. . D. Renzo, J. D. Rosny, M. Debbah, M.-S. Alouini and R. Zhang, "Wireless Communications Through Reconfigurable Intelligent Surfaces," *IEEE Access*, vol. 7, pp. 116753-116773, 2019.
- [22] K. Asmoro and S. Y. Shin, "RIS Grouping Based Index Modulation for 6G Telecommunications," *IEEE Wireless Communications Letters*, vol. 11, no. 11, pp. 2410-2414, 2022.
- [23] E. Basar, "Reconfigurable Intelligent Surface-Based Index Modulation: A New Beyond MIMO Paradigm for 6G," *IEEE Transactions on Communications*, vol. 68, no. 5, pp. 3187-3196, 2020.
- [24] X. Yuan, Y.-J. A. Zhang, Y. Shi, W. Yan and H. Liu, "Reconfigurable-Intelligent-Surface Empowered Wireless Communications: Challenges and Opportunities," *IEEE Wireless Communications*, vol. 28, no. 2, pp. 136-143, April 2021.
- [25] Y. Liu, X. Liu, X. Mu, T. Hou, J. Xu, M. D. Renzo and N. Al-Dhahir, "Reconfigurable Intelligent Surfaces: Principles and Opportunities," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1546-1577, 2021.
- [26] C. Pan, H. Ren, K. Wang, J. F. Kolb, M. El Kashlan, M. Chen, M. D. Renzo, Y. Hao, J. Wang, A. L. Swindlehurst, X. You and L. Hanzo, "Reconfigurable Intelligent Surfaces for 6G Systems: Principles, Applications, and Research Directions," *IEEE Communications Magazine*, vol. 59, no. 6, pp. 14-20, June 2021.
- [27] V. Velez, J. P. Pavia, N. Souto, P. Sebastião and A. Correia, "A Generalized Space-Frequency Index Modulation Scheme for Downlink MIMO Transmissions With Improved Diversity," *IEEE Access*, vol. 9, pp. 118996-119009, 2021.
- [28] J. P. Pavia, V. Velez, B. Brogueira, N. Souto and A. Correia, "Precoded Generalized Spatial Modulation for Downlink MIMO Transmissions in Beyond 5G Networks," *MDPI Applied Sciences*, vol. 10, no. 18, p. 6617, 2020.
- [29] M. Wen, "A Survey on Spatial Modulation in Emerging Wireless Systems: Research Progresses and Applications," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 9, pp. 1949-1972, 2019.
- [30] X. Cheng, M. Zhang, M. Wen and L. Yang, "Index Modulation for 5G: Striving to Do More with Less," *IEEE Wireless Communications*, vol. 25, no. 2, pp. 126-132, April 2018.
- [31] H. A. Lopes and N. Souto, "Iterative signal detection for large scale GSM-MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 8, p. 7734–7738, August 2018.
- [32] B. Shim, S. Kwon and B. Song, "Sparse Detection With Integer Constraint Using Multipath Matching Pursuit," *IEEE Communications Letters*, vol. 18, no. 10, p. 1851–1854, 2014.
- [33] Q. Spencer, A. Swindlehurst and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, p. 461–471, February 2004.

- [34] V. Velez, J. P. Pavia, C. Rita, C. Gonçalves, N. Souto, P. Sebastião and A. Correia, "System-Level Assessment of a C-RAN based on Generalized Space–Frequency Index Modulation for 5G New Radio and Beyond," *Applied Sciences*, vol. 12, no. 3, pp. 1592-1617, February 2022.
- [35] R. T. Rodoshi, K. Taewoon and W. Choi, "Resource Management in Cloud Radio Access Network: Conventional and New Approaches," *Sensors*, vol. 9, no. 20, p. 2708, 2020.
- [36] "Cloud/Centralized Radio Access Network (C-RAN)," Artiza Networks, [Online]. Available: <https://www.artizanetworks.com/resources/tutorials/cran.html>. [Accessed 3 09 2023].
- [37] J. P. Pavia, V. Velez, R. Ferreira, N. Souto, M. Ribeiro, J. Silva and R. Dinis, "Low Complexity Hybrid Precoding Designs for Multiuser mmWave/THz Ultra Massive MIMO Systems," *Sensors MDPI*, vol. 21, no. 18, p. 6054, 2021.
- [38] J. P. Pavia, V. Velez, N. Souto, M. Ribeiro, P. Sebastião and A. Correia, "System-Level Assessment of Low Complexity Hybrid Precoding Designs for Massive MIMO Downlink Transmissions in 3 Beyond 5G Networks," *Applied sciences*, vol. 12, no. 6, p. 2812, 2022.
- [39] A. Shojaeifard, "ISG RIS Activity Report 2022," ETSI, [Online]. Available: <https://www.etsi.org/committee-activity/activity-report-ris>. [Accessed 28 09 2023].
- [40] E. Basar, "Reconfigurable Intelligent Surface-Based Index Modulation: A New Beyond MIMO Paradigm for 6G," *IEEE Transactions on Communications*, vol. 68, no. 5, pp. 3187-3196, 2020.
- [41] V. Velez, J. P. Pavia, N. Souto, P. Sebastião and A. Correia, "Performance Assessment of a RIS-empowered C-RAN operating at the mmWave/THz bands," *IEEE Access*, vol. 11, pp. 49625-49638, 2023.

Precoded Generalized Spatial Modulation for Downlink MIMO Transmissions in Beyond 5G Networks

João Pedro Pavia ^{1,2,*}, Vasco Velez ^{1,2}, Bernardo Brogueira ^{1,2}, Nuno Souto ^{1,2}
and Américo Correia ^{1,2}

¹ Department of Information Science and Technology, ISCTE—Instituto Universitário de Lisboa, Av. das Forças Armadas, 1649-026 Lisbon, Portugal; Bernardo_Brogueira@iscte-iul.pt (B.B.); Vasco_Velez@iscte-iul.pt (V.V.); Nuno.Souto@iscte-iul.pt (N.S.); américo.correia@iscte.pt (A.C.)

² Radio Systems Group, Instituto de Telecomunicações, Av. Rovisco Pais, 1049-001 Lisbon, Portugal

* Correspondence: Joao_Pedro_Pavia@iscte-iul.pt

Received: 20 August 2020; Accepted: 16 September 2020; Published: 22 September 2020

Abstract: The design of multiple input multiple output (MIMO) schemes capable of achieving both high spectral and energy efficiency constitutes a challenge for next-generation wireless networks. MIMO schemes based on generalized spatial modulations (GSM) have been widely considered as a powerful technique to achieve that purpose. In this paper, a multi-user (MU) GSM MIMO system is proposed, which relies on the transmission of precoded symbols from a base station to multiple receivers. The precoder's design is focused on the removal of the interference between users and allows the application of single-user GSM detection at the receivers, which is accomplished using a low-complexity iterative algorithm. Link level and system level simulations of a cloud radio access network (C-RAN) comprising several radio remote units (RRUs) were run in order to evaluate the performance of the proposed solution. Simulation results show that the proposed GSM MU-MIMO approach can exploit efficiently a large number of antennas deployed at the transmitter. Moreover, it can also provide large gains when compared to conventional MU-MIMO schemes with identical spectral efficiencies. In fact, regarding the simulated C-RAN scenario with perfect channel estimation, system level results showed potential gains of up to 155% and 139% in throughput and coverage, respectively, compared to traditional cellular networks. The introduction of imperfect channel estimation reduces the throughput gain to 125%.

Keywords: B5G; generalized spatial modulation (GSM); precoder design; massive multiple input multiple output (MIMO); quadrature amplitude modulation (QAM) constellations

1. Introduction

Considering the technological advances over the last decades, the next generation of wireless communications is expected to follow this trend with a significant increase in system robustness (SR), spectral efficiency (SE) and energy efficiency (EE). In recent years, new emerging techniques have appeared in order to meet the increasingly challenging requirements of beyond fifth generation (B5G) communication systems, such as non-orthogonal multiple access (NOMA) schemes like Signaling Aided Sparse Code Multiple Access [1] or index modulations (IM). IM has received significant attention due to its ability to activate a subset of certain elements of communication resources, namely antennas, subcarriers, and slots [2,3]. Generalized spatial modulation (GSM) constitutes a particular case of IM, which is suitable for large scale multiple input multiple output (MIMO) antennas schemes enabling

greater EE with ease of implementation [4,5]. Considering that the information is encoded in the combination of active antennas and, also in the modulated symbols transmitted in the active antennas, GSM can also achieve a greater SE than single antenna communications.

GSM can be considered as a compromise between conventional MIMO and simple radio frequency (RF) transmissions, since only a subset of the available transmission antennas is active for a certain period of time, thus reducing the number of RF chains required. Several detectors have been reported in the literature for single user scenarios. The authors of [6] proposed a minimal average square block error detector (OB-MMSE) that can achieve a close to optimal performance, while its required complexity is much lower when compared to other detectors. This detector uses an algorithm that sorts the possible transmit antenna combinations (TAC), followed by the detection in sequence of the possible signal vector for each TAC using block minimum mean square error (MMSE). A termination threshold must be applied in order to reduce the number of tested TACs. Although this detector is able to achieve near-optimal performance, it can incur in substantial complexity in large scenarios. A different GSM iterative detector is proposed in [7], which is based on dividing the problem of the maximum likelihood detection (MLD) into a sequence of simpler steps, such as the minimization of the unrestricted Euclidean distance, the projection of the elements onto the signal constellation and the projection onto the set of valid active antenna combinations. This approach allows a substantial complexity reduction when compared with the optimal MLD while still achieving near-optimal performance.

Although the wide range of precoding schemes referred to in the literature for MIMO systems considering both uplink and downlink scenarios [8–12], there is a significant imbalance between the number of approaches aimed at those scenarios for GSM-based schemes. In fact, there are very few studies that have extended the use of GSM to downlink MU [13].

Despite describing a system for scalable video broadcast communications in [3], the proposed scheme also considered the use of GSM for multiple users. However, the removal of inter-user interference is made at the receiver, which demands a large number of antennas at the users. A better suited alternative for dedicated links relies on removing inter-user interference at the transmitter through a precoder. This approach is often applied for conventional MU-MIMO as presented in [14], where the authors describe a precoder that accomplishes block diagonalization (BD) of the equivalent channel matrix. The proposed BD precoding guarantees zero inter-user interference and can be thought of as a generalization of channel inversion. Despite the similarity of the approach proposed by the authors in [15], their method cannot only provide improved bit error rate (BER) and throughput performances, but also additional diversity gain by adopting a partial nulling technique for the generalized block diagonalization (GBD). A few precoded schemes have been introduced for spatial modulations (SM) and GSM since then. A new precoder scheme for the downlink of MU-SM systems was proposed in [16], which exploits the channel status information (CSI) at the base station (BS). Here, a precoding matrix is computed, which allows the MU downlink system to be broken down into several independent single user SM systems. A precoded scheme designed for multi-user (MU) GSM systems was reported in [17], with the aim of eliminating all inter-user interference while maintaining the antenna selection features of GSM, which means that only some of the antennas are active, while the rest are silenced. Both proposals of [16,17] are limited in terms of spectral efficiency, since the first one was only defined for SM, while the later was designed specifically for a version of GSM, where the M -quadrature amplitude modulations (M-QAM) symbols are the same in all active antennas.

CSI is fundamental in channel estimation process in order to enable uplink and downlink transmissions in MIMO systems. However, the channel estimation for downlink transmissions on massive MIMO systems operating at frequency division duplexing (FDD) represents a very complex problem, since it is unfeasible for practical applications [18,19]. Time Division Duplexing (TDD) represents an interesting solution that can be used as alternative in order to overcome the aforementioned problem in context of downlink transmission in FDD systems. Considering the use of TDD mode, it is possible to exploit the channel reciprocity, which allows the estimation of the downlink channel by the base station through the uplink channel information. In the uplink scenario, orthogonal pilot signals are sent from the users to the base station, and based on that, signals at the base station will estimate the CSI to the user equipment (UE). After accomplishing this task, the base station beamforms the downlink data towards the UE. Considering that there is a limited number of orthogonal pilots that can be reused between cells, a pilot contamination issue may appear and become a critical problem for

massive MIMO channel estimation [20]. In order to overcome this issue and others, such as the increasing amount of required hardware and computational complexity cost due to use of large number of required antennas in those schemes, several channel estimation algorithms have been developed over last few years [21–23]. The success of the channel estimation process affects the performance of massive MIMO schemes [24] and, such as, should also be taken into account in the system evaluation. It is important to highlight that even though we will not cover in this paper, massive MIMO systems such as the GSM schemes addressed here are prone to several hardware impairments such as non-linear distortions from power amplifier, I/Q imbalance, sampling jitter, and finite-resolution quantization in analog digital converters (ADCs) [20]. To reduce the impact of these effects on the overall performance of the system, compensation algorithms can be developed to mitigate the impairments.

Another issue that must be considered with the introduction of 5G and beyond is the extreme densification of the network, which requires an increase in the network capacity [25,26]. Poor cell-edge coverage and throughput are the most limiting factors of 4G cellular radio access network (RAN). Some research has been dedicated to decrease inter-cell interference by base station coordination and coordinated beamforming [27,28]. Coordinated multi-point (CoMP) transmission or reception is one of the key techniques in 5G that mitigates inter-cell interference (ICI) from neighboring cells, providing higher spectral efficiency and coverage. CoMP indeed extends the cell coverage area and improves cell edge throughput. Joint processing coordinated multipoint transmission (JP-CoMP) requires the clustering of neighboring cells and cooperative transmission within each cluster. Clustering algorithms can be static and dynamic, centralized or distributed [29,30]. Static clustering relies on a predetermined fixed base station cluster. Each static clustering algorithm utilizes different strategies to determine the efficient cluster formation. The network then decides on base station clusters. Dynamic clustering adapts to network changes, where the usual methods are designed based on centralized control on the network, which requires extensive information sharing. In our study, we only consider static clustering based on channel state information (CSI). The techniques mentioned above are essential to improve the overall spectral and energy efficiencies and also increase the throughput and coverage gains, when compared to traditional cellular networks [31].

Motivated by the work above, in this paper we provide a study on MU-MIMO systems, where GSM symbols are transmitted simultaneously to multiple users (differences between the proposed approach and a conventional MU-MIMO assumed as reference are shown in Table 1). To increase the SE of the transmission, different modulated symbols are sent on different (virtual) antennas, where high-order M-QAM constellation with sizes reaching $M = 1024$ symbols are considered. To remove inter-user interference and transform the MU transmission into several independent SU links, a BD precoder is applied at the BS, while a modified and improved version of the low-complexity SU GSM detector presented in [7] is used at the receiver.

The influence of imperfect channel estimation on the performance of this massive MIMO GSM-based system is also analyzed. Link level simulations show that the presented GSM MU-MIMO approach can provide substantial performance gains over conventional MU-MIMO. Additionally, system level simulations show that deployments based on cloud-RAN (C-RAN) comprising several radio remote units (RRUs) can achieve large throughput and coverage gains over traditional cellular

Table 1. Comparison between the reference conventional multi-user multiple input multiple output (MU-MIMO) and the proposed generalized spatial modulations multi-user multiple input multiple output (GSM MU-MIMO).

	Conventional MU-MIMO	GSM MU-MIMO
Precoder	<ul style="list-style-type: none"> Block Diagonalization 	<ul style="list-style-type: none"> Block Diagonalization
Detector	<ul style="list-style-type: none"> conventional single user MIMO proposed ADMM-based receiver 	<ul style="list-style-type: none"> Single-user GSM detector capable of operating in undetermined scenarios (sMMP, OB-MMSE, ...) proposed ADMM-based receiver
Information Mapping	<ul style="list-style-type: none"> Modulated Symbols 	<ul style="list-style-type: none"> Modulated Symbols (Virtual) Antenna Indices
Possibilities for improving Spectral Efficiency	<ul style="list-style-type: none"> Increase modulation order Increase number of transmit antennas → additional transmit power needed 	<ul style="list-style-type: none"> Increase modulation order Increase number of transmit antennas → no additional transmit power needed

networks. The paper is organized as follows: Section 2 presents the model for the MU GSM system, Section 3 presents the transmitter and receiver structure followed by the numerical results obtained in Section 4. Finally, the conclusions are outlined in Section 5.

Notation: Matrices and vectors are denoted by uppercase and lowercase boldface letters, respectively, $(\cdot)^T$ and $(\cdot)^H$ denote the transpose and conjugate transpose, $\lfloor \cdot \rfloor$ is the floor function, $\binom{N}{k}$ denotes the number of combinations of N symbols taken k at a time, and $\text{supp}(\mathbf{x})$ returns the set of indices of nonzero elements in \mathbf{x} (i.e., the support of \mathbf{x}).

2. System Model

Let us consider a downlink MU-MIMO system where a BS transmits simultaneously to N_u users. The BS is equipped with N_{tx} antennas and each user has N_{rx} antennas, as illustrated in Figure 1. We assume that the signal can be represented as $\mathbf{s} = [s_0^T, \dots, s_{N_u-1}^T]^T$, where $\mathbf{s}_k \in \mathbb{C}^{N_s \times 1}$ contains the information transmitted to user k and $N_s \leq N_{tx}/N_u$.

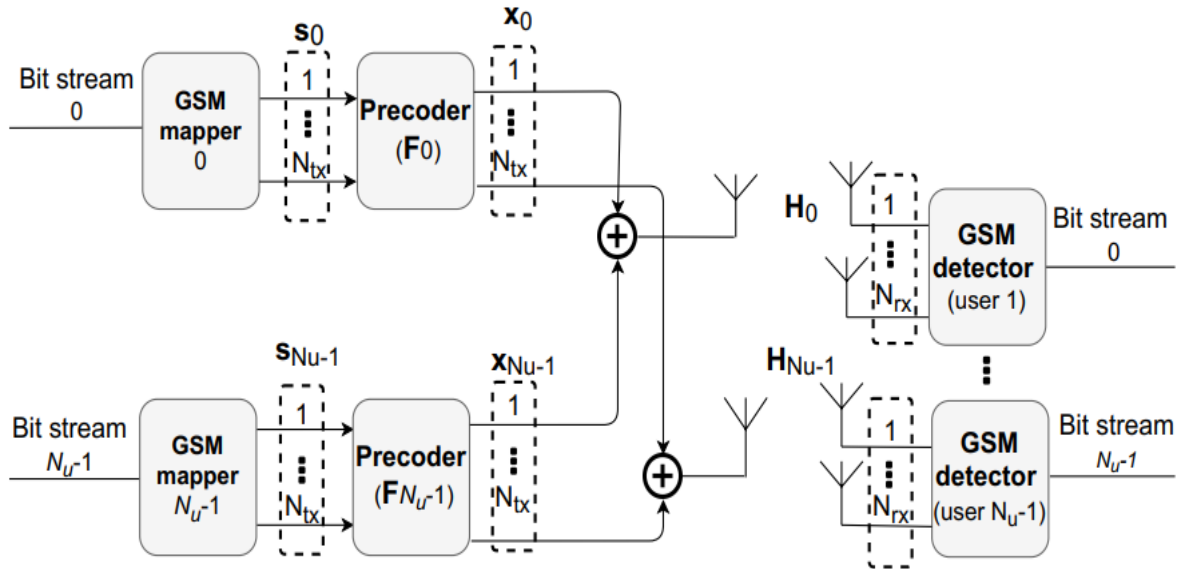


Figure 1. Transmitter and receiver scheme.

Considering that GSM is being used, only N_a positions of \mathbf{s}_k are nonzero. These correspond to active indexes, which carry M-QAM modulated symbols. The signal vector of each user can be written as

$$\mathbf{s}_k = [\dots, 0, s_k^0, 0, \dots, 0, s_k^{N_a-1}, 0, \dots]^T \quad (1)$$

where $s_k^j \in \mathcal{A}$ ($j = 0, \dots, N_a - 1$) with \mathcal{A} denoting an M-QAM complex valued constellation set.

According to this model, the information will be divided in such a way that part of the data are used to select an active index (AI) from a total $N_{\text{comb}} = 2^{\lfloor \log_2(\frac{N_s}{N_a}) \rfloor}$ AI combinations (AICs) available per user. The remaining data are mapped onto N_a complex-valued M-QAM symbols. The resulting SE is then

$$N_{\text{bits}} = N_u \left\lfloor \log_2 \left(\frac{N_s}{N_a} \right) \right\rfloor + N_u N_a \log_2(M) \quad (2)$$

bits per channel use (bpcu).

3. Transmitter and Receiver Structure

In this section the transmitter and receiver structures' design will be addressed. The receiver design will be based on the alternating direction method of the multipliers (ADMM), which will be explained further in Section 3.2.

3.1. Transmitter Design

Channel state information at the transmitter (CSIT) will be used to pre-process the symbols through a linear precoder $\mathbf{F} = [\mathbf{F}_0, \dots, \mathbf{F}_{N_u-1}]$, where $\mathbf{F}_k \in \mathbb{C}^{N_{tx} \times N_s}$. Considering that the transmitted signal propagates through a flat fading channel, the baseband signal received by user k can be written as

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{H}_k \sum_{\substack{j=0 \\ j \neq k}}^{N_u-1} \mathbf{x}_j + \mathbf{n}_k. \quad (3)$$

where

$$\mathbf{x} = \sum_{k=0}^{N_u-1} \mathbf{x}_k = \sum_{k=0}^{N_u-1} \mathbf{F}_k \mathbf{s}_k. \quad (4)$$

In this expression, $\mathbf{H}_k \in \mathbb{C}^{N_{rx} \times N_{tx}}$ corresponds to the channel matrix for the link between the BS and user k and $\mathbf{n}_k \in \mathbb{C}^{N_{rx} \times 1}$ is the noise vector with samples taken according to a zero-mean circularly symmetric Gaussian distribution with covariance $2\sigma^2 \mathbf{I}_{N_{rx}}$. The first term in (4) is related to the desired signal and the second one is the interference caused by the other users' signals. Moreover, the multiuser interference can be eliminated by using a BD method as proposed in [9]. Following this approach, the equivalent overall channel matrix $\mathbf{H}\mathbf{F}$, with $\mathbf{H} = [\mathbf{H}_0^T, \dots, \mathbf{H}_{N_u-1}^T]^T$, will become block diagonal. A simple BD precoder without any power loading optimization is assumed in this paper, with each precoder matrix \mathbf{F}_k designed so as to enforce that $\mathbf{H}_i \mathbf{F}_k = 0$ for all $i \neq k$. This particular condition can be satisfied using vectors selected from the null space of matrix $\tilde{\mathbf{H}}_k$, which is defined as

$$\tilde{\mathbf{H}}_k = [\mathbf{H}_0^T, \dots, \mathbf{H}_{k-1}^T, \mathbf{H}_{k+1}^T, \dots, \mathbf{H}_{N_u-1}^T]^T. \quad (5)$$

$\tilde{\mathbf{H}}_k$ corresponds to the concatenation of the channel matrices between the BS and all users except user k . An orthonormal basis for the null space of $\tilde{\mathbf{H}}_k$ can be found by computing its singular value decomposition (SVD) as

$$\tilde{\mathbf{H}}_k = \tilde{\mathbf{U}}_k \tilde{\mathbf{\Lambda}}_k [\tilde{\mathbf{V}}_k^{(1)} \tilde{\mathbf{V}}_k^{(0)}]^H, \quad (6)$$

where $\tilde{\mathbf{U}}_k$ is the matrix with the left-singular vectors and $\tilde{\mathbf{\Lambda}}_k$ is a rectangular diagonal matrix containing the nonzero singular values. $\tilde{\mathbf{V}}_k^{(1)}$ and $\tilde{\mathbf{V}}_k^{(0)}$ contain the right singular vectors corresponding to the nonzero singular values and the null singular values, respectively. \mathbf{F}_k is obtained from $\tilde{\mathbf{V}}_k^{(0)}$ by selecting its first N_s columns. In this case, the signal arriving at each receiver reduces to

$$\mathbf{y}_k = \hat{\mathbf{H}}_k \mathbf{s}_k + \mathbf{n}_k, \quad (7)$$

where $\hat{\mathbf{H}}_k = \mathbf{H}_k \mathbf{F}_k$ is the equivalent single user channel seen by the receiver.

3.2. Receiver Design

Considering the system model and the BD precoder described in the previous sections, each receiver will have to apply simple single user GSM detection. This can be seen as an attempt to solve the MLD problem related to receiver k , which is formulated as

$$\min_{\mathbf{s}} f(\mathbf{s}_k) \triangleq \|\mathbf{y}_k - \hat{\mathbf{H}}_k \mathbf{s}_k\|_2^2 \quad (8)$$

$$\text{subject to } \mathbf{s}_k \in \mathcal{A}_0^{N_s} \quad (9)$$

$$\text{supp}(\mathbf{s}_k) \in \mathbb{S}, \quad (10)$$

where $\mathcal{A}_0 \stackrel{\text{def}}{=} \mathcal{A} \cup \{0\}$ and \mathbb{S} denotes the set of valid AICs, which has a size of N_{comb} . Solving this non-convex problem directly would require excessive or even unfeasible computational complexity for moderate to large problem settings. To tackle the problem, we adopt instead a similar approach to the one we applied in [7], which is based on the idea of using ADMM as an heuristic for splitting a complex problem into a sequence of simpler ones (as addressed in [32]). Being an heuristic based approach, there will be no guarantee that the resulting algorithm will converge to the solution of the original MLD problem. While this means that the detector will be suboptimal, it will require a much lower computational cost. Following a similar derivation to the one provided in [7], we can arrive at the iterative detection algorithm shown in Table 2 which can be used in each GSM receiver.

Table 2. Iterative GSM detection algorithm for each user k .

1:	Input: $\mathbf{u}^0, \mathbf{w}^0, \mathbf{x}^0, \mathbf{z}^0, \hat{\mathbf{H}}_k, \mathbf{y}_k, \rho_x, \rho_z, Q$
2:	$f_{\text{best}} = \infty.$
3:	$\Phi \leftarrow (\hat{\mathbf{H}}_k^H \hat{\mathbf{H}}_k + (\rho_x + \rho_z) \mathbf{I}_{N_s})^{-1}.$
4:	for $t = 0, 1, \dots, Q-1$ do
5:	$\mathbf{s}_k^{(t+1)} \leftarrow \Phi (\hat{\mathbf{H}}_k^H \mathbf{y}_k + \rho_x (\mathbf{x}^{(t)} - \mathbf{u}^{(t)}) + \rho_z (\mathbf{z}^{(t)} - \mathbf{w}^{(t)})).$
6:	$\mathbf{x}^{(t+1)} \leftarrow \prod_D (\mathbf{s}_k^{(t+1)}).$
7:	$\mathbf{z}^{(t+1)} \leftarrow \prod_{\mathcal{A}_0^{N_s}} (\mathbf{s}_k^{(t+1)} + \mathbf{w}^{(t)}).$
8:	$I \leftarrow \text{supp}(\mathbf{x}^{(t+1)}).$
9:	$\hat{\mathbf{s}}_I^{\text{candidate}} \leftarrow 0, \hat{\mathbf{s}}_I^{\text{candidate}} \leftarrow \prod_{\mathcal{A}^{N_u}} (\mathbf{s}_I^{(t+1)})$
10:	if $f(\hat{\mathbf{s}}^{\text{candidate}}) < f_{\text{best}}$ then
11:	$\hat{\mathbf{s}}_{k,I} \leftarrow 0, \hat{\mathbf{s}}_{k,I} \leftarrow \hat{\mathbf{s}}_I^{\text{candidate}}.$
12:	$f_{\text{best}} = f(\hat{\mathbf{s}}^{\text{candidate}}).$
13:	end if
14:	$\mathbf{u}^{(t+1)} \leftarrow \mathbf{u}^{(t)} + \mathbf{s}_k^{(t+1)} - \mathbf{x}^{(t+1)}.$
15:	$\mathbf{w}^{(t+1)} \leftarrow \mathbf{w}^{(t)} + \mathbf{s}_k^{(t+1)} - \mathbf{z}^{(t+1)}.$
16:	end for.
17:	Output: $\hat{\mathbf{s}}_k$

In this table, $\mathbf{u}, \mathbf{w} \in \mathbb{C}^{N_{\text{tx}}/N_u \times 1}$ are scaled dual variables and ρ_x and ρ_z are penalty parameters associated to constraints (9) and (10). A careful tuning of these parameters will ensure that the algorithm reaches a good performance during its execution. In the algorithm, Q is the maximum number of iterations, $\prod_D(\cdot)$ denotes the projection onto set $\mathcal{D} = \{\mathbf{s} : \text{supp}(\mathbf{s}) \in \mathbb{S}\}$, and $\prod_{\mathcal{A}_0^{N_s}}(\cdot)$ is the projection over $\mathcal{A}_0^{N_s}$. The projection over set \mathcal{D} can be accomplished by keeping the N_u largest magnitude

elements whose indices also match a valid antenna combination, whereas $\prod_{\mathcal{A}_0^{N_s}}(\cdot)$ can be computed as simple rounding of each component to the closest element in \mathcal{A}_0 .

Although heuristic-based approaches as the one adopted in the proposed GSM detector can reach a solution faster, it may not be the optimal one. Therefore, it is not guaranteed that the algorithm will converge to the optimal solution of the original MLD problem (which is nonconvex). To increase the chances of finding an optimal solution and to improve the performance of the GSM detector, we present several different strategies. The first method is the simplest and consists of restarting the algorithm multiple times by using different initializations [32] for the variables \mathbf{u}^0 , \mathbf{w}^0 , $\mathbf{x}^{u(0)}$, \mathbf{z}^0 required by the algorithm. Another improvement strategy that we propose relies on checking at the end of the algorithm if any of the P neighboring candidates result in an improvement of $f(\hat{\mathbf{s}}^{candidate})$. These P neighbors can be selected amongst those with the closest supports using the algorithm presented in Table 3. A last possible refinement method that we consider consists of re-solving the MLD problem with the support set fixed according to the candidate point $\hat{\mathbf{s}}_k$ generated by the main algorithm.

Table 3. Solution refinement algorithm based on a closest neighbor search for user k .

1:	Input: $\hat{\mathbf{H}}_k, \mathbf{y}_k, \hat{\mathbf{s}}, \mathbf{s}_k^{(Q)}, f_{best}, P$
2:	$\bar{D}_0 = D, I \leftarrow \text{supp}(\hat{\mathbf{s}})$
3:	for $p = 1, \dots, P$ do
4:	$\bar{D}_p = \bar{D}_{p-1} \setminus \{\mathbf{s} : \text{supp}(\mathbf{s}) = I\}.$
5:	$\mathbf{x} \leftarrow \prod_{\bar{D}_p}(\mathbf{s}_k^{(Q)})$
6:	$I \leftarrow \text{supp}(\mathbf{x}).$
7:	$\hat{\mathbf{s}}_I^{candidate} \leftarrow 0, \hat{\mathbf{s}}_I^{candidate} \leftarrow \prod_{\mathcal{A}^{N_s}}(\mathbf{s}_I^{(Q)})$
8:	if $f(\hat{\mathbf{s}}^{candidate}) < f_{best}$ then
9:	$\hat{\mathbf{s}}_{k,I} \leftarrow 0, \hat{\mathbf{s}}_{k,I} \leftarrow \hat{\mathbf{s}}_I^{candidate}.$
10:	$f_{best} = f(\hat{\mathbf{s}}^{candidate}).$
11:	end if
12:	end for.
13:	Output: $\hat{\mathbf{s}}_k$

In this case, the resulting formulation becomes a conventional MIMO detection problem which can also be approximated by a simple projected MMSE estimate, i.e., as

$$\hat{\mathbf{s}}_I^{candidate} = \prod_{\mathcal{A}^{N_s}} \left(\left(\hat{\mathbf{H}}_{k,I}^H \hat{\mathbf{H}}_{k,I} + 2\sigma^2 \mathbf{I}_{N_u} \right)^{-1} \hat{\mathbf{H}}_{k,I} \mathbf{y}_k \right). \quad (11)$$

We refer to this third approach as the MMSE polishing step. In terms of computational complexity, the BD precoding requires the computation of N_u SVDs, which is the step with the heaviest cost resulting in a complexity order of $O(N_u^4 N_{rx}^3 + N_u^2 N_{tx}^2 N_{rx})$. This cost is supported by the BS which typically can have higher computational capabilities. More critical is the required computational complexity at the users. Regarding the receiver, the s-update step (line 5 of Table 2) has the highest cost as it involves an $N_{tx}/N_u \times N_{tx}/N_u$ matrix inversion (although it is only computed in the beginning of the algorithm). Considering a fixed number of iterations, the total complexity order is $O((N_{tx}/N_u)^3)$. For comparison, the complexity order of MLD is $O(N_{comb} M^{N_s})$ (with $N_{comb} = 2^{\lceil \log_2 \binom{N_s}{N_a} \rceil}$), of a linear MMSE is $O((N_{tx}/N_u)^3)$, of OB-MMSE [33] is $O(N_{comb} N_a^3)$ and of multipath matching pursuit with slicing (sMMP)

[34] is $O\left(\frac{N_{tx}}{N_u} N_{rx} (1-T^k) / (1-T) + (N_a^2 N_{rx} + N_a^3) T^k\right)$ (T is the number of child candidates expanded at each iteration). Therefore, the proposed approach has a similar complexity order to the linear MMSE. Note that the complexity of OB-MMSE does not grow exponentially with the signal constellation size, M , like in the case of MLD, but it still depends on $N_{comb} = 2^{\left\lfloor \log_2 \left(\frac{N_s}{N_a} \right) \right\rfloor}$ which can restrict its use when a large number of bits are conveyed on antenna indices.

4. Numerical Results

In this section, we present numerical simulations, both link level and system level. Link performance results, namely, block error rate (BLER), are used as input by the system level simulator. The system is illustrated in Figure 2, where the C-RAN is comprised of 19 radio remote units (RRUs) connected through fiber to a central unit (CU), each RRU with $N = 60$ active pedestrian users. Each RRU consists of three transmission and reception points (TRP), each one equipped with $N_{tx, total} = 256$ antennas while users have N_{rx} antennas (i.e., each RRU corresponds to a BS according to the system model presented in Section 2). The RRUs array configuration corresponds to cylindrical arrays: $16 \times 16 \times 3$, where the separation between antennas of the array is half wavelength [35].

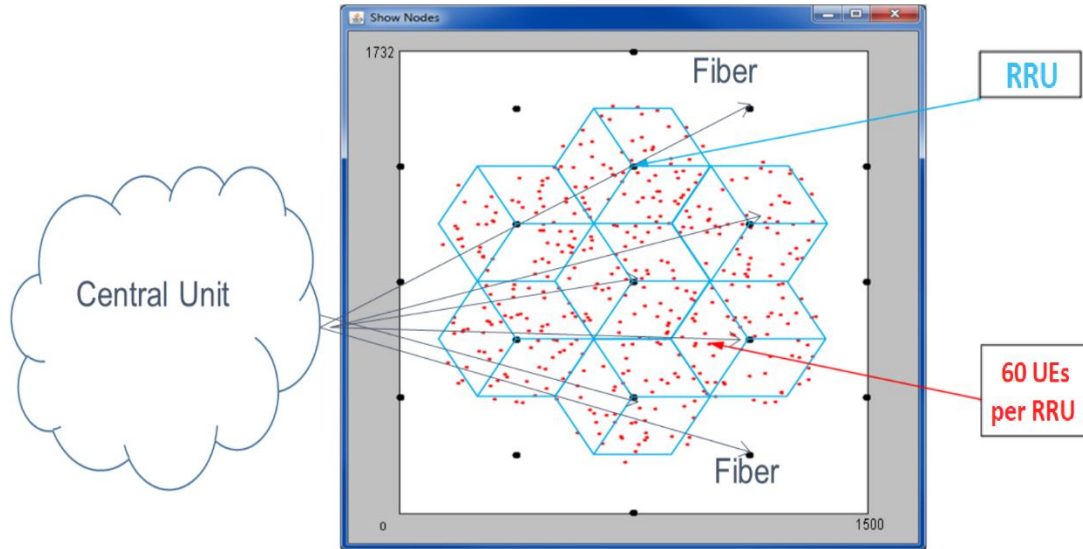


Figure 2. Cloud radio access network (C-RAN) considered for the system level simulations. RRU = radio remote unit.

