



UNIVERSIDADE  
ESTADUAL DE LONDRINA

---

JAIME LAELSON JACOB

**EFICIÊNCIA DE RECURSOS SOB EQUIDADE DE TAXA EM  
SISTEMAS NOMA, MIMO-NOMA E LSA-MIMO PDMA**

---

Londrina  
2022

JAIME LAELSON JACOB

**EFICIÊNCIA DE RECURSOS SOB EQUIDADE DE TAXA EM  
SISTEMAS NOMA, MIMO-NOMA E LSA-MIMO PDMA**

Tese de Doutorado apresentada ao Programa de Pós-Graduação Associado em Engenharia Elétrica UEL/UTFPR-CP, da Universidade Estadual de Londrina em cumprimento aos requisitos para a obtenção do grau de Doutor em Engenharia Elétrica. Área de Concentração: Sistemas de Telecomunicações

Orientador: Professor Dr. Taufik Abrão

Londrina  
2022

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UEL

J15e      Jacob, Jaime Laelson.  
Eficiência de Recursos Sob Equidade de Taxa em Sistemas NOMA, MIMO-NOMA e LSA-MIMO PDMA / Jaime Laelson Jacob. - Londrina, 2022. 209 f. : il.

Orientador: Taufik Abrão.  
Dissertação (Mestrado em Engenharia Elétrica) - Universidade Estadual de Londrina, Centro de Tecnologia e Urbanismo, Programa de Pós-Graduação em Engenharia Elétrica, 2022.  
Inclui bibliografia.

1. Acesso Múltiplo Não-Ortogonal - Tese. 2. Múltipla Entrada Múltipla Saída - Tese. 3. Eficiência Energética - Tese. 4. Eficiência Espectral - Tese. I. Abrão, Taufik . II. Universidade Estadual de Londrina. Centro de Tecnologia e Urbanismo. Programa de Pós-Graduação em Engenharia Elétrica. III. Título.

CDU 62

JAIME LAELSON JACOB

**EFICIÊNCIA DE RECURSOS SOB EQUIDADE DE TAXA EM  
SISTEMAS NOMA, MIMO-NOMA E LSA-MIMO PDMA**

Tese de Doutorado apresentada ao Programa de Pós-Graduação Associado em Engenharia Elétrica UEL/UTFPR-CP, da Universidade Estadual de Londrina em cumprimento aos requisitos para a obtenção do grau de Doutor em Engenharia Elétrica. Área de Concentração: Sistemas de Telecomunicações

**BANCA EXAMINADORA**

---

Orientador: Prof. Dr. Taufik Abrão  
Universidade Estadual de Londrina - UEL

---

Prof. Dr. Flávio Henrique Teles Vieira  
Universidade Federal de Goiás – UFG

---

Prof. Dr. Francisco Rafael Marques Lima  
Universidade Federal do Ceará – UFC

---

Prof. Dr. José Carlos Marinello Filho  
Universidade Tecnológica Federal do Paraná -  
UTFPR - Campus Cornélio Procópio

---

Profª. Dra. Diana Pamela Moya Osorio  
University of Oulu

Londrina, 29 de novembro de 2022.

# Agradecimentos

Agradeço ao professor, orientador e colega de trabalho, professor Doutor Taufik Abrão, pela confiança, pelo apoio e pelas contribuições durante a realização do trabalho.

Aos professores que fizeram parte da minha banca de Qualificação e Defesa agradeço por todas as relevantes contribuições para a melhoria do trabalho.

Agradeço ao Professor Doutor Cristiano Magalhaes Panazio da Escola Politécnica da Universidade de São Paulo, à professora Doutora Cristiane A. Pendeza Martinez e ao professor Doutor André Luis Machado Martinez ambos da Universidade Tecnológica Federal do Paraná de Cornélio Procópio pela ajuda na pesquisa com geração de co-autoria de artigos publicados, frutos deste trabalho.

Agradeço aos colegas do laboratório Telecom & DSP-UEL que de alguma forma ajudaram na pesquisa.

Agradeço a Ingrid Hildegard Vogler pela ajuda nos textos e pelo apoio e incentivo que contribuíram em muito para a conclusão desse trabalho.

Agradeço aos meus filhos Adolfo Lincon Jacob, Gabriel Angelo Jacob e Caroline Vieira Jacob pelo inestimável apoio emocional e carinho.

# Resumo

Analisa-se neste trabalho o compromisso entre a eficiência energética (EE - *Energy Efficiency*) e eficiência espectral (SE - *Spectral Efficiency*) de múltiplo acesso não ortogonal (NOMA - *Non-Orthogonal Multiple Access*), sistemas de múltiplas entradas e múltiplas saídas (MIMO - *Multiple Input Multiple Output*) combinados aos esquemas NOMA e MIMO-NOMA, bem como ao esquema de acesso múltiplo por divisão de padrões (PDMA - *Pattern Division Multiple Acces*) e ao MIMO massivo (mMIMO - *Massive MIMO*) em enlaces de descida (DL - *Downlink*) sob a restrição de equidade de taxa. Este compromisso entre EE e SE, denominado eficiência de recursos (RE - *Resource Efficiency*) é analisado para diferentes números de usuários por célula no NOMA. Obtêm-se o número de usuários por *cluster* na configuração MIMO-NOMA que resulta na máxima EE a partir da equidade de máxima taxa. Combinamos também o sistema mMIMO ao PDMA para oferecer suporte eficiente aos serviços de banda larga móvel avançada (eMBB - *Enhanced Mobile Broadband*), Comunicação Massiva do Tipo Máquina (mMTC - *Massive Machine Type Communication*), e Comunicações Ultra-Confiáveis e de Baixa Latência (uRLLC - *Ultra-Reliable and Low Latency Communications*) em sistemas da quinta geração (5G - *Fifth Generation*) e além da 5G (B5G - *Beyond 5G*); seleciona-se matrizes padrão que melhor atendem os requisitos heterogêneos desses serviços.

A primeira parte desse trabalho trata da otimização da proporção de potência alocada para cada usuário nos sistemas NOMA. Na segunda parte, investigamos o compromisso EE-SE em sistemas MIMO-NOMA, considerando o DL, sob a restrição de mesma taxa entre os usuários. Essa restrição visa garantir a máxima equidade de taxa para todos os usuários ativos no sistema. Na terceira parte, o compromisso SE-EE foi formulado como um problema de otimização multi-objetivo (MOO - *Multi-Objective Optimization*) para o sistema MIMO-NOMA mantendo-se as restrições definidas anteriormente. O problema de otimização foi resolvido primordialmente adotando-se o método de escalarização ( $\epsilon$ -C -  $\epsilon$ -Constraint) combinado a técnicas de programação não linear (NLP - *Nonlinear Programming*), particularmente a programação quadrática sequencial (SQP- *Sequential Quadratic Programming*) e o método de Dinkelback (DK). Finalmente, foram analisadas e comparadas várias matrizes PDMA disponíveis na literatura com características favoráveis ao atendimento dos requisitos impostos pelos diferentes serviços 5G, com uso da EE e outras métricas para a seleção da matriz padrão que melhor atende aos quesitos conflitantes em sistemas mMIMO PDMA.

Resultados numéricos revelam que é possível identificar a melhor matriz padrão do sistema mMIMO PDMA que atenda requisitos conflitantes de cada um dos modos de uso 5G separadamente ou simultaneamente. No sistema NOMA, pode-se encontrar o ponto de equilíbrio entre EE e a soma de taxas para cada cenário do sistema, bem como o melhor

ponto de operação de RE. Mostra-se ainda que o número total de usuários atinge a máxima EE em cada configuração *cluster*-usuários analisada. Resultados para a relação EE-SE no sistema MIMO-NOMA demonstraram a habilidade do método  $\epsilon$ -C em encontrar diversidade de soluções na fronteira de Pareto.

**Palavras Chaves:** Múltiplo Acesso Não Ortogonal (NOMA), Múltiplas Entradas e Múltiplas Saídas (MIMO), Eficiência Espectral (SE), Eficiência Energética (EE), Eficiência de Recurso (RE), Acesso Múltiplo por Divisão de Padrões (PDMA), Equidade de Taxa.

# Abstract

This work analyzes the tradeoff between energy efficiency (EE) and spectral efficiency (SE) of non-orthogonal multiple access (NOMA), multiple input and multiple output systems (MIMO) combined with NOMA and MIMO-NOMA schemes, as well as pattern division multiple access (PDMA) scheme and massive MIMO (mMIMO) on downlinks (DL) under rate fairness constraint. This compromise between EE and SE, called resource efficiency (RE) is analyzed for different numbers of users per cell in NOMA. The number of users per cluster in the MIMO-NOMA configuration that results in maximum EE is obtained from maximum rate fairness. We also combine the mMIMO system with PDMA to efficiently support enhanced mobile broadband (eMBB), massive machine type communication (mMTC) and Ultra-Reliable and Low Latency Communications (uRLLC) services in fifth generation (5G) and beyond 5G (B5G) systems. We selected pattern matrices from the literature that best meet the heterogeneous requirements of these services.

The first part of this work deals with the optimization of the proportion of power allocated to each user in NOMA systems. In the second part, we investigate the EE-SE compromise in MIMO-NOMA systems, considering the DL, under the same rate constraint for all users. This restriction is intended to ensure maximum rate fairness for all active users in the system. In the third part, the SE-EE compromise was formulated as a multi-objective optimization (MOO) problem for the MIMO-NOMA system, keeping the previously defined constraints. The optimization problem was solved primarily by adopting the scalarization method ( $\epsilon$ -C –  $\epsilon$ -Constraint) combined with non-linear programming (NLP) techniques, particularly the sequential quadratic programming (SQP) and the Dinkelback (DK) method. Finally, several PDMA matrices available in the literature were analyzed and compared with favorable characteristics to meet the requirements imposed by different 5G services, using EE and other metrics for the selection of the pattern matrix that best meets the conflicting questions in mMIMO PDMA systems.

Numerical results reveal that it is possible to identify the best mMIMO PDMA system pattern matrix that meets conflicting requirements of each of the 5G usage modes separately or simultaneously. In the NOMA system, one can find the balance point between EE and the sum rate for each system scenario, as well as the best RE operating point. It is also shown that the total number of users reaches the maximum EE in each cluster-users configuration analyzed. Results for the EE-SE relation in the MIMO-NOMA system demonstrated the ability of the  $\epsilon$ -C method to find diversity of solutions on the Pareto front.

**Keywords:** Non-Orthogonal Multiple Access (NOMA), Multiple Input and Multiple Output (MIMO), Spectral Efficiency (SE), Energy Efficiency (EE), Resource Efficiency (RE), Successive Interference Cancellation (SIC), Rate Fairness.

# Lista de figuras

Figura 1.1 – Classificação dos esquemas de múltiplo acesso existentes. . . . .	3
Figura 1.2 – Exemplo de padrão PDMA. . . . .	4
Figura 2.1 – NOMA com $M$ usuários e de raio $\mathfrak{R}_D$ . . . . .	17
Figura 2.2 – Conjunto fechado para $\sigma_n^2 = 1\mu$ W e $M = 4$ usuários com distâncias $d_1 = 1,0$ km, $d_2 = 0,7$ km, $d_3 = 0,5$ km e $d_4 = 0,3$ km. . . . .	21
Figura 3.1 – Sistema celular com uma BS com $A$ antenas, $A \cdot K$ usuários com $A$ <i>clusters</i> , $K$ usuários/ <i>clusters</i> e cada usuário tem $U$ antenas ( $U \geq A$ ). . . . .	25
Figura 3.2 – Conjunto fechado para $\alpha = 3$ , $\sigma_n^2 = 1\mu$ W, $A = 2$ <i>clusters</i> e $K = 3$ usuários/ <i>cluster</i> com distâncias $d_{11} = 0,08$ km, $d_{12} = 0,15$ km, $d_{13} = 0,28$ km, $d_{21} = 0,07$ km, $d_{22} = 0,15$ km e $d_{23} = 0,27$ km. . . . .	31
Figura 3.3 – Conjunto fechado de soluções $\beta_{a,k}$ 's $\times$ EE $\times R^*$ , considerando os valores de parâmetros: potência $P_{BS} = 1$ W/antena BS e $P_K = 100$ mW/antena UE; $\alpha = 2$ ; $\sigma_n^2 = 1\mu$ W, $A = 2$ <i>clusters</i> e $K = 3$ UE's/ <i>cluster</i> , com distâncias $d_{11} = 80$ m, $d_{12} = 150$ m, $d_{13} = 280$ m, $d_{21} = 70$ m, $d_{22} = 150$ m e $d_{23} = 270$ m. . . . .	35
Figura 4.1 – Esquema de transmissão mMIMO PDMA. Cada AC é equipado com $N_T$ antenas formando $N$ feixes, com $N_T \geq N$ ; ademais, cada UG contém $K$ usuários, sendo $N_T \leq KN_R$ e o filtro espacial (SF - <i>spatial filter</i> ) $\mathbf{v}_{nk}$ para cada UG definido na seção 4.1.4. . . . .	45
Figura 4.2 – Resultado da soma das combinações da constelação de dois usuários. (a) e (b) representam os resultados para a modulação QPSK, enquanto (c) e (d) para a modulação 16-QAM. Temos (a) e (c) sem as correções de ambiguidade e (b) e (d) com as correções de ambiguidade ao aplicarmos a mudança de fase $\varphi = \pi/4$ e fator escalar de potência $\mu = 0.5$ em (4.23). . . . .	50
Figura 5.1 – Distribuição de usuários em um sistema NOMA com 4 usuários. . . . .	59
Figura 5.2 – Taxas de dados ( $R$ ) e soma das taxas ( $R_\Sigma$ ) <i>versus</i> potência alocada considerada perda de percurso ( $\alpha = 2$ ). As taxas não ótimas foram obtidas com a proporção de potência dada pela Eq. (5.1). . . . .	61
Figura 5.3 – $R_\Sigma$ <i>versus</i> $P$ para $\alpha = 2$ e 3. As taxas não ótimas foram obtidas com a proporção de potência dada pela Eq. (5.1). . . . .	62
Figura 5.4 – Proporção de potência ótima e comportamento da soma das taxas para os três cenários e diferentes expoentes de perda de percurso: a) $\beta_{opt}^* \times R$ ; b) $\beta_{opt}^* \times P_{opt}^*$ e $R \times P_{opt}^*$ . . . . .	63
Figura 5.5 – Os pontos ótimos considerando $\alpha = 2$ . . . . .	65
Figura 5.6 – Compromisso entre EE e SE para alguns Raios $\mathfrak{R}_D$ e $\alpha = 2$ . . . . .	66
Figura 5.7 – Compromisso entre EE e SE para alguns Raios $\mathfrak{R}_D$ e $\alpha = 3$ . . . . .	66

Figura 5.8 – Compromisso entre EE e SE para alguns Raios $\mathfrak{R}_D$ e $\alpha = 4, 5$ . . . . .	67
Figura 5.9 – Distribuição de usuários em um sistema NOMA com 6 usuários conforme Tabela 5.4. . . . .	68
Figura 5.10–As taxas de bits <i>versus</i> potência alocada considerando perda de percurso LoS ( $\alpha = 2$ ) e 6 usuários. As taxas de ICR foram obtidas com a proporção de potência dada pela Eq. (3.18) e taxas ERPA pelo problema (3.19). (a) Cenário 1, Figura 5.9.(a); (b) Cenário 1, Figura 5.9.(b); (c) Cenário 2, Figura 5.9.(c); e (d) Cenário 2, Figura 5.9.(d). . . . .	70
Figura 5.11– $R_\Sigma$ <i>versus</i> $P$ and $\frac{R_\Sigma \text{ERPA}}{R_\Sigma \text{ICR}}$ <i>versus</i> $P_{opt}^*$ to $\alpha = 2$ and 3. A proporção de potência foi obtida pela distribuição ICR, Eq. (3.18), e a distribuição ERPA pelo problema (3.19). . . . .	71
Figura 5.12– $\beta_{ak}^*$ <i>versus</i> taxa de dados, problema (3.19). . . . .	72
Figura 5.13– $EE_{opt}^* \times R_\Sigma$ , $P_{opt}^* \times R_\Sigma$ and $EE_{opt}^* \times P_{opt}^*$ , problema (3.25) para $\alpha = 2$ e 6 usuários. A distribuição ERPA (solução para o problema (3.23) foi destacada com marcadores de estrela verde. . . . .	74
Figura 5.14– $EE \times R_\Sigma$ e $P \times R$ (equidade) para 2, 3 e 4 usuários/ <i>cluster</i> , $\alpha = 2, 3$ e 3, 5 e $\mathfrak{R}_D = 50$ m, problema (3.23). . . . .	76
Figura 5.15– $EE \times R_\Sigma$ e $P \times R$ (equidade) para 2, 3 e 4 usuários/ <i>cluster</i> , $\alpha = 2, 3$ e 3, 5 e $\mathfrak{R}_D = 100$ m, problema (3.23). . . . .	78
Figura 5.16– $EE \times R_\Sigma$ e $P \times R$ (equidade) para 2, 3 e 4 usuários/ <i>cluster</i> , $\alpha = 2, 3$ e 3, 5 e $\mathfrak{R}_D = 200$ m, problema (3.23). . . . .	78
Figura 5.17–Comparação do tempo decorrido e NMSE alcançado pelos três métodos BOO propostos $\epsilon$ –C-SQP, $\epsilon$ –C-DK, e $\epsilon$ –C-HC, sob convergência total considerando três cenários diferentes (valores de $r_D$ ). . . . .	84
Figura 5.18–Frente de Pareto obtida pelos três métodos de BOO $\epsilon$ –C-SQP, $\epsilon$ –C-DK e $\epsilon$ –C-HC para diferentes cenários (valores de $r_D$ ); expoentes de perda de percurso $\alpha = 3$ e $\alpha = 3, 5$ . . . . .	85
Figura 5.19– $EE \times SE$ obtido para o método $\epsilon$ –C-SQP sob diferentes valores de raio de célula $r_D$ , considerando expoentes de perda de percurso $\alpha = 2, 3$ e 3, 5. . . . .	86
Figura 5.20–Compromisso entre EE-SE e orçamento de potência considerando o método $\epsilon$ -C-SQP; 2 usuários/ <i>cluster</i> ; $\alpha = 3$ e $r_D = 50$ m. . . . .	87
Figura 5.21–EE máximo extraído da frente de Pareto contra SE para o método $\epsilon$ –C-SQP. Raio da célula $r_D = 100$ m; diferente número de usuários para 2, 3 e 4 usuários/ <i>cluster</i> e expoentes de perda de percurso $\alpha$ . . . . .	87
Figura 5.22–Frente de Pareto obtida pelos dois métodos BOO $\epsilon$ –C-SQP e WS-SQP para diferentes cenários (valores $r_D$ ); expoente de perda de percurso $\alpha = 3$ . . . . .	89

Figura 5.23–Cenários A, B e C (máxima-equidade) para o problema de otimização de EE em DL, Eq. (4.45). . . . .	90
Figura 5.24– $EE^* \times R$ para $\alpha = 3$ . Os valores ótimos de EE foram obtidos resolvendo Eq. (4.45). . . . .	94
Figura 5.25–Valores obtidos de $EE^*$ máximo e a respectiva igual taxa de bits $R$ resolvendo o problema de otimização (4.45) sob um número fixo de usuários $K$ de acordo com cada matriz padrão. <u>Cenário A</u> foi adotado para avaliar a matriz padrão Modelo 10; <u>Cenário B</u> os Modelos 1, 2, 4, 5, 7, 8, 11 e 14; <u>Cenário C</u> os Modelos 3, 6, 9, 12, 13 e Modelo 15. . . .	95

# Lista de tabelas

Tabela 5.1 – Parâmetros Adotados para os cenários de Simulação. . . . .	59
Tabela 5.2 – Equidade $\mathcal{F}_{it}$ obtida da Eq. (2.17). . . . .	62
Tabela 5.3 – Ponto de Compromisso entre EE e $R_{\Sigma}$ Para alguns $\mathfrak{R}_D$ e Usuários com Distribuição Uniforme. . . . .	66
Tabela 5.4 – Parâmetros Adotados para os cenários de Simulação. . . . .	69
Tabela 5.5 – Distribuição ERPA para $\alpha = 2$ e seis usuários obtida de acordo com os gráficos da Figura 5.13. . . . .	75
Tabela 5.6 – Combinação de $K$ usuários por <i>cluster</i> com $A$ <i>clusters</i> que resulta em $KA$ usuários em ordem decrescente de ganho de canal. . . . .	77
Tabela 5.7 – Parâmetros Adotados Para os Cenários de Simulação. . . . .	79
Tabela 5.8 – Figura de mérito para avaliação de desempenho dos algoritmos EE-SE MIMO-NOMA; expoente de perda de percurso $\alpha = 3$ . . . . .	82
Tabela 5.9 – Número ótimo de usuários $M^*$ para atingir o EE ótimo, considerando o raio da célula $r_D = 100$ m, número diferente de usuários/ <i>cluster</i> $K$ e usuários por célula $M$ . . . . .	86
Tabela 5.10–Figura de mérito obtida com o método WS-SQP para o problema EE-SE MIMO-NOMA; expoente de perda de percurso $\alpha = 3$ . . . . .	88
Tabela 5.11–Parâmetros Adotados Para os Cenários de Simulação. . . . .	90
Tabela 5.12–Matrizes padrão do projeto PDMA de diferentes dimensões (SUN et al., 2020), (REN et al., 2016), (LI et al., 2021). . . . .	91
Tabela 5.13–Parâmetros a serem analisados para os serviços eMBB. . . . .	93
Tabela 5.14–Parâmetros a serem analisados para os serviços mMTC. . . . .	94
Tabela 5.15–Parâmetros a serem analisados para os serviços uRLLC. . . . .	95
Tabela 5.16–Parâmetros a serem analisados para os serviços híbridos eMBB e uRLLC. . . . .	96

# Lista de abreviaturas e siglas

1G	- Primeira Geração - <i>First-Generation</i> .
2G	- Segunda Geração - <i>Second-Generation</i> .
3G	- Terceira Geração - <i>Third-Generation</i> .
4G	- Quarta Geração - <i>Fourth-Generation</i> .
5G	- Quinta Geração - <i>Fifth Generation</i> .
AC	- Aglomerados de Antenas - <i>Antenna Clusters</i>
ALM	- Método Lagrangiano Aumentado - <i>Augmented Lagrangian Method</i> .
B5G	- Além da 5G - <i>Beyond 5G</i>
AWGN	- Ruído Gaussiano Complexo Branco Aditivo - <i>Additive White Complex Gaussian Noise</i> .
BDM	- Multiplexação por Divisão de Bits - <i>Bit Division Multiplexing</i> .
BOO	- Otimização Bi-Objetiva - <i>Bi-Objective Optimization</i> .
BPSK	- Chaveamento de Deslocamento de Fase Binária - <i>Binary Phase Shift Keying</i> .
BS	- Estação Rádio Base - <i>Base Station</i> .
CDMA	- Acesso Múltiplo por Divisão de Código - <i>Code-Division Multiple Access</i> .
CSI	- Informação de Estado do Canal - <i>Channel State Information</i> .
DL	- Enlace de Descida - <i>Downlink</i> .
EE	- Eficiência Energética - <i>Energy Efficiency</i> .
eMBB	- Banda Larga Móvel Avançada - <i>Enhanced Mobile Broadband</i> .
ERPA	- Alocação de Potência Ótima com Taxa Igual - <i>Equal-Rate optimal Power Allocation</i> .
GA	- Algoritmo Genético - <i>Genetic Algorithm</i> .
HC	- Escalonado - <i>Hill Climbing</i> .
ICR	- Inverso da Resposta do Canal - <i>Inverse of the Channel Response</i> .
IPI	- Interferência entre Padrões - <i>Inter-Patterns Interference</i> .
LDS	- Espalhamento de Baixa Densidade - <i>Low-Density Spreading</i> .

LSA	- Antena de Grande Escala <i>Large-Scale Antenna</i> .
MC-NOMA	- Múltiplo Acesso Não Ortogonal de Múltiplas Portadoras - <i>Multi-Carrier Non-Orthogonal Multiple Access</i> .
MIMO	- Múltiplas Entradas e Saídas - <i>Multiple Input Multiple Output</i> .
mMTC	- Comunicação Massiva do Tipo Máquina - <i>Massive Machine Type Communication</i> .
MOO	- Otimização Multi-Objetivo - <i>Multi-Objective Optimization</i> .
MRW	- Peso Máximo da linha - <i>Maximum Row Weight</i> .
MUSA	- Acesso Compartilhado Multiusuário - <i>Multi-user Shared Access</i> .
NCMINP	- Programação Não Linear Inteira Mista Não Convexa - <i>Non-Convex Mixed-Integer Nonlinear Programming</i> .
NLP	- Programação Não Linear - <i>Nonlinear Programming</i> .
NOMA	- Múltiplo Acesso Não Ortogonal - <i>Non-Orthogonal Multiple Access</i> .
NP	- Polinômio Não Determinístico - <i>Nondeterministic Polynomial</i> .
OCDMA	- Múltiplo Acesso por Divisão de Código Óptico - <i>Optical Code Division Multiple Access</i> .
OF	- Fator de Sobrecarga - <i>Overload Factor</i> .
OFDMA	- Acesso por Multiplexação por Divisão de Frequência Ortogonal - <i>Orthogonal Frequency Division Multiplexing Access</i> .
OMA	- Múltiplo Acesso Ortogonal - <i>Orthogonal Multiple Access</i> .
PA	- Alocação de Potência - <i>Power Allocation</i> .
PDMA	- Múltiplo Acesso por Divisão de Padrão - <i>Pattern Division Multiple Access</i> .
PEF	- Fator de Avaliação de Desempenho - <i>Performance Evaluation Factor</i> .
PSMA	- Múltiplo Acesso no Domínio da Potência e do Código Esparsa - <i>Power Domain Sparse Code Multiple Access</i> .
QAM	- Modulação em Amplitude de Quadratura - <i>Quadrature amplitude modulation</i> .
QoS	- Qualidade de Serviço - <i>Quality of Service</i> .
QPSK	- Chaveamento de Deslocamento de Fase em Quadratura - <i>Quadrature Phase Shift Keying</i> .
RE	- Eficiência de Recursos - <i>Resource Efficiency</i> .

RF	- Rádio Frequência - <i>Radio Frequency</i> .
RV	- Valores de Recursos - <i>Resource Values</i> ).
SCMA	- Acesso Múltiplo por Código Esperso - <i>Sparse Code Multiple Access</i> .
SCUs	- Usuários de Pequenas Células - <i>Small Cell Users</i> .
SE	- Eficiência Espectral - <i>Spectral Efficiency</i> .
SIC	- Cancelamento de Interferência Sucessiva - <i>Successive Interference Cancellation</i> .
SIMO	- Entrada Única e Múltipla Saída - <i>Single Input and Multiple Output</i> .
SINR	- Relação entre o Sinal e Sinal Interferente mais Ruído - <i>Signal-to-Interference-plus-Noise Ratio</i> .
SOO	- Otimização de Objetivo Único - <i>Single-Objective Optimization</i> .
SPGA	- Algoritmo de Geração de População Superior - <i>Superior Population Generation Algorithm</i> .
SQP	- Programação Quadrática Sequencial - <i>Sequential Quadratic Programming</i> .
UEs	- Equipamentos do Usuário - <i>User Equipments</i> .
UL	- Enlace de Subida - <i>Uplink</i> .
UG	- Grupo de Usuários - <i>User Group</i> .
uRLLC	- Comunicações Ultra-Confíáveis e de Baixa Latência - <i>Ultra-Reliable and Low Latency Communications</i> .
WS	- Soma Ponderada - <i>Weighted Sum</i> .
ZFBF	- Formação de Feixe Forçando a Zero - <i>Zero-Forcing Beam-Forming</i> .

# Lista de símbolos

## Sistema NOMA

$\mathcal{R}_D$	Raio da Célula.
$M$	Número de Usuários na Célula.
$g_i$	Ganho do Canal com Desvanecimento Rayleigh do $i$ -ésimo Usuário.
$\alpha$	Expoente de Perda de Percurso.
$d_i$	Distância do $i$ -ésimo Usuário à BS.
$h_i$	Canal com Desvanecimento Rayleigh e a Perda de Percurso.
$x$	Sinal de Transmissão pelo Usuário na BS com Aplicação do NOMA.
$s_i$	Mensagem do $i$ -ésimo Usuário.
$P$	Potência Total de Transmissão.
$\beta_i$	Coefficiente de Alocação de Potência do $i$ -ésimo Usuário.
$z_i$	Ruído Gaussiano Branco Aditivo Complexo.
$y_m$	Sinal Recebido do $m$ -ésimo Usuário.
$\sigma_n^2$	Variância do Ruído AWGN.
$\gamma_m$	Relação SINR do $m$ -ésimo Usuário ( $1 \leq m \leq (M - 1)$ ).
$\gamma_M$	Potência SINR do Último Usuário Próximo da BS.
$R_m$	Taxa de Transmissão do $m$ -ésimo Usuário.
$R_s$	Soma das Taxas Transmitidas por Todos os Usuários.
EE	Eficiência Energética.
$R$	Taxa Única para Todos os Usuários.
$P_t$	Potência Total de RF.
$\rho$	Ineficiência do Amplificador de Potência RF.
$P_c$	Potência do Circuito no Transmissor.
RE	Eficiência de Recursos.

SE	Eficiência Espectral.
$\xi_0$	Fator de Ponderação em [W] Controla os Pesos de EE e SE.
$\mathcal{F}_j$	Índice de Equidade de Jain.
$\mathcal{F}_{it}$	Índice de Equidade Baseada na Teoria da Informação.
$\Delta_m^2$	Quadrado dos Desvios das Taxas do $m$ -ésimo Usuário.
$\hat{\Delta}^2$	Estimativa dos Desvios da Taxa Quadrada Média.
$\hat{R}^2$	Taxa Quadrada Média.
$R_m^f$	Taxa de Equidade para o $m$ -ésimo Usuário.
$R_c$	Taxa Cumulativa.

### Sistema MIMO-NOMA

$A$	Número de Antenas na BS.
$K$	Número de Usuários por <i>Cluster</i> .
$U$	Número de Antenas por Usuário.
$\mathbf{x}$	Vetor dos Sinais Codificados Transmitidos pelos Usuários na BS.
$s_{A,K}$	Sinal de Transmissão do $K$ -ésimo Usuário do $A$ -ésima <i>Cluster</i> .
$s_A$	Soma dos Sinais Transmitidos pela BS aos Usuários do $A$ -ésima <i>Cluster</i> .
$\hat{\mathbf{s}}$	Vetor dos Sinais Transmitidos pela BS.
$\mathbf{W}$	Matriz de Pré-Codificação.
$\mathbf{I}_A$	Matriz Identidade $A \times A$ .
$\beta_{a,k}$	Coefficiente de Proporção de Potência do $a$ -ésimo <i>Cluster</i> e $k$ -ésimo Usuário.
$z_{a,k}$	Vetor AWGN do $a$ -ésimo <i>Cluster</i> e $k$ -ésimo Usuário.
$d_{a,k}$	Distância do $a$ -ésimo <i>Cluster</i> e $k$ -ésimo Usuário.
$\mathbf{w}_\ell$	Vetor da $\ell$ -ésima Coluna da Matriz $\mathbf{W}$ .
$\mathbf{Q}_{a,k}$	Matriz de Canal de Desvanecimento Rayleigh do $a$ -ésimo <i>Cluster</i> e $k$ -ésimo Usuário.

$\mathbf{H}_{a,k}$	Matriz de Canal com Desvanecimento e Perda de Percurso do $a$ -ésimo <i>Cluster</i> e $k$ -ésimo Usuário.
$\mathbf{y}_{a,k}$	Vetor do Sinal Recebido do $a$ -ésimo <i>Cluster</i> e $k$ -ésimo Usuário.
$\mathbf{v}_{a,k}$	Vetor de Detecção do $a$ -ésimo <i>Cluster</i> e $k$ -ésimo Usuário.
$\mathbf{h}_{j,ak}$	Vetor da $j$ -ésima Coluna de $\mathbf{H}_{a,k}$ .
$\bar{\mathbf{H}}_{a,k}$	Matriz $\mathbf{H}_{a,k}$ sem a Coluna $\mathbf{h}_{a,ak}$ .
$\mathbf{U}_{a,k}$	Matriz dos Vetores Singulares à Esquerda de $\bar{\mathbf{H}}_{a,k}$ após Decompor em Valores Singulares.
$\mathbf{g}_{a,k}$	Vetor de Combinação de Proporção Máxima.
$\gamma_{a,k}$	SINR para o $k$ -ésimo Usuário, $a$ -ésimo <i>Cluster</i> ( $1 \leq a \leq A$ e $2 \leq k \leq K$ ).
$\gamma_{a,1}$	Relação SINR para o Primeiro Usuário no $a$ -ésimo <i>Cluster</i> .
$R_{a,k}$	Taxa de Dados do $a$ -ésimo <i>Cluster</i> e $k$ -ésimo Usuário.
$R_{\Sigma}$	Soma da Taxa de Dados Transmitidas pelos Usuários.
$P_K$	Potência de Operação do Circuito de Cada Antena de Usuário.
$P_{BS}$	Potência de Operação do Circuito de Cada Antena da BS.
$P_{ij}$	Potência ICR para o $j$ -ésimo Usuário Pertencente ao $i$ -ésimo <i>Cluster</i> .
$q^n$	Valor Fixo de EE para $n$ -ésima Iteração.

### Sistema mMIMO PDMA

$N_T$	Número de Antenas de transmissão na BS.
$N_R$	Número de Antenas de recepção por usuário.
$N$	Número de Feixes.
$K$	Número de Usuários.
$\mathbf{H}_k$	Matriz de Canal do $k$ -ésimo Usuário do UG.
$\mathbf{Q}_k$	Matriz de canal com desvanecimento Rayleigh.
$\alpha$	Expoente de Perda de Percurso.
$d_k$	Distância do $k$ -ésimo Usuário até a BS.

$\mathbf{b}_n$	Vetor ZFBF do $n$ -ésimo Feixe.
$\mathbf{H}_C$	Matriz de Canal do Usuário Selecionado.
$\mathbf{B}_C$	Matriz Precodificação ZFBF.
$\mathbf{B}$	Matriz Precodificação ZFBF Ajustada.
$\mathbf{s}$	Vetor Símbolo de Transmissão.
$\mathbf{P}$	Matriz de Alocação de Potência.
$\mathbf{G}$	Matriz de Alocação de Feixe.
$\mathbf{t}$	Vetor de Sinal Sobreposto após o Mapeamento de Padrão.
$\mathbf{x}$	Vetor de Transmissão do Sinal AC.
$\mathbf{y}_k$	Vetor do Sinal Recebido do $k$ -ésimo Usuário.
$\mathbf{w}_k$	Vetor do Ruído Gaussiano Aditivo Branco do $k$ -ésimo usuário.
$\mathbf{G}_{\text{PDMA}}^{N \times K}$	Matriz Padrão PDMA de dimensões $N \times K$ .
$\iota$	Fator de Sobrecarga.
$v_j$	Sinal de Transmissão do $j$ -ésimo RV.
$x_k$	Símbolo da Modulação do $k$ -ésimo Usuário.
$\mu$	Fator de Escala de Potência.
$\varphi$	Fator de Mudança de Fase.
$\lambda^N$	Número de Padrões PDMA possíveis.
$r$	Repetibilidade.
SSC	Correlação da Soma Quadrática ao Quadrado.
DIV	Ordem de Diversidade.
$d_{\text{max}}^w$	Peso Máximo Ponderado da Linha.
$\Lambda$	Fator de Avaliação de Desempenho.
$\mathbf{V}_k$	Matriz do Filtro Espacial SF.
$\mathbf{v}_{nk}$	Vetor da $n$ -ésima coluna de $\mathbf{V}_k$ .
$z_{nk}$	Resultado Escalar após Suprimir a Interferência entre Feixes.

$u_{nk}$	Ganho de Canal Normalizado Equivalente entre o $k$ -ésimo usuário e AC.
$q_{nk}$	Soma da Interferência entre Feixes e Ruído após a Normalização.
<b>U</b>	Ganho de Canal Normalizado Equivalente entre UG e AC.
$\gamma_{nk}$	Relação Sinal-Interferência+Ruído (SINR) do $k$ -ésimo Usuário no $n$ -ésimo feixe.

# Sumário

	Lista de figuras . . . . .	x
	Lista de tabelas . . . . .	xi
	Lista de abreviaturas e siglas . . . . .	xiv
	Lista de símbolos . . . . .	xx
1	<b>INTRODUÇÃO</b> . . . . .	1
1.1	Introdução . . . . .	1
1.1.1	Sistema NOMA . . . . .	5
1.1.2	Sistemas MIMO-NOMA . . . . .	6
1.1.3	Compromisso entre EE-SE: Eficiência de Recurso (RE) . . . . .	8
1.1.4	Sistemas mMIMO PDMA . . . . .	10
1.2	Motivação, Justificativa e Objetivos . . . . .	12
1.3	Estrutura do Trabalho de Tese . . . . .	14
1.4	Submissões e Publicações . . . . .	15
1.4.1	Artigos Publicados . . . . .	15
1.4.2	Artigo em Processo de Re-submissão . . . . .	16
2	<b>SISTEMA NOMA</b> . . . . .	17
2.1	Modelo do Sistema NOMA . . . . .	17
2.1.1	Otimização do Sistema NOMA para o Enlace de Descida . . . . .	19
2.1.1.1	Otimização de Potência para o NOMA . . . . .	19
2.1.1.2	Otimização da Alocação de Potência do Usuário a partir da Perspectiva EE no NOMA . . . . .	20
2.1.1.3	Eficiência de Recursos (RE) . . . . .	23
2.1.1.4	Equidade para o NOMA . . . . .	23
3	<b>SISTEMA MIMO-NOMA</b> . . . . .	25
3.1	Modelo do Sistema MIMO-NOMA . . . . .	25
3.1.1	Otimização do Sistema MIMO-NOMA para o Enlace de Descida . . . . .	28
3.1.1.1	Equidade de Taxa para os Usuários MIMO-NOMA . . . . .	29
3.1.1.2	Alocação de Potência Baseada no Inverso da Resposta do Canal (ICR) no MIMO-NOMA . . . . .	29
3.1.1.3	Alocação de Potência Ótima com Taxa Igual (ERPA) em MIMO-NOMA . . . . .	30

3.1.1.4	Problema de Otimização da Eficiência Energética no MIMO-NOMA	32
3.2	Abordagem de Otimização Bi-Objetivo EE-SE	36
3.2.1	Alocação de Recursos em Sistemas MIMO-NOMA	36
3.2.1.1	Otimização Multiobjetivo: Definições	37
3.2.1.2	Formulação do EE-SE Biobjetivo em Sistemas MIMO-NOMA	37
3.2.2	Resolvendo o EE-SE em MIMO-NOMA Sob Equidade de Taxa	38
3.2.3	Métodos de Otimização Implementados	40
3.2.3.1	Método Escalarização com DK	40
3.2.3.2	Método Escalarização com SQP	41
3.2.3.3	Método Escalarização Heurístico Ganancioso	41
3.2.4	Comparação com Método de Escalarização de Soma Ponderada	42
4	SISTEMA MMIMO PDMA PARA SERVIÇOS EMBB, MMTTC E URLLC	44
4.1	Modelo do sistema mMIMO PDMA	44
4.1.1	Projeto de Matrizes Padrão para Modos de Uso 5G	47
4.1.1.1	Descrição e Características da Matriz Padrão	47
4.1.1.2	Fator de Sobrecarga das Matrizes Padrão	48
4.1.1.3	Projeto de Matriz Padrão PDMA Estendida	49
4.1.2	Critérios de Seleção e Avaliação da Matriz Padrão	51
4.1.2.1	Viabilidade da Matriz Padrão	51
4.1.2.2	Desafios na Otimização de Matrizes Padrão para Modos de Uso Híbridos em 5G	51
4.1.3	Critérios de Avaliação da Matriz Padrão	52
4.1.3.1	Repetibilidade	52
4.1.3.2	Correlação da Soma Quadrática ao Quadrado	53
4.1.3.3	Ordem de Diversidade	53
4.1.3.4	Peso Máximo Ponderado da Linha	53
4.1.3.5	Fator de Avaliação de Desempenho	54
4.1.4	Deteção Híbrida com Filtro Espacial e SIC no Receptor	54
4.1.5	Problema de Otimização de Eficiência Energética	56
5	RESULTADOS E DISCUSSÕES	58
5.1	Resultados Numéricos para Sistemas NOMA	58
5.1.1	Desempenho ERPA <i>versus</i> ICA em Sistemas NOMA	58
5.1.2	Equidade sob Sistemas NOMA	62
5.1.3	Pontos de Operação Ótimos para Eficiência Energética no NOMA	64
5.1.4	Ponto Ótimo da Eficiência de Recursos ( $EE \times SE$ ) em Função do Número de Usuários NOMA	64

5.2	Resultados Numéricos para MIMO-NOMA . . . . .	66
5.2.1	Políticas Propostas ERPA $\times$ ICR em Cenários Realistas MIMO-NOMA . . . . .	67
5.2.2	Equidade de Taxa em Sistemas MIMO-NOMA . . . . .	73
5.2.3	Ponto de Operação Ótimo para Eficiência Energética Sob Diferentes Cenários . . . . .	73
5.2.4	Ponto Ótimo de Eficiência de Recursos em Função do Número de <i>Clusters</i> e Usuários Uniformemente Distribuídos . . . . .	75
5.3	Resultados Numéricos para o MIMO-NOMA com Otimização Biobjetiva . . . . .	79
5.3.1	Detalhes de Implantação . . . . .	79
5.3.1.1	SQP $\times$ DK na Etapa 5 do Algoritmo 3 . . . . .	80
5.3.2	Comparação do Compromisso entre EE-SE em Cenários Realísticos MIMO-NOMA . . . . .	81
5.3.3	Resultados da Comparação com Método de Escalarização de Soma Ponderada . . . . .	88
5.4	Resultados para Max-EE em Sistemas mMIMO PDMA com Serviços eMBB, mMTC e uRLLC . . . . .	89
5.4.1	Seleção da Matriz Padrão PDMA . . . . .	90
5.4.2	Análise das Matrizes Padrão PDMA para o eMBB . . . . .	92
5.4.3	Matrizes padrão PDMA para mMTC . . . . .	93
5.4.4	Matrizes padrão PDMA para uRLLC . . . . .	94
5.4.5	Matrizes padrão PDMA para transmissão híbrida com serviços eMBB e uRLLC . . . . .	96
6	RELAÇÃO ENTRE OS TRABALHOS . . . . .	97
6.1	Variantes NOMA sob Equidade de Taxa . . . . .	97
7	CONCLUSÕES E PERSPECTIVAS . . . . .	100
7.1	Sugestões de Trabalhos Futuros . . . . .	102
	REFERÊNCIAS . . . . .	103
	ANEXOS . . . . .	111
	ANEXO A – NONORTHOGONAL MULTIPLE ACCESS SYSTEMS OPTIMIZATION TO ENSURE MAXIMUM FAIRNESS TO USERS . . . . .	112

ANEXO B – ENERGY AND SPECTRAL EFFICIENCIES TRADE-OFF IN MIMO-NOMA SYSTEM UN- DER USER-RATE FAIRNESS AND VARIA- BLE USER PER CLUSTER . . . . .	127
ANEXO C – NON-LINEAR BIOBJECTIVE EE-SE OPTI- MIZATION FOR NOMA-MIMO SYSTEMS UNDER USER-RATE FAIRNESS AND VARIA- BLE NUMBER OF USERS PER CLUSTER	142
ANEXO D – PDMA LARGE-SCALE MIMO ENERGY EF- FICIENCY MAXIMIZATION FOR EMBB, MMTC AND URLLC . . . . .	158

# 1 Introdução

## 1.1 Introdução

O rápido crescimento do volume de dados móveis, junto à demanda contínua por sistemas celulares de maior capacidade, os quais geram dados a partir de *smartphones* e novos dispositivos móveis, bem como a recente demanda por dispositivos de acesso à Internet das coisas (IoT - *Internet of Things*) e sistemas avançados de assistência à direção de veículos, fazem com que haja um aumento fenomenal na procura por sistemas de acesso sem fio em todo o mundo.

Devido ao seu sucesso, desde o início da implantação dos sistemas de comunicação sem fio existe um aumento continuado da demanda. Com isso, houve a necessidade crescente de sua evolução, conhecido hoje por sistemas de primeira geração (1G - *First-Generation*), segunda geração (2G - *Second-Generation*), terceira geração (3G - *Third-Generation*), quarta geração (4G - *Fourth-Generation*) e 5G. O principal fator para a evolução dos sistemas de comunicação sem fio ao longo de suas diferentes gerações foi propor novas técnicas de múltiplo acesso, incluindo acesso múltiplo por divisão de frequência (FDMA - *Frequency-Division Multiple Access*) para 1G, acesso múltiplo por divisão de tempo (TDMA - *Time-Division multiple access*) para 2G, acesso múltiplo por divisão de código (CDMA - *Code-Division Multiple Access*) para 3G e acesso múltiplo por divisão de frequência ortogonal (OFDMA - *Orthogonal Frequency-Division Multiple Access*) para 4G. Todas essas técnicas são de múltiplo acesso ortogonal (OMA - *Orthogonal Multiple Access*), ou seja, os recursos sem fio são compartilhados ortogonalmente entre vários usuários em diferentes domínios, como tempo, frequência e código. A ortogonalidade dos elementos de recursos alocados reduz a interferência de múltiplo acesso e possibilita o uso de receptores de baixa complexidade eficientes. As técnicas OMA acima mencionadas não são suficientes para responder à crescente demanda de conectividade das gerações futuras, uma vez que o número de usuários suportados é limitado pelo número de elementos de recursos ortogonais disponíveis (REBHI et al., 2021).

O NOMA é uma das técnicas de acesso de rádio promissoras para o aprimoramento de desempenho em comunicações celulares de próxima geração B5G. Comparado ao OMA, que é uma técnica de acesso múltiplo ortogonal de alta capacidade bem conhecida, o NOMA oferece um conjunto de benefícios desejáveis, incluindo maior eficiência de espectro. Existem diferentes tipos de técnicas NOMA, incluindo NOMA no domínio de potência e domínio de código (ISLAM et al., 2017). Este trabalho de tese se explora o princípio do NOMA no domínio das potências, com a possibilidade de agregação de superposição de sinais no transmissor e cancelamento de interferência sucessiva no receptor. Vários

pesquisadores demonstraram que o NOMA pode ser usado de forma eficaz para atender aos requisitos de taxa de dados no nível de rede e do ponto de vista da experiência do usuário das tecnologias 5G. Este trabalho de tese propõe esquemas de transmissão NOMA no domínio da potência, sob a perspectiva da equidade de taxas, enfatizando a análise de desempenho dos usuários de borda, estratégias de alocação de potência para se atingir a máxima equidade de taxa<sup>1</sup> e esquemas de emparelhamento de usuários nos diferentes *clusters* NOMA.

De fato, em contraste com os esquemas convencionais de acesso múltiplo ortogonal OMA, o sistema NOMA permite a transmissão simultânea de vários usuários por meio de codificação de sobreposição com diferentes níveis de potência (SAITO et al., 2013). Os esquemas NOMA são divididos em duas categorias de multiplexação: a) multiplexação no domínio de potência PD-NOMA (SAITO et al., 2013), (HIGUCHI; KISHIYAMA, 2013); b) multiplexação no domínio de código CD-NOMA. Nesta tese foi adotado o NOMA no domínio da potência, portanto, por facilidade e conveniência, o acrônimo NOMA será usado para se referir ao NOMA no domínio da potência.

Entre as configurações de multiplexação no domínio de código estão as de acesso múltiplo com espalhamento de baixa densidade (LDS - *Low-Density Spreading*) (HOSHYAR et al., 2008) que pode ser combinado com o acesso múltiplo por divisão de código (CDMA - *Code-Division Multiple Access*) (LDS-CDMA) e OFDMA (LDS-OFDMA), acesso múltiplo por código esparso (SCMA - *Sparse Code Multiple Access*) (NIKOPOUR; BALIGH, 2013), acesso compartilhado multiusuário (MUSA - *Multi-User Shared Access*) (YUAN et al., 2016) e múltiplo acesso passível de cancelamento de interferência sucessiva (SAMA - *Successive Interference Cancellation Amenable Multiple Access*) (DAI et al., 2014). Existe ainda outra configuração que pode ser realizada tanto no domínio de potência como no domínio de código é a de múltiplo acesso no domínio da potência e do código esparso (PSMA - *Power Domain Sparse Code Multiple Access*) (MOLTAFET et al., 2018). A Figura 1.1 apresenta a classificação para os principais esquemas de múltiplo acesso existentes. Outros esquemas de múltiplo acesso, como PDMA e multiplexação por divisão de bits (BDM - *Bit Division Multiplexing*) (HUANG et al., 2014) também foram propostos na literatura (DAI et al., 2015).

O sistema NOMA de múltiplo acesso emergiu como um esquema promissor para futuros esquemas de transmissão via rádio B5G. O sistema NOMA permite que os usuários na borda da célula seja atendido nas mesmas condições dos usuários próximos à estação rádio base (BS - *Base Station*). Essa característica do sistema NOMA permite que todos os usuários sejam atendidos com a mesma taxa ou com a máxima equidade de taxa. A especificação da tecnologia 5G determina o fornecimento de 1Gbps para os seus usuários

---

<sup>1</sup>Justiça de taxa para ambos os tipos de usuários de borda e usuários localizados próximos à BS (internos).

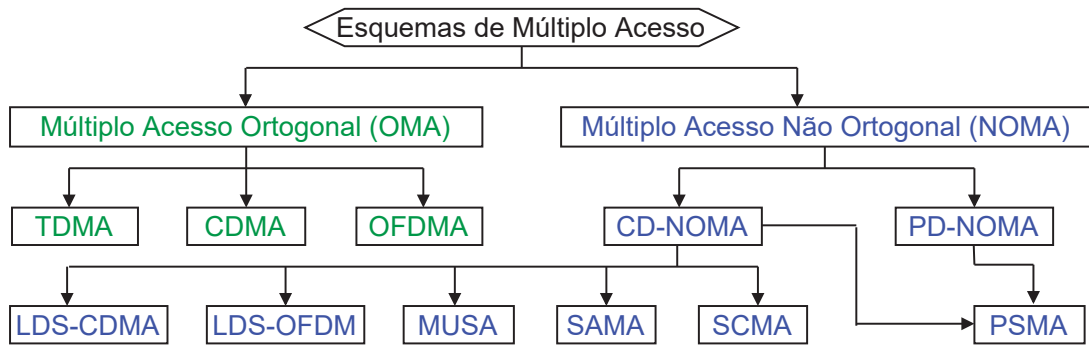


Figura 1.1 – Classificação dos esquemas de múltiplo acesso existentes.

operando no modo eMBB. Com a aplicação da equidade sugerida por esse trabalho, todos os usuários podem ser atendidos com essa taxa inclusive os usuários de borda de célula ou com menor ganho de canal. Por isso, todos os trabalhos desenvolvidos nesta tese buscam obter a máxima equidade de taxa; neste sentido, o esquema de transmissão NOMA se apresentou adequado para atingir essa meta. O NOMA alcança eficiências espectrais superiores por superposição de códigos (SC - *Superposition Coding*) no transmissor e cancelamento de interferência sucessiva (SIC - *Successive Interference Cancellation*) no receptor (BENJEBBOUR et al., 2013). Além disso, a evolução das redes sem fio para o 5G coloca novos desafios e níveis elevados para a EE, uma vez que toda a rede será ultra-densa. Com um aumento extremo no número de nós de infraestrutura, o consumo total de energia pode simplesmente atingir níveis inaceitáveis. Embora a energia substancial seja basicamente consumida pelo *hardware*, o NOMA possui uma capacidade inerente de adequar sua estratégia de transmissão de acordo com o tráfego e as condições do estado do canal (CSI - *Channel State Information*) dos usuários. Assim, pode atingir um bom ponto de operação, onde tanto a SE quanto a EE atingem o compromisso EE-SE ótimo.

Muitos trabalhos de pesquisa tecnológica se preocupam em projetar e implementar a técnica NOMA, bem como resolver vários problemas técnicos inerentes a esse novo método. A literatura recente demonstra que o sistema NOMA é compatível com comunicações cooperativas, retransmissão ou ainda com sistemas que utilizam de MIMO para aumentar significativamente os ganhos de desempenho e/ou confiabilidade (ISLAM et al., 2017). Pesquisadores introduziram uma técnica onde a taxa de dados baixa e qualidade ruim de serviços experimentados na extremidade da célula podem ser facilmente melhoradas adotando-se estruturas de transmissão baseadas em NOMA. Da mesma forma que no MIMO, existem outros esquemas de transmissão que podem ser combinados ao NOMA. Um desses sistemas bastante explorado na literatura envolve o PDMA, conforme discutimos na sequência.

Recentemente, o esquema PDMA foi introduzido nos cenários 5G e B5G (DAI et al., 2014), (ZENG et al., 2015), (LI et al., 2021). Ao contrário do esquema NOMA convencional, o PDMA adota segmentação de padrões para separar os sinais do usuário

no transmissor/receptor, mas igualmente ao NOMA, aplica-se SIC no lado do receptor (DAI et al., 2014), (CHEN et al., 2017), (LI et al., 2017). De fato, o PDMA é um esquema NOMA, recentemente proposto para atender aos inúmeros requisitos do 5G-B5G, podendo esses padrões serem realizados em diferentes domínios de recursos (SUN et al., 2020), (JIANG et al., 2019). O PDMA distingue vários sinais de usuário no transmissor através do uso de domínio de espaço, domínio de potência e domínio de código, individualmente ou conjuntamente e simultaneamente, possibilitando assim que vários usuários compartilhem os mesmos recursos de tempo-frequência (LI et al., 2020). Nesta tese analisamos o PDMA no domínio da potência e no domínio do feixe (*beam space*) aplicado a sistemas de múltiplas antenas.

Representamos um exemplo da distribuição dos valores de recursos (RV - *Resource Value*) e da matriz padrão PDMA na Figura 1.2. As colunas da Figura 1.2 a) representam

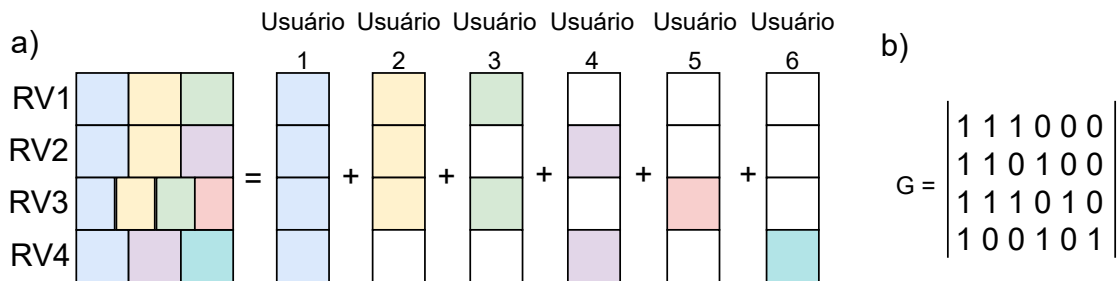


Figura 1.2 – Exemplo de padrão PDMA.

os usuários e as linhas representam os valores de recursos. Portanto, temos 6 usuários e 4 valores de recursos. Considere que o recurso do exemplo é o feixe gerado por antenas. Os quadrados coloridos representam os usuários ativos nos respectivos feixes. Na primeira linha RV1 ou feixe 1 somente os usuários 1, 2 e 3 estão ativos e na segunda linha RV2 ou feixe 2 somente os usuários 1, 2 e 4 estão ativos e assim por diante. Os usuários em cada feixe podem usar o NOMA para a distribuição de potência. A matriz padrão PDMA é representada na A Figura 1.2.b), sendo os usuários ativos em cada feixe identificados com o valor 1 e os usuários inativos com 0. Ao implantar a multiplexação de vários domínios, o PDMA pode aproveitar ao máximo os recursos sem fio para aumentar a soma de taxas com complexidade computacional acessível. Recentemente, receptores cada vez mais avançados têm sido propostos para operar em esquemas PDMA. Embora o PDMA seja um candidato promissor para futuros sistemas de comunicação, ainda é necessária uma pesquisa abrangente e completa sobre o padrão e o projeto do transceptor (JIANG et al., 2019). Com o intuito de explorar mais esse sistema, investigamos alguns aspectos e aplicabilidade do esquema de transmissão PDMA sob os diferentes modos de uso em 5G (uRLLC, mMTC e eMBBC). Tendo em vista propor um sistema de transmissão que enriqueça as características do esquema PDMA convencional, apresentamos na sequência um sistema de transmissão PDMA modificado promissor para 5G que utilizamos nesse trabalho.

As matrizes padrão PDMA para sistema mMIMO são uma das principais tecnologias promissoras para sistemas de comunicações móveis B5G. Os sistemas mMIMO foram propostos para melhorar significativamente a soma das taxas do sistema com graus de liberdade extras que facilitam a obtenção de elevados ganhos de multiplexação espacial e diversidade de transmissão (WANG et al., 2016; JIANG et al., 2019). Exemplos de sistemas mMIMO estão em (SANTOS et al., 2022; ALVES; ABRÃO, 2022) e MIMO extra grande (XL-MIMO - *Extra-Large* MIMO) em (SOUZA et al., 2021; UBIALI et al., 2021). Tais arranjos podem fornecer recursos espaciais suficientes para um grande número de dispositivos conectados simultaneamente nos chamados cenários lotados (*crowded*). Atualmente, a composição do mMIMO e PDMA tem recebido cada vez mais atenção, por agregar as vantagens de ambos os sistemas e ter como consequência um aumento no ganho de desempenho (LI et al., 2017; JIANG et al., 2019).

Nas próximas subseções descreve-se o estado da arte e uma revisão bibliográfica para as técnicas NOMA, MIMO-NOMA e mMIMO PDMA para uma melhor organização da nomenclatura específica, caracterização e entendimento dos sistemas analisados neste trabalho de tese. Os respectivos modelos de sistema e problemas de otimização tratados nesta tese são formulados e discutidos nos Capítulos 2, 3 e 4.

### 1.1.1 Sistema NOMA

O NOMA é uma técnica promissora para tecnologias B5G por otimizar o uso da largura de banda, sendo que alguns usuários podem compartilhar diferentes proporções de potência, ou seja, aqueles com menor ganho de canal receberão maior proporção. Como resultado, usuários na borda da célula são tratados nas mesmas condições que os outros usuários em sistemas celulares sem fio (TSE; PRAMOD, 2005).

O esquema NOMA permite que vários usuários compartilhem recursos de tempo e frequência na mesma camada espacial via domínio de potência ou multiplexação de domínio de código (DAI et al., 2015). Este trabalho de tese usa exclusivamente o NOMA no domínio da potência, ou seja, os usuários compartilham a mesma banda com diferentes proporções da potência total entre esses usuários. Neste caso, maior proporção de potência de transmissão é destinada aos usuários com condições de canais piores. Esses usuários decodificam suas próprias mensagens tratando as informações dos outros como ruído. Os usuários com melhores condições de canal utilizam a técnica de SIC para cancelar a respectiva interferência gerada pelos demais usuários emparelhados. Esses usuários primeiro detectam mensagens de outros usuários emparelhados com condição do CSI baixo e, em seguida, detectam sua própria mensagem removendo informações dos outros usuários (DING et al., 2016a).

O domínio de potência usa diferentes níveis de potência entre usuários para transmitir no mesmo canal simultaneamente (SAYED-AHMED et al., 2018). O receptor NOMA

extraí seus sinais usando o SIC para eliminar a interferência entre os usuários (ALI et al., 2017). Ele explora a perda de percurso diferente entre os usuários para separar os sinais, aumentando assim a eficiência espectral SE e a equidade dos usuários, aplicando um esquema para atingir a capacidade desejada no DL (SHIN et al., 2017).

Para alcançar esse objetivo, alguns estudos são desenvolvidos pela pesquisa de melhores condições do uso de um sistema NOMA. Podemos maximizar a taxa de transmissão de dois usuários em um *cluster* (KIM et al., 2013), ou seja, a soma da capacidade de dois usuários, a partir do uso da técnica *beamforming* na BS. Ou ainda, maximizar a soma das taxas dos usuários no NOMA ao melhorar alguns parâmetros do sistema, aplicando-se um algoritmo de seleção do vetor de *beamforming* com menor interferência e que minimize o consumo de potência; implemente uma política de seleção dos usuários de forma a reduzir interferência e obtenção da melhor alocação de potência desses usuários para determinados número de usuários por *cluster* (CARVALHO et al., 2020). Além disso, temos a configuração em que os usuários informam seus ganhos de canal para a BS, ao qual a BS classifica e ordena-os para determinar a aplicação do NOMA nos usuários que possuem maior diferença de ganho e geometria adequada, enquanto que os outros usuários continuam operando no modo OMA (CHEN et al., 2014).

### 1.1.2 Sistemas MIMO-NOMA

As comunicações MIMO têm o potencial de obter ganhos significativos na taxa de transmissão do sistema (ALI et al., 2017). Portanto, em um sistema MIMO-NOMA DL, várias antenas de recepção de diferentes usuários com diferentes ganhos de canal são agrupadas em diferentes *clusters* NOMA com base na disparidade de potências recebidas. Assim, os equipamentos do usuário (UEs - *User Equipment*) em cada *cluster* funcionam com base na detecção SIC.

Observe que, em (CHOI, 2014), os autores não consideram a distribuição de potência ótima. Alternativamente, os autores consideram o NOMA combinado com a tecnologia MIMO multiusuário para mostrar o ganho na taxa de transmissão dos usuários na borda da célula. Em vez de uma BS, podemos usar duas BSs e aplicar o código Alamouti combinado com a técnica SIC para recuperar o sinal na recepção (CHOI, 2014).

A aplicação da tecnologia MIMO em NOMA é promissora porque o uso da técnica MIMO fornece alta taxa de dados e baixa taxa de erro por meio de ganho de multiplexação e diversidade espacial, respectivamente (KANG et al., 2020), (DING et al., 2016b). Uma nova estrutura MIMO-NOMA para transmissão DL e enlace de subida (UL - *Uplink*) aplicando o conceito de alinhamento de sinal foi proposta em (DING et al., 2016b). Um novo esquema de implementação MIMO-NOMA UL com cancelamento de interferência de grupo é mostrado na literatura como um esquema melhor do que transmissão OMA e alinhamento de sinal (WANG et al., 2018). Em (LIU et al., 2016), é demonstrado

analiticamente e através de simulações, a superioridade do MIMO-NOMA sobre o sistema MIMO convencional baseado em OMA em termos da soma da capacidade do canal e também no que diz respeito à soma da capacidade ergódica. Além disso, a maximização da capacidade ergódica do sistema MIMO-NOMA com a restrição da potência máxima transmitida e da taxa mínima para o usuário desvanecido tem sido proposta mostrando sua superioridade sobre o MIMO-OMA tradicional (SUN et al., 2015). Uma técnica de formação de feixe foi sugerida em (ALI et al., 2017) para cancelar a interferência inter *cluster* em sistemas MIMO-NOMA. Os autores também propõem uma solução de alocação de potência (PA - *Power Allocation*) dinâmica intra *cluster* e inter *cluster* para maximizar a capacidade geral da célula e minimizar a energia total consumida pelo agrupamento de usuários móveis. A PA também foi investigada por meio de uma abordagem de otimização linear para otimizar conjuntamente vetores formadores de feixe e coeficientes de alocação de energia para um *cluster* MIMO-NOMA em (DING; CAI, 2018).

Uma importante forma de aplicação da técnica MIMO no sistema NOMA foi considerada para que um novo projeto de pré-codificação e detecção de matrizes para sistemas MIMO-NOMA e seu desempenho fossem analisados (DING et al., 2016a). Para um número genérico de usuários por *cluster* em um sistema MIMO-NOMA DL, os esquemas de EE e PA foram propostos para converter um problema complexo de EE-PA em um problema SE-PA equivalente (SAYED-AHMED et al., 2018). Uma estratégia de PA para garantir que o sistema MIMO-NOMA atinja uma alta taxa individual para cada usuário foi comparada ao PA arbitrário no sistema MIMO-OMA (ZENG et al., 2017a). Foi demonstrado que o MIMO-NOMA poderia alcançar melhor equidade por meio de PA apropriada. O artigo (AL-OBIEDOLLAH et al., 2019) fez uma análise do compromisso entre a soma da taxa e a equidade usando a otimização simples e multi-objetivo (MOO - *Multi-Objective Optimization*) com base no projeto de *beamforming* para um sistema de entrada única e múltiplas saídas (SIMO - *Single Input and Multiple Output*) com NOMA. Para alcançar uma melhor compensação, um algoritmo para selecionar usuários fortes e fracos com base na razão de equidade foi desenvolvido em (AL-WANI et al., 2019) para um sistema NOMA de enlace de descida com formação de feixe de forçamento a zero (*zero forcing*), a fim de melhorar a compensação de taxa-equidade em sistemas NOMA.

Outra abordagem consiste em separar os usuários em três categorias: usuários sozinhos, usuários fracos e usuários fortes, e analisar a taxa de transmissão e relação sinal-ruído (SNR - *Signal to Noise Ratio*) e sinal interferente mais ruído (SINR - *Signal-to-Interference-plus-Noise Ratio*) SNR/SINR para as três categorias, enfatizando a Qualidade de Serviço (QoS - *Quality of Service*) dos usuários fracos (DHAKAL et al., 2019). Neste trabalho, foi analisado o caso de dois usuários por *cluster* apenas. Além disso, em (ZHANG et al., 2017) foi desenvolvida a otimização da EE, sendo o *ponto verde operacional* obtido, ou seja, o ponto de operação do sistema ideal que maximiza a EE geral; no entanto, esses trabalhos não consideram a equidade de taxa entre os usuários. Observe que o

compromisso entre EE e SE já foi estudado recentemente em (LIU et al., 2017), mas os autores consideram apenas a potência como uma variável analisada, não consideram o número de usuários como uma variável de otimização. Além disso, a maximização da taxa mínima para cada usuário foi analisada em (BUI et al., 2019), mas apesar de mencionar a equidade este conceito não foi sistematicamente explorado na obtenção de novos resultados. O artigo em (NI et al., 2018) analisa o compromisso entre EE e SE considerando a equidade.

### 1.1.3 Compromisso entre EE-SE: Eficiência de Recurso (RE)

O problema de otimização da SE sob equidade de taxa de usuário e configuração NOMA apresenta limitações e desafios. Melhorar o SE unilateralmente em geral leva a um alto consumo de energia e, assim, aumenta o custo geral de operação da rede. Usando o máximo de energia possível para maximizar SE, a EE geral é drasticamente reduzida (JI et al., 2018). SE e EE são duas métricas importantes para melhorar a eficiência de recursos (RE - *Resource Efficiency*) em sistemas de comunicação sem fio com menores custos de energia por bit transferido e maiores taxas de dados. No entanto, ambas as métricas EE e SE são frequentemente conflitantes, ou seja, aumentar uma métrica leva a uma redução na outra, e vice-versa (UDALCOVS et al., 2017). Portanto, uma melhor abordagem de otimização consiste em lidar com o compromisso entre as métricas SE e EE, adotando uma descrição de MOO. De fato, recentemente, vários trabalhos na literatura têm se dedicado a resolver problemas de MOO relacionados a redes de comunicação. Outros trabalhos que adotam abordagens de MOO incluem (SASIKUMAR; JAYAKUMARI, 2020; LIU et al., 2018; ZHANG et al., 2018; RUAN et al., 2019; ZUO et al., 2018; MARTINEZ; ABRÃO, 2020; SONG et al., 2018). Em (SASIKUMAR; JAYAKUMARI, 2020), uma nova técnica, ou seja, o algoritmo de geração de população superior (SPGA - *Superior Population Generation Algorithm*), modifica o algoritmo genético (GA - *Genetic Algorithm*) regular para incorporar as capacidades das técnicas convencionais. O algoritmo foi desenvolvido para resolver o problema de otimização do compromisso entre SE-EE para redes heterogêneas de 5G, que usa tecnologias futuras importantes, como agregação de portadoras e rádio cognitivo.

O compromisso EE-SE no DL de sistemas OFDMA é investigado em (LIU et al., 2018), considerando o custo de estimativa de canal e o efeito correspondente de CSI no SE e EE. O problema é formulado como uma MOO para determinar a potência ótima de transmissão do piloto, potência de transmissão de dados e atribuição de subportadora. O problema MOO original é transformado em um problema de otimização de objetivo único (SOO - *Single-Objective Optimization*), pertencentes à classes de problemas de programação não linear inteira mista não convexa (NCMINP - *Non-Convex Mixed-Integer Nonlinear Programming*) e polinômio não determinístico complexo (NP-complexo - *Hard Nondeterministic Polynomial*). Para resolvê-lo, os autores propõem um algoritmo eficiente,

adotando uma abordagem de subproblema de otimização convexa alternada na região de baixa potência, bem como uma conversão aproximada e método de derivação e limitação (*Branch-and-Bound*) na região de alta potência. Os resultados da simulação validam o desempenho atingível em termos de otimização do compromisso entre EE-SE.

Os autores em (ZHANG et al., 2018) investigam a compensação EE-SE em redes heterogêneas baseadas em DL NOMA (HetNets). O problema de compensação EE-SE é formulado como um problema de otimização bi-objetiva (BOO - *Bi-Objective Optimization*). Sob as restrições de potência máxima de transmissão de usuários de pequenas células (SCUs - *Small Cell Users*) e requisitos de taxa mínima, tanto EE quanto SE são maximizados. Para tornar o MOO tratável, um método de soma ponderada é usado para transformá-lo em um problema SOO, o que permite que os provedores de rede sintonizem dinamicamente o fator de peso para se adaptar aos diferentes requisitos de projeto de EE e SE. Para resolver o SOO transformado, um algoritmo de alocação de SCU é proposto para determinar os SCUs múltiplos para cada subcanal, com base no qual um algoritmo de determinação de fator proporcional de potência para SCUs e um novo algoritmo de alocação de potência para subcanais são propostos. Os resultados numéricos mostram que os algoritmos convergem rapidamente e demonstram que os valores de EE e SE são soluções ótimas de Pareto e superiores às obtidas em esquemas convencionais de OMA.

O compromisso entre EE-SE em um enlace de descida de uma transmissão MIMO massiva de uma única célula com CSI estatístico, com o objetivo de otimizar o sistema RE é investigado em (RUAN et al., 2019). Além disso, em (ZUO et al., 2018), os autores propõem um novo esquema de alocação de recursos de baixa complexidade para obter uma compensação EE-SE superior em sistemas NOMA com múltiplas operadoras com a restrição de equidade do usuário. O objetivo é maximizar as métricas de EE e SE por meio do número de usuários e da alocação de energia entre eles. A rede, por sua vez, possui diferentes demandas de comunicação em diferentes cenários. Portanto, é mais importante aumentar o SE no tráfego de pico, mas por outro lado, fora do horário de pico, torna-se mais importante priorizar a EE. Como resultado, é fundamental descrever o problema de RE como um problema de otimização MOO.

Em (MARTINEZ; ABRÃO, 2020) a compensação EE-SE é discutida pela formulação de um problema de otimização bi-objetivo em redes de múltiplo acesso por divisão de código óptico (OCDMA - *Optical Code Division Multiple Access*), enquanto as soluções são comparadas ao implantar a soma ponderada (WS - *Weighted Sum*) e procedimentos de escalarização de restrição ( $\epsilon$ -C) são combinados com programação não linear como método lagrangiano aumentado (ALM - *Augmented Lagrangian Method*) e método de programação quadrática sequencial (SQP - *Sequential Quadratic Programming*) para resolver os subproblemas gerados. Em (SONG et al., 2018), os autores investigam o problema de alocação de recursos para alcançar compensações SE-EE superiores sob requisitos de taxa de usuário

mínima em sistemas híbridos de múltiplo acesso não ortogonal de múltiplas portadoras (MC-NOMA - *Multi-Carrier NOMA*); tais sistemas incorporam os modos NOMA e OMA em análise de estrutura unificada. Em primeiro lugar, a compensação SE-EE é formulada como um problema de MOO com restrições de exigência de taxa mínima. Considerando a natureza de não convexidade do problema MOO, ele é convertido em um problema SOO utilizando o método de Tchebycheff ponderado. A decomposição dual lagrangiana e a programação convexa sequencial são aplicadas para resolver o problema SOO equivalente.

### 1.1.4 Sistemas mMIMO PDMA

Não podemos confundir múltiplo acesso por divisão de percurso (PDMA - *Path-Division Multiple Access*) (ZENG; ZHANG, 2016), (WANG et al., 2018), (JI et al., 2019), (TANG et al., 2020), (LI et al., 2021b) com múltiplo acesso por divisão de padrões (PDMA - *Pattern Division Multiple Access*), (DAI et al., 2014), (ZENG et al., 2015), (LI et al., 2021). Ambos possuem a mesma sigla, mas com significados e aplicações de sistemas diferentes. O primeiro é um esquema de multiplexação que cria pequenas cadeias de radiofrequência com altos ganhos em um sistema MIMO com antenas de lentes para ondas milimétricas (*mmWave*) cuja resposta segue uma função *sinc*. Essa multiplexação espacial é gerada para banda estreita e banda larga em comunicações de ondas milimétricas, onde fluxos de dados paralelos são transmitidos por diferentes caminhos de propagação com processamento de caminho único. No segundo contexto de interesse aqui, o PDMA significa o desenho da matriz padrão que define os usuários e os recursos empregados para transmitir os sinais dos usuários, gerando domínios temporais, espaciais e de codificação. Neste trabalho, utilizamos o domínio da potência e do feixe das antenas no PDMA. Outra variação recentemente publicada do PDMA é o acesso múltiplo por divisão de padrão holográfico (HDMA - *Holographic-Pattern Division Multiple Access*), (DENG et al., 2022). A principal característica do HDMA é usar o padrão holográfico obtido da superfície holográfica reconfigurável (RHS - *Reconfigurable Holographic Surface*). O RHS usa os elementos de radiação metamateriais para construir um padrão holográfico baseado no princípio da interferência holográfica. O objetivo deste padrão holográfico é controlar a amplitude de radiação da onda de vazamento em cada elemento de forma eficaz para alcançar as direções de transmissão de sinal desejadas.

Considerando o contexto de PDMA tratado em nosso trabalho, os autores do (ZENG et al., 2021) propõem o uso de PDMA com uma estratégia de PA otimizada. Eles estudam o esquema PA para maximizar a soma das taxas de um sistema PDMA de DL com informações de estado de canal CSI perfeitas e imperfeitas. Os autores aplicaram matrizes padrão existentes na literatura para desenvolver os resultados numéricos e corroborar o PDMA proposto com o esquema PA. Além disso, um esquema de alocação de recursos que utiliza o processo de decisão restrito de Markov (MCDP - *Markov Constrained Decision*

*Process*) baseado em rede Q profunda (DQN - *Deep Q-Network*) é proposto em (LI et al., 2021a) para maximizar o EE médio da rede de informações sem fio simultâneas e transferência de potência (SWIPT - *Simultaneous Wireless Information and Power Transfer*) baseado em PDMA com convergência rápida. No entanto, tal análise foi realizada em uma determinada matriz padrão sem levar em consideração suas características e possível otimização.

A tecnologia PDMA tem muitas aplicações em serviços e exemplos de uso em 5G e B5G, como eMBB, mMTC e uRLLC. Os serviços de aplicativos eMBB exigem alta largura de banda para suportar filmes em 4K e 8K que podem conter trabalho ou entretenimento na nuvem. No cenário eMBB, o PDMA aumenta a capacidade multiusuário, proporcionando mais equidade contra o efeito perto-distante enquanto melhora a experiência do usuário em redes ultradensas (VAEZI et al., 2019). Para pequenos pacotes eMBB, o PDMA pode aumentar transferência da soma de taxas e reduzir a latência e a sobrecarga quando combinado com protocolos sem concessão (GF - *Grant-Free*). Além disso, os serviços uRLLC possuem requisitos extremamente rígidos em relação à taxa de transferência de dados, atraso e confiabilidade. Exemplos de serviços uRLLC são automação industrial, cirurgia de telemedicina, automação de distribuição de rede inteligente, segurança de transporte e veículo aéreo não tripulado (UAV - *Unmanned Aerial Vehicle*). Além disso, no modo de uso uRLLC, a transmissão GF pode ser aplicada para obter latência ultrabaixa. Embora recursos físicos dedicados possam ser alocados para UE uRLLC com o objetivo de alcançar alta confiabilidade, a SE resulta em baixos valores, especialmente para serviços aperiódicos ou baseados em rajadas. Para obter alta confiabilidade, baixa latência e alta eficiência de espectro ao mesmo tempo, o método de transmissão PDMA representa uma escolha adequada e pode ser aplicado para aumentar a confiabilidade em uma rede de comunicação sujeita a colisões (VAEZI et al., 2019). A aplicação de PDMA permite a multiplexação eficiente de serviços eMBB, mMTC e uRLLC para melhorar ainda mais a utilização dos blocos de recursos.

Os serviços mMTC contêm um grande número de dispositivos IoT com baixa quantidade de dados enviados e alta tolerância de latência, exigindo baixo custo de fabricação e longa duração da bateria, como em aplicativos de cidade inteligente (*Smart City*) (CHEN et al., 2019). Em um cenário mMTC, uma BS precisa fornecer uma conexão a um grande número de terminais. Portanto, o principal desafio para um cenário mMTC é de como lidar efetivamente com conexões densas com restrições de potência (VAEZI et al., 2019). De fato, o PDMA é um método muito competitivo para resolver o problema da conectividade densa, juntamente com a exigência de ampla cobertura do sistema. Uma caracterização sucinta desses serviços pode ser apresentada da seguinte forma: (a) o eMBB suporta conexões estáveis com taxas de dados de pico muito altas, bem como taxas moderadas para usuários na borda da célula; (b) o mMTC suporta um grande número de dispositivos IoT, que são ativos apenas esporadicamente e enviam pequenas cargas

de dados; (c) o uRLLC suporta transmissões de baixa latência de pequenas cargas com confiabilidade muito alta a partir de um conjunto limitado de terminais, que são ativos de acordo com padrões normalmente especificados por eventos externos, como alarmes (POPOVSKI et al., 2018). Para que o padrão PDMA atenda melhor ao modo de uso do eMBB, a taxa de transferência do usuário deve ser maximizada. Por outro lado, no modo de uso mMTC, o fator de sobrecarga (OF - *Overload Factor*) no PDMA deve ser maximizado, enquanto no cenário uRLLC, para atingir simultaneamente baixa latência e alta confiabilidade, a interferência entre padrões (IPI - *Inter-Patterns Interference*) deve ser mitigada, implantando padrões com OFs baixos e minimizando a correlação da soma quadrada (SSC - *Sum Squared Correlation*) (SUN et al., 2020).

Em (SUN et al., 2020), os autores estabelecem critérios para obtenção das matrizes padrão PDMA em UL e DL para os três cenários de modo de uso 5G, apresentando uma simulação numérica de desempenho para cada serviço; os autores consideram apenas uma antena de transmissão. Por outro lado, (LI et al., 2017) propõe o sistema mMIMO PDMA em cenário de transmissão DL com detecção híbrida onde um filtro espacial (SF - *Spatial Filter*) é usado no receptor com o SIC para separar sinais de múltiplos domínios sobrepostos. Além disso, os autores formulam o problema de maximização da soma de taxas para obter a política ótima de mapeamento de padrões. Os resultados da simulação mostram que o esquema mMIMO PDMA apresenta um ganho de desempenho de soma de taxas substancial sobre os sistemas OMA e NOMA. A abordagem apresentada em (JIANG et al., 2019) é semelhante à apresentada em (LI et al., 2017); a diferença está na adição da etapa de otimização de alocação de feixe. As simulações revelam que a combinação das duas etapas de otimização apresenta melhor desempenho. É importante ressaltar que em (JIANG et al., 2019) e (LI et al., 2017) os autores não consideram aplicativos 5G, como URLLC, mMTC e eMBB. Até o momento, há uma carência na literatura em analisar as matrizes padrão PDMA que melhor atendem aos requisitos do serviço 5G.

## 1.2 Motivação, Justificativa e Objetivos

Com a alta demanda de serviços em telecomunicações, cada vez mais tem-se uma crescente escassez de recursos, principalmente energia e espectro. Com o objetivo de compensar essa escassez, neste trabalho de tese busca-se obter uma maior EE e maior SE nos diferentes sistemas de comunicação para 5G e B5G analisados. Note-se no entanto, uma vez que esses fatores de recursos são antagônicos, deve-se obter o melhor compromisso entre EE e SE denominado eficiência de recurso RE de forma a atender as diferentes necessidades e quesitos dos diferentes modos de uso em 5G e B5G, *i.e.*, URLLC, mMTC e eMBB.

Equidade de taxa ou simplesmente equidade no contexto deste trabalho de tese é

um conceito importante quando se trata de sistemas 5G. Existem muitos trabalhos que analisam o comportamento da equidade em sistemas 5G. Porém, diferentemente do que existe na literatura, em nosso trabalho revelamos o comportamento dos sistemas NOMA, MIMO-NOMA e mMIMO PDMA para a configuração DL operando simultaneamente sob a máxima equidade para o usuário (incluindo usuários de borda) e máxima eficiência energética do sistema para a obtenção de políticas de alocação de proporção de potência para cada usuário. No projeto de otimização faz-se necessário encontrar a potência total mínima e o fator de proporção de potência para cada usuário ativo. Depois disso, a EE máxima do sistema, a taxa de dados igualitária, a potência total e a distribuição da proporção de potência entre os usuários dos *clusters* devem ser determinadas sob um específico consumo de energia do sistema. Os procedimentos de otimização devem adotar as mesmas restrições e todos os usuários devem ser submetidos à mesma taxa de dados.

Descrevem-se a seguir os objetivos específicos perseguidos no decorrer do desenvolvimento desta tese de doutorado.

1. Obtenção do número de usuários por *cluster* e por célula que apresente a máxima eficiência energética com atendimento à restrição de equidade da taxa de dados para todos os usuários em sistemas NOMA e MIMO-NOMA.
2. Estabelecimento da melhor política de alocação de potência nos sistemas NOMA e MIMO-NOMA em termos de coeficientes de alocação de potência capaz de garantir a mesma taxa (abordagem de equidade) para todos os usuários.
3. Maximização da EE e da equidade de taxa (índice de equidade de Jain) entre todos os usuários no sistema NOMA.
4. Obtenção do melhor compromisso entre EE-SE e portanto da melhor eficiência de recursos em sistemas NOMA.
5. Otimização conjunta da EE e da soma das taxas com equidade máxima para assim determinar o melhor compromisso entre EE-SE, obtendo o total de *clusters* na célula que resulta em EE máxima nas diferentes configurações do sistema MIMO-NOMA.
6. Estabelecimento do compromisso SE-EE em sistemas MIMO-NOMA operando sob equidade máxima de taxa sob diferentes números de usuários por *cluster* a partir da formulação do problema de otimização MOO.
7. Investigação das técnicas de programação não linear (NLP - *Nonlinear Programming*), particularmente SQP e DK, que combinadas a métodos de escalarização, especialmente  $\epsilon$ -C, sejam capazes de atingir o melhor compromisso EE-SE em sistemas MIMO-NOMA.

8. Análise e comparação dos métodos de otimização híbridos que combinem os métodos heurísticos gulosos  $\epsilon$ -C e escalonado (HC - *Hill Climbing*), assim como uma comparação com o método de escalarização WS para operarem em sistemas MIMO-NOMA.
9. Avaliação sistemática das matrizes padrão PDMA que atendam adequadamente aos requisitos e características dos serviços 5G visando a otimização do sistema mMIMO PDMA de forma a obter a taxa de bits operacional que resulte na máxima EE do sistema.

### 1.3 Estrutura do Trabalho de Tese

Este texto de tese de Doutorado está organizado da seguinte forma:

Capítulo 1: Introdução.

Este capítulo fornece a caracterização dos problemas investigados e os respectivos objetivos traçados.

Capítulo 2: Sistema NOMA.

A modelagem matemática do sistema NOMA é estabelecida ao longo desse capítulo, assim como o desenvolvimento dos problemas de otimização são formulados.

Capítulo 3: Sistema MIMO-NOMA.

A modelagem matemática do sistema MIMO-NOMA é estabelecida ao longo desse capítulo, assim como o desenvolvimento dos problemas de otimização mono-objetiva e bi-objetiva formulados.

Capítulo 4: Sistema mMIMO PDMA.

Este capítulo fornece a modelagem matemática do sistema mMIMO PDMA e dos parâmetros das matrizes padrão para auxiliar na identificação daquela que melhor atende as especificações de cada um dos serviços 5G (eMBB, mMTC e uRLLC).

Capítulo 5: Resultados e Discussões.

Neste capítulo, são apresentados os resultados e as discussões dos métodos e procedimentos propostos aplicados aos mais variados cenários práticos de interesse, tendo em vista corroborar a formulação e conjecturas propostas.

Capítulo 6: Relação Entre os Trabalhos.

O objetivo deste capítulo é mostrar o que há em comum entre os temas abordados.

Capítulo 7: Conclusões e Perspectivas.

Neste capítulo são apresentadas as principais conclusões e direções para continuidade do trabalho de pesquisa tendo em vista os resultados da pesquisa alcançados até o momento.

## 1.4 Submissões e Publicações

Esta seção elenca os trabalhos científicos (artigos completos) desenvolvidos durante o período de doutorado, divididos em artigos publicados e submetidos.

### 1.4.1 Artigos Publicados

- P1. **Título:** "NOMA Systems Optimization to Ensure Maximum Fairness to Users"  
**Autores:** Jaime Laelson Jacob, Taufik Abrão  
**Jornal:** *ETT - Transactions on Emerging Telecommunications Technologies*  
**Data da Publicação:** Fevereiro 2020  
**Disponível em:** <<https://doi.org/10.1002/ett.3875>>
- P2. **Título:** "Energy and Spectral Efficiencies Trade-off in MIMO-NOMA System under User-rate Fairness and Variable User per *Cluster*"  
**Autores:** Jaime Laelson Jacob, Cristiano Magalhães Panazio, Taufik Abrão  
**Jornal:** *PHYCOM - Physical Communication*  
**Data da Publicação:** Maio 2021.  
**Disponível em:** <<https://doi.org/10.1016/j.phycom.2021.101348>>
- P3. **Título:** "Non-Linear Biobjective EE-SE Optimization for NOMA-MIMO Systems under User-rate Fairness and Variable Number of Users per *Cluster*"  
**Autores:** Jaime Laelson Jacob, Cristiane A. Pendeza Martinez, André Luís Machado Martinez, Taufik Abrão  
**Jornal:** *AEUE - International Journal of Electronics and Communications*  
**Data da Publicação:** Agosto 2021.  
**Disponível em:** <<https://doi.org/10.1016/j.aeue.2021.153870>>

### 1.4.2 Artigo em Processo de Re-submissão

P4. **Título:** "PDMA Large-Scale MIMO Energy Efficiency Maximization for eMBB, mMTC and uRLLC"

**Autores:** Jaime Laelson Jacob, Taufik Abrão

**Jornal:** *AEUE - International Journal of Electronics and Communications*

**Data da Submissão:** 09 de Junho de 2022.

**Re-submissão prevista:** Janeiro de 2023.

**Disponível:** ANEXO [D](#).

## 2 Sistema NOMA

Neste capítulo desenvolvemos uma formulação matemática para a descrição de funcionamento do sistema NOMA. Para o esquema de transmissão NOMA operando no domínio de potência, resolve-se o problema de alocação de potência e máxima eficiência energética com a restrição de mesma taxa de dados entre todos os usuários (de borda e internos). A Figura 2.1 esquematiza um sistema celular NOMA multi-usuário, contendo  $M$  usuários localizados em um setor da célula circular de raio  $\mathfrak{R}_D$ .

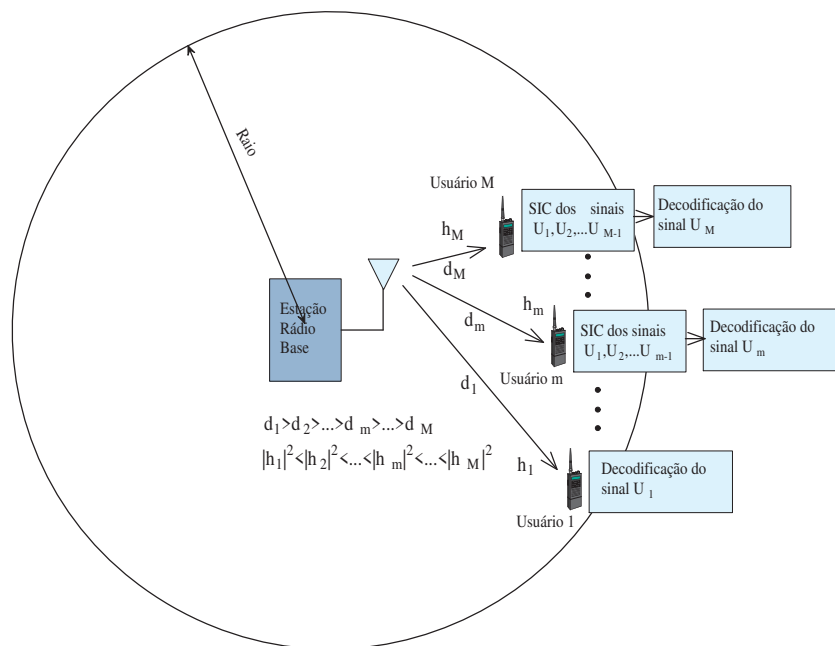


Figura 2.1 – NOMA com  $M$  usuários e de raio  $\mathfrak{R}_D$ .

### 2.1 Modelo do Sistema NOMA

Considere um sistema celular com cenário de transmissão com enlace de descida, sendo uma BS localizada no centro do disco com raio  $\mathfrak{R}_D$  e  $M$  usuários alinhados, Figura 2.1. Todos os terminais estão equipados com uma única antena. O canal entre os usuários e a estação base é dado por (DING et al., 2014)

$$h_i = \frac{g_i}{\sqrt{1 + d_i^\alpha}} \quad \text{para } i = 1, 2, \dots, M \quad (2.1)$$

no qual  $g_i$  representa o ganho do canal com desvanecimento Rayleigh,  $\alpha$  é o expoente de perda de percurso e  $d_i$  representa a distância do usuário à BS. Sem perda de generalidade, os canais são classificados como  $|h_1|^2 \leq |h_2|^2 \leq \dots \leq |h_M|^2$ . O sinal de transmissão na BS

com aplicação da técnica NOMA é dado por

$$x = \sum_{i=1}^M \sqrt{\beta_i P} s_i \quad (2.2)$$

no qual  $s_i$  é a mensagem do  $i$ -ésimo usuário,  $P$  é a potência de transmissão e  $\beta_i$  é o coeficiente de alocação de potência, ou seja,  $\beta_1 \geq \beta_2 \geq \dots \geq \beta_M$ . Há a superposição de  $M$  sinais para seus  $M$  usuários via divisão no domínio da potência. O sinal recebido no  $m$ -ésimo usuário é dado por

$$y_m = h_m \sum_{i=1}^M \sqrt{\beta_i P} s_i + z_m \quad (2.3)$$

no qual  $z_i$  é ruído Gaussiano branco aditivo complexo (AWGN - *Additive White Complex Gaussian Noise*) i.i.d. com média zero e variância de  $\sigma_n^2$ . No modelo de sistema apresentado aqui, apenas o usuário 1 não usa a técnica SIC na detecção de sinal, conseqüentemente os demais usuários são tratados como ruído. A técnica SIC é inicialmente aplicada ao usuário 2, na qual a interferência causada pelo sinal do usuário 1 sobre o usuário 2 é totalmente removida. O privilégio do usuário mais próximo do BS, ou seja, o  $M$ -ésimo usuário é ter os sinais de interferência totalmente extraídos. Para calcular a capacidade, é necessário primeiro calcular a relação de potência SINR recebida. Da Eq. (2.3) podemos obter o SINR genérico para  $1 \leq m \leq (M - 1)$  dado por

$$\gamma_m = \frac{P|h_m|^2\beta_m}{P|h_m|^2\sum_{i=m+1}^M\beta_i + \sigma_n^2}, \quad m = 1, \dots, M - 1 \quad (2.4)$$

e para o último usuário (próximo da BS) resulta:

$$\gamma_M = \frac{P|h_M|^2\beta_M}{\sigma_n^2} \quad (2.5)$$

capacidade do canal é dada por  $\log_2(1 + \text{SINR})$  portanto a capacidade dos usuários é dada por (DING et al., 2014)

$$R_m = \log_2(1 + \gamma_m), \quad m = 1, \dots, M \quad [\text{bits/s/Hz}]. \quad (2.6)$$

Neste trabalho, presume-se que a interferência no primeiro usuário foi completamente cancelada ao usar a técnica SIC.

A EE é dada pela razão entre a soma das taxas alcançáveis por todos os usuários do sistema e a energia total consumida, dada por

$$\text{EE} = \frac{R_\Sigma}{P_t + MP_c} = \frac{MR}{P_t + MP_c}, \quad \left[ \frac{\text{bits}}{\text{J} \cdot \text{Hz}} \right] \quad (2.7)$$

no qual  $R_\Sigma = \sum_{m=1}^M R_m = MR$  é a soma das taxas transmitidas por todos os usuários dada pela Eq. (2.6) onde todos os usuários têm a mesma taxa, ou seja,  $R_m = R$ ; a potência

total da frequência de rádio (RF - *Radio Frequency*) consumida é dada por:

$$P_t = \varrho \cdot \sum_{i=1}^M \beta_i P \quad [\text{W}] \quad (2.8)$$

no qual  $\varrho \geq 1$  é a ineficiência do amplificador de potência de RF, assumida igual para todos os amplificadores que alimentam cada antena Tx;  $\beta_i$  é o coeficiente de alocação de potência, e  $P_c$  é a potência do circuito necessária para o processamento do sinal no lado do transmissor, assumindo um valor fixo por usuário.

## 2.1.1 Otimização do Sistema NOMA para o Enlace de Descida

Nesta seção, a otimização do sistema NOMA é feita de duas maneiras. No primeiro processo de otimização pretende-se buscar a menor potência total de transmissão e no segundo processo a máxima EE, ambos os processos de otimização sob a mesma restrição de equidade de taxa de dados para todos os usuários. Na sequência, descrevemos o problema da eficiência de recursos (RE) e o da equidade de taxas envolvendo usuários de borda e aqueles localizados próximos à BS.

### 2.1.1.1 Otimização de Potência para o NOMA

A formulação para obter a proporção da potência total a uma taxa mínima para cada usuário, em que a expressão para as proporções da potência total foi obtida da Eq. (2.6), são as seguintes:

$$\begin{aligned} \min_{(P, \beta_i)} \quad & P \\ \text{Sujeito a:} \quad & \text{C1 } \sum_{j=1}^M \beta_j = 1 \\ & \text{C2 } \beta_j > 0; P > 0; R > 0 \\ & \text{C3 } \beta_m = (2^R - 1) \sum_{i=m+1}^M \beta_i + (2^R - 1) \frac{\sigma_n^2}{P|h_m|^2} \text{ para } m = 1, \dots, M-1 \\ & \text{C4 } \beta_M = (2^R - 1) \frac{\sigma_n^2}{P|h_M|^2} \end{aligned} \quad (2.9)$$

Na restrição C1 garantimos que a soma das potências dos usuários resulta em 100%. Para garantir que a proporção de potência de cada usuário, potência total e a taxa de dados sejam positivas acrescenta-se a restrição C2 na Eq. (2.9). Finalmente, as restrições C3 e C4 definem as equações para a obtenção das proporções de potência de cada usuário. Para que a Eq. (2.9) seja factível precisamos provar que o conjunto de restrições é convexo. Para isso, a restrição C1 pode ser desenvolvida substituindo as restrições C3 e C4 na restrição C1 sujeita às restrições de C2 para as quais a seguinte equação é obtida.

$$f = \sum_{m=1}^{M-1} \left( (2^R - 1) \sum_{i=m+1}^M \beta_i + (2^R - 1) \frac{\sigma_n^2}{P|h_m|^2} \right) + (2^R - 1) \frac{\sigma_n^2}{P|h_M|^2} - 1 \quad (2.10)$$

O Hessiano da Eq. (2.10) deve ser semi-definida positiva, ou seja,  $\nabla^2 f \succcurlyeq 0$  para ser convexa (BOYD; VANDENBERGHE, 2004). Portanto, os autovalores do Hessiano devem ser maiores ou iguais a zero. Depois de aplicar o Hessiano e aplicar a definição de autovalores na Eq. (2.10), obtemos o seguinte polinômio característico

$$-\lambda^M + \frac{2(2^R - 1)\sigma_n^2 \lambda^{M-1}}{\sum_{m=1}^{M-1} |h_m|^2 P^3} + \frac{2(2^R - 1)\sigma_n^2 \lambda^{M-1}}{P^3 |h_M|^2} = 0. \quad (2.11)$$

Colocamos  $\lambda^{M-1}$  em evidência, com isso, temos que existem  $M - 1$  autovalores iguais a zero e portanto o resultado é dado pela equação

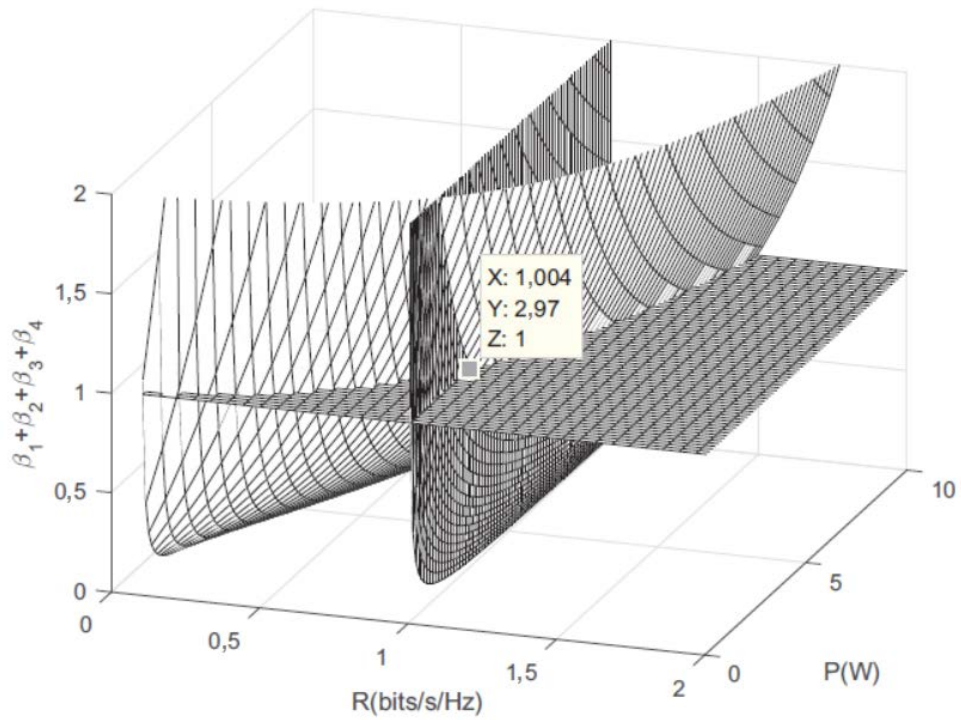
$$\lambda = \frac{2(2^R - 1)\sigma_n^2}{P^3} \left( \frac{1}{\sum_{m=1}^{M-1} |h_m|^2} + \frac{1}{|h_M|^2} \right) \quad (2.12)$$

Como a variável  $P$  é positiva, assim como as constantes  $R$ ,  $\sigma_n^2$  e  $|h_i|^2$ , os autovalores são positivos ou iguais a zero e a equação em questão é semi-definida positiva e, portanto, as restrições formam um conjunto convexo.