The system level block diagram can be found in references [36,37]. This simulator is based on the one described in [37]. In the system level simulator, there are general parameters that must be defined, such as network layout and antenna parameters. The setup used considers several of the same parameters adopted in the case study presented in section 7.7 of [38] for the deployment of a Massive MIMO based outdoor network. Our system level simulator considers the 3D urban macro 3D-Uma scenario [36], where the BSs are mounted above rooftop levels of surrounding buildings with antenna height: 25 m and pedestrians height: 1.5–2.5 m. To each pedestrian is assigned line-of-sight (LOS) or non-line-of-sight (NLOS) propagation conditions, depending on the distance to RRU. It is generated correlated large-scale and small-scale parameters to create channel coefficients and pathloss and shadowing are applied with $\sigma_{SF} = 7.8dB$. For the NLOS pathloss distance, we have $PL = 32.4 + 20\log(f_c) + 30\log(d_{3D})$ dB, where d_{3D} is the distance in meters [36]. Other simulator parameters are: carrier frequency $f_c = 3.5GHz$, maximum TRP transmit power 46dBm, receiver spectral noise power density $-174dBm/Hz$, cyclic prefix overhead 5%, pilots/TRP = 15 and arrays with uni-polarized antennas. We choose the 5G NR numerology 1 and slot configuration parameters taken from

[39]: the bandwidth is $B_t = 20\text{MHz}$ with normal CP where the subcarrier spacing is 30 kHz and 28 OFDM symbols are transmitted in every subframe of 1ms. Each user feedbacks all CSI and signal-to-interference-plus-noise ratio (SINR) to TRPs. The static clustering technique partitions the network into three adjacent RRU sets where each user is served by at least one RRU, while the others perform inter-user interference. The RRU inter-site distance is 433 m corresponding to a radius of 250 m.

4.1. Link Level Simulations

The first simulation results had the objective of evaluating the behavior of the iterative GSM receiver and of the overall proposed GSM MU-MIMO transmitter/receiver scheme. Figures 3 and 4 present the results of BER performance versus the signal-to-noise ratio (SNR) in dB of the proposed GSM MU-MIMO system with $N_{tx} = 255$, $N_{rx} = 10$, $N_u = 15$, $N_s = 17$ and $N_a = 2$, which corresponds to a spectral efficiency of 23 bpcu/user for 256-QAM and 27bpcu/user for 1024-QAM.

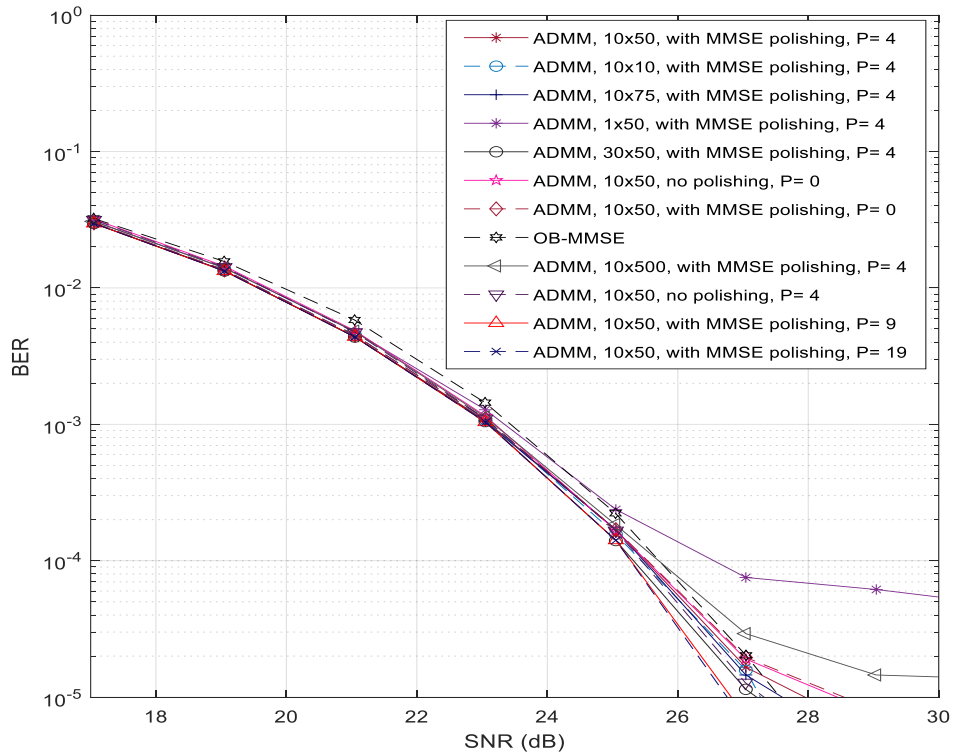


Figure 3. Bit error rate (BER) performance of ADMM in a MU-MIMO scenario with $N_{tx} = 255$, $N_{rx} = 10$, $N_u = 15$, $N_s = 17$ and $N_a = 2$, 256-QAM.

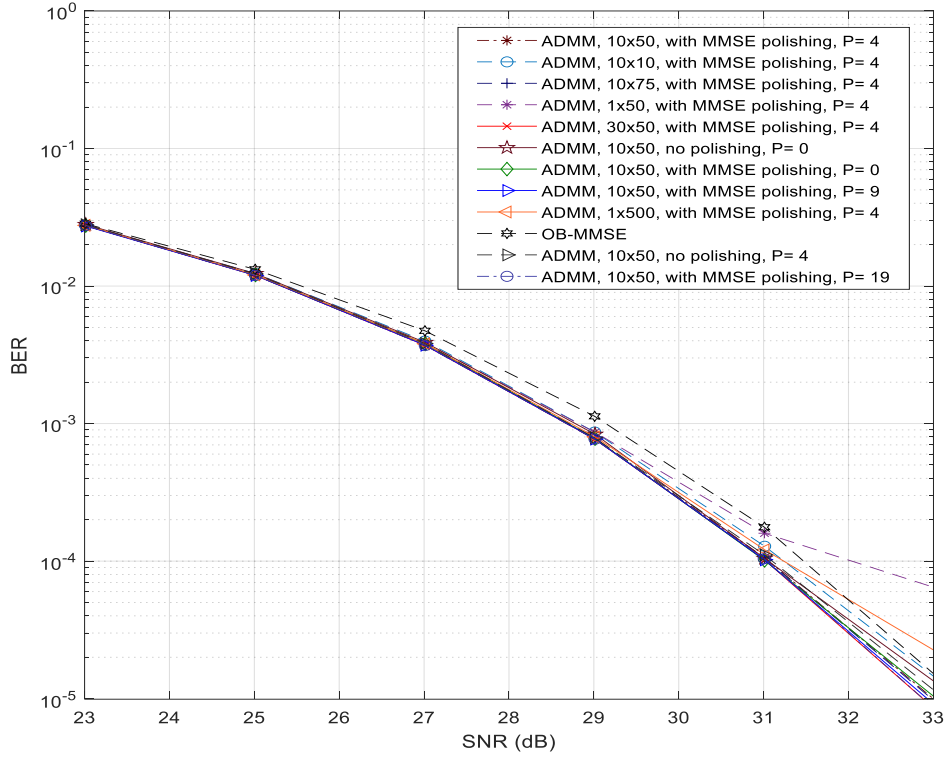


Figure 4. BER performance of ADMM in a MU-MIMO scenario with $N_{tx} = 255$, $N_{rx} = 10$, $N_u = 15$, $N_s = 17$ and $N_a = 2$, 1024-QAM.

The expression $n_1 \times n_2$, mentioned in the legend of both figures, denotes that the receiver algorithm was ran with n_1 restarts and n_2 iterations. The type of polishing applied as well as the number of neighbors is also shown. Besides the expected improvement when using more iterations, it can be observed that by increasing the number of algorithm restarts, we can achieve a better system performance. When considering the 1×500 and 10×50 cases, which have the same total number of iterations, it is clear that the best results are achieved by the case with more restarts (10×50). Considering the scenarios where MMSE polishing is used and those where it is not, one can observe that those where polishing is applied have better performance. We also studied the impact of changing the number of neighbors on the performance of the algorithm and we concluded that the greater the number of neighbors, the better the performance will be (see the cases where $P = 1, 4, 9$ and 19). Moreover, the combination of the three proposed improvement strategies for the ADMM receiver lead to a better performance than the usage of the individual approaches. Globally, the proposed ADMM algorithm tends to lead to better results when compared to the case where the well-known OB-MMSE receiver (which we included as benchmark) is used [6].

Our next goal is to provide a comparison between a conventional BD precoded MU-MIMO [9] and the proposed GSM MU-MIMO. Figure 5 shows the results for two different configurations. The first case concerns a comparison between the precoded GSM MU-MIMO with $N_{tx} = 160$, $N_{rx} = 6$, $N_u = 10$, $N_s = 16$, $N_a = 1$, 16-QAM and the conventional BD precoded MU-MIMO with $N_{tx} = 60$, $N_{rx} = 6$, $N_u = 10$, $N_s = 1$ and 256-QAM, both with a spectral efficiency of 8 bpcu/user. In the second case, we present a comparison between the precoder based on GSM MU-MIMO with $N_{tx} = 90$, $N_{rx} = 8$, $N_u = 10$, $N_s = 9$, $N_a = 3$, quadrature phase-shift keying (QPSK) and the precoder based on conventional MU-MIMO with $N_{tx} = 80$, $N_{rx} = 8$, $N_u = 10$, $N_s = 3$ and 16-QAM, both with a spectral efficiency of 12bpcu/user. Regarding the GSM MU-MIMO scheme, results with the proposed receiver as well as other alternative ones are included, namely with a linear MMSE and with the sMMP (from [34]). In the case of sMMP, a lower number of child nodes ($T = 3$) was adopted for 12 bpcu due to the very high computational complexity when operating with higher values in this scenario.

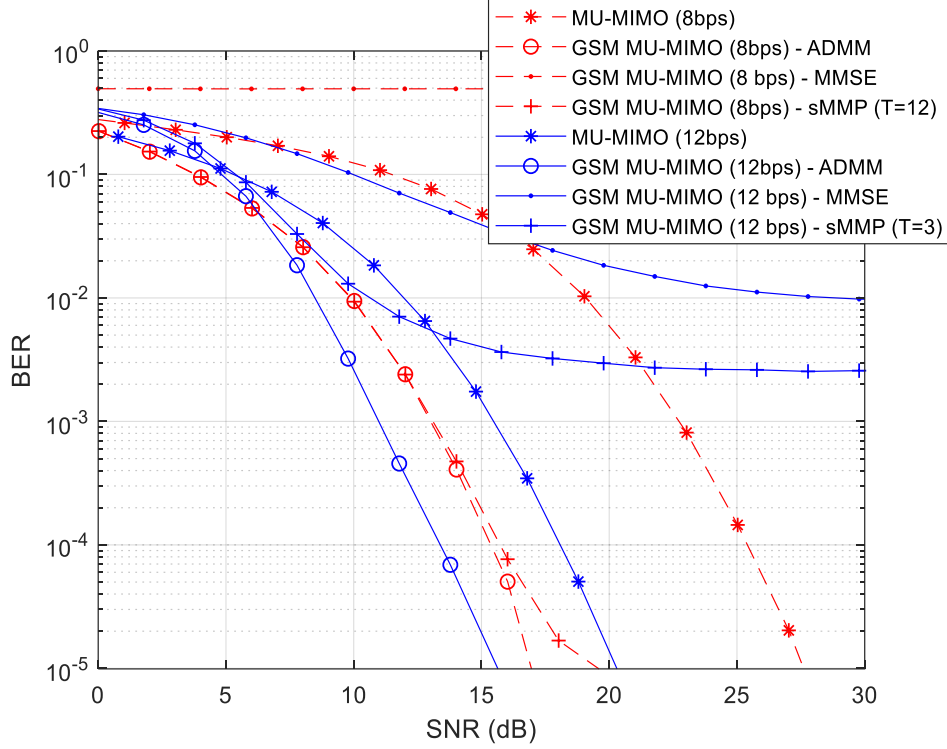


Figure 5. BER performance of a precoder based on GSM MU-MIMO and a precoder based on conventional MU-MIMO.

In the results, it is clear that the proposed ADMM receiver achieves the best results when compared against sMMP and MMSE. In the case of MMSE, it simply cannot correctly detect the information (8 bps) or it has a high irreducible BLER (12 bps). This is due to the fact that from a receiver point of view, both scenarios correspond to underdetermined systems ($N_{rx} < N_s$), which a simple MMSE has a high difficulty to cope with. Through this figure, it can also be seen that the GSM MU-MIMO precoder with the proposed ADMM receiver achieves a better performance when compared to the conventional MU-MIMO precoder (which also uses the same receiver). When we focus on the curve's behavior for a 10^{-4} BER considering a 8bpcu/user scenario, the GSM MU-MIMO shows a gain of about 10 dB over the conventional MU-MIMO. Moving on to the 12bpcu/user scenario and maintaining the BER at 10^{-4} , the GSM MU-MIMO presents a gain of about 5 dB over the conventional MU-MIMO. These results suggest that GSM MU-MIMO can be a potential alternative to increase the SE of the system when compared with the adoption of higher-level modulations in conventional MU-MIMO.

A second set of simulations were performed in order to analyze the block error rate (BLER) performance versus the energy per symbol to noise power spectral density (E_s/N_0) in dB of the proposed GSM MU-MIMO system. These results are required for the system level evaluation in the next subsection. Both perfect channel estimation and imperfect channel estimation curves are presented. For the imperfect channel estimation results we adopted the same model as in [40]. Denoting the CSI available at the transmitter as $\bar{\mathbf{H}}$ and the CSI error as \mathbf{H}_{error} we can express the channel matrix as $\mathbf{H} = \bar{\mathbf{H}} + \mathbf{H}_{error}$. The entries of $\bar{\mathbf{H}}$ and \mathbf{H}_{error} are drawn assuming complex Gaussian distributions of $\mathcal{CN}(0, \sigma_h^2)$ and $\mathcal{CN}(0, \sigma_{h_{error}}^2)$, with $\sigma_h^2 + \sigma_{h_{error}}^2 = 1$. In the presented results we considered $\sigma_{h_{error}} = 0.03$. In our simulations, a minimum of 25,000 blocks were transmitted for computing each BLER result. In Figures 6 and 7, we have $N_{sc} = 256$ subcarriers, $N_{tx} = 17$ antennas/user, $N_{rx} = 16$ antennas/user and $N_u = 15$ users. The number of active antennas are $N_a = 2$ and $N_a = 3$, respectively. The case $N_a = 3$ and 1024-QAM corresponds to a spectral efficiency of 39bpcu/user. The peak bit rate per user achieved assuming 5G NR numerology 1 is 279.552Mbps. This means that 1bpcu/user is equivalent to bit rate of 7.168Mbps.

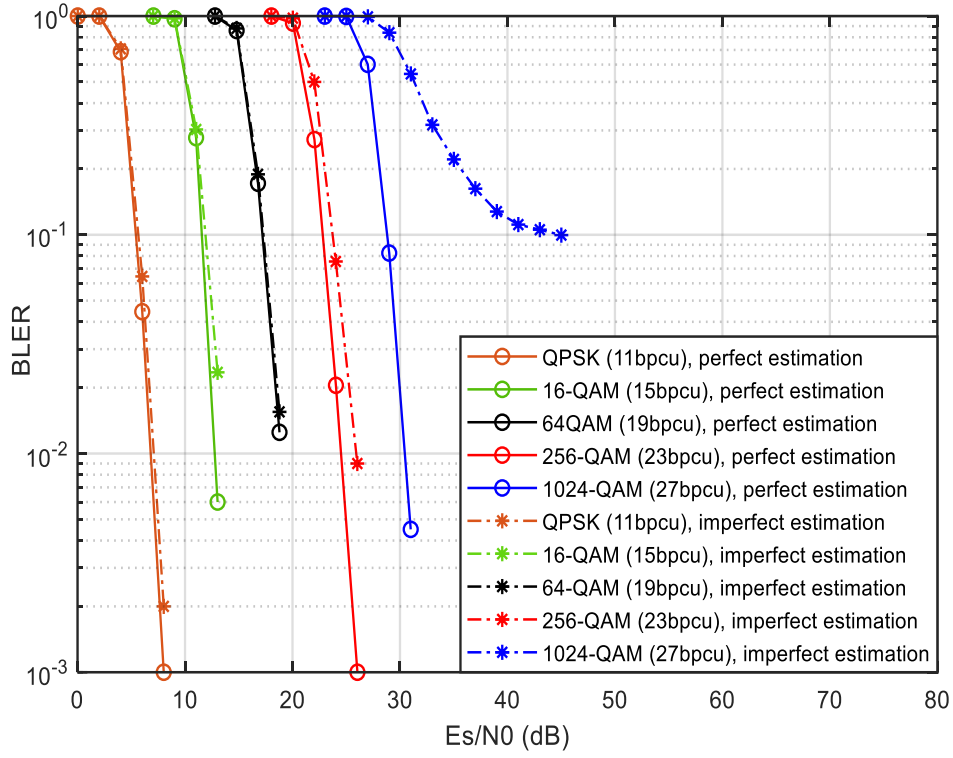


Figure 6. Block error rate (BLER) vs. E_s/N_0 for $N_a = 2$, $N_{sc} = 256$, $N_u = 15$, $N_{tx} = 17$ antennas/user, and $N_{rx} = 16$ antennas/user.

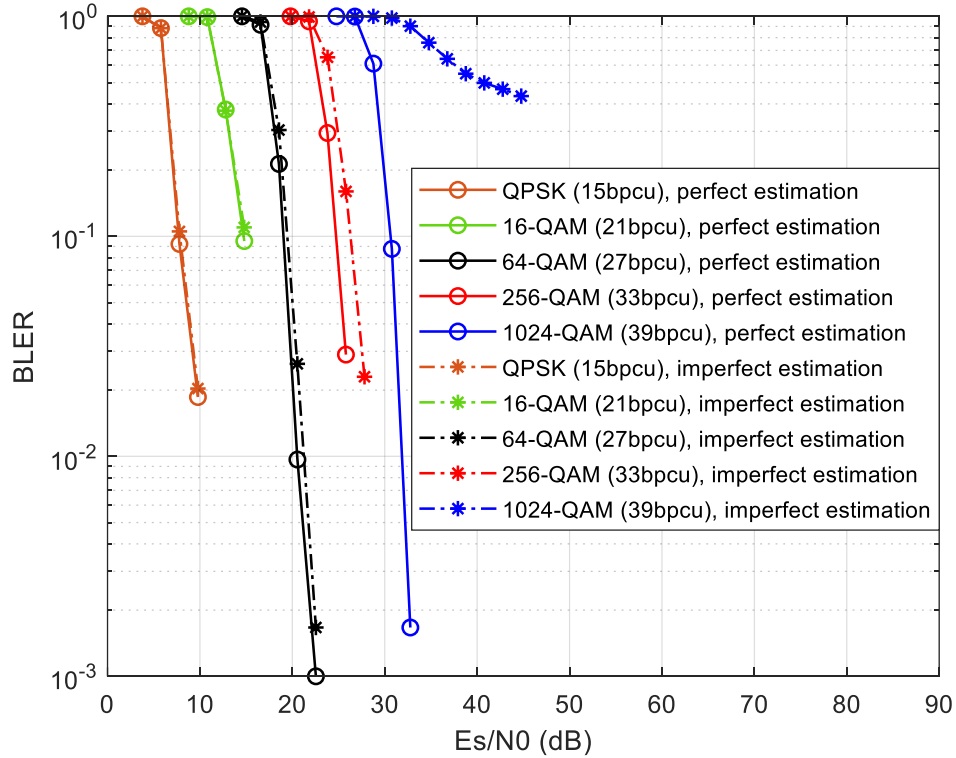


Figure 7. Perfect and imperfect estimation of BLER vs. E_s/N_0 for $N_a = 3$, $N_{sc} = 256$, $N_u = 15$, $N_{tx} = 17$ antennas/user, and $N_{rx} = 16$ antennas/user.

Doubling $N_{sc} = 256$ to $N_{sc} = 512$ doubles the spectral efficiency to 78 bpcu/user use which is equivalent to a bit rate of 14.336Mbps. In both figures, the BLER of GSM MU-MIMO is presented versus

(E_s/N_0) in dB, for five uniform M-QAM constellations namely, $M \in \{4, 16, 64, 256, 1024\}$. As expected, independently of N_a , higher values of M require higher values of E_s/N_0 (dB) to reach the reference $\text{BLER} = 10^{-1}$. In Figure 6, 1024-QAM with perfect estimation requires an additional 24dB of E_s/N_0 compared to 4-QAM(QPSK). With imperfect channel estimation there is an additional 15dB penalty to reach $\text{BLER} = 10^{-1}$ in the detection of 1024-QAM (it has a higher sensitivity to channel estimation errors).

In Figure 7, for $N_a = 3$, it is clear the higher sensitivity of 1024-QAM, as one can notice that with imperfect estimation there is the emergence of a BLER floor. The other modulations only reveal small or negligible degradation.

4.2. System Level Simulations

Using the BLER results described previously, several system level simulations were performed. The signal-to-noise ratio in dB used in the system level simulations is obtained using $\text{SNR} = (E_s/N_0) + 10 \log(R_s/B)$ dB, where R_s is the total transmitted symbol rate per antenna and user, B is the total bandwidth (we considered 20MHz), and E_s/N_0 is the ratio of symbol energy to noise spectral density in dB. Values of E_s/N_0 are obtained from the link level BLER results.

In Figure 8, we have chosen $N_a = 2$ with perfect estimation, and computed the SNR values corresponding to the $\text{BLER} = 10^{-1}$ so as to obtain the coverage results vs. the percentage of transmitted carrier power. Based on the parameters $N_u = 15$, $N_{tx} = 17$ transmit antennas/user and $N_{rx} = 16$ receive antennas/user there are a total of 255 active antennas at each TRP. The coverage of each of the five different M-QAM constellations and the arithmetic average of the coverage of all constellations (labelled as AllQAM) is presented for two different clusterings. In the present cellular topology, RRUs correspond to base stations, and each user is served by one RRU while the other RRUs generate inter-interference when transmitting towards their users. The label 1C means that the cluster contains one RRU. According to BLER performance results of Figure 6, it is expected that the 1024-QAM constellation has the minimum coverage due to more demanding signal-to-noise ratio, while 4-QAM has the maximum coverage for 100% of the transmitted carrier power. Only users close to RRUs are able to decode correctly 1024-QAM symbols, whereas 4-QAM symbols are decoded everywhere. We can check in Figure 8 that only for 100% of transmitted carrier power, the average coverage of all constellations reaches 71.5% of the area. The remaining coverage curves correspond to clustering where the network is partitioned into three adjacent RRU sets and each user is served by three RRUs (labelled as 3C). It is clear that there is an improved coverage obtained for all constellations, which is due to a much lower inter-interference between RRUs. Now, the average coverage of all constellations for 100% of carrier power is 99.6%, which corresponds to a coverage gain of 139%.

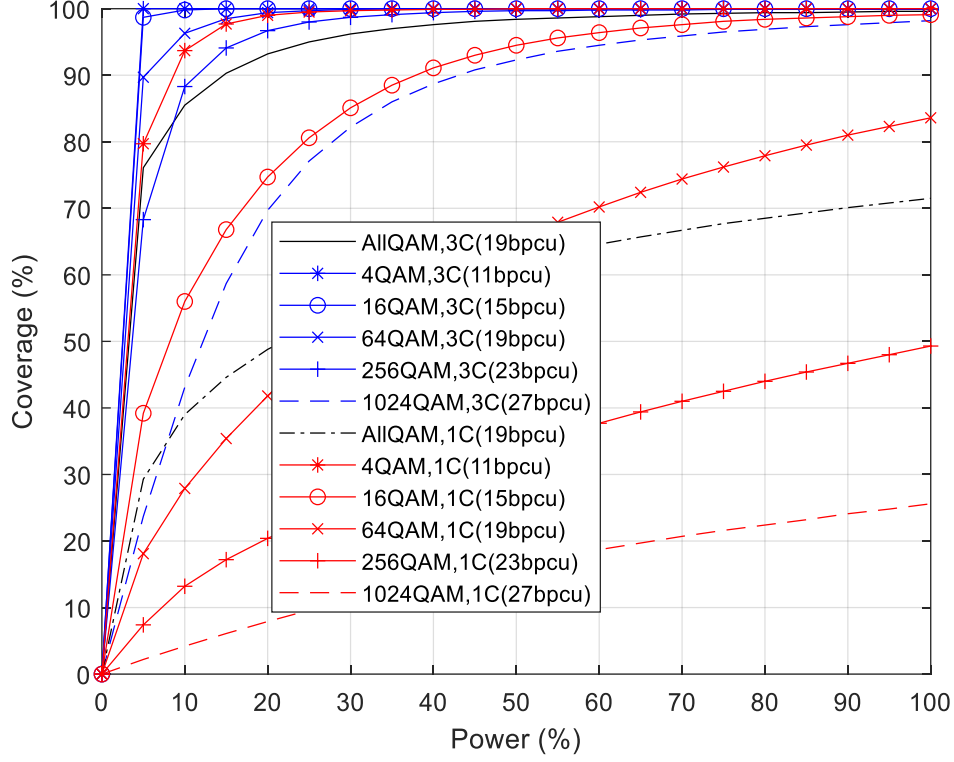


Figure 8. Coverage vs percentage of transmitted power for $N_a = 2$, $N_{sc} = 256$, $N_u = 15$, $N_{tx} = 17$ antennas/user and $N_{rx} = 16$ antennas/user.

In Figure 9, the average throughput results corresponding to the coverage results of the previous figure are presented. Considering the impact of clustering in which only one RRU (1C) is transmitting, the constellation with the lowest throughput is 1024-QAM and the highest result is achieved for 64-QAM. The latter presents a better tradeoff of coverage and spectral efficiency when compared to the use of 16-QAM and 4-QAM constellations. For 15 users and 100% of transmitted carrier power, the average throughput obtained with all constellations is 1306.5 Mbps (87.1Mbps/user). For the clustering where three RRUs (3C) transmit to each user, the constellation with the highest throughput is 1024-QAM, followed by 256-QAM, 64-QAM, 16-QAM, and 4-QAM. There is a match between the spectral efficiency (bpcu) of each constellation and the corresponding average throughput achieved at C-RAN system level. In fact, the same type of match is also observed for the curve of all constellations with spectral efficiency of 19 bpcu and the average throughput achieved at RAN. Notice that the spectral efficiency of 19 bpcu corresponds to a 64-QAM constellation and is equal to the average spectral efficiency of all constellations. For 15 users and 100% of transmitted power, the average throughput of all constellations is 2031.0 Mbps (135.4 Mbps/user), which gives a throughput gain of 155%.

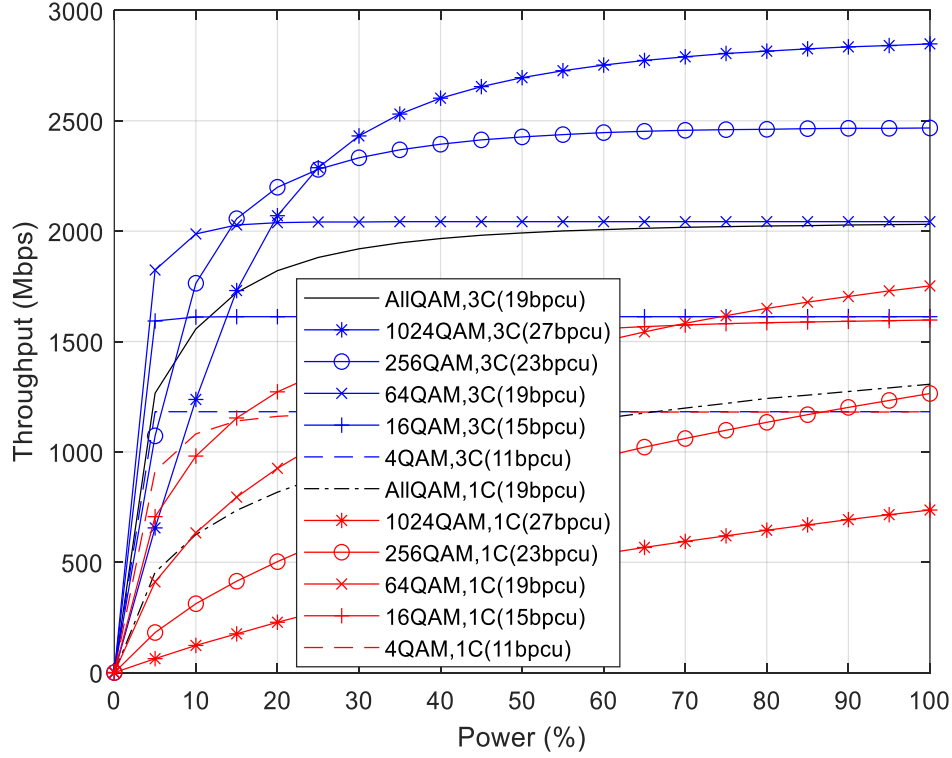


Figure 9. Average throughput vs percentage of transmitted power for $N_a = 2$, $N_{sc} = 256$, $N_u = 15$, $N_{tx} = 17$ antennas/user and $N_{rx} = 16$ antennas/user.

Figure 10 presents the throughput averaged over all users uniformly distributed, for the C-RAN scenario where three RRUs (3C) transmit to each user. The parameters of the previous figures are kept the same, namely $N_{tx} = 17$ antennas/user, $N_{rx} = 16$ antennas/user, $N_{sc} = 256$, and we vary the number of users N_u from 1 up to 15, considering 100% of transmitted carrier power. We consider that the channel estimation is perfect. We can confirm that the BD MU precoding used at the RRUs and the ADMM receivers are operating as expected because every throughput curve is a straight line with slope dependent of the constellation but independent of N_u . The increase in throughput depends on the spectral efficiency. We present two set of results. For $N_a = 2$, the minimum is 11 bpcu (4-QAM) and the maximum is 27 bpcu (1024-QAM). The second set of performance curves have $N_a = 3$, starting from 15 bpcu (4-QAM) up to a maximum of 39 bpcu (1024-QAM). We observe the same throughput results for 15 bpcu with 16-QAM and $N_a = 2$ or 4-QAM with $N_a = 3$. The throughput results are almost the same between the average of all constellations with $N_a = 2$ (19 bpcu) or $N_a = 3$ (27 bpcu) and the 64-QAM constellation having the same spectral efficiencies. This can be explained because the average coverage of all constellations is only slightly lower than the coverage of 64-QAM. The ratio of throughput results for the average of all constellations with $N_a = 3$ (27 bpcu) compared to those of $N_a = 2$ (19 bpcu) is $2875.8/2031.0 = 1.42$, the same as the expected ratio $(27 \text{ bpcu}/19 \text{ bpcu}) = 1.42$.

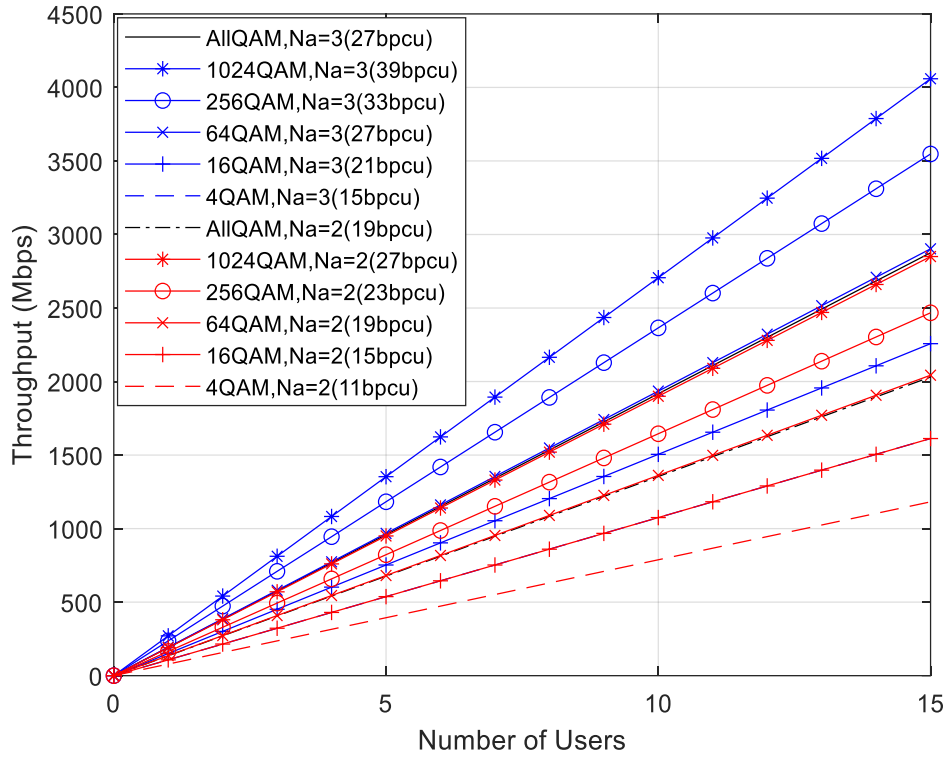


Figure 10. Average throughput vs number of users for $N_{sc} = 256$, $N_{tx} = 17$ antennas/user and $N_{rx} = 16$ antennas/user with perfect channel estimation.

Figure 11 considers the same parameters of Figure 10 but the channel estimation is imperfect instead of perfect. Some performance degradation due to imperfect channel estimation can be observed. The throughput results are not anymore the same between the average of all constellations with $N_a = 2$ (19 bpcu) or $N_a = 3$ (27 bpcu) and the 64-QAM constellation having the same spectral efficiencies. Indeed, the simulation results indicate that the throughput of the average of all constellations is lower than the 16-QAM constellation results with $N_a = 2$ (15 bpcu) or $N_a = 3$ (21 bpcu). For both numbers of active antennas, the throughput results for 1024-QAM become the lowest instead of the highest and for $N_a = 3$, the throughput is zero (does not attain a BLER of 10^{-1} as observed previously). There is an obvious decrease in the simulated throughput results compared to the expected results based on the constellation bpcu. The ratio of throughput results for the average of all constellations with $N_a = 3$ (27 bpcu) compared to those of $N_a = 2$ (19 bpcu) is $2062.5/1546.5 = 1.33$, lower than the expected ratio (27 bpcu/19 bpcu) = 1.42. The comparison between Figures 10 and 11 indicates that the throughput reduction due to imperfect channel estimation for $N_a = 2$ is $(1-1546.5/2031.0) = 0.24$ and for $N_a = 3$ is $(1-2062.5/2875.8) = 0.28$. Therefore, the throughput reduction due to imperfect estimation increases with the number of GSM active antennas which was expected based on the BLER results of Figures 6 and 7.