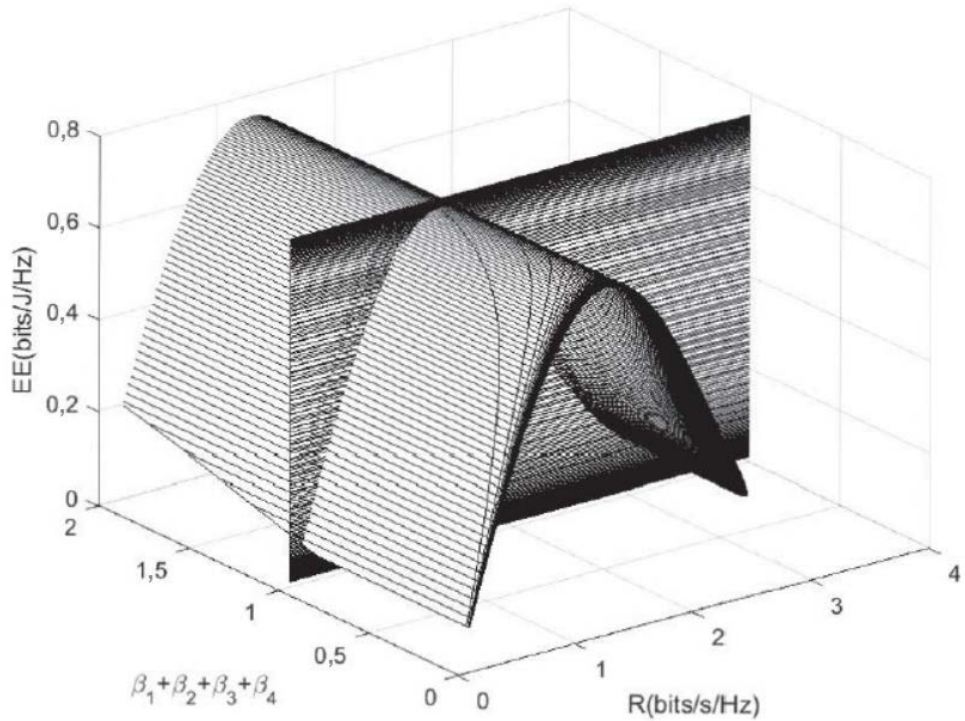
As restrições formam uma coleção de conjuntos convexos e, portanto, a interseção desses conjuntos é convexa. Isso pode ser confirmado pela Figura 2.2.(a) na qual temos a intersecção não-nula composta pelo plano gerado pela restrição C1 na horizontal, pelo plano vertical em que todos os usuários apresentam a mesma taxa  $R = 1$  bit/s/Hz e por último pelo plano resultante da soma da proporção de potência total para diferentes valores de taxa  $R$  e potência  $P$ . De acordo com essa figura, as restrições definidas em (2.9) formam um conjunto fechado, desta forma o problema de otimização apresentado tem uma solução com ótimo global. Na mesma figura pode-se observar o subespaço dado pela condição de  $\beta_1 + \beta_2 + \beta_3 + \beta_4 = 1$  e pela taxa de cada usuário  $R = 1$  bits/s/Hz. Nesse caso, a solução é obtida para a potência total  $P \approx 2,97$  W, entretanto, os valores da proporção de potência em questão não podem ser obtidos tão facilmente.

### 2.1.1.2 Otimização da Alocação de Potência do Usuário a partir da Perspectiva EE no NOMA

A Eq. (2.7) é quase-concava (ZHANG et al., 2017), portanto, a função objetivo deve ser maximizada. Usamos o método de Dinkelbach para transformar uma função objetivo quase côncava em uma côncava (TANG et al., 2019). O problema (2.13) pode ser resolvido como indicado no Algoritmo 1, ao qual  $U_R = MR$  e  $U_T = \varrho \sum_{j=1}^M \beta_j P + MP_c$ . O problema de otimização EE para obter a alocação de potência ótima  $P$ , taxa de dados  $R$



(a)  $\beta$ 's  $\times$  Taxa  $\times$  Potência



(b)  $EE \times \beta$ 's  $\times$  Taxa, com potência  $P_c = 1W$

Figura 2.2 – Conjunto fechado para  $\sigma_n^2 = 1\mu W$  e  $M = 4$  usuários com distâncias  $d_1 = 1,0$  km,  $d_2 = 0,7$  km,  $d_3 = 0,5$  km e  $d_4 = 0,3$  km.

e a proporção de potência  $\beta_i$  para todos os usuários pode ser formulado como

$$\begin{aligned}
& \max_{(P, \beta_j, R)} MR - q^{(n)} (\rho \beta_j P + MP_c) \\
\text{Sujeito a: } & \text{C1 } \sum_{j=1}^M \beta_j = 1 \\
& \text{C2 } \beta_j > 0; P > 0; R > 0 \\
& \text{C3 } \beta_m = (2^R - 1) \sum_{i=m+1}^M \beta_i + (2^R - 1) \frac{\sigma_n^2}{P|h_m|^2} \text{ para } m = 1, \dots, M-1 \\
& \text{C4 } \beta_M = (2^R - 1) \frac{\sigma_n^2}{P|h_M|^2}
\end{aligned} \tag{2.13}$$

Devemos ressaltar aqui que além de  $P$  e  $\beta_i$  serem variáveis de otimização, a taxa  $R$  também é uma variável de otimização. Portanto, ao analisar a convexidade das restrições em relação à potência total  $P$ , proporção da potência total entre os usuários  $\beta_i$  e a taxa de dados  $R$ , pode-se concluir que o conjunto de restrições (C1) a (C4) na Eq. (2.13) é não convexo, pois os autovalores do Hessiano para essas variáveis resultam em valores complexos.

---

**Algoritmo 1:** EE-NOMA – Alocação iterativa de proporção de potência com base no método Dinkelback.

---

```

1 início
2    $n = 0$  e  $q^{(n)} = 0$ 
3   enquanto  $\varepsilon > 0$  faça
4     Para um dado  $q^{(n)}$ , resolver as Eq. (2.13) e/ou (2.14) para obter primeiro a
       alocação de potência  $P^{(n)}$ , proporção de potência dos usuários  $\beta_k^{(n)}$  e taxa
       de dados  $R^{(n)}$  e/ou segundo a alocação de potência  $P^{(n)}$  e a proporção de
       potência dos usuários  $\beta_k^{(n)}$ ;
5     se  $U_R(P^{(n)}, R^{(n)}, \beta_k^{(n)}) - q^{(n)} U_T(P^{(n)}, R^{(n)}, \beta_k^{(n)}) \leq \varepsilon$  então
6        $P^* = P^{(n)}, R^* = R^{(n)}, q^* = q^{(n)}, \beta_k^* = \beta_k^{(n)}$ ;
7       para  $k = 1, \dots, K$ 
8     senão
9       Definir  $n = n + 1$  e  $q^{(n)} = \frac{U_R(P^{(n-1)}, R^{(n-1)}, \beta_k^{(n-1)})}{U_T(P^{(n-1)}, R^{(n-1)}, \beta_k^{(n-1)})}$ ;
10    fim
11  fim
12 fim

```

---

No entanto, ao tornar constante a taxa de dados  $R$  mantendo a potência total  $P$  e a proporção da potência total  $\beta_i$  entre os usuários como variáveis de otimização, o problema na Eq. (2.13) torna-se côncava quando a Eq. (2.10) é obtida. Como provado, anteriormente este conjunto de restrições com tais variáveis é convexo. Assim, a equação do problema (2.13) é reformulada como Eq. (2.14). Usamos o método de Dinkelbach para transformar uma função objetiva quase côncava em uma côncava, conforme Algoritmo 1

da equação do problema (2.14).

$$\begin{aligned} \max_{(P, \beta_j)} \quad & MR - q^{(n)} \left( \varrho \sum_{j=1}^M \beta_j P + MP_c \right) \\ \text{Sujeito a:} \quad & (C1), (C2), (C3), (C4). \end{aligned} \quad (2.14)$$

A Figura 2.2.(b) mostra o conjunto de soluções para o problema de otimização EE. Este conjunto é formado pela soma das proporções de potência igual a 1, ou seja,  $\beta_1 + \beta_2 + \beta_3 + \beta_4 = 1$  e a interseção do outro plano definido pela EE com a taxa do usuário. A EE é resultado dos diferentes valores da taxa dos usuários e da potência total assumidos.

Uma vez que as métricas SE e EE são antagônicas, torna-se necessário o estabelecimento de um critério de equilíbrio (ou importância relativa) entre as eficiências EE e SE. Para isto, definimos a métrica RE, a qual combina ou seja pondera a importância das métricas EE e SE. Detalhamos a definição e as implicações da métrica RE na seção seguinte.

### 2.1.1.3 Eficiência de Recursos (RE)

Em projetos de sistemas convencionais, em vez de focar no SE ou EE separadamente, é muito mais efetivo equilibrar ou ponderar a importância da SE e EE atingíveis, adotando-se a métrica de RE (HUANG et al., 2018). O SE é definido para um sistema de célula única com  $M$  usuários de antena única como  $SE = \sum_{i=1}^M R_i$  (HUANG et al., 2018). A RE pode ser expressa como uma soma ponderada do EE e o SE:

$$RE = \xi \cdot EE + SE \quad (2.15)$$

no qual  $\xi$  é o fator de ponderação em [W], controlando os pesos de EE e SE no projeto. Portanto, quando  $\xi = 0$  a expressão (2.15) se reduz a SE.

Observe-se que o critério de projeto focado na equidade de taxa impacta no ponto de operação do sistema NOMA guiado pela métrica RE. A igualdade de taxas entre os usuários resulta da busca pela máxima equidade de taxas entre os usuários de borda e internos. Para entender melhor a métrica de equidade, na próxima seção é apresentado a equação de Jain que mede a equidade dos usuários no sistema NOMA.

### 2.1.1.4 Equidade para o NOMA

A equidade de taxa pode ser obtida a partir do *índice de equidade de Jain* (JAIN et al., 1984), (WANG et al., 2017):

$$\text{Índice de Equidade de Jain: } \mathcal{F}_j = \frac{\left( \sum_{k=1}^M R_k \right)^2}{M \sum_{k=1}^M R_k^2}. \quad (2.16)$$

No contexto do projeto de otimização deste trabalho, a mesma taxa alvo foi adotada para todos os usuários; portanto, verifica-se pela Eq. (2.16) que a equidade calculada resulta em  $\mathcal{F}_j = 1$  que representa seu valor máximo. Para fins de ilustração, vamos considerar o seguinte exemplo: assumindo  $M = 4$  usuários, com  $R_1 = R_2 = R_3 = 1$  e  $R_4 = 0.7$  resulta em  $(\sum_{k=1}^4 R_k)^2 = 13.69$  e  $4 \cdot \sum_{k=1}^4 R_k^2 = 13.96$ , então  $\mathcal{F}_j = 0.98$  muito próximo de 1, visto que apenas um usuário não tem uma taxa igual aos outros usuários.

Uma forma alternativa de definir equidade é baseada puramente nas considerações da teoria da informação (GUI et al., 2019) que pode ser expressa por:

$$\mathcal{F}_{it} = 1 - \frac{\hat{\Delta}^2}{\hat{R}^2} \quad (2.17)$$

no qual  $\hat{\Delta}^2 = \frac{1}{M-1} \sum_{m=1}^M \Delta_m^2$  é a estimativa imparcial dos desvios da taxa quadrática média e  $\hat{R}^2 = \frac{R_\Sigma^2}{M}$  é a taxa quadrática média. O  $\Delta_m^2$  é o quadrado dos desvios das taxas reais do usuário de sua taxa de equidade ou a taxa igual para todos os usuários, conforme segue

$$\Delta_m^2 = (R_m^f - R_m)^2, \quad m = 1, 2, \dots, M \quad (2.18)$$

A taxa de equidade para cada usuário pode ser representada por

$$R_m^f = \log_2 \left( 1 + \frac{P |h_m| \beta_m}{\sigma_n^2} \right) \cdot \frac{R_\Sigma}{R_c} \quad (2.19)$$

ao qual  $R_\Sigma$  é a soma das taxas dos usuários e finalmente a taxa cumulativa é dada por

$$R_c = \sum_{m=1}^M \log_2 \left( 1 + \frac{P |h_m| \beta_m}{\sigma_n^2} \right). \quad (2.20)$$

## 3 Sistema MIMO-NOMA

Neste capítulo desenvolvemos uma formulação matemática para a descrição de funcionamento dos sistemas MIMO-NOMA. Para os esquemas de transmissão MIMO-NOMA operando no domínio de potência, resolve-se o problema de alocação de potência e máxima eficiência energética com a restrição de equidade de taxa de dados entre todos os usuários (de borda e internos). Adicionalmente, a otimização BOO para o problema de compromisso EE-SE em sistemas MIMO-NOMA foi desenvolvida.

### 3.1 Modelo do Sistema MIMO-NOMA

Considere uma transmissão DL em um sistema de células MIMO combinado com um sistema NOMA no qual a BS está equipada com  $A$  antenas, conforme ilustrado na Figura 3.1.

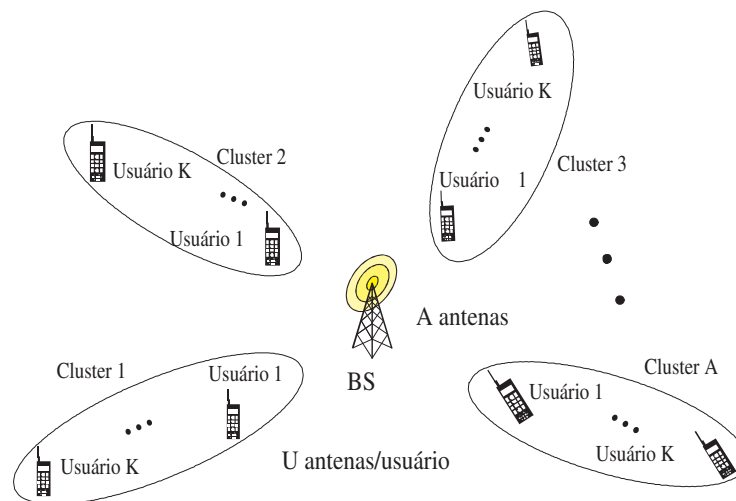


Figura 3.1 – Sistema celular com uma BS com  $A$  antenas,  $A \cdot K$  usuários com  $A$  clusters,  $K$  usuários/clusters e cada usuário tem  $U$  antenas ( $U \geq A$ ).

Para ter um cenário NOMA, o sistema celular é dividido em  $A$  clusters agrupados aleatoriamente. Existem  $K \geq 2$  usuários para cada cluster, onde os usuários têm  $U$  antenas cada um. Aqui, assumimos  $U \geq A$ <sup>1</sup> para tornar o sistema viável. Todos os clusters possuem

<sup>1</sup>A condição  $U \geq A$  pode ser justificada pelo surgimento das redes 5G ultradensas implementadas a partir do conceito de micro BS (reduzida área de cobertura) de baixo custo e potência. Essas BS podem apresentar configurações de antenas contendo o mesmo número ou mesmo um número menor de antenas em relação aos dispositivos do usuário de baixa mobilidade e equipados com dezenas de antenas. Isto se deve ao crescimento e rápida popularização de *smartphones*, *tablets* e dos dispositivos e sensores da internet das coisas, empregados em casas e edifícios inteligentes, etc (WANG et al., 2016; ZENG et al., 2018).

o mesmo número de usuários e todos os usuários pertencem à mesma célula (sistema de célula única); portanto, sem perda de generalidade, o número total de usuários na célula é selecionado para ser múltiplo do número de usuários por *cluster*, ou seja, igual a  $K \cdot A$ . O sinal transmitido pelo BS ao usuário pode ser descrito por (ZENG et al., 2018)

$$\mathbf{x} = \hat{\mathbf{s}} \quad (3.1)$$

no qual o vetor de dados  $\hat{\mathbf{s}} \in \mathbb{C}^{A \times 1}$  pode ser escrito como:

$$\hat{\mathbf{s}} = \begin{bmatrix} \sqrt{P\beta_{1,1}}s_{1,1} + \cdots + \sqrt{P\beta_{1,k}}s_{1,k} + \cdots + \sqrt{P\beta_{1,K}}s_{1,K} \\ \vdots \\ \sqrt{P\beta_{a,1}}s_{a,1} + \cdots + \sqrt{P\beta_{a,k}}s_{a,k} + \cdots + \sqrt{P\beta_{a,K}}s_{a,K} \\ \vdots \\ \sqrt{P\beta_{A,1}}s_{A,1} + \cdots + \sqrt{P\beta_{A,k}}s_{A,k} + \cdots + \sqrt{P\beta_{A,K}}s_{A,K} \end{bmatrix} = \begin{bmatrix} \hat{s}_1 \\ \vdots \\ \hat{s}_a \\ \vdots \\ \hat{s}_A \end{bmatrix}. \quad (3.2)$$

ao qual  $s_{a,k}$  e  $\beta_{a,k}$  representam o sinal e os coeficientes de proporção de potência para o  $a$ -ésimo *cluster* e  $k$ -ésimo usuário, respectivamente, e  $P$  indica a potência total transmitida pela BS. Além disso, os *coeficientes de proporção de potência* devem satisfazer

$$\sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} = 1. \quad (3.3)$$

Portanto, para obter equidade na taxa de dados entre todos os usuários NOMA, aqueles usuários com menor ganho de canal em BS receberão  $\beta_{a,k}$  maior. Justamente, tais valores serão obtidos a partir de processos de otimização.

Consequentemente, no DL, o sinal recebido por um determinado usuário  $k$  pertencente ao  $a$ -ésimo *cluster* é dado por

$$\mathbf{y}_{a,k} = \mathbf{H}_{a,k}\hat{\mathbf{s}} + \mathbf{z}_{a,k}, \quad \text{no qual } \mathbf{H}_{a,k} = \frac{\mathbf{Q}_{a,k}}{\sqrt{1 + d_{a,k}^\alpha}}. \quad (3.4)$$

Na Eq. (3.4),  $\mathbf{Q}_{a,k} \in \mathbb{C}^{U \times A}$  representa uma matriz de canal de desvanecimento Rayleigh,  $\alpha$  é o expoente de perda de percurso,  $d_{a,k}$  representa a distância do  $k$ -ésimo usuário que pertence ao  $a$ -ésimo *cluster* da BS, e  $\mathbf{z}_{a,k}$  é um vetor AWGN, ao qual  $\mathbf{z}_{a,k} \in \mathbb{C}^{U \times 1}$  e  $\mathcal{CN}(0, \sigma^2 \mathbf{I})$ . O sinal recebido deve ser multiplicado pelo vetor *filtro de detecção*  $\mathbf{v}_{a,k}$  para remover completamente a interferência inter *cluster* (ZENG et al., 2017b). Dessa forma, pode-se identificar o termo do *cluster* interferente:

$$\mathbf{v}_{a,k}^H \mathbf{y}_{a,k} = \mathbf{v}_{a,k}^H \mathbf{h}_{a,ak} \sum_{k=1}^K \sqrt{P\beta_{a,k}}s_{a,k} + \underbrace{\sum_{i=1, i \neq a}^A \mathbf{v}_{a,k}^H \mathbf{h}_{i,ak} \hat{s}_i}_{\text{clusters interferentes}} + \mathbf{v}_{a,k}^H \mathbf{z}_{a,k} \quad (3.5)$$

no qual  $\hat{s}_i$  representa o  $i$ -ésimo elemento de  $\hat{\mathbf{s}}$  na Eq. (3.2), e  $\mathbf{h}_{\ell,ak}$  é a  $\ell$ -ésima coluna da matriz  $\mathbf{H}_{a,k}$ . Esta coluna representa a antena que atende o  $i$ -ésimo *cluster* dentre  $A$  *clusters*.

Para remover a interferência de outros *clusters*, a matriz de detecção deve obedecer à condição  $\mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_i = 0$  para qualquer  $i \neq a$ . Esta restrição é reduzida a  $\mathbf{v}_{a,k}^H \bar{\mathbf{H}}_{a,k} = 0$  no qual  $\bar{\mathbf{H}}_{a,k} = [\mathbf{h}_{1,ak} \cdots \mathbf{h}_{a-1,ak} \mathbf{h}_{a+1,ak} \cdots \mathbf{h}_{A,ak}]$  e  $\mathbf{h}_{j,ak}$  é a  $j$ -ésima coluna de  $\mathbf{H}_{a,k}$ . Temos (DING et al., 2016a)

$$\mathbf{v}_{a,k} = \mathbf{U}_{a,k} \mathbf{g}_{a,k}, \quad (3.6)$$

no qual  $\mathbf{U}_{a,k}$  é a matriz dos vetores singulares à esquerda de  $\bar{\mathbf{H}}_{a,k}$  após a decomposição em valores singulares diferentes de zero e o vetor de combinação de proporção máxima (MRC), expresso como

$$\mathbf{g}_{a,k} = \frac{\mathbf{U}_{a,k}^H \mathbf{h}_{a,ak}}{\|\mathbf{U}_{a,k}^H \mathbf{h}_{a,ak}\|}. \quad (3.7)$$

Depois de remover a interferência dos outros *clusters*, a Eq. (3.5) pode ser reescrita como

$$\mathbf{v}_{a,k}^H \mathbf{y}_{a,k} = \mathbf{v}_{a,k}^H \mathbf{h}_{a,ak} \sum_{k=1}^K \sqrt{P \beta_{a,k}} \mathbf{s}_{a,k} + \mathbf{v}_{a,k}^H \mathbf{z}_{a,k} \quad (3.8)$$

Sem perda de generalidade, os ganhos efetivos do canal são ordenados como (DING et al., 2016a)

$$|\mathbf{v}_{a,1}^H \mathbf{h}_{a,a1}|^2 \geq \cdots \geq |\mathbf{v}_{a,K}^H \mathbf{h}_{a,aK}|^2, \quad \forall a \in \{1, \dots, A\} \quad (3.9)$$

e os coeficientes de proporção de potência são ordenados da seguinte forma

$$\beta_{1,1} \leq \cdots \leq \beta_{1,K}. \quad (3.10)$$

A BS ainda precisa ser informada sobre o ganho de canal para em seguida calcular a ordem escalar do ganho de canal efetivo para implementar o NOMA conforme mostrado em (3.9).

Na recepção, cada usuário aplica SIC para remover a interferência de outros usuários com pior ganho de canal. De maneira genérica, os sinais dos usuários interferentes  $(a, k+1), \dots, (a, K)$  são removidos pelo usuário  $(a, k)$ . Consideramos que todas as interferências são removidas porque usamos um SIC perfeito. Desta forma, a relação SINR para o  $k$ -ésimo usuário no  $a$ -ésimo *cluster* é dada por (DING et al., 2016a)

$$\gamma_{a,k} = \frac{P |\mathbf{v}_{a,k}^H \mathbf{h}_{a,ak}|^2 \beta_{a,k}}{P \sum_{j=1}^{k-1} \beta_{a,j} |\mathbf{v}_{a,k}^H \mathbf{h}_{a,ak}|^2 + |\mathbf{v}_{a,k}|^2 \sigma^2} \quad (3.11)$$

para  $1 \leq a \leq A$  e  $2 \leq k \leq K$ . O primeiro usuário resulta no seguinte SINR:

$$\gamma_{a,1} = \frac{P |\mathbf{v}_{a,1}^H \mathbf{h}_{a,a1}|^2 \beta_{a,1}}{|\mathbf{v}_{a,1}|^2 \sigma^2} \quad (3.12)$$

enquanto a taxa de dados normalizada ou eficiência espectral para o  $k$ -ésimo usuário do  $a$ -ésimo *cluster* NOMA alcançado por um usuário em um determinado *cluster* resulta:

$$R_{a,k} = \log_2(1 + \gamma_{a,k}) \quad \left[ \frac{\text{bits}}{\text{s} \cdot \text{Hz}} \right] \quad (3.13)$$

A eficiência energética é dada pela razão entre a soma das taxas alcançáveis pelos usuários e a energia total consumida no sistema:

$$EE = \frac{R_{\Sigma}}{P_{\text{RF}} + P_{\text{C}}} = \frac{K \cdot A \cdot R}{P_{\text{RF}} + K \cdot U \cdot A \cdot P_{\text{K}} + A \cdot P_{\text{BS}}} \left[ \frac{\text{bits}}{\text{J} \cdot \text{Hz}} \right]. \quad (3.14)$$

ao qual  $P_{\text{C}}$  é o consumo de energia do circuito necessária para o processamento do sinal, resfriamento, etc, nos transmissores e receptores, assumido ser um valor fixo por usuário e por antena. Para simplificar, apenas as partes  $P_{\text{K}}$  e  $P_{\text{BS}}$  foram explicitamente consideradas, onde a potência necessária para operar o circuito de cada antena de usuário é  $P_{\text{K}}$ , e para operar os componentes do circuito de cada antena da BS é  $P_{\text{BS}}$ . Além disso, o *consumo total de energia de RF* é dado por:

$$P_{\text{RF}} = \varrho \cdot \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} P = \varrho \cdot P \quad [\text{W}] \quad (3.15)$$

no qual  $\varrho \geq 1$  é a ineficiência do amplificador de potência de RF, assumida igual para todos eles alimentando cada antena Tx. Adotando a largura de banda normalizada, a *soma das taxas* normalizada é calculada como o total das taxas de dados transmitidas por todos os usuários sob equidade máxima. Neste caso, todos os usuários atingem a mesma taxa de dados ( $R_{ak} = R$ ) e, conseqüentemente:

$$R_{\Sigma} = \sum_{a=1}^A \sum_{k=1}^K R_{a,k} = \sum_{a=1}^A \sum_{k=1}^K R = K \cdot A \cdot R. \quad (3.16)$$

para um total de  $K$  usuários/*cluster* e um total de  $A$  *clusters* na célula analisada.

### 3.1.1 Otimização do Sistema MIMO-NOMA para o Enlace de Descida

Nesta seção, realizamos a otimização do sistema MIMO-NOMA de duas maneiras diferentes. Primeiro, obtivemos a potência total mínima e a taxa igual para todos os usuários, usando a estratégia de alocação de potência ótima com taxa igual (ERPA - *Equal-Rate optimal Power Allocation*). Em segundo lugar, calculamos a otimização de EE para encontrar a taxa de dados igual  $R$  que resulta na EE máxima. O conceito de equidade é apresentado na próxima subseção para melhor entender o processo de obtenção de seu valor máximo, seguido pelo problema de otimização de potência e otimização de EE. Na literatura encontramos a otimização de eficiência energética de forma a obter uma taxa mínima entre os usuários não importando o fato de se ter usuários com taxa bastante superior (ALI et al., 2016). Essa discrepância entre os usuários faz com que exista uma baixa equidade de taxas entre os usuários. Em nosso sistema garantimos que as taxas de dados entre os usuários tenham equidade máxima.

### 3.1.1.1 Equidade de Taxa para os Usuários MIMO-NOMA

A equidade pode ser definida pelo *índice de equidade de Jain* (JAIN et al., 1984), (WANG et al., 2017):

$$\mathcal{F}_{\text{Jain}} = \frac{\left( \sum_{a=1}^A \sum_{k=1}^K R_{a,k} \right)^2}{A \cdot K \sum_{a=1}^A \sum_{k=1}^K R_{a,k}^2}. \quad (3.17)$$

No projeto de otimização deste trabalho, adota-se taxa de dados igualitária para todos os usuários; portanto, da Eq. (3.17), a equidade de Jain deve resultar seu valor máximo,  $\mathcal{F}_{\text{Jain}} = 1$ .

### 3.1.1.2 Alocação de Potência Baseada no Inverso da Resposta do Canal (ICR) no MIMO-NOMA

A técnica usada para definir a potência proporcional para cada usuário em (EL-SAYED et al., 2016) é dada pelo inverso da resposta do canal (ICR - *Inverse of the Channel Response*). Usamos esta técnica para comparar a soma das taxas com taxas igualitárias entre todos os usuários obtidos na próxima Subseção (3.1.1.3). A atribuição de potência para o  $j$ -ésimo usuário pertencente ao  $i$ -ésimo *cluster* é definida como:

$$P_{ij} = \underbrace{\left( \sum_{a=1}^A \sum_{k=1}^K \frac{1}{\frac{|\mathbf{v}_{a,k}|^2}{|\mathbf{v}_{a,k}^H \mathbf{h}_{a,ak}|^2}} \right)^{-1}}_{\beta'_{ij}} \cdot \left( \frac{|\mathbf{v}_{i,j}|^2}{|\mathbf{v}_{i,j}^H \mathbf{h}_{i,ak}|^2} \right)^{-1} \cdot P. \quad (3.18)$$

Na Eq. (3.18), o termo  $\beta'_{ij}$  é constituído pelo produto do inverso da soma do inverso do ganho do canal de todos os usuários da célula no primeiro parêntese e o inverso do ganho de canal do usuário de interesse no segundo parêntese. Esta equação representa a porção de potência de cada usuário dada pelo inverso do ganho do canal normalizado  $\beta'_{ij}$  multiplicado pela potência total. Nesse caso, a proporção de potência será maior para usuários com menor ganho de canal e menor para usuários com maior ganho de canal.

A Eq. (3.17) nos permite confirmar que ao adotarmos a igualdade entre as taxas dos usuários da célula do sistema MIMO-NOMA garantimos que a equidade será máxima, ou seja, igual a 1. Portanto não existe a necessidade de usarmos no projeto de otimização da potência total ou da EE. Para termos um comparativo entre a efetividade do nosso projeto de otimização NOMA baseado no critério de equidade de taxa com outro método disponível na literatura (EL-SAYED et al., 2016), adotamos a Eq. (3.18) que estabelece a proporção de potência para cada usuário como o inverso do ganho de canal. Os projetos de otimização por nós desenvolvidos serão apresentados na sequência.

### 3.1.1.3 Alocação de Potência Ótima com Taxa Igual (ERPA) em MIMO-NOMA

As proporções de potência a serem consideradas para cada usuário e *cluster* tendo em vista atingir a potência total mínima com taxa igual para todos os usuários MIMO-NOMA com a estratégia ERPA é dada por

$$\begin{aligned}
& \min_{(P, \beta_{a,k})} P \\
\text{s.a: } & \text{C.1 } \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} = 1 \\
& \text{C.2 } \beta_{a,k} > 0; P > 0 \\
& \text{C.3 } \beta_{a,k} = (2^R - 1) \frac{\sigma^2 |\mathbf{v}_{a,k}|^2}{P |\mathbf{v}_{a,k}^H \mathbf{h}_{a,ak}|^2} + (2^R - 1) \sum_{i=1}^{k-1} \beta_{a,i}, \quad \text{com } \begin{cases} 2 \leq k \leq K \\ 1 \leq a \leq A \end{cases} \\
& \text{C.4 } \beta_{a,1} = \frac{(2^R - 1) \sigma^2 |\mathbf{v}_{a,1}|^2}{P |\mathbf{v}_{a,1}^H \mathbf{h}_{a,ak}|^2} \quad \text{para } 1 \leq a \leq A
\end{aligned} \tag{3.19}$$

no qual o conjunto de restrições é convexo. Assim, vamos mostrar que as restrições C.1-C.4. podem ser incorporadas à função custo convexa. O C.1 pode ser desenvolvido substituindo as restrições C.3 e C.4 na restrição C.1 sujeito às restrições de C.2, obtendo a seguinte equação:

$$\begin{aligned}
f = & \sum_{a=1}^A \sum_{k=2}^K \left[ (2^R - 1) \sum_{i=1}^{k-1} \beta_{a,i} + (2^R - 1) \frac{\sigma^2 |\mathbf{v}_{a,k}|^2}{P |\mathbf{v}_{a,k}^H \mathbf{h}_{a,ak}|^2} \right] + \\
& \sum_{a=1}^A (2^R - 1) \frac{\sigma^2 |\mathbf{v}_{a,1}|^2}{P |\mathbf{v}_{a,1}^H \mathbf{h}_{a,ak}|^2} - 1
\end{aligned} \tag{3.20}$$

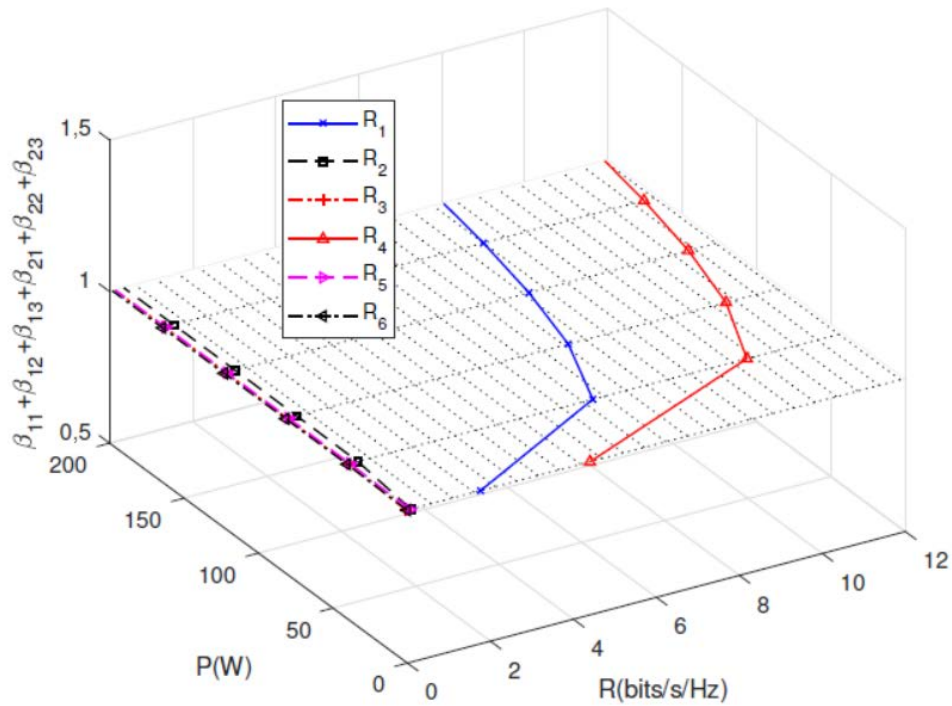
Portanto, para ser convexo, o Hessiano da Eq. (3.20) deve ser semi-definido positivo, ou seja,  $\nabla^2 f \succcurlyeq 0$  (BOYD; VANDENBERGHE, 2004). Portanto, todos os autovalores do Hessiano devem ser maiores ou iguais a zero. Após obter o Hessiano e aplicar a definição dos autovalores na Eq. (3.20), o seguinte polinômio característico pode ser obtido:

$$-\lambda^M + \frac{2(2^R - 1)\sigma^2}{P^3} \sum_{a=1}^A \sum_{k=2}^K \frac{|\mathbf{v}_{a,k}|^2}{|\mathbf{v}_{a,k}^H \mathbf{h}_{a,ak}|^2} \lambda^{M-1} + \frac{2(2^R - 1)\sigma^2}{P^3} \sum_{a=1}^A \frac{|\mathbf{v}_{a,1}|^2}{|\mathbf{v}_{a,1}^H \mathbf{h}_{a,ak}|^2} \lambda^{M-1} = 0 \tag{3.21}$$

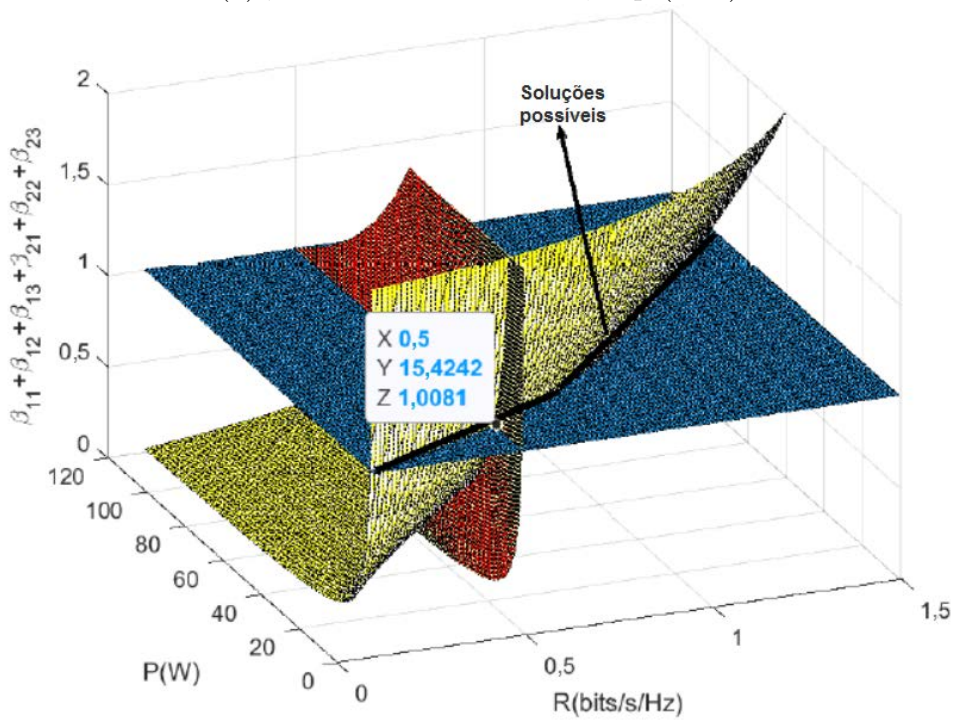
Colocando  $\lambda^{M-1}$  em evidência, existem  $M - 1$  autovalores iguais a zero e, portanto, o valor obtido para o último é:

$$\lambda = \frac{2(2^R - 1)\sigma^2}{P^3} \left( \sum_{a=1}^A \sum_{k=2}^K \frac{|\mathbf{v}_{a,k}|^2}{|\mathbf{v}_{a,k}^H \mathbf{h}_{a,ak}|^2} + \sum_{a=1}^A \frac{|\mathbf{v}_{a,1}|^2}{|\mathbf{v}_{a,1}^H \mathbf{h}_{a,ak}|^2} \right) \tag{3.22}$$

Como a variável  $P$  é positiva, assim como as constantes  $R$ ,  $\sigma^2$ ,  $\frac{|\mathbf{v}_{a,k}|^2}{|\mathbf{v}_{a,k}^H \mathbf{h}_{a,ak}|^2}$  e  $\frac{|\mathbf{v}_{a,1}|^2}{|\mathbf{v}_{a,1}^H \mathbf{h}_{a,ak}|^2}$ , os autovalores são positivos ou iguais a zero. Portanto, Eq. (3.20) é positiva semi-definida.



(a)  $\beta$ 's  $\times$  Potência  $\times$  Taxa, Eq. (3.18).



(b)  $\beta$ 's  $\times$  Taxa  $\times$  Potência, problema (3.19).

Figura 3.2 – Conjunto fechado para  $\alpha = 3$ ,  $\sigma_n^2 = 1\mu$  W,  $A = 2$  clusters e  $K = 3$  usuários/cluster com distâncias  $d_{11} = 0,08$  km,  $d_{12} = 0,15$  km,  $d_{13} = 0,28$  km,  $d_{21} = 0,07$  km,  $d_{22} = 0,15$  km e  $d_{23} = 0,27$  km.

Na Figura 3.2.(a), os valores de  $R_1$ ,  $R_2$  e  $R_3$  são as taxas dos usuários 1, 2 e 3 do cluster 1 e  $R_4$ ,  $R_5$  e  $R_6$  dos usuários 1, 2 e 3 do cluster 2, obtidos aplicando-se o princípio ICR, Eq. (3.18). Observe que neste cenário de amostra, os dois clusters têm distâncias

semelhantes entre os usuários, e os valores dos respectivos  $R_{a,k}$  estão próximos. Portanto, aplicando (3.18), os valores das taxas de dados  $R_{a,k}$  não resultam em equidade máxima. Aplicando a formulação ERPA e resolvendo o problema de otimização em (3.19), são obtidos os valores ótimos de  $\beta_{a,k}$  e  $P$  que minimizam a potência de transmissão geral de acordo com a Figura 3.2.(b) enquanto mantém-se a taxa de dados igualitária para todos os usuários (curva de "possíveis soluções"). De fato, para uma taxa de bits fixa  $R$ , a potência total é minimizada pela função objetivo sem violar qualquer uma das C.1-C.4 restrições do problema. As restrições formam uma coleção de conjuntos convexos e, portanto, a interseção desses conjuntos é um subconjunto convexo. Conforme a Figura 3.2.(b), as restrições definidas em (3.19) formam um conjunto fechado. Desta forma, o problema de otimização apresentado tem uma solução global factível. Figura 3.2.(b) representa o subespaço dado pela condição  $\beta_{11} + \beta_{12} + \beta_{13} + \beta_{21} + \beta_{22} + \beta_{23} = 1$  no eixo  $z$ , identificado pelo plano horizontal, e neste exemplo, o eixo  $X$  tem o valor de  $R = 0,5$  bits/s/Hz para cada usuário e é representado pelo plano vertical. A solução viável é a interseção dos três planos, resultando em um valor de  $P \approx 15,42$  W para a potência total  $P$  no eixo  $Y$ . No entanto, os valores de proporção de potência  $\beta_{a,k}$  em questão não podem ser obtidos tão facilmente em condições práticas de operação do sistema MIMO-NOMA. É necessário usar uma ferramenta de otimização; especificamente, adotamos a função `fmincon`, disponível no *Toolbox Optimization* do simulador matemático MatLab<sup>®</sup>. Trata-se de um *solver* que resolve problemas de otimização não-lineares com ou sem restrições. Detalhes adicionais são apresentados no Capítulo 5. O traço na Figura 3.2.(b) mostra as soluções viáveis que podem ser dadas pelo problema (3.19). O marcador na figura identifica os valores otimizados para a taxa de dados no eixo  $x$ , a potência total no eixo  $y$  e  $\sum_{a=1}^A \sum_{k=1}^K \beta_{a,k}$  no eixo  $z$ .

#### 3.1.1.4 Problema de Otimização da Eficiência Energética no MIMO-NOMA

As variáveis de otimização para o problema da eficiência energética consideradas são: a) taxa de dados; b) potência total; c) proporção da potência entre os usuários em que a função objetivo é dada pela Eq. (3.14). Essa equação é quase côncava (ZHANG et al., 2017). Usamos o método iterativo de Dinkelbach para transformar a função objetivo do tipo fracional quase-côncava em uma série de funções objetivo côncavas (TANG et al., 2019), conforme expresso pelo problema (3.23). Portanto, o projeto de otimização de EE no sistema MIMO-NOMA para se obter a alocação ótima de potências  $P$ , na qual resulte em valores idênticos de taxa de dados  $R$  (equidade de taxa máxima) e proporção das potências dos usuários nos respectivos *clusters*,  $\beta_{a,k}$ , pode ser definido pelo seguinte

problema equivalente:

$$\begin{aligned}
& \max_{(P, \beta_{a,k}, R)} K \cdot A \cdot R - q^{(n)} \left( \varrho \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} P + K \cdot U \cdot A \cdot P_K + A \cdot P_{BS} \right) \\
\text{s. a: } & \text{C.1 } \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} = 1 \\
& \text{C.2 } \beta_{a,k} > 0; P > 0; R > 0 \\
& \text{C.3 } \beta_{a,k} = (2^R - 1) \frac{\sigma^2 |\mathbf{v}_{a,k}|^2}{P |\mathbf{v}_{a,k}^H \mathbf{h}_{a,ak}|^2} + (2^R - 1) \sum_{i=1}^{k-1} \beta_{a,i}, \quad \text{com } \begin{cases} 2 \leq k \leq K \\ 1 \leq a \leq A \end{cases} \\
& \text{C.4 } \beta_{a,1} = \frac{(2^R - 1) \sigma^2 |\mathbf{v}_{a,1}|^2}{P |\mathbf{v}_{a,1}^H \mathbf{h}_{a,ak}|^2} \text{ para } 1 \leq a \leq A
\end{aligned} \tag{3.23}$$

no qual  $q^{(n)}$  é um valor fixo de EE na  $n$ -ésima iteração da solução. Um pseudo-código para a maximização da eficiência energética iterativa MIMO-NOMA com base no método Dinkelback descrito em (3.23) é descrito no Algoritmo 2, sendo o valor da EE na  $n$ -ésima iteração determinado pela potência e alocação de taxas obtidos na iteração anterior como:

$$q^{(n)} = \frac{U_R(P^{(n-1)}, R^{(n-1)}, \beta_{a,k}^{(n-1)})}{U_T(P^{(n-1)}, R^{(n-1)}, \beta_{a,k}^{(n-1)})} = \frac{K \cdot A \cdot R^{(n-1)}}{\varrho \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k}^{(n-1)} P^{(n-1)} + K \cdot U \cdot A \cdot P_K + A \cdot P_{BS}} \tag{3.24}$$

---

**Algoritmo 2:** EE-MIMO-NOMA – Alocação iterativa de proporção de potência com base no método Dinkelback.

---

```

1 início
2    $n = 0$  e  $q^{(n)} = 0$ 
3   enquanto  $\varepsilon > 0$  faça
4     Para um dado  $q^{(n)}$ , resolver (3.23) para obter  $P^{(n)}$  e  $R^{(n)}$  e  $\beta_{a,k}$ 
5     se  $U_R(P^{(n)}, R^{(n)}, \beta_{a,k}^{(n)}) - q^{(n)} U_T(P^{(n)}, R^{(n)}, \beta_{a,k}^{(n)}) \leq \varepsilon$  então
6        $P^* = P^{(n)}, R^* = R^{(n)}, q^* = q^{(n)}, \beta_{a,k}^* = \beta_{a,k}^{(n)}$ ;
7       para  $a = 1, \dots, A$  e  $k = 1, \dots, K$ 
8         senão
9           Definir  $n = n + 1$  e  $q^{(n)} = \frac{U_R(P^{(n-1)}, R^{(n-1)}, \beta_{a,k}^{(n-1)})}{U_T(P^{(n-1)}, R^{(n-1)}, \beta_{a,k}^{(n-1)})}$ ;
10      fim
11    fim
12 fim

```

---

Ao analisar-se a convexidade das restrições com as variáveis potência total, proporção da potência total entre os usuários e taxa de dados, verificou-se que não é convexa, pois os autovalores do Hessiano com essas variáveis resultam em um valor complexo. No entanto, fixando-se a taxa de dados<sup>1</sup>  $R$ , mantendo-se a potência total  $P$  e a proporção

<sup>1</sup>Taxa igualitária para todos os usuários MIMO-NOMA tendo em vista atingir a equidade de taxa  $R$  para todos os usuários após  $n = 1, 2, \dots, N$  iterações.

da potência total  $\beta_{a,k}$  entre os usuários como variáveis de otimização, a convexidade do conjunto de restrições se mantém, da mesma forma que na Eq. (3.20). Portanto, conforme apontado anteriormente, o conjunto de restrições sob essas condições modificadas resulta em um conjunto convexo.

Como consequência, o problema de maximização de EE-MIMO-NOMA sob o problema de equidade da taxa de usuário utilizando o método iterativo de Dinkelbach pode ser reformulado como:

$$\begin{aligned} \max_{(P, \beta_{a,k})} \quad & K \cdot A \cdot R^* - \mathbf{q}^{(n)} \left( \varrho \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} P + K \cdot U \cdot A \cdot P_K + A \cdot P_{BS} \right) \\ \text{Sujeito a:} \quad & \text{C.1, C.2, C.3 e C.4} \end{aligned} \quad (3.25)$$

No problema Eq. (3.25), novamente, o método de Dinkelbach foi implantado para transformar a função objetivo quase-côncava em uma côncava. O Algoritmo 3 descreve a maximização de EE adotando-se uma taxa de dados fixa. Neste problema reduzido, o valor de EE é fixado na  $n$ -ésima iteração, sendo definido como:

$$\mathbf{q}^{(n)} = \frac{U_R(P^{(n-1)}, \beta_{a,k}^{(n-1)})}{U_T(P^{(n-1)}, \beta_{a,k}^{(n-1)})} = \frac{K \cdot A \cdot R^*}{\varrho \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k}^{(n-1)} P^{(n-1)} + K \cdot U \cdot A \cdot P_K + A \cdot P_{BS}}, \quad (3.26)$$

enquanto a taxa igualitária de dados  $R$  é predefinida, fixa de valor  $R^*$ .

---

**Algoritmo 3:** EE-MIMO-NOMA – Alocação iterativa de potência com base no método Dinkelbach e taxa de dados igualitária.

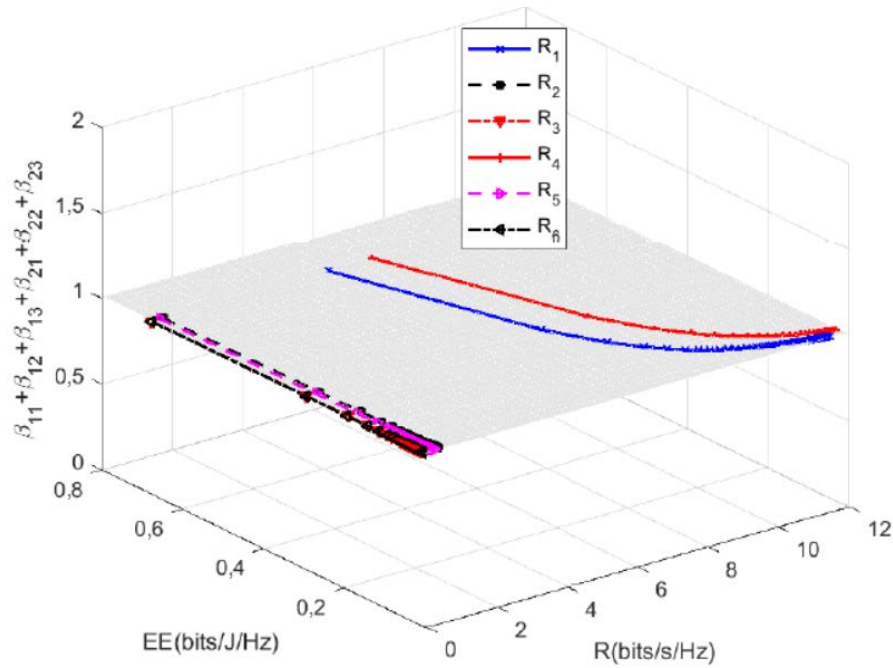
---

```

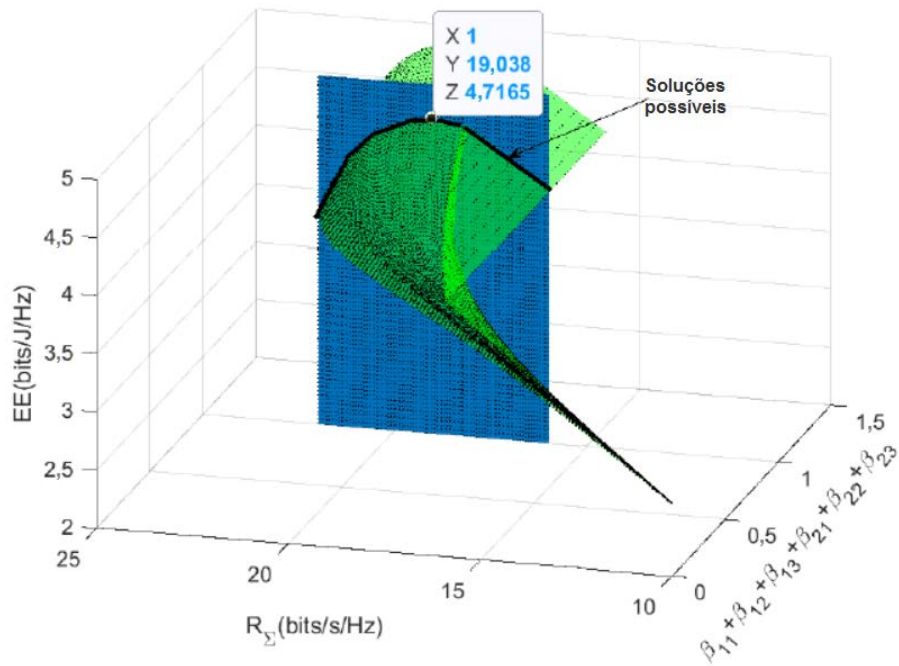
1 início
2    $n = 0$ ,  $\mathbf{q}^{(n)} = 0$ , e  $R = R^*$  (Meta de taxa de dados igualitária)
3   enquanto  $\varepsilon > 0$  faça
4     Para um dado  $\mathbf{q}^{(n)}$ , resolver (3.25) para obter  $P^{(n)}$  e  $\beta_{a,k}$ 
5     se  $U_R(P^{(n)}, \beta_{a,k}^{(n)}) - \mathbf{q}^{(n)} U_T(P^{(n)}, \beta_{a,k}^{(n)}) \leq \varepsilon$  então
6        $P^* = P^{(n)}$ ,  $\mathbf{q}^* = \mathbf{q}^{(n)}$ ,  $\beta_{a,k}^* = \beta_{a,k}^{(n)}$  para  $a = 1, \dots, A$  e  $k = 1, \dots, K$ 
7     senão
8       Definir  $n = n + 1$  e  $\mathbf{q}^{(n)} = \frac{U_R(P^{(n-1)}, \beta_{a,k}^{(n-1)})}{U_T(P^{(n-1)}, \beta_{a,k}^{(n-1)})}$ 
9     fim
10  fim
11 fim
```

---

O Algoritmo 2 obtém o valor da taxa de bits que apresenta o EE máximo diretamente. O Algoritmo 3 obtém o EE máximo para cada taxa de bits fixa. Assim, realizamos a varredura de toda uma gama de taxas de bits fixas candidatas e, em seguida, usamos algum método para refinamento do resultado próximo ao pico do EE, a fim de obter a taxa de bits solução que resulta na máxima EE no Algoritmo 3. Portanto, adotamos o Algoritmo 2 para encontrar a taxa de bits que resulta no máximo EE. Embora o algoritmo 2 não



(a)  $\beta$ 's  $\times$  EE  $\times$  Taxa de dados  $R$ , problema (3.18).