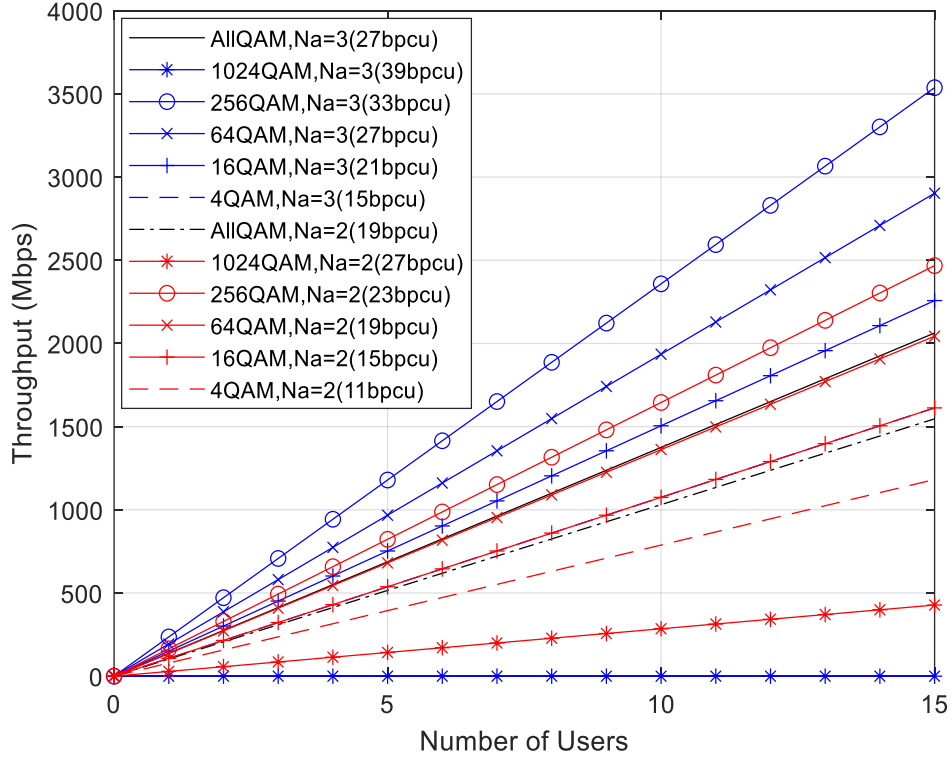


Figure 11. Average throughput vs number of users for $N_{sc} = 256$, $N_{tx} = 17$ antennas/user, and $N_{rx} = 16$ antennas/user with imperfect channel estimation.

Figure 12 presents the cumulative distribution function (CDF) of a RRU with three TRPs, each TRP with $N_{tx} = 255$ active antennas serving 60 users each with $N_{rx} = 16$ antennas. The CDF of this figure corresponds to the case of 100% of carrier transmission power. We consider only the C-RAN scenario with clusters of three RRUs (3C), with curves for both perfect channel estimation and imperfect channel estimation cases. As expected, only for 1024-QAM there is an obvious difference in CDF results due to imperfect estimation compared to perfect estimation. For the other constellations, there are almost the same CDF results which is in agreement with the BLER results of Figure 6. The receiving throughput of all users exceeds 2.5 Gbps, 3.5 Gbps, 4.5 Gbps, 5.5 Gbps, and 6.5 Gbps for 10% of the users with 4-QAM, 16-QAM, 64-QAM, 256-QAM, and 1024-QAM (perfect estimation), respectively. For 50% of the users, the throughput received corresponds to the performance results presented on Figure 9. Only less than 10% of users receive a throughput level lower than 100Mbps, with the exception of 1024-QAM users with imperfect estimation.

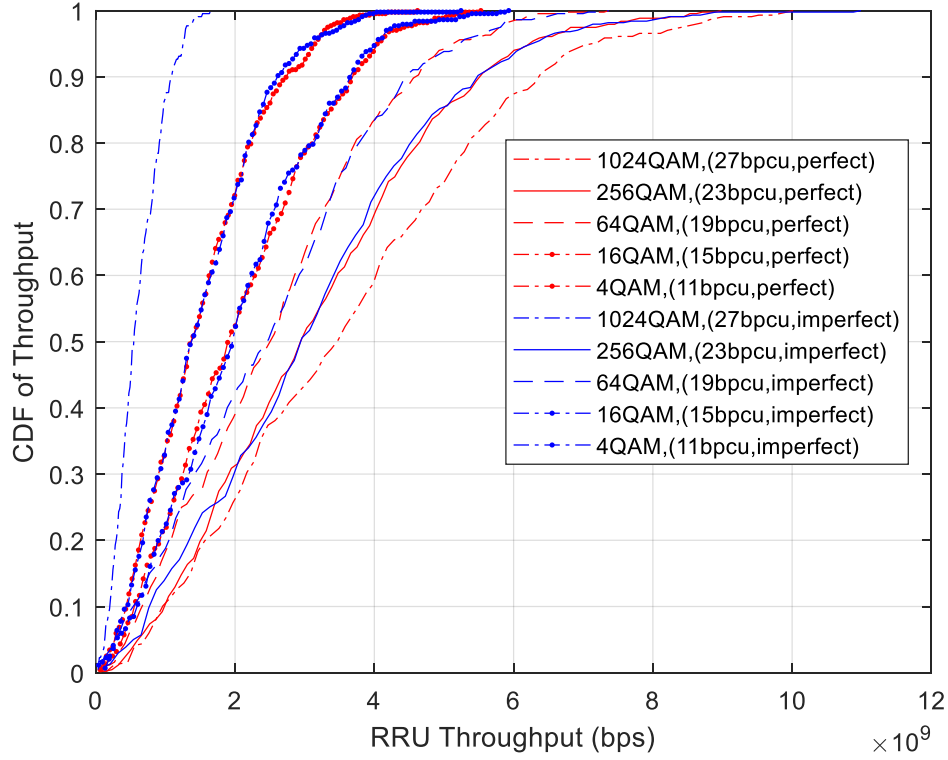


Figure 12. CDF of throughput for 100% transmitted power for $N_a = 2$, $N_{sc} = 256$, $N_u = 15$, $N_{tx} = 17$ antennas/user, and $N_{rx} = 16$ antennas/user.

Table 4 summarizes the average throughput per user with perfect and imperfect channel estimation for C-RAN with clusters 1C and 3C and the corresponding throughput gain. The maximum throughput gain is 1.55 and the minimum is 1.25.

Table 4. Average throughput per user with perfect and imperfect channel estimation and the corresponding throughput gain. CSI = channel status information.

Na	CSI	1C	3C	Gain
2	perfect	87.1Mbps	135.4Mbps	1.55
2	imperfect	82.5Mbps	103.1Mbps	1.25
3	perfect	125.4Mbps	191.7Mbps	1.53
3	imperfect	103.4Mbps	137.5Mbps	1.33

5. Conclusions

In this paper, a novel MIMO system where GSM symbols are transmitted simultaneously to multiple users has been described. By combining large antenna settings at the BS with high-order M-QAM constellations, the proposed approach is capable of improving the spectral efficiency and energy efficiency. A precoder is applied at the BS to completely remove inter-user interference, while a reduced complexity iterative SU GSM detector is implemented at each receiver. Simulation results show that the proposed approach can achieve a very competitive and very promising performance compared to conventional MU-MIMO systems with identical SE. In fact, system level results based on a C-RAN scenario with multiple RRU showed potential gains of up to 155% in throughput and 139% in coverage when compared to traditional cellular networks. The introduction of imperfect channel estimation reduces the throughput gain to 125%. Future work will include a thorough evaluation of the impact of several hardware impairments (such as phase-noise, non-linear distortion, and I/Q imbalances) and robust mitigation algorithms.

Author Contributions: Conceptualization, J.P.P., A.C., B.B. and N.S.; methodology, J.P.P., A.C., B.B. and N.S.; software, J.P.P., A.C., V.V., B.B. and N.S.; validation, J.P.P., A.C., B.B., V.V. and N.S.; formal analysis, J.P.P., A.C., B.B., V.V. and N.S.; investigation, J.P.P., A.C., B.B., V.V. and N.S.; resources, A.C. and N.S.; data curation, J.P.P., B.B., V.V., N.S. and A.C.; writing—original draft preparation, J.P.P., A.C., B.B. and N.S.; writing—review and editing, J.P.P., A.C., B.B., V.V. and N.S.; visualization, J.P.P., A.C., B.B., V.V. and N.S.; supervision, A.C. and N.S.; project administration, A.C. and N.S.; funding acquisition, A.C. and N.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by the ISCTE-University Institute of Lisbon under the grant ISTA-BMPDCTI-2019 and by the FCT/MCTES through national funds and when applicable co-funded by EU funds under the project UIDB/50008/2020.

Acknowledgments: The authors acknowledge the support of ISCTE—University Institute of Lisbon, Fundação para a Ciência e Tecnologia (FCT) and Instituto de Telecomunicações (IT), as described above in funding.

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Yuan, W.; Wu, N.; Zhang, A.; Huang, X.; Li, Y.; Hanzo, L. Iterative Receiver Design for FTN Signaling Aided Sparse Code Multiple Access. *IEEE Trans. Wirel. Commun.* **2020**, *19*, 915–928.
2. Cheng, X.; Zhang, M.; Wen, M.; Yang, L. Index Modulation for 5G: Striving to Do More with Less. *IEEE Trans. Wirel. Commun.* **2018**, *25*, 126–132.
3. Correia, A.; Souto, N.; Sebastiao, P.; Gomez-Barquero, D.; Fuentes, M. Broadcasting Scalable Video with Generalized Spatial Modulation in Cellular Networks. *IEEE Access* **2020**, *8*, 22136–22144.
4. Datta, T.; Eshwaraiah, H.; Chockalingam, A. Generalized Space-and-Frequency Index Modulation. *IEEE Trans. Veh. Technol.* **2016**, *65*, 4911–4924.
5. Murtala, S.; Holoubi, T.; Muchena, N.; Mohaisen, M.; Choi, K. On the Performance of the Multiple Active Antenna Spatial Modulation with 3-Dimensional Constellation. *Appl. Sci.* **2020**, *10*, 3718.
6. Jiang, Y.; Xu, Y.; Xie, Y.; Hong, S.; Wu, X. Low-Complexity Detection Scheme for Generalized Spatial Modulation. *J. Commun.* **2016**, *11*, 726–732.
7. Lopes, H.; Souto, N. Iterative Signal Detection for Large-Scale GSM-MIMO Systems. *IEEE Trans. Veh. Technol.* **2018**, *67*, 7734–7738.
8. Castaneda, E.; Silva, A.; Gameiro, A.; Kountouris, M. An Overview on Resource Allocation Techniques for Multi-User MIMO Systems. *IEEE Commun. Surv. Tutor.* **2017**, *19*, 239–284.
9. Zu, K.; de Lamare, R.; Haardt, M. Generalized Design of Low-Complexity Block Diagonalization Type Precoding Algorithms for Multiuser MIMO Systems. *IEEE Trans. Commun.* **2013**, *61*, 4232–4242.
10. Lopes, P.; Gerald, J. Leakage-based precoding algorithms for multiple streams per terminal MU-MIMO systems. *Digit. Signal Process.* **2018**, *75*, 38–44.
11. Guerreiro, J.; Dinis, R.; Montezuma, P.; Marques da Silva, M. On the Achievable Performance of Nonlinear MIMO Systems. *IEEE Commun. Lett.* **2019**, *23*, 1725–1729.
12. Chataut, R.; Akl, R. Massive MIMO Systems for 5G and beyond Networks—Overview, Recent Trends, Challenges, and Future Research Direction. *Sensors* **2020**, *20*, 2753.
13. Souto, N.; Correia, A. Frequency Domain Equalization for Single and Multiuser Generalized Spatial Modulation Systems in Time Dispersive Channels. *IEEE Wirel. Commun. Lett.* **2020**, *9*, 316–320.
14. Spencer, Q.; Swindlehurst, A.; Haardt, M. Zero-Forcing Methods for Downlink Spatial Multiplexing in Multiuser MIMO Channels. *IEEE Trans. Signal Process.* **2004**, *52*, 461–471.
15. Lee, W.; Ro, J.; You, Y.; Hwang, D.; Song, H. An Enhanced Precoder for Multi User Multiple-Input Multiple-Output Downlink Systems. *Appl. Sci.* **2020**, *10*, 4547.
16. Li, X.; Zhang, Y.; Xiao, L.; Xu, X.; Wang, J. A novel precoding scheme for downlink multi-user spatial modulation system. In Proceedings of the IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), London, UK, 8–11 September 2013; pp. 1361–1365.
17. Zhang, H.; Ma, M.; Shao, Z. Multi-User Linear Precoding for Downlink Generalized Spatial Modulation Systems. *IEEE Commun. Lett.* **2020**, *24*, 212–216.

18. Marques da Silva, M.; Dinis, R.; Guerreiro, J. Energy-Efficient and Spectrally-Efficient MIMO. In *Wiley 5G Ref. Online*, **2019**, 1–18, doi:10.1002/9781119471509.w5GRef017.
19. Marques da Silva, M.; Dinis, R.; Guerreiro, J. Implicit Pilots for an Efficient Channel Estimation in Simplified Massive MIMO Schemes with Precoding. *Int. J. Antennas Propag.* **2019**, *2019*, 1–11.
20. Gustavsson, U.; Sanchez-Perez, C.; Eriksson, T.; Athley, F.; Durisi, G.; Landin, P.; Hausmair, K.; Fager, C.; Svensson, L. On the impact of hardware impairments on massive MIMO. In *Proceedings of the IEEE Globecom Workshops (GC Wkshps)*, Austin, TX, USA, 8–12 December 2014.
21. Souto, N.; Dinis, R. Efficient MIMO Detection for High-Order QAM Constellations in Time Dispersive Channels. In *Proceedings of the IEEE International Conference on Communications (ICC)*, Kansas City, MO, USA, 20–24 May 2018.
22. Marques da Silva, M.; Dinis, R.; Guerreiro, J. A Low Complexity Channel Estimation and Detection for Massive MIMO Using SC-FDE. *Telecom* **2020**, *1*, 3–17.
23. Guerreiro, J.; Dinis, R.; Montezuma, P. Analytical Performance Evaluation of Precoding Techniques for Nonlinear Massive MIMO Systems with Channel Estimation Errors. *IEEE Trans. Commun.* **2018**, *66*, 1440–1451.
24. Mokhtari, Z.; Sabbaghian, M.; Dinis, R. A Survey on Massive MIMO Systems in Presence of Channel and Hardware Impairments. *Sensors* **2019**, *19*, 164.
25. Tullberg, H.; Popovski, P.; Li, Z.; Uusitalo, M.; Høglund, A.; Bulakci, O.; Fallgren, M.; Monserrat, J. The METIS 5G System Concept: Meeting the 5G Requirements. *IEEE Commun. Mag.* **2016**, *54*, 132–139.
26. Cerwall, P. Ericsson Mobility Report. June 2019. Available online: <https://www.ericsson.com/49d1d9/assets/local/mobility-report/documents/2019/ericsson-mobility-report-june-2019.pdf> (accessed on 30 July 2020).
27. 3rd Generation Partnership Project (3GPP). TR 36819 v11.2.0, Coordinated Multi-Point Operation for LTE Physical Layer Aspects, Release 11. September 2013. Available online: http://www.3gpp.org/ftp/Specs/archive/36_series/36.819/36819-b20.zip (accessed on 30 July 2020).
28. Karakayali, M.; Foschini, G.; Valenzuela, R. Advances in smart antennas—Network coordination for spectrally efficient communications in cellular systems. *IEEE Wirel. Commun.* **2006**, *13*, 56–61.
29. Zhang, J.; Chen, R.; Andrews, J.; Ghosh, A.; Heath, R. Networked MIMO with clustered linear precoding. *IEEE Trans. Wirel. Commun.* **2009**, *8*, 1910–1921.
30. Dewa, G.; Park, C.; Sohn, I. Distributed Cell Clustering Based on Multi-Layer Message Passing for Downlink Joint Processing Coordinated Multipoint Transmission. *Appl. Sci.* **2020**, *10*, 5154.
31. Jabbar, S.; Li, Y. Analysis and Evaluation of Performance Gains and Tradeoffs for Massive MIMO Systems. *Appl. Sci.* **2016**, *6*, 268.
32. Boyd, S.; Parikh, N.; Chu, E. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers; In *Foundations and Trends in Machine Learning Series*; Now Publishers Inc.: Delft, The Netherlands, 2010; pp. 1–122.
33. Xiao, Y.; Yang, Z.; Dan, L.; Yang, P.; Yin, L.; Xiang, W. Low-Complexity Signal Detection for Generalized Spatial Modulation. *IEEE Commun. Lett.* **2014**, *18*, 403–406.
34. Shim, B.; Kwon, S.; Song, B. Sparse Detection with Integer Constraint Using Multipath Matching Pursuit. *IEEE Commun. Lett.* **2014**, *18*, 1851–1854.
35. Van Chien, T.; Björnson, E. Massive MIMO Communications. In *5G Mobile Communications*; Xiang, W., Zheng, K., Shen, X., Eds.; Springer: Berlin, Germany, 2017, doi: http://dx.doi.org/10.1007/978-3-319-34208-5_4.
36. 3rd Generation Partnership Project (3GPP). TR 38.901 v. 14.3.0. “5G: Study on Channel Model for Frequencies from 0.5 to 100 GHz”. Release 14. May 2017. Available online: http://www.3gpp.org/ftp/Specs/archive/38_series/38.901/38901-e11.zip (accessed on 9 September 2020).
37. Correia, A.; Silva, M.M. Link and system level simulation for MIMO. In *MIMO Processing for 4G and Beyond: Fundamentals and Evolution*; Silva, M.M., Monteiro, F.A., Eds.; CRC Press: Boca Raton, FL, USA, 2014.
38. Björnson, E.; Hoydis, J.; Sanguinetti, L. Massive MIMO Networks: Spectral, Energy and Hardware Efficiency. 2017. Available online: <https://www.nowpublishers.com/article/Details/SIG-093> (accessed on 9 September 2020).
39. 3rd Generation Partnership Project (3GPP). TS 38211 v15.2.0. 5G/NR Physical Channels and Modulation, Release 15. July 2018. Available online: http://www.3gpp.org/ftp/Specs/archive/38_series/38.221/38211-f20.zip (accessed on 30 July 2020).

40. Stavridis, A.; Sinanovic, S.; Di Renzo, M.; Haas, H. Transmit precoding for receive spatial modulation using imperfect channel knowledge. In Proceedings of the IEEE 75th Vehicular Technology Conference (VTC Spring), Yokohama, Japan, 6–9 May 2012.



© 2020 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).



Low Complexity Hybrid Precoding Designs for Multiuser mmWave/THz Ultra Massive MIMO Systems

João Pedro Pavia ^{1,3,*}, Vasco Velez ^{1,3}, Renato Ferreira ^{1,3}, Nuno Souto ^{1,3}, Marco Ribeiro ^{1,3}, João Silva ^{1,3} and Rui Dinis ^{2,3}

- ¹ Department of Information Science and Technology, ISCTE-Instituto Universitário de Lisboa, 1649-026 Lisboa, Portugal; Vasco_Velez@iscte-iul.pt (V.V.); Renato_Ferreira@iscte-iul.pt (R.F.); Nuno.Souto@iscte-iul.pt (N.S.); Marco.Ribeiro@iscte-iul.pt (M.R.); Joao.Silva@iscte-iul.pt (J.S.)
- ² Instituto de Telecomunicações, 1049-001 Lisboa, Portugal
- ³ Department of Electrical and Computer Engineering, Faculty of Science and Technology, Universidade Nova de Lisboa, 1099-085 Lisboa, Portugal; rdinis@fct.unl.pt
- * Correspondence: jpavia@lx.it.pt

Citation: Pavia, J.P.; Velez, V.; Ferreira, R.; Souto, N.; Ribeiro, M.; Silva, J.; Dinis, R. Low Complexity Hybrid Precoding Designs for Multiuser mmWave/THz Ultra Massive MIMO Systems. *Sensors* **2021**, *21*, 6054. <https://doi.org/10.3390/s21186054>

Academic Editor: Luis Castedo Ribas

Received: 14 July 2021

Accepted: 8 September 2021

Published: 9 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Millimeter-wave and terahertz technologies have been attracting attention from the wireless research community since they can offer large underutilized bandwidths which can enable the support of ultra-high-speed connections in future wireless communication systems. While the high signal attenuation occurring at these frequencies requires the adoption of very large (or the so-called ultra-massive) antenna arrays, in order to accomplish low complexity and low power consumption, hybrid analog/digital designs must be adopted. In this paper we present a hybrid design algorithm suitable for both mmWave and THz multiuser multiple-input multiple-output (MIMO) systems, which comprises separate computation steps for the digital precoder, analog precoder and multiuser interference mitigation. The design can also incorporate different analog architectures such as phase shifters, switches and inverters, antenna selection and so on. Furthermore, it is also applicable for different structures, namely fully-connected structures, arrays of subarrays (AoSA) and dynamic arrays of subarrays (DAoSA), making it suitable for the support of ultra-massive MIMO (UM-MIMO) in severely hardware constrained THz systems. We will show that, by using the proposed approach, it is possible to achieve good trade-offs between spectral efficiency and simplified implementation, even as the number of users and data streams increases.

Keywords: millimeter wave (mmWave); Terahertz (THz); multiuser ultra-massive-MIMO; hybrid precoding and combining; antenna arrays

1. Introduction

Over the last few years, significant advances have been made to provide higher-speed connections to users in wireless networks, with

several novel technologies being proposed to achieve this objective. However, future generations of communication systems will have to fulfil more demanding requirements that cannot be met by the methods adopted in today's communications systems. This motivates the exploration of other candidate technologies, like the millimeter wave (mmWave) and Terahertz (THz) bands, where many applications that require ultra-high data rates can be designed. These bands offer great underutilized bandwidths and also allow for a simplified implementation of large antenna arrays, which are crucial to combat the severe signal attenuation and path losses that occurs at these frequencies [1-4]. The first applications of the THz band were limited to imaging and sensing due to the unavailability of efficient devices that can work on these frequencies. However, recent advances in the field of THz devices give us strong indications that THz communications will be feasible in the near future [3]. These technologies (THz systems in particular), are expected to ease the spectrum limitations of today's systems. They face several issues, such as the reflection and scattering losses through the transmission path, the high dependency between distance and frequency of channels in the THz band and the need for controllable time-delay phase shifters, since the phase shift will vary with frequencies based on the signal traveling time, which will also affect the system performance. These limitations require not only the proper system design, but also the definition of a set of strategies to enable communications [5,6].

The exploration of the potentialities of millimeter and sub-millimeter wavelengths is closely related to the paradigm of using very large arrays of antennas in beamforming architectures. This gives rise to so-called ultra-massive multiple-input multiple-output (UM-MIMO) systems. Still, to achieve the maximum potential of these systems, it is necessary to consider the requirements and the challenges related not only to the channel characteristics but also to the hardware component, especially regarding THz circuits [5,7,8]. Considering that high complexity and power usage are pointed out as major constraints of large-antenna systems, it is unfeasible to implement UM-MIMO schemes with a dedicated RF chain per antenna element in the mmWave and THz bands. Therefore, instead of fully-digital precoders and combiners, it becomes crucial to adopt hybrid digital/analog architectures as these require a reduced number of dedicated RF chains. By adopting this type of design, the signal processing is split into two separate parts: a low-dimensional digital part and a complementary analog part. This approach can enable a substantial reduction in the overall circuit complexity and power consumption [9]. By adopting a proper problem formulation, the analog design part can then be reduced to a simple projection operation in a flexible precoding or combining algorithm that can cope with different architectures, as we proposed in [10,11]. Despite the ultra-wide bandwidths available in mmWave and THz bands, and besides considering the problem of distance limitation, MIMO systems should take into account the operation in frequency selective channels [12]. To make the development of hybrid schemes for these systems a reality, it is necessary to handle the fading caused by multiple propagation paths typical in these types of channels [13]. Therefore, solutions inspired on multi-carrier schemes, such as orthogonal frequency division multiplexing (OFDM), are often adopted to address such problems [14].

Spectral Efficiency (SE) of point-to-point transmissions is a major concern in SingleUser (SU) and MultiUser (MU) systems. To achieve good

performances, it is necessary to develop algorithms that are specially tailored to the architecture of these systems. Several hybrid precoding schemes have been proposed in the literature [15-19]. The authors of [15] proposed two algorithms for low-complexity hybrid precoding and beamforming for MU mmWave systems. Even though they assume only one stream per user, i.e., the number of data streams (N_s) is equal to the number of users (N_u), it is shown that the algorithms achieve interesting results when compared to the fully-digital solution. The concept of precoding based on adaptive RF-chain-to-antenna was only introduced in [16] for SU scenarios, but showed promising results. In [17], a nonlinear hybrid transceiver design relying on Tomlinson–Harashima precoding was proposed. Their approach only considers Fully-Connected (FC) architectures but can achieve a performance close to the fully-digital transceiver. A Kalman-based Hybrid Precoding method was proposed for MU scenarios in [18]. While designed for systems with only one stream per user and based on fully-connected structures, the performance of the algorithm is competitive with other existing solutions. A hybrid MMSE-based precoder and combiner design with low complexity was proposed in [19]. The algorithm is designed for MU-MIMO systems in narrowband channels, and it presents lower complexity and better results when compared to Kalman’s precoding. Most of the hybrid solutions for mmWave systems aim to achieve near-optimal performance using FC structures, resorting to phase shifters or switches. However, the difficulty of handling the hardware constraint imposed by the analog phase shifters or by switches in the THz band is an issue that limits the expected performance in terms of SE.

Array-of-Subarrays (AoSAs) structures have gained particular attention over the last few years as a more practical alternative to FC structures, especially for the THz band. In contrast to FC structures, in which every RF chain is connected to all antennas via an individual group of phase shifters (which is prohibitive for higher frequencies), the AoSA approach allows us to have each RF chain connected to only a reduced subset of antennas. The adoption of a disjointed structure with fewer phase shifters reduces the system complexity, the power consumption and the signal power loss. Moreover, all of the signal processing can be easily carried out at the subarray level by using an adequate number of antennas [6].

Following the AoSA approach, it was shown in [20] that, to balance SE and power consumption in THz communications, adaption and dynamic control capabilities should be included in the hybrid precoding design. Therefore, Dynamic Arrays-of-Subarrays (DAoSAs) architectures could be adopted. The same authors proposed a DAoSA hybrid precoding architecture which can intelligently adjust the connections between RF chains and subarrays through a network of switches. Their results showed that it is possible to achieve a good trade-off for the balancing between the SE and power consumption.

Within the context of multiuser downlink scenarios, the authors of [21] studied some precoding schemes considering THz massive MIMO systems for Beyond 5th Generation (B5G) networks. Besides showing the impact on EE and SE performance, carrier frequency, bandwidth and antenna gains, three different precoding schemes were evaluated and compared. It was observed that the hybrid precoding approach with baseband Zero Forcing for multiuser interference mitigation (HYB-ZF) achieved much better results than an Analog-only Beamsteering (AN-

BST) scheme with no baseband precoder. In fact, this approach was capable of better approaching the upper bound defined by the singular value decomposition precoder (SVD-UB). The other relevant conclusion is that the design of precoding algorithms should be adapted to the communication schemes. While considering all the specific constraints may allow the maximization of the system performance of the system, formulating and solving the corresponding optimization problem may not be so simple. Motivated by the work above, in this paper we developed an algorithm for hybrid precoding design which can accommodate different low-complexity architectures suitable for both mmWave and THz MU-MIMO systems. It is based on the idea of accomplishing a near-optimal approximation of the fully-digital precoder for any configuration of antennas, RF chains and data streams through the application of the alternating direction method of multipliers (ADMM) [22]. ADMM is a well-known and effective method for solving convex optimization problems but can also be a powerful heuristic for several non-convex problems [22,23]. To use it effectively within the context of MU-MIMO, THE proper formulation of the hybrid design problem as a multiple constrained matrix factorization problem is first presented. Using the proposed formulation, an iterative algorithm comprising several reduced complexity steps is obtained.

The main contributions of this paper can be summarized as follows:

- We propose a hybrid design algorithm with near fully-digital performance, where the digital precoder, analog precoder and multiuser interference mitigation are computed separately through simple closed-form solutions. Even though the hybrid design algorithm is developed independently of a specific channel or antenna configuration, it is particularly suitable for mmWave and THz systems where, on the one hand, very large antenna arrays are required to overcome distance limitations but, on the other hand, current hardware constraints in terms of cost and power consumption make the adoption fully-digital precoders/combiners with one dedicated RF chain per antenna element unviable. Whereas our previous work [10] also proposed a hybrid design algorithm for mmWave, it did not address multiuser systems, and in particular the MIMO broadcast channel. Therefore, it does not include any step for inter-user interference mitigation within its design. As we show here, for this multiuser channel, the hybrid design method must also deal with the residual inter-user interference as it can degrade system performance, particularly at high Signal Noise Ratios (SNRs);
- Due to the separability of the different steps (analog precoder, digital precoder and interference suppression), the proposed algorithm can incorporate different architectures, making it suitable for supporting UM-MIMO in severely hardware-constrained systems typical in the THz band. Unlike [10], where we only considered the adoption of phase shifters, in this paper we present explicit solutions for some of the most common architectures, namely FC, AoSA and DAoSA structures based in either Unquantized Phase Shifters (UPS), Quantized Phase Shifters (QPS), Switches (Swi), Switches and Inverters (SI), Antenna Selection (AS) or Double Phase Shifters (DPS);
- To cope with the large bandwidths available in mmWave/THz bands, where practical MIMO systems likely have to operate in frequency selective channels, the proposed hybrid design considers

the application in a multicarrier context, where the same analog precoder is applied at different frequencies;

- We explicitly show how the proposed design can be applied to a DAoSAs approach where a reduced number of switches are inserted at each AoSA panel, which allows the connections to the RF chains to be dynamically adjusted. Through extensive simulations, it is shown that our proposed solution is capable of achieving good trade-offs between spectral efficiency, hardware complexity and power consumption, proving to be a suitable solution for the deployment of UM-MIMO, especially in hardware-constrained THz systems.

The paper is organized as follows: Section II presents the adopted system model. The adopted formulation of the hybrid design problem for the MU-MIMO scenario and the proposed algorithm are described in detail in Section III, which includes the implementation of the algorithm for different analog architectures. Performance results are then presented in Section IV. Finally, the conclusions are outlined in Section V.

Notation: Matrices and vectors are denoted by uppercase and lowercase boldface letters, respectively. The superscript $(.)^T$ and $(.)^H$ denote the transpose and conjugate transpose of a matrix/vector, $\|\cdot\|_p$ is the ℓ_p -norm of a vector, $\|\cdot\|_0$ is its cardinality (i.e., the number of non-zero elements in a vector which is sometimes referred to as the ℓ_0 -norm in the literature) and \mathbf{I}_n is the $n \times n$ identity matrix.

2. System Model

In this section, we present the system and channel models adopted for the design of the hybrid precoding algorithm. Let us consider the OFDM base system illustrated in Figure 1. In this case, we have a mmWave/THz hybrid multiuser MIMO system, where a base station (BS) is equipped with N_{tx} antennas and transmits to N_u users equipped with N_{rx} antennas over F carriers, as can be seen in Figure 1. On each subcarrier, N_s data streams are transmitted to each user which are represented as $\mathbf{s}_k = [\mathbf{s}_{k,1}^T \dots \mathbf{s}_{k,N_u}^T]^T$, with $\mathbf{s}_{k,u} \in \mathbb{C}^{N_s \times 1}$. Instead of a fully-digital design which would require a dedicated RF chain per antenna element, both the precoder and combiner comprise separate digital and analog processing blocks. This approach allows for the use of reduced digital blocks with only a few RF chains, which are complemented by the analog blocks, that can be supported solely on networks of phase shifters and switches. Since the analog precoder (combiner) is located after (before) the IFFT (FFT) blocks, it is shared between the different subcarriers, as in [24,25]. Regarding the analog precoder and combiner, which are represented by matrices $\mathbf{F}_{RF} \in \mathbb{C}^{N_{tx} \times N_{RF}^{tx}}$ and $\mathbf{W}_{RF_u} \in \mathbb{C}^{N_{rx} \times N_{RF}^{rx}}$ with $u=1, \dots, N_u$, it is assumed that $N_u N_s \leq N_{RF}^{tx} \leq N_{tx}$ and $N_s \leq N_{RF}^{rx} \leq N_{rx}$, where N_{RF}^{tx} and N_{RF}^{rx} are the number of RF chains at the BS and each user, respectively. The received signal model at subcarrier k after the combiner can be written as

$$\mathbf{Y}_{k,u} = \sqrt{\rho_u} \mathbf{W}_{BB_{k,u}}^H \mathbf{W}_{RF_u}^H \mathbf{H}_{k,u} \mathbf{F}_{RF} \mathbf{F}_{BB_k} \mathbf{s}_k + \mathbf{W}_{BB_{k,u}}^H \mathbf{W}_{RF_u}^H \mathbf{n}_{k,u}, \quad (1)$$

where $\mathbf{H}_{k,u} \in \mathbb{C}^{N_{rx} \times N_{tx}}$ is the frequency domain channel matrix (assumed to be perfectly known at the transmitter and receiver) between the base station and the u^{th} receiver at subcarrier k . Vector $\mathbf{n}_{k,u} \in \mathbb{C}^{N_{rx} \times 1}$ contains independent zero-mean circularly symmetric Gaussian noise samples with covariance $\sigma_n^2 \mathbf{I}_{N_{rx}}$ and ρ_u denotes the average received power. The digital baseband precoders and combiners are denoted by $\mathbf{F}_{BB_k} \in \mathbb{C}^{N_{RF}^{tx} \times N_u N_s}$ and $\mathbf{W}_{BB_{k,u}} \in \mathbb{C}^{N_{RF}^{rx} \times N_s}$, respectively. Regarding the channel model, it is important to note that even though the mmWave and THz bands share a few commonalities, the THz channel has several peculiarities that distinguish it from the mmWave channel. For example, the very high scattering and diffraction losses in the THz band will typically result in a much sparser channel in the angular domain with fewer multipath components (typically less than 10) [21]. Furthermore, the gap between the line of sight (LOS) and non-line of sight (NLOS) components tends to be very large, making it often LOS-dominant with NLOS-assisted [25]. An additional aspect relies on the much larger bandwidth of THz signals which can suffer performance degradation due to the so-called beam split effect, where the transmission paths squint into different spatial directions depending on the subcarrier frequency [20]. In light of this, in this paper we consider a clustered wideband geometric channel, which is commonly adopted both in the mmWave [15] and THz literature [20,25-38]. However, it should be noted that the hybrid precoding/combining approach proposed in this paper is independent of a specific MIMO channel. In this case, the frequency domain channel matrices can be characterized as

$$\mathbf{H}_{k,u} = \gamma \left(\alpha_u^{LOS} \mathbf{a}_r(\phi_u^{r,LOS}, \theta_u^{r,LOS}) \mathbf{a}_t(\phi_u^{t,LOS}, \theta_u^{t,LOS})^H + \sum_{i=1}^{N_{cl}} \sum_{l=1}^{N_{ray}} \alpha_{i,l,u} \mathbf{a}_r(\phi_{i,l,u}^r, \theta_{i,l,u}^r) \mathbf{a}_t(\phi_{i,l,u}^t, \theta_{i,l,u}^t)^H \right) e^{-j2\pi\tau_{i,u}f_k}, \quad (2)$$

where N_{cl} denotes the scattering clusters with each cluster i having a time delay of $\tau_{i,u}$ and N_{ray} propagations paths. α_u^{LOS} $\alpha_{i,l,u}$ are the complex gains of the LOS component and of the l th ray from cluster i . Index u is the user ($u = 1, \dots, N_u$), $f_k = f_c + \frac{B}{F} \left(k - 1 - \frac{F-1}{2} \right)$ ($k = 1, \dots, F$) is the k^{th} subcarrier frequency, B is the bandwidth, f_c is the central frequency and γ is a normalizing factor such that $\mathbb{E}[\|\mathbf{H}_{k,u}\|_F^2] = N_{tx} N_{rx}$. Vectors $\mathbf{a}_t(\phi_{i,l,u}^t, \theta_{i,l,u}^t)$ and $\mathbf{a}_r(\phi_{i,l,u}^r, \theta_{i,l,u}^r)$ represent the transmit and receive antenna array responses at the azimuth and elevation angles of $(\phi_{i,l,u}^t, \theta_{i,l,u}^t)$ and $(\phi_{i,l,u}^r, \theta_{i,l,u}^r)$, respectively. Vectors $\mathbf{a}_t(\phi_u^{t,LOS}, \theta_u^{t,LOS})$ and $\mathbf{a}_r(\phi_u^{r,LOS}, \theta_u^{r,LOS})$ have similar meanings but refer to the LOS path angles $(\phi_u^{t,LOS}, \theta_u^{t,LOS})$ and $(\phi_u^{r,LOS}, \theta_u^{r,LOS})$. By carefully selecting the parameters of the channel model we can make it depict a mmWave or a THz channel.