(b) EE  $\times$   $\beta$ 's  $\times$  Taxa, problema (3.23).

Figura 3.3 – Conjunto fechado de soluções  $\beta_{a,k}$ 's  $\times$  EE  $\times R^*$ , considerando os valores de parâmetros: potência  $P_{BS} = 1$  W/antena BS e  $P_K = 100$  mW/antena UE;  $\alpha = 2$ ;  $\sigma_n^2 = 1\mu W$ ,  $A = 2$  clusters e  $K = 3$  UE's/cluster, com distâncias  $d_{11} = 80$  m,  $d_{12} = 150$  m,  $d_{13} = 280$  m,  $d_{21} = 70$  m,  $d_{22} = 150$  m e  $d_{23} = 270$  m.

garanta convergência ao ótimo global porque o problema é não convexo, ele consegue obter desempenho tão bom quanto o obtido pelo Algoritmo 3. A convergência do Algoritmo 2 resulta expedita e factível para todas as configurações de sistema analisadas neste capítulo.

Os resultados numéricos mostraram-se realmente próximos ao ótimo, conforme discutido na Seção 5.2. Por conveniência, vamos chamá-lo de solução de ponto ótimo.

Os valores  $R_{a,k}$  foram obtidos para cada usuário na Figura 3.3.(a) a partir das respectivas proporções de potência dadas pela Eq. (3.18), e foram usados para calcular os valores de EE como na Eq. (3.14). Como esperado, as taxas de bits para cada usuário resultaram em valores diferentes. Por outro lado, o problema de otimização apresentado pela Eq. (3.23) calcula o valor ótimo da potência total  $P$  e a taxa de bits  $R$  que maximiza o EE, que gera Figura 3.3.(b). EE é maximizado pela função objetivo sem violar as quatro restrições do problema. Elaborando mais, a Figura 3.3.(b) revela o conjunto de soluções do problema de otimização apontado pela intersecção das duas superfícies no gráfico, conforme indicado pela curva de "soluções possíveis": a soma das proporções de potência igual a 1, ou seja,  $\beta_{11} + \beta_{12} + \beta_{13} + \beta_{21} + \beta_{22} + \beta_{23} = 1$ , no eixo x, e o outro plano definido por EE no eixo z *versus* usuários da soma da taxa no eixo y para cada valor da soma dos  $\beta$ 's. A intersecção entre os planos é a solução possível para o problema dado pela Eq. (3.23). Os valores para EE máximo e equidade máxima para este exemplo são aproximadamente EE= 4,72 bits/J/Hz e  $R = 3,17$  bits/s/Hz ou  $R_{\Sigma} = 19,04$  bits/s/Hz. A curva com soluções viáveis e exemplo de soluções são destacadas no gráfico da Figura 3.3.(b), identificado como:  $\sum_{a=1}^A \sum_{k=1}^K \beta_{a,k}$  no eixo x, a soma das taxas dos usuários no eixo y e EE no eixo z.

## 3.2 Abordagem de Otimização Bi-Objetivo EE-SE

Nesta seção, métodos de otimização BOO e diferentes estratégias para resolver o problema da RE em MIMO-NOMA, ou seja, o problema de compensação EE-SE sob equidade de taxa de usuários máxima e número de usuários variável por *cluster* é desenvolvido usando principalmente o Método de escalarização  $\epsilon$ -Constraint ( $\epsilon$ -C) combinado com técnicas de otimização de programação não linear (NLP - *Nonlinear Programming*), denominada programação quadrática sequencial (SQP - *Sequential Quadratic Programming*), bem como o emprego do método iterativo de Dinkelback (DK) para problemas de otimização quasi-côncavos fracionais.

### 3.2.1 Alocação de Recursos em Sistemas MIMO-NOMA

Problemas de alocação de recursos, principalmente considerando a otimização de EE e SE simultaneamente, resultam em metas de otimização BOO em vez de apenas uma. Esses objetivos, em geral, são conflitantes e raramente existe uma solução ótima para todos os objetivos simultaneamente. As características da RE, ou seja, o compromisso EE-SE sob a perspectiva da otimização MOO, foram exploradas recentemente em (MARTINEZ; ABRÃO, 2020), considerando o contexto de comunicação óptica, bem como em cenários de

comunicação sem fio (JACOB; ABRÃO, 2020). Portanto, a métrica RE pode ser expressa como uma soma ponderada de EE e SE dada por (2.15) no qual  $\xi$  controla a prioridade da EE em relação ao SE em um projeto BOO. Além disso, o fator  $\xi$  expresso em [J/s] atua como um fator fisicamente consistente na Eq. (2.15), enquanto equilibra a importância relativa entre as duas métricas de otimização.

### 3.2.1.1 Otimização Multiobjetivo: Definições

Sejam  $\mathbb{R}^n$  e  $\mathbb{R}^m$  espaços vetoriais euclidianos referidos ao espaço de decisão e ao espaço objetivo. Seja  $\mathcal{X} \subset \mathbb{R}^n$  um conjunto viável não vazio e seja  $f$  uma função objetivo com valor vetorial  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  composto por  $m$  funções objetivos com valor real,  $\mathbf{f} = (f_1, \dots, f_m)$ , onde  $f_k : \mathbb{R}^n \rightarrow \mathbb{R}$  para  $k = 1, \dots, m$ . Um problema de MOO pode ser formulado como em (EHRGOTT; RUZIKA, 2008), (MARTINEZ; ABRÃO, 2020):

$$\underset{\mathbf{x} \in \mathcal{X}}{\text{maximize}} \quad \{(f_1(\mathbf{x}), \dots, f_m(\mathbf{x})); \mathbf{x} \in \mathcal{X}\} \quad (3.27)$$

Assumimos que o conjunto  $\mathcal{X}$  seja dado implicitamente na forma de restrições, no qual:

$\mathbf{x} = \{x_1, x_2, \dots, x_n\} \in \mathbb{R}^n$ : vetor de variáveis de decisão;

$z_i = f_i(\mathbf{x})$ :  $i$ -ésimo função objetiva a ser maximizada,  $i = 1, \dots, m$ ;

$\emptyset \neq \mathcal{X} \subseteq \mathbb{R}^n$ : decisão ou espaço viável.

**Definição 1. *Dominação.*** Uma solução  $\mathbf{x}^{(1)}$  domina outra solução  $\mathbf{x}^{(2)}$ , no qual  $\mathbf{x}^{(1)}, \mathbf{x}^{(2)} \in \mathcal{X}$ , se as seguintes condições forem atendidas:

1. A solução  $\mathbf{x}^{(1)}$  não é pior do que a solução  $\mathbf{x}^{(2)}$  em nenhum dos objetivos, ou seja,  $f_i(\mathbf{x}^{(1)}) \geq f_i(\mathbf{x}^{(2)})$  para todos  $i = 1, \dots, m$ .
2. A solução  $\mathbf{x}^{(1)}$  é estritamente melhor do que a solução  $\mathbf{x}^{(2)}$  em pelo menos um componente  $i \in \{1, \dots, m\}$ , ou seja,  $f_i(\mathbf{x}^{(1)}) > f_i(\mathbf{x}^{(2)})$ .

A dominância da solução  $\mathbf{x}^{(1)}$  sobre a solução  $\mathbf{x}^{(2)}$  é denotada por:  $\mathbf{x}^{(1)} \succ \mathbf{x}^{(2)}$ .

**Definição 2. *Solução Eficiente.*** Uma solução  $\mathbf{x}^* \in \mathcal{X}$  é chamada eficiente ou ótima no sentido de Pareto, se não houver outra solução  $\mathbf{x} \in \mathcal{X}$ , de modo que  $\mathbf{x} \succ \mathbf{x}^*$ .

### 3.2.1.2 Formulação do EE-SE Biobjetivo em Sistemas MIMO-NOMA

Neste trabalho, estamos interessados no problema de eficiência de recursos em termos de otimização do compromisso entre EE-SE em redes MIMO-NOMA. A compensação EE-SE pode ser formulada com o objetivo de maximizar EE e SE simultaneamente.

Pode-se formulá-la como um problema de BOO da seguinte maneira:

$$\begin{aligned}
 & \underset{P, \beta_{a,k}, R}{\text{maximize}} \quad SE \quad \left[ \frac{\text{bit}}{\text{s}\cdot\text{Hz}} \right] \quad \text{e} \quad \underset{(P, \beta_{a,k}, R)}{\text{maximize}} \quad EE \quad \left[ \frac{\text{bit}}{\text{J}\cdot\text{Hz}} \right] \\
 \text{s.a.:} \quad & \text{C.1} \quad \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} = 1 \\
 & \text{C.2} \quad \beta_{a,k} = (2^R - 1) \frac{\sigma^2 |\mathbf{v}_{a,k}|^2}{P |\mathbf{v}_{a,k}^H \mathbf{h}_{a,ak}|^2} + (2^R - 1) \sum_{i=1}^{k-1} \beta_{a,i}, \quad \text{sendo} \quad \begin{cases} 2 \leq k \leq K \\ 1 \leq a \leq A \end{cases} \quad (3.28) \\
 & \text{C.3} \quad \beta_{a,1} = \frac{(2^R - 1) \sigma^2 |\mathbf{v}_{a,1}|^2}{P |\mathbf{v}_{a,1}^H \mathbf{h}_{a,ak}|^2} \quad \text{para} \quad 1 \leq a \leq A. \\
 & \text{C.4} \quad \beta_{a,k} > 0; P > 0; R > 0
 \end{aligned}$$

A abordagem tradicional para resolver o MOO é por métodos de escalarização, que envolvem a formulação de um problema de objetivo único relacionado ao MOO, usando uma função escalar de valor real. Às vezes, o conjunto viável do MOO é ainda mais restrito por novas funções de restrição relacionadas às funções objetivo do MOO, além das novas variáveis introduzidas. Ressalte-se ainda que em abordagens de escalarização, é fundamental considerar as condições necessárias e suficientes para se atingir as soluções ótimas em termos de soluções MOO fracamente, estritamente ou adequadamente eficientes.

### 3.2.2 Resolvendo o EE-SE em MIMO-NOMA Sob Equidade de Taxa

O método de otimização MOO  $\epsilon$ -C será empregado para resolver o problema do compromisso EE-SE em sistemas MIMO-NOMA. Este método tem como ideia central escolher um objetivo único a ser otimizado, enquanto os outros objetivos conflitantes são tratados como restrições de desigualdade, permitindo a implantação a partir de abordagens clássicas de otimização com objetivo único. Desta forma, o método  $\epsilon$ -C não aborda diretamente um peso apropriado entre as métricas EE e SE, como ocorre no método de escalonamento WS descrito na Seção 3.2.4, para determinar soluções na fronteira de Pareto (MARTINEZ; ABRÃO, 2020). Assim, aplicando a técnica  $\epsilon$ -C, o problema RE MIMO-NOMA original apresentado em (3.28) pode ser reformulado da seguinte forma:

$$\begin{aligned}
 & \underset{P, \beta_{a,k}, R}{\text{maximize}} \quad \frac{\sum_{a=1}^A \sum_{k=1}^K R_{a,k}}{\varrho \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} P + K \cdot U \cdot A \cdot P_K + A \cdot P_{BS}} \\
 \text{s.t.} \quad & \text{C.1} \quad \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} = 1 \\
 & \text{C.2} \quad \beta_{a,k} = (2^R - 1) \frac{\sigma^2 |\mathbf{v}_{a,k}|^2}{P |\mathbf{v}_{a,k}^H \mathbf{h}_{a,ak}|^2} + (2^R - 1) \sum_{i=1}^{k-1} \beta_{a,i} \quad \text{no qual} \quad \begin{cases} 2 \leq k \leq K \\ 1 \leq a \leq A \end{cases} \quad (3.29) \\
 & \text{C.3} \quad \beta_{a,1} = \frac{(2^R - 1) \sigma^2 |\mathbf{v}_{a,1}|^2}{P |\mathbf{v}_{a,1}^H \mathbf{h}_{a,ak}|^2} \quad \text{para} \quad 1 \leq a \leq A \\
 & \text{C.4} \quad \beta_{a,k} > 0; P > 0; R > 0 \\
 & \text{C.}\epsilon \quad \sum_{a=1}^A \sum_{k=1}^K R_{a,k} \geq \epsilon
 \end{aligned}$$

no qual  $\epsilon \in [\text{SE}(\overline{P, \beta_{a,k}, R}), \max \text{SE}]$ . Os valores de  $\epsilon$  variam do limite inferior da restrição (C. $\epsilon$ ) aos valores que SE pode assumir na frente de Pareto. O valor mínimo  $\text{SE}(\overline{P, \beta_{a,k}, R})$  é determinado pela obtenção de  $(\overline{P, \beta_{a,k}, R})$  solução do problema de EE. Esta escolha para o valor mínimo é feita porque qualquer outro vetor  $(P, \beta_{a,k}, R)$  cujo valor de SE é menor que esse também terá um valor menor que EE; portanto, este vetor será dominado pelo vetor solução  $(\overline{P, \beta_{a,k}, R})$  e por consequência não se caracteriza como uma solução de Pareto. Assim, consideramos  $\epsilon$  variando no intervalo  $\epsilon \in [\text{SE}(\overline{P, \beta_{a,k}, R}), \max \text{SE}]$  para indicar o vetor onde ocorre o máximo EE.

O Teorema 1 apresenta a condição necessária e suficiente para soluções globais de (3.29) estarem na fronteira de Pareto do problema de compromisso entre EE-SE.

**Teorema 1.** *Uma solução global para o problema definido em (3.29) é uma solução na fronteira de Pareto para o problema de compensação EE-SE se, e somente se, a seguinte igualdade for mantida no vetor de solução.*

$$\sum_{a=1}^A \sum_{k=1}^K R_{a,k} = \epsilon \quad (3.30)$$

*Prova:*

Considere  $\mathbf{p} = (P, \beta_{a,k}, R)$  e  $\mathbf{p}_\epsilon^* \in \varphi$ , no qual

$$\varphi \subseteq \mathbb{R}^n = \{[p_i]_{1 \times K} | \mathbf{p}_{\min} \leq \mathbf{p}_i \leq \mathbf{p}_{\max}; \beta_{a,k} > 0; P > 0; R_{a,k} > 0\}. \quad (3.31)$$

( $\Rightarrow$ ) Considere  $\mathbf{p}_\epsilon^*$  um ponto na fronteira de Pareto, devemos mostrar que este valor é a solução do problema (3.29) para qualquer escolha de  $\epsilon$ . Suponha que  $\mathbf{p}_\epsilon^*$  não seja uma solução global para o problema  $\epsilon = \text{SE}_m(\mathbf{p}_\epsilon^*)$ . Então, deve haver outra solução global para este problema, então

$$\text{EE}(\overline{\mathbf{p}}) > \text{EE}(\mathbf{p}_\epsilon^*) \quad \text{e} \quad \text{SE}(\overline{\mathbf{p}}) > \text{SE}(\mathbf{p}_\epsilon^*) \quad (3.32)$$

o que, por sua vez, contradiz o fato de que  $\mathbf{p}_\epsilon^*$  está na fronteira de Pareto. Então  $\mathbf{p}_\epsilon^*$  é uma solução global do problema (3.29) com  $\epsilon = \text{SE}_m(\mathbf{p}_\epsilon^*)$ , consequentemente  $\mathbf{p}_\epsilon^*$  satisfaz a igualdade.

( $\Leftarrow$ ) Se  $\mathbf{p}^*$  é a solução global do problema (3.29), em que vale

$$\epsilon = \text{SE}_m(\mathbf{p}_\epsilon^*), \quad (3.33)$$

então não há outra solução para que  $\text{EE}(\overline{\mathbf{p}}) \geq \text{EE}(\mathbf{p}_\epsilon^*)$  e  $\text{SE}(\overline{\mathbf{p}}) \geq \text{SE}(\mathbf{p}_\epsilon^*)$ , então  $\mathbf{p}_\epsilon^*$  está na fronteira de Pareto.  $\square$

Eq. (3.30) no Teorema 1 indica que a restrição de desigualdade (C. $\epsilon$ ) no problema (3.29) é satisfeita como uma igualdade na solução pertencente à frente de Pareto.

O método  $\epsilon$ -C é resumido no Algoritmo 4, no qual o vetor  $\mathbf{p}^{(k)}$  denota a solução vetorial do problema (3.29) na  $k$ -ésima iteração. Como a dificuldade em aplicar o método  $\epsilon$ -C é que qualquer solução de Pareto encontrada deve satisfazer a restrição de desigualdade (C. $\epsilon$ ) como igualdade. Uma desvantagem do método WS consiste em pesos diferentes que podem implicar em soluções iguais conforme relatado em (MARTINEZ; ABRÃO, 2020). Ao contrário do método WS, o método  $\epsilon$ -C não considera diretamente RE em seu objetivo, mas ao definir os valores apropriados para  $\epsilon$ , impõe-se o peso exato que SE terá sobre a métrica RE; assim, resolvendo o problema em (3.28) também se pode obter soluções para RE. Além disso, uma motivação para usar o método  $\epsilon$ -C é que diferentes valores para  $\epsilon$  implicam em soluções diferentes. Assim, no método  $\epsilon$ -C operando sob uma distribuição uniforme de valores  $\epsilon$ , pode-se esperar uma distribuição uniforme de soluções na fronteira de Pareto.

---

**Algoritmo 4:**  $\epsilon$ -C-NOMA-MIMO –  $\epsilon$ -Restrição para o compromisso entre EE-SE em MIMO-NOMA

---

```

1 Escolher:  $\delta \in (0, 1)$ .
2 Calcular:  $\max SE$  e  $\mathbf{p}^{(0)} = (\overline{P}, \beta_{a,k}, \overline{R})$ ;
3 Definir:  $k = 1$  e  $\epsilon = SE(\overline{P}, \beta_{a,k}, \overline{R})$ ;
4 enquanto  $\epsilon \leq \max SE$  faça
5    $\mathbf{p}^{(k)} \leftarrow$  Resolver problema (3.29) pelo método HC, DK ou SQP;
6    $\epsilon \leftarrow \epsilon + \delta (\max SE - SE(\overline{P}, \beta_{a,k}, \overline{R}))$ ;
7   Salvar  $\mathbf{p}^{(k)}$ ;
8    $k = k + 1$ ;
9 fim
10 Saída:  $\mathbf{p}^{(0)}, \dots, \mathbf{p}^{(k)}$ ;

```

---

### 3.2.3 Métodos de Otimização Implementados

Apresentamos os métodos de otimização não lineares implementados para resolver o procedimento da linha 5 do Algoritmo 4. Além disso, discutimos como os métodos DK, HC e SQP foram implementados para resolver esse problema de otimização intrincado.

#### 3.2.3.1 Método Escalarização com DK

Aplicamos o método DK conforme discutido na Seção 3.1.1.4 e detalhado no Algoritmo 3, para resolver a alocação de potência específica no problema EE dentro de  $\epsilon$ -C-NOMA do Algoritmo 4. Observe que a restrição C. $\epsilon$  juntamente com o problema (3.14) como aparece no problema (3.29) é uma restrição convexa. Este fato permite o uso do Algoritmo 3 na resolução do Passo 5 do Algoritmo 4.

### 3.2.3.2 Método Escalarização com SQP

O método SQP determina uma aproximação para a solução de (3.29) através do limite das soluções de uma sequência de problemas quadráticos pré-definidos. No contexto do compromisso entre EE-SE em MIMO-NOMA, para simplificar a notação consideramos nesta subseção  $\mathbf{x} = (P, \beta_{a,k}, R)$  (e não mais por  $\mathbf{p}$ ),  $f$  representa a função EE,  $\mathbf{h}(\mathbf{x}) = 0$  representa as restrições C.1, C.2 e C.3, e  $g(\mathbf{x}) \geq 0$  representa a restrição C.ε. Em resumo, o método SQP em cada iteração substitui  $EE$  por uma aproximação quadrática da função Lagrangiana  $\mathcal{L}$ , definida por:

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \mu) = f(\mathbf{x}) + \sum_i \lambda_i h_i(\mathbf{x}) - \mu g(\mathbf{x}). \quad (3.34)$$

Portanto, as restrições não lineares são substituídas por aproximações lineares. Assim, na  $k$ -ésima iteração, o método SQP (NOCEDAL; WRIGHT, 1996) resolve o seguinte problema de programação quadrática:

$$\begin{aligned} \underset{\mathbf{d} \in \mathbb{R}^K}{\text{maximize}} \quad & f(\mathbf{x}^{(k)}) + \nabla f(\mathbf{x}^{(k)})^T \mathbf{d} + \frac{1}{2} \mathbf{d}^T \nabla^2 \mathcal{L}(\mathbf{x}^{(k)}, \boldsymbol{\lambda}^{(k)}, \mu^{(k)}) \mathbf{d} \\ \text{Sujeito a:} \quad & -\nabla g(\mathbf{x}^{(k)})^T \mathbf{d} - g(\mathbf{x}^{(k)}) \geq 0, \\ & \nabla h_i(\mathbf{x}^{(k)})^T \mathbf{d} + h_i(\mathbf{x}^{(k)}) = 0, \quad \forall i. \end{aligned} \quad (3.35)$$

Atualizando para a próxima solução,  $\mathbf{x}^{(k+1)}$  é obtido projetando  $\mathbf{x}^{(k)} + \bar{\alpha}_k \mathbf{d}^{(k)}$  no conjunto  $\Omega$  definido pela restrição C.2,  $\bar{\alpha}_k \in (0, 1]$  é o tamanho do passo determinado para garantir as propriedades de convergência do método,  $\mathbf{d}$  é a direção da pesquisa. Uma descrição mais detalhada do método SQP pode ser encontrada em (NOCEDAL; WRIGHT, 1996). Um pseudocódigo para o método baseado em SQP resolvendo o compromisso entre EE-SE em MIMO-NOMA é descrito em Algoritmo 5.

---

**Algoritmo 5:** SQP-NOMA – Programação quadrática sequencial para o compromisso entre EE-SE no MIMO-NOMA

---

- 1 **Escolher:** ponto inicial  $(\mathbf{x}^{(0)}, \boldsymbol{\mu}^{(0)}, \boldsymbol{\lambda}^{(0)})$  e tolerância  $\zeta > 0$ ;
  - 2 **Definir:**  $k = 0$ ;
  - 3 **enquanto**  $\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| > \zeta$  **faça**
  - 4     **Calcular:**  $f(\mathbf{x}^{(k)})$ ,  $\nabla f(\mathbf{x}^{(k)})$ ,  $\nabla^2 \mathcal{L}(\mathbf{x}^{(k)}, \mathbf{h}(\mathbf{x}^{(k)}), \nabla \mathbf{h}(\mathbf{x}^{(k)}), g(\mathbf{x}^{(k)})$  e  $\nabla g(\mathbf{x}^{(k)})$ ;
  - 5     **Resolver:** (3.35) para obter  $\mathbf{d}^{(k)}$  e  $\boldsymbol{\mu}^{(k+1)}$ ;
  - 6      $\mathbf{x}^{(k+1)} \leftarrow \mathcal{P}(\mathbf{x}^{(k)} + \bar{\alpha}_k \mathbf{d}^{(k)})$ ; no qual  $\mathcal{P}$  é o operador de projeção ortogonal no conjunto  $\Omega$ ;
  - 7 **fim**
  - 8 **Saída:**  $\mathbf{x}^* = \mathbf{x}^{(k)}$ .
- 

### 3.2.3.3 Método Escalarização Heurístico Ganancioso

O algoritmo HC é um método ganancioso de otimização de busca local que começa com uma solução chamada nó atual e procura em vizinhos próximos, onde a distância é

---

**Algoritmo 6:** HC-NOMA – HC para o compromisso entre EE-SE no MIMO-NOMA
 

---

```

1 Escolher:  $\mathbf{x} = (P^0, \beta_{a,k}^0, R^0)$ ;
2 Definir: Solução = 0; Iteração = 0; Max-Iteração > 0; Penalidade =  $10^8$ ;
    $0 < \bar{\alpha}_k \leq 1$  (tamanho do passo);
3 enquanto Solução = 0 e Iteração  $\leq$  Max-Iteração faça
4   | Iteração = Iteração + 1;
5   | Gerar vizinhos;
6   | se Qualquer vizinho tem valor melhor de  $F$  do que de  $x$  então
7   |   |  $\mathbf{x}$  é atualizado como o vizinho com melhor custo para  $F$ ;
8   | senão
9   |   | Solução = 1;
10  | fim
11 fim
Saída:  $\mathbf{x}$ .

```

---

controlada pelo tamanho do passo, a fim de encontrar um nó melhor; se não houver um nó melhor associado, será possível encontrar uma solução para o problema. Portanto, se o algoritmo detecta um nó vizinho melhor do que o nó atual, ele transforma o vizinho como a nova aproximação para a solução atual e, em seguida, itera até atingir uma localização ainda melhor, com base nas distâncias, seguindo um critério de terminação, ou ter alcançado o número máximo de iterações pré-estabelecidas (STOREY, 1962), (RUSSELL; NORVIG, 2012), (JOHNSON; JACOBSON, 2002). Um pseudo-código para o algoritmo HC adaptado à compensação EE-SE em sistemas de comunicação MIMO-NOMA é descrito no Algoritmo 6. A função custo  $F(\mathbf{x})$  com restrições incorporadas, implementada no Algoritmo 6, é definida por:

$$F(\mathbf{x}) = \text{EE}(\mathbf{x}) + \rho \left\{ \sum [h_i(\mathbf{x})]^2 + \sum [\max(0, g_i(\mathbf{x}))]^2 \right\}. \quad (3.36)$$

no qual  $\rho$  é o parâmetro de penalidade associado à violação das restrições do problema, enquanto as funções  $h$  e  $g$  no problema EE-SE NOMA-MIMO foram discutidas na Seção 3.2.3.2. A função custo (3.36) admite parâmetro de penalidade constante, requerendo apenas que seja um valor elevado mas limitado para inibir a violação das restrições que constituem o problema (CAO; CHEN, 2004).

### 3.2.4 Comparação com Método de Escalarização de Soma Ponderada

Nesta subseção, comparamos nosso método de escalonamento  $\epsilon$ -C com o método de otimização WS proposto em (MARTINEZ; ABRÃO, 2020). De fato, na subseção anterior 5.3.2 o método  $\epsilon$ -C foi extensivamente avaliado combinando-se com os métodos SQP, DK e HC. Aqui, avaliamos o método de escalonamento WS combinado com o método SQP nos mesmos cenários de sistema e canal que aqueles realizados com  $\epsilon$ -C na Subseção 5.3.2.

O método WS redimensiona as funções objetivo EE e SE do problema BOO original em (3.28) como um problema de objetivo único, particularmente como uma combinação convexa de tais objetivos. Portanto, dado  $\tau \in [0, 1]$ , pode-se considerar a combinação convexa:

$$(1 - \tau) \cdot \text{EE} + \frac{\tau}{\bar{P}_{\text{RF}} + P_C} \text{SE},$$

no qual  $\bar{P}_{\text{RF}}$  é calculado no vetor de solução  $(\bar{P}, \bar{\beta}_{a,k}, \bar{R})$  do problema EE, conforme discutido na Seção 3.2.2. Assim, dividimos SE por  $\bar{P}_{\text{RF}} + P_C$  para que as combinações possíveis de EE e SE a serem maximizadas tenham a mesma ordem de magnitude. O problema do compromisso entre EE-SE bi-objetivo original pode ser descrito pelo método WS como o seguinte problema SOO:

$$\begin{aligned} & \underset{P, \beta_{a,k}, R}{\text{maximize}} && (1 - \tau) \frac{\sum_{a=1}^A \sum_{k=1}^K R_{a,k}}{\varrho \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} P + K \cdot A \cdot P_K + A \cdot P_{\text{BS}}} + \tau \frac{\sum_{a=1}^A \sum_{k=1}^K R_{a,k}}{\bar{P}_{\text{RF}} + P_C} \\ & \text{s.t.} && \text{C.1} \quad \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} = 1 \\ & && \text{C.2} \quad \beta_{a,k} = (2^R - 1) \frac{\sigma^2 |\mathbf{v}_{a,k}|^2}{P |\mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_a|^2} + (2^R - 1) \sum_{i=1}^{k-1} \beta_{a,i}, \text{ no qual } \begin{cases} 2 \leq k \leq K \\ 1 \leq a \leq A \end{cases} \quad (3.37) \\ & && \text{C.3} \quad \beta_{a,1} = \frac{(2^R - 1) \sigma^2 |\mathbf{v}_{a,1}|^2}{P |\mathbf{v}_{a,1}^H \mathbf{H}_{a,1} \mathbf{w}_a|^2} \text{ para } 1 \leq a \leq A. \\ & && \text{C.4} \quad \beta_{a,k} > 0; P > 0; R > 0 \end{aligned}$$

O método de escalarização WS para resolver o problema de otimização do compromisso entre EE-SE pode ser resumido no Algoritmo 7. Soluções globais de (3.37), assumindo  $\tau \in [0, 1]$ , pertencem ao limite de Pareto do problema BOO original em (3.28). Detalhes sobre o método WS aplicado ao compromisso entre EE-SE em redes ópticas podem ser encontrados em (MARTINEZ; ABRÃO, 2020).

---

**Algoritmo 7: WS-MIMO-NOMA** – Soma ponderada para o compromisso entre EE-SE em MIMO-NOMA.

---

- 1 **Escolher:**  $\delta \in (0, 1)$ .
  - 2 **Calcular:**  $\mathbf{p}^0 = (\bar{P}, \bar{\beta}_{a,k}, \bar{R})$ ;
  - 3 **Definir:**  $k = 1$  e  $\tau = \delta$ ;
  - 4 **enquanto**  $\tau \leq 1$  **faça**
  - 5      $\mathbf{p}^k \leftarrow$  Resolver problema (3.37) pelo método SQP;
  - 6      $\tau \leftarrow \tau + \delta$ ;
  - 7     Salvar  $\mathbf{p}^k$ ;
  - 8      $k = k + 1$ ;
  - 9 **fim**
  - 10 **Saída:**  $\mathbf{p}^0, \dots, \mathbf{p}^k$ ;
- 

Um pseudo-código para o compromisso EE-SE MIMO-NOMA resolvido pelo método WS é apresentado no Algoritmo 7. O vetor  $\mathbf{p}^k$  denota a solução vetorial para o problema (3.37) na  $k$ -ésima iteração.

# 4 Sistema mMIMO PDMA para Serviços eMBB, mMTC e uRLLC

Neste capítulo, discute-se a viabilidade de aplicação de um esquema de transmissão inovador denominado PDMA multiusuário, operando simultaneamente sob os diferentes modos de uso eMBB, mMTC e uRLLC em sistemas 5G com grande número de antenas mMIMO. Tal esquema de transmissão é baseado no projeto da matriz padrão, cujas características e especificidades são descritas na subseção 4.1.1. Neste trabalho, adaptamos as matrizes padrão disponíveis na literatura, quando necessário, para operarem no enlace de descida do sistema mMIMO PDMA, tendo em vista avaliar seus parâmetros e identificar aquela matriz padrão que melhor atende as especificações de cada um dos serviços 5G (eMBB, mMTC e uRLLC), ou ainda simultaneamente a combinação de dois modos de uso.

## 4.1 Modelo do sistema mMIMO PDMA

Um grande número de antenas na estação rádio-base BS pode ser espalhado ao longo da área geográfica da célula, nomeadamente MIMO distribuído em larga escala<sup>1</sup>, ou implementado de forma centralizada na BS, conhecido como matrizes de grande escala LSA (WANG et al., 2016). Como mostrado na Figura 4.1, em nosso estudo, uma BS equipada com mMIMO transmitindo para múltiplos usuários no DL usando os princípios de PDMA é considerada como em (LI et al., 2017), (JIANG et al., 2019). Cada usuário está equipado com  $N_R$  antenas. Supõe-se que existam diversas antenas na BS atendendo cada *cluster* de usuários, denominado *antenas de cluster* (AC - *antenna clusters*). Localizados na BS, cada AC é equipado com  $N_T$  antenas que formam  $N$  feixes, no qual

$$N_T \geq N. \quad (4.1)$$

Todos os usuários são divididos em vários *grupo de usuários* (UGs - *User Groups*) e cada UG contém  $K$  usuários, sendo  $N_T \leq KN_R$ . Um AC deve cobrir um UG com

$$N \leq K \leq 2^N - 1. \quad (4.2)$$

Sem perda de generalidade, simplificamos o cenário em que um AC se comunica com um UG. Na Figura 4.1 temos AC representado pelo feixe 1 que forma o UG com os usuários 1 e 2 enquanto o feixe 2 forma o UG com os usuários 1 e 3, então cada AC se comunica com um UG. Ainda de acordo com a Figura 4.1, as primeiras antenas de 1 a  $N_T$  à direita suportam o feixe 1 enquanto as antenas de 1 a  $N_T$  à esquerda suportam o feixe 2.

---

<sup>1</sup>Também conhecido como sistemas livre de células.

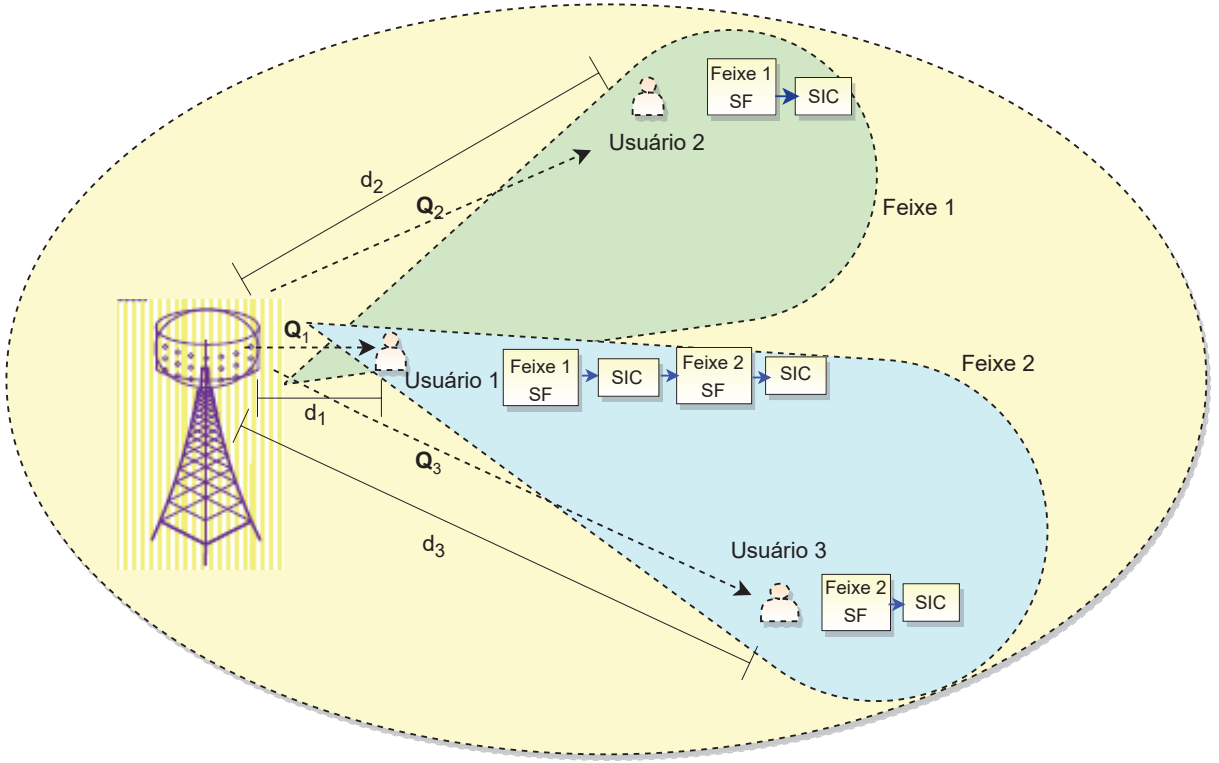


Figura 4.1 – Esquema de transmissão mMIMO PDMA. Cada AC é equipado com  $N_T$  antenas formando  $N$  feixes, com  $N_T \geq N$ ; ademais, cada UG contém  $K$  usuários, sendo  $N_T \leq KN_R$  e o filtro espacial (SF - *spatial filter*)  $\mathbf{v}_{nk}$  para cada UG definido na seção 4.1.4.

Seja  $\mathbf{H}_k \in \mathbb{C}^{N_R \times N_T}$  a matriz do canal entre o AC e o  $k$ -ésimo usuário do UG onde

$$\mathbf{H}_k = \frac{\mathbf{Q}_k}{\sqrt{1 + d_k^\alpha}} \quad (4.3)$$

e  $\mathbf{Q}_k \in \mathbb{C}^{N_R \times N_T}$  representa uma matriz de canal com desvanecimento Rayleigh;  $\alpha$  é o expoente de perda de percurso e  $d_k$  representa a distância do  $k$ -ésimo usuário até a BS.

Por simplicidade de análise, assumiu-se que a BS possui conhecimento perfeito do *status* do canal (hipótese CSI perfeitas). Ademais, por conveniência, usaremos formação de feixe (*precoding*) por forçamento a zero (ZFBBF - *Zero-Forcing Beamforming*) no transmissor (JIANG et al., 2019). Seja  $\mathbf{b}_n \in \mathbb{C}^{N_T \times 1}$  a representação do vetor ZFBBF do  $n$ -ésimo feixe, sendo gerado com base no CSI do usuário selecionado entre os usuários cobertos pelo  $n$ -ésimo feixe representado pelo índice  $\text{sel}_n$ . O conjunto de pares de feixes e o usuário selecionado correspondente pode ser expresso como

$$\Omega = \{(1, \text{sel}_1), (2, \text{sel}_2), \dots, (n, \text{sel}_n), \dots, (N, \text{sel}_N)\} \quad (4.4)$$

Assim, a matriz de canal composta pelos usuários selecionados é definida como

$$\mathbf{H}_C = [\mathbf{H}_{\text{sel}_1}^H, \mathbf{H}_{\text{sel}_2}^H, \dots, \mathbf{H}_{\text{sel}_N}^H]^H, \quad \text{com dimensões } \mathbf{H}_C \in \mathbb{C}^{N_R \cdot N \times N_T} \quad (4.5)$$

e a matriz composta ZFBF pode ser expressa como

$$\mathbf{B}_C = \mathbf{H}_C^H (\mathbf{H}_C \mathbf{H}_C^H)^{-1}. \quad (4.6)$$

Nós particionamos a matriz ZFBF composta na forma  $\mathbf{B}_C = [\mathbf{B}_1, \mathbf{B}_2, \dots, \mathbf{B}_n, \dots, \mathbf{B}_N]$  e então o vetor ZFBF do  $n$ -ésimo feixe pode ser expresso da seguinte forma

$$\mathbf{b}_n = \mathbf{B}_n \mathbf{1}^{N_R \times 1}, \quad (4.7)$$

no qual  $\mathbf{1}^{N_R \times 1}$  denota um vetor coluna de dimensão  $N_R$  com todos os elementos iguais a 1. Portanto, vamos representar a matriz ZFBF por  $\mathbf{B} \in \mathbb{C}^{N_T \times N}$  que pode ser expressa como

$$\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n, \dots, \mathbf{b}_N]. \quad (4.8)$$

O sistema PDMA na Figura 4.1 usa mapeamento de padrões no transmissor com alocação de energia e domínios de alocação de feixe para sobrepor os sinais. Considere que  $\mathbf{s} \in \mathbb{C}^{K \times 1}$  representa o vetor símbolo de transmissão para o UG com  $\mathbf{s} \sim \mathcal{CN}(0, \mathbf{I}_K)$ ,  $[\mathbf{s}]_k = s_k$  o símbolo de transmissão para o  $k$ -ésimo usuário. Seja  $\mathbf{P} \in \mathbb{R}^{N \times K}$  a matriz de alocação de potência, cujo  $n, k$ -ésimo elemento de  $\sqrt{\mathbf{P}}$  é definido por:

$$[\mathbf{P}^{1/2}]_{nk} = \sqrt{p_{nk}} \quad (4.9)$$

sendo  $p_{nk}$  a potência de transmissão alocada ao  $k$ -ésimo usuário no  $n$ -ésimo feixe. Consideremos que a matriz binária  $\mathbf{G} \in [0, 1]^{N \times K}$  representa a *matriz padrão de alocação de feixe*:

$$[\mathbf{G}]_{nk} = g_{nk} \quad (4.10)$$

no qual o fator de alocação de feixe para o  $k$ -ésimo usuário no  $n$ -ésimo feixe assume os valores  $g_{nk} \in \{0, 1\}$ . Quando  $g_{nk} = 1$  significa que o  $k$ -ésimo usuário está coberto pelo  $n$ -ésimo feixe e para  $g_{nk} = 0$  este usuário não está coberto por este feixe. O resultado do vetor de sinal sobreposto após o mapeamento de padrão pode ser expresso por

$$\mathbf{t} = (\mathbf{G} \circ \mathbf{P}^{1/2}) \mathbf{s}, \quad (4.11)$$

no qual  $\circ$  representa a operação do produto de Hadamard<sup>1</sup>. O  $n$ -ésimo elemento do vetor de sinal sobreposto após o mapeamento de padrões é definido como:

$$[\mathbf{t}]_n = t_n = \sum_{k=1}^K g_{nk} \sqrt{p_{nk}} s_k, \quad n = 1, 2, \dots, N. \quad (4.12)$$

No transmissor, temos  $\mathbf{t} \in \mathbb{C}^{N \times 1}$  representando o vetor do sinal sobreposto após o processo de mapeamento de padrões. Seja  $\mathbf{x} \in \mathbb{C}^{N_T \times 1}$  o vetor do sinal de transmissão AC que pode ser expresso da seguinte forma

$$\mathbf{x} = \mathbf{B}\mathbf{t}. \quad (4.13)$$

<sup>1</sup>Para duas matrizes de mesma dimensão  $I \times J$ , o produto de Hadamard é uma matriz de mesma dimensão dos operandos, com o  $(i, j)$ -ésimo elemento dado por  $[\mathbf{G} \circ \mathbf{P}]_{i,j} = [\mathbf{G}]_{i,j} [\mathbf{P}]_{i,j}$ .

No receptor, temos  $\mathbf{y}_k \in \mathbb{C}^{N_R \times 1}$  que representa o vetor do sinal recebido do  $k$ -ésimo usuário que pode ser expresso da seguinte forma

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x} + \mathbf{w}_k \quad (4.14)$$

no qual  $\mathbf{w}_k \sim \mathcal{CN}(0, \sigma_k^2 \mathbf{I}_{N_R})$  é o vetor do ruído Gaussiano aditivo branco do  $k$ -ésimo usuário cujos elementos têm média zero e variância  $\sigma_k^2$ .

## 4.1.1 Projeto de Matrizes Padrão para Modos de Uso 5G

### 4.1.1.1 Descrição e Características da Matriz Padrão

A matriz padrão PDMA, dimensão  $N \times K$ , é composta pelos elementos 0s e 1s, sendo representada por  $\mathbf{G}_{\text{PDMA}}^{N \times K}$ . Os elementos 0s indicam que o usuário não pode usar os valores de recursos (RV - *Resource Values*) correspondentes, que no nosso modelo RV são os feixes, e os elementos 1s indicam que o usuário pode usar o RV correspondente. A coluna e a linha de  $\mathbf{G}_{\text{PDMA}}$  representam usuários e RV, respectivamente. A dimensão e o nível de esparsidade são propriedades da matriz padrão PDMA que contribuem para a complexidade e desempenho do sistema (CHEN et al., 2017). Portanto, a matriz padrão é projetada com os seguintes princípios (VAEZI et al., 2019), (KONG et al., 2016):

- Um padrão com mais elementos 1s no padrão fornece uma ordem mais alta de diversidade.
- O número de usuários com elementos 1s em cada linha deve ser igual.
- Os padrões selecionados devem ter tantas ordens de diversidade diferentes quanto possível.

Suponha que os  $K$  usuários sejam mapeados em elementos de recursos físicos ortogonais  $N$  com base na matriz padrão PDMA. Com base nos requisitos especificados acima, o número máximo de usuários suportados com  $N$  elementos de recurso pode ser calculado por

$$\binom{N}{1} + \binom{N}{2} + \cdots + \binom{N}{N} = 2^N - 1 \quad (4.15)$$

no qual  $\binom{N}{K}$  representa o conjunto de todas as combinações dos  $K$  usuários de um conjunto de  $N$  RVs.

Como exemplo de valores para a Eq. (4.15) pode-se desenhar as estruturas PDMA para  $N = 2, 3$  e 4 da seguinte forma:

$$\mathbf{G}_{\text{PDMA}}^{2 \times 3} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ \underbrace{\begin{pmatrix} 2 \\ 2 \end{pmatrix}}_{=1} & \underbrace{\begin{pmatrix} 2 \\ 1 \end{pmatrix}}_{=2} \end{bmatrix} \quad (4.16)$$

$$\mathbf{G}_{\text{PDMA}}^{3 \times 7} = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ \underbrace{\begin{pmatrix} 3 \\ 3 \end{pmatrix}}_{=1} & \underbrace{\begin{pmatrix} 3 \\ 2 \end{pmatrix}}_{=3} & \underbrace{\begin{pmatrix} 3 \\ 1 \end{pmatrix}}_{=3} \end{bmatrix} \quad (4.17)$$

$$\mathbf{G}_{\text{PDMA}}^{4 \times 15} = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ \underbrace{\begin{pmatrix} 4 \\ 4 \end{pmatrix}}_{=1} & \underbrace{\begin{pmatrix} 4 \\ 3 \end{pmatrix}}_{=4} & \underbrace{\begin{pmatrix} 4 \\ 2 \end{pmatrix}}_{=6} & \underbrace{\begin{pmatrix} 4 \\ 1 \end{pmatrix}}_{=4} \end{bmatrix} \quad (4.18)$$

Especificamente, considere uma matriz padrão  $3 \times 5$  usado como exemplo, que é mostrado a seguir

$$\mathbf{G}_{\text{PDMA}}^{3 \times 5} = \begin{matrix} & \text{Us.1} & \text{Us.2} & \text{Us.3} & \text{Us.4} & \text{Us.5} \\ \text{RV1} & \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \end{bmatrix} \\ \text{RV2} & \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \end{bmatrix} \\ \text{RV3} & \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \end{bmatrix} \end{matrix} \quad (4.19)$$

no qual o elemento 1 indica que os dados estão mapeados para o RV especificado, o elemento 0 representa o oposto. A matriz da Eq. (4.19) mostra o processo de mapeamento de recursos para PDMA. Assumimos que cada usuário recebe um padrão PDMA. Após a codificação, vários fluxos de dados são transmitidos simultaneamente por diferentes RVs de acordo com a matriz padrão. Como podemos ver na matriz (4.19), os dados do usuário 1 são mapeados para o primeiro, segundo e terceiro RV. Os dados do usuário 2 são mapeados para o primeiro e segundo RV e assim por diante.

#### 4.1.1.2 Fator de Sobrecarga das Matrizes Padrão

O OF é definido como a razão entre o número de colunas e o número de linhas em uma matriz padrão PDMA (VAEZI et al., 2019) dada por

$$\iota = \frac{K}{N} \quad (4.20)$$

O OF reflete o número excessivo de usuários multiplexados nos mesmos elementos de recurso PDMA em comparação com o esquema OMA. Tomando como exemplo  $N = 2$  número do recursos e  $K = 3$  número do usuários, o OF é então  $\iota = \frac{K}{N} = 150\%$ , o que significa que o PDMA suporta uma vez e meia o número de usuários em relação ao OMA.

Para obter um determinado fator de sobrecarga, basta variar as dimensões da matriz padrão PDMA. À medida que o tamanho da matriz aumenta, maiores ganhos de diversidade e um aumento na complexidade de detecção também são fornecidos e, por outro lado, há uma melhoria no desempenho (CHEN et al., 2017). A partir de um OF, a matriz padrão deve ser dimensionada para obter o melhor compromisso entre complexidade e desempenho.

#### 4.1.1.3 Projeto de Matriz Padrão PDMA Estendida

Primeiro, temos que definir a ambiguidade da matriz padrão e depois apresentar a matriz padrão PDMA estendida. Começamos com um exemplo onde temos três usuários e dois valores de recurso RV, representados por  $\mathbf{G}_{\text{PDMA}}^{2 \times 3}$  (VAEZI et al., 2019). Os sinais a serem transmitidos com este RV são dados por

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad (4.21)$$

no qual  $v_j$  é o sinal de transmissão no  $j$ -ésimo RV, e  $x_k$  é o símbolo da modulação do  $k$ -ésimo usuário. O sinal a ser transmitido em cada RV da Eq. (4.21) é dado pela combinação linear dos símbolos obtidos na modulação:

$$\begin{aligned} v_1 &= x_1 + x_2 \\ v_2 &= x_1 + x_3. \end{aligned} \quad (4.22)$$

Vamos supor que os três usuários transmitam com modulação BPSK. Se os usuários 1 e 2, portanto  $x_1$  e  $x_2$ , transmitem o símbolo  $-1$  e o usuário 3, portanto  $x_3$ , o símbolo  $+1$ . A combinação dos sinais transmitidos seriam  $-2$  para o  $v_1$  do RV 1 e  $0$  para o  $v_2$  do RV 2 conforme Eq. (4.22). Para um canal sem ruído, o receptor do RV 1 recebe o  $-2$ , com isso o receptor pode deduzir que o símbolo do RV 1 é  $[-1, -1]$ . O receptor do RV 2 recebe o  $0$ , portanto o receptor pode recuperar o símbolo do RV 2 como  $[+1, -1]$  ou  $[-1, +1]$ , desta forma é impossível recuperar os símbolos transmitidos de forma correta e precisa neste caso.

Com este exemplo, podemos ver que a constelação combinada tem uma distribuição não uniforme que resulta em uma combinação que leva à ambiguidade. O problema da ambiguidade pode ser resolvido adicionando um fator escalar e uma mudança de fase à matriz padrão PDMA. Antes de somar os dois símbolos, o fator de escala de potência

e o fator de deslocamento de fase devem ser inseridos corretamente de acordo com o mapeamento (VAEZI et al., 2019):

$$v = \sqrt{\mu}e^{j\varphi}x_1 + \sqrt{1-\mu}x_2 \quad (4.23)$$

no qual  $\mu$  é o fator escalar de potência e  $\varphi$  é o fator de mudança de fase.

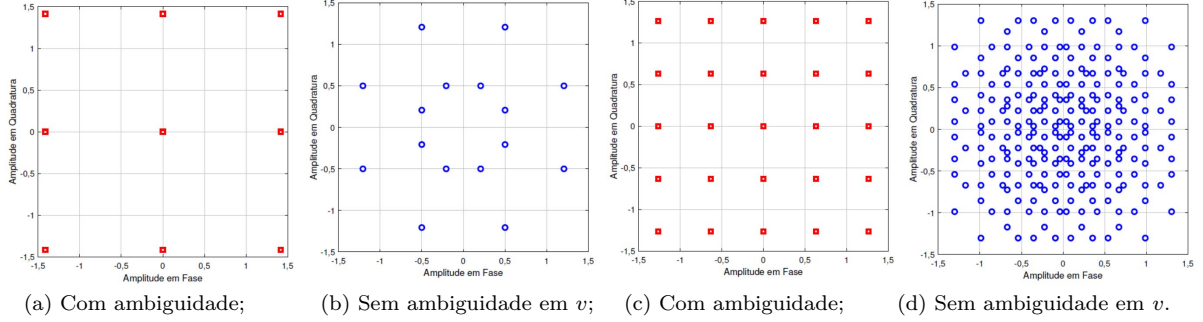


Figura 4.2 – Resultado da soma das combinações da constelação de dois usuários. (a) e (b) representam os resultados para a modulação QPSK, enquanto (c) e (d) para a modulação 16-QAM. Temos (a) e (c) sem as correções de ambiguidade e (b) e (d) com as correções de ambiguidade ao aplicarmos a mudança de fase  $\varphi = \pi/4$  e fator escalar de potência  $\mu = 0.5$  em (4.23).

Considere o exemplo da Figura 4.2 onde  $\varphi = \pi/4$  e  $\mu = 0,5$  foram adotados para dois usuários somados conforme (4.23) para modulação QPSK e modulação 16-QAM. Na Figura 4.2 (a) e (c), temos a soma dos símbolos sem as correções da ambiguidade, no qual pode-se ver que existem alguns resultados dessas somas nas constelações QPSK e 16-QAM, respectivamente, que resultam em zero. Na Figura 4.2 b) e d) com a correção da ambiguidade, o mesmo não ocorre, podendo-se concluir que a ambiguidade foi resolvida. Além disso, essa nova distribuição da constelação somada se aproxima de uma distribuição Gaussiana. Conseqüentemente, há uma maximização da capacidade do canal de ruído Gaussiano branco aditivo AWGN. Isso significa que a introdução de um fator de mudança de fase resulta em um ganho na capacidade do canal.

O fator escalar de potência e o fator de mudança de fase podem ser incorporadas diretamente na matriz padrão PDMA. O valor 1 é substituído pelo valor complexo que corresponde ao fator escalar de potência e ao fator de mudança de fase na matriz padrão PDMA. Essa matriz agora é chamada de matriz de padrão PDMA estendida. Para o exemplo apresentado acima, a matriz padrão estendida é dada por

$$\mathbf{G}_{\text{E-PDMA}}^{2 \times 3} = \begin{bmatrix} \mu_{11}e^{-j\varphi_{11}} & \mu_{12}e^{-j\varphi_{12}} & 0 \\ \mu_{21}e^{-j\varphi_{21}} & 0 & \mu_{23}e^{-j\varphi_{23}} \end{bmatrix} \quad (4.24)$$

no qual  $\mu_{jk}$  e  $\varphi_{jk}$  é o fator de escala de potência e fator de mudança de fase no  $j$ -ésimo RV do  $k$ -ésimo usuário.

Para obter o valor ideal do fator escalar de potência e fator de deslocamento de fase, é necessário conhecer o número de usuários e a constelação adotada.

### 4.1.2 Critérios de Seleção e Avaliação da Matriz Padrão

Esta seção fornece uma metodologia simples para a seleção de matriz PDMA operando sob uso híbrido nos modos uRLLC, mMTC e eMBB. A matriz padrão usada na transmissão DL PDMA deve ser restrita pela potência total de transmissão disponível na BS, limitada por (SUN et al., 2020):

$$\sum_{k=1}^K \sum_{n=1}^N |g_{nk}|^2 = 1 \quad (4.25)$$

Essa limitação garante que a potência total de transmissão da BS não seja maior que o recurso de potência máxima de transmissão disponível na BS para o DL. Além disso, consideramos:

$$\mathbf{G}^0 = \lceil \mathbf{G}_{\text{PDMA}}^{N \times K} \rceil = [\mathbf{g}_{:,1}^0, \mathbf{g}_{:,2}^0, \dots, \mathbf{g}_{:,K}^0] = [\mathbf{g}_{1,:}^{0T}, \mathbf{g}_{2,:}^{0T}, \dots, \mathbf{g}_{N,:}^{0T}]^T \quad (4.26)$$

no qual o símbolo  $\lceil \cdot \rceil$  representa o arredondamento para o primeiro valor inteiro mais alto;  $\mathbf{g}_{:,k}^0 = [g_{1,k}^0, g_{2,k}^0, \dots, g_{N,k}^0]$  é o vetor coluna para o  $k$ -ésimo usuário; de forma similar,  $\mathbf{g}_{n,:}^0 = [g_{n,1}^0, g_{n,2}^0, \dots, g_{n,K}^0]$  é o vetor linha para o  $n$ -ésimo RV associado à matriz  $\mathbf{G}^0$ . Por definição, caso os valores sejam complexos a parte real e a imaginária são arredondadas.

#### 4.1.2.1 Viabilidade da Matriz Padrão

Existem  $\lambda^N$  padrões PDMA possíveis para uma matriz padrão PDMA, sendo  $\lambda$  o número de valores possíveis de  $g_{n,k}$ . Sabendo que  $K$  é o número de colunas necessárias para um determinado valor de OF, pode-se escolher  $K$  padrões PDMA entre  $\lambda^N$  candidatos para construir uma matriz padrão PDMA. Para este caso, o número de combinações possíveis é dado por (REN et al., 2016):

$$C_{\lambda^N}^K = \binom{\lambda^N}{K} = \frac{\lambda^N!}{(\lambda^N - K)!K!} \quad (4.27)$$

Assumindo que  $N$  ou  $\lambda$  crescem,  $C_{\lambda^N}^K$  é muito grande para se encontrar por busca exaustiva a matriz padrão PDMA ideal ou ótima.

#### 4.1.2.2 Desafios na Otimização de Matrizes Padrão para Modos de Uso Híbridos em 5G

O número de combinações possíveis para gerar a matriz padrão pode ser calculado pela Eq. (4.27). Pode-se concluir que a gama de combinações possíveis é extremamente grande, o que torna praticamente impossível a geração de forma exaustiva. Assim, evitamos tal abordagem, a qual requer grandes esforços computacionais; optamos pela adoção e análise de matrizes padrão disponíveis na literatura. Observe que as matrizes padrão disponíveis na literatura são subótimas, dado que a quantidade de combinações possíveis

para se encontrar a matriz padrão PDMA é muito elevada, conforme Eq. (4.27). Neste trabalho, tais matrizes padrão foram analisadas e comparadas adotando-se diversas figuras de mérito, tendo em vista validar nossa metodologia de busca da melhor matriz padrão quase ótima que atenda aos requisitos dos diferentes modos de uso (uRLLC, mMTC e eMBB) em serviços 5G.

A dificuldade em definir a matriz padrão que atende simultaneamente mais de um modo de uso 5G reside nas características antagônicas dos modos de uso. Como exemplo, entre os modos de uso mMTC e uRLLC, o primeiro modo deve atender o maior número possível de terminais do tipo máquina, o que requer uma matriz padrão de maior OF, enquanto uma matriz padrão mais adequada para o serviço uRLLC deve resultar em uma menor OF para atender ao requisito de baixa latência (SUN et al., 2020). Pode-se concluir que dificilmente uma matriz padrão atenderá plenamente todos os requisitos de modos de uso simultaneamente, *i.e.* quando operando simultânea e conjuntamente dois ou três modos. Nesses casos, um modo de uso terá que tolerar uma certa deterioração inevitável em um de seus requisitos. Assim, nossa abordagem é encontrar a matriz padrão PDMA entre as disponíveis na literatura que forneça o melhor desempenho combinado, considerando o sistema 5G operando em modos de uso híbridos (dois ou três) simultaneamente.

### 4.1.3 Critérios de Avaliação da Matriz Padrão

Existem critérios diferentes, mas complementares, para avaliar direta e indiretamente a matriz PDMA, incluindo: a) *Repetibilidade*; b) *Correlação*; c) *Diversidade*; d) *Fator de desempenho*; e) *Peso máximo ponderado da linha*; f) *Eficiência energética do sistema*; g) *Eficiência espectral*; h) *Eficiência de recursos*. Na sequência, discutimos cinco dos oito critérios nominados anteriormente.

#### 4.1.3.1 Repetibilidade

O parâmetro relacionado à interferência entre usuários é denominado *repetibilidade*,  $r$ . Tal parâmetro expressa a repetibilidade dos parâmetros binários correspondentes à matriz padrão. Ao reduzir seu valor, a interferência entre os usuários também é reduzida. A repetibilidade é dada por

$$r = \max[\text{REP}(\mathbf{g}_1^0, \mathbf{g}_2^0, \dots, \mathbf{g}_K^0)] \quad (4.28)$$

no qual  $\text{REP}(\cdot)$  é o número de padrões repetidos. Para melhor entendimento, considere um exemplo simples em que temos  $\mathbf{G} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$  ao aplicarmos a Eq. (4.28) resulta  $r = 1$  porque somente na primeira coluna existe a repetição. Se nas outras colunas houvessem repetições seria escolhido o valor da coluna com maior contagem de repetições.

### 4.1.3.2 Correlação da Soma Quadrática ao Quadrado

Outro parâmetro a ser considerado é a *correlação de soma quadrática ao quadrado* (SSC - *Quadratic Sum Squared Correlation*) (REN et al., 2016). Para maximizar o número de usuários acessando o sistema, o SSC deve ser minimizado, sendo calculado da seguinte forma

$$\text{SSC} = \sum_{k=1}^K \sum_{k'=k+1}^K |\mathbf{g}_k^H \mathbf{g}_{k'}|^2 \quad (4.29)$$

na transmissão DL. Sua redução faz com que a esparsidade aumente (SUN et al., 2020).

### 4.1.3.3 Ordem de Diversidade

A ordem de *diversidade* opera de acordo com a convergência do detector. Com o aumento da diversidade, é possível aumentar a velocidade de convergência do detector. A ordem de diversidade do sistema baseado em PDMA pode ser determinada como (LI et al., 2021)

$$\text{DIV} = \max \left( \sum_{n=1}^N \mathbf{g}_n^{0T} \right) \quad (4.30)$$

### 4.1.3.4 Peso Máximo Ponderado da Linha

O peso máximo ponderado da linha (MRW - *Weighted Maximum Row Weight*) é indicado por (SUN et al., 2020):

$$d_{\max}^w = \sum_{n=1}^N \left[ \frac{1}{2} \sum_{k=1}^K [|g_{n,k}|^2] \left( \sum_{k=1}^K [|g_{n,k}|^2] - 1 \right) \right] \quad (4.31)$$

Observe que a redução sistemática de  $d_{\max}^w$  minimiza a interferência inter-usuário, consequentemente, reduzindo a complexidade da detecção multi-usuário no lado receptor. Quanto menor o valor de MRW, melhor o desempenho do receptor. Para ilustrar o cálculo desta métrica, vamos considerar novamente a matriz padrão da Eq. (4.21); a matriz padrão obtida para modulação QPSK com ambiguidade, Figura 4.2.(a), é dada por

$$\begin{bmatrix} 1+i & -1-i & 0 \\ 1+i & 0 & -1-i \end{bmatrix}, \quad (4.32)$$

e o MRW correspondente é  $d_{\max}^w = 18$ . Para a mesma matriz, mas agora adotando-se uma matriz padrão inequívoca, dada por

$$\begin{bmatrix} 0.7071i & -0.7071 - 0.7071i & 0 \\ 0.7071i & 0 & -0.7071 - 0.7071i \end{bmatrix}, \quad (4.33)$$

como na Figura 4.2.(b), resulta em um valor MRW reduzido de  $d_{\max}^w = 1$ .

#### 4.1.3.5 Fator de Avaliação de Desempenho

Um fator de avaliação de desempenho (PEF - *Performance Evaluation Factor*) abrangente para a matriz padrão PDMA pode ser definido como (LI et al., 2021):

$$\Lambda = \frac{\text{SSC} \cdot d_{\max}^W}{\text{DIV}}. \quad (4.34)$$

Deseja-se reduzir os valores de PEF da matriz padrão a fim de melhorar o desempenho do sistema PDMA. Para ilustrar, consideremos o mesmo exemplo da Subseção 4.1.3.4. A matriz padrão com ambiguidade na Eq. (4.32), com  $\text{SSC} = 8$ ,  $d_{\max}^W = 18$  e  $\text{DIV} = 4$ , resultou em  $\Lambda = 36$ . Assim, comparando-se com a matriz inequívoca, a Eq. (4.33) com  $\text{SSC} = 1$ ,  $d_{\max}^W = 1$  e  $\text{DIV} = 2$  resultou em  $\Lambda = 0.5$ , indicando uma significativa melhoria geral do desempenho do sistema PDMA.

#### 4.1.4 Detecção Híbrida com Filtro Espacial e SIC no Receptor

No receptor, o processo de detecção é usado para recuperar as informações originalmente transmitidas. Neste capítulo, assumimos que o CSI é perfeitamente conhecido no transmissor e no receptor. A combinação de SIC com diversidade espacial/*beamforming* gera esquemas de detecção híbridos, que adicionalmente podem explorar a diversidade no domínio da potência através do NOMA, conforme esquematizado na Figura 4.1. O esquema de detecção híbrido associado é composto pelo SF que suprime a interferência entre feixes causada pela multiplexação no domínio do feixe espacial e pelo SIC que remove a interferência intra-feixe causada pela multiplexação no domínio da potência.

Primeiro, vamos considerar a matriz SF dada por  $\mathbf{V}_k \in \mathbb{C}^{N_R \times N}$  para o  $k$ -ésimo usuário e  $\mathbf{v}_{nk}$  o vetor da  $n$ -ésima coluna de  $\mathbf{V}_k$ . Nesta etapa, o SF é executado no vetor de sinal recebido para suprimir a interferência entre feixes. A resultante escalar é representada por  $z_{nk}$  e pode ser expressa da seguinte forma

$$z_{nk} = \mathbf{v}_{nk}^H \mathbf{y}_k = \underbrace{\mathbf{v}_{nk}^H \mathbf{H}_k \mathbf{b}_n t_n}_{\text{Interferência Intra-feixe}} + \underbrace{\mathbf{v}_{nk}^H \mathbf{H}_k \sum_{\substack{i=1 \\ i \neq n}}^N \mathbf{b}_i t_i}_{\text{Interferência Inter-feixe}} + \mathbf{v}_{nk}^H \mathbf{w}_k, \quad (4.35)$$

$$n = 1, 2, \dots, N \quad \text{e} \quad k = 1, 2, \dots, K,$$

no qual o primeiro termo do lado direito representa a combinação de informação desejada e interferência intra-feixe, enquanto os outros dois termos representam interferência inter-feixe e ruído, respectivamente.

A matriz SF  $\mathbf{V}_k$  pode ser construída com base nos critérios MMSE ou ZF. Suponha que  $\mathbf{V}_k$  calculado com base nos critérios MMSE (HIGUCHI; KISHIYAMA, 2013) como

segue

$$\mathbf{V}_k = \min_{\tilde{\mathbf{v}}_k} \mathbb{E} \left\{ \left\| \mathbf{t} - \tilde{\mathbf{V}}_k^H \mathbf{y}_k \right\|_2^2 \right\} \quad (4.36)$$

$$= (\mathbf{H}_k \mathbf{B} \mathbf{A} \mathbf{B}^H \mathbf{H}_k^H + \sigma_k^2 \mathbf{I}_{N_R})^{-1} \mathbf{H}_k \mathbf{B} \mathbf{A}, \quad k = 1, 2, \dots, K, \quad (4.37)$$

no qual  $\mathbf{A} \triangleq \mathbb{E} \{ \mathbf{t} \mathbf{t}^H \}$  com seu elemento sendo expresso como  $[\mathbf{A}]_{i,j} = \sum_{k=1}^K g_{i,k} g_{j,k} \sqrt{p_{i,k} p_{j,k}}$ . A expressão para  $\mathbf{V}_k$  desenvolvida com base no critério ZF (JIANG et al., 2019) tem a seguinte forma

$$\mathbf{V}_k = \mathbf{H}_k \mathbf{B} (\mathbf{B}^H \mathbf{H}_k^H \mathbf{H}_k \mathbf{B})^{-1}, \quad k = 1, 2, \dots, K. \quad (4.38)$$

Em segundo lugar, considere a potência agregada de interferência e ruído entre feixes em  $z_{nk}$  normalizada em um. A normalização é aplicada ao sinal escalar recebido para atender à condição de implementação do SIC. Seja ainda  $u_{nk}$  o ganho de canal normalizado equivalente entre o  $k$ -ésimo usuário e o *cluster* de antenas (AC), expresso por (JIANG et al., 2019):

$$u_{nk} = \sqrt{\frac{|\mathbf{v}_{nk}^H \mathbf{H}_k \mathbf{b}_n|^2}{\sum_{\substack{i=1 \\ i \neq n}}^N |\mathbf{v}_{nk}^H \mathbf{H}_k \mathbf{b}_i|^2 + \sigma_k^2 \|\mathbf{v}_{nk}\|^2}}, \quad n = 1, 2, \dots, N \quad \text{e} \quad k = 1, 2, \dots, K. \quad (4.39)$$

Consequentemente, a Eq. (4.36) pode ser reescrita da seguinte forma

$$z_{nk} = u_{nk} \sum_{i=1}^K g_{ni} \sqrt{p_{ni}} s_i + q_{nk}, \quad n = 1, 2, \dots, N \quad \text{e} \quad k = 1, 2, \dots, K, \quad (4.40)$$

no qual  $q_{nk}$  representa a soma da interferência entre feixes e do ruído após a normalização, sendo  $\mathbb{E} [ |q_{nk}|^2 ] = 1$ . A matriz dada por  $\mathbf{U} \in \mathbb{R}^{N \times K}$  corresponde ao ganho de canal normalizado equivalente entre UG e AC onde podemos representar  $[\mathbf{U}]_{nk} = u_{nk}$ . Assim, o canal MIMO entre o  $k$ -ésimo usuário e o AC pode ser degradado em canais de entrada única e saída única (SISO - *Single-Input Single-Output*) após a normalização (HIGUCHI; KISHIYAMA, 2013), que atende a condição de implementação do SIC.

O SIC é empregado sobre o sinal escalar normalizado recebido com o intuito de remover a interferência intra-feixe. Também é usado para detectar símbolos subtraindo os símbolos detectados de usuários transmitidos com maior potência para facilitar a detecção de usuários transmitidos com menor potência. Assumimos que os usuários são classificados em ordem crescente de ganho de canal normalizado, então  $u_{ni} \leq u_{nj}$  é válido se  $1 \leq i \leq j \leq K$ . Como resultado, o  $j$ -ésimo usuário pode detectar corretamente seu símbolo de sinal apesar da interferência do  $i$ -ésimo usuário. Consequentemente, a SINR do  $k$ -ésimo usuário no  $n$ -ésimo feixe pode ser expressa por (SUN et al., 2020):

$$\gamma_{nk} = \begin{cases} \frac{|u_{nk}|^2 \cdot |g_{nk}|^2 \cdot p_{nk}}{1 + |u_{nk}|^2 \cdot \sum_{i=k+1}^K |g_{ni}|^2 \cdot p_{ni}}, & n = 1, 2, \dots, N \quad \text{e} \quad k = 1, 2, \dots, K-1, \\ |u_{nk}|^2 \cdot |g_{nk}|^2 \cdot p_{nk}, & n = 1, 2, \dots, N \quad \text{e} \quad k = K \end{cases} \quad (4.41)$$

Portanto, a soma das taxas dos ACs pode ser expressa da seguinte forma

$$R_{\Sigma} = \sum_{n=1}^N \sum_{k=1}^K \log_2(1 + \gamma_{nk}) \quad \left[ \frac{\text{bits}}{s \cdot \text{Hz}} \right]. \quad (4.42)$$

Finalmente, a eficiência energética é dada pela razão entre a somatória das taxas de dados alcançada pela potência total consumida no sistema:

$$\text{EE} = \frac{R_{\Sigma}}{P_{RF} + P_c} = \frac{R_{\Sigma}}{P_{RF} + N_R \cdot K \cdot N \cdot P_K + N \cdot N_T \cdot P_{BS}} \quad \left[ \frac{\text{bits}}{\text{J} \cdot \text{Hz}} \right] \quad (4.43)$$

onde  $P_c$  é o consumo de potência dos circuitos necessários ao processamento de sinal, resfriamento, etc, nos transmissores e receptores. Para simplificar, apenas as partes do consumo de potência  $P_K$  e  $P_{BS}$  foram explicitamente consideradas, sendo  $P_K$  a potência necessária para operar os circuitos de cada antena do usuário móvel e  $P_{BS}$  a potência consumida para operar os componentes do circuito de cada antena na BS. Além disso, o consumo total de potência de RF é escrito como:

$$P_{RF} = \varrho \cdot \sum_{n=1}^N \sum_{k=1}^K \beta_{nk} \cdot P = \varrho \cdot P \quad [W] \quad (4.44)$$

no qual  $\varrho \geq 1$  é a ineficiência do amplificador de potência de RF, assumida a mesma para todos que alimentam cada antena de transmissão e  $\beta_{nk}$  é a proporção de potência para o  $k$ -ésimo usuário no  $n$ -ésimo feixe. A soma das potências entre os usuários é dada por  $\sum_{n=1}^N \sum_{k=1}^K \beta_{nk} \cdot P$  no qual  $\sum_{n=1}^N \sum_{k=1}^K \beta_{nk} = 1$  e  $p_{nk} = \beta_{nk} \cdot P$ .

#### 4.1.5 Problema de Otimização de Eficiência Energética

O projeto de otimização de EE no sistema mMIMO-PDMA para obter a alocação de potência ótima  $P$ , sob valores iguais de taxa de dados  $R$  (equidade de taxa) e proporção de potência  $\beta_{nk}$  para todos os usuários pode ser definido por um problema equivalente da seguinte forma:

$$\begin{aligned} & \max_{(P, \beta_{a,k})} \frac{R_{\Sigma}}{P_{RF} + N_R \cdot K \cdot N \cdot P_K + N \cdot N_T \cdot P_{BS}} \\ \text{s.t.} \quad & \text{C.1} \quad \sum_{n=1}^N \sum_{k=1}^K \beta_{nk} = 1 \\ & \text{C.2} \quad \beta_{nk} = \frac{(2^R - 1)}{u_{n,k}^2 |g_{nk}|^2 P} + \frac{(2^R - 1)}{|g_{nk}|^2} \sum_{i=k+1}^K |g_{ni}|^2 \beta_{ni}, \quad \text{no qual} \quad \begin{cases} 1 \leq n \leq N \\ 1 \leq k \leq K - 1 \end{cases} \quad (4.45) \\ & \text{C.3} \quad \beta_{nK} = \frac{(2^R - 1)}{u_{nK}^2 |g_{nK}|^2 P}, \quad \text{para todo } 1 \leq n \leq N \\ & \text{C.4} \quad P > 0 \quad \text{e} \quad \beta_{nk} > 0 \end{aligned}$$

A função objetivo em (4.45) não é convexa em  $P$  e  $\beta_{nk}$  simultaneamente, apesar das restrições C.1–C.4 formar um conjunto convexo. A prova de convexidade das restrições é

---

**Algoritmo 8:** SQP-PDMA – Programação quadrática sequencial para a otimização da EE no sistema mMIMO PDMA.

---

- 1 **Escolher:** ponto inicial  $(\mathbf{x}^{(0)}, \boldsymbol{\mu}^{(0)}, \boldsymbol{\lambda}^{(0)})$  e tolerância  $\zeta > 0$ ;
  - 2 **Definir:**  $k = 0$ ;
  - 3 **enquanto**  $\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| > \zeta$  **faça**
  - 4     **Calcular:**  $f(\mathbf{x}^{(k)})$ ,  $\nabla f(\mathbf{x}^{(k)})$ ,  $\nabla^2 \mathcal{L}(\mathbf{x}^{(k)})$ ,  $\mathbf{h}(\mathbf{x}^{(k)})$ ,  $\nabla \mathbf{h}(\mathbf{x}^{(k)})$ ,  $g(\mathbf{x}^{(k)})$  e  $\nabla g(\mathbf{x}^{(k)})$ ;
  - 5     **Resolver:** (3.35) para obter  $\mathbf{d}^{(k)}$  e  $\boldsymbol{\mu}^{(k+1)}$ ;
  - 6      $\mathbf{x}^{(k+1)} \leftarrow \mathcal{P}(\mathbf{x}^{(k)} + \bar{\alpha}_k \mathbf{d}^{(k)})$ ; no qual  $\mathcal{P}$  é o operador de projeção ortogonal no conjunto  $\Omega$ ;
  - 7 **fim**
  - 8 **Saída:**  $\mathbf{x}^* = \mathbf{x}^{(k)}$ .
- 

semelhante àquela fornecida em (JACOB et al., 2021b), disponível no Anexo B; portanto, omitida aqui por conveniência.

Para resolver o problema (4.45) de forma eficiente, modificamos ligeiramente a função objetivo tendo em vista implantar convenientemente a técnica de SQP como em (JACOB et al., 2021a), descrita aqui na subseção 3.2.3.2 e no Algoritmo 8, visando obter uma solução quase ótima sob convergência rápida apesar da não-convexidade da função objetivo.

## 5 Resultados e Discussões

Este capítulo refere-se aos resultados de simulação e discussões decorrentes da metodologia proposta para a obtenção do compromisso entre EE e SE, para um número fixo de usuários por *cluster* com máxima equidade de taxa e máxima eficiência energética ou, alternativamente, máxima eficiência de recursos (RE, isto é, compromisso EE-SE), para os sistemas NOMA, MIMO-NOMA e mMIMO PDMA operando no enlace de descida. A ferramenta utilizada para a otimização foi a função MatLab<sup>®</sup> *fmincon*. Adotou-se faixa de valores crescentes (do menor ao maior) dos parâmetros de sistema, tendo em vista caracterizar completamente o comportamento do sistema. Desta forma, pode-se traçar o comportamento de um sistema de comunicação sem fio multiusuário NOMA SISO, MIMO e MIMO massivo considerando diferentes cenários práticos de interesse.

### 5.1 Resultados Numéricos para Sistemas NOMA

Os valores dos parâmetros e os cenários usados nas simulações do sistema NOMA são descritos na Tabela 5.1. A técnica utilizada foi o algoritmo SQP. Ao contrário do cenário 4, onde os usuários têm distâncias aleatórias seguindo uma distribuição uniforme para os respectivos  $\mathfrak{R}_D$ , nos cenários 1, 2 e 3 de raios 0, 34; 0, 80 e 1, 80 km, respectivamente, os usuários possuem posições fixas e diferentes distâncias da BS.

A Figura 5.1 mostra a distribuição dos usuários em relação à estação base em uma célula. Para comparar a soma das taxas com taxas igualitárias entre os usuários, usamos aqui a técnica de atribuição de potência entre os usuários proporcional ao inverso da resposta do canal (ICA - *inverse of channel response*) (EL-SAYED et al., 2016), dada por:

$$P_i = \frac{\sum_{k=1}^M |h_k|^2}{|h_i|^2} P, \quad i = 1, \dots, M. \quad (5.1)$$

#### 5.1.1 Desempenho ERPA *versus* ICA em Sistemas NOMA

Com o objetivo de demonstrar a eficácia da otimização ERPA proposta seguindo o princípio de equidade de taxa (mesma taxa para todos os usuários), a Figura 5.2 mostra a respectiva somatória ótima e não ótima, considerando quatro usuários com taxas fixas iguais a 0,5; 1; 1,5; 2 e 2,5 bits/s/Hz e potências médias ótimas para o cenário 1 com  $\alpha = 2$  para cada usuário. A estratégia de ERPA é comparada com a obtida pela Eq. (5.1)

Tabela 5.1 – Parâmetros Adotados para os cenários de Simulação.

Parâmetro	Valor
Número de Usuários na Célula	$M = 2$ a $12$
Expoente de Perda de Percurso	$\alpha = [2, 0; 3, 0; 4, 5]$
Desvanecimento (termo curto)	$h_i \in \mathcal{CN}(0, 1)$
Potência do Ruído Branco	$\sigma_n^2 = 0,1 \mu\text{ W}$
Orçamento de Potência	$P = 120\text{ W}$
Meta da Taxa de Usuário Igual	$R \in [0, 1; 0, 5; 1; 1, 5; 2; 2, 5; 3]$ [bit/s/Hz]
Potência da Circuitaria	$P_c = 250$ $\left[\frac{\text{mW}}{\text{usuário}}\right]$
Ineficiência do Amplificador de Potência RF	$\rho = 1,4$
Número de Amostras	$10^4$
Ensaio da Simulação Monte-Carlo	$5$
<b>Cenários: Distribuição dos Usuários Dentro da Célula</b>	
Cen. 1 – $\sum d_i = 1,0$ [km]	$d_1 = 0,34, d_2 = 0,29, d_3 = 0,22, d_4 = 0,15.$
Cen. 2 – $\sum d_i = 2,0$ [km]	$d_1 = 0,8, d_2 = 0,6, d_3 = 0,4, d_4 = 0,2.$
Cen. 3 – $\sum d_i = 4,0$ [km]	$d_1 = 1,8, d_2 = 1,2, d_3 = 0,7, d_4 = 0,3.$
Cen. 4 – $\mathfrak{R}_D$ [m] ( Distrib. Uniform.)	$50, 100, 200, 300, 400$
<b>Estratégia de Alocação de Potência</b>	
Alocação Inversa de Potência (ICA)	$P_i$ definido por Eq. (5.1)
Alocação de Potência Ótima para a Taxa Igual (ERPA)	Subseção 2.1.1.1

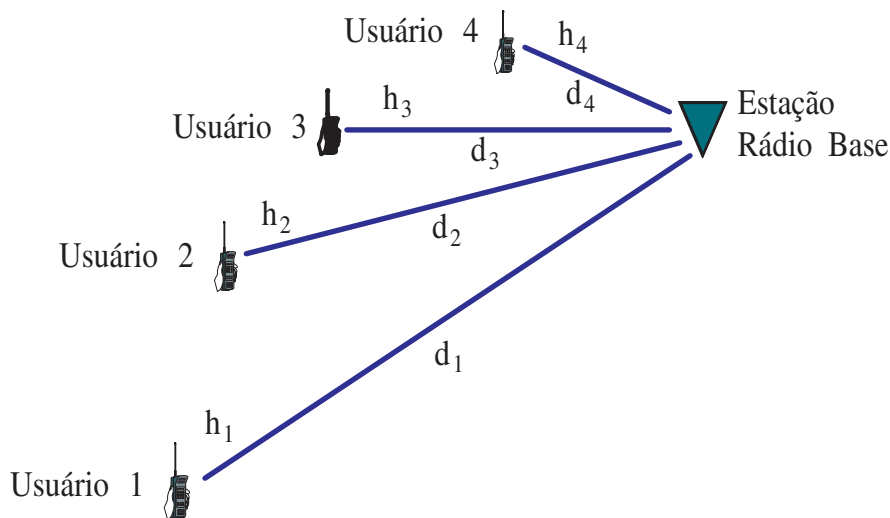


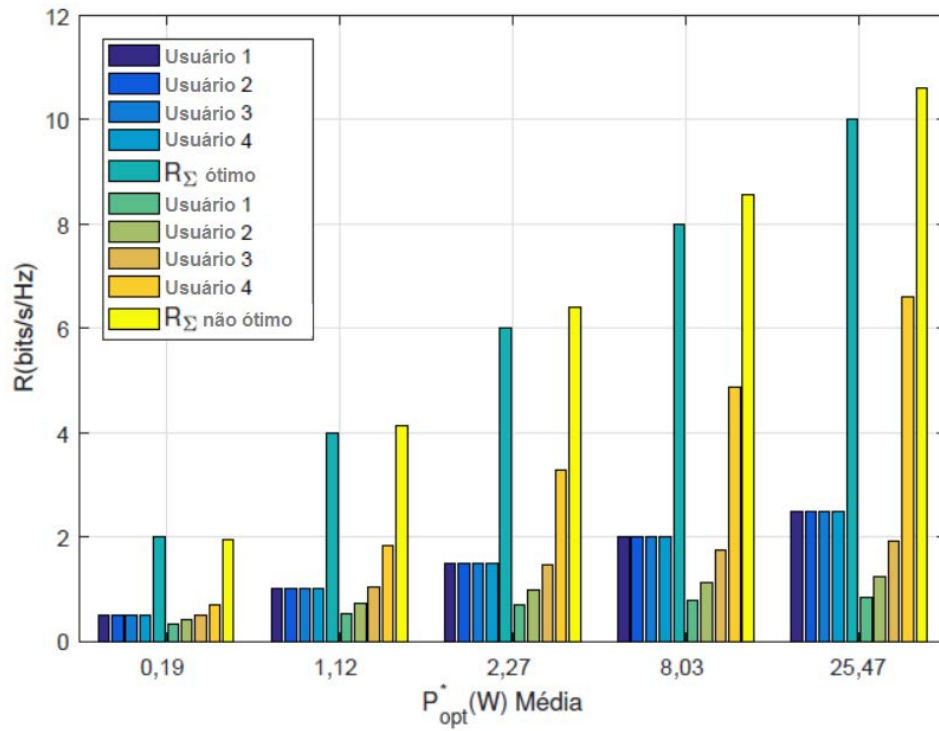
Figura 5.1 – Distribuição de usuários em um sistema NOMA com 4 usuários.

com a mesma potência total. Observa-se que apenas o sistema ótimo manteve as mesmas taxas para todos os usuários ao custo de reduzir marginalmente a taxa de consumo em cada cenário de potência total. No sistema não ótimo, quando comparado ao sistema ótimo, o usuário 4 que está mais próximo da BS, recebe uma maior proporção de potência (não ótima) do que a proporção de potência ótima; com isso, causa maior interferência nos demais usuários e conseqüentemente reduz suas taxas, aumentando a própria taxa e resultando na somatória de taxas um pouco maiores. Observe que, para garantir a equidade, o usuário mais forte (mais próximo) perde parcialmente a taxa, enquanto os usuários distantes ganham taxa ao alocar mais energia para os usuários de borda. Esse mecanismo ocorre para todos os cenários de taxas de potência analisados. Este é o preço pago para se atingir equidade de taxas entre os usuários internos e os de borda. De fato, para manter todos os usuários na taxa de  $R = 2$  bits/s/Hz no cenário 1 e  $\alpha = 2$ , a potência total necessária é 8,03 W, Figura 5.2, e para o mesmo caso no cenário 2, a potência total necessária é 23,48 W. Ou seja, a potência total deve ser 2,9 vezes maior para o cenário 2, porque os usuários estão mais distantes.

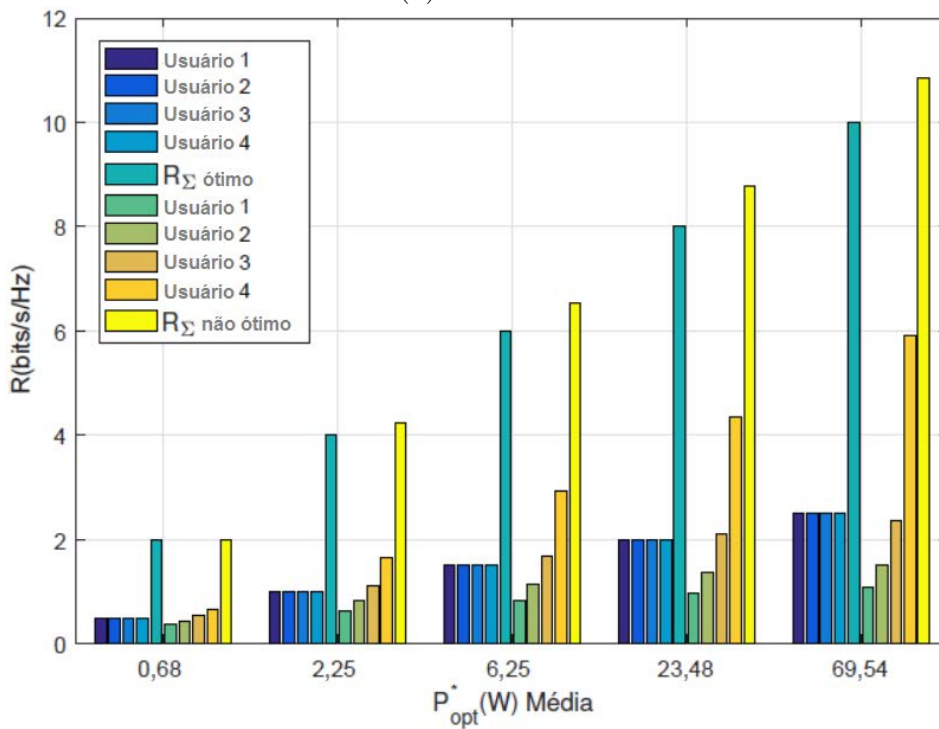
A Figura 5.3 mostra a soma das taxas para distribuição ótima e não ótima de potências totais, para cenários com distâncias diferentes entre usuários, e para  $\alpha = 2$  e 3. Com a mesma potência total, as somas das taxas não ótimas são um pouco maiores para a distribuição de potência não ótima porque os usuários mais próximos, com distâncias  $d_2$ ,  $d_3$  e  $d_4$  têm uma proporção de potência maior e conseqüentemente maior taxa do que o usuário mais distante, com distância  $d_1$ . Por exemplo, no ambiente com  $\alpha = 2$  e cenário 3, a soma das taxas não ótimas com um orçamento de energia médio de  $P_{opt}^* \approx 59,7$  W é  $R_\Sigma = 8,8$  bits/s/Hz enquanto para a soma das taxas ótimas é  $R_\Sigma = 8,0$  bits/s/Hz para o mesmo orçamento de energia de 59,7 W.

Na Figura 5.4.(a), temos a variação das proporções de potência dos usuários  $\beta_{opt}^*$  em função da taxa de cada usuário dada por  $R$  em bits/s/Hz. As proporções de potência para usuários do cenário 2 com expoente de perda de percurso  $\alpha = 2$  e cenário 1 com  $\alpha = 3$  têm uma diferença marginal, pois o aumento das distâncias dos usuários no primeiro é compensado pelo expoente de perda de percurso aumentado do segundo. Para os usuários do cenário 1 e do expoente de perda de percurso  $\alpha = 2$ , a proporção de potência dos usuários é menor devido à diminuição do efeito de perda de percurso.

Para se ter uma visão de como as proporções de potência  $\beta_i$  se comportam em relação à potência total  $P$  e à taxa fixa  $R$  de cada usuário temos a Figura 5.4.(b). Para uma taxa fixa de  $R = 1,5$  bits/s/Hz por usuário, a potência total ótima é em torno de 2,3 W; 6,2 W e 21,0 W para os cenários 1, 2 e 3, respectivamente, enquanto a proporção de potência para usuários 1 está em torno de  $\beta_1 = 0,7$ ; 0,71 e 0,73. Observe que a mesma proporção de potência, mas com potência total diferente e a mesma taxa, ocorrem devido às diferenças de distância entre os usuários, resultando no mesmo para os três cenários.



(a) Cenário 1



(b) Cenário 2

Figura 5.2 – Taxas de dados ( $R$ ) e soma das taxas ( $R_{\Sigma}$ ) versus potência alocada considerada perda de percurso ( $\alpha = 2$ ). As taxas não ótimas foram obtidas com a proporção de potência dada pela Eq. (5.1).

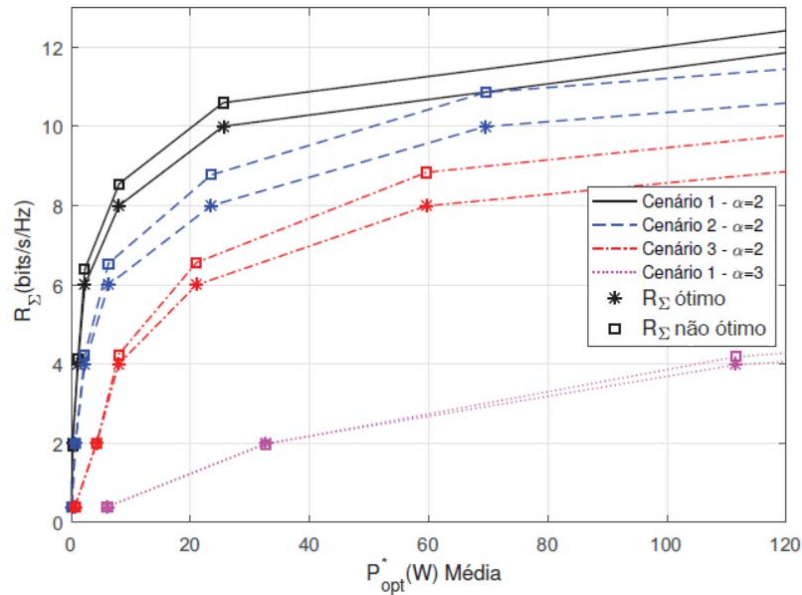


Figura 5.3 –  $R_{\Sigma}$  versus  $P$  para  $\alpha = 2$  e 3. As taxas não ótimas foram obtidas com a proporção de potência dada pela Eq. (5.1).

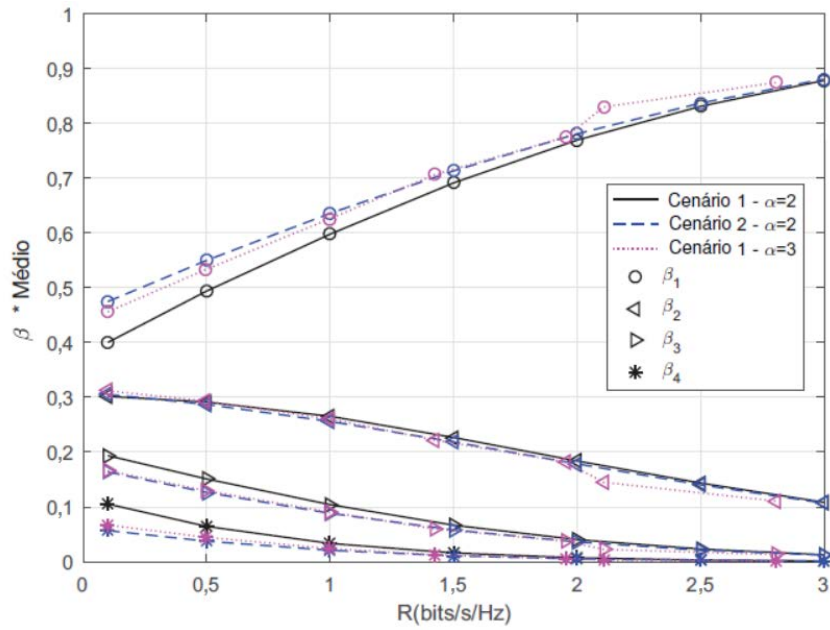
### 5.1.2 Equidade sob Sistemas NOMA

O índice de equidade (*fairness*)  $\mathcal{F}_{it}$  definido em (2.17) foi aplicado à estratégia ERPA-NOMA proposta considerando os quatro cenários diferentes. Conforme mostrado na Tabela 5.2, os resultados do índice de equidade para diferentes cenários e taxas-alvo estão muito próximos a 1<sup>1</sup>, quer seja ao aplicarmos a definição de equidade baseada na teoria da informação (2.17), quer seja ao aplicarmos o índice de equidade de Jain (2.16), demonstrando a eficácia do método de otimização proposto. Comparando os cenários 1 a 3 para  $\alpha = 2$ , pode-se concluir que o aumento no índice de equidade é gradual, aproximando-se de 1 quando a soma da distância dos usuários aumenta (cenário 3), pois o desvio quadrático estatístico entre as taxas diminui. Além disso, o índice de equidade  $\mathcal{F}_{it}$  aumenta quando  $\alpha > 2$  pelo mesmo motivo. Finalmente,  $\mathcal{F}_{it}$  diminui lentamente de forma constante quando a taxa igual por usuário  $R$  aumenta, demonstrando a dificuldade em atingir metas de QoS vinculadas mais restritivas.

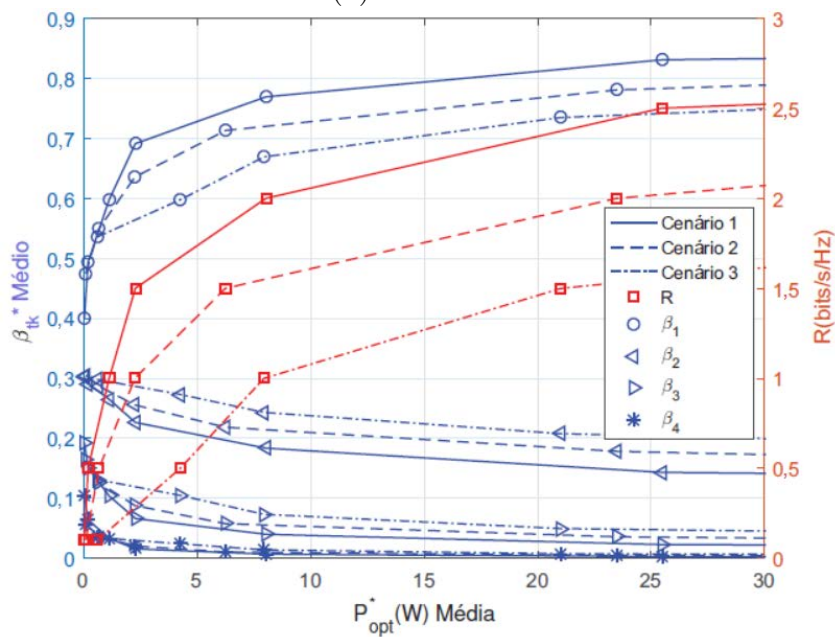
Tabela 5.2 – Equidade  $\mathcal{F}_{it}$  obtida da Eq. (2.17).

R(bits/s/Hz)	$\alpha = 2$			$\alpha = 3$
	Cenário 1	Cenário 2	Cenário 3	Cenário 1
1,0	0,947	0,968	0,976	0,953
1,5	0,943	0,962	0,970	0,948
2,0	0,941	0,958	0,966	0,945
2,5	0,940	0,955	0,962	0,943
3,0	0,939	0,952	0,960	0,942

<sup>1</sup>Dado que as taxas finais atingidas pelos usuários de borda e internos são todas iguais, o índice de equidade resulta igual a 1.



(a)  $\alpha = 2$  e  $3$ .



(b)  $\alpha = 2$ .

Figura 5.4 – Proporção de potência ótima e comportamento da soma das taxas para os três cenários e diferentes expoentes de perda de percurso: a)  $\beta_{opt}^* \times R$ ; b)  $\beta_{opt}^* \times P_{opt}^*$  e  $R \times P_{opt}^*$ .

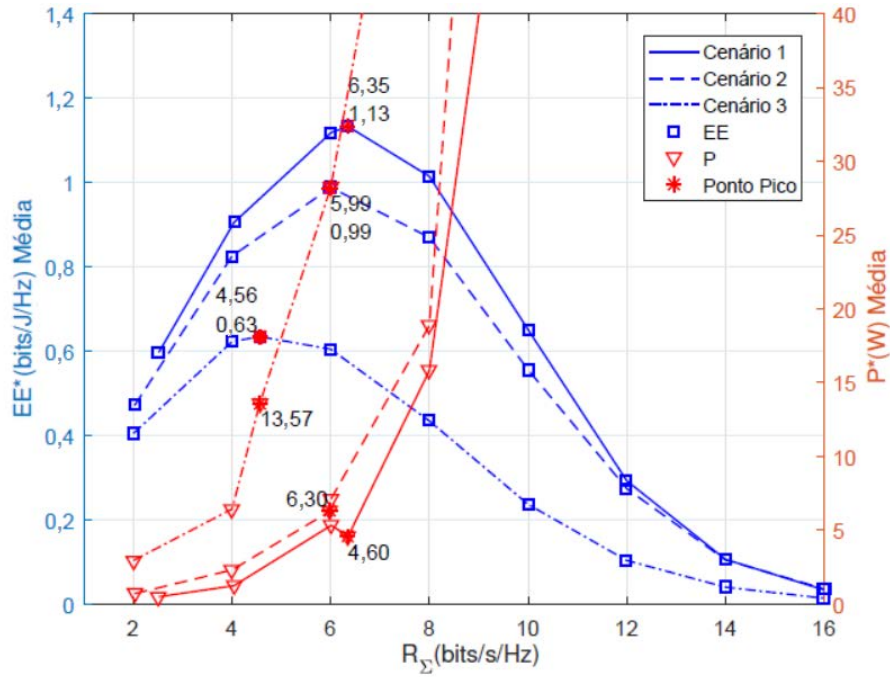
### 5.1.3 Pontos de Operação Ótimos para Eficiência Energética no NOMA

O ponto ótimo para a eficiência energética ( $EE^*$ ) é representado na Figura 5.5.(a) e (b) para os três cenários de valor fixo e considerando o canal com  $\alpha = 2$ . Para o cenário 1, o ponto de operação  $EE^* = 1,13$  bits/J/Hz resulta em uma taxa de  $R = 6,35$  bits/s/Hz para um consumo total de energia de  $P = 4,60$  W. Nos cenários 2 e 3, temos taxas de  $R = 5,99$  bits/s/Hz e  $R = 4,56$  bits/s/Hz para uma potência total de  $P = 6,30$  W e  $P = 13,57$  W, que resultou em  $EE^* = 0,99$  bits/J/Hz e  $EE^* = 0,63$  bits/J/Hz, respectivamente. No cenário 1 (usuários próximos), menos energia é necessária para atingir a taxa especificada, de forma que a energia muito baixa atinge a taxa mais alta e, conseqüentemente  $EE^*$  maior, como pode ser visto no gráfico da Figura 5.5.(a) e (b) nos pontos do asterisco. De fato, o ponto de  $EE^*$  para o cenário 1 é maior do que os usuários dos cenários 2 e 3, uma vez que os usuários do cenário 1 estão mais próximos da BS, o que pode ter corroborado para o comportamento da Figura 5.5.(b). Tendo em vista avaliar o potencial de cada técnica, assumimos a hipótese de SIC perfeito, com isso, descartamos as configurações que resultassem potências recebidas entre os usuários na mesma célula NOMA muito semelhantes. Isso ocorre quando os terminais móveis na mesma célula NOMA estão localizados geometricamente muito próximos e os ganhos semelhantes.