Considering Gaussian signaling, the spectral efficiency achieved by the system for the transmission to MS- u in subcarrier k is [28]

$$\mathbf{R}_{k,u} = \log_2 \left| \mathbf{I}_{N_{RF}^{rx}} + \mathbf{R}_u^{-1} \mathbf{W}_{BB,k,u}^H \mathbf{W}_{RF,u}^H \mathbf{H}_{k,u} \mathbf{F}_{BB,k,u} \times \mathbf{F}_{BB,k,u}^H \mathbf{F}_{RF,k,u}^H \mathbf{H}_{k,u}^H \mathbf{W}_{RF,u} \mathbf{W}_{BB,k,u} \right|, \quad (3)$$

where $\mathbf{R}_{k,u}$ is the covariance matrix of the total inter-user interference plus noise at MS- u , which is characterized by

$$\mathbf{R}_{k,u} = \mathbf{W}_{BB,k,u}^H \mathbf{W}_{RF,u}^H (\mathbf{H}_{k,u} \sum_{j \neq u} \mathbf{F}_{RF,j} \mathbf{F}_{BB,k,j}^H \mathbf{F}_{BB,k,j}^H \mathbf{F}_{RF,j}^H \mathbf{H}_{k,u}^H + \sigma^2 \mathbf{I}_{N_{rx}}) \mathbf{W}_{RF,u} \mathbf{W}_{BB,k,u}. \quad (4)$$

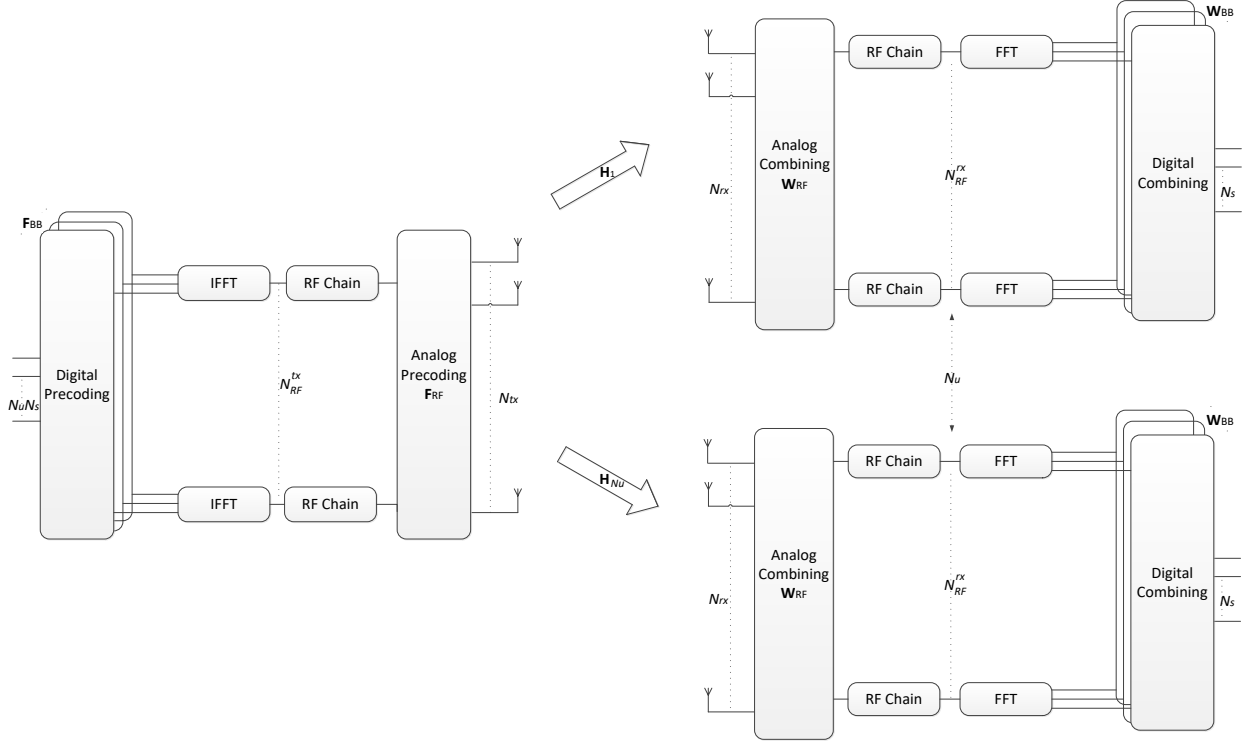


Figure 1. A multiuser OFDM mmWave/THz MIMO system with hybrid precoding.

3. Proposed Hybrid Design Algorithm

In this section, we will introduce the algorithm for the hybrid precoding problem and show how it can be adapted to different architectures. Although we will focus on the precoder design, a similar approach can be adopted for the combiner. However, since our design assumes that inter-user interference suppression is applied at the transmitter, only single-user detection is required at the receiver and therefore the algorithm reduces to the one described in [10].

3.1. Main Algorithm

Although there are several problem formulations for the hybrid design proposed in the literature, one of the most effective relies on the minimization of the Frobenius norm of the difference between the fully-digital precoder and the hybrid precoder [22,29-31]. In this paper we follow this matrix approximation-based approach. First, we compute fully-digital precoders, which we assume to be designed so as to enforce zero inter-user interference, using, for example, the block-diagonalization (BD) approach described in [32]. Using the BD procedure, we obtain one

different digital precoder matrix for each subcarrier, $\mathbf{F}_{\text{opt}_k}$ (with $k=1, \dots, F$, and $\mathbf{F}_{\text{opt}_k} = [\mathbf{F}_{\text{opt}_{k,1}}, \dots, \mathbf{F}_{\text{opt}_{k,N_u}}]$), which satisfy $\mathbf{H}_{k,u'} \mathbf{F}_{k,u} = \mathbf{0}$ for all $u' \neq u$ ($u, u' = 1, \dots, N_u$), thus guaranteeing no inter-user interference. Using these digital precoder matrices, we design the hybrid digital/analog precoder by solving a matrix approximation problem formulated as

$$\min_{\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}_k}} \sum_{k=1}^F \left\| \mathbf{F}_{\text{opt}_k} - \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_k} \right\|_F^2 \quad (5)$$

$$\text{subject to } \mathbf{F}_{\text{RF}} \in \mathcal{C}_{N_{\text{tx}} \times N_{\text{RF}}^{\text{tx}}} \quad (6)$$

$$\left\| \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_k} \right\|_F^2 = N_u N_s \quad (7)$$

where Equation (7) enforces the transmitter's total power constraint and $\mathcal{C}_{N_{\text{tx}} \times N_{\text{RF}}^{\text{tx}}}$ is the set of feasible analog precoding matrices, which is defined according to the adopted RF architecture (it will be formally defined for several different architectures in the next subsection). Matrix $\mathbf{F}_{\text{opt}_k}$ denotes the fully-digital precoder. Even if $\mathbf{F}_{\text{opt}_k}$ is selected in order to cancel all interference between users, the hybrid design resulting as a solution of Equations (5)–(7) will correspond to an approximation and, as such, residual inter-user interference will remain. To avoid the performance degradation that will result from this, an additional constraint can be added to the problem formulation, namely

$$\sum_{\substack{u'=1 \\ u' \neq u}}^{N_u} \mathbf{H}_{k,u'} \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_{k,u}} = \mathbf{0}, \quad k=1, \dots, F, \quad u=1, \dots, N_u \quad (8)$$

where $\mathbf{F}_{\text{BB}_{k,u}} = \mathbf{F}_{\text{BB}_k}[:, (u-1)N_s + 1 : uN_s]$. This restriction is equivalent to enforcing $\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_{k,u}}$ to lie in the null space of $\bar{\mathbf{H}}_{k,u} \in \mathbb{C}^{(N_u-1)N_{\text{rx}} \times N_{\text{tx}}}$ ($\bar{\mathbf{H}}_{k,u}$ is a matrix corresponding to \mathbf{H}_k with the N_{rx} lines of user u removed) which we denote as $\mathcal{N}(\bar{\mathbf{H}}_{k,u})$. The overall optimization problem can be then expressed as

$$\min_{\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}_k}} \sum_{k=1}^F \left\| \mathbf{F}_{\text{opt}_k} - \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_k} \right\|_F^2 \quad (9)$$

$$\text{subject to } \mathbf{F}_{\text{RF}} \in \mathcal{C}_{N_{\text{tx}} \times N_{\text{RF}}^{\text{tx}}} \quad (10)$$

$$\left\| \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_k} \right\|_F^2 = N_u N_s \quad (11)$$

$$\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_{k,u}} \in \mathcal{N}(\bar{\mathbf{H}}_{k,u}), \quad k=1, \dots, F, \quad u=1, \dots, N_u. \quad (12)$$

To derive a hybrid precoder/design algorithm that can cope with the different RF architectures, we can integrate the RF constraint directly into the objective function of the optimization problem. This can be accomplished through the addition of an auxiliary variable, \mathbf{R} , combined with the use of the indicator function. The indicator function for a generic

set \mathcal{A} is defined as $I_{\mathcal{A}}(\mathbf{x})$, returning 0 if $\mathbf{x} \in \mathcal{A}$ and $+\infty$ otherwise. A similar approach can be adopted for integrating the other constraints, Equations (11) and (12), also into the objective function. The optimization problem can then be rewritten as

$$\begin{aligned} \min_{\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}_k}, \mathbf{R}, \mathbf{B}_k, \mathbf{F}_{\text{aprox}_{k,u}}} \quad & \sum_{k=1}^F \left\| \mathbf{F}_{\text{opt}_k} - \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_k} \right\|_F^2 + I_{C_{N_{\text{tx}} \times N_{\text{RF}}^{\text{tx}}}}(\mathbf{R}) + \\ & + \sum_{k=1}^F I_{\left\| \cdot \right\|_F^2 = N_u N_s}(\mathbf{B}_k) + \sum_{k=1}^F \sum_{u=1}^{N_u} I_{\mathcal{N}(\bar{\mathbf{H}}_{k,u})}(\mathbf{F}_{\text{aprox}_{k,u}}) \end{aligned} \quad (13)$$

$$\text{subject to } \mathbf{R} = \mathbf{F}_{\text{RF}} \quad (14)$$

$$\mathbf{B}_k = \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_k} \quad (15)$$

$$\mathbf{F}_{\text{aprox}_{k,u}} = \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_k}, \quad (16)$$

where $\mathbf{F}_{\text{aprox}_k} = [\mathbf{F}_{\text{aprox}_{k,1}}, \dots, \mathbf{F}_{\text{aprox}_{k,N_u}}]$. The augmented Lagrangian function (ALF) for Equations (13)–(16) can be written as

$$\begin{aligned} L_{\rho, \eta, \mu}(\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}, \mathbf{R}, \mathbf{B}, \mathbf{F}_{\text{aprox}}, \mathbf{\Lambda}, \mathbf{\Psi}, \mathbf{\Gamma}) = & \sum_{k=1}^F \left\| \mathbf{F}_{\text{opt}_k} - \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_k} \right\|_F^2 + I_{C_{N_{\text{tx}} \times N_{\text{RF}}^{\text{tx}}}}(\mathbf{R}) + \\ & + \sum_{k=1}^K I_{\left\| \cdot \right\|_F^2 = N_u N_s}(\mathbf{B}_k) + \sum_{k=1}^F \sum_{u=1}^{N_u} I_{\mathcal{N}(\bar{\mathbf{H}}_{k,u})}(\mathbf{F}_{\text{aprox}_{k,u}}) + 2 \text{Re}\{\text{tr}(\mathbf{\Lambda}^H (\mathbf{F}_{\text{RF}} - \mathbf{R}) + \\ & + \text{tr}(\mathbf{\Psi}_k^H \sum_{k=1}^F (-\mathbf{B}_k + \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_k})) + \text{tr}(\mathbf{\Gamma}_k^H \sum_{k=1}^F (-\mathbf{F}_{\text{aprox}_k} + \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_k}))\} + \rho \left\| \mathbf{F}_{\text{RF}} - \mathbf{R} \right\|_F^2 + \\ & + \eta \sum_{k=1}^F \left\| -\mathbf{B}_k + \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_k} \right\|_F^2 + \mu \sum_{k=1}^F \left\| -\mathbf{F}_{\text{aprox}_k} + \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_k} \right\|_F^2, \end{aligned} \quad (17)$$

where $\mathbf{\Lambda} \in \mathbb{C}^{N_{\text{tx}} \times N_{\text{RF}}}$, $\mathbf{\Psi} \in \mathbb{C}^{N_{\text{tx}} \times N_{\text{RF}}}$ and $\mathbf{\Gamma} \in \mathbb{C}^{N_{\text{tx}} \times N_{\text{RF}}}$ are dual variables and ρ , η , μ are penalty parameters. After some straightforward algebraic manipulation and working with scaled dual variables $\mathbf{U} = \mathbf{\Lambda}/\rho$, $\mathbf{W}_k = \mathbf{\Psi}/\eta$ and $\mathbf{Z}_k = \mathbf{\Gamma}_k/\mu$ we can rewrite the ALF as

$$\begin{aligned} L_{\rho, \eta, \mu}(\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}, \mathbf{R}, \mathbf{B}, \mathbf{F}_{\text{aprox}}, \mathbf{U}, \mathbf{W}, \mathbf{Z}) = & \sum_{k=1}^K \left\| \mathbf{F}_{\text{opt}_k} - \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_k} \right\|_F^2 + I_{C_{N_{\text{tx}} \times N_{\text{RF}}^{\text{tx}}}}(\mathbf{R}) + \\ & + \sum_{k=1}^F I_{\left\| \cdot \right\|_F^2 = N_u N_s}(\mathbf{B}_k) + \sum_{k=1}^F \sum_{u=1}^{N_u} I_{\mathcal{N}(\bar{\mathbf{H}}_{k,u})}(\mathbf{F}_{\text{aprox}_{k,u}}) + \rho \left\| \mathbf{F}_{\text{RF}} - \mathbf{R} + \mathbf{U} \right\|_F^2 - \rho \left\| \mathbf{U} \right\|_F^2 + \\ & + \eta \sum_{k=1}^F \left\| -\mathbf{B}_k + \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_k} + \mathbf{W}_k \right\|_F^2 - \eta \sum_{k=1}^F \left\| \mathbf{W}_k \right\|_F^2 + \\ & + \mu \sum_{k=1}^F \left\| -\mathbf{F}_{\text{aprox}_k} + \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_k} + \mathbf{Z}_k \right\|_F^2 - \mu \sum_{k=1}^F \left\| \mathbf{Z}_k \right\|_F^2. \end{aligned} \quad (18)$$

In the following, we apply ADMM [22] as a heuristic for solving problem Formulation (13)–(16). To accomplish this, we can apply the gradient ascent to the dual problem involving the ALF, which allows us to obtain an iterative precoding algorithm comprising the following sequence of steps. We start with the minimization of the ALF over \mathbf{F}_{RF} for iteration $t+1$ defined as

$$\mathbf{F}_{\text{RF}}^{(t+1)} = \min_{\mathbf{F}_{\text{RF}}} L_{\rho, \eta, \mu}(\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}^{(t)}, \mathbf{R}^{(t)}, \mathbf{B}^{(t)}, \mathbf{F}_{\text{aprox}}^{(t)}, \mathbf{U}^{(t)}, \mathbf{W}^{(t)}, \mathbf{Z}^{(t)}), \quad (19)$$

which can be obtained from

$$\nabla_{\mathbf{F}_{\text{RF}}^H} L_{\rho, \eta, \mu}(\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}^{(t)}, \mathbf{R}^{(t)}, \mathbf{B}^{(t)}, \mathbf{F}_{\text{aprox}}^{(t)}, \mathbf{U}^{(t)}, \mathbf{W}^{(t)}, \mathbf{Z}^{(t)}) = 0 \quad (20)$$

leading to the closed form expression

$$\begin{aligned} \mathbf{F}_{\text{RF}}^{(t+1)} = & \left[\sum_{k=1}^F [\mathbf{F}_{\text{opt}_k} + \eta(\mathbf{B}_k^{(t)} - \mathbf{W}_k^{(t)}) + \mu(\mathbf{F}_{\text{aprox}_k}^{(t)} - \mathbf{Z}_k^{(t)})] \times \mathbf{F}_{\text{BB}_k}^{(t)H} + \rho(\mathbf{R}^{(t)} - \mathbf{U}^{(t)}) \right] \times \\ & \times [(1 + \eta + \mu) \sum_{k=1}^F \mathbf{F}_{\text{BB}_k}^{(t)} \mathbf{F}_{\text{BB}_k}^{(t)H} + \rho \mathbf{I}_{N_{\text{RF}}^{\text{tx}}}]^{-1}. \end{aligned} \quad (21)$$

After obtaining the expression for \mathbf{F}_{RF} , $\mathbf{F}_{\text{BB}}^{(t+1)}$ can be found by following the same methodology. In this case the minimization is expressed as

$$\mathbf{F}_{\text{BB}}^{(t+1)} = \min_{\mathbf{F}_{\text{BB}}} L_{\rho, \eta, \mu}(\mathbf{F}_{\text{RF}}^{(t)}, \mathbf{F}_{\text{BB}}, \mathbf{R}^{(t)}, \mathbf{B}^{(t)}, \mathbf{F}_{\text{aprox}}^{(t)}, \mathbf{U}^{(t)}, \mathbf{W}^{(t)}, \mathbf{Z}^{(t)}), \quad (22)$$

which can be obtained from

$$\nabla_{\mathbf{F}_{\text{BB}}^H} L_{\rho, \eta, \mu}(\mathbf{F}_{\text{RF}}^{(t)}, \mathbf{F}_{\text{BB}}, \mathbf{R}^{(t)}, \mathbf{B}^{(t)}, \mathbf{F}_{\text{aprox}}^{(t)}, \mathbf{U}^{(t)}, \mathbf{W}^{(t)}, \mathbf{Z}^{(t)}) = 0 \quad (23)$$

and leads to the closed form expression

$$\begin{aligned} \mathbf{F}_{\text{BB}_k}^{(t+1)} = & (1 + \eta + \mu) \left(\mathbf{F}_{\text{RF}}^{(t+1)H} \mathbf{F}_{\text{RF}}^{(t+1)} \right)^{-1} \mathbf{F}_{\text{RF}}^{(t+1)H} \cdot \\ & \cdot \left(\mathbf{F}_{\text{opt}_k} + \eta(\mathbf{B}_k^{(t)} - \mathbf{W}_k^{(t)}) + \mu(\mathbf{F}_{\text{aprox}_k}^{(t)} - \mathbf{Z}_k^{(t)}) \right), k = 1, \dots, F. \end{aligned} \quad (24)$$

The next steps consist of the minimization over \mathbf{R} and \mathbf{B}_k . The minimization of Equation (18) with respect to \mathbf{R} and \mathbf{B}_k can be written as

$$\begin{aligned} \mathbf{R}^{(t+1)} = & \min_{\mathbf{R}} \left\{ \mathbf{I}_{C_{N_{\text{tx}} \times N_{\text{RF}}}}(\mathbf{R}) + \rho \left\| \mathbf{F}_{\text{RF}}^{(t+1)} - \mathbf{R} + \mathbf{U}^{(t)} \right\|_F^2 \right\} \\ = & \Pi_{C_{N_{\text{tx}} \times N_{\text{RF}}}} \left(\mathbf{F}_{\text{RF}}^{(t+1)} + \mathbf{U}^{(t)} \right), \end{aligned} \quad (25)$$

and

$$\begin{aligned} \mathbf{B}_k^{(t+1)} = & \min_{\mathbf{B}_k} \left\{ \mathbf{I}_{\|\mathbf{F}\|_F^2 = N_u N_s}(\mathbf{B}_k) + \eta \left\| \mathbf{F}_{\text{RF}}^{(t+1)} \mathbf{F}_{\text{BB}_k}^{(t+1)} - \mathbf{B}_k + \mathbf{W}^{(t)} \right\|_F^2 \right\} \\ = & \Pi_{\|\mathbf{F}\|_F^2 = N_u N_s} \left(\mathbf{F}_{\text{RF}}^{(t+1)} \mathbf{F}_{\text{BB}_k}^{(t+1)} + \mathbf{W}^{(t)} \right), k = 1, \dots, F, \end{aligned} \quad (26)$$

where $\Pi_{\mathcal{C}_{a \times b}}(\cdot)$ and $\Pi_{\|\cdot\|_F^2 = N_u N_s}(\cdot)$ denote the projection onto set $\mathcal{C}_{a \times b}$ and onto the set of matrices whose squared Frobenius norm is $N_u N_s$, respectively. While the former projection depends on the adopted analog architecture and will be explained in the next subsection, the second projection is simply computed as

$$\mathbf{B}_k^{(t+1)} = \frac{\left(\mathbf{F}_{\text{RF}}^{(t+1)} \mathbf{F}_{\text{BB}_k}^{(t+1)} + \mathbf{W}^{(t)} \right) \sqrt{N_u N_s}}{\left\| \mathbf{F}_{\text{RF}}^{(t+1)} \mathbf{F}_{\text{BB}_k}^{(t+1)} + \mathbf{W}^{(t)} \right\|_F^2}. \quad (27)$$

The minimization of (18) with respect of $\mathbf{F}_{\text{aprox}_{k,u}}$ can be written as

$$\begin{aligned} \mathbf{F}_{\text{aprox}_{k,u}}^{(t+1)} &= \min_{\mathbf{F}_{\text{aprox}_{k,u}}} \{ \mathbf{I}_{\mathcal{N}(\bar{\mathbf{H}}_{k,u})} (\mathbf{F}_{\text{aprox}_{k,u}}) + \mu \left\| \mathbf{F}_{\text{RF}}^{(t+1)} \mathbf{F}_{\text{BB}_{k,(u-1)N_s+1:uN_s}}^{(t+1)} - \mathbf{F}_{\text{aprox}_{k,u}} + \mathbf{Z}_{k,(u-1)N_s+1:uN_s}^{(t)} \right\|_F^2 \} \\ &= \Pi_{\mathcal{N}(\bar{\mathbf{H}}_{k,u})} \left(\mathbf{F}_{\text{RF}}^{(t+1)} \mathbf{F}_{\text{BB}_{k,(u-1)N_s+1:uN_s}}^{(t+1)} + \mathbf{Z}_{k,(u-1)N_s+1:uN_s}^{(t)} \right), \end{aligned} \quad (28)$$

which also involves a projection, $\Pi_{\mathcal{N}(\bar{\mathbf{H}}_{k,u})}(\cdot)$, but in this case onto the null-space of $\bar{\mathbf{H}}_{k,u}$. Let us use \mathbf{A} to denote $\mathbf{A} = \mathbf{F}_{\text{RF}}^{(t+1)} \mathbf{F}_{\text{BB}_{k,u}}^{(t+1)} + \mathbf{Z}_{k,(u-1)N_s+1:uN_s}^{(t)}$. The procedure to compute the projection of matrix \mathbf{A} onto the null-space of $\bar{\mathbf{H}}_{k,u}$ can be formulated as another optimization problem, which can be expressed as

$$\min \sum_{i=1}^{N_s} \left\| \mathbf{A}_{:,i} - \mathbf{X}_{:,i} \right\|_F^2 \quad (29)$$

$$\text{subject to } \bar{\mathbf{H}}_{k,u} \mathbf{X}_{:,i} = 0. \quad (30)$$

The general solution for this problem is presented in [29] corresponding to

$$\mathbf{X}_{:,i} = \left(\mathbf{I}_{N_{\text{tx}}} - \bar{\mathbf{H}}_{k,u}^H \left(\bar{\mathbf{H}}_{k,u} \bar{\mathbf{H}}_{k,u}^H \right)^{-1} \bar{\mathbf{H}}_{k,u} \right) \mathbf{A}_{:,i}, \quad i = 1, \dots, N_s. \quad (31)$$

Reordering the column vectors in the original matrix form results in the final expression which can be rewritten as

$$\begin{aligned} \mathbf{X} &= \left(\mathbf{I}_{N_{\text{tx}}} - \bar{\mathbf{H}}_{k,u}^H \left(\bar{\mathbf{H}}_{k,u} \bar{\mathbf{H}}_{k,u}^H \right)^{-1} \bar{\mathbf{H}}_{k,u} \right) \mathbf{A} \\ &= \left(\mathbf{I}_{N_{\text{tx}}} - \bar{\mathbf{V}}_{k,u}^{(1)} \left(\bar{\mathbf{V}}_{k,u}^{(1)} \right)^H \right) \mathbf{A}. \end{aligned} \quad (32)$$

In this expression, $\bar{\mathbf{V}}_{k,u}^{(1)}$ denotes the matrix containing the right singular vectors corresponding to the non-zero singular values associated with the singular value decomposition (SVD) given by $\bar{\mathbf{H}}_{k,u} = \bar{\mathbf{U}}_{k,u} \bar{\mathbf{\Lambda}}_{k,u} \left[\bar{\mathbf{V}}_{k,u}^{(1)} \bar{\mathbf{V}}_{k,u}^{(0)} \right]^H$. Therefore, to compute matrix \mathbf{X} , one can perform a single value decomposition of $\bar{\mathbf{H}}_{k,u}$ and then use this to remove

the projection of \mathbf{A} onto the row space of $\bar{\mathbf{H}}_{k,u}$. Finally, the expressions for the update of dual variables \mathbf{U} , \mathbf{W} and \mathbf{Z} are given by

$$\mathbf{U}^{(t+1)} = \mathbf{U}^{(t)} + \mathbf{F}_{\text{RF}}^{(t+1)} - \mathbf{R}^{(t+1)}, \quad (33)$$

$$\mathbf{W}_k^{(t+1)} = \mathbf{W}_k^{(t)} + \mathbf{F}_{\text{RF}}^{(t+1)} \mathbf{F}_{\text{BB}_k}^{(t+1)} - \mathbf{B}_k^{(t+1)}, \quad (34)$$

$$\mathbf{Z}_k^{(t+1)} = \mathbf{Z}_k^{(t)} + \mathbf{F}_{\text{RF}}^{(t+1)} \mathbf{F}_{\text{BB}_k}^{(t+1)} - \mathbf{F}_{\text{aprox}_k}^{(t+1)}. \quad (35)$$

Appropriate values for the penalty parameters can be obtained in a heuristic manner by performing numerical simulations. Regarding the initialization and termination of the algorithm, the same approach described in [10] can be adopted. The whole algorithm is summarized in Table 1. In this table, Q denotes the maximum number of iterations. The projection operation is the only step specific to the implemented architecture, as will be explained in the next subsection.

Table 1. General Iterative Hybrid Design Algorithm.

1: Input: $\mathbf{F}_{\text{opt}_k}$, $\mathbf{F}_{\text{RF}}^{(0)}$, $\mathbf{F}_{\text{BB}_k}^{(0)}$, $\mathbf{R}^{(0)}$, $\mathbf{B}_k^{(0)}$, $\mathbf{F}_{\text{aprox}_{k,u}}^{(0)}$, ρ , Q
2: for $t = 0, 1, \dots, Q-1$ do
3: Compute $\mathbf{F}_{\text{RF}}^{(t+1)}$ using (21).
4: Compute $\mathbf{F}_{\text{BB}_k}^{(t+1)}$ using (24), for all $k = 1, \dots, F$.
5: Compute $\mathbf{R}^{(t+1)}$ using (25).
6: Compute $\mathbf{B}_k^{(t+1)}$ using (26), for all $k = 1, \dots, F$.
7: Compute $\mathbf{F}_{\text{aprox}_{k,u}}^{(t+1)}$ using (28), for all $k = 1, \dots, F$ and $u = 1, \dots, N_u$.
8: Update $\mathbf{U}^{(t+1)}$ using (33).
9: Update $\mathbf{W}_k^{(t+1)}$ using (34), for all $k = 1, \dots, F$.
10: Update $\mathbf{Z}_k^{(t+1)}$ using (35), for all $k = 1, \dots, F$.
11: end for.
12: $\hat{\mathbf{F}}_{\text{RF}} \leftarrow \mathbf{R}^{(Q)}$.
13: $\hat{\mathbf{F}}_{\text{BB}_k} \leftarrow \left(\hat{\mathbf{F}}_{\text{RF}}^H \hat{\mathbf{F}}_{\text{RF}} \right)^{-1} \hat{\mathbf{F}}_{\text{RF}}^H \mathbf{F}_{\text{aprox}_k}^{(Q)}$, for all $k = 1, \dots, F$.
14: $\hat{\mathbf{F}}_{\text{BB}_k} \leftarrow \sqrt{N_u N_s} \left\ \hat{\mathbf{F}}_{\text{BB}_k}^H \hat{\mathbf{F}}_{\text{BB}_k} \right\ _F^{-1} \hat{\mathbf{F}}_{\text{BB}_k}$.
15: Output: $\hat{\mathbf{F}}_{\text{RF}}$, $\hat{\mathbf{F}}_{\text{BB}}$.

3.2. Analog RF Precoder/Combiner Structure

The projection required for obtaining matrix \mathbf{R} in step 5 of the precoding algorithm has to be implemented according to the specific analog beamformer [6,20,33-37]. This makes the proposed scheme very generic, allowing it to be easily adapted to different RF architectures. In the following, we will consider a broad range of architectures that can be adopted for the RF precoder for achieving reduced complexity and power consumption implementations. We will consider FC, AoSA and DAoSA structures as illustrated in Figure 2, where we assume single phase shifters (SPS). Besides SPS, we will also consider other alternative implementations for these structures, as illustrated in Figure 3 for AoSA. The different solutions either rely on selectors, switches, inverters or phase shifters, or combinations of these. The overall analog structure is

defined as a combination of one of the architectures in Figure 2 with either SPS or one of the alternatives illustrated in Figure 3.

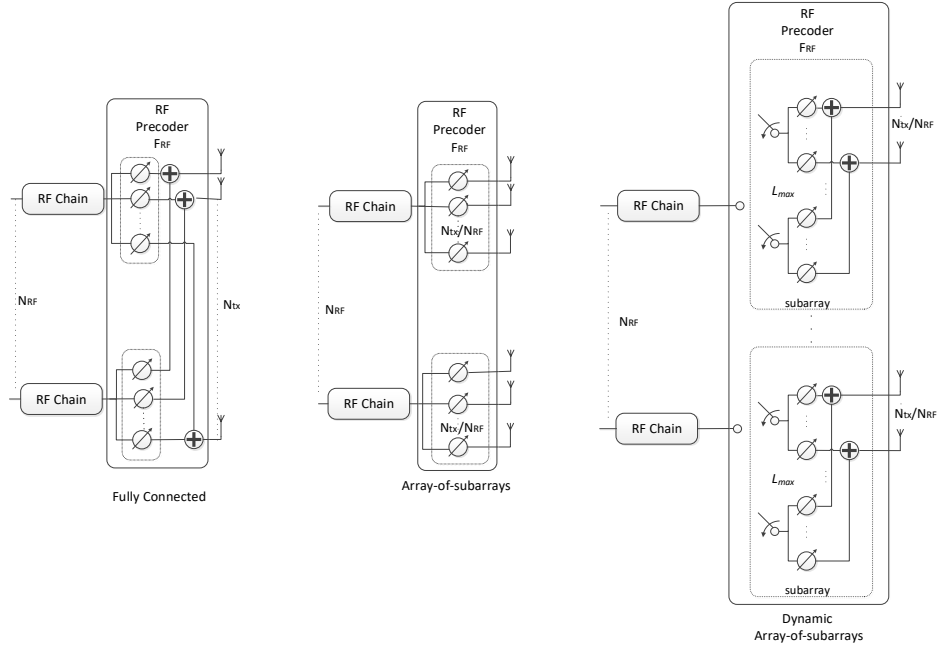


Figure 2. Different precoder architectures for a mmWave/THz MIMO system based on phase shifters.

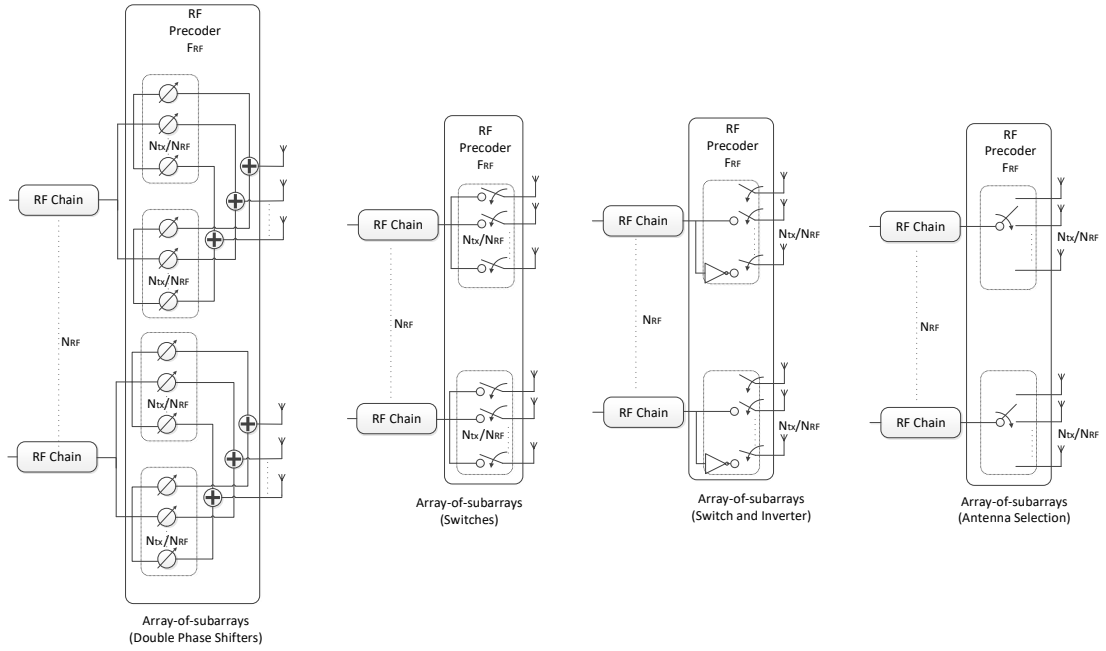


Figure 3. Alternative implementations to single phase shifters based on array-of-subarrays for a mmWave/THz MIMO system.

(1) Unquantized Phase Shifters (UPS)

In the first case, we consider the use of infinite resolution phase shifters which, while being ideal, are often used as a reference benchmark. For this architecture the RF constraint set is given by

$$\mathcal{C}_{a \times b} = \left\{ \mathbf{X} \in \mathbb{C}^{a \times b} : |X_{i,j}| = 1 \right\} \quad (36)$$

and the corresponding projection can be performed simply using

$$\mathbf{R}^{(t+1)} = \left(\mathbf{F}_{\text{RF}}^{(t+1)} + \mathbf{W}^{(t)} \right) \oslash \left| \mathbf{F}_{\text{RF}}^{(t+1)} + \mathbf{W}^{(t)} \right|, \quad (37)$$

where \oslash denotes the Hadamard (i.e., element-wise) division.