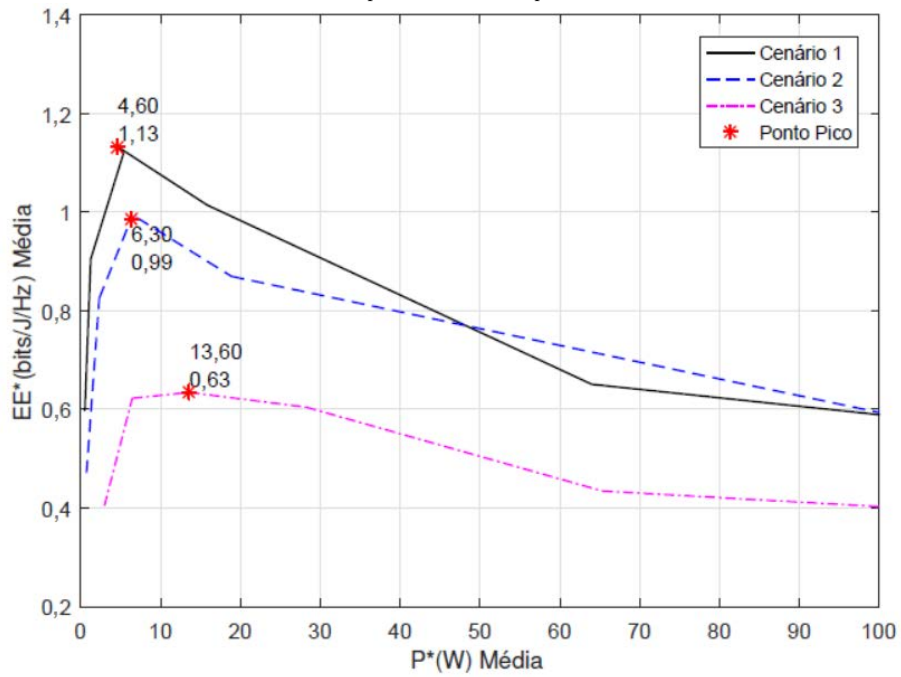
### 5.1.4 Ponto Ótimo da Eficiência de Recursos ( $EE \times SE$ ) em Função do Número de Usuários NOMA

Nas Figuras 5.6, 5.7 e 5.8 pode-se ver os valores ótimos para eficiência energética, soma das taxas e potência total em função do número de usuários  $M$  da perspectiva do compromisso entre  $EE \times SE$ . Nesse caso, a soma da taxa  $R_\Sigma$  e a potência total  $P$  são consideradas como variáveis no processo de otimização da Eq. (2.13). Para a obtenção das curvas, considerou-se que os usuários possuem distribuição uniforme dentro da célula de raio  $\mathfrak{R}_D$  e  $P_c = 250$  mW por usuário (DEVROYE, 1986). Consideramos  $10^4$  amostras para cada número de usuários  $M$ . A Tabela 5.3 resumiu os pontos relevantes dos resultados numéricos obtidos nas Figuras 5.6, 5.7 e 5.8, onde o melhor compromisso entre  $EE$  e  $R_\Sigma$  é selecionado para cada cenário. Essa compensação representa o número de usuários que obtêm a melhor eficiência de recursos do sistema (RE) para  $\xi = 1,8$  na Eq. (2.15).

Para os diferentes raios  $\mathfrak{R}_D$  e  $\alpha$  considerados na Tabela 5.3, a melhor RE foi obtida para a configuração de dois usuários por célula. O  $EE$  e  $SE$  reduzem com o aumento dos raios das células e conseqüentemente o aumento das potências ótimas obtidas devido às distâncias dos usuários serem maiores. Esses valores obtidos são as melhores configurações para obter a maior eficiência de recursos do sistema NOMA nos cenários apresentados.

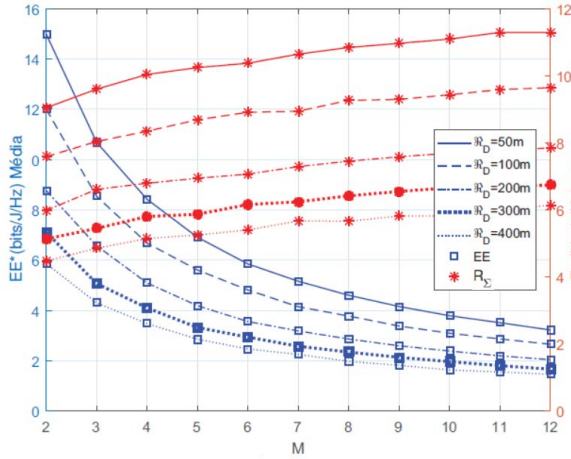


(a)  $EE_{opt} \times R_{\Sigma}$  e  $P_{opt} \times R_{\Sigma}$

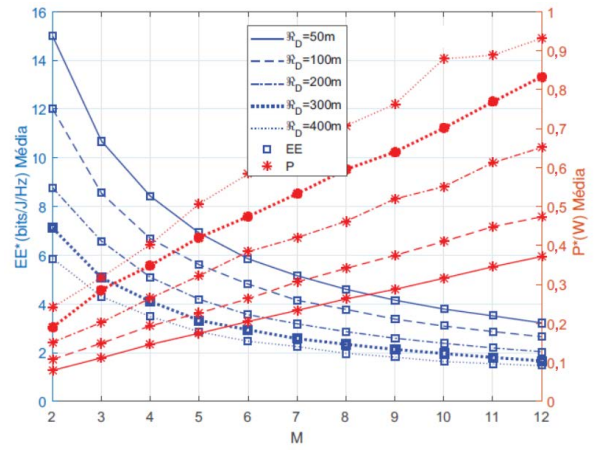


(b)  $EE_{opt} \times P_{opt}$

Figura 5.5 – Os pontos ótimos considerando  $\alpha = 2$ .

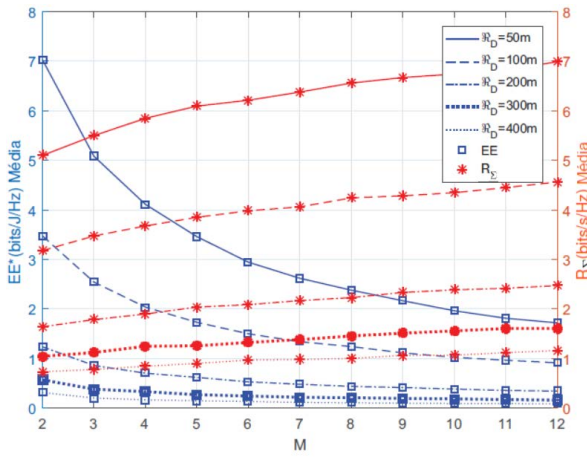


(a)  $EE_{opt}^* \times M$  e  $R_{\Sigma} \times M$

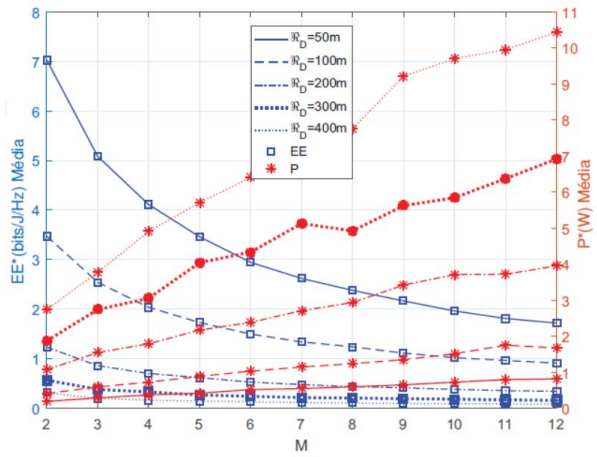


(b)  $EE_{opt}^* \times M$  e  $P_{opt}^* \times M$

Figura 5.6 – Compromisso entre EE e SE para alguns Raios  $\mathfrak{R}_D$  e  $\alpha = 2$ .



(a)  $EE_{opt}^* \times M$  e  $R_{\Sigma} \times M$



(b)  $EE_{opt}^* \times M$  e  $P_{opt}^* \times M$

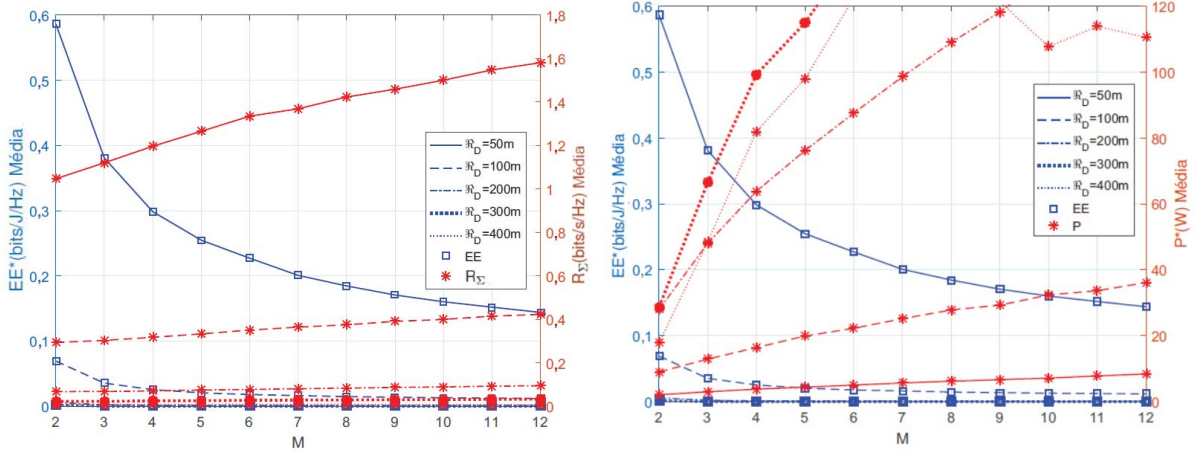
Figura 5.7 – Compromisso entre EE e SE para alguns Raios  $\mathfrak{R}_D$  e  $\alpha = 3$ .

Tabela 5.3 – Ponto de Compromisso entre EE e  $R_{\Sigma}$  Para alguns  $\mathfrak{R}_D$  e Usuários com Distribuição Uniforme.

$\mathfrak{R}_D$	$\alpha = 2$				$\alpha = 3$				$\alpha = 4, 5$			
	Us.	EE	$R_{\Sigma}$	P	Us.	EE	$R_{\Sigma}$	P	Us.	EE	$R_{\Sigma}$	P
50	2	15,00	9,05	0,08	2	7,03	5,10	0,20	2	0,59	1,05	2,11
100	2	12,01	7,59	0,11	2	3,48	3,18	0,41	2	0,07	0,29	9,03
200	2	8,77	5,98	0,15	2	1,24	1,64	1,07	2	0,006	0,07	27,94
300	2	7,12	5,13	0,19	2	1,04	0,57	1,87	2	0,0005	0,02	28,65
400	2	4,48	5,86	0,24	2	0,32	0,73	2,74	2	0,0005	0,008	17,92

## 5.2 Resultados Numéricos para MIMO-NOMA

Os valores dos parâmetros usados nas simulações são descritos na Tabela 5.4. Para essa configuração também foi utilizado o algoritmo SQP (NOCEDAL; WRIGHT, 1996) por ser um método iterativo poderoso para otimização não linear. Nos cenários 1, 2 e 3, os usuários têm uma distância fixa da BS, sendo que os índices fora dos parênteses indicam


 (a)  $EE_{opt}^* \times M$  e  $R_\Sigma \times M$ 

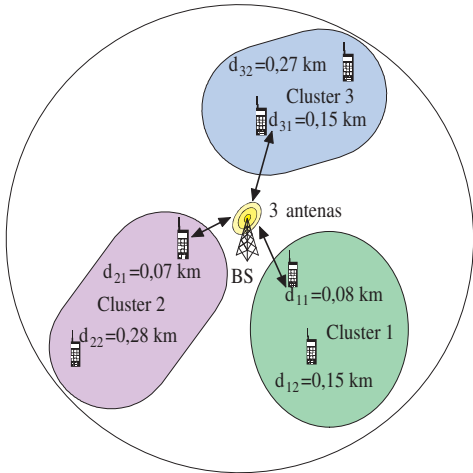
 (b)  $EE_{opt}^* \times M$  e  $P_{opt}^* \times M$ 

 Figura 5.8 – Compromisso entre EE e SE para alguns Raios  $R_D$  e  $\alpha = 4, 5$ .

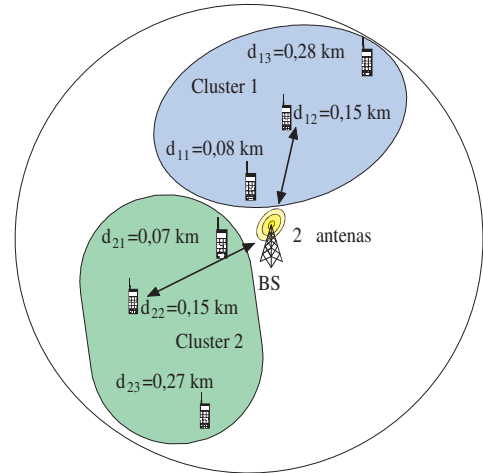
configuração com 3 *clusters* e 2 usuários/*cluster*, enquanto os índices com parênteses indicam configuração com 2 *clusters* e 3 usuários/*cluster*; entretanto, os usuários no cenário 4 têm uma distribuição uniforme para o respectivo  $R_D$ . A soma das distâncias entre usuários e BS aumenta progressivamente nos cenários 1, 2 e 3 e os raios de cobertura para cada cenário também, sendo iguais a 0,28; 0,60 e 0,90 km, respectivamente. Os cenários 1, 2 a 3 da Tabela 5.4 são mostrados na Figura 5.9. Os 3 cenários do lado esquerdo com as letras (a), (c) e (e) possuem uma distribuição de 3 *clusters* com 2 usuários/*cluster* e do lado direito com as letras (b), (d) e (f) uma distribuição de 2 *clusters* com 3 usuários/*cluster*. Em cada cenário da Figura 5.9 são mostrados os *clusters*, as distâncias dos usuários a BS e o número de antenas nas BSs igual ao número de *clusters*.

## 5.2.1 Políticas Propostas ERPA $\times$ ICR em Cenários Realistas MIMO-NOMA

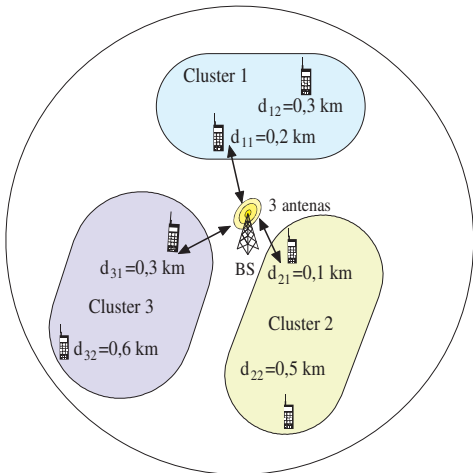
Demonstramos que existe realmente um problema de otimização para obter a mesma taxa para todos os usuários. Assumimos os cenários 1 e 2 com  $\alpha = 2$ . Figura 5.10.(a) e (c) mostram o ERPA (barras à esquerda) e a soma de ICR (barras à direita), considerando seis usuários com taxas fixas iguais a 1, 2, 3 e 4 bits/s/Hz para 3 *clusters* com 2 usuários/*cluster*. Figura 5.10.(b) e (d) mostram as taxas 0,5; 1; 1,5; 2 e 2,5 bits/s/Hz para 2 *clusters* com 3 usuários/*cluster*. A estratégia ERPA distribuída é comparada com a obtida pela Eq. (3.18) com a mesma potência total. Observamos que apenas o sistema ERPA manteve as mesmas taxas para todos os usuários ao custo de piores somas de taxas. Na estratégia ICR, os demais usuários recebem uma proporção maior de potência que causa maior interferência em outros usuários e consequentemente reduz as taxas dos outros usuários e aumenta sua própria taxa. O usuário mais forte (mais próximo) perde a taxa



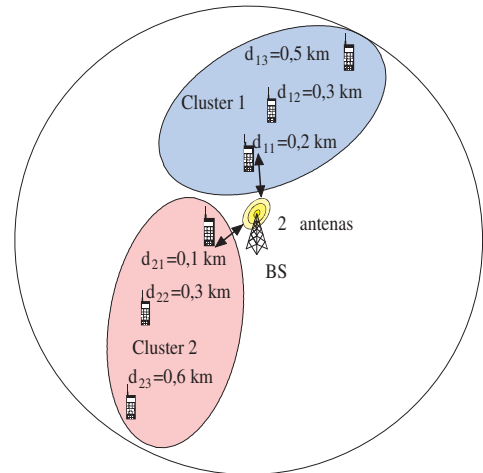
(a) Cenário 1 - 2 usuários/*cluster*.



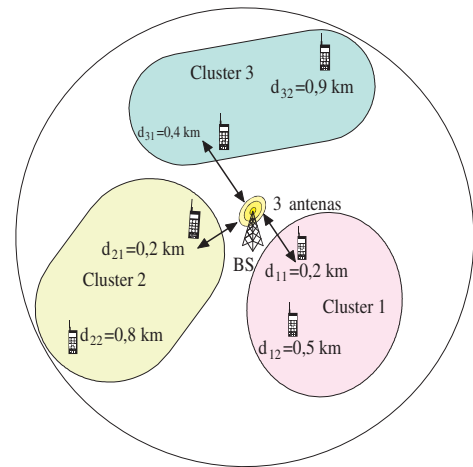
(b) Cenário 1 - 3 usuários/*cluster*.



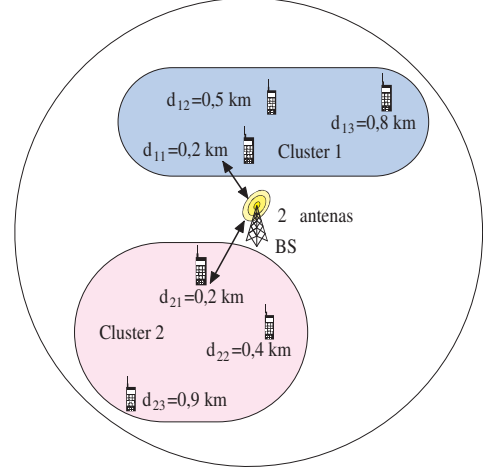
(c) Cenário 2 - 2 usuários/*cluster*.



(d) Cenário 2 - 3 usuários/*cluster*.



(e) Cenário 3 - 2 usuários/*cluster*.



(f) Cenário 3 - 3 usuários/*cluster*.

Figura 5.9 – Distribuição de usuários em um sistema NOMA com 6 usuários conforme Tabela 5.4.

Tabela 5.4 – Parâmetros Adotados para os cenários de Simulação.

Parameter	Value
#Usuários	$M \in \{4; 24\}$
Expoente de Perda de Percurso	$\alpha = [2, 0; 3, 0; 3, 5]$
Desvanecimento (termo curto)	$q_{a,k} \in \mathcal{CN}(0, 1)$
Potência do Ruído Branco	$\sigma_n^2 = 0, 1 \mu \text{ W}$
Orçamento de Potência	$P = 100 \text{ W}$
# clusters	$A \in \{2; 12\}$
usuários/cluster	$K \in \{2, 3, 4\}$
# antenas por usuário	$U \geq A$ , com $U \in \{2; 12\}$ .
Meta de Igual Taxa de Usuário Normalizada	$R^* \in [0, 1; 0, 5; 1; 1, 5; 2; 2, 5; 3; 3, 5; 4]$ [bits/s/Hz]
Potência da Circuitaria	$P_{\text{BS}} = 1 \text{ W/antena BS}$ $P_{\text{K}} = 100 \text{ mW/antena usuário}$
Ineficiência do Amplificador de Potência RF	$\varrho = 1, 4$
# Realizações MCS	$10^4$
<b>Cenários: Distribuição do usuário dentro da célula: 6 usuários, 2 ou (3) clusters</b>	
Cen. 1 – $\sum_{a=1}^A \sum_{k=1}^K d_{ak} = 1, 0$	$d_{11(11)} = 0, 08$ , $d_{12(12)} = 0, 15$ , $d_{22(13)} = 0, 28$ , $d_{21(21)} = 0, 07$ , $d_{31(22)} = 0, 15$ , $d_{32(23)} = 0, 27$ [km]
Cen. 2 – $\sum_{a=1}^A \sum_{k=1}^K d_{ak} = 2, 0$	$d_{11(11)} = 0, 20$ , $d_{12(12)} = 0, 30$ , $d_{22(13)} = 0, 50$ , $d_{21(21)} = 0, 10$ , $d_{31(22)} = 0, 30$ , $d_{32(23)} = 0, 60$ [km]
Cen. 3 – $\sum_{a=1}^A \sum_{k=1}^K d_{ak} = 3, 0$	$d_{11(11)} = 0, 20$ , $d_{12(12)} = 0, 50$ , $d_{22(13)} = 0, 80$ , $d_{21(21)} = 0, 20$ , $d_{31(22)} = 0, 40$ , $d_{32(23)} = 0, 90$ [km]
Scen. 4 – $\mathfrak{R}_D$ (distrib. uniform.)	0, 05, 0, 10, 0, 20 [km]
<b>Estratégias de Alocação de Potência</b>	
Inverso da Resposta do Canal (ICR)	$P_{ij}$ da Eq. (3.18)
Alocação de Potência Ótima para Taxa Igual (ERPA)	Subseção 3.1.1.3

parcialmente para garantir a equidade, enquanto os usuários distantes ganham taxa ao alocar mais energia para eles. Esse mecanismo aconteceu para todas as taxas analisadas e esse é o preço que se paga para ter equidade entre os usuários. De fato, para manter uma taxa de  $R = 2$  bits/s/Hz para todos os usuários no cenário 1 com 3 clusters com 2 usuários/cluster e  $\alpha = 2$ , a potência total necessária é 0,08 W, como mostrado na Figura 5.10.(a). Para o mesmo caso, a Figura 5.10.(b) mostra a potência total necessária de 1,42 W para 2 clusters com 3 usuários/cluster. Assim, a potência total precisa ser 17,75 vezes maior para 2 clusters com 3 usuários/cluster, porque há maior diversidade ou capacidade de multiplexação no caso de 3 clusters com 2 usuários/cluster, além do fato de que a interferência em cada cluster é menor.

A Figura 5.11 mostra a soma das taxas para a distribuição ERPA e ICR das potências totais, nos cenários 1, 2 e 3, para  $\alpha = 2$  e 3. As somas das taxas de ICR  $R_{\Sigma}$  são maiores do que as distribuições ERPA para a mesma potência total. Justifica-se pelo fato de que as distribuições do ERPA são mais restritivas do que as do ICR. Essa diferença

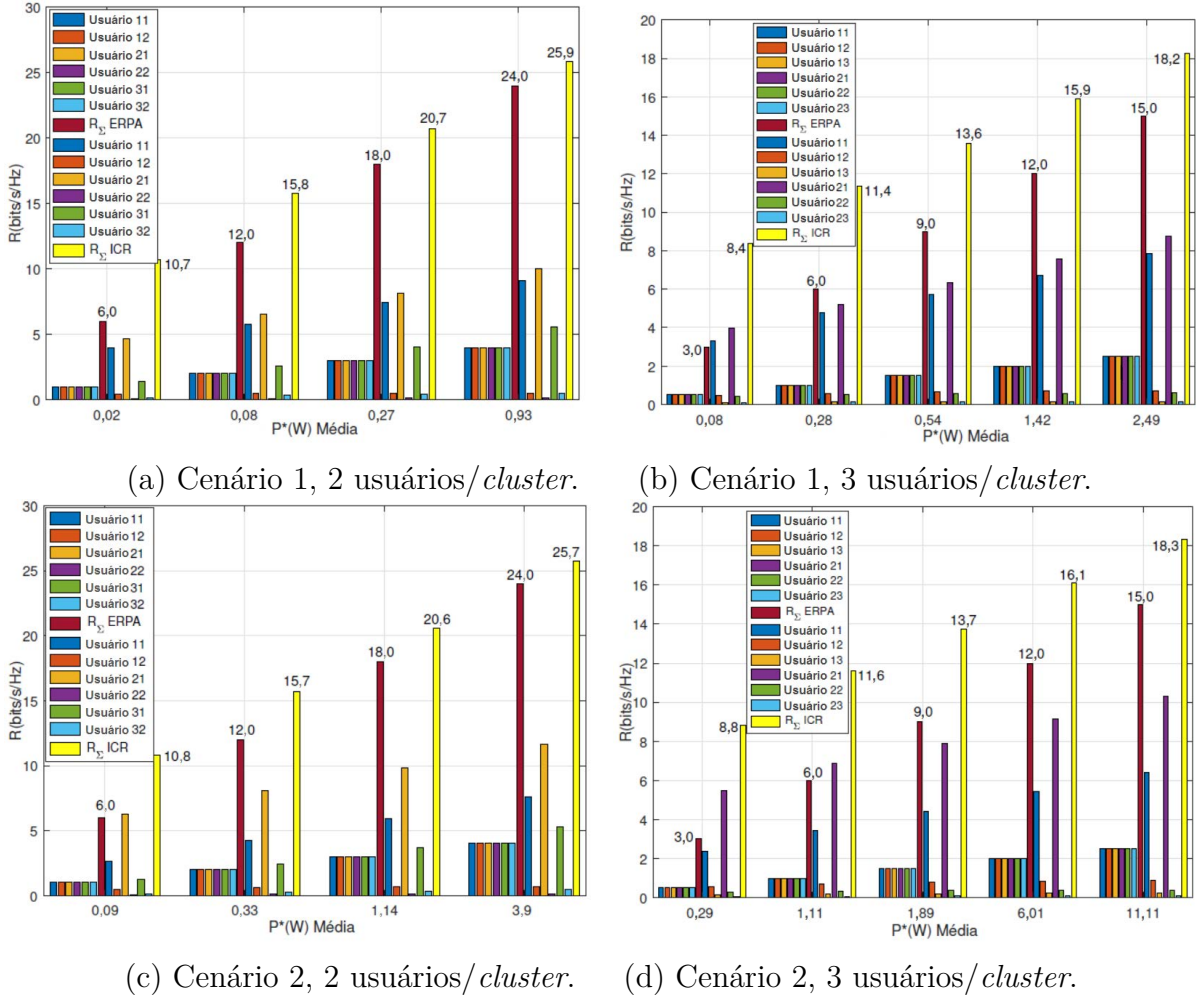
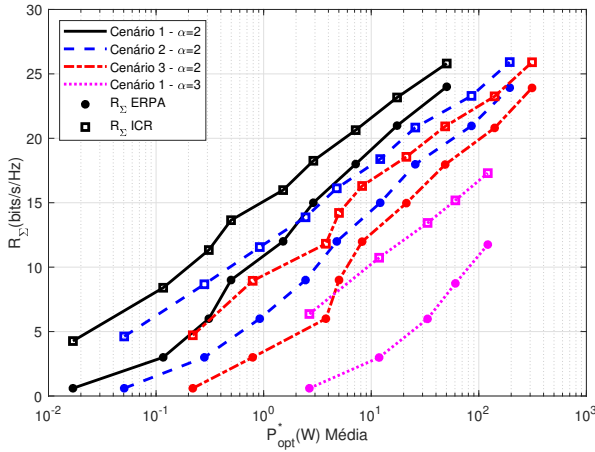
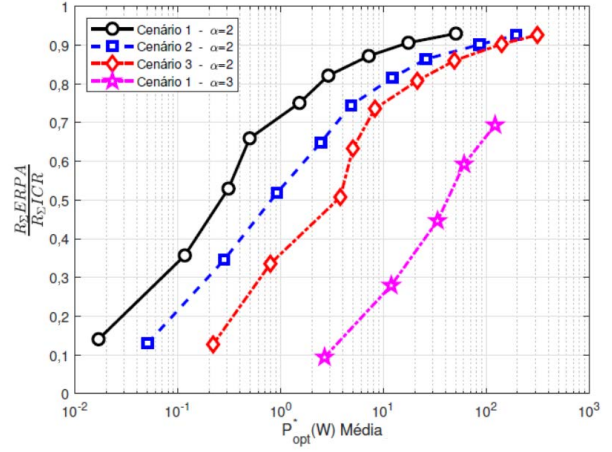


Figura 5.10 – As taxas de bits *versus* potência alocada considerando perda de percurso LoS ( $\alpha = 2$ ) e 6 usuários. As taxas de ICR foram obtidas com a proporção de potência dada pela Eq. (3.18) e taxas ERPA pelo problema (3.19). (a) Cenário 1, Figura 5.9.(a); (b) Cenário 1, Figura 5.9.(b); (c) Cenário 2, Figura 5.9.(c); e (d) Cenário 2, Figura 5.9.(d).

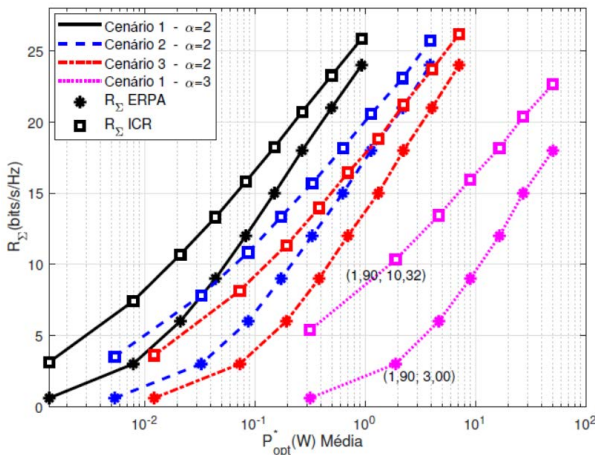
diminui à medida que a potência aumenta, porque a potência não é mais um fator restritivo. Há maior interferência dentro do *cluster* para sistemas com 3 usuários/*cluster*, Figura 5.11.(a), portanto, requer alta potência para atingir a mesma taxa que sistemas com 2 usuários/*cluster*, Figura 5.11.(c). Por exemplo, no ambiente com  $\alpha = 3$  e cenário 1, a soma das taxas de distribuição ICR com um orçamento de potência média de  $P_{opt}^* \approx 1,90$  W é  $R_{\Sigma} = 10,32$  bits/s/Hz, enquanto a soma das taxas de distribuição do ERPA é  $R_{\Sigma} = 3,00$  bits/s/Hz para o mesmo orçamento de potência de 1,90 W para 2 usuários/*cluster*, Figura 5.11.(c). A razão entre a soma das taxas de ERPA e ICR aumenta à medida que a soma das taxas aumentam, conforme mostrado na Figura 5.11.(b) e (d). As inclinações das curvas são semelhantes. A proporção entre a soma das taxas do ERPA e do ICR é próxima a noventa por cento para altas potências. Os cenários mais restritivos precisam de mais potência para atingir este nível. Para 3 usuários/*cluster*, a potência necessária para estabilização é maior porque há maior interferência entre os usuários no *cluster*.



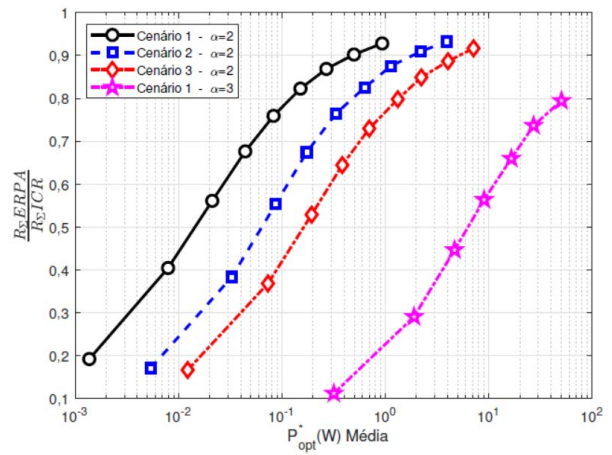
(a) 3 usuários/*cluster*, Figura 5.9 (b), (d) e (f).



(b) 3 usuários/*cluster*.



(c) 2 usuários/*cluster*, Figura 5.9 (a), (c) e (e).

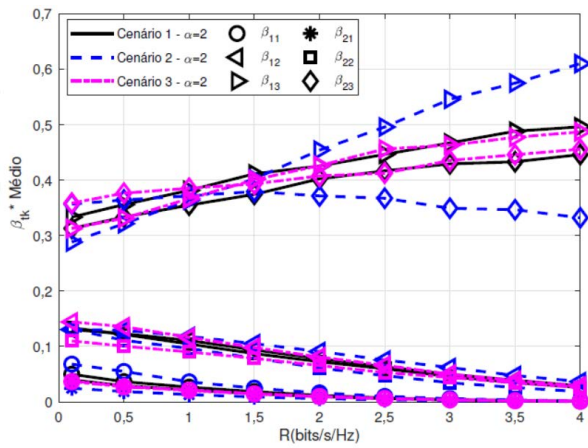


(d) 2 usuários/*cluster*

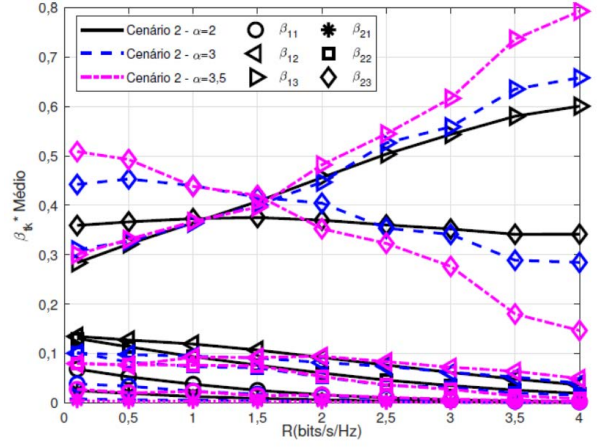
Figura 5.11 –  $R_{\Sigma}$  versus  $P$  and  $\frac{R_{\Sigma}ERPA}{R_{\Sigma}ICR}$  versus  $P_{opt}^*$  to  $\alpha = 2$  and 3. A proporção de potência foi obtida pela distribuição ICR, Eq. (3.18), e a distribuição ERPA pelo problema (3.19).

A Figura 5.12 mostra a variação das proporções de potência dos usuários  $\beta_{opt}^*$  em função da mesma taxa para os seis usuários dada por  $R$  em bits/s/Hz. A Figura 5.12.(a) e (c) mostra os resultados para  $\alpha = 2$  na configuração de 3 usuários/*cluster* e 2 usuários/*cluster*, respectivamente, nos cenários 1, 2 e 3. Figura 5.12.(b) e (d) mostra os resultados do cenário 2 para  $\alpha = 2, 3$  e 3,5 na configuração de 3 usuários/*cluster* e 2 usuários/*cluster*, respectivamente.

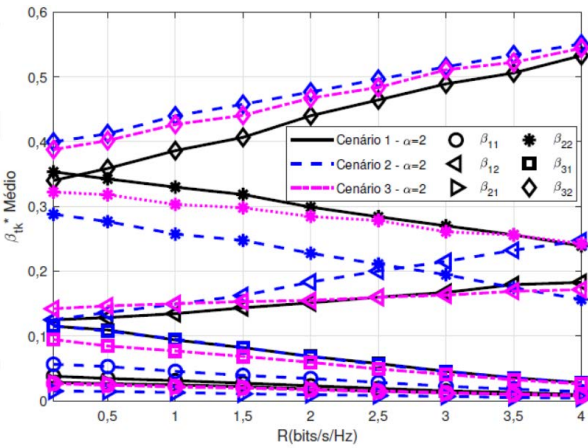
Nos cenários 1 e 3, a distribuição de potência entre os usuários assume valores próximos porque os usuários têm uma distribuição com similar proporção de distâncias na célula, Figura 5.12.(a). No *cluster* 1 do cenário 2, os usuários têm valores marginais de perda de percurso e, conseqüentemente, ganho de canal próximo. Como os ganhos estão próximos, o usuário com o menor ganho de canal tem um aumento na proporção de potência, taxa de dados e potência total. Valores próximos da proporção de potência total causam mais interferência entre os usuários. No *cluster* 2 do cenário 2, as proporções das



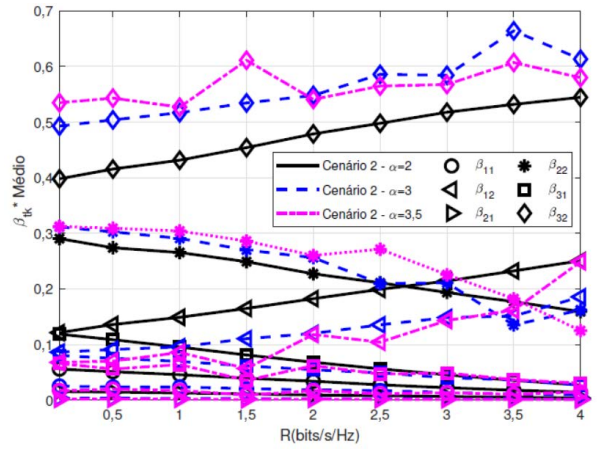
(a) 3 usuários/*cluster*, (Figura 5.9 (b), (d) e (f)).



(b) 3 usuários/*cluster*.



(c) 2 usuários/*cluster*, (Figura 5.9.(a), (c) e (e)).



(d) 2 usuários/*cluster*.

Figura 5.12 –  $\beta_{ak}^*$  versus taxa de dados, problema (3.19).

distâncias entre os usuários são maiores do que nos cenários 1 e 3 e no *cluster* 1 do cenário 2. Isso leva a um aumento na proporção de potência para o usuário com a maior perda de percurso à medida que a taxa de dados e a energia total dos usuários aumentam. Pode ser justificado pela necessidade do usuário mais crítico ter mais potência para manter a mesma taxa de dados, devido à alta interferência de usuários do mesmo *cluster* na detecção do sinal.

Na Figura 5.12.(a) e (b), há maior interferência entre os usuários dentro do *cluster*, para a situação de 3 usuários/*cluster*, o que resulta em uma proporção mais pronunciada de potência para usuários com menor ganho de canal. Essa proporção de energia aumenta a taxa de dados. Esse processo ocorre ao contrário para usuários com maior ganho de canal. Portanto, há pouca variação nas proporções de potência com o aumento da taxa de usuários neste *cluster*. Os usuários do *cluster* 2 no cenário 2 têm distâncias proporcionalmente maiores entre eles.

A proporção de potência para o usuário com maior perda de percurso  $\beta_{23}$  resulta em um valor maior da proporção de potência para valores baixos de taxa de dados  $R$ ,

quando comparado com o usuário que possui a maior perda de percurso do *cluster* 1 no cenário 2 dado por  $\beta_{13}$ , Figura 5.12.(b). A maior proporção de potência para o usuário mais distante predomina neste caso. O usuário com a maior perda de percurso no *cluster* 1 dada por  $\beta_{13}$  terá uma maior proporção de potência associada ao aumento da taxa de dados  $R$ . O predomínio de interferências sofridas pelo usuário com maior perda de percurso é causado por usuários com distâncias menores. Os valores da proporção de potência total  $\beta_{23}$  são maiores para valores baixos de  $R$  devido à predominância de usuários com maior expoente de perda de percurso  $\alpha$ . Os valores de  $\beta_{13}$  aumentam com o incremento de  $R$  devido ao predomínio da interferência do usuário com o ganho do canal de valores próximos.

Na Figura 5.12.(c), a proporção de potência total  $\beta_{22}$  diminui quando a taxa de dados  $R$  aumenta em *clusters* com distâncias maiores. Em contraste, as proporções de potência total  $\beta_{12}$  e  $\beta_{32}$  aumentam quando a taxa de dados  $R$  aumenta em *clusters* com distâncias mais curtas. Há um predomínio de interferência entre usuários de um mesmo *cluster* com distâncias menores. Na Figura 5.12.(d), aumentamos o expoente de perda do percurso  $\alpha$ , e o comportamento foi semelhante ao da Figura 5.12.(b) e (c). A curva moveu-se para cima devido ao aumento na taxa de dados e aos maiores valores de proporção de potência total.

## 5.2.2 Equidade de Taxa em Sistemas MIMO-NOMA

A equidade foi calculada pela Eq. (3.17). Uma vez que garantimos que todos os usuários têm a mesma taxa  $R$ , a equidade máxima resultou em  $\mathcal{F}_j = 1$ . Na Figura 5.10, os valores de potência total  $P$ , bem como a proporção de potência  $\beta$  para cada usuário na Figura 5.12 foram obtidos a fim de garantir o mesma taxa para todos os usuários e, portanto, a equidade máxima. Percebe-se que os valores das taxas dos usuários são realmente as mesmas, o que mostra a precisão do processo de otimização.

## 5.2.3 Ponto de Operação Ótimo para Eficiência Energética Sob Diferentes Cenários

Os pontos ERPA, Eq. (3.25), para EE são mostrados na Figura 5.13 para três cenários de valor fixo para  $\alpha = 2$ . Os pontos ERPA, Eq. (3.23) da Figura 5.13 estão resumidos na Tabela 5.5. Para o cenário 2, as taxas de  $R = 19,30$  e  $R = 11,30$  bits/s/Hz foram obtidas para as potências totais de  $P = 1,21$  e  $P = 1,92$  W, o que resultaram EE= 2,99 e EE= 1,91 bits/J/Hz com 2 usuários/*cluster* e 3 usuários/*cluster*, respectivamente. Isso é aproximadamente três vezes menos energia para 1,56 vezes mais taxa e EE para 2 usuários/*cluster*. Na mesma análise, os resultados do cenário 3 mostram uma redução de aproximadamente vinte por cento em todos os parâmetros, mas com a mesma proporcionalidade na comparação para 2 e 3 usuários/*cluster*. No cenário 1, obtemos taxas

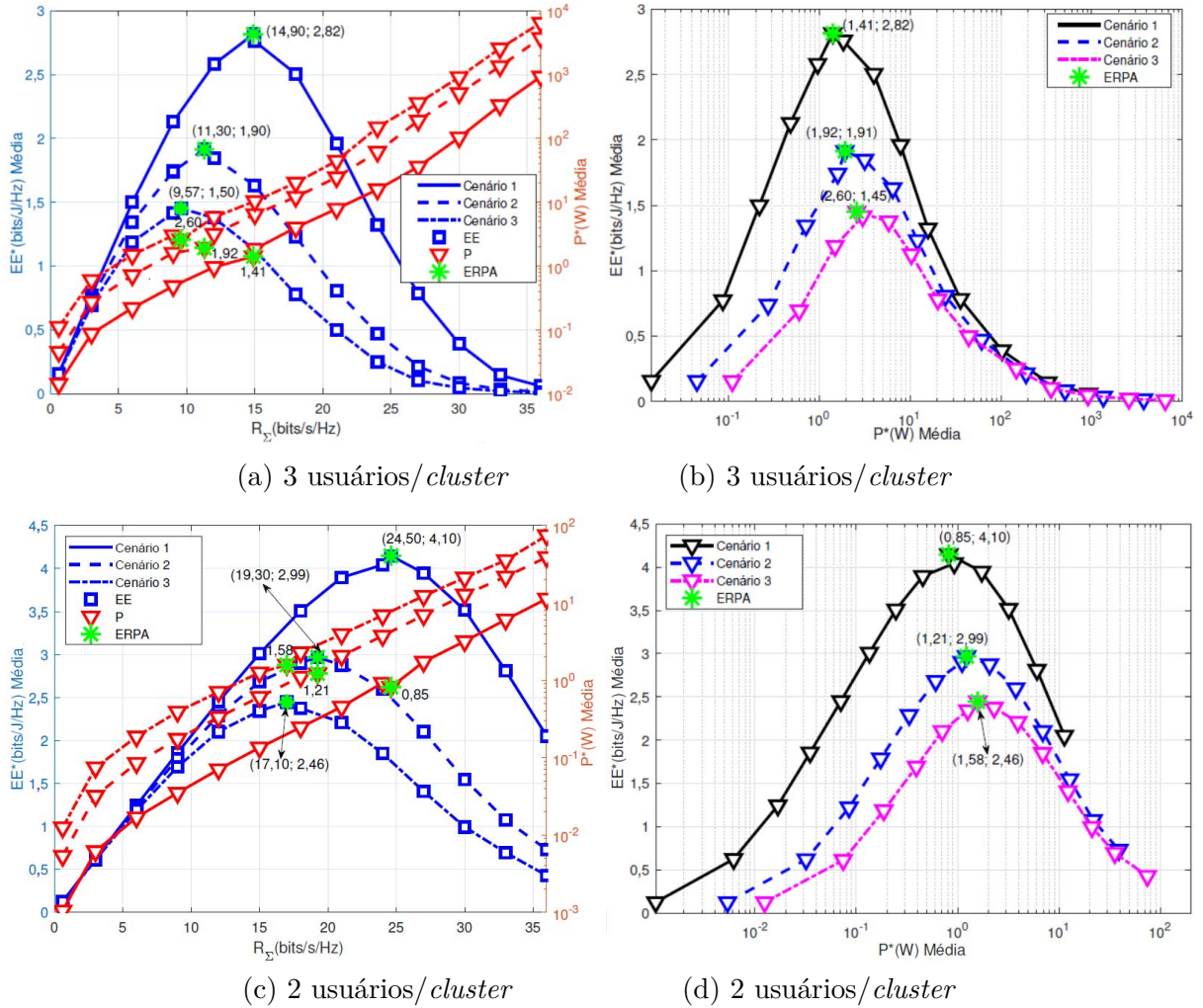


Figura 5.13 –  $EE^*_{opt} \times R_{\Sigma}$ ,  $P^*_{opt} \times R_{\Sigma}$  and  $EE^*_{opt} \times P^*_{opt}$ , problema (3.25) para  $\alpha = 2$  e 6 usuários. A distribuição ERPA (solução para o problema (3.23)) foi destacada com marcadores de estrela verde.

de  $R = 24, 50$  e  $14, 90$  bits/s/Hz e respectivas potências totais de  $P = 0, 85$  e  $1, 41$  W, o que resultaram em  $EE = 4, 10$  e  $2, 82$  bits/J/Hz com 2 usuários/cluster e 3 usuários/cluster, respectivamente. Isso é aproximadamente sessenta por cento menos energia para dobrar a taxa e EE para 2 usuários/cluster. Para usuários mais próximos no cenário 1, menos energia é necessária para atingir a taxa especificada, de forma que uma energia muito baixa atinge a taxa mais alta e, conseqüentemente, maior EE, como pode ser visto nos pontos no asterisco da Figura 5.13. O ponto EE ERPA Eq. (3.23) de usuários no cenário 1 é maior do que os usuários dos cenários 2 e 3 porque os usuários do cenário 1 estão mais próximos ou com menos perda de percurso e com maior evidência de potência total de acordo com a Figura 5.13.(b) e (d). Existem mais usuários interferentes em um mesmo cluster para a situação de 3 usuários/cluster, portanto, a taxa de dados e a eficiência energética são menores nestes casos, conforme mostrado na Tabela 5.5. EE e taxa de dados diminuem quando passamos dos cenários 1 para 2 e 3 porque nos últimos cenários os usuários têm menos ganho de canal. Ao aumentar o número de usuários por cluster de 2 para 3, há uma

Tabela 5.5 – Distribuição ERPA para  $\alpha = 2$  e seis usuários obtida de acordo com os gráficos da Figura 5.13.

Cenário	2 usuários por <i>cluster</i>			3 usuários por <i>cluster</i>		
	$EE_{\text{opt}}^*$	$P_{\text{opt}}^*$	$R_{\Sigma}^*$	$EE_{\text{opt}}^*$	$P_{\text{opt}}^*$	$R_{\Sigma}^*$
<b>1</b>	4,10	0,85	24,50	2,82	1,41	14,90
<b>2</b>	2,99	1,21	19,30	1,90	1,92	11,30
<b>3</b>	2,46	1,58	17,10	1,45	2,60	9,57

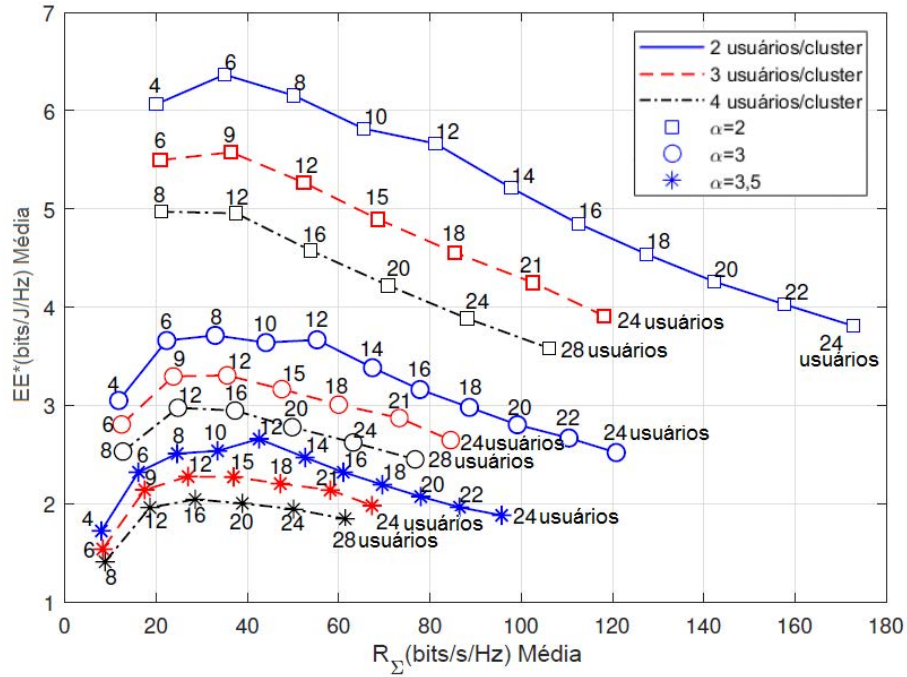
redução na distribuição do ERPA de EE e  $R$  e um aumento em  $P$ . A interferência entre os usuários aumenta com o aumento do número de usuários por *cluster*, conforme Tabela 5.5. Os pontos gerados pelo ERPA Eq. (3.23) são consistentes com a curva gerada pelo ERPA Eq. (3.25). Isso mostra a robustez dos resultados. Os resultados máximos de EE da Eq. (3.23) coincidem com o máximo da Eq. (3.25). Isso nos permite usar Eq. (3.23) para obter os pontos máximos de EE nos gráficos a seguir.

### 5.2.4 Ponto Ótimo de Eficiência de Recursos em Função do Número de *Clusters* e Usuários Uniformemente Distribuídos

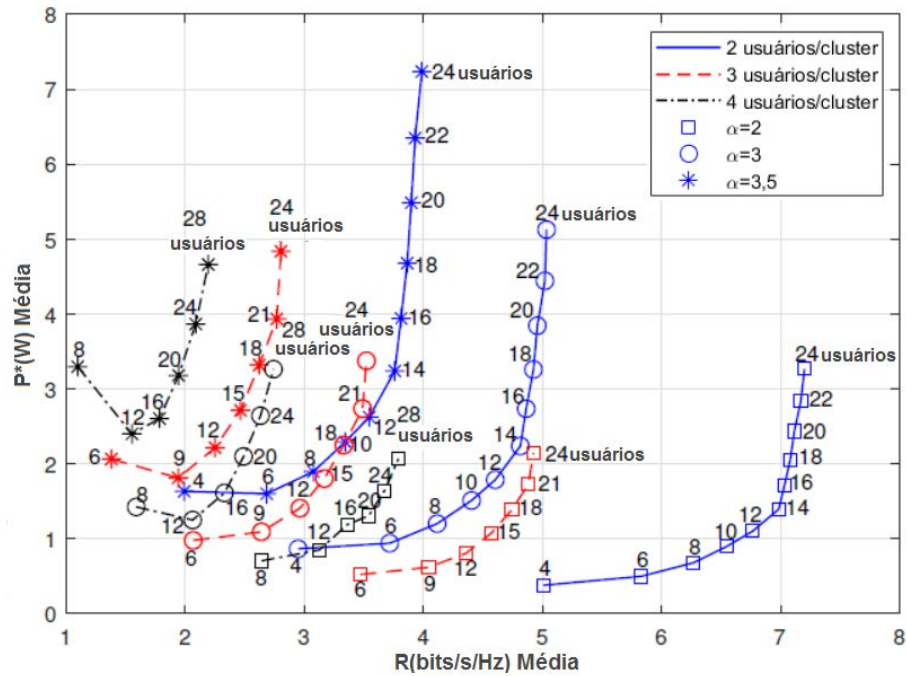
Nesta seção, a *eficiência de recursos* RE, ou seja, o compromisso entre  $EE \times R_{\Sigma}$  ótimo é analisado em termos do número de *clusters*  $A$ . O problema de otimização foi resolvido para obter cada ponto no gráfico e os pontos foram interligados para identificar os pontos com as mesmas características de simulação. Adotamos  $U = A$ , ou seja, o número de antenas de cada usuário e na BS é igual ao número de *clusters*. Consideramos  $10^4$  realizações MCS para cada número de usuários  $M$ . Para definir a combinação de usuários para o NOMA, o ganho de canal de usuários de  $KA$  foi ordenado de forma decrescente. Consideramos o cenário no qual as distâncias mudam aleatoriamente com distribuição uniforme obtida pelo cenário 4. Os parâmetros e cenários considerados estão na Tabela 5.4.

Os usuários pertencentes a cada *cluster* foram definidos de acordo com a Tabela 5.6, com base na referência (ELGHARIANI; ZOLTOWSKI, 2016). Por exemplo, para  $K = 2$  usuários/*cluster* e  $A = 3$  *clusters*, resulta em 6 usuários ordenados em ordem decrescente de ganho de canal, de modo que temos a combinação de usuários (1, 4), (2, 5) e (3, 6) com os *clusters* 1, 2 e 3, respectivamente.

Nas Figuras 5.14, 5.15, 5.16, consideramos o raio da célula  $\mathfrak{R}_D = 50$  m,  $\mathfrak{R}_D = 100$  m e  $\mathfrak{R}_D = 200$  m, respectivamente. Os valores ótimos para EE, soma das taxas e potência total foram obtidos resolvendo o problema de otimização proposto pela Eq. (3.23). Neste caso, EE, soma das taxas  $R_{\Sigma}$  e potência total  $P$  foram consideradas como variáveis no processo de otimização para garantir a taxa igual para todos os usuários NOMA. Nesta configuração, todos os usuários foram uniformemente distribuídos (DEVROYE, 1986)



(a)  $EE_{opt}^* \times R_\Sigma$



(b)  $P^* \times R$  (fairness)

Figura 5.14 –  $EE \times R_\Sigma$  e  $P \times R$  (equidade) para 2, 3 e 4 usuários/*cluster*,  $\alpha = 2, 3$  e 3,5 e  $\mathfrak{R}_D = 50$  m, problema (3.23).

Tabela 5.6 – Combinação de  $K$  usuários por *cluster* com  $A$  *clusters* que resulta em  $KA$  usuários em ordem decrescente de ganho de canal.

Usuários/ <i>cluster</i>	<i>Cluster</i>				
	# 1	# 2	# 3	...	# A
1	1	2	3	...	A
2	A + 1	A + 2	A + 3	...	2A
⋮	⋮	⋮	⋮	⋮	⋮
k	(k - 1)A + 1	(k - 1)A + 2	(k - 1)A + 3	...	kA
⋮	⋮	⋮	⋮	⋮	⋮
K	(K - 1)A + 1	(K - 1)A + 2	(K - 1)A + 3	...	KA

dentro da célula de raio  $\mathfrak{R}_D$ , enquanto o consumo de energia do circuito adotado foi  $P_{BS} = 1$  W por antena de BS e  $P_K = 100$  mW por antena de usuário.

O problema de otimização (3.23) foi simulado para um número total de usuários na célula  $M$ , número de usuários por *cluster*  $K$ , raio da célula  $\mathfrak{R}_D$  e expoente de perda de percurso  $\alpha$  para valores diferentes para o cenário 4 e os parâmetros da Tabela 5.4. Desse modo, o EE máximo e o correspondente  $R_{\Sigma}$  foram obtidos para a equidade máxima em cada ponto dos gráficos nas Figuras 5.14.(a), 5.15.(a) e 5.16.(a). As respectivas potências  $P$  e os  $R$  correspondentes de equidade ou a taxa de dados para cada usuário em cada ponto dos gráficos estão nas Figuras 5.14.(b), 5.15.(b) e 5.16.(b).

O valor máximo de EE para  $\alpha = 2$  e 2 usuários/*cluster* resultou em 6 usuários na célula, para 3 usuários/*cluster* em 9 usuários na célula e para 4 usuários/*cluster* em 12 usuários na célula, o que equivale a 3 *clusters* na célula nos três casos. As curvas obtidas para  $\alpha = 2$  e 2, 3 e 4 usuários/*cluster*, nas Figuras 5.14.(b), 5.15.(b) e 5.16.(b), corroboram este resultado, pois os valores de potência aumentam exponencialmente, enquanto equidade  $R$  aumenta marginalmente para mais de 6, 9 e 12 usuários na célula nos três casos.

Da mesma forma para  $\alpha = 3$  e  $\mathfrak{R}_D = 50$  m, o valor máximo de EE resultou 8 usuários na célula com 4 *clusters* de 2 usuários/*cluster* e 9 e 12 usuários na célula com 3 *clusters* de 3 e 4 usuários/*cluster* respectivamente, conforme mostrado na Figura 5.14.(a). Como podemos ver na Figura 5.14.(b), a potência aumentou exponencialmente para 8 usuários na célula com 2 usuários/*cluster*, para 9 usuários na célula com 3 usuários/*cluster* e para 12 usuários na célula com 4 usuários/*cluster*, enquanto a equidade  $R$  aumentou marginalmente. Portanto, o EE resultou em valores menores após esses valores.

Na Figura 5.15.(a), EE máximo e  $\mathfrak{R}_D = 100$  m resultaram em 12 usuários na célula correspondente a 6 *clusters* de 2 usuários/*cluster*, 12 usuários na célula correspondente a 4 *clusters* de 3 usuários/*cluster* e 16 usuários na célula correspondendo a 4 *clusters* de 4 usuários/*cluster*. Além disso, na Fig 5.16.(a), EE máximo e  $\mathfrak{R}_D = 200$  m também resultou em 12 usuários na célula correspondente a 6 *clusters* de 2 usuários/*cluster*, mas 18 usuários na célula correspondente a 6 *clusters* de 3 usuários/*cluster* e 24 usuários na

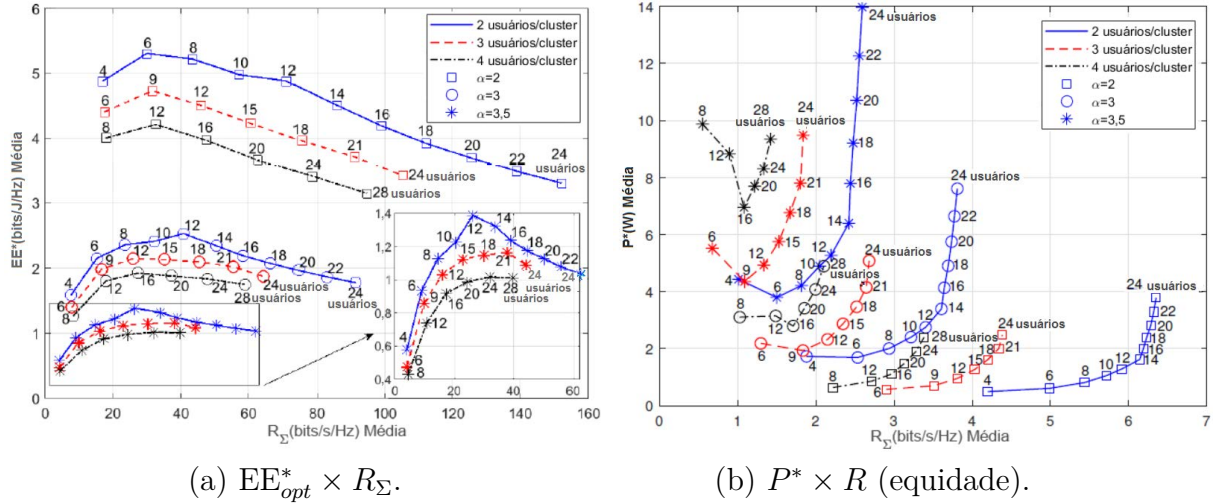


Figura 5.15 –  $EE \times R_{\Sigma}$  e  $P \times R$  (equidade) para 2, 3 e 4 usuários/*cluster*,  $\alpha = 2, 3$  e 3, 5 e  $\mathfrak{R}_D = 100m$ , problema (3.23).

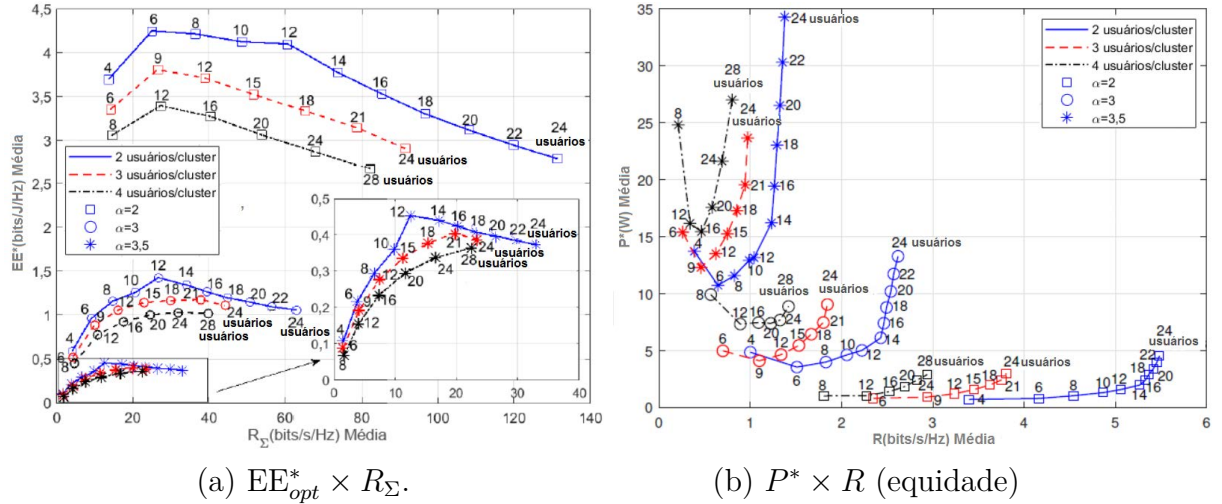


Figura 5.16 –  $EE \times R_{\Sigma}$  e  $P \times R$  (equidade) para 2, 3 e 4 usuários/*cluster*,  $\alpha = 2, 3$  e 3, 5 e  $\mathfrak{R}_D = 200m$ , problema (3.23).

célula correspondente a 6 *clusters* de 4 usuários/*cluster*. Nas Figuras 5.15.(b) e 5.16.(b), observamos um crescimento exponencial para a potência total e um aumento marginal para a equidade de  $R$  a partir do número de usuários que atingiram máximo EE. Isso é consistente com os resultados apresentados anteriormente nas Figuras 5.15.(a) e 5.16.(a), e mostra que o EE máximo com equidade máxima para 2, 3 e 4 usuários/*cluster* depende do raio da célula e do expoente de perda de percurso.

Finalmente, o número de usuários na célula para EE máximo e  $\alpha = 3, 5$  resultou em  $M = 12$  ou 6 *clusters* na célula com 2 usuários/*cluster* para os raios das três células em análise nas Figuras 5.14.(a), 5.15.(a) e 5.16.(a). Na Figura 5.14.(a), para 3 e 4 usuários/*cluster*, resultou em 12 e 16 usuários em 4 *clusters* na célula em ambos os casos, respectivamente. Para 3 usuários/*cluster*, o EE máximo obtido foi de 21 usuários em 7 *clusters* com raio de  $\mathfrak{R}_D = 100$  e  $\mathfrak{R}_D = 200$  m de acordo com 5.15.(a) e 5.16.(a) em análise.

Os resultados para 4 usuários/*cluster* foram 24 usuários em 6 *clusters* na célula com raio  $\mathfrak{R}_D = 100$  m de acordo com a Figura 5.15.(a). Para o raio celular de  $\mathfrak{R}_D = 200$  m, resultou em 28 usuários em 7 *clusters* na célula de acordo com a Figura 5.16.(a). Os resultados mostrados nas Figuras 5.14.(b), 5.15.(b) e 5.16.(b) confirmam os valores exibidos.

### 5.3 Resultados Numéricos para o MIMO-NOMA com Otimização Biobjetiva

Os valores dos parâmetros usados nas simulações são definidos na Tabela 5.7. Consideramos os cenários 1, 2, e 3 no qual os usuários estão uniformemente distribuídos dentro da área do anel de raio  $r_0$  e  $r_D$  [metros], Figura 3.1.

Tabela 5.7 – Parâmetros Adotados Para os Cenários de Simulação.

Parâmetros	Valores
#Usuários	$M \in \{4; 40\}$
Expoente de Perda de Percurso	$\alpha = [2, 0; 3, 0; 3, 5]$
Desvanecimento (termo curto)	$q_{a,k} \in \mathcal{CN}(0, 1)$
Potência do Ruído Branco	$\sigma_n^2 = 0.1 \mu \text{ W}$
Orçamento de Potência	$P = 100 \text{ W}$
# <i>clusters</i>	$A \in \{2; 12\}$
usários/ <i>cluster</i>	$K \in \{2, 3, 4\}$
# antenas por usuário	$U \geq A$ , com $U \in \{2; 12\}$ .
Meta para Igual Taxa de Usuário Normalizada	$R^* \in [0, 1; 0, 5 : 0, 5 : 4, 0]$ [bits/s/Hz]
Potência de Circuitaria	$P_{\text{BS}} = 1 \text{ W/antena (BS)}$ $P_K = 100 \text{ mW/antena por usuário}$
Ineficiência do Amplificador de Potência RF	$\rho = 1, 4$
<b>Métodos de Otimização: Parâmetros</b>	
Limites $\epsilon$ -C	$SE(P, \beta_{a,k}, R) \leq \epsilon \leq \max SE$
Heurística Ambiciosa <i>Hill Climbing</i>	$\rho = 10^8$ (penalidade); Max-Iteração= $10^4$ $\bar{\alpha}_k = 0, 0008$ (tamanho do passo)
<b>Cenários: Distribuição dos usuários dentro da célula <math>\sim \mathcal{U}[r_0, r_D]</math></b>	
Cenário 1	$r_0 = 10 \text{ [m]}$ e $r_D = 50 \text{ [m]}$
Cenário 2	$r_0 = 10 \text{ [m]}$ e $r_D = 100 \text{ [m]}$
Cenário 3	$r_0 = 10 \text{ [m]}$ e $r_D = 200 \text{ [m]}$
# Realizações MCS	$10^4$
<b>Estratégia de Alocação de Potência</b>	
Alocação de Potência para Igual Taxa (ERPA)	Subseção 3.1.1.3

#### 5.3.1 Detalhes de Implantação

Como o problema  $EE$  é fracionário e a função objetivo de  $SE$  é côncava (Seção 3.2.2), além dos resultados adequados do método SQP obtidos em um problema semelhante

(MARTINEZ; ABRÃO, 2020), pode-se esperar que a solução obtida no Passo 5 do Algoritmo 4 seja global. Porém, apenas um bom desempenho de DK ou SQP nos subproblemas gerados pelo método  $\epsilon$ -C não é suficiente para garantir as soluções de frente de Pareto. Portanto, para completar a análise numérica obtida com o método  $\epsilon$ -C, avaliamos o erro quadrático médio normalizado (NMSE - *Normalized Mean Squared Error*).

Do Teorema 1, pode-se medir a qualidade da solução obtida usando o Algoritmo 4, na  $k$ -ésima iteração, expressando o quão longe a solução atual ( $\epsilon$ ) é da solução ótima final  $SE(\mathbf{x}^*)$ , usando a solução  $\mathbf{x}^*$  determinada na etapa 4 do algoritmo. Esta qualidade da solução é quantificada pelo NMSE quando o equilíbrio é alcançado. Numericamente, a qualidade da solução obtida por um algoritmo de otimização EE já está definida por

$$\text{NMSE} = \mathbb{E} \left[ \frac{\|\text{SE}(\mathbf{x}^*) - \epsilon\|^2}{\|\epsilon\|^2} \right], \quad (5.2)$$

no qual  $\|\cdot\|^2$  denota a distância quadrática euclidiana e  $\mathbb{E}[\cdot]$  é o operador de esperança estatística.

### 5.3.1.1 SQP $\times$ DK na Etapa 5 do Algoritmo 3

Aplicamos os métodos SQP e DK para resolver o problema de otimização (3.29) na etapa 5 do Algoritmo 4. O método DK conforme apresentado no Algoritmo 2 usa um método de *penalidade quadrática* para resolver o problema em (3.23). Para implementar o método SQP, implementamos a função `fmincon` do Matlab. Além disso, o cálculo do gradiente foi implementado por diferenças finitas, enquanto o subproblema quadrático foi resolvido usando um método de pontos interiores (NOCEDAL; WRIGHT, 1996) aplicado a subprogramas quadráticos convexos.

**Critério de parada:** Para os métodos DK e SQP, a seguinte condição foi considerada como critério de parada:

$$\zeta = \|\mathbf{x}^k - \mathbf{x}^{k-1}\| < 10^{-6}, \quad \text{com } \mathbf{x}^k \text{ um ponto factível.}$$

**Subetapas DK  $\times$  SQP:** A escolha dos métodos DK e SQP para fins de comparação deve-se à simplicidade de definição e às boas propriedades teóricas de ambos os métodos, especificamente para DK em problemas de otimização fracionária onde o método converge para a solução geral. Esses métodos requerem a solução de uma sequência mais simples de problemas de PNL do que o original. Considerando a resolução sequencial do problema (3.29), aplicamos DK e SQP, de forma que o primeiro problema use multiplicadores nulos (SQP). Para a resolução dos demais subproblemas sequenciais, o ponto de partida é obtido resolvendo o problema anterior; além disso, no SQP, os multiplicadores iniciais e parâmetros de penalidade são assumidos como os multiplicadores e parâmetros de penalidade obtidos com a solução dos subproblemas anteriores. Finalmente, no método  $\epsilon$ -C,

conforme mencionado anteriormente, um desafio é determinar um intervalo apropriado para o parâmetro  $\epsilon$  conforme estabelecido no Teorema 1. Assim, devemos considerar  $\epsilon$  variando no intervalo  $\epsilon \in [\text{SE}(\overline{P, \beta_{a,k}, R}), \max \text{SE}]$ , tendo em vista obter o ponto no qual ocorre a máximo EE. Assim, a função objetivo EE foi mantida, enquanto SE foi definido como uma restrição, conforme descrito no Algoritmo 4.

### 5.3.2 Comparação do Compromisso entre EE-SE em Cenários Realísticos MIMO-NOMA

A comparação dos métodos  $\epsilon$ -C-DK e  $\epsilon$ -C-SQP e  $\epsilon$ -C-HC é amplamente discutida considerando os cenários 1, 2 e 3 com foco na análise de desempenho do compromisso entre EE-SE em sistemas MIMO-NOMA com equidade de taxa de dados e número variável de usuários por *cluster*.

Vamos considerar os valores dos parâmetros descritos anteriormente na Tabela 5.7 para diferentes  $r_D \in [50; 100; 200]$  [m] valores associados aos cenários 1 e 3, respectivamente, e diferentes expoentes de perda de percurso. O desempenho dos três algoritmos EE-SE-NOMA-MIMO é comparado na Tabela 5.8 em termos de tempo de execução (tempo), número médio percentual de sucesso de convergência (CS - *Convergence Success*) e NMSE para o algoritmo  $\epsilon$ -C, número de avaliações de função (FunC - *Function Evaluations*), e número distinto de soluções na Frente de Pareto (S-PF - *Solutions in Pareto Front*), considerando diferentes números de usuários e número de usuários por *cluster*.

Assim, para o compromisso entre EE-SE MIMO-NOMA, pode-se observar que o método DK gera subproblemas com um grau de complexidade semelhante ao problema abordado diretamente pelo método SQP; portanto, pode-se esperar um desempenho de tempo decorrido maior de  $\epsilon$ -C-DK em relação ao  $\epsilon$ -C-SQP. Essa complexidade computacional crescente varia de 100 vezes (1 *cluster* e 2 usuários) a 235 vezes (3 *clusters* e 6 usuários de configuração MIMO-NOMA). Além disso, foi observado que quando os cenários MIMO-NOMA se tornam mais complexos, por exemplo, 3 *clusters* e 6 usuários, o método  $\epsilon$ -C-DK não alcançou convergência total em todas as execuções. Porém, nas outras configurações menos complexas, quando a dimensão do problema de otimização é reduzida, o método  $\epsilon$ -C-DK sempre atinge convergência total em todas as execuções, como o método  $\epsilon$ -C-SQP. Como esperado pelo Teorema 1, a restrição  $C.\epsilon$  é satisfeita como igualdade em todos os casos avaliados, como  $\epsilon$ -C-SQP. É importante notar que o método baseado em busca local ambicioso HC resultou ineficiente na resolução de problemas de compensação BOO EE-SE para todos os três cenários de distribuição de usuários na célula e o número de usuários e configurações de *clusters* NOMA: para o mais complexo (três *clusters* e seis usuários), o aumento no tempo decorrido exigido pelo  $\epsilon$ -C-HC varia de duas a seis vezes

Tabela 5.8 – Figura de mérito para avaliação de desempenho dos algoritmos EE-SE MIMO-NOMA; expoente de perda de percurso  $\alpha = 3$ .

Método	$r_D$ [m]	CS [%]	Tempo [seg.]	NMSE	$R^*$ [bps/Hz]	FunC	S-PF
<b>1 cluster e 2 usuários</b>							
$\epsilon$ -C-DK	50	100	160,043556	9,0350e-07	2,8982	1,4011e+6	11
$\epsilon$ -C-SQP		100	1,2163	1,1234e-17	2,9740	8928	11
$\epsilon$ -C-HC		100	127,464792	0,0070	<b>3,4612</b>	4,0396e+6	11
$\epsilon$ -C-DK	100	100	109,003807	1,0804e-06	1,4479	6,7499e+5	11
$\epsilon$ -C-SQP		100	1,1529	2,9737e-17	1,8296	9046	11
$\epsilon$ -C-HC		100	132,431144	0,0132	<b>2,4335</b>	4,6132e+6	11
$\epsilon$ -C-DK	200	100	79,314762	4,7637e-06	0,5808	5,4086e+5	11
$\epsilon$ -C-SQP		100	1,1211	4,4042e-17	1,2096	9246	11
$\epsilon$ -C-HC		100	252,629958	0,0077	<b>2,0623</b>	6,5318e+6	11
<b>2 cluster and 4 usuários</b>							
$\epsilon$ -C-DK	50	100	367,431154	7,4973e-06	<b>3,7934</b>	2,9768e+6	11
$\epsilon$ -C-SQP		100	1,3553	1,3395e-17	3,7915	12174	11
$\epsilon$ -C-HC		100	630,490926	0,0101	3,5929	1,3880e+7	11
$\epsilon$ -C-DK	100	100	260,128037	1,4427e-06	2,0865	1,5796e+6	11
$\epsilon$ -C-SQP		100	1,3469	3,9909e-17	<b>2,4385</b>	12620	11
$\epsilon$ -C-HC		100	789,395572	0,0234	2,2856	2,0171e+7	11
$\epsilon$ -C-DK	200	100	146,104273	1,5602e-06	<b>1,6552</b>	1,2693e+6	11
$\epsilon$ -C-SQP		100	1,3934	7,2783e-17	1,4033	12898	11
$\epsilon$ -C-HC		100	1150,971378	0,0110	1,3273	2,4700e+7	11
<b>3 cluster and 6 usuários</b>							
$\epsilon$ -C-DK	50	100	545,3777	8,0740e-7	4,0584	1,9246e+6	11
$\epsilon$ -C-SQP		100	2,3179	5,4295e-17	4,3099	18058	11
$\epsilon$ -C-HC		100	1037,2966	0,0080	<b>4,7923<sup>†</sup></b>	1,2149e+7	11
$\epsilon$ -C-DK	100	100	285,8444	6,738e-7	3,0794	1,4711e+6	11
$\epsilon$ -C-SQP		100	2,3476	1,1659e-16	3,3178	20002	11
$\epsilon$ -C-HC		100	1661,6796	0,0145	<b>3,8713</b>	1,9571e+7	11
$\epsilon$ -C-DK	200	100	264,7762	9,0229e-6	1,7936	1,4498e+6	11
$\epsilon$ -C-SQP		100	2,3907	9,7081e-16	1,9476	20080	11
$\epsilon$ -C-HC		100	1383,3902	0,0140	<b>2,5699</b>	1,7240e+7	11
$\epsilon$ -C-DK	300	100	287,2231	0,0011	1,2122	1,5383e+6	11
$\epsilon$ -C-SQP		100	2,2659	3,9103e-16	1,2734	19172	11
$\epsilon$ -C-HC		100	1652,4219	0,0145	<b>1,8702</b>	1,2149e+7	11
$\epsilon$ -C-DK	400	100	532,9553	0,0334	1,0171	2,7164e+6	11
$\epsilon$ -C-SQP		100	2,4897	6,7735e-16	0,9118	19274	11
$\epsilon$ -C-HC		100	1123,8642	0,0140	<b>1,5050</b>	1,2417e+7	11

<sup>†</sup>Meta mais alta para igual taxa de usuário normalizada.

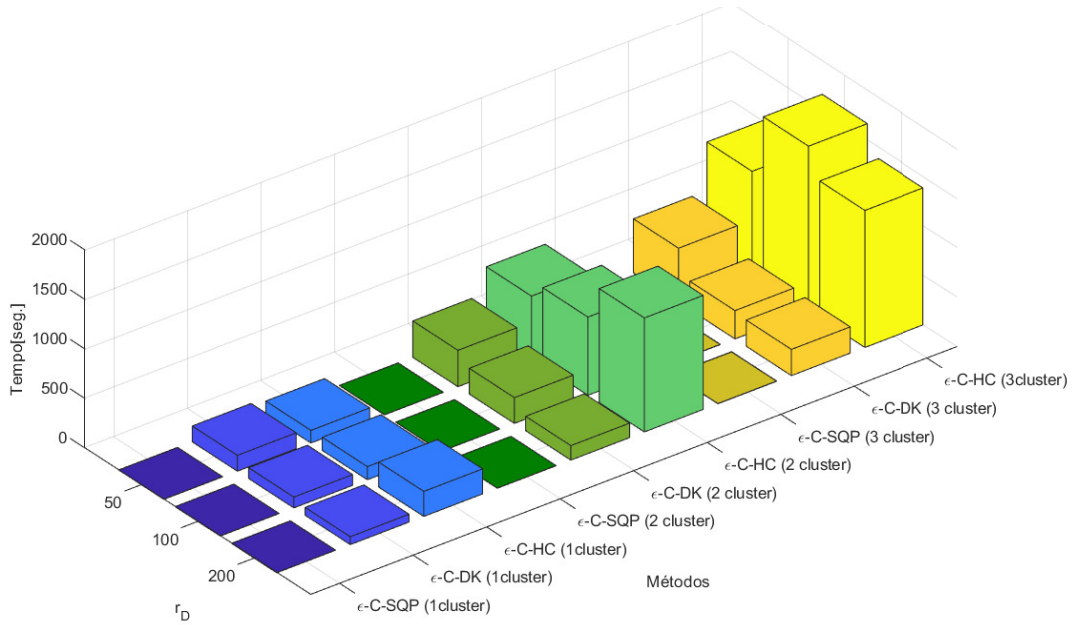
em relação ao método  $\epsilon$ -C-DK, e  $\approx 500$  vezes maior do que o tempo de execução necessário para o método  $\epsilon$ -C-SQP, enquanto a precisão NMSE do  $\epsilon$ -C-HC é substancialmente pior em relação ao  $\epsilon$ -C-SQP e método  $\epsilon$ -C-DK. Na verdade, a superioridade do método  $\epsilon$ -C-SQP pode ser confirmada comparando a complexidade (tempo decorrido) e NMSE dos três métodos propostos nas Figuras 5.17.(a) e 5.17.(b), respectivamente. Assim, SQP exige um número menor de FunC, o método de otimização  $\epsilon$ -C-SQP iterativo atingiu a melhor escolha em termos de complexidade e desempenho.

A Figura 5.18 mostra a frente de Pareto obtida com os três métodos BOO considerados nas soluções de compromisso entre EE-SE no sistema MIMO-NOMA. As soluções obtidas foram próximas às determinadas pelos métodos  $\epsilon$ -C-DK e  $\epsilon$ -C-SQP, o que reforça o bom desempenho de ambos os métodos propostos. É importante notar que para  $r_D = 50$  m, o método  $\epsilon$ -C-HC teve desempenho semelhante aos outros dois métodos iterativo-analíticos propostos; entretanto, para  $r_D = 100$  e  $200$  m uma pequena degradação nas soluções de frente de Pareto que podem ser observadas claramente. Esta degradação é relatada na Tabela 5.8, através dos valores de precisão NMSE obtidos por  $\epsilon$ -C-HC. Com isso, pode-se concluir que o método  $\epsilon$ -C-HC apresenta desempenho inferior em relação ao método  $\epsilon$ -C-DK e  $\epsilon$ -C-SQP para  $r_D = 100$  e  $200$  m, porém atinge o desempenho similar sob a condição de cobertura com  $r_D = 50$  m.

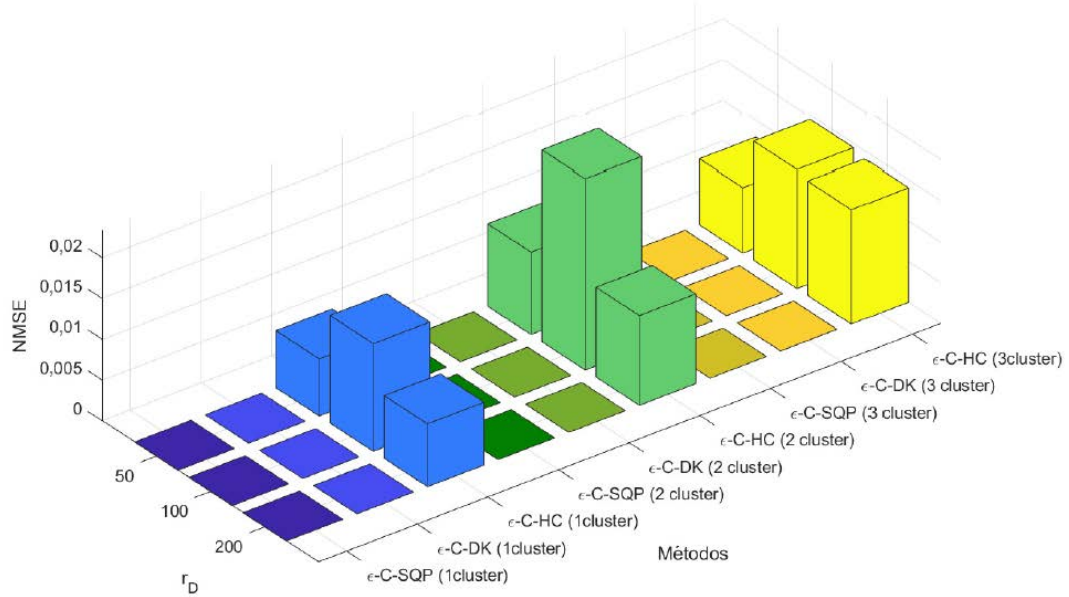
A Figura 5.19 evidencia a interdependência EE e SE, considerando os parâmetros de sistema adotados. Pode-se notar que a curva obtida está em forma de sino no domínio considerado. Além disso, as curvas revelam a natureza conflitante da otimização SE e EE simultâneas, ou seja, o SE máximo é atingível depois que o EE atinge seu nível máximo e diminui monotonicamente. Na verdade, os pontos que satisfazem a solução de compromisso bi-objetivo EE-SE podem ser identificados nos segmentos das curvas entre os pontos onde os máximos de EE e SE ocorrem, representando as soluções da frente de Pareto.

Figura 5.20 plota o SE *versus* EE e a potência total alocada ( $\Sigma P$ ) para o método de otimização EE-SE iterativo-analítico mais eficaz, o  $\epsilon$ -C-SQP. O EE associado na Figura 5.20 diminui com o aumento de SE, enquanto a potência correspondente aumenta para atingir a taxa de usuário igual  $R^*$  para todos os usuários em toda a área da célula. Isso ocorre porque um aumento no SE impacta negativamente o requisito de potência total, que cresce exponencialmente no sistema e, portanto, mais recursos de potência devem ser alocados para satisfazer a restrição de equidade de taxa de bits ou a mesma taxa de dados dos usuários.

A partir da frente de Pareto, as curvas médias de  $EE \times SE$  foram obtidas na Figura 5.21 para o método  $\epsilon$ -C-SQP, considerando  $r_D = 100$  m. Este gráfico mostra que o máximo (ótimo) EE é alcançado com diferentes números de usuários na faixa de  $M \in \{4; 40\}$  usuários,  $K \in [2; 3; 4]$  usuários/*cluster* e considerando três expoentes de perda de percurso diferentes  $\alpha \in [2; 3; 3,5]$  e raio da célula de  $r_D = 100$  m. Lembrando



(a) Tempo decorrido para os três métodos de BOO propostos.



(b) NMSE para os três métodos de BOO propostos.

Figura 5.17 – Comparação do tempo decorrido e NMSE alcançado pelos três métodos BOO propostos  $\epsilon$ -C-SQP,  $\epsilon$ -C-DK, e  $\epsilon$ -C-HC, sob convergência total considerando três cenários diferentes (valores de  $r_D$ ).

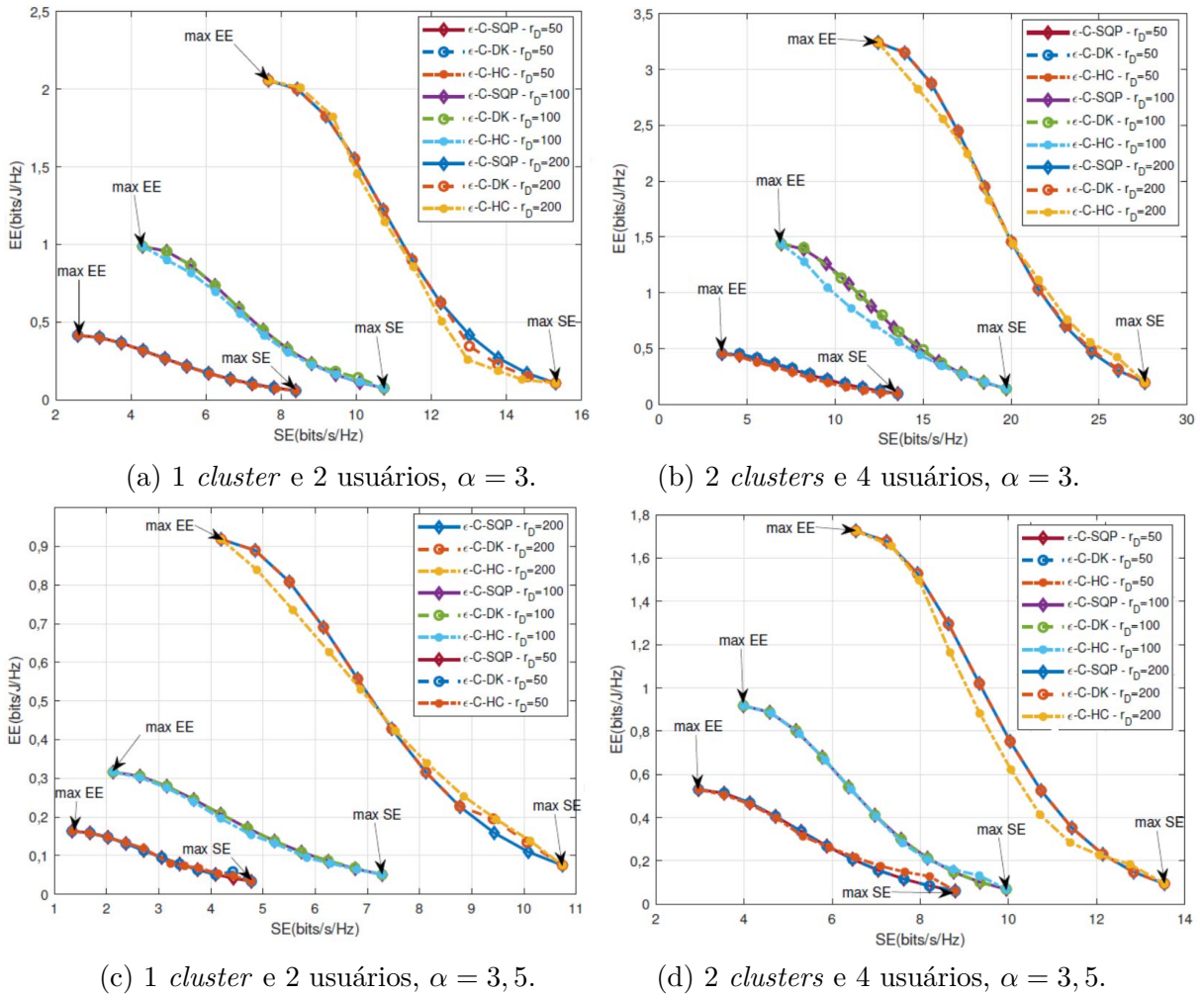


Figura 5.18 – Frente de Pareto obtida pelos três métodos de BOO  $\epsilon$ -C-SQP,  $\epsilon$ -C-DK e  $\epsilon$ -C-HC para diferentes cenários (valores de  $r_D$ ); expoentes de perda de percurso  $\alpha = 3$  e  $\alpha = 3, 5$ .

que para cada ponto de uma mesma curva, todos os usuários possuem a mesma taxa e, portanto, equidade máxima. De fato, a partir dessas curvas, é possível obter o número ótimo de usuários na célula que resulta no máximo EE para diferentes expoentes de perda de percurso do canal e o número de usuários/*cluster* em configurações MIMO-NOMA. Por exemplo, para  $\alpha = 2$  o  $EE^* = 6$  [bps/Hz] é obtido em  $M^* = 4$  usuários quando  $K = 2$  usuários/*cluster*. Tal EE máximo diminui para  $EE^* \approx 1,03$  [bps/Hz] em  $K = 4$  usuários/*cluster*, que é alcançado com  $M^* = 20$  usuários na célula. A Tabela 5.9 resume o número ótimo de usuários  $M^*$  para atingir EE máximo sob um número distinto de usuários por *cluster* e expoentes de perda de percurso. Em geral, quanto maior o número de usuários por *cluster* e o expoente de perda de percurso, menor é o EE ótimo.

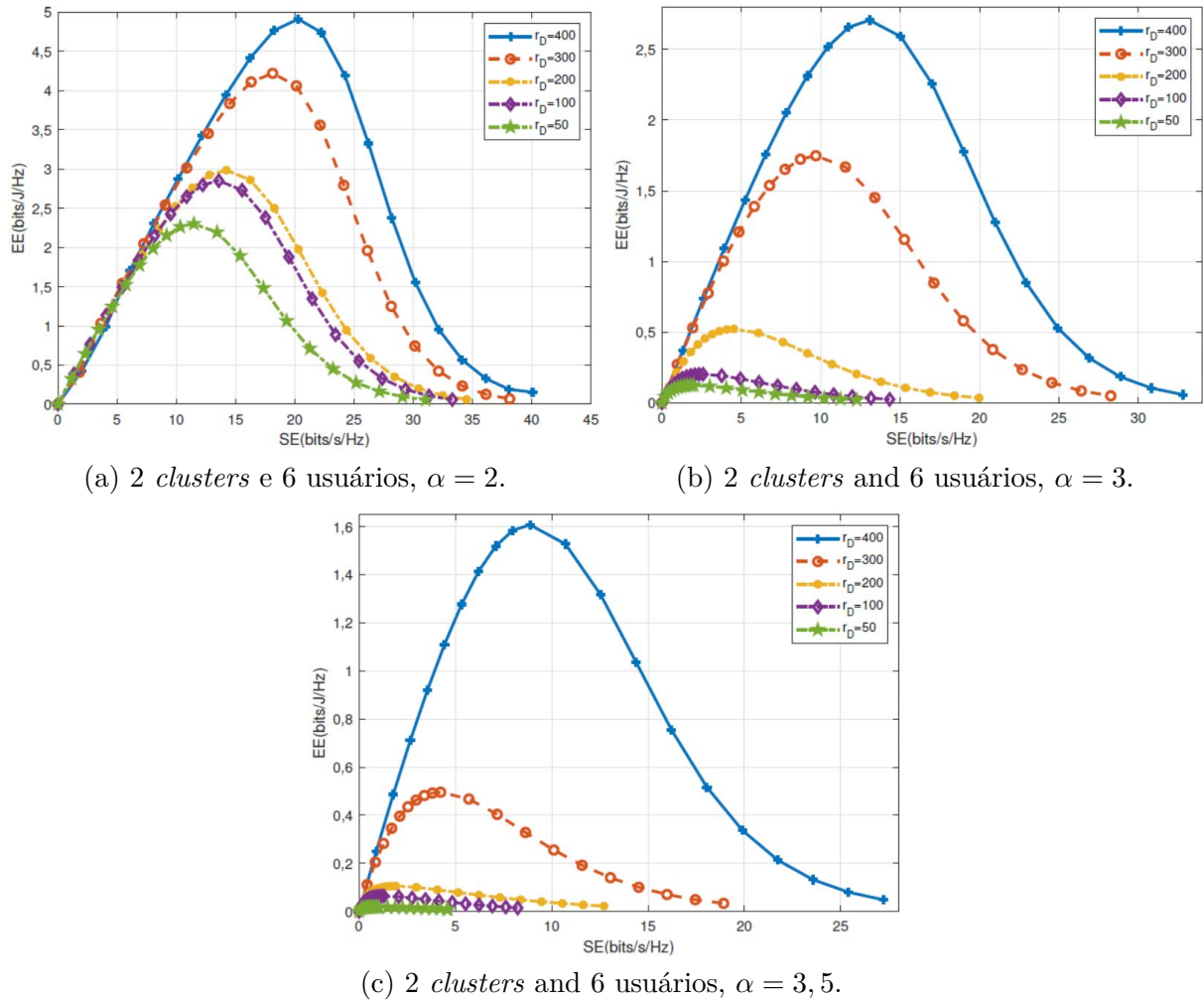


Figura 5.19 – EE  $\times$  SE obtido para o método  $\epsilon$ -C-SQP sob diferentes valores de raio de célula  $r_D$ , considerando expoentes de perda de percurso  $\alpha = 2, 3$  e  $3,5$ .

Tabela 5.9 – Número ótimo de usuários  $M^*$  para atingir o EE ótimo, considerando o raio da célula  $r_D = 100$  m, número diferente de usuários/*cluster*  $K$  e usuários por célula  $M$ .

Coef. Perda Percurso	$K = 2$	$K = 3$	$K = 4$
$\alpha = 2,0$	4	6	8
$\alpha = 3,0$	8	9	12
$\alpha = 3,5$	12	15	24

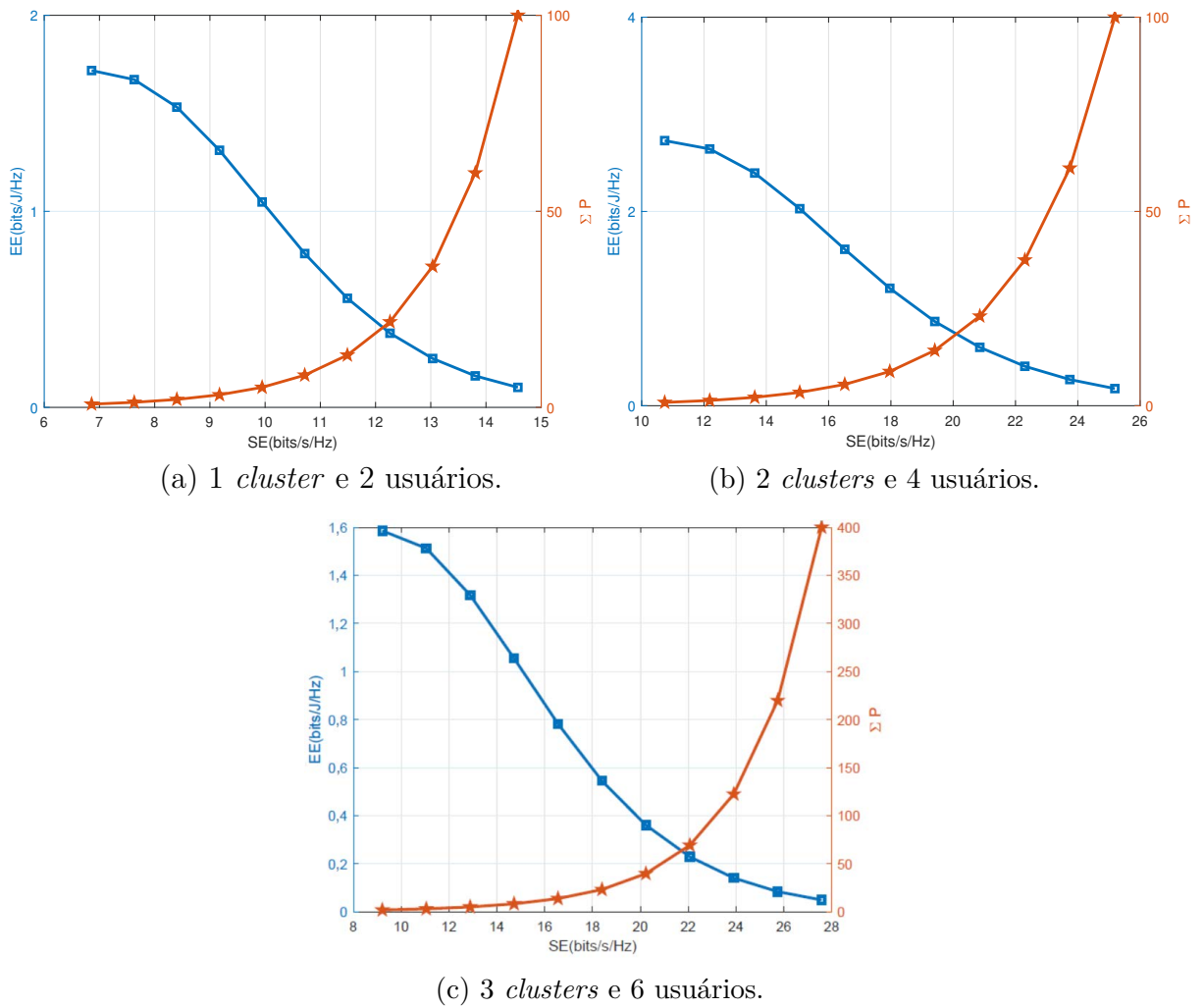


Figura 5.20 – Compromisso entre EE-SE e orçamento de potência considerando o método  $\epsilon$ -C-SQP; 2 usuários/*cluster*;  $\alpha = 3$  e  $r_D = 50$  m.

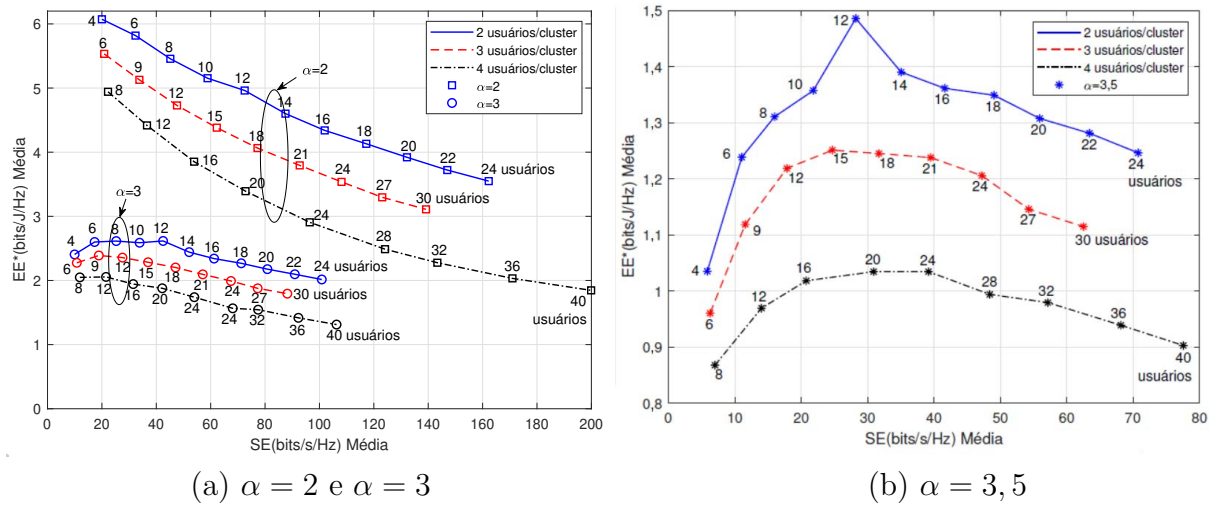


Figura 5.21 – EE máximo extraído da frente de Pareto contra SE para o método  $\epsilon$ -C-SQP. Raio da célula  $r_D = 100$  m; diferente número de usuários para 2, 3 e 4 usuários/*cluster* e expoentes de perda de percurso  $\alpha$ .