(2) Quantized Phase Shifters (QPS)

The second case considers a more realistic scenario, in which phase shifters can be digitally controlled with N_b bits. These devices allow the selection of 2^{N_b} different quantized phases and the RF constraint set becomes

$$\mathcal{C}_{a \times b} = \left\{ \mathbf{X} \in \mathbb{C}^{a \times b} : X_{i,j} = e^{2\pi k i / 2^{N_b}}, k=0, \dots, 2^{N_b} - 1 \right\}. \quad (38)$$

The implementation of the projection in line 5 of Table 1 can be obtained as the following element-wise quantization

$$R_{i,j}^{(t+1)} = e^{\min_{k=0, \dots, 2^{N_b}-1} \left\{ \text{angle} \left(F_{\text{RF},i,j}^{(t+1)} + W_{i,j}^{(t)} \right) - 2\pi k / 2^{N_b} \right\}}, i=1, \dots, N_{\text{tx}}, j=1, \dots, N_{\text{RF}}^{\text{tx}}. \quad (39)$$

Phase shifters are typically one of the best solutions for analog processing blocks but also have a higher implementation cost and power consumption, especially when they have high resolutions.

(3) Double Phase Shifters (DPS)

Another appealing architecture relies on the use of double phase shifters (DPS) since these remove the constant modulus restriction on the elements of \mathbf{F}_{RF} , following the idea in [37]. The main difference between SPS and DPS structures relies on the number of phase shifters in use to compose each connection from an RF chain to a connected antenna element, which in this case is doubled. Even though it increases the implementation complexity and power consumption, this solution can increase the spectral efficiency and approach the performance of the fully-digital one [37]. In this case, the projection can be implemented element-wise simply as

$$R_{i,j}^{(t+1)} = \left(F_{\text{RF},i,j}^{(t+1)} + W_{i,j}^{(t)} \right) - e^{i \cdot \text{angle} \left(F_{\text{RF},i,j}^{(t+1)} + W_{i,j}^{(t)} \right)} \times \max \left(0, \left| F_{\text{RF},i,j}^{(t+1)} + W_{i,j}^{(t)} \right| - 2 \right) \quad (40)$$

Similarly to other architectures, DPS can be used not only in the fully-connected approach but also in the AoSA and DAoSA cases, replacing the constant modulus setting operation.

(4) Switches (Swi)

In an architecture based on switches, each of the variable phase shifters can be replaced by a switch which typically consumes less power [34]. This simplification results in a network of switches connecting each RF chain to the antennas. The RF constraint set can be represented as

$$\mathcal{C}_{a \times b} = \left\{ \mathbf{X} \in \mathbb{R}^{a \times b} : X_{i,j} = 0 \text{ or } X_{i,j} = 1 \right\} \quad (41)$$

(set of matrices having solely '0's or '1's as elements) and the projection can be implemented element-wise as

$$R_{i,j}^{(t+1)} = 1/2 + 1/2 \cdot \text{sign}\left(2\text{Re}\left[F_{\text{RF},i,j}^{(t+1)} + W_{i,j}^{(t)}\right] - 1\right). \quad (42)$$

(5) Switches and Inverters (SI)

Assuming that $N_b = 1$, then each variable phase shifter of the QPS architecture can be replaced by a pair of switched lines, including also an inverter. The corresponding constraint set can be reduced to

$$\mathcal{C}_{a \times b} = \{\mathbf{X} \in \mathbb{R}^{a \times b} : X_{i,j} = \pm 1\} \quad (43)$$

and the implementation of the projection simplifies to

$$R_{i,j}^{(t+1)} = \text{sign}\left(\text{Re}\left[F_{\text{RF},i,j}^{(t+1)} + W_{i,j}^{(t)}\right]\right). \quad (44)$$

(6) Antenna Selection (AS)

The simplest scenario that we can consider corresponds to an architecture, where each RF chain can be only connected to a single antenna (and vice-versa). Antenna selection is a low-cost low-complexity alternative where only a specified subset of antennas are active at any given time [34]. The RF constraint set will comprise a matrix with only one non-zero element per column and per row, i.e.,

$$\mathcal{C}_{a \times b} = \{\mathbf{X} \in \mathbb{R}^{a \times b} : X_{i,j} = 0 \text{ or } X_{i,j} = 1, \|X_{i,:}\|_0 = 1, \|X_{:,j}\|_0 = 1\}. \quad (45)$$

In the definition, $\|\cdot\|_0$ represents the cardinality of a vector (sometimes referred to as the ℓ_0 -norm in the literature). Defining $\mathbf{X} = \mathbf{F}_{\text{RF}}^{(t+1)} + \mathbf{W}^{(t)}$, the projection can be approximately implemented by setting all the elements in \mathbf{X} as 0 except for $X_{t_j,j} = 1$, where t_j is the row position with the highest real component in column j :

$$t_j = \arg \max_{i=1, \dots, N_{\text{tx}}} \{\text{Real}[X_{i,j}]\}. \quad (46)$$

The computation of t_j is performed for all columns $j = 1, \dots, N_{\text{RF}}^{\text{tx}}$, sorted by descending order in terms of highest real components. It should be noted that during this operation, the same row cannot be repeated.

(7) Array-of-Subarrays (AoSAs)

Within the context of UM-MIMO, one of the most appealing architectures for keeping the complexity acceptable relies on the use of AoSA, where each RF chain is only connected to one or more subsets of antennas (subarrays). Denoting the number of subarrays as n_{SA} , which is typically set as $n_{\text{SA}} = N_{\text{RF}}$, and the size of each subarray as $N_{\text{tx}}^{\text{SA}}$, then we have $N_{\text{tx}}^{\text{SA}} = \frac{N_{\text{tx}}}{n_{\text{SA}}} = \frac{N_{\text{tx}}}{N_{\text{RF}}}$. To limit the complexity of the architecture, each RF chain can connect to a maximum of L_{max} consecutive subarrays. In this case, the RF constraint set comprises matrices where each column has a maximum of L_{max} blocks of $N_{\text{tx}}^{\text{SA}}$ constant modulus elements, with all the remaining elements being zero. Defining $\mathbf{X} = \mathbf{F}_{\text{RF}}^{(t+1)} + \mathbf{W}^{(t)}$, the projection

can be implemented by setting all the elements in \mathbf{X} as 0 except for the subblocks in each column j which fulfill

$$\left\| \mathbf{X}_{\{(j-1)N_{tx}^{SA} + (i-1)N_{tx}^{SA} + 1 : (j-1)N_{tx}^{SA} + i \cdot N_{tx}^{SA} - 1\} \bmod N_{tx} + 1, j\}} \right\|_1 > \frac{N_{tx}^{SA}}{2} \quad (47)$$

with $i = 1, \dots, L_{\max}$ and $j = 1, \dots, N_{RF}$. In this case, the corresponding elements of \mathbf{R} are set as $R_{i,j}^{(t+1)} = (X_{i,j}) / |X_{i,j}|$, assuming UPS in these connections. Clearly, the phase shifters can be replaced by any of the other alternatives presented previously.

(8) Dynamic Array-of-Subarrays (DAoSAs)

As a variation of the previous AoSA architecture, we also consider an implementation where each subarray can be connected to a maximum of L_{\max} RF chains (which can be non-adjacent). In this case, the constraint set comprises matrices where each $N_{tx}^{SA} \times N_{RF}$ component submatrix contains a maximum of L_{\max} columns with constant modulus elements. The rest of the matrix contains only zeros. In this case, starting with $\mathbf{X} = 0$, the projection can be obtained by selecting the L_{\max} columns of

$$\left\| \mathbf{X}_{\{(j-1)N_{tx}^{SA} + (i-1)N_{tx}^{SA} + 1 : (j-1)N_{tx}^{SA} + i \cdot N_{tx}^{SA} - 1\} \bmod N_{tx} + 1, j\}} \right\|_1 > \frac{N_{tx}^{SA}}{2} \quad (48)$$

where $j = 1, \dots, n_{SA}$ with the largest ℓ_1 -norm and setting the corresponding elements of \mathbf{R} as $R_{i,j}^{(t+1)} = (X_{i,j}) / |X_{i,j}|$, assuming the use of UPS. Care must be taken to guarantee that at least one sub-block will be active in every column of \mathbf{R} . Similarly to the AoSA, the phase shifters can be replaced by any of the other presented alternatives.

3.3. Complexity

In the proposed algorithm, the $\mathbf{F}_{RF}^{(t+1)}$ and $\mathbf{F}_{BB}^{(t+1)}$ updates (steps 3 and 4 in Table 1) are defined using closed-form expressions that encompass several matrix multiplications, sums and an $N_{RF} \times N_{RF}$ matrix inverse (with an assumed complexity order of $\mathcal{O}(N_{RF}^3)$). These steps require a complexity order of $\mathcal{O}(QN_u N_s N_{RF} N_{tx} + F^{-1} Q N_{RF}^2 N_{tx})$ and $\mathcal{O}(QN_u N_s N_{RF} N_{tx} + Q N_{RF}^2 N_{tx})$, respectively.

The $\mathbf{R}^{(t+1)}$ update (step 5) involves simple element-wise division (assuming UPS) with $\mathcal{O}(Q N_{RF} N_{tx})$ while variable $\mathbf{B}_k^{(t+1)}$ (step 6) comprises a Frobenius norm computation with $\mathcal{O}(Q N_u N_s N_{RF} N_{tx})$. Step 7, the $\mathbf{F}_{\text{aprox}_{k,u}}^{(t+1)}$ update, has a complexity order of $\mathcal{O}(Q N_{tx}^2 N_u N_s + N_u^3 N_{tx} N_{rx}^2 + N_u^4 N_{rx}^3)$, whereas the dual variables updates (steps 8–10) have a complexity of $\mathcal{O}(Q N_u N_s N_{RF} N_{tx} + Q N_{RF}^2 N_{tx})$. Therefore, keeping only the dominant terms, the overall complexity order for the proposed algorithm is $\mathcal{O}(Q(N_{tx}^2 N_u N_s + N_{RF}^2 N_{tx}) + N_u^3 N_{tx} N_{rx}^2 + N_u^4 N_{rx}^3)$. Table 2 presents the total complexity order of the proposed method and compares it against other existing low-complexity alternatives, namely AM-based [15], LASSO-based Alt-Min (SPS and DPS) [14] and element-by-element (EBE) [20]

algorithms. Taking into account that in UM-MIMO, N_{tx} will tend to be very large, this means the algorithms with higher complexity, and the one proposed in this paper, will typically be EBE due to the terms $\mathcal{O}(QN_{tx}^2)$ and $\mathcal{O}(QN_{tx}^2 N_u N_s)$. It is important to note, however, that while the computational complexity of these two design methods may be higher, both algorithms can be applied to simple AoSA/DAoSA architectures. In particular, the proposed approach directly supports structures with lower practical implementation complexity (and are more energy-efficient) such as those based on switches. Furthermore, in a single-user scenario, the interference cancellation step of the proposed algorithm is unnecessary, and the complexity reduces to $\mathcal{O}(Q(N_u N_s N_{RF} N_{tx} + N_{RF}^2 N_{tx}))$. Regarding the other algorithms, they have similar complexities. However, the AM-based algorithm is designed for single stream scenarios whereas the others consider multiuser multi-stream scenarios.

Table 2. Overall Complexity of Different Hybrid Precoding Algorithms (per subcarrier).

AM—Based	
Operation	Complexity Order
Overall [15]	$\mathcal{O}(Q(N_u N_s N_{RF} N_{tx} + N_{RF}^2 N_u N_s + F^{-1} N_{RF}^3) + F^{-1} N_{RF}^2 N_{tx} + N_u^3 N_s^3)$
LASSO—Based Alt-Min (SPS)	
Operation	Complexity Order
Overall [14]	$\mathcal{O}(Q(N_u N_s N_{RF} N_{tx} + N_{RF}^2 N_u N_s + F^{-1} N_{RF}^3) + N_u^2 N_s N_{RF} N_{tx} + N_u^4 N_s^3)$
ADMM	
Operation	Complexity Order
Overall [10]	$\mathcal{O}(Q(N_s N_{RF} N_{tx} + N_{RF}^2 N_{tx}))$
EBE	
Operation	Complexity Order
Overall [20]	$\mathcal{O}(QN_{tx}^2)$
Proposed	
Operation	Complexity Order
\mathbf{F}_{RF}	$\mathcal{O}(QN_u N_s N_{RF} N_{tx} + F^{-1} QN_{RF}^2 N_{tx})$
\mathbf{F}_{BB}	$\mathcal{O}(QN_u N_s N_{RF} N_{tx} + QN_{RF}^2 N_{tx})$
\mathbf{R}	$\mathcal{O}(QN_{RF} N_{tx})$
\mathbf{B}	$\mathcal{O}(QN_u N_s N_{RF} N_{tx})$
$\mathbf{F}_{\text{aprox}}$	$\mathcal{O}(QN_{tx}^2 N_u N_s + N_u^3 N_{tx} N_{rx}^2 + N_u^4 N_{rx}^3)$
$\mathbf{U}, \mathbf{W}, \mathbf{Z}$	$\mathcal{O}(QN_u N_s N_{RF} N_{tx} + QN_{RF}^2 N_{tx})$
Overall	$\mathcal{O}(Q(N_{tx}^2 N_u N_s + N_{RF}^2 N_{tx}) + N_u^3 N_{tx} N_{rx}^2 + N_u^4 N_{rx}^3)$

4. Numerical Results

In this section, the performance of the proposed algorithm will be evaluated and compared against other existing alternatives from the literature, considering multiuser MIMO systems. We consider that both the transmitter and receivers are equipped with uniform planar arrays (UPAs) with $\sqrt{N_{tx}} \times \sqrt{N_{tx}}$ antenna elements at the transmitter and $\sqrt{N_{rx}} \times \sqrt{N_{rx}}$ at the receiver. The respective array response vectors are given by

$$\mathbf{a}_{t/r}(\phi_{i,l,u}^{t/r}, \theta_{i,l,u}^{t/r}) = \frac{1}{\sqrt{N_{tx/rx}}} \times \left[1, \dots, e^{j\frac{2\pi}{\lambda}d(p \sin \phi_{i,l,u}^{t/r} \sin \theta_{i,l,u}^{t/r} + q \cos \theta_{i,l,u}^{t/r})}, \dots, e^{j\frac{2\pi}{\lambda}d((\sqrt{N_{tx/rx}}-1) \sin \phi_{i,l,u}^{t/r} \sin \theta_{i,l,u}^{t/r} + (\sqrt{N_{tx/rx}}-1) \cos \theta_{i,l,u}^{t/r})} \right]^T, \quad (49)$$

where $p, q = 0, \dots, \sqrt{N_{tx/rx}} - 1$ are the antenna indices, λ is the signal wavelength and d is the inter-element spacing, which we assume to be $d = \lambda/2$. We consider a sparse channel with limited scattering where $N_{ray} = 4$ and $N_{cl} = 6$. The angles of departure and arrival were selected according to a Gaussian distribution whose means are uniformly distributed in $[0, 2\pi]$ and whose angular spreads are 10 degrees. While we include a few results for a NLOS channel, which is often considered a possible scenario in mmWave communications [14,15,25,31], we also present results for a channel with a LOS component which is more realistic, especially in the THz band. In the scenarios with a LOS component, a ratio of $E[|\alpha_u^{LOS}|^2] / \sum_{i=1}^{N_{cl}} \sum_{l=1}^{N_{ray}} E[|\alpha_{i,l,u}|^2] = 10$ is assumed (in this case we are admitting very weak NLOS paths compared to LOS which is typical in the THz band [27]). A fully-digital combiner was considered at each receiver and all simulation results were computed with 5000 independent Monte Carlo runs.

4.1. Fully-Connected Structures

First, we evaluate the performance assuming a fully-connected structure. Simulation results for a scenario where a base station with $N_{tx} = 100$ antennas transmits a single data stream ($N_s = 1$) to $N_u = 4$ users with $N_{rx} = 4$ antennas are shown in Figure 4 for $F = 1$ and Figure 5 for $F = 64$. The number of RF chains in the transmitter (N_{RF}^{tx}) is equal to $N_u N_s$. Besides our proposed precoder, several alternative precoding schemes are compared against the fully-digital solution, namely the LASSO-Based Alt-Min, the AM-Based and ADMM-Based precoding [10,14,15].

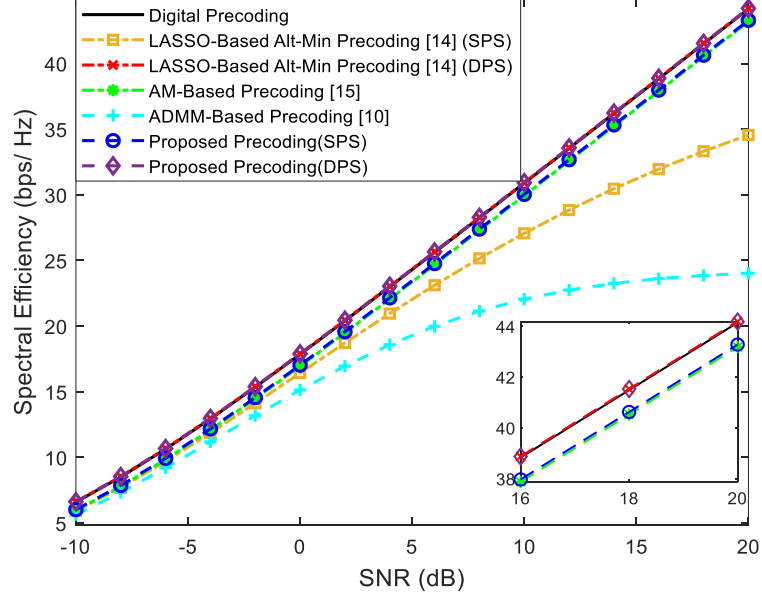


Figure 4. Spectral efficiency versus SNR achieved by different methods with $N_u = 4$, $N_s = 1$, $N_{RF}^{tx} = 4$, $F = 1$, $N_{tx} = 100$ and $N_{rx} = 4$ (only NLOS).

It can be observed that when $F = 1$, only the LASSO-Based Alt-Min with SPS and the ADMM-Based precoder from [10] (which does not remove the inter-user interference) lie far from fully-digital precoder. All the others achieve near optimum results and, in fact, can even match them when adopting DPS (proposed approach and LASSO-based Alt-Min). As explained in Section II, whereas for $F = 1$ we have \mathbf{F}_{BB} and \mathbf{F}_{RF} designed for that specific carrier, when $F = 64$, \mathbf{F}_{RF} has to be common to all subcarriers. While this reduces the implementation complexity, it also results in a more demanding restriction that makes the approximation of \mathbf{F}_{opt_k} (Problems (5)–(7)) to become worse. Additionally, when this approximation worsens, there can also be increased interference between users. Therefore, it can be observed in the results of Figure 5, that the gap between the fully-digital precoder and all the different hybrid algorithms is substantially wider. Still, the proposed precoder manages to achieve the best results.

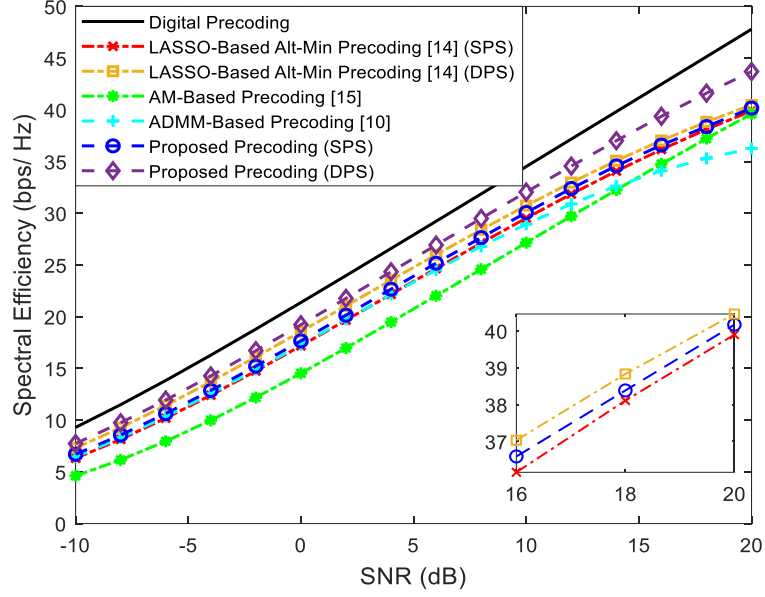


Figure 5. Spectral efficiency versus SNR achieved by different methods with $N_u = 4$, $N_s = 1$, $N_{RF}^{tx} = 8$, $F = 64$, $N_{tx} = 100$ and $N_{rx} = 4$ (only NLOS).

Given the performances of the different approaches, it is important to keep in mind that the AM-based precoding algorithm has the lowest performance in wideband but also one of the lowest computational complexities (see Table 2 of Section 3.3). In general, the proposed precoding algorithm is the one that can achieve better results at the cost of some additional computational complexity. Later on, we will address strategies based on lower complexity architectures that will allow for reducing the power consumption associated with its complexity.

In Figure 6, we consider a scenario where the BS employs a larger array with $N_{tx} = 256$ antennas to transmit $N_s = 2$ simultaneous streams to each user, where $N_u = 2$. To better fit this scenario to a typical communication in the THz band, we consider the existence of a LOS component, a center frequency of $f_c = 300\text{GHz}$ and a bandwidth of $B = 15\text{GHz}$ (it is important to note that the beam split effect is also considered in the channel model). The AM precoder from [15] requires a single stream per user and thus was not included in the figure. It is important to note that in this scenario the use of $F = 64$ with only 4 RF chains results in a more demanding restriction that makes the approximation to $\mathbf{F}_{\text{opt}_k}$ (Equation (5)) more difficult, thus widening the gap between all the schemes and the fully-digital curve. Still, the LASSO-based Alt-Min precoding schemes present a performance substantially lower when compared to the proposed approaches. Furthermore, the best performance is achieved with the use of double phase shifters, as expected. Once again, in comparing the curves of the proposed precoder against the ADMM-based precoder from [10], the advantage of adopting an interference-cancellation-based design over a simple matrix approximation one is clear.

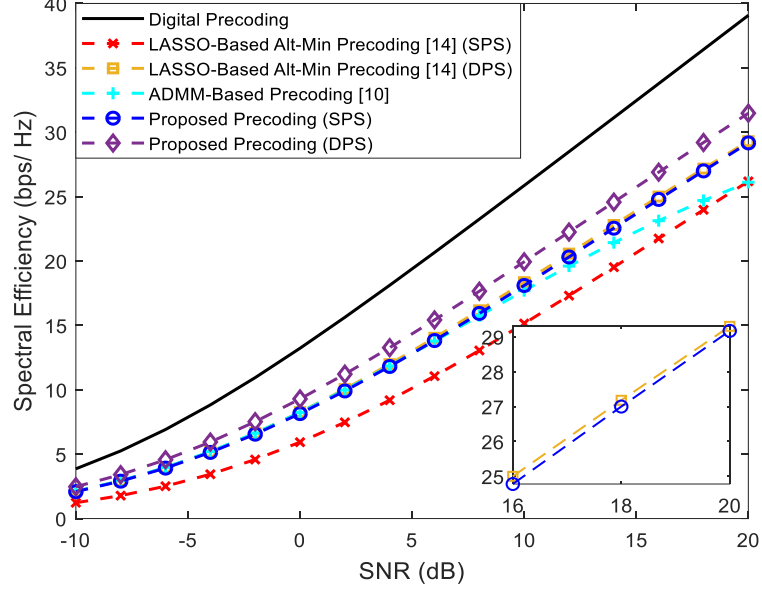


Figure 6. Spectral efficiency versus SNR achieved by different methods with $N_u = 2$, $N_s = 2$, $N_{RF}^{tx} = 4$, $F = 64$, $N_{tx} = 256$ and $N_{rx} = 4$ (with LOS component).

4.2. Reduced Complexity Architectures

Next, we will focus on the adoption of different reduced complexity architectures according to the typologies presented in Section 3.2. The objective is to evaluate the performance degradation when simpler architectures are adopted.

Figure 7 considers a scenario in which we have more than one data stream ($N_s = 2$) being sent from the BS to each user ($N_u = 4$) in a system with $N_{RF}^{tx} = N_u N_s$, $F = 1$, $N_{tx} = 256$ and $N_{rx} = 4$. We considered the same penalty parameters configuration: $\rho = 0.05$, $\mu = 1$ and $\eta = \rho$. This figure is placed in a perspective of simplifying the implementation of the analog precoder but keeping a fully-connected structure. We can observe that the versions based on DPS and single UPS achieve the best results, as expected. Considering the more realistic QPS versions, the results can worsen but it is visible that it is not necessary to use high resolution phase shifters since, with only 3 bits resolution, the results are already very close to the UPS curve. It can also be observed that the simplest of the architectures, AS, results in the worst performance but the SE improves when the antenna selectors are replaced by a network of switches, or, even better, if branches with inverters are also included.

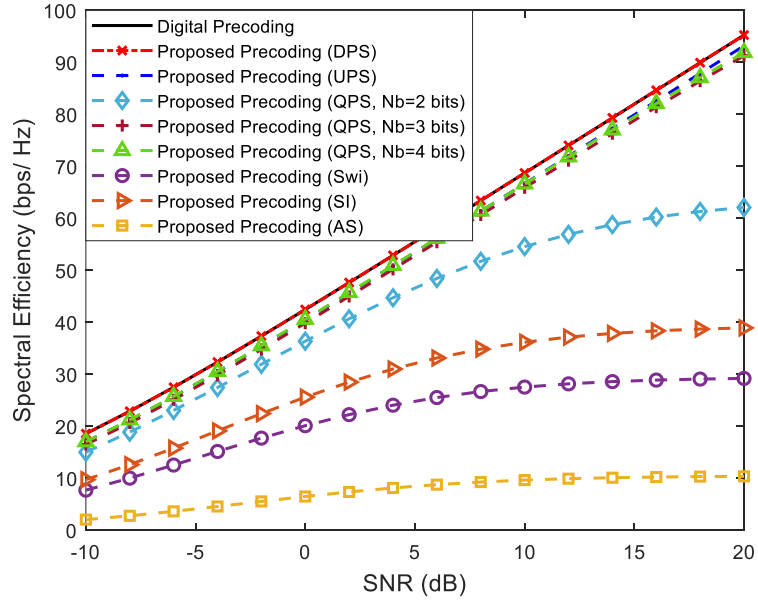


Figure 7. Spectral efficiency versus SNR achieved by the proposed precoder using different fully-connected architectures for $N_u = 4$, $N_s = 2$, $N_{RF}^{tx} = 8$, $F = 1$, $N_{tx} = 256$ and $N_{rx} = 4$ (only NLOS).

In Figure 8, we intend to simplify the implementation even further with the adoption of AoSAs. In this case we considered that the maximum number of subarrays that can be connected to an RF chain (L_{max}) is only one. This imposes a very demanding restriction on matrix \mathbf{F}_{RF} since most of it will be filled with zeros, thus substantially deteriorating the approximation to \mathbf{F}_{opt_k} (Equation (5)). The scenario is the same in Figure 7 but considers the existence of a LOS component with a few weak NLOS paths.

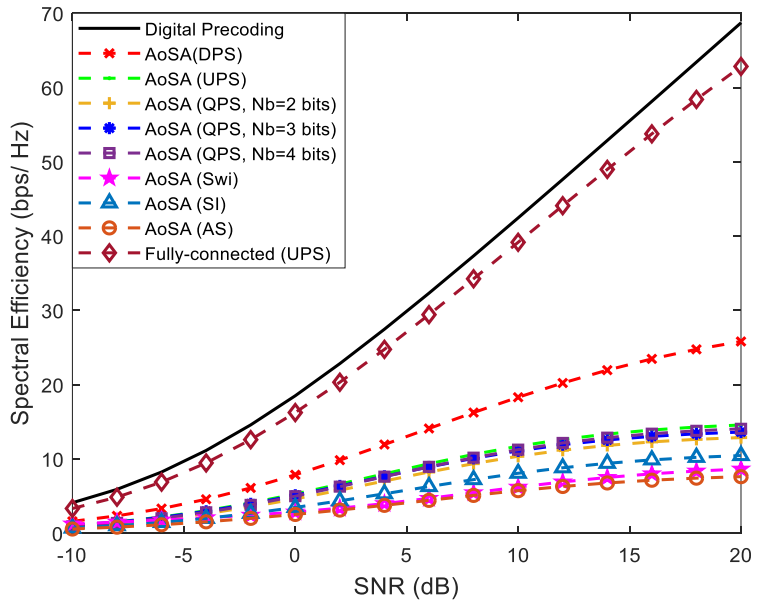


Figure 8. Spectral efficiency versus SNR achieved by the proposed precoder using different AoSA architectures with $L_{max} = 1$, $N_u = 4$, $N_s = 2$, $N_{RF}^{tx} = 8$, $F = 1$, $N_{tx} = 256$ and $N_{rx} = 4$ (with LOS component).

In fact, hereafter, the existence of an LOS component is assumed for the remaining figures of the paper in order to fit the AoSA/DAoSA results to a more typical scenario in the THz band. We can observe that for AoSA structures, the degradation of the SE is notorious, since all candidate versions present worse results when compared to the corresponding fully-connected design and are all far from the fully-digital solution. To reduce the large performance loss due to the adoption of a simple AoSA architecture, we can allow the dynamic connection of more subarrays to each RF chain by adopting a DAoSA structure, as introduced in Section 3.2.

In Figure 9, we study the effect of increasing the maximum number of subarrays that can be connected to an RF chain (L_{max}) in the performance of these schemes. Each subarray has a size of 32 antennas (n_i). Curves assuming the use of SPS as well as of DPS are included. It can be observed that the increase in the number of connections to subarrays, L_{max} , has a dramatic effect on the performance, resulting in a huge improvement by simply going from $L_{max} = 1$ to $L_{max} = 2$. Increasing further to $L_{max} = 4$, the results become close to the fully-connected case, showing that the DAoSA can be a very appealing approach for balancing spectral efficiency with hardware complexity and power consumption. Combining the increase in L_{max} with the adoption of DPS can also improve the results but the gains become less pronounced for $L_{max} > 1$. It is important to note that the penalty parameters can be fine-tuned for different system configurations.

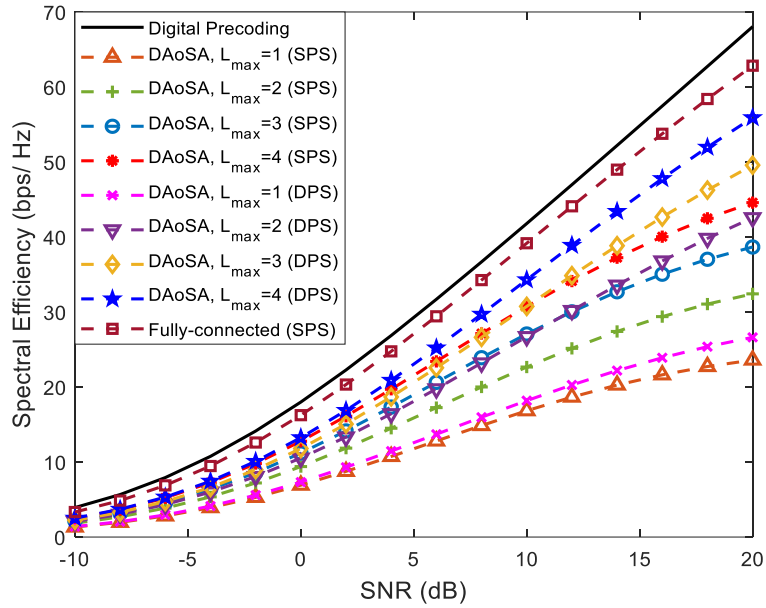


Figure 9. Spectral efficiency versus SNR achieved by the proposed precoder considering an architecture based on DAoSAs and the variation of the maximum number of subarrays that can be connected to an RF chain (L_{max}) for $N_u = 4$, $N_s = 2$, $N_{RF}^{tx} = 8$, $F = 1$, $N_{tx} = 256$ and $N_{rx} = 4$ (with LOS component).

One of the objectives of adopting these low-complexity solutions is to reduce overall power consumption. Based on [20], we can calculate the total power consumption of each precoding scheme using

$$P_C = P_{BB} N_{BB} + (P_{DAC} + P_{OS} + P_M) N_{RF}^{tx} + P_{PA} N_{tx} + P_{PC} N_{tx} + P_{PS} N_{PS} + P_{SWI} N_{SWI} + P_{tx}, \quad (50)$$

where PBB is the power of the baseband block (with NBB = 1), PDAC is the power of a DAC, POS is the power of an oscillator, PM is the power of a mixer, PPA is the power of a power amplifier, PPC is the power of a power combiner, PPS is the power of a phase shifter, PSWI is the power of a switch and P_{tx} denotes the transmit power. The N_x variable represents the number of elements of each device used in the precoder configuration.

Based on the values provided in [20,38] for the power consumption of individual devices in the 300 GHz band, we adopt the following values: P_{BB} = 200 mW, P_{DAC} = 110 mW, P_{OS} = 4 mW, P_M = 22 mW, P_{PA} = 60 mW, P_{PC} = 6.6 mW, P_{SWI} = 24 mW and P_T = 100 mW.

Regarding the phase shifters, we assume values of P_{PS} = 10, 20, 40 and 100 mW for 1, 2, 3 and 4 quantization bits. Considering the same configuration scenario as Figures 7–9 with N_u = 4, N_s = 2, N_{RF}^{tx} = N_uN_s, F = 1 and N_{tx} = 256, we provide the values of power consumption for different precoder configurations in Table 3.

Table 3. Power Consumption for Different Implementations of the Proposed Precoder for N_u = 4, N_s = 2, N_{RF}^{tx} = 8, F = 1 and N_{tx} = 256.

Precoder		Estimated Power Consumption
Fully-Connected	DPS	428.04
	UPS	223.24
	QPS (N _b = 2)	59.4
	QPS (N _b = 3)	100.36
	SWI	67.59
	SI	38.92
DAoSA SPS	L _{max} = 1	28.87
	L _{max} = 2	39.30
	L _{max} = 3	49.73
	L _{max} = 4	60.17
DAoSA DPS	L _{max} = 1	39.11
	L _{max} = 2	59.78
	L _{max} = 3	80.45
	L _{max} = 4	101.13

For the fully-connected structure with UPS, we assumed that P_{PS} = 100 mW, which corresponds to quantized phase shifters with N_b = 4 bits [48]. For the remaining phase-shifter-based precoder structures, we assumed that P_{PS} = 40 mW, which corresponds to quantized phase shifters with N_b = 3 bits, since with only 3 bits resolution the results are already very close to the UPS curve (see Figure 7). As can be seen from this table, the use of architectures based on DAoSAs allows us to reduce considerably the amount of power that is consumed by the precoder. In fact, we can reduce the amount of consumed power up to 55% if we consider a precoder scheme based on DAoSA with DPS and L_{max} = 4 versus an FC structure precoder based on UPS, with only a small performance penalty (Figure 9). This reduction increases to 73% if the DPS structure is replaced by an SPS one.

In the particular case of architectures based on quantized phase shifters, we observed that by decreasing the number of quantization bits, it is possible to substantially reduce the power consumption without excessively compromising the complexity (as seen in Figure 7). This conclusion is corroborated by [20,38], since the architectures based on low

resolution QPS, AoSAs and DAoSAs present superior energy efficiency when compared to the fully-connected structure with UPS.

In Figures 10 and 11, we provide a comparison between our proposed precoder and the EBE precoder from [20], considering an architecture based on DAoSAs (with SPS) and a scenario configuration similar to Figure 9, i.e., with $N_s = 2$, $N_{RF}^{tx} = 8$, $F = 1$, $N_{tx} = 256$ and $N_{rx} = 4$. These figures present various curves where the maximum number of subarrays that can be connected to an RF chain, L_{max} , is changed. Figure 10 refers to an SU scenario ($N_u = 1$) whereas Figure 11 corresponds to an MU scenario with $N_u = 4$. In the SU case, the proposed precoder achieves results very close to the fully-digital precoder, even with only $L_{max} = 2$. Compared to the proposed algorithm, EBE shows a wider gap even though it has a smaller complexity (as presented in Table 2 of Section 3.3).

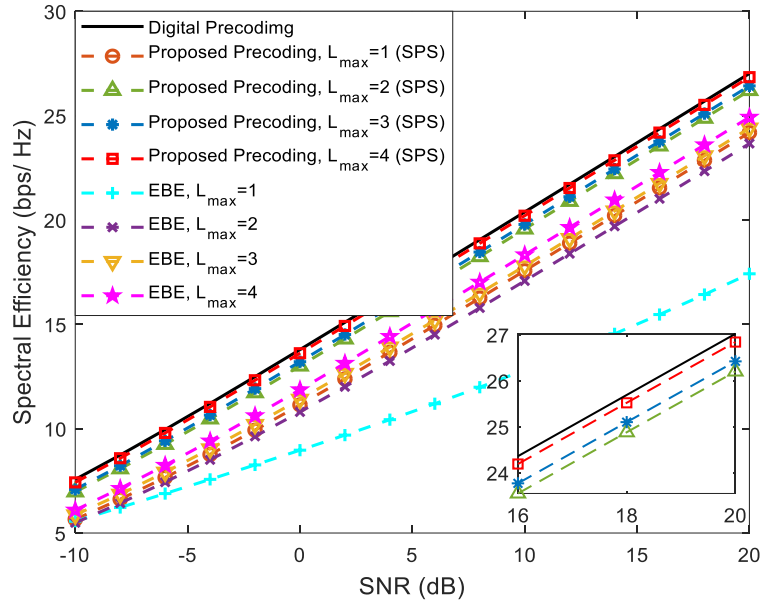


Figure 10. Spectral efficiency versus SNR achieved by the proposed precoder and by the EBE algorithm considering an architecture based on DAoSAs and the variation of the maximum number of subarrays that can be connected to an RF chain (L_{max}) for $N_u = 1$, $N_s = 2$, $N_{RF}^{tx} = 8$, $F = 1$, $N_{tx} = 256$ and $N_{rx} = 4$ (with LOS component).

When we increase the number of users from $N_u = 1$ to $N_u = 4$, we can clearly observe that the EBE algorithm suffers a substantial degradation compared to the proposed solution which can be explained by the lack of inter-user interference cancellation (it was not specifically designed for MU scenarios).

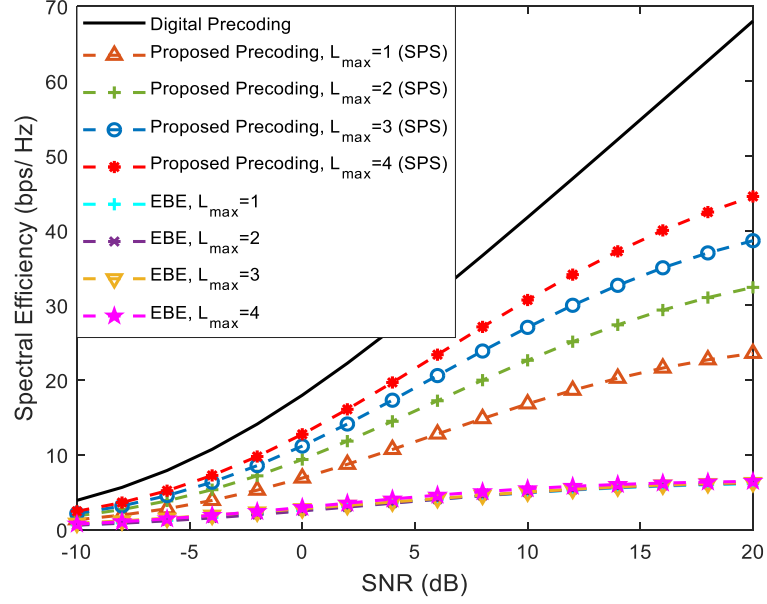


Figure 11. Spectral efficiency versus SNR achieved by the proposed precoder and by the EBE algorithm considering an architecture based on DAoSAs and the variation of the maximum number of subarrays that can be connected to an RF chain (L_{\max}) for a mmWave/THz system with $N_u = 4$, $N_s = 2$, $N_{RF}^{tx} = 8$, $F = 1$, $N_{tx} = 256$ and $N_{rx} = 4$ (with LOS component).

Even though a sub-6 GHz system often adopts fully-digital processing [39], where each antenna element has a dedicated RF chain, it is possible to apply the proposed hybrid design algorithm to a sub-6 GHz channel since it is independent of a specific MIMO channel (as are the other alternative algorithms that we used as benchmarks and which are targeted at solving the matrix approximation problem). To exemplify, Figure 12 presents the simulated results obtained for the same scenario of Figure 4 but considering an ideal uncorrelated channel which approximates a rich scattering environment that is typical in sub-6 GHz bands. It can be observed that the proposed approach displays similar behavior to the ones in the upper-bands channel, showing that it can also be used for this particular type of channel (even though it may require a higher number of RF chains to achieve a good approximation of the fully-digital solution in some scenarios, due to the channel not being sparse, as noted in [40]). It is important to highlight that even though the proposed approach can be applied to other channels, the algorithm was designed with the aim of dealing with architectures with a very large number of antennas and with large hardware constraints, making it especially interesting for mmWave and THz.

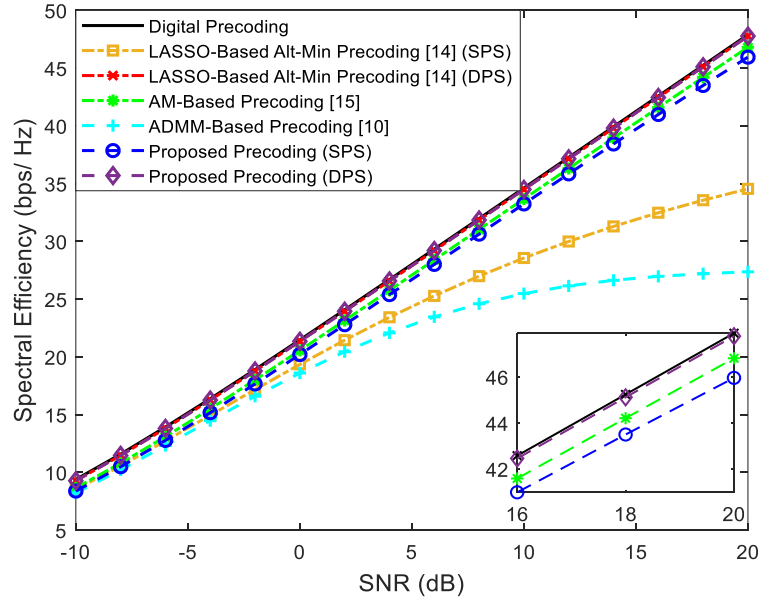


Figure 12. Spectral efficiency versus SNR achieved by different methods for a mmWave/THz MIMO-OFDM system with $N_u = 4$, $N_s = 1$, $N_{RF}^{tx} = 4$, $F = 1$, $N_{tx} = 100$ and $N_{rx} = 4$ considering an uncorrelated channel.

While we have shown how the proposed approach can deal with several relevant types of analog precoders/combiners, it is important to note that there are other alternative structures that have been recently proposed in the literature. For example, some authors have considered precoding paradigms based on time-delayer structures for THz systems [27,41]. One of the most notorious is the Delay Phase Precoding (DPD), which consists in the use of a Time-Delay (TD) network between the RF chains and the traditional phase shifters network in order to convert phase-controlled analog precoding into delay-phase-controlled analog precoding. The main advantage related to this type of precoding is that the time delays in the TD network are carefully designed to generate frequency-dependent beams which are aligned with the spatial directions over the whole bandwidth [41]. While we do not address the adoption of time-delay structures in this paper, it should be possible to derive a projection algorithm that simultaneously takes into account the constraints imposed in both analog-precoding steps: time-delay networks and frequency-independent phase shifters.

5. Conclusions

In this paper, we proposed an iterative algorithm for hybrid precoding design which is suitable for multiuser UM-MIMO systems operating in mmWave and THz bands. The adopted approach replaces unfeasible fully-digital precoders/combiners relying on a dedicated RF chain per antenna with a hybrid architecture comprising low dimension digital blocks with only a few RF chains which are complemented by analog blocks supported solely on networks of phase shifters and switches. The proposed hybrid design algorithm is based on the approximation of the fully-digital approach whose problem formulation is split into a sequence of smaller subproblems with closed-form solutions and can work with a broad range of configuration of antennas, RF chains and data streams. The separability of the design process allows for the adaptability of the algorithm to different architectures, making it suitable for implementation with low-complexity AoSA and DAoSA structures, which are particularly relevant for the deployment of UM-MIMO in hardware-constrained THz systems. It was shown that good trade-offs between SE and hardware implementation complexity can in fact be achieved by the proposed algorithm for several different architectures. Numerical results showed that the use of architectures based on DAoSAs allows us to reduce considerably the amount of power that is consumed at the precoder. In fact, in a reference scenario, we showed that it was possible to reduce the amount of consumed power up to 55% if we consider a precoder scheme based on DAoSAs with DPS and $L_{\max} = 4$ versus an FC structure based on UPS, with only a small performance penalty. This reduction increases to 73% if the DPS structure is replaced by an SPS one.

Author Contributions: Conceptualization, J.P.P. and N.S.; methodology, J.P.P., N.S. and R.D.; software, J.P.P. and N.S.; validation, J.P.P., N.S., M.R. and R.D.; formal analysis, J.P.P., V.V., R.F. and N.S.; investigation, J.P.P., V.V., R.F. and N.S.; resources, N.S. and M.R.; data curation, J.P.P.; writing—original draft preparation, J.P.P. and N.S.; writing—review and editing, J.P.P., V.V., R.F., N.S., M.R., J.S. and R.D.; visualization, J.P.P. and N.S.; supervision, N.S. and M.R.; project administration, N.S., M.R., J.S. and R.D.; funding acquisition, N.S. and R.D. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the FCT—Fundação para a Ciência e Tecnologia under the grant 2020.05621.BD. The authors also acknowledge the funding provided by FCT/MCTES through national funds and when applicable co-funded EU funds under the project UIDB/50008/2020.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rappaport, T.S.; Sun, S.; Mayzus, R.; Zhao, H.; Azar, Y.; Wang, K.; Wong, G.; Schulz, J.K.; Samimi, M.; Gutierrez, F. Millimeter Wave Mobile Communications for 5G Cellular: It Will Work! *IEEE Access* **2013**, *1*, 335–349, <https://doi.org/10.1109/access.2013.2260813>.

2. Uwaechia, A.N.; Mahyuddin, N.M. A Comprehensive Survey on Millimeter Wave Communications for Fifth-Generation Wireless Networks: Feasibility and Challenges. *IEEE Access* **2020**, *8*, 62367–62414, <https://doi.org/10.1109/access.2020.2984204>.
3. Akyildiz, I.F.; Kak, A.; Nie, S. 6G and Beyond: The Future of Wireless Communications Systems. *IEEE Access* **2020**, *8*, 133995–134030, <https://doi.org/10.1109/access.2020.3010896>.
4. Tan, J.; Dai, T. THz Precoding for 6G: Applications, Challenges, Solutions, and Opportunities. *arXiv Prepr.* **2020**, arXiv:2005.10752. Available online: <https://arxiv.org/abs/2005.10752> (accessed on 2 July 2021).
5. Sameddeen, H.; Alouini, M.-S.; Al-Naffouri, T.Y. An Overview of Signal Processing Techniques for Terahertz Communications. *Proc. IEEE* **2021**, *PP*, 1–38, <https://doi.org/10.1109/jproc.2021.3100811>.
6. Lin, C.; Li, G.Y.L. Terahertz Communications: An Array-of-Subarrays Solution. *IEEE Commun. Mag.* **2016**, *54*, 124–131, <https://doi.org/10.1109/mcom.2016.1600306cm>.
7. Ahmed, I.; Khammari, H.; Shahid, A.; Musa, A.; Kim, K.S.; De Poorter, E.; Moerman, I. A Survey on Hybrid Beamforming Techniques in 5G: Architecture and System Model Perspectives. *IEEE Commun. Surv. Tutorials* **2018**, *20*, 3060–3097, <https://doi.org/10.1109/comst.2018.2843719>.
8. Chataut, R.; Akl, R. Massive MIMO Systems for 5G and beyond Networks—Overview, Recent Trends, Challenges, and Future Research Direction. *Sensors* **2020**, *20*, 1–35.
9. Sohrabi, F.; Yu, W. Hybrid Digital and Analog Beamforming Design for Large-Scale Antenna Arrays. *IEEE J. Sel. Top. Signal Process.* **2016**, *10*, 501–513, <https://doi.org/10.1109/jstsp.2016.2520912>.
10. Souto, N.N.; Silva, J.; Pavia, J.; Ribeiro, M. An alternating direction algorithm for hybrid precoding and combining in millimeter wave MIMO systems. *Phys. Commun.* **2019**, *34*, 165–173.
11. Pavia, J.P.; Souto, N.; Ribeiro, M.; Silva, J.; Dinis, R. Hybrid Precoding and Combining Algorithm for Reduced Complexity and Power Consumption Architectures in mmWave Communications. In Proceedings of the 2020 IEEE 91st Vehicular Technology Conference: VTC2020-Spring, Antwerp, Belgium, 25–28 May 2020; pp. 1–5, <https://doi.org/10.1109/vtc2020-spring48590.2020.9128553>.
12. Guan, K.; Li, G.; Kuerner, T.; Molisch, A.F.; Peng, B.; He, R.; Hui, B.; Kim, J.; Zhong, Z. On Millimeter Wave and THz Mobile Radio Channel for Smart Rail Mobility. *IEEE Trans. Veh. Technol.* **2016**, *66*, 5658–5674, <https://doi.org/10.1109/tvt.2016.2624504>.
13. Alkhateeb, A.; Mo, J.; Gonzalez-Prelcic, N.; Heath, R. MIMO Precoding and Combining Solutions for Millimeter-Wave Systems. *IEEE Commun. Mag.* **2014**, *52*, 122–131, <https://doi.org/10.1109/mcom.2014.6979963>.
14. Yu, X.; Zhang, J.; Letaief, K.B. Alternating minimization for hybrid precoding in multiuser OFDM mmWave systems. In Proceedings of the 2016 50th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 6–9 November 2016; pp. 281–285, <https://doi.org/10.1109/acssc.2016.7869042>.
15. Yuan, H.; An, J.; Yang, N.; Yang, K.; Duong, T.Q. Low Complexity Hybrid Precoding for Multiuser Millimeter Wave Systems Over Frequency Selective Channels. *IEEE Trans. Veh. Technol.* **2018**, *68*, 983–987, <https://doi.org/10.1109/tvt.2018.2880787>.
16. Liu, F.; Kan, X.; Bai, X.; Du, R.; Liu, H.; Zhang, Y. Hybrid precoding based on adaptive RF-chain-to-antenna connection for millimeter wave MIMO systems. *Phys. Commun.* **2020**, *39*, 100997, <https://doi.org/10.1016/j.phycom.2019.100997>.
17. Xu, K.; Cai, Y.; Zhao, M.; Niu, Y.; Hanzo, L. MIMO-Aided Nonlinear Hybrid Transceiver Design for Multiuser Mmwave Systems Relying on Tomlinson-Harashima Precoding. *IEEE Trans. Veh. Technol.* **2021**, *70*, 6943–6957, <https://doi.org/10.1109/tvt.2021.3087651>.
18. Vizziello, A.; Savazzi, P.; Chowdhury, K.R. A Kalman Based Hybrid Precoding for Multi-User Millimeter Wave MIMO Systems. *IEEE Access* **2018**, *6*, 55712–55722, <https://doi.org/10.1109/access.2018.2872738>.
19. Elmagzoub, H.M. On the MMSE-based multiuser millimeter wave MIMO hybrid precoding design. *Int. J. Commun. Syst.* **2020**, *33*, e4409, <https://doi.org/10.1002/dac.4409>.
20. Yan, L.; Han, C.; Yuan, J. A Dynamic Array-of-Subarrays Architecture and Hybrid Precoding Algorithms for Terahertz Wireless Communications. *IEEE J. Sel. Areas Commun.* **2020**, *38*, 2041–2056, <https://doi.org/10.1109/jsac.2020.3000876>.

21. Busari, S.; Huq, K.; Mumtaz, S.; Rodriguez, J. Terahertz Massive MIMO for Beyond-5G Wireless Communication. In Proceedings of the ICC 2019-2019 IEEE International Conference on Communications (ICC), Shanghai, China, 11–22 March 2019.
22. Boyd, S.; Parikh, N.; Chu, E.; Peleato, B.; Eckstein, J. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *Found. Trends Mach. Learn.* **2010**, *3*, 1–122, <https://doi.org/10.1561/22000000016>.
23. Wang, Y.; Yin, W.; Zeng, J. Global Convergence of ADMM in Nonconvex Nonsmooth Optimization. *J. Sci. Comput.* **2018**, *78*, 29–63, <https://doi.org/10.1007/s10915-018-0757-z>.
24. Alkhateeb, A.; Heath, R.W. Frequency Selective Hybrid Precoding for Limited Feedback Millimeter Wave Systems. *IEEE Trans. Commun.* **2016**, *64*, 1801–1818.
25. Yuan, H.; Yang, N.; Yang, K.; Han, C.; An, J. Hybrid Beamforming for Terahertz Multi-Carrier Systems Over Frequency Selective Fading. *IEEE Trans. Commun.* **2020**, *68*, 6186–6199, <https://doi.org/10.1109/tcomm.2020.3008699>.
26. Tarboush, S.; Sardeddeen, H.; Chen, H.; Loukil, M.H.; Jemaa, H.; Alouini, M.S.; Al-Naffouri, T.Y. TeraMIMO: A Channel Simulator for Wideband Ultra-Massive MIMO Terahertz Communications. *arXiv Prepr.* **2021**, arXiv:2104.11054. Available online: <https://arxiv.org/abs/2104.11054> (accessed on 2 July 2021).
27. Lin, C.; Li, G.Y.; Wang, L. Subarray-Based Coordinated Beamforming Training for mmWave and Sub-THz Communications. *IEEE J. Sel. Areas Commun.* **2017**, *35*, 2115–2126, <https://doi.org/10.1109/jsac.2017.2720038>.
28. Nguyen, D.D.; Le, L.; Le-Ngoc, T.; Heath, R. Hybrid MMSE Precoding and Combining Designs for mmWave Multiuser Systems. *IEEE Access* **2017**, *5*, 19167–19181.
29. Bertsekas, D. *Nonlinear Programming*; Athena Scientific: Belmont, MA, USA, 2016.
30. El Ayach, O.; Rajagopal, S.; Abu-Surra, S.; Pi, Z.; Jr, R.W.H. Spatially Sparse Precoding in Millimeter Wave MIMO Systems. *IEEE Trans. Wirel. Commun.* **2014**, *13*, 1499–1513, <https://doi.org/10.1109/twc.2014.011714.130846>.
31. Yu, X.; Shen, J.-C.; Zhang, J.; Letaief, K.B. Alternating Minimization Algorithms for Hybrid Precoding in Millimeter Wave MIMO Systems. *IEEE J. Sel. Top. Signal Process.* **2016**, *10*, 485–500, <https://doi.org/10.1109/jstsp.2016.2523903>.
32. Spencer, Q.; Swindlehurst, A.L.; Haardt, M. Zero-Forcing Methods for Downlink Spatial Multiplexing in Multiuser MIMO Channels. *IEEE Trans. Signal Process.* **2004**, *52*, 461–471, <https://doi.org/10.1109/tsp.2003.821107>.
33. Mendez-Rial, R.; Rusu, C.; Gonzalez-Prelcic, N.; Alkhateeb, A.; Heath, R. Hybrid MIMO Architectures for Millimeter Wave Communications: Phase Shifters or Switches? *IEEE Access* **2016**, *4*, 247–267.
34. Lee, J.; Lee, Y.H. AF relaying for millimeter wave communication systems with hybrid RF/baseband MIMO processing. In Proceedings of the 2014 IEEE International Conference on Communications (ICC), Sydney, Australia, 10–14 June 2014; pp. 5838–5842, <https://doi.org/10.1109/icc.2014.6884253>.
35. Payami, S.; Ghoraishi, M.; Dianati, M.; Sellathurai, M. Hybrid Beamforming with a Reduced Number of Phase Shifters for Massive MIMO Systems. *IEEE Trans. Veh. Technol.* **2018**, *67*, 4843–4851, <https://doi.org/10.1109/tvt.2018.2807921>.
36. Tian, M.; Zhang, J.; Zhao, Y.; Yuan, L.; Yang, J.; Gui, G. Switch and Inverter Based Hybrid Precoding Algorithm for mmWave Massive MIMO System: Analysis on Sum-Rate and Energy-Efficiency. *IEEE Access* **2019**, *7*, 49448–49455, <https://doi.org/10.1109/access.2019.2910094>.
37. Yu, X.; Zhang, J.; Letaief, K.B. Doubling Phase Shifters for Efficient Hybrid Precoder Design in Millimeter-Wave Communication Systems. *J. Commun. Inform. Netw.* **2019**, *4*, 51–67.
38. Yan, L.; Han, C.; Yang, N.; Yuan, J. Dynamic-subarray with Quantized- and Fixed-phase Shifters for Terahertz Hybrid Beamforming. In Proceedings of the GLOBECOM 2020-2020 IEEE Global Communications Conference, Taipei, Taiwan, 7–11 December 2020; pp. 1–6, <https://doi.org/10.1109/globecom42002.2020.9348113>.
39. Li, Z.; Zhang, C.; Lu, I.-T.; Jia, X. Hybrid Precoding Using Out-of-Band Spatial Information for Multi-User Multi-RF-Chain Millimeter Wave Systems. *IEEE Access* **2020**, *8*, 50872–50883, <https://doi.org/10.1109/access.2020.2979712>.

40. Park, S.; Alkhateeb, A.; Jr., R.W.H. Dynamic Subarrays for Hybrid Precoding in Wideband mmWave MIMO Systems. *IEEE Trans. Wirel. Commun.* **2017**, *16*, 2907–2920, <https://doi.org/10.1109/twc.2017.2671869>.
41. Tan, J.; Dai, L. Delay-Phase Precoding for THz Massive MIMO with Beam Split. In Proceedings of the 2019 IEEE Global Communications Conference (GLOBECOM), Big Island, Hawaii, USA, 9–13 December 2019; pp. 1–6, <https://doi.org/10.1109/globecom38437.2019.9014304>.

System-Level Assessment of Low Complexity Hybrid Precoding Designs for Massive MIMO Downlink Transmissions in Beyond 5G Networks

João Pedro Pavia ^{1,2,*}, Vasco Velez ^{1,2}, Nuno Souto ^{1,2}, Marco Ribeiro ^{1,2}, Pedro Sebastião ^{1,2} and Américo Correia ^{1,2}

¹ Department of Information Science and Technology, ISCTE-Instituto Universitário de Lisboa, 1649-026 Lisbon, Portugal; Vasco_Velez@iscte-iul.pt (V.V.); Nuno.Souto@iscte-iul.pt (N.S.); Marco.Ribeiro@iscte-iul.pt (M.R.); Pedro.Sebastiao@iscte-iul.pt (P.S.); Americo.Correia@iscte-iul.pt (A.C.)

² Instituto de Telecomunicações, 1049-001 Lisbon, Portugal

* Correspondence: jpavia@lx.it.pt (J.P.P.)

Citation: Pavia, J.P.; Velez, V.; Souto, N.; Ribeiro, M.; Sebastião, P.; Correia, A. System-Level Assessment of Low Complexity Hybrid Precoding Designs for Massive MIMO Downlink Transmissions in Beyond 5G Networks. *Appl. Sci.* **2022**, *12*, 2812. <https://doi.org/10.3390/app12062812>

Academic Editor: Mário Marques da Silva

Received: 3 February 2022

Accepted: 7 March 2022

Published: 9 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: The fast growth experienced by the telecommunications field along the last decades have been motivating the academy and the industry to invest in the design, testing and deployment of new evolutions of wireless communication systems. Terahertz (THz) communication represents one of the possible technologies to explore in order to achieve the desired achievable rates above 100Gbps and the extremely low latency required in many envisioned applications. Despite the potentialities, it requires proper system design, since working at the THz band brings a set of challenges such as the reflection and scattering losses through the transmission path, the high dependency with distance and the severe hardware constraints. One key approach for overcoming some of these challenges relies on the use of massive/ultramassive antenna arrays combined with hybrid precoders based on fully connected phase-shifters architectures or partially connected architectures, such as arrays of subarrays (AoSAs) or dynamic AoSAs (DAoSAs). Through this strategy, it is possible to obtain very high-performance gains while drastically simplifying the practical implementation and reducing the overall power consumption of the system when compared to a fully digital approach. Although this type of solutions have been previously proposed to address some of the limitations of mmWave/THz communications, a lack between link level and system level analysis is commonly verified. In this paper, we present a thorough system level assessment of a cloud radio access network (C-RAN) for beyond 5G (B5G) systems where the access points (APs) operate in the mmWave/THz bands, supporting multi-user MIMO (MU-MIMO) transmission with massive/ultra-massive antenna arrays combined with low-complexity hybrid precoding architectures. Results showed that the C-RAN deployments in two indoor office scenarios for the THz were capable of achieving good throughput and coverage performances, with only a small compromise in terms of gains when adopting reduced complexity hybrid precoders. Furthermore, we observed that the indoor mixed office scenario can provide higher throughput and coverage performances independently of the cluster size, when compared to the indoor open office scenario.

Keywords: Terahertz (THz); multiuser Ultra-Massive-MIMO; hybrid precoder design; system level simulation; cloud radio access network (C-RAN); Beyond 5G (B5G); 6G.

1. Introduction

In the recent years we have been witnessing the increasing deployment of the fifth generation of wireless communications (5G). This generation represents a significant mark in the way we communicate, enabling new applications that would be otherwise infeasible with the technology of previous generations. Therefore, several new technologies were incorporated into 5G which were crucial to achieve all the requirements that are needed for the operation of these systems. However, the telecommunications field is experiencing a fast growth along the last decades, which have been motivating the academy and the industry to invest in the design, testing and deployment of the next generation of wireless communications (6G). Within this context, terahertz (THz) communications have been attracting more and more attention, being referred by research community as one of the most promising research fields on the topic, not only for the availability of spectrum but also because of the achievable rates that can be offered to the users [1, 2].

THz systems are becoming feasible due to the recent advances in the field of THz devices, and they are expected to ease the spectrum limitations of today's systems [3]. However, there are several issues that can affect the system performance, such as the reflection and scattering losses through the transmission path, the high dependency between range and frequency of channels at the THz band and also the need for controllable time-delay phase shifters. Such limitations require not only the proper system design, but also the definition of a set of strategies to enable communications [4, 5]. According to the literature, some of the challenges for beyond 5G networks (B5G) are the fabrication of plasmonic nano array antennas, channel estimation, precoding, signal detection, beamforming, and beamsteering [6]. To overcome the distance limitation in THz communications, one can take advantage of the very large antenna arrays that can be implemented at these bands while minimizing interference between multiple users. However, when working with massive/ultra-massive arrays of antennas, a fully digital precoder is often not feasible and hybrid designs must be used instead. Regarding hybrid precoder, it is necessary to adopt the most adequate architecture in order to maintain the intended trade-off between performance and complexity. Fully-connected (FC) structures based on phase shifters, arrays of subarrays (AoSAs) and dynamic arrays of subarrays (DAoSAs) represent some of the most referred architectures in the literature [7-11]. From the energy efficiency (EE) and power consumption perspective which are particularly relevant and the mmWave/THz bands, the use of partially connected (PC) structures such as AoSAs and DAoSAs, can be more efficient than fully connected structures. In particular, DAoSAs represent a much more appealing solution, since they can offer a good compromise (in terms of performance) between fully connected structures which have a higher implementation complexity (especially with a massive number of antennas) and AoSAs which are lighter but can suffer significant degradation in performance. When concerning about this topic, one can

find on the literature a wide range of schemes for multiple input multiple output (MIMO) systems considering both uplink and downlink scenarios [7-11], but there is a significant imbalance between the number of approaches aimed at those scenarios for hybrid precoding at higher frequencies. The authors of [12] proposed two algorithms for low-complexity hybrid precoding and beamforming for multi-user (MU) mmWave systems. Even though they assume only one stream per user, i.e., the number of data streams (N_s) is equal to the number of users (N_u), it is shown that the algorithms achieve interesting results when compared to the fully-digital solution. The concept of precoding based on adaptive RF-chain-to-antenna was only introduced in [13] for single user (SU) scenarios but showed promising results. In [14], a nonlinear hybrid transceiver design relying on Tomlinson–Harashima precoding was proposed. Their approach only considers FC architectures but can achieve a performance close to the fully-digital transceiver. Most of the hybrid solutions for mmWave systems aim to achieve near-optimal performance using FC structures, resorting to phase shifters or switches. However, the difficulty of handling the hardware constraint imposed by the analog phase shifters or by switches in the THz band is an issue that limits the expected performance in terms of SE. In [15] the authors proposed a low complexity design based on the alternating direction method of multipliers (ADMM) that can approximate the performance of a hybrid precoder to the fully-digital performance. It is targeted to the millimeter wave (mmWave)/THz bands and can incorporate different architectures at the analog component of the precoder, making it suitable for supporting ultra-massive MIMO (UM-MIMO) in severely hardware-constrained systems that are typical at these bands. There are some recent experimental demonstrations of hybrid precoding schemes in for MIMO mmWave communications as mentioned in [16, 17]. However, these implementations still have limitations in the size of the adopted arrays and number of RF chains. Regarding the applications of these schemes in the THz band, the technology is still in early stages and there are still no UM-MIMO implementations with hybrid precoding/beamforming. Nevertheless, there are already some simpler SISO and MIMO implementations in THz as the ones mentioned in [18, 19].

When transitioning from a link level perspective to a system level analysis, we must deal with several other issues such as distance limitation, signal-to-noise ratio (SNR) degradation, weak coverage areas and blind zones which are particularly relevant when working in the high frequency spectrum [11, 20, 21]. This is the main reason why network and resource allocation planning are crucial when deploying these systems for cellular communications, independently of the scenario under study [11, 22]. There are few system-level evaluations of mmWave schemes in the literature, such as the ones mentioned in [23, 24], but beyond all the proposals in the field of system level analysis very few examples have been extended to 5G New Radio (5G NR) standard of the 3rd Generation Partnership Project (3GPP) and for the THz level (e.g. 100 GHz) as we cover in this paper. This standard suggests the use of deterministic cluster delay lines (CDLs) for link-level simulations, which requires the definition of average angles of departure (AODs) and arrival (AOAs) and also a tapped delay line (TDL). For system-level simulations, a full three-dimensional (3D) modelling of a radio channel is recommended since this type of analysis requires a statistical approach [25, 26].

Motivated by the work above, in this paper we study a cloud radio access network (C-RAN) for 5G and beyond systems which is based on the adoption of low complexity hybrid precoding designs for massive and ultra-massive multi-user MIMO (MU-MIMO) schemes operating in the mmWave/THz bands. As far as the authors are aware there are no previous studies similar to the one presented in this paper focusing on system-level evaluation of robust hybrid algorithms with different architectures (FC, AoSA, DAoSA, etc.) at the THz band. The C-RAN studied in this paper assumes that each access point (AP) uses the precoder to remove the multi-user interference (MUI) generated at the receivers, breaking the MU communication into equivalent small SU links which enables a lower complexity at the receiver side. We consider a virtualized C-RAN, where the network determines which APs are to be associated with each terminal. The cell moves with and always surrounds the terminal in order to provide a cell-center experience throughout the entire network. Each terminal designated as user equipment (UE) is served by its preferred set of APs. The actual serving set for a UE may contain one or multiple APs and the terminal's data is partially or fully available at some cluster with potential serving APs. The AP controller (Central Processor) will accommodate each UE with its preferred cluster and transmission mode at every communication instance while considering load and channel state information (CSI) knowledge associated with the cluster of APs [27].

The main contributions of this paper can be summarized as follows:

- Thorough system level assessment of a virtualized C-RAN with two clusters sizes, namely, size 1 and 3, where the APs operate in the mmWave/THz bands with massive/ultra-massive antenna arrays combined with low-complexity hybrid precoding architectures. The system level simulations were performed based on link level results between the APs and multiple terminals, where it is considered that the ADMM algorithm from [15] is applied for hybrid precoding design at the transmitter side.
- System level evaluation of the proposed C-RAN in two 5G NR three-dimension (3D) scenarios, namely, the indoor mixed office (InD-MO) and indoor open office (InD-OO).
- System level assessment of the proposed based C-RAN with the fifth numerology of 5G NR in the 3D indoor mixed office scenario with different parameters, such as, number of transmitted antennas per user and number of subcarriers, with the results benchmarked against two alternative MU-MIMO schemes.
- System level evaluation demonstrates that low-complexity hybrid precoding-based C-RAN deployments in an indoor scenario can enable the practical implementation of those schemes, which rely on massive/ultra-massive antenna arrays to combat distance limitation and minimize the MUI. While these hybrid designs sacrifice some performance, significant throughput performance and coverage improvements can still be achieved over typical cellular networks.

In Table 1, we present a list of acronyms adopted along the text in order to improve the readability of the paper. The paper is organized as follows: Section 2 presents the model for the low-complexity hybrid precoding system and the system level scenario that is considered in the evaluation. Section 3 presents and discusses the system level simulations results, whereas the conclusions are outlined in Section 4.

Table 1. Table of acronyms.

Acronym	Designation
3D	Three-dimensional
3GPP	The 3rd Generation Partnership Project
5G	The fifth generation of wireless communications
5G NR	5G New radio
6G	The sixth generation of wireless communications
ADMM	Alternating direction method of multipliers
AM	Alternating minimization
AOA	Average angle of arrival
AOD	Average angle of departure
AoSA	Array of subarrays
AP	Access point
B5G	Beyond 5G
BER	Bit error rate
BLER	Block error rate
CDL	Cluster delay line
C-RAN	Cloud radio access network
CSI	Channel state information
DaoSA	Dynamic array of subarrays
DPS	Double phase shifters
EC	Energy consumption
EE	Energy efficiency
FC	Fully-connected
InD-MO	Indoor-mixed office
InD-OO	Indoor-pen office
LASSO	Least absolute shrinkage and selection operator
LOS	Line-of-sight
MIMO	Multiple Input Multiple Output
MU	Multi-user
MUI	Multi-user interference
MU-MIMO	Multi-user MIMO
mmWave	Millimeter wave
NLOS	Non-line of sight
PC	Partially-connected
QPS	Quantized phase shifters
UM-MIMO	Ultra-massive MIMO
SE	Spectral efficiency
SNR	Signal-to-noise ratio
SPS	Single phase shifters
SU	Single-user
TDL	Tapped delay line
UE	User equipment
UPA	Uniform planar array
TRP	Transmission and reception point
THz	Terahertz

Notation: Matrices and vectors are denoted by uppercase and lowercase boldface letters, respectively. $(\cdot)^T$ and $(\cdot)^H$ denote the transpose and conjugate transpose of a matrix/vector, $\|\cdot\|_p$ is the ℓ_p -norm of a vector, $\|\cdot\|_0$ is its cardinality, $\lfloor \cdot \rfloor$ is the floor function and \mathbf{I}_n is the $n \times n$ identity matrix.