### 5.3.3 Resultados da Comparação com Método de Escalarização de Soma Ponderada

Tabela 5.10 apresenta os resultados numéricos associados aplicando o método WS-SQP para resolver o compromisso entre EE-SE em MIMO-NOMA. Pode-se observar que o método WS gera problemas mais complexos a serem resolvidos pelo método SQP em comparação ao método  $\epsilon$ -C. O método  $\epsilon$ -C-SQP alcançou convergência em todas as execuções, para as quais o método atinge soluções distintas, e, como esperado pelo Teorema 1, a restrição (C. $\epsilon$ ) no problema (3.29) é igualmente satisfeito em todos os casos avaliados, verificando os valores NMSE obtidos na Seção 5.3.2. Assim, para os métodos WS-SQP e  $\epsilon$ -C-SQP, pode-se concluir pelo mesmo comportamento em relação ao sucesso de convergência (SN), mas o número de soluções distintas na frente de Pareto (S-PF) em relação àquelas obtidos por WS-SQP são inferiores a  $\epsilon$ -C-SQP. Especificamente, WS-SQP levou mais tempo e consumiu mais avaliações FunC para alcançar a convergência em comparação com o método  $\epsilon$ -C-SQP.

Tabela 5.10 – Figura de mérito obtida com o método WS-SQP para o problema EE-SE MIMO-NOMA; expoente de perda de percurso  $\alpha = 3$ .

Método	$r_D$ [m]	SN [%]	Tempo [seg.]	$R^*$ [bps/Hz]	FunC	S-PF
<b>1 cluster e 2 usuários</b>						
WS-SQP	50	100	1,4610	3,5113	14077	6
WS-SQP	100	100	1,3420	2,2626	13032	7
WS-SQP	200	100	1,5191	1,3445	13711	8
<b>2 clusters and 4 usuários</b>						
WS-SQP	50	100	2,5017	2,9938	22374	5
WS-SQP	100	100	1,5666	1,6799	21050	6
WS-SQP	200	100	1,5398	1,1583	20392	7
<b>3 clusters and 6 usuários</b>						
WS-SQP	50	100	2,5882	3,9422	32346	4
WS-SQP	100	100	2,2437	2,8636	31798	6
WS-SQP	200	100	2,1345	1,9545	30772	6

A Figura 5.22 mostra a frente de Pareto obtida para ambos os métodos de escalarização considerados no problema do compromisso entre EE-SE MIMO-NOMA. Nota-se também que os métodos WS-SQP repetiram soluções e, como esperado devido à sua construção, as soluções não estão uniformemente distribuídas na curva. Além disso, as soluções obtidas pelo WS-SQP foram próximas ao método  $\epsilon$ -C-SQP, o que reforça o bom desempenho das abordagens propostas.

Nossos extensos resultados numéricos indicam que para o método de escalonamento WS ser considerado como uma opção para o problema de compensação EE-SE em sistemas

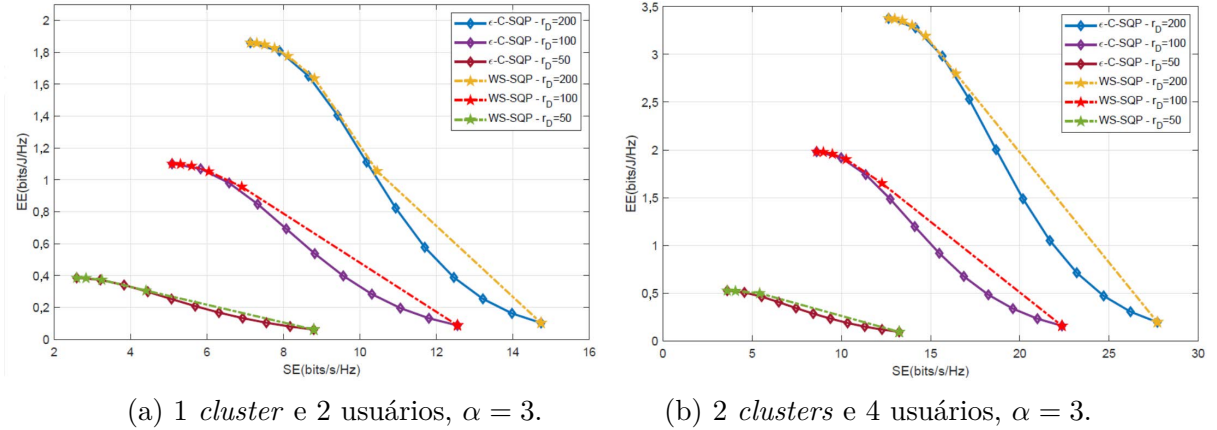


Figura 5.22 – Frente de Pareto obtida pelos dois métodos BOO  $\epsilon$ -C-SQP e WS-SQP para diferentes cenários (valores  $r_D$ ); expoente de perda de percurso  $\alpha = 3$ .

MIMO-NOMA, é necessário considerar um maior número de variações para o parâmetro  $\tau$  a fim de determinar mais pontos na fronteira de Pareto. No entanto, essa estratégia implica em custos computacionais adicionais do método WS-SQP; portanto, o método teria um desempenho ainda menor a este respeito do que o método  $\epsilon$ -C-SQP.

## 5.4 Resultados para Max-EE em Sistemas mMIMO PDMA com Serviços eMBB, mMTC e uRLLC

Esta seção busca obter a matriz padrão que melhor atende aos requisitos para cada um dos serviços eMBB, mMTC e uRLLC no sistema mMIMO PDMA. Para isso, consideraremos o cálculo de diversos parâmetros envolvendo as matrizes padrão como repetibilidade  $r$ , correlação quadrática da soma ao quadrado SSC, diversidade DIV, peso máximo ponderado das linhas  $d_{\max}^w$  e o fator de avaliação de desempenho  $\Lambda$ . Três cenários diferentes foram avaliados, cenários A, B e C, conforme esboçado na Figura 5.23. O problema de otimização foi resolvido para obter o EE máximo sob cada matriz PDMA padrão considerando que o critério *todos os usuários têm a mesma taxa* ou máxima equidade é selecionado para o sistema mMIMO PDMA. Ao analisar esses dados, pode-se obter a matriz padrão que melhor atende a cada tipo de serviço (eMBB, mMTC e uRLLC).

Os valores dos parâmetros adotados na simulação numérica estão descritos na Tabela 5.11. O número de antenas na BS que formam cada feixe e para cada usuário são definidos como  $N_T = 16$  e  $N_R = 4$ , respectivamente. Os usuários ficam ao redor da BS com um raio de até 200 metros. Os cenários A, B e C definem as distâncias fixas dos usuários à BS de acordo com o número de usuários por feixe da matriz padrão, conforme Figura 5.23.

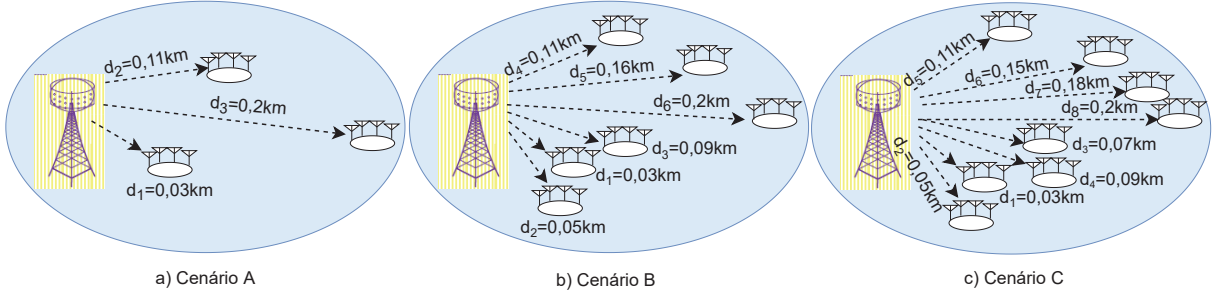


Figura 5.23 – Cenários A, B e C (máxima-equidade) para o problema de otimização de EE em DL, Eq. (4.45).

Tabela 5.11 – Parâmetros Adotados Para os Cenários de Simulação.

Parâmetro	Valor
#Usuários	$K \in \{3; 6; 8\}$
Expoente da perda de percurso	$\alpha = 3,0$
Potência do Ruído Branco	$\sigma_n^2 = -100$ dBm
Potência <i>Budget</i>	$P = 100$ W
# Feixes	$N = 2, 3$ e $4$
Usuários/feixe	$K \in \{2, 3, 4, 5\}$
Antenas/feixe	$N_T = 16$
# antenas por usuário	$N_R = 4$
Meta de igual taxa entre usuários normalizada	$R \in [4; 45]$ [bits/s/Hz]
Potência dos circuitos	$P_{BS} = 1$ W/antena (BS) $P_K = 100$ mW/antena por usuário
Ineficiência do amplificador de potência de RF	$\rho = 1, 4$
<b>Cenários:</b> Distribuição dos usuários dentro da célula: 3, 6 or 8 usuários, 2, 3 ou 4 feixes	
Cen. A – 3 usuários (Figura 5.23 (a))	$d_1 = 0,03, d_2 = 0,11, d_3 = 0,20$ [km]
Cen. B – 6 usuários (Figura 5.23 (b))	$d_1 = 0,03, d_2 = 0,05, d_3 = 0,09,$ $d_4 = 0,11, d_5 = 0,16, d_6 = 0,20$ [km]
Cen. C – 8 usuários (Figura 5.23 (c))	$d_1 = 0,03, d_2 = 0,05,$ $d_3 = 0,07, d_4 = 0,09, d_5 = 0,11,$ $d_6 = 0,15, d_7 = 0,18, d_8 = 0,20$ [km]

### 5.4.1 Seleção da Matriz Padrão PDMA

As matrizes padrão do Modelo 1 ao Modelo 9 na Tabela 5.12 foram geradas a partir de matrizes existentes em (SUN et al., 2020), enquanto as matrizes do Modelo 10 ao Modelo 12 foram geradas a partir de (LI et al., 2021), e do Modelo 13 ao Modelo 15 de (REN et al., 2016). Além disso, para simplicidade de análise, outras matrizes PDMA encontradas na literatura em que os valores da figura de mérito  $\nu$ , SSC,  $d_{\max}^W$ , DIV,  $\Lambda$  coincidem totalmente com as do Modelo 1 ao Modelo 15 não foram incluídas em nossas análises, pois apesar dos diferentes *designs* de matrizes padrão, elas produzem o mesmo desempenho no sistema. Originalmente, as matrizes padrão do Modelo 7 ao Modelo 15 na Tabela 5.12 atendem às condições para operar em UL; entretanto, neste trabalho todas as

Tabela 5.12 – Matrizes padrão do projeto PDMA de diferentes dimensões (SUN et al., 2020), (REN et al., 2016), (LI et al., 2021).

Modelo	Matriz Padrão	$\iota$	SSC	$d_{\max}^W$	DIV	$\Lambda$
$G_{\text{Modelo1}}^{4 \times 6}$	$\begin{bmatrix} 0,4082 & 0 & 0 & 0 & 0,4082 & 0 \\ 0 & 0,4082i & 0 & 0 & 0 & 0,4082i \\ 0 & 0 & 0,4082 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0,4082i & 0 & 0 \end{bmatrix}$	150%	0,06	2	1	0,12
$G_{\text{Modelo2}}^{4 \times 6}$	$\begin{bmatrix} 0,408 & 0 & 0 & 0,236 & 0,236 & 0 \\ 0 & 0,408i & 0 & 0,236i & 0,236 & 0 \\ 0 & 0 & 0,408 & 0,236 & 0,236i & 0 \\ 0 & 0 & 0 & 0 & 0 & 0,408i \end{bmatrix}$	150%	0,06	9	3	0,18
$G_{\text{Modelo3}}^{4 \times 8}$	$\begin{bmatrix} 0,25 & 0,25i & 0 & 0 & 0,354 & 0 & 0 & 0 \\ 0,25i & 0 & 0,25 & 0 & 0 & 0,354i & 0 & 0 \\ 0 & 0,25 & 0 & 0,25i & 0 & 0 & 0,354 & 0 \\ 0 & 0 & 0,25i & 0,25 & 0 & 0 & 0 & 0,354i \end{bmatrix}$	200%	0,08	12	2	0,48
$G_{\text{Modelo4}}^{4 \times 6}$	$\begin{bmatrix} 0,2911 & 0 & 0,2276 & 0,2276i & 0,2633 & 0 \\ 0 & 0 & 0,2276i & 0,2276 & 0,2633i & 0,2633 \\ 0 & 0 & 0,2276 & 0,2276i & 0,2633 & 0,2633i \\ 0 & 0,2911i & 0,2276i & 0,2276 & 0 & 0,2633 \end{bmatrix}$	150%	0,10	24	4	0,60
$G_{\text{Modelo5}}^{4 \times 6}$	$\begin{bmatrix} 0,2692 & 0 & 0,1860 + 0,1867i & 0,2104i & 0,1999 + 0,2005i & 0 \\ 0 & 0 & 0,2104 & 0,2124 + 0,1603i & 0,2434 & 0,2434i \\ 0 & 0 & 0,1860 + 0,1867i & 0,2104 & 0,2434i & 0,2434 \\ 0 & 0,2692i & 0,2104i & 0,2124 + 0,1603i & 0 & 0,1999 + 0,2005i \end{bmatrix}$	150%	0,13	24	4	0,78
$G_{\text{Modelo6}}^{4 \times 8}$	$\begin{bmatrix} 0,2911 & 0,2633 & 0,2276 & 0,2633 & 0 & 0 & 0 & 0 \\ 0,2276i & 0,2276i & 0 & 0 & 0,2276i & 0,2633i & 0 & 0 \\ 0,2276 & 0 & 0,2633i & 0 & 0,2633 & 0 & 0,2276 & 0 \\ 0 & 0 & 0 & 0,2276 & 0 & 0,2276i & 0,2633 & 0,2911i \end{bmatrix}$	200%	0,11	24	3	0,88
$G_{\text{Modelo7}}^{4 \times 6}$	$\begin{bmatrix} 0,4083 & 0 & 0,2042 & 0,2042i & 0,2356 & 0 \\ 0 & 0 & 0,2042i & 0,2042 & 0,2356i & 0,2356 \\ 0 & 0 & 0,2042 & 0,2042i & 0,2356 & 0,2356i \\ 0 & 0,4083i & 0,2042i & 0,2042 & 0 & 0,2356 \end{bmatrix}$	150%	0,09	24	4	0,54
$G_{\text{Modelo8}}^{4 \times 6}$	$\begin{bmatrix} 0,4082 & 0 & 0,1021 + 0,1768i & 0,1445 + 0,1445i & 0,1180 + 0,2041i & 0 \\ 0 & 0 & 0,1021 + 0,1768i & 0,2041i & 0,2356 & 0,2356 \\ 0 & 0 & 0,2041 & 0,1445 + 0,1445i & 0,2356 & 0,2356 \\ 0 & 0,4082i & 0,1021 + 0,1768i & 0,2041 & 0 & 0,1180 + 0,2041i \end{bmatrix}$	150%	0,12	24	4	0,72
$G_{\text{Modelo9}}^{4 \times 8}$	$\begin{bmatrix} 0,204 & 0,25 & 0,25i & 0,25 & 0 & 0 & 0 & 0 \\ 0,204i & 0,25i & 0 & 0 & 0,25i & 0,25 & 0 & 0 \\ 0,204 & 0 & 0,25 & 0 & 0,25 & 0 & 0,25i & 0 \\ 0 & 0 & 0 & 0,25i & 0 & 0,25i & 0,25 & 0,3536 \end{bmatrix}$	200%	0,10	24	3	0,80
$G_{\text{Modelo10}}^{2 \times 3}$	$\begin{bmatrix} -0,4082 & -0,5774i & 0 \\ -0,4082i & 0 & -0,5774 \end{bmatrix}$	150%	0,11	2	2	0,11
$G_{\text{Modelo11}}^{3 \times 6}$	$\begin{bmatrix} 0,2357 & 0 & 0 & 0 & 0,4082i & -0,4082 \\ 0,2357 & 0,2887 & 0 & 0,4082 & 0 & 0 \\ -0,2357 & 0,2887 & 0,4082 & 0 & 0 & 0 \end{bmatrix}$	200%	0,09	9	3	0,27
$G_{\text{Modelo12}}^{4 \times 8}$	$\begin{bmatrix} 0,1768i & -0,2041i & 0 & -0,25 & -0,25i & 0 & 0 & -0,3536i \\ -0,1768i & 0 & -0,25i & 0 & 0 & 0 & -0,3536 & 0 \\ -0,1768 & 0,2041i & 0,25 & 0 & -0,25 & 0 & 0 & 0 \\ -0,1768 & 0,2041 & 0 & 0,25i & 0 & 0,3536 & 0 & 0 \end{bmatrix}$	200%	0,08	25	4	0,50
$G_{\text{Modelo13}}^{4 \times 8}$	$\begin{bmatrix} 0,3536 & 0 & 0 & 0,204 & -0,1022 + 0,1768i & -0,1022 - 0,1768i & -0,1022 - 0,1768i & 0 \\ 0 & 0,3536 & 0 & 0,204 & 0,204 & 0,204 & -0,204 & 0 \\ 0 & 0 & 0,3536 & 0,204 & 0,204 & 0,204 & 0,204 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,3536 \end{bmatrix}$	200%	0,08	30	3	0,80
$G_{\text{Modelo14}}^{4 \times 6}$	$\begin{bmatrix} 0,4082 & 0 & -0,1021 + 0,1768i & -0,1768 - 0,1021i & -0,1445 + 0,1445i & -0,1768 + 0,1021i \\ 0 & 0,4082 & -0,1021 - 0,1768i & 0,1021 + 0,1768i & -0,2041i & 0,1021 - 0,1768i \\ 0 & 0 & 0,2041 & -0,2041i & 0,1445 + 0,1445i & 0,2041i \\ 0 & 0 & -0,1021 + 0,1768i & -0,1021 + 0,1768i & -0,2041 & -0,1021 - 0,1768i \end{bmatrix}$	150%	0,13	32	4	1,04
$G_{\text{Modelo15}}^{4 \times 8}$	$\begin{bmatrix} 0,2041 & 0 & 0,2041 & 0,25 & 0,25 & 0 & 0 & 0,3537 \\ 0 & 0,2041 & 0,2041 & 0,25 & 0 & 0,25 & 0 & 0 \\ 0,2041 & 0,2041 & 0,2041 & 0 & 0 & 0,25 & 0,25 & 0 \\ 0,2041 & 0,2041 & 0 & 0 & 0,25 & 0 & 0,25 & 0 \end{bmatrix}$	200%	0,14	32	3	1,49

matrizes padrão foram geradas para satisfazer a condição de *potência total de transmissão disponível* na BS, Eq. (4.25). Esta condição foi garantida em todos os três cenários A, B e C satisfazendo o sistema operando em DL.

As condições de operação das matrizes PDMA no DL são diferentes das do UL. De fato, para os cenários DL, os projetos de matriz PDMA existentes na literatura ainda são notavelmente escassos. Assim, para atender a essa necessidade, as matrizes padrão projetadas originalmente para operar no UL foram aqui adaptadas para operar na direção DL, aplicando a condição expressa na Eq. (4.25). Esta adaptação foi feita simplesmente normalizando o valor do lado esquerdo dos elementos da matriz da Eq. (4.25).

Os gráficos da Figura 5.24 revelam o comportamento da eficiência energética ótima  $EE^*$  com o aumento da taxa de usuário  $R$ . Os pontos  $EE^*$  obtidos em tais gráficos são o resultado da solução EE de imparcialidade do problema em (4.45). Para resolução desse problema foi aplicado o Algoritmo 8 conforme descrito na Subseção 3.2.3.2 com o uso da função `fmincon` do aplicativo MatLab. Observe que nestas curvas os resultados de máximo  $EE^*$  sob condição de taxa de usuário igual (equidade), mostrado na Figura 5.24 e resumido na Figura 5.25, apresentam recursos de diferentes eficiências, *i.e.*  $RE = EE^* \times R$  desempenho. Os valores das figuras de mérito  $\nu$ , **SSC**,  $d_{\max}^W$ , **DIV** e  $\mathbf{\Lambda}$  expressos nas tabelas desta seção de resultados foram obtidos a partir das Eq. (4.20), (4.29), (4.31), (4.30) e (4.34), respectivamente.

Nas próximas subseções são analisadas as matrizes padrão que atendem aos requisitos dos respectivos serviços dos modos de uso eMBB, mMTC e uRLLC em sistemas 5G.

## 5.4.2 Análise das Matrizes Padrão PDMA para o eMBB

Na Tabela 5.13 representamos a eficiência de recursos ideal (RE) como:

$$RE^* = EE_{\max}^* \cdot R \quad \left[ \frac{\text{bits}^2}{\text{watts}} \right] \quad (5.3)$$

para compor os valores de  $EE_{\max}^*$  e  $R$ . Observe que  $RE^*$  é obtido para o máximo de  $EE^*$ , ou seja,  $EE_{\max}^*$ , que vem do resultado de  $EE^*$  resolvendo (4.45) para cada valor de  $R$  com um número fixo de usuários  $K$  de acordo com cada matriz padrão; o comportamento de  $EE^*$  é revelado na Figura 5.24. Portanto, para usar PDMA em serviços eMBB é necessário maximizar a taxa de transmissão. Neste trabalho, otimizamos o sistema mMIMO PDMA para operar para todas as matrizes padrão nos cenários A, B e C sob  $EE$  máximo na condição de mesma taxa entre usuários ou máxima equidade.

As matrizes padrão que apresentaram a maior taxa de bits para o  $EE$  máximo estão resumidas na Tabela 5.13 e obtidas diretamente da Figura 5.24.(a) e Figura 5.25. **Cenário A** (Figura 5.23) é adotado na avaliação da matriz padrão Modelo 10; **Cenário B**

para avaliação do Modelo 1, Modelo 2, Modelo 4, Modelo 5, Modelo 7, Modelo 8, Modelo 11 e Modelo 14; e **Cenário C** para avaliação das matrizes padrão Modelo 3, Modelo 6, Modelo 9, Modelo 12, Modelo 13 e Modelo 15.

Tabela 5.13 – Parâmetros a serem analisados para os serviços eMBB.

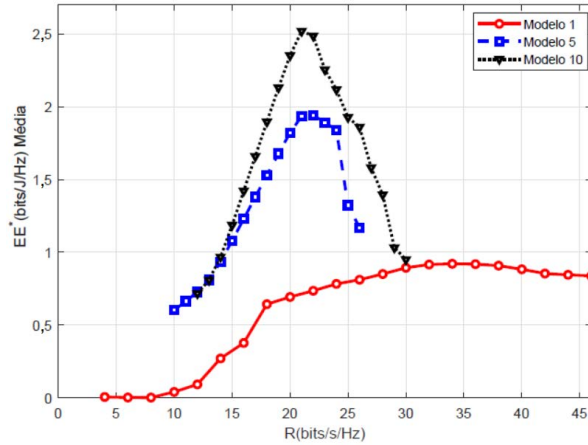
PDMA Modelo	$R$ $\left[\frac{\text{bits}}{\text{s}\cdot\text{Hz}}\right]$	$EE_{\max}$ $\left[\frac{\text{bits}}{\text{W}}\right]$	$RE^*$ $\left[\frac{\text{bits}^2}{\text{W}}\right]$	#Us.	$\iota$	SSC	$d_{\max}^W$	DIV	$\Lambda$
<b>1<sup>†</sup></b>	34	0,92	31,28	6	150%	0,06	2	1	0,12
5	22	1,94	42,68	6	150%	0,13	24	4	0,78
10	21	2,52	52,92	3	150%	0,11	2	2	0,11

<sup>†</sup> Matriz padrão que melhor atende aos requisitos dos serviços eMBB.

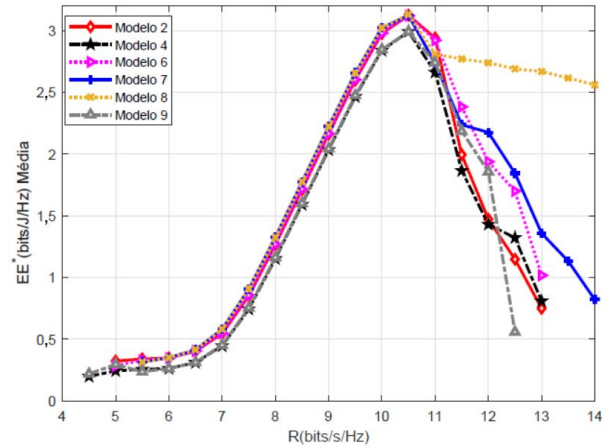
Embora as matrizes padrão Modelo 1 e Modelo 5 tenham resultado em valores máximos de EE mais altos, a matriz padrão que resultou na maior taxa de bits de usuários para o EE máximo foi a matriz Modelo 1, apesar de  $RE^*$  ser menor. A matriz padrão Modelo 5 não é preferida para o serviço eMBB porque possui o parâmetro  $d_{\max}^W = 24$ , portanto resulta em alta complexidade, conforme inferido da Tabela 5.12, e resumido na Tabela 5.13. Excluindo o Modelo 5, temos que decidir entre as matrizes Modelo 1 e Modelo 10. O Modelo 1 tem uma taxa de bits muito maior para EE máximo do que o Modelo 10. Além disso, o parâmetro SSC da matriz padrão Modelo 1 é 0,06 e para o Modelo 10 é  $SSC = 0,11$ , o que resulta na maior esparsidade alcançada pela matriz padrão Modelo 1. Além disso, a matriz Modelo 1 pode atender até 6 usuários enquanto a matriz Modelo 10 atende até 3 usuários, aumentando as vantagens da matriz Modelo 1. Assim, concluímos que a matriz Modelo 1 é a matriz padrão que melhor se adequa ao serviço eMBB, conforme destacado na Tabela 5.13.

### 5.4.3 Matrizes padrão PDMA para mMTC

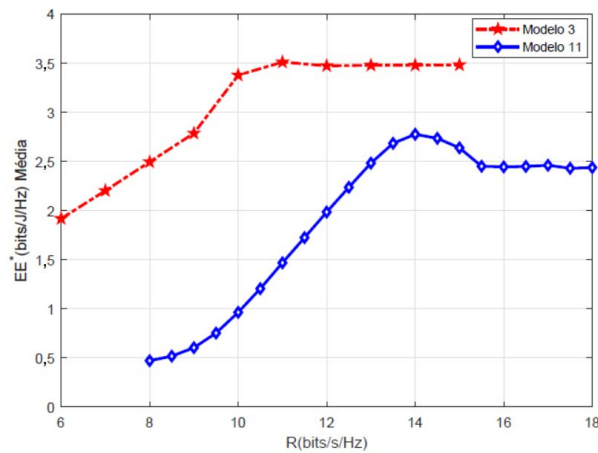
Para avaliar as matrizes padrão para serviços mMTC, deve-se considerar um alto valor de OF e uma alta dimensão de matriz. É claro que quanto maior o OF melhor o serviço mMTC atende a sua demanda. A alta dimensão é essencial para o serviço mMTC porque eles precisam se conectar a um grande número de terminais enquanto aceitam alta latência. Analisando as matrizes candidatas pode-se selecionar as matrizes Modelos 3, 6 e 9 como os maiores OFs e dimensões com  $\iota = 200\%$  e 8 usuários de acordo com a Tabela 5.12 e, por conveniência, também apresentado na Tabela 5.14. De fato, entre eles, o Modelo 3 é a matriz padrão com o melhor trade-off EE-SE, resultando no produto max-EE e taxa de bits mais alta (com equidade máxima), isto é, máximo RE conforme Tabela 5.14 e Figura 5.25. A matriz padrão Modelo 3 resultou no EE máximo mais alto de 3,51 [bits/W]. Ao considerar a composição entre  $EE^*$  e  $R$  o correspondente  $RE^*$  para a matriz Modelo 3 resultou em 38,61 [bits<sup>2</sup>/W], o maior valor entre os resultados da Tabela 5.14. Pode-se concluir que a matriz padrão Modelo 3 é a melhor entre as candidatas a matriz padrão PDMA da Tabela 5.12 para o serviço mMTC.



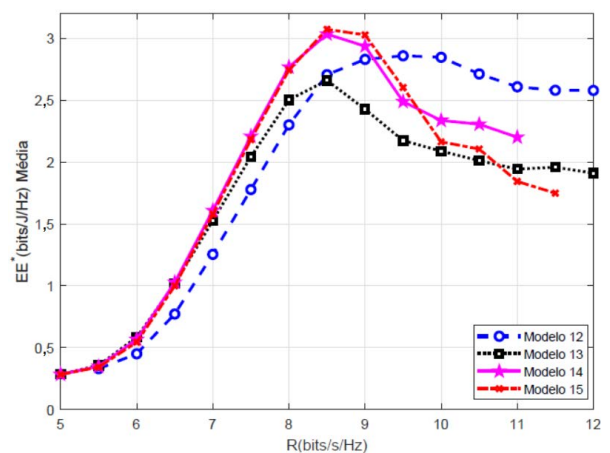
a) Modelos 1, 5 ( $K = 6$  usuários) e 10 ( $K = 3$ )



b) Modelos 2, 4, 7, 8 ( $K = 6$ ), e 6, 9 ( $K = 8$ ).



c) Modelo 3 ( $K = 8$ ) e 11 ( $K = 6$ )



d) Modelos 12, 13, 15 ( $K = 8$ ), e 14 ( $K = 6$ ).

Figura 5.24 –  $EE^* \times R$  para  $\alpha = 3$ . Os valores ótimos de EE foram obtidos resolvendo Eq. (4.45).

Tabela 5.14 – Parâmetros a serem analisados para os serviços mMTC.

PDMA Modelo	$R$ $\left[\frac{\text{bits}}{\text{s}\cdot\text{Hz}}\right]$	$EE_{\max}$ $\left[\frac{\text{bits}}{\text{W}}\right]$	$RE^*$ $\left[\frac{\text{bits}^2}{\text{W}}\right]$	#Us.	$\iota$	SSC	$d_{\max}^W$	DIV	$\Lambda$
<b>3<sup>†</sup></b>	11,0	3,51	38,61	8	200%	0,08	12	2	0,48
6	10,5	3,12	32,76	8	200%	0,08	12	2	0,48
9	10,5	2,99	31,40	8	200%	0,08	12	2	0,48

<sup>†</sup> Matriz padrão que melhor atende aos requisitos dos serviços mMTC.

### 5.4.4 Matrizes padrão PDMA para uRLLC

Para serviços uRLLC, uma matriz padrão PDMA deve minimizar a interferência enquanto mantém a dispersão. Isto é conseguido com: a) baixo OF; b) correlação de soma quadrada minimizada (SSC); c) minimizar a complexidade  $d_{\max}^W$ ; d) consequentemente reduzindo o fator de avaliação de desempenho (PEF). As matrizes padrão com o menor OF são as de Modelo 1, Modelo 2, Modelo 4, Modelo 7 e Modelo 10, conforme mostrado na Tabela 5.12 e por conveniência na Tabela 5.15. Dentre essas matrizes, as de menor PEF são as matrizes padrão Modelo 1 e Modelo 10. Além disso, analisando as curvas obtidas

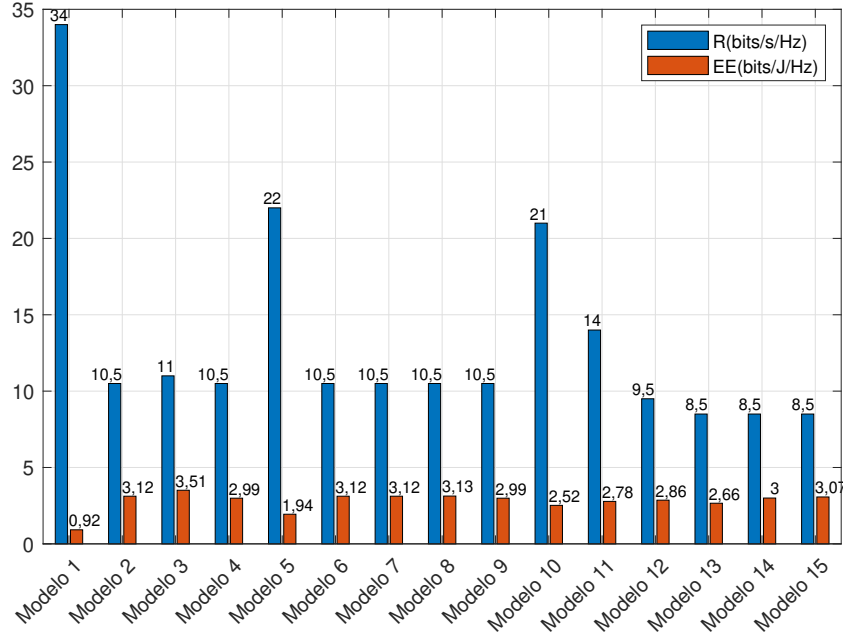


Figura 5.25 – Valores obtidos de  $EE^*$  máximo e a respectiva igual taxa de bits  $R$  resolvendo o problema de otimização (4.45) sob um número fixo de usuários  $K$  de acordo com cada matriz padrão. Cenário A foi adotado para avaliar a matriz padrão Modelo 10; Cenário B os Modelos 1, 2, 4, 5, 7, 8, 11 e 14; Cenário C os Modelos 3, 6, 9, 12, 13 e Modelo 15.

Tabela 5.15 – Parâmetros a serem analisados para os serviços uRLLC.

PDMA Modelo	$R$ $\left[\frac{\text{bits}}{\text{s}\cdot\text{Hz}}\right]$	$EE_{\max}$ $\left[\frac{\text{bits}}{\text{W}}\right]$	$RE^*$ $\left[\frac{\text{bits}^2}{\text{W}}\right]$	#Us.	$\iota$	SSC	$d_{\max}^W$	DIV	$\Lambda$
1	34,0	0,92	31,28	6	150%	0,06	2	1	0,12
2	10,5	3,12	32,76	6	150%	0,06	9	3	0,18
4	10,5	2,99	31,40	6	150%	0,10	24	4	0,60
7	10,5	3,12	32,76	6	150%	0,09	24	4	0,54
<b>10<sup>†</sup></b>	21,0	2,52	52,92	3	150%	0,11	2	2	0,11

<sup>†</sup> Matriz padrão que melhor atende aos requisitos dos serviços uRLLC.

após o processo de otimização de EE máximo e imparcialidade máxima considerando esta matriz padrão, pode-se obter o  $EE_{\max} = 2,52$  [bits/W] para uma taxa de bits de 21 [bits/s/Hz] para o Modelo 10 e apenas 0,92 [bits/W] e 34 [bits/s/Hz] para o Modelo 1, conforme mostrado na Figura 5.24 (a) e resumido na Figura 5.25. Embora a matriz padrão Modelo 1 tenha uma taxa maior ao compor o  $EE^*$  e  $R$ , essa matriz resulta em 31,28 enquanto a matriz padrão Modelo 10 resulta em 52,92, portanto, nesse quesito, a matriz padrão Modelo 10 é muito maior. O que representa uma alta taxa e EE para esta matriz padrão.

Considerando que todas as matrizes padrão da Tabela 5.15 possuem o mesmo OF, aquela que simultaneamente garante a baixa latência (menor dimensão), e garante o melhor desempenho (menor PEF) é a matriz padrão Modelo 10. O PEF para matriz padrão Modelo 1 é 0,12 enquanto para matriz padrão Tipo 10 resulta em 0,11 que é menor principalmente porque possui maior diversidade, Tabela 5.15. Em conclusão, a

matriz padrão que melhor atende aos requisitos de serviço uRLLC é a de Modelo 10.

### 5.4.5 Matrizes padrão PDMA para transmissão híbrida com serviços eMBB e uRLLC

O serviço eMBB requer uma alta taxa de transmissão, enquanto o serviço uRLLC é altamente restrito em termos de atraso e confiabilidade. De acordo com as Tabelas 5.13 e 5.16 entre os candidatos a matriz padrão que atendem a esses requisitos é apenas o Modelo 10. Esta é uma matriz que suporta totalmente o serviço uRLLC, mas suporta parcialmente o serviço eMBB. Isso ocorre porque a taxa máxima de  $EE^*$  para a matriz padrão Modelo 10 é 21 [bits/s/Hz], enquanto a matriz padrão selecionada para o serviço eMBB foi Modelo 1 com máximo  $EE^*$  de 34 [bits/s/Hz]. Por outro lado, a matriz Modelo 10 tem melhor desempenho, pois seu valor de PEF é menor. A taxa de 21 [bits/s/Hz] para a matriz padrão Modelo 10 ainda é alta em comparação com as outras opções na Tabela 5.12 e Figura 5.25. Essa taxa é a terceira maior dentre as taxas obtidas para o máximo  $EE^*$  considerando todas as quinze matrizes padrão. Em conclusão, a matriz de padrão Modelo 10 deve ser selecionada para funcionar em um modo de uso híbrido duplo eMBB e uRLLC.

Tabela 5.16 – Parâmetros a serem analisados para os serviços híbridos eMBB e uRLLC.

PDMA Modelo	$R$ $\left[\frac{\text{bits}}{\text{s}\cdot\text{Hz}}\right]$	$EE_{\max}$ $\left[\frac{\text{bits}}{\text{W}}\right]$	$RE^*$ $\left[\frac{\text{bits}^2}{\text{W}}\right]$	#Us.	$\iota$	SSC	$d_{\max}^W$	DIV	$\Lambda$
<b>10<sup>†</sup></b>	21,0	2,52	52,92	3	150%	0,11	2	2	0,11

<sup>†</sup> Matriz padrão que melhor atende aos requisitos dos serviços híbridos eMBB e uRLLC.

## 6 Relação entre os Trabalhos

O objetivo desse capítulo consiste em demonstrar a interconexão e a coesão das três partes dos trabalhos discutidos nos capítulos anteriores, correlacionando os resultados e as contribuições dos diferentes sistemas NOMA sob equidade de taxa máxima analisados.

### 6.1 Variantes NOMA sob Equidade de Taxa

Este trabalho de Tese explorou a análise de diferentes variantes do sistema do NOMA sob a perspectiva da equidade de taxa máxima. O trabalho buscou explorar e analisar sistematicamente o compromisso entre a EE e a SE em sistemas NOMA combinados ao MIMO, também a mMIMO, bem como ao PDMA em enlaces de descida sob a restrição de equidade de taxa, de forma que todos os usuários atinjam a mesma (e a máxima) taxa de dados.

O compromisso EE-SE, denominado RE é analisado para diferentes números de usuários por célula NOMA. Os resultados foram obtidos para os sistemas NOMA, MIMO-NOMA e mMIMO PDMA. Esses sistemas possuem em comum a técnica NOMA, a qual explora o domínio das potências para criar uma nova categoria de diversidade. Para uma adequada análise das soluções, foram propostos e desenvolvidos diferentes problemas de otimização focados na obtenção da máxima EE sob a máxima equidade de taxa dos usuários em transmissão no DL. Combinamos também o sistema mMIMO ao PDMA para oferecer suporte eficiente aos serviços eMBB, mMTC e uRLLC em sistemas 5G. Seleccionamos matrizes padrão a partir da literatura que melhor se adéquam e atendem aos requisitos heterogêneos desses serviços.

Confirmou-se que para a obtenção da equidade de máxima taxa faz-se necessário garantir que as proporções de potência dos usuários sejam otimizadas de forma a garantir que a taxa de bit seja igual para todos os usuários. As variáveis do problema de otimização correspondente incluem a potência total, a proporção de potência de cada usuário e a equidade de taxa.

No sistema NOMA básico analisado SISO, a BS e os usuários possuem uma única antena. Com essa análise preliminar foi possível: formular o problema de otimização associado, obter um procedimento expedito para determinar analiticamente a máxima EE e a respectiva taxa de equidade para cada número de usuários na célula, determinando o valor que resulta na máxima EE e a respectiva taxa de equidade através do procedimento de inspeção gráfica ou numérica. Para os diferentes cenários (canal e sistema) analisados, a máxima EE com máxima equidade foi atingida com o número mínimo de usuários NOMA

na célula: dois usuários na célula; portanto uma célula com dois usuários NOMA.

Ao agregarmos o sistema MIMO à técnica NOMA, o sistema foi capaz de suportar mais de um *cluster*. Neste contexto, a análise agregou mais uma variável, o número de usuários por *cluster*. Nesse caso, analisamos o número de usuários por célula o qual resulta na máxima EE com equidade de máxima taxa para diferentes números de usuários por *cluster* MIMO-NOMA respeitando as mesmas restrições já mencionadas. Os resultados numéricos e de simulação permitiram corroborar que o número de usuários por célula, que resulta na máxima EE do sistema com a máxima equidade, depende conjuntamente de três fatores: a) número de usuários por *cluster*; b) do expoente de perda de percurso; c) do raio (cobertura) da célula.

Ainda para o sistema MIMO-NOMA propusemos uma otimização BOO para o compromisso EE-SE; resolvemos o problema BOO a partir de diferentes técnicas de otimização utilizando a mesma função objetivo e restrições dos casos anteriores apenas com adaptações para a otimização biobjetiva. Traçamos as curvas para as diferentes técnicas com as soluções do Pareto *front*; com isso, pode-se constatar que a convergência e a precisão da técnica de otimização convexa quadrática SQP se sobressai em relação às demais técnicas de otimização adotadas.

Considerando o sistema mMIMO PDMA, adotou-se o critério de escolha da matriz padrão que melhor atendesse aos requisitos dos serviços 5G operando sob os modos de uso eMBB, mMTC e/ou uRLLC. Para isso foram selecionados quinze matrizes padrão e adaptadas, quando necessário, para operarem no modo de enlace de descida (DL). Na sequência foram calculados diversos parâmetros dessas matrizes padrão e verificado a que melhor atendesse cada quesito de serviços ou da sua combinação. Entre os critérios adotados está a maximização da EE sob equidade de taxa com as mesmas restrições adotadas nos sistemas anteriores. Adicionalmente, a partir dos resultados de simulação, analisamos o parâmetro eficiência de recursos, neste contexto dado pela métrica  $RE^* = EE_{\max}^* \times R$ , produto entre a  $EE_{\max}^*$  obtida no processo de otimização para um  $R$  fixo usado neste produto e oriundo da equidade de taxa. Assim, quantificamos o ganho de RE, primeiramente, analisando o serviço eMBB. Selecionamos três matrizes padrão as quais apresentam a EE máxima com a maior taxa de bit. Uma delas apresenta alta complexidade, tendo sido descartada. Entre as outras duas, uma delas apresentou melhor resultado em duas métricas e portanto tal matriz padrão foi selecionada para atender os usuários operando no modo de uso eMBB.

Outro modo de uso 5G analisado sob a perspectiva da seleção de matrizes PDMA foi o mMTC. Foram selecionadas as três matrizes padrão, dentre mais de quinze disponíveis na literatura, que pudessem atender o maior número de usuários e maior fator de sobrecarga. Analisando as métricas de qualidade aplicadas às três matrizes padrão, somente uma das métricas resultou com valores diferentes: o  $RE^*$ . Com isso, foi selecionada a matriz padrão

de maior  $RE^*$  para o serviço mMTC.

Na sequência, o critério de seleção da matriz PDMA foi aplicado ao serviço uRLLC. Foram selecionadas cinco matrizes padrão as quais apresentaram o menor fator de sobrecarga. Entre elas foram escolhidas as duas que apresentaram o menor fator de desempenho PEF. Finalmente, dentre as duas matrizes foi selecionada a matriz padrão que resultou na maior  $RE^*$  e menor número de usuários para o serviço uRLLC.

Finalmente, aplicou-se o critério de seleção da matriz padrão para acomodar a operação simultânea dos modos eMBB/uRLLC. Para esse modo de uso híbrido foi selecionado a matriz definida para o serviço uRLLC e que também aparecesse na lista de matrizes selecionadas pelo serviço eMBB. Como resultado, obtivemos uma matriz PDMA dentre as 15 disponíveis aquela que atendesse todos os requisitos do serviço uRLLC, tendo sido capaz de atender parcialmente os quesitos do serviço eMBB.

## 7 Conclusões e Perspectivas

O objeto principal deste trabalho de tese de doutorado foi atingido. Obteve-se o número de usuários por *clusters* via simulação computacional, o qual resulta na máxima eficiência energética com a restrição da mesma taxa de dados para todos os usuários em sistemas MIMO e MIMO-NOMA. Analisou-se a alocação de proporção de potência para cada usuário em sistemas NOMA e MIMO-NOMA ambos para o DL. No projeto de otimização obteve-se a potência total mínima e o fator de proporção de potência ótimos para cada usuário ativo. Além disso, a EE máxima do sistema, a taxa de dados igualitária, a potência total e a distribuição da proporção de potência entre os usuários dos *clusters* foram determinados sob um específico consumo de energia do circuito. Em ambos os procedimentos de otimização adotou-se as mesmas restrições e todos os usuários foram submetidos à mesma taxa de dados.

Para o sistema NOMA com dois ou mais usuários obteve-se a proporção de potência distribuída de maneira ótima em cada usuário para uma determinada taxa de transmissão. Para dois ou mais usuários também foi possível otimizar a potência total para a mesma taxa mínima. Utilizamos as técnicas de SQP que apresentam convergência rápida. Os resultados encontrados neste trabalho levam à potência total mínima para dois usuários ou mais para atingir a capacidade definida. A mesma metodologia de otimização foi executada para determinar a taxa máxima igualitária  $R$  para cada usuário em cada cenário de maior eficiência energética EE. Nesses projetos, a equidade, ou seja, a mesma taxa para todos os usuários, foi garantida para todos os usuários. Também encontramos o ponto do compromisso entre EE e a soma das taxas, ou seja, o sistema operando no ponto de melhor RE. Para que esses resultados sejam mais realistas, faz-se necessário adicionar mais restrições, uma vez que apenas a capacidade de Shannon para representar a taxa alcançável foi considerada.

O problema de otimização de recursos no enlace de descida do sistema MIMO-NOMA foi resolvido em termos da proporção de energia distribuída otimamente entre cada usuário para uma condição de equidade máxima que garante potência total mínima e EE máxima. A descrição do problema de otimização do vetor de potências mostrou-se expedito na obtenção da potência total mínima ótima e na distribuição adequada da proporção de potência entre os usuários NOMA. O problema de otimização de EE do sistema formulado foi analisado em duas partes. Primeiro, a EE máxima e a equidade máxima foram obtidas para distâncias fixas e taxas de dados para cada usuário. Os pontos de ajuste ideais para a taxa de dados variável quase otimizada comportaram-se de forma muito semelhante a uma análise de taxa de dados fixa. Na segunda parte, para um determinado raio de célula, o processo de análise considerou as distâncias obtidas aleatoriamente com distribuição

uniforme dos usuários em dois, três e quatro usuários por *cluster*. A taxa de dados foi uma variável de otimização obtida para um número diferente de *clusters* na célula. O objetivo era encontrar o número aproximado de *clusters* na célula que resulta em EE máximo ou, alternativamente, equidade máxima. Para os valores adotados de usuários por *cluster* ( $K$ ), número de *clusters* ( $A$ ), número de antenas BS por usuário ( $U \geq A$ ), expoente de perda de percurso ( $\alpha$ ), e cenários de cobertura de célula ( $\mathfrak{R}_D$ ), os resultados numéricos indicaram que o número de *clusters* por célula que maximiza o EE e o número de usuários com taxa de dados igual dependem do número de usuários por *cluster*, do fator do expoente de perda de percurso e do raio da célula. Notavelmente, com relação ao expoente de perda de percurso do canal, um valor geral de  $\alpha$  impacta (aumenta) o número de *clusters* necessários para atingir a maior eficiência energética do sistema, bem como para obter a equidade máxima.

As formulações analíticas e os resultados numéricos obtidos demonstraram que o projeto de otimização do compromisso EE-SE pode ser tratado adequadamente através do método  $\epsilon$ -C associado ao SQP dentro de uma abordagem bi-objetivo dada a natureza conflitante entre as métricas de EE e SE sob o ponto de vista de eficiência de recursos no sistema MIMO-NOMA. Além disso, os métodos  $\epsilon$ -C-DK e  $\epsilon$ -C-SQP foram capazes de encontrar soluções EE-SE na frente de Pareto. Como esperado, pode-se verificar através de extensos testes numéricos realizados que o método  $\epsilon$ -C foi capaz de encontrar as diferentes soluções na fronteira de Pareto.

Os valores obtidos com o método  $\epsilon$ -C-SQP vêm das características das funções componentes do compromisso bi-objetivo entre EE-SE. O método se mostrou mais adequado para esse tipo de problema de otimização. Além disso, os três métodos  $\epsilon$ -C-SQP,  $\epsilon$ -C-DK e  $\epsilon$ -C-HC derivados de  $\epsilon$ -C validaram o Teorema 1. Assim, se a restrição  $C.\epsilon$  for satisfeita como igualdade, então as soluções encontradas estão na fronteira de Pareto, significando que tais soluções são soluções eficientes.

As simulações para a otimização do compromisso entre EE-SE em sistemas MIMO-NOMA demonstraram que o sistema é mais eficiente quando há grandes diferenças entre o ganho de canal dos usuários no mesmo *cluster*. Isto significa que o princípio do domínio da potência no MIMO-NOMA pode ser explorado com vantagem. Além disso, um alto expoente de perda de percurso representando canais NLoS é uma maneira de realizar a diversidade de potência NOMA. Como resultado, para o ponto máximo de operação EE, maior valor do expoente de perda de percurso implica em um número maior de usuários com a mesma equidade de taxa de dados na célula. Em contraste, para um número menor de usuários por *cluster*, o EE máximo atingível aumenta.

Analisamos também as matrizes padrão que melhor atendem aos requisitos de serviços 5G como eMBB, mMTC e uRLLC. Além dos diferentes parâmetros das matrizes padrão, também analisamos a distribuição de energia entre os usuários que resulta em uma

taxa de uso igual para todos os usuários, atingindo a máxima equidade e o máximo EE, simultaneamente. Após o perfil das matrizes padrão, analisamos e definimos cada matriz que atende aos requisitos de cada modo de usuário dos serviços eMBB, mMTC e uRLLC.

Para o serviço eMBB um dos requisitos essenciais é a alta taxa de dados, enquanto para o serviço mMTC é primordial atender um número elevado de dispositivos do tipo máquina; finalmente, para