2. System Model

2.1. Transmitter and receiver model

Let us consider a mmWave/THz hybrid MU-MIMO system, where an AP is equipped with N_{tx} antennas that transmit to N_u users simultaneously over F carriers. Each user is equipped with N_{tx} antennas that transmits to N_u users simultaneously. Each user is equipped with N_{rx} antennas over F carriers, as the one described in Figure 1. N_s data streams are transmitted to each user and to each subcarrier, which can be represented as $\mathbf{s}_k = [\mathbf{s}_{k,1}^T \dots \mathbf{s}_{k,N_u}^T]^T$, with $\mathbf{s}_{k,u} \in \mathbb{C}^{N_s \times 1}$. Since a fully digital design would require a dedicated RF chain per antenna element, both the digital and analog processing blocks of the precoder and combiner are separated. By following this approach, it is possible to use reduced digital blocks with only a few radio frequency (RF) chains, which can be complemented by the analog blocks that are based on networks of phase-shifter and switches solely. It is assumed that $N_u N_s \leq N_{RF}^{tx} \leq N_{tx}$ and $N_s \leq N_{RF}^{rx} \leq N_{rx}$, where N_{RF}^{tx} and N_{RF}^{rx} are the number of RF chains at the AP and each user, respectively. The received signal model at user u and subcarrier k after the combiner can be written as

$$\mathbf{Y}_{k,u} = \sqrt{\rho_u} \mathbf{W}_{BB,k,u}^H \mathbf{W}_{RF,u}^H \mathbf{H}_{k,u} \mathbf{F}_{RF} \mathbf{F}_{BB,k} \mathbf{s}_k + \mathbf{W}_{BB,k,u}^H \mathbf{W}_{RF}^H \mathbf{n}_{k,u}, \quad (1)$$

where $\mathbf{H}_{k,u} \in \mathbb{C}^{N_{rx} \times N_{tx}}$ is the frequency domain channel matrix (assumed to be perfectly known at the transmitter and receiver) between the AP and the u^{th} receiver at subcarrier k . $\mathbf{F}_{RF} \in \mathbb{C}^{N_{tx} \times N_{RF}^{tx}}$ and $\mathbf{W}_{RF,u} \in \mathbb{C}^{N_{rx} \times N_{RF}^{rx}}$ represent the analog precoder and combiner, with $u = 1, \dots, N_u$, ρ_u denotes the average received power and vector $\mathbf{n}_{k,u} \in \mathbb{C}^{N_{rx} \times 1}$ contains independent zero-mean circularly symmetric Gaussian noise samples with covariance $\sigma_n^2 \mathbf{I}_{N_{rx}}$. The digital baseband precoders are denoted by $\mathbf{F}_{BB,k} \in \mathbb{C}^{N_{RF}^{tx} \times N_u N_s}$ and the combiners by $\mathbf{W}_{BB,k,u} \in \mathbb{C}^{N_{RF}^{rx} \times N_s}$. In order to maintain the complexity of the implementation limited, the analog component is the same for all subcarriers, which means that \mathbf{F}_{RF} and \mathbf{W}_{RF} will be the same for all subcarriers.

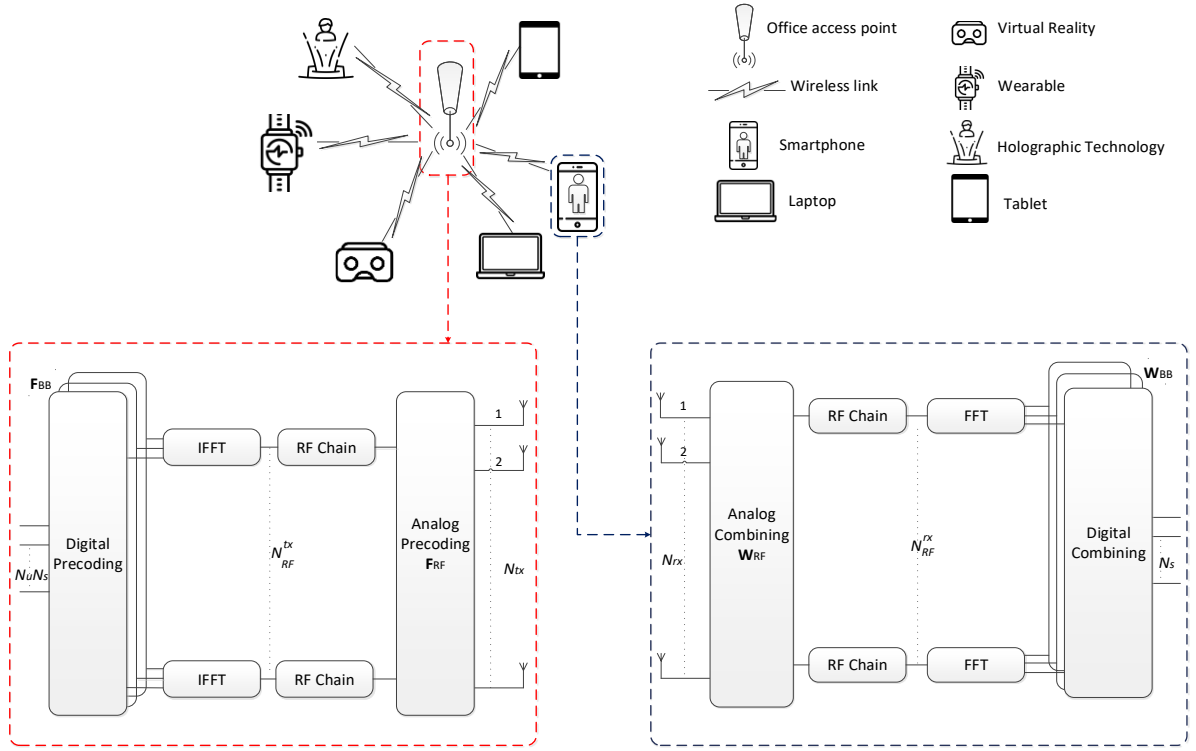


Figure 1. Transmitter and receiver structure of a multiuser OFDM mmWave/THz MIMO system with hybrid precoding for an InD scenario.

Considering the major issues related to spectral efficiency (SE), energy consumption and hardware implementation for THz pointed in introduction, in this study we considered the architectures depicted in Figure 2 [15,28-31]. The FC architecture is the closest to the performance of the digital architecture, but it is also the most power consuming of the hybrid architectures. This is the main motivation for the development of partially connected architectures (AoSAs and DAoSAs) with the aim of trying to achieve similar performances to FC structures, but with lower energy consumption. The AoSAs architectures, especially the ones based on phase shifters, can be divided into two categories, namely, the ones based on single phase shifters (SPS) and the ones based on double phase shifters (DPS). Increasing the number of phase shifters connected to each antenna results in an improved performance but entails a higher energy consumption. On the other hand, DAoSAs represent a much more appealing solution due the good compromise between performance and energy consumption (EC). However, the choice of the most suitable architecture depends on several aspects, which are related to the calculation of the \mathbf{F}_{RF} and \mathbf{F}_{BB} matrices. When considering a FC structure with SPS, the \mathbf{F}_{RF} will be dense with all elements having unit amplitude. Nevertheless, if we consider the case of AoSA architecture, the \mathbf{F}_{RF} elements also have unit amplitude, but the matrix structure is

$$\mathbf{F}_{\text{RF}} = \text{blkdiag}\{\mathbf{f}^{\text{RF},1}, \dots, \mathbf{f}^{\text{RF},N_{\text{RF}}}\} = \begin{bmatrix} \mathbf{f}^{\text{RF},1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \mathbf{f}^{\text{RF},N_{\text{RF}}} \end{bmatrix}, \quad (2)$$

with $\mathbf{f}^{\text{RF},i} = \begin{bmatrix} f_1^{\text{RF},i} & \dots & f_{\frac{N_{\text{tx}}}{N_{\text{RF}}}}^{\text{RF},i} \end{bmatrix} \in \mathcal{U}_{N_{\text{tx}}/N_{\text{RF},i}}$ for $i=1, \dots, N_{\text{RF}}$. In the case of the DAOSA architecture, the matrix is similar to case of AoSA but can have a

number of non-null columns in each row of up to the maximum number of subarrays that can be connected to an RF chain (L_{max}). If we consider DPS-based architectures, the amplitude of the elements becomes less than or equal to 2.

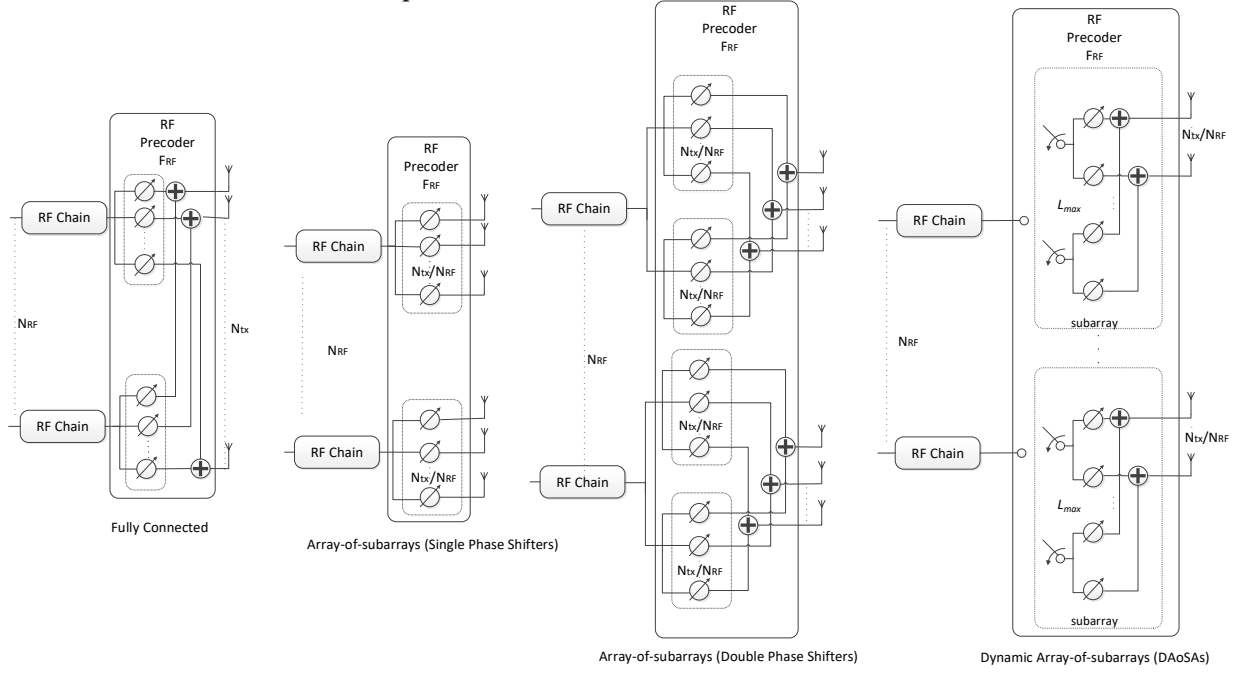


Figure 2. Different precoder architectures for a mmWave/THz MIMO system based on phase shifters.

Several schemes have been proposed in the literature to calculate the matrices \mathbf{F}_{RF} and \mathbf{F}_{BB} , such as [12, 29]. The approach suggested in [15], can calculate \mathbf{F}_{RF} and \mathbf{F}_{BB} by approximation of the digital \mathbf{F} for any of these four structures. The overall optimization problem can be then expressed as

$$\min_{\mathbf{F}_{RF}, \mathbf{F}_{BB_k}} \sum_{k=1}^F \left\| \mathbf{F}_{opt_k} - \mathbf{F}_{RF} \mathbf{F}_{BB_k} \right\|_F^2 \quad (3)$$

$$\text{subject to } \mathbf{F}_{RF} \in \mathcal{C}_{N_{tx} \times N_{RF}^{tx}} \quad (4)$$

$$\left\| \mathbf{F}_{RF} \mathbf{F}_{BB_k} \right\|_F^2 = N_u N_s. \quad (5)$$

In this formulation matrix \mathbf{F}_{opt_k} denotes the fully-digital precoder and the equation (5) enforces the transmitter's total power constraint and $\mathcal{C}_{N_{tx} \times N_{RF}^{tx}}$ is the set of feasible analog precoding matrices, which is defined according to the adopted RF architecture. In order to enforce $\mathbf{F}_{RF} \mathbf{F}_{BB_{k,u}}$ to lie in the null space of $\bar{\mathbf{H}}_{k,u} \in \mathbb{C}^{(N_u-1)N_{tx} \times N_{tx}}$, we write the following restriction to the overall optimization problem expressed in (3) to (5)

$$\mathbf{F}_{RF} \mathbf{F}_{BB_{k,u}} \in \mathcal{N}(\bar{\mathbf{H}}_{k,u}), \quad (6)$$

with $k=1, \dots, F$ and $u=1, \dots, N_u$. $\bar{\mathbf{H}}_{k,u}$ is a matrix corresponding to \mathbf{H}_k with the N_{rx} lines of user u removed which we denote as $\mathcal{N}(\bar{\mathbf{H}}_{k,u})$.

Other RF constraints can be directly integrated into the objective function of the optimization problem in order to cope with the different RF architectures.

2.2. Channel model

Even though mmWave and THz bands share a few similarities, the THz channel presents several unique features that differentiate it from the mmWave channel. In the THz band, the very high scattering and diffraction losses tend to result in a much sparser channel in the angular domain with fewer multipaths components (typically less than 10) [20]. Because of the referred phenomenon, the gap between the line of sight (LOS) and non-line of sight (NLOS) components tends to be very large, which often makes LOS-dominant with NLOS-assisted [28]. An additional aspect relies on the much larger bandwidth of THz signals which can suffer performance degradation due to the so-called beam split effect, where the transmission paths squint into different spatial directions depending on the subcarrier frequency [29]. In light of this, in this paper we consider a clustered wideband geometric channel, which is commonly adopted both in mmWave [12] and THz literature [5, 6, 30, 31]. However, it should be noted that the hybrid precoding/combining approach proposed in this paper is independent of a specific MIMO channel. In this case the frequency domain channel matrices can be characterized as

$$\mathbf{H}_{k,u} = \gamma \left(\alpha_u^{LOS} \mathbf{a}_r(\phi_u^{r,LOS}, \theta_u^{r,LOS}) \mathbf{a}_t(\phi_u^{t,LOS}, \theta_u^{t,LOS})^H + \sum_{i=1}^{N_{cl}} \sum_{l=1}^{N_{ray}} \alpha_{i,l,u} \mathbf{a}_r(\phi_{i,l,u}^r, \theta_{i,l,u}^r) \mathbf{a}_t(\phi_{i,l,u}^t, \theta_{i,l,u}^t)^H \right) e^{-j2\pi\tau_{i,u}f_k}, \quad (7)$$

where N_{cl} denotes the number of scattering clusters with each cluster i , having a time delay of $\tau_{i,u}$ and N_{ray} is the number of propagation paths per cluster. α_u^{LOS} and $\alpha_{i,l,u}$ are the complex gains of the LOS component and of the l^{th} ray from cluster i . Index u is the user ($u=1, \dots, N_u$), $f_k = f_c + \frac{B}{F} \left(k - 1 - \frac{F-1}{2} \right)$ ($k=1, \dots, F$) is the k^{th} subcarrier frequency, B is the bandwidth, f_c is the central frequency and γ is a normalizing factor such that $E[\|\mathbf{H}_{k,u}\|_F^2] = N_{tx} N_{rx}$. Vectors $\mathbf{a}_t(\phi_{i,l,u}^t, \theta_{i,l,u}^t)$ and $\mathbf{a}_r(\phi_{i,l,u}^r, \theta_{i,l,u}^r)$ represent the transmit and receive antenna array responses at the azimuth and elevation angles of $(\phi_{i,l,u}^t, \theta_{i,l,u}^t)$ and $(\phi_{i,l,u}^r, \theta_{i,l,u}^r)$, respectively. Vectors $\mathbf{a}_t(\phi_u^{t,LOS}, \theta_u^{t,LOS})$ and $\mathbf{a}_r(\phi_u^{r,LOS}, \theta_u^{r,LOS})$ have similar meanings but refer to the LOS path angles $(\phi_u^{t,LOS}, \theta_u^{t,LOS})$ and $(\phi_u^{r,LOS}, \theta_u^{r,LOS})$. By carefully selecting the parameters of the channel model it is possible to make it depict a mmWave or a THz channel. As represented in (7), this channel model includes both LOS and NLOS components. In the case of the NLOS components, we consider complex gaussian distributed paths gains, as in [32]. Radio propagation measurements have shown that this type of clustered based channel models can yield good agreements with real channel behavior at mmWave and subTHz frequencies [33].

2.3. System level scenario

The indoor hotspot deployment scenario focuses on small cells and high user density in buildings. This scenario, described in Table 2, represents InDs with a total area of 120 m×50 m. There are 12 small tri-sector cells which are deployed with an inter-site distance (ISD) of 20 m. In this case the AP antenna height is 3 m. The coverage radius of the APs is $R=6.7$ m. The carrier frequency option is 100 GHz (THz waves). The bandwidth chosen is $B_f=400$ MHz corresponding to numerology five of 5G NR. Up to 16 carriers can be aggregated up to 6.4 GHz of bandwidth. It is important to note that in case of the number of users and wireless devices increasing relative to the numbers considered here, the expected behavior of the overall system will remain the same as long as carrier aggregation is used, and they will not cause inter-interference as long as they operate in different bands. A total of 15 users per AP are distributed uniformly and all users are indoors with 3 km per hour velocity as can be seen in Figure 3. Full buffer model is assumed. Our 3D simulation channel model considers the 5G NR indoor office wireless propagation environment in terms of physical aspects of mmWave and THz waves.

Table 2. Evaluation parameters for indoor-office scenario.

Parameters		Indoor – office
Layout	Room size (W×L×H)	120 mx50 mx3 m
	ISD	20m
AP antenna height h_{AP}		3 m (ceiling)
UE location	LOS/NLOS	LOS and NLOS
	Height h_{UE}	1.5 m
UE mobility (horizontal plane only)		3 km/h
Min. AP - UE distance (2D)		0
UE distribution (horizontal)		Uniform

For this particular scenario in any considered operating frequency, the number of clusters, number of subarrays (scatterers) per cluster, and the positions of the clusters are determined by the detailed steps and procedures of [25].

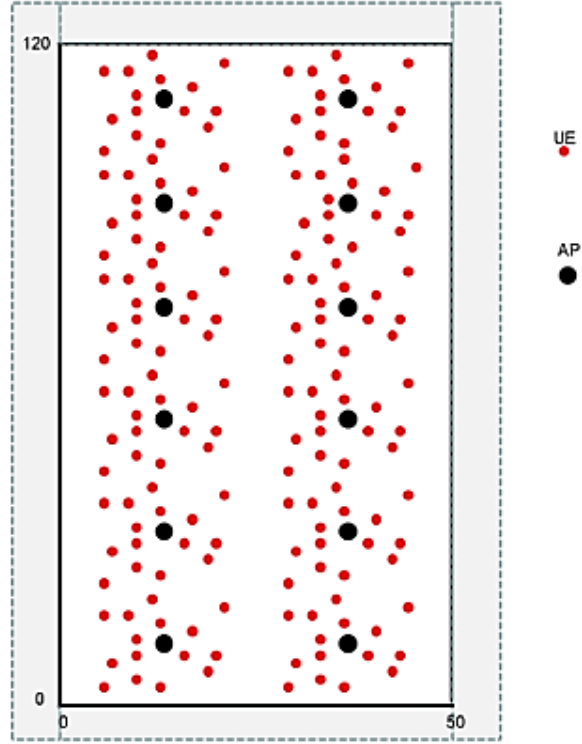


Figure 3. Indoor Office scenario layout.

According to the 5G 3GPP 3D channel models, the number of clusters and scatterers are determined using the Poisson and uniform distributions with specific parameters. Since, we extend the operating frequency range up to 100 GHz, specific multi-antenna solutions and techniques need to be employed depending on the utilized spectrum. For higher frequency bands, the transmission is characterized by a considerable signal attenuation that limits the network coverage. To overcome this limitation, one of the key features is the adoption of a very large number of multi-antenna elements having a given aperture to increase the transmission/reception capability of MU-MIMO and beamforming. Since, managing transmissions in higher frequency bands is complicated, beam management is necessary to establish the correspondence between the directions of the transmitter and the receiver-side beams by identifying the most suitable beam pair for both downlink and uplink.

3. Numerical Results

In this section we present the numerical assessment of both the link and system level of a massive/ultra-massive MU-MIMO downlink scheme operating in the mmWave/THz band integrated into a 5G NR system, where the APs are based on the low complexity ADMM-based hybrid precoding designs. Link level results are presented in terms of bit error rate (BER) and measure the performance of the signal across the entire communication chain, from transmitter to receivers. Both the link and system level diagrams can be found in references [15, 27].

3.1. Link Level Simulations

Considering that both the transmitter and receivers are equipped with uniform planar arrays (UPAs) with $\sqrt{N_{tx}} \times \sqrt{N_{tx}}$ antenna elements

at the transmitter and $\sqrt{N_{rx}} \times \sqrt{N_{rx}}$ at the receiver, the respective array response vectors are given by

$$\mathbf{a}_{t/r}(\phi_{i,l,u}^{t/r}, \theta_{i,l,u}^{t/r}) = \frac{1}{\sqrt{N_{tx/rx}}} \times \left[1, \dots, e^{j\frac{2\pi}{\lambda}d(p \sin \phi_{i,l,u}^{t/r} \sin \theta_{i,l,u}^{t/r} + q \cos \theta_{i,l,u}^{t/r})}, \dots, e^{j\frac{2\pi}{\lambda}d((\sqrt{N_{tx/rx}}-1) \sin \phi_{i,l,u}^{t/r} \sin \theta_{i,l,u}^{t/r} + (\sqrt{N_{tx/rx}}-1) \cos \theta_{i,l,u}^{t/r})} \right]^T, \quad (8)$$

where λ is the signal wavelength, d is the inter-element spacing (a $d = \lambda/2$ is assumed) and $p, q = 0, \dots, \sqrt{N_{tx/rx}} - 1$ are the antenna indices. We consider a sparse channel with limited scattering where $N_{ray} = 1$ and $N_{cl} = 9$. The angles of departure and arrival were selected according to a Gaussian distribution whose means are uniformly distributed in $[0, 2\pi]$ and whose angular spreads are 10 degrees. The results are presented for a NLOS and LOS channel. This last case considers a ratio of $E[\alpha_u^{LOS}]^2 / \sum_{i=1}^{N_{cl}} \sum_{l=1}^{N_{ray}} E[\alpha_{i,l,u}]^2 = 10$, which means we are admitting very weak NLOS paths when compared to LOS.

In a first approach, we compared several solutions available in the literature concerning the hybrid precoding, such as AM [12], LASSO [34] and ADMM [15] based precoding. In Figure 4, we assumed a scenario with $N_{tx}=100$, $N_{rx}=4$, $N_{RF}^{tx} = 12$, $N_u=4$ and $N_s=2$, in which we change the number of subcarriers. To ensure a fair comparison, all schemes have a SE close to 2 bits per channel use (bpcu) per user. In this study, we considered that the hybrid precoding algorithms (LASSO and ADMM) can be applied to an architecture based on SPS or DPS, as proposed by [35].

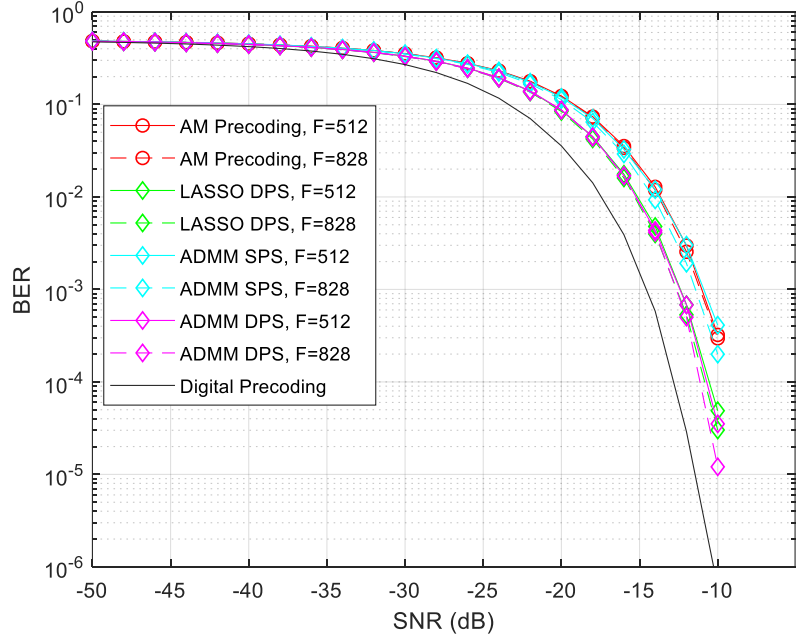


Figure 4. BER performance versus SNR achieved by different methods by changing the number of subcarriers and considering a system with $N_{tx}=100$, $N_{rx}=4$, $N_{RF}^{tx} = 12$, $N_u=4$ and $N_s=2$ (only NLOS component).

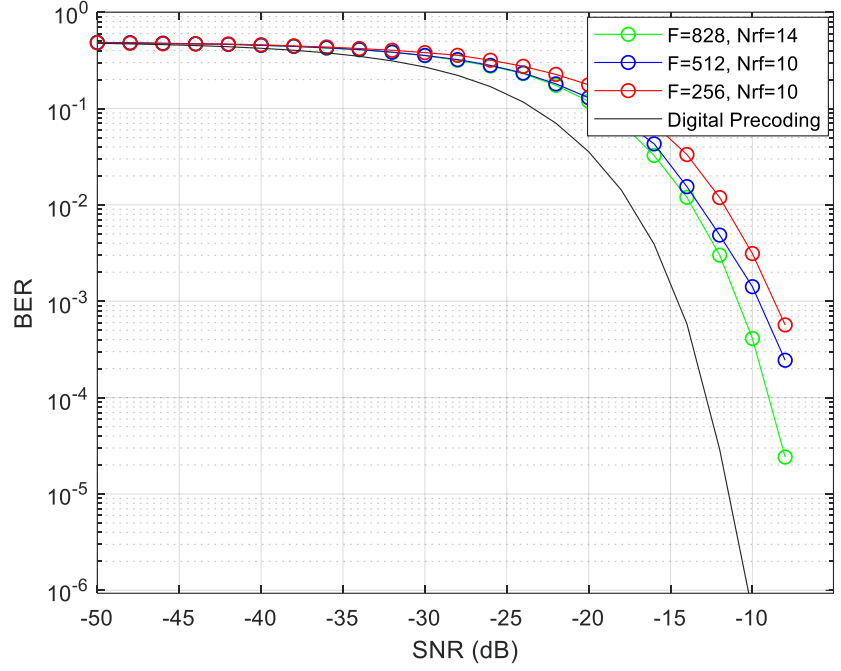


Figure 5. BER performance versus SNR achieved by the proposed precoder with $N_{tx}=256$, $N_{rx}=4$, $N_u=4$, and $N_s=2$, by changing the number of subcarriers and the number of RF chains (only NLOS component).

It can be observed that the ADMM based on a DPS scheme outperforms both versions of AM and LASSO (SPS/DPS), and the gains tend to be slightly greater when the symbols are distributed over a larger number of subcarriers (F). However, the larger the number of subcarriers the greater will be the number of the required RF chains to maintain a good performance. In general, the ADMM precoding algorithm is the one that can achieve the best results at the cost of some additional computational complexity and it is also the most flexible since it can cope with different architectures, as explained in [15]. At the next subsection of the paper, we will compare these three methods based on system level simulations.

In order to understand how many RF chains are necessary to maintain a good performance as the number of subcarriers increases, we simulated the ADMM algorithm for a scenario with $N_{tx}=256$, $N_{rx}=4$, $N_u=4$ and $N_s=2$ and we changed the number of subcarriers, as can be seen in Figure 5. By maintaining the same number of RF chains as the number of subcarriers increases, some performance degradation results as can be seen when we compare the curves of $F=256$ and $F=512$. In fact, in the case of $F = 828$, 10 RF chains were insufficient to provide acceptable performance and therefore 14 RF chains were employed.

After studying the performance of the ADMM precoder considering the variation of the number subcarriers, F , and RF chains, it is important to understand the main differences in terms of performance when considering architectures based on SPS, DPS as well as the impact of quantized phase shifters (QPS). The results are shown in Figure 6. It can be seen that a better performance can be achieved when considering a design based on a DPS over the SPS. It is important to remember, however, that the architecture based on DPS provides a better

performance but requires the use of twice the number of phase shifters in the implementation.

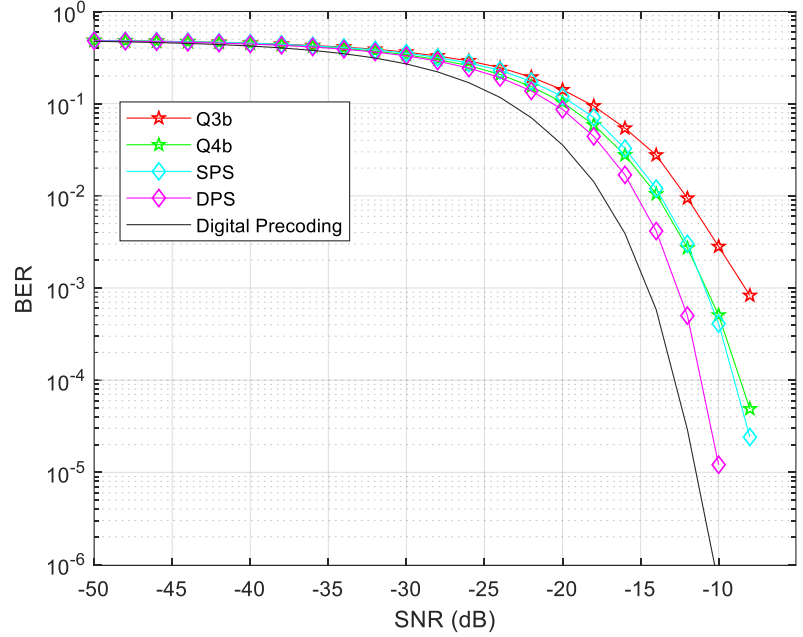


Figure 6. BER performance versus SNR achieved by the proposed precoder with $N_{tx}=256$, $N_{rx}=4$, $N_u=4$, $N_{RF}^{tx} = 14$, $F=828$ and $N_s=2$, considering several versions of the algorithm (only NLOS component).

In this same figure, we include QPS curves, which corresponds to the SPS architecture with quantized phase shifters. It can be observed that, as expected, there is a performance loss when comparing the QPS curves against the ideal SPS, but this degradation is greatly reduced when using phase shifters with only 4 bits of quantization. Beyond, the choice of the parameters of the scenario, the BER performance can be improved by increasing the number of iterations of precoding algorithms.

3.2. System Level Simulations

Bit and block error rate (BER/BLER) results obtained from link level simulations are used as input for the system level evaluation. As presented in the system model section, we consider an indoor office virtualized C-RAN, where the network determines which APs are to be associated with each UE. The total number of sites with APs is 12, which are equally spaced. Each site with APs consists of three transmission and reception points (TRPs), each one equipped with UPAs with N_{tx} antennas, while UEs also have one UPA with N_{rx} antennas. The number of antennas N_{tx} of each array is 100 or 256 antennas whereas the separation between antennas of the array is half wavelength. The signal-to-noise ratio (SNR) in dB considered in the system level simulations is obtained from $SNR = (E_s/N_0) + 10 \log(R_s/B)$ dB, where R_s is the total transmitted symbol rate per antenna and user, B is the total bandwidth, and E_s/N_0 is the ratio of symbol energy to noise spectral density in dB. Values of E_s/N_0 are obtained from the link level BLER results. The BLER can be expressed using the BER, according to the following expression

$$BLER=1-(1-BER)^L, \quad (9)$$

where L is the length of the block in bits. L can be calculated as $L = 2 \times F$, since the selected modulation is quadrature phase shift keying (QPSK)

and each subcarrier transports a symbol with 2 bits. By maintaining the reference BLER (BLER_{ref}) equal to 0.1 as the F increases, we must decrease the respective reference BER (BER_{ref}).

We considered the fifth numerology of 5G NR with spacing between the subcarriers of 480 KHz. The transmission time interval (TTI) of this numerology is 31.25 μs and the total bandwidth is $B=400$ MHz. Up to 16 carriers can be aggregated if more bandwidth and higher binary rates, according to the 5G NR specifications, are to be achieved. We used the 3D InD-MO channel model specified by 3GPP [25] which has a LOS probability (P_{LOS}) equal to 1 for a distance lower than 1.2 m between the AP and terminals. The probability of NLOS components is $P_{NLOS} = 1 - P_{LOS}$ and depends on the distance between the AP and the terminals. Previously, we started our study by comparing several solutions available in the literature concerning hybrid precoding algorithms in terms of BER versus SNR.

Based on those results, we decided to analyze the performance of a system, where the APs operate in the mmWave/THz bands with massive/ultra-massive antenna arrays combined with those different hybrid precoding architectures. For this evaluation, first we start by considering that the cluster size is equal to 1. When RAN cluster size is one (1C), we have the traditional cellular network where each AP generates inter-site interference. When RAN cluster size is 3, the network is partitioned into three adjacent site sets and each user is served at the same time by three TRPs generating much less inner interference. It is important to note that the throughput presented in the following figures concerns the average throughput value of the various UEs moving randomly throughout the 200 seconds of simulation. Initially, the terminals are placed uniformly distributed inside each sector served by an UPA antenna. The movement of the terminals is "Random waypoint around AP". At every 0.5 ms of simulation the SNR at all terminals is calculated. Those blocks which are received with $\text{SNR} > \text{SNR}_{target}$ are considered in the throughput calculation. The blocks that are received with $\text{SNR} < \text{SNR}_{target}$ are not counted in the throughput due to the throughput definition itself. SNR_{target} is the SNR value for $\text{BLER}_{ref}=0.1$. In practical terms, each block received with $\text{SNR} < \text{SNR}_{target}$ would be retransmitted until it is correctly received. However, we do not consider this situation in the simulations.

Figure 7 is based on the link level simulation of Figure 4 presented in the previous subsection. In general, it is observed that the greater the number of subcarriers, the greater will be the throughput. As expected, the digital precoder is the one which presents the best performance in terms of throughput as the number of subcarriers and users increases. However, it is possible to approach its performance by considering the use of hybrid algorithms and their different architectures. The ADMM based on a SPS scheme is the one, which presents the worst performance for a lower number of subcarriers, but when this parameter is increased it obtains a similar performance to the AM precoding algorithm. The LASSO precoding algorithm with DPS and $F=512$ can obtain a slightly higher performance when compared to the ADMM DPS with $F=512$. However, by increasing the number of subcarriers we observe that both algorithms obtain very similar performances. The blue curves represent the theoretical curves of average throughput value expected to be obtained

based on the formula $\text{Throughput} = R_{b_{\max}} (1 - \text{BLER}_{\text{ref}})$, where $\text{BLER}_{\text{ref}} = 0.1$ and $R_{b_{\max}}$ is the maximum binary transmission rate considered for the two cases $F=512$ and $F=828$. Contrary to what it can be seen for $F=512$, a loss of performance of the algorithms for $F=828$ when compared to the curve of average throughput value expected for the case of $F=828$, is observed. Moreover, we observe that with a fixed number of RF chains, the throughput of hybrid schemes tends to be closer to digital with $F=512$ than with $F=828$. When considering the digital precoding for $F=828$ and 3C, a maximum SE of 1.9 bps/Hz/user can be obtained. Note that, this fact is also verified in the later graphs.

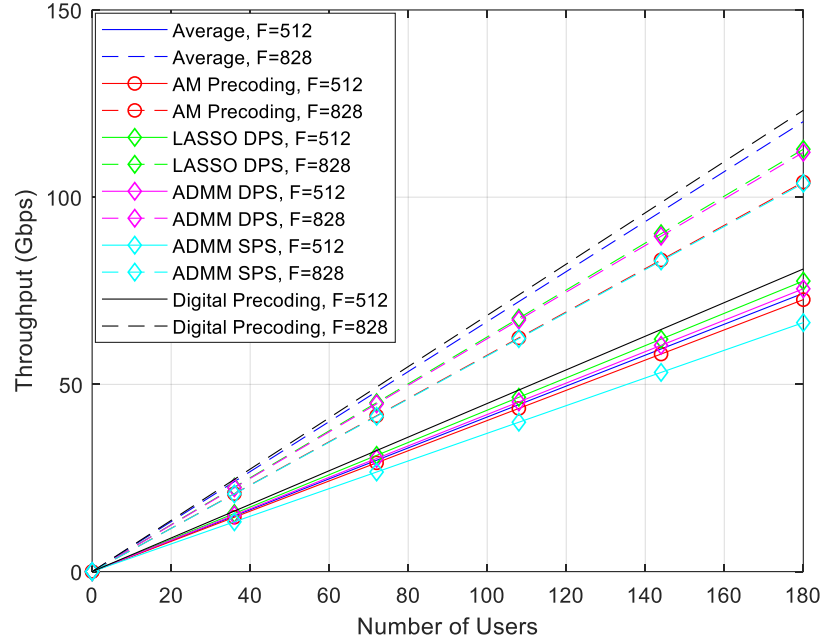


Figure 7. Throughput (Gbps) vs number of users for different methods with $N_{tx}=100$, $N_{rx}=4$, $N_{RF}^{tx} = 12$ and $N_s=2$, by changing the number of subcarriers (only NLOS component).

Following the study on the required number of RF chains to accommodate an increasing number of subcarriers discussed in Figure 5, it is necessary to understand how the system behaves when considering different RAN cluster sizes. In Figure 8, we can conclude that by increasing the number of RF chains for higher values of F allow us to reach greater levels of throughput. Furthermore, when the cluster size increases to 3 (3C), as the number of users increases, the difference becomes more notorious due to the lesser inner interference caused by the network partitioning. When we have $N_u=180$, the achieved throughput increases from 105.7 Gbps to 133.3 Gbps if the cluster size triples. It must be noted that for smaller blocks, such as $F=16$ up to 128, the system performance is independent of cluster size. When F presents higher values, we observed that the performance difference of 1C and 3C increases with block size.

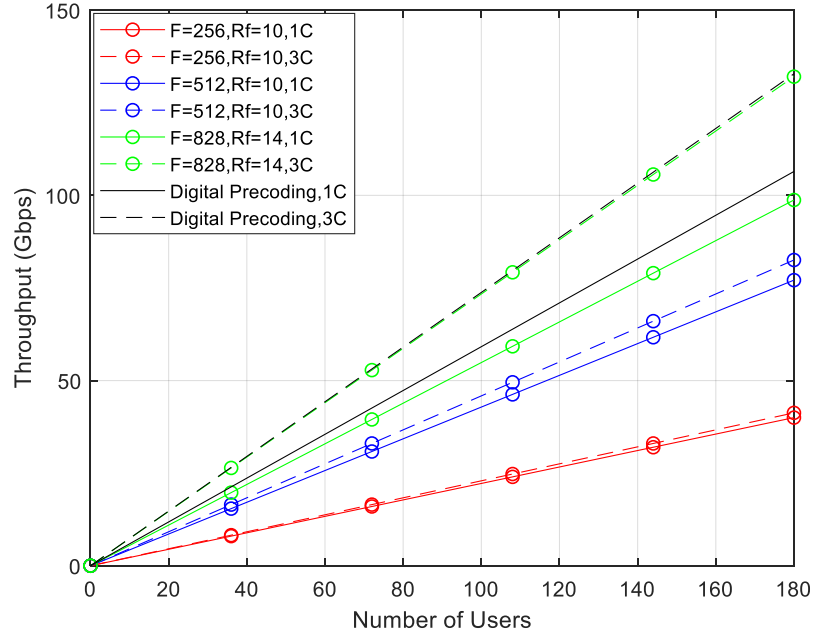


Figure 8. Throughput (Gbps) vs number of users for different number of subcarriers, RF chains and cluster sizes, considering the proposed precoder with $N_{tx}=256$, $N_{rx}=4$ and $N_s=2$ (only NLOS component).

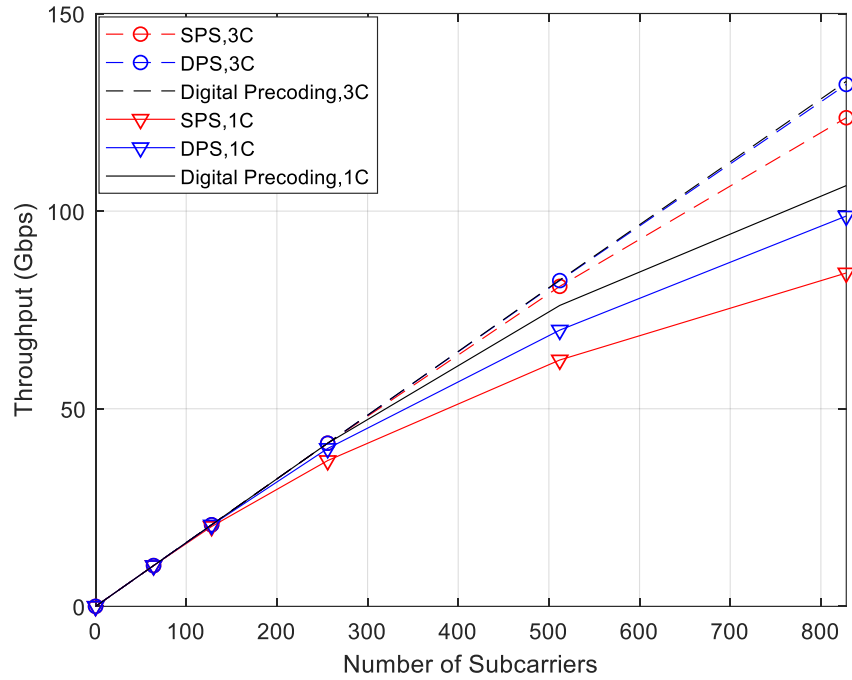


Figure 9. Throughput (Gbps) vs number of subcarriers for different versions of the algorithm (SPS and DPS) and cluster sizes, considering the proposed precoder with $N_{tx}=256$, $N_{rx}=4$, $N_{RF}^{tx}=14$, $N_s=2$ and $N_u=180$ (with LOS component).

At this point, we already know that the throughput increases as the number of subcarriers increases and the ADMM precoding algorithm constitutes as an interesting alternative, since it can present a good performance and can cope with different architectures, facilitating their implementation. However, it is worth to understand how the system behaves when the AP antennas are combined with different architectures based on phase shifters, as can be seen in Figure 9. Independently of the

adopted architecture, it is observed that the throughput increases as the number of subcarriers increases. However, the adoption of a DPS architecture can provide a higher throughput when compared to an architecture based on SPS. This difference is even more notorious when we increase the cluster size from 1 to 3. Considering a $F=828$ and 1C the version of the algorithm based on SPS can reach a throughput of 85.79 Gbps whereas the one based on DPS can reach a throughput of 99.99 Gbps. When we have 3C, both versions of the algorithm can obtain a significant improvement over the values obtained for 1C. The first one can reach a throughput around 127.3 Gbps and the latter one reaches 132 Gbps. In fact, the version based on a DPS architecture with 3C is the one who can closely approach the digital precoding curve.

After previously analyzing two different solutions based on phase shifters, we now evaluate the performance of the system in terms of throughput and coverage with more realistic phase shifters, i.e., phase shifters with quantization. Figure 10 is based on the link level simulation of Figure 6 presented in the previous subsection. From this figure, we observe that, as expected, by increasing the number of quantization (QPS curves) it is possible to approach the performance of the ideal unquantized version of the algorithm (SPS curve). Furthermore, it can also be seen that by increasing the cluster size, the throughput can almost double when the number of subcarriers is large. Moreover, we observe that by increasing the cluster size from 1C to 3C that the hybrid precoders curves get closer to the digital precoder curve. A lower throughput loss, due to the quantization effect, is also observed.

Following the conclusions about the previous figure concerning the impact of quantized phases shifter in the hybrid architectures, we decided to study the system performance using the same hybrid architecture in the InD-MO and the InD-OO scenarios. The throughput per user results of Figure 11 and the average coverage results of Figure 12, were simulated considering a precoder based on QPS with $N_b = 3$ and 4 bits. It can be observed in Figure 11 for the InD-MO scenario, that by increasing the cluster size the throughput will improve and if we consider the effect of using more quantization bits the difference becomes more notorious. When the cluster size increases from 1 to 3, we can observe at 100 mW that the throughput doubles in both cases of quantization. If we consider the InD-OO scenario, the obtained gains present a similar behavior as the power increases but the throughput we can reach will be significantly lower. The difference between these two scenarios lies in the relative weight of the LOS and NLOS components, as the attenuation losses associated with them are the same in both scenarios. In the InD-MO scenario the P_{LOS} is 1 up to 1.2 m and decreases to 0.368 at 6.5 m.

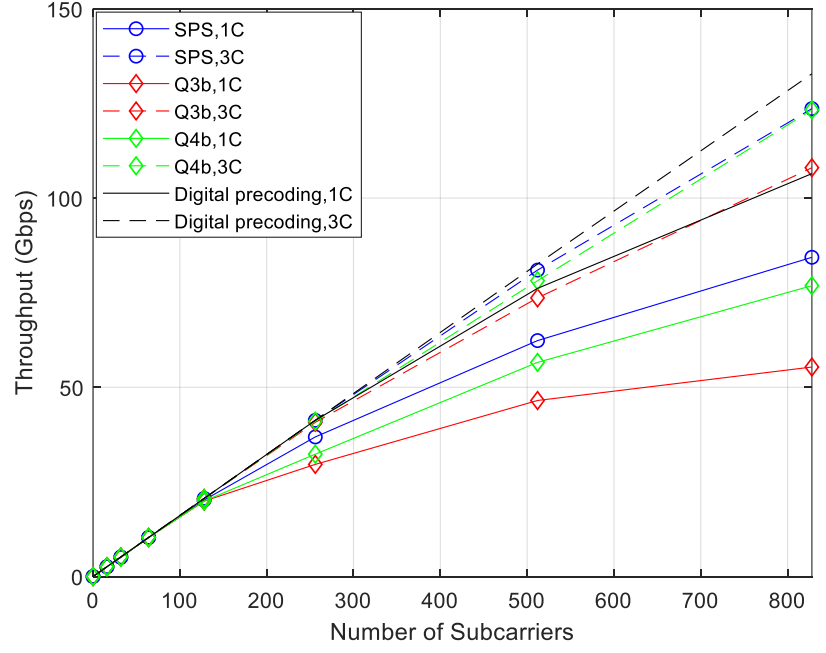


Figure 10. Throughput (Gbps) vs number of subcarriers for different versions of the algorithm (with and without quatization) and cluster sizes, considering the proposed precoder with $N_{tx}=256$, $N_{rx}=4$, $N_{RF}^{tx} = 14$, $N_s=2$ and $N_u=180$ (only NLOS component).

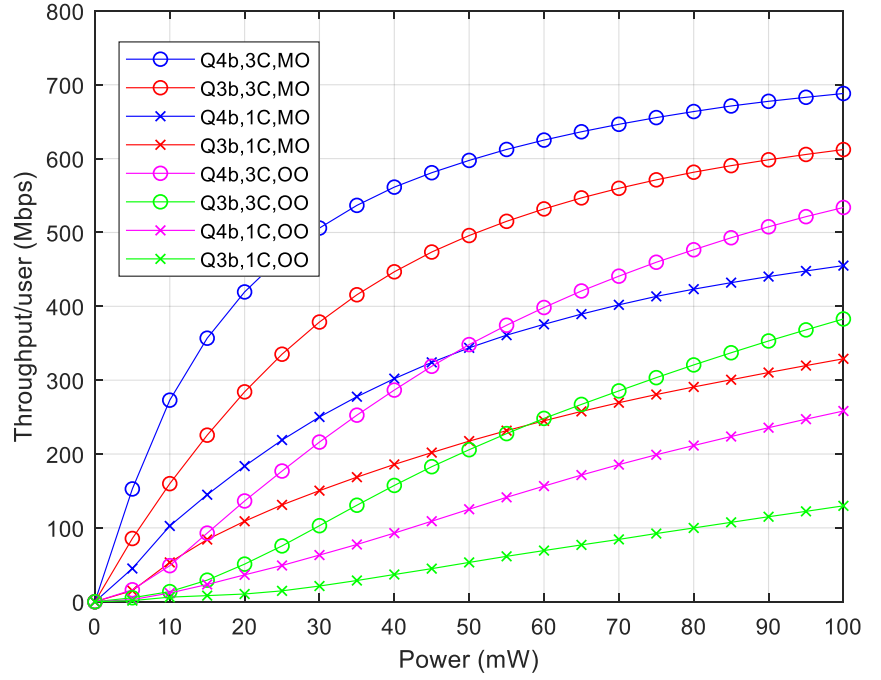


Figure 11. Throughput (Mbps) versus mean value of power required by each user (mW), considering the proposed precoder with $N_{tx}=256$, $N_{rx}=4$, $N_{RF}^{tx} = 14$, $N_s=2$, $F=828$ and the quantization effect (with LOS component).

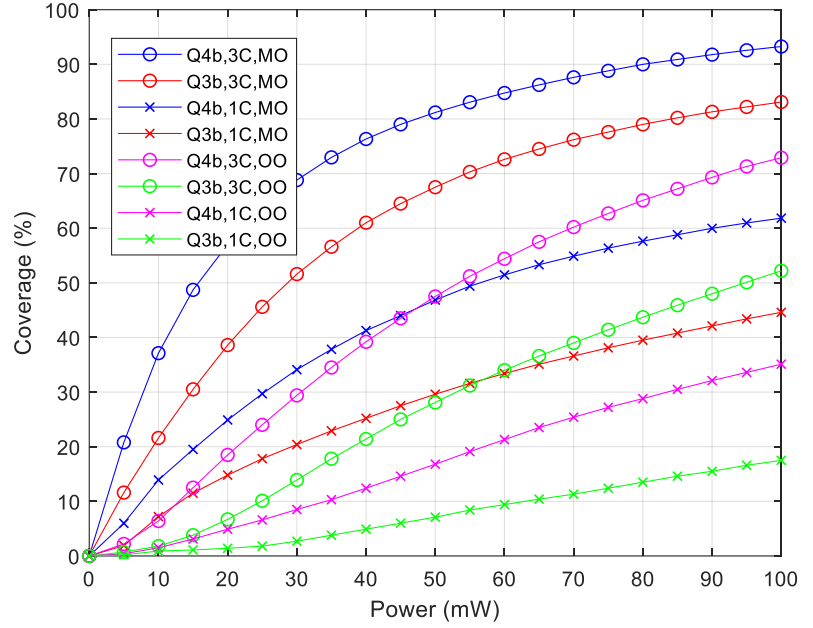


Figure 12. Coverage (%) versus mean value of power required by each user (mW), considering the proposed precoder with $N_{tx}=256$, $N_{rx}=4$, $N_{RF}^{tx} = 14$, $N_s=2$, $F=828$ and the quantization effect (with LOS component).

In the InD-OO scenario the P_{LOS} is 1 up until 5.0 m and decreases to 0.538 at 49.0 m. Knowing that the P_{NLOS} is equal to $1 - P_{LOS}$, we realize that in the InD-MO scenario the weight of the NLOS component will be greater than that of the LOS component. Since the attenuation losses of NLOS are much higher than those of the LOS component, the received power of the APs decreases rapidly with distance. Therefore, in the case of the InD-OO scenario, the weight of the LOS component is greater than the one of the NLOS component, so the transmitted power of the APs decreases very slowly with distance generating a strong inter-site interference, which affects the maximum throughput and coverage that can be reached. Moreover, the channel can influence the results since with a stronger LOS component the multipaths (NLOS component) become weaker, which worsens the spatial multiplexing and the spatial diversity effects. Figure 12 exhibits a similar behaviour when compared to the previous figure, since the greater the cluster size and the number of quantization bits, the greater will be the achieved coverage. With 3 and 4 bits of quantization at 100 mW and when the cluster size increases from 1 to 3, we obtain a significant improvement which almost doubles the initial coverage. If we consider the InD-OO scenario, the obtained gains in terms of coverage present a similar behavior as the power increases but we can no longer obtain more than 73% of coverage ($N_b=4$ bits and 3C).

In order to understand the influence of the variation of the transmitted power (P_t) of the APs on the system-level simulation results, we considered two InD scenarios with the proposed precoder based on a DPS architecture with $N_{tx}=256$, $N_{rx}=4$, $N_{RF}^{tx} = 14$, $N_s=2$ and $F=828$. In Table 3, we present the results of average throughput per user for both InD-OO and InD-MO scenarios when we vary the transmitted power (P_t) of each AP. We considered that $P_t \in \{10, 100, 1000\}$ mW. As expected, based on Figures 11 and 12, the performance of scenario MO is higher than OO due

to less inter-cell interference. For the same reason cluster 3C presents higher performance than cluster 1C. We can check that independently of the scenario and cluster size there is an optimum value of P_i , which is 100 mW. The highest values of throughput occur for 100 mW. This value of P_i was considered in all system level simulation of this paper. It corresponds to the best tradeoff between transmitted power and inter-cell interference. For cluster size one (1C) and InD-OO, the performance with 10 mW tends to be higher than 1000 mW due to the high inter-interference of this cluster. However, for cluster size (3C) occurs the opposite because of lower inter-interference of this cluster.

Table 3. Influence of the variation of the transmitted power of the APs on the system results.

Transmitted Power (mW)	Cluster size	Average Throughput (Mbps)	
		MO	OO
10	1C	542.6	388.3
10	3C	725.6	648.2
100	1C	551.6	411.6
100	3C	733.4	662.8
1000	1C	526.3	376.3
1000	3C	722.3	665.8

To reduce the large performance loss due to the adoption of a simple AoSA architecture, we can allow the dynamic connection of more subarrays to each RF chain by adopting a DAoSA structure and study the impact of the use of SPS and DPS architectures, from a system level perspective. The goal of these structures is to try to reach a compromise between fully connected structures which are more complex to implement (especially with a massive number of antennas) and AoSAs ($L_{max}=1$) which are lighter but typically suffer a large performance loss.

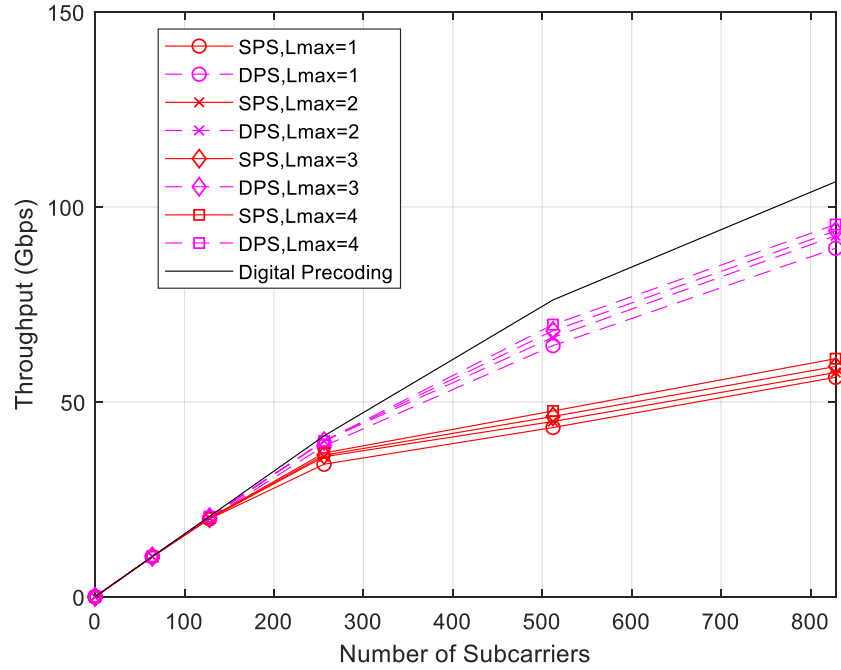


Figure 13. Throughput (Gbps) vs number of subcarriers for different versions of the algorithm (SPS or DPS) as a function of L_{max} , considering the proposed

precoder with $N_{tx}=256$, $N_{rx}=4$, $N_{RF}^{tx} = 14$, $N_s=2$ and $N_u=180$ (with LOS component) and 1C.

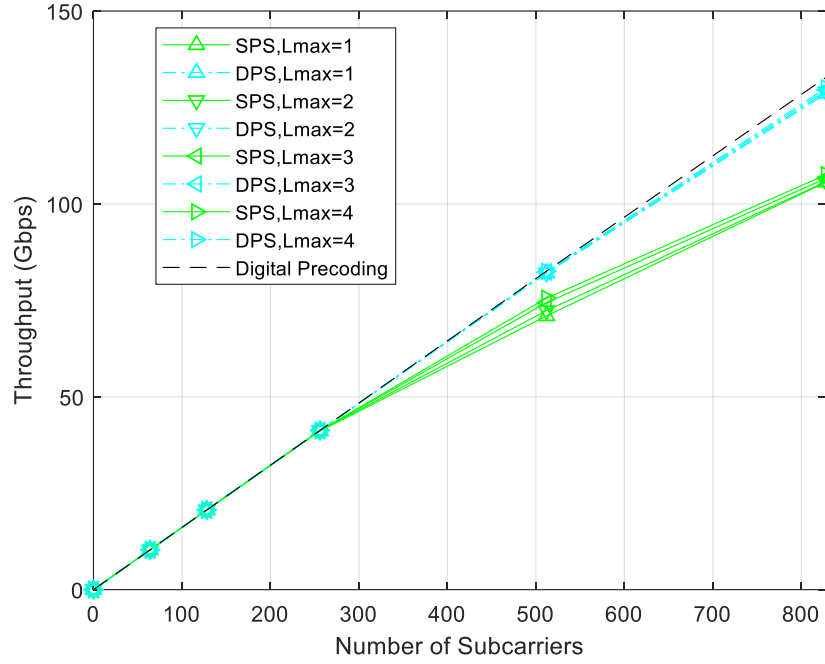


Figure 14. Throughput (Gbps) vs number of subcarriers for different versions of the algorithm (SPS or DPS) as a function of L_{max} , considering the proposed precoder with $N_{tx}=256$, $N_{rx}=4$, $N_{RF}^{tx} = 14$, $N_s=2$ and $N_u=180$ (with LOS component) and 3C.

It possible to conclude, through Figure 13 and Figure 14, that the performance of the system can be improved as the L_{max} increases. The use of DPS instead of using SPS architectures can provide higher throughputs but the difference between the gains tend to reduce when the cluster size increases. When we have 1C the gains of the DPS over the SPS can surpass 50% and with 3C we obtain gains around 20%. In general, by combining the increase in L_{max} with the adoption of DPS and the cluster size, it is possible to improve the results, but the gains become less pronounced for $L_{max} > 1$. A performance close to the digital precoding can be achieved when considering the use of a DPS scheme for $L_{max}=4$ and a cluster size of 3. When compared to the cluster size 1C, with 3C it is observed that the throughput loss is lesser when adopting the hybrid schemes instead of the digital one.

4. Conclusions

In this paper we described a cloud radio access network aimed at operating at the mmWave/THz bands in beyond 5G systems where the access points support multi-user MU-MIMO transmission using massive/ultra-massive antenna arrays. In order to make the implementation of this scheme feasible, low-complexity hybrid precoding based on several analog architectures, namely fully connected, array of subarrays and dynamic array of subarrays combined with single or double phase shifters, was considered. Numerical evaluation of both the link and system level for the proposed cloud radio access network downlink scheme integrated into a beyond 5G system showed that it is possible to obtain a performance close to the digital precoding when

considering scenarios with a larger number of subcarriers and users. The use of partially connected architectures (array of subarrays and dynamic array of subarrays) when compared with fully connected structures tends to achieve good performances in terms of the trade-off between spectral efficiency and energy consumption.

Assessment of the cloud radio access network deployments in two indoor office scenarios at the THz band showed on the capability of achieving significant improvements in terms of throughput performance and coverage over typical cellular networks. Indoor-mixed office scenario provides higher throughput and coverage performances in comparison to the indoor-open office scenario independently of the cluster size, due to less inter-site interference. The variation of the transmitted power of the access points can impact the system results, since in order to obtain a better performance an optimal value of transmitted power around 100 mW must be considered. Furthermore, it was observed that the use of a larger cluster size, namely using 3C instead of 1C, tends to result in a lower throughput penalty when replacing a fully digital implementation by a lower complexity hybrid scheme (fully connected or dynamic array of subarrays).

Author Contributions: Conceptualization, J.P.P., A.C. and N.S.; methodology, J.P.P., A.C. and N.S.; software, J.P.P., A.C., V.V. and N.S.; validation, J.P.P., A.C., V.V., M.R., P.S. and N.S.; formal analysis, J.P.P., A.C., V.V., M.R., P.S. and N.S.; investigation J.P.P., A.C., V.V., M.R., P.S. and N.S.; resources, A.C., M. R., P.S. and N.S.; data curation, J.P.P. and A.C.; writing—original draft preparation J.P.P., A.C., V.V., M.R., P.S. and N.S.; writing—review and editing, J.P.P., A.C., V.V., M.R., P.S. and N.S.; visualization, J.P.P. and A.C.; supervision, A.C., M.R. and N.S.; project administration, A.C. and N.S.; funding acquisition, A.C. and N.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the FCT—Fundação para a Ciência e Tecnologia under the grant 2020.05621.BD. The authors also acknowledge the funding provided by FCT/MCTES through national funds and when applicable co-funded EU funds under the project UIDB/50008/2020.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dang, S.; Amin, O.; Shihada, B. and Alouini, M. S. “What should 6G be?,” *Nature Electronics* **2020**, 3, 1, 20–29.
2. Saad, W.; Bennis, M. and Chen, M. “A Vision of 6G Wireless Systems: Applications, Trends, Technologies, and Open Research Problems,” *IEEE Network* **2020**, 34, 3, 134–142.
3. Akyildiz, I.F.; Kak, A.; Nie, S. 6G and Beyond: The Future of Wireless Communications Systems. *IEEE Access* **2020**, 8, 133995–134030.
4. Srieddeen, H.; Alouini, M.-S.; Al-Naffouri, T.Y. An Overview of Signal Processing Techniques for Terahertz Communications. *Proc. IEEE* **2021**, 1–38.
5. Lin, C.; Li, G.Y.L. Terahertz Communications: An Array-of-Subarrays Solution. *IEEE Commun. Mag.* **2016**, 54, 124–131.
6. Faisal, A.; Srieddeen, H.; Dahrouj, H.; Al-Naffouri, T.; Alouini, M. Ultramassive MIMO Systems at Terahertz Bands: Prospects and Challenges. *IEEE Vehicular Technology Magazine* **2020**, 15, 33–42.

7. Castaneda, E.; Silva, A.; Gameiro, A.; Kountouris, M. An Overview on Resource Allocation Techniques for Multi-User MIMO Systems. *IEEE Commun. Surv. Tutor.* **2017**, *19*, 239–284.
8. Zu, K.; de Lamare, R.; Haardt, M. Generalized Design of Low-Complexity Block Diagonalization Type Precoding Algorithms for Multiuser MIMO Systems. *IEEE Trans. Commun.* **2013**, *61*, 4232–4242.
9. Lopes, P.; Gerald, J. Leakage-based precoding algorithms for multiple streams per terminal MU-MIMO systems. *Digit. Signal Process.* **2018**, *75*, 38–44.
10. Guerreiro, J.; Dinis, R.; Montezuma, P.; Marques da Silva, M. On the Achievable Performance of Nonlinear MIMO Systems. *IEEE Commun. Lett.* **2019**, *23*, 1725–1729.
11. Chataut, R.; Akl, R. Massive MIMO Systems for 5G and beyond Networks—Overview, Recent Trends, Challenges, and Future Research Direction. *Sensors* **2020**, *20*, 2753.
12. Yuan, H.; An, J.; Yang, N.; Yang, K.; Duong, T.Q. Low Complexity Hybrid Precoding for Multiuser Millimeter Wave Systems Over Frequency Selective Channels. *IEEE Trans. Veh. Technol.* **2018**, *68*, 983–987.
13. Liu, F.; Kan, X.; Bai, X.; Du, R.; Liu, H.; Zhang, Y. Hybrid precoding based on adaptive RF-chain-to-antenna connection for millimeter wave MIMO systems. *Phys. Commun.* **2020**, *39*, 100997.
14. Xu, K.; Cai, Y.; Zhao, M.; Niu, Y.; Hanzo, L. MIMO-Aided Nonlinear Hybrid Transceiver Design for Multiuser Mmwave Systems Relying on Tomlinson-Harashima Precoding. *IEEE Trans. Veh. Technol.* **2021**, *70*, 6943–6957.
15. Pavia, J.; Velez, V.; Ferreira, R.; Souto, N.; Ribeiro, M.; Silva, J.; Dinis, R. Low Complexity Hybrid Precoding Designs for Multiuser mmWave/THz Ultra Massive MIMO Systems. *Sensors* **2021**, *21*, 6054.
16. Payami, S.; Khalily, M.; Araghi, A.; Loh, T.; Cheadle, D.; Nikitopoulos, K.; Tafazolli, R. Developing the First mmWave Fully-Connected Hybrid Beamformer With a Large Antenna Array. *IEEE Access* **2020**, *8*, 141282–141291.
17. Loh, T.; Cheadle, D.; Payami, S.; Khalily, M.; Nikitopoulos, K.; Tafazolli, R. Experimental Evaluation of a Millimeter-wave Fully-Connected Hybrid Beamformer with a Large Antenna Array. *2021 15th European Conference on Antennas and Propagation (EuCAP)* **2021**.
18. Li, X.; Yu, J.; Wang, K.; Kong, M.; Zhou, W.; Zhu, Z.; Wang, C.; Zhao, M.; Chang, G. 120 Gb/s Wireless Terahertz-Wave Signal Delivery by 375 GHz-500 GHz Multi-Carrier in a 2×2 MIMO System. *Journal of Lightwave Technology* **2019**, *37*, 606–611.
19. Jia, S.; Zhang, L.; Wang, S.; Li, W.; Qiao, M.; Lu, Z.; Idrees, N.; Pang, X.; Hu, H.; Zhang, X.; Oxenlowe, L.; Yu, X. 2×300 Gbit/s Line Rate PS-64QAM-OFDM THz Photonic-Wireless Transmission. *Journal of Lightwave Technology* **2020**, *38*, 4715–4721.
20. Busari, S.; Huq, K.; Mumtaz, S.; Rodriguez, J. Terahertz Massive MIMO for Beyond-5G Wireless Communication. In Proceedings of the ICC 2019-2019 IEEE International Conference on Communications (ICC), Shanghai, China, 11–22 March **2019**.
21. Akyildiz, I.F.; Han, C.; Hu, Z.; Nie, S.; Jornet, J.M. TeraHertz Band Communication: An Old Problem Revisited and Research Directions for the Next Decade. *arXiv* **2021**, arXiv: 2112.13187v1. Available online: <https://arxiv.org/abs/2112.13187v1> (accessed on 2 July 2021).
22. Godinho, A.; Fernandes, D.; Soares, G.; Pina, P.; Sebastião, P.; Correia, A.; Ferreira, L.S. A Novel Way to Automatically Plan Cellular Networks Supported by Linear Programming and Cloud Computing. *Appl. Sci.* **2020**, *10*, 3072.
23. Jao, C. et al., "WiSE: A System-Level Simulator for 5G Mobile Networks," in *IEEE Wireless Communications* **2018**, *25*, 2, 4–7.
24. Riviello, D.G.; Di Stasio, F.; Tuninato, R. Performance Analysis of Multi-User MIMO Schemes under Realistic 3GPP 3-D Channel Model for 5G mmWave Cellular Networks. *Electronics* **2022**, *11*, 330.
25. 3rd Generation Partnership Project (3GPP). TR 38.901 v14.1.1, Study on channel model for frequencies from 0.5 to 100 GHz, Release 14. July **2017**. Available online: www.3gpp.org/ftp/Specs/archive/38_series/38.901/38901-g10.zip (accessed on 01 November 2021).
26. Bechta, K.; Kelner, J.; Ziolkowski, C.; Nowosielski, L. Inter-Beam Co-Channel Downlink and Uplink Interference for 5G New Radio in mm-Wave Bands. *Sensors* **2021**, *21*, 793.
27. Zaidi, S.; Ben Smida, O.; Affes, S.; Vilaipornsawai, U.; Zhang, L. and Zhu, P. User-Centric Base-Station Wireless Access Virtualization for Future 5G Networks. *IEEE Transactions on Communications*. **2019**, *67*, 7, 5190–5202.
28. Yuan, H.; Yang, N.; Yang, K.; Han, C.; An, J. Hybrid Beamforming for Terahertz Multi-Carrier Systems Over Frequency Selective Fading. *IEEE Trans. Commun.* **2020**, *68*, 6186–6199.

29. Yan, L.; Han, C.; Yuan, J. A Dynamic Array-of-Subarrays Architecture and Hybrid Precoding Algorithms for Terahertz Wireless Communications. *IEEE J. Sel. Areas Commun.* **2020**, *38*, 2041–2056.
30. Han, C.; Yan, L.; Yuan, J. Hybrid Beamforming for Terahertz Wireless Communications: Challenges, Architectures, and Open Problems. *IEEE Wireless Communications* **2021**, *28*, 198–204.
31. Ning, B.; Zhongbao, T.; Chen, Z.; Han, C.; Yuan, J.; Li, S. TeraMIMO Prospective Beamforming Technologies for Ultra-Massive MIMO in Terahertz Communications: A Tutorial. arXiv **2021**, arXiv: 2107.03032v2. Available online: <https://arxiv.org/abs/2107.03032v2> (accessed on 2 January 2022).
32. Basar, E.; Yildirim, I. SimRIS Channel Simulator for Reconfigurable Intelligent Surface-Empowered Communication Systems, 2020 IEEE Latin-American Conference on Communications (LATINCOM), 2020, pp. 1–6.
33. Ju, S.; Xing, Y.; Kanhere, O.; Rappaport, T. S. Millimeter Wave and Sub-Terahertz Spatial Statistical Channel Model for an Indoor Office Building. arXiv 2021, arXiv:2103.17127v1. Available online: <https://doi.org/10.48550/arXiv.2103.17127> (accessed on 26 February 2022).
34. Yu, X., Zhang, J. and Letaief, K., "Alternating minimization for hybrid precoding in multiuser OFDM mmWave systems", 2016 50th Asilomar Conference on Signals, Systems and Computers, **2016**.
35. Yu, X., Zhang, J. and Letaief, K. B., "Doubling Phase Shifters for Efficient Hybrid Precoder Design in Millimeter-Wave Communication Systems", *Journal of Communications and Information Networks* **2019**, *4*, pp. 51–67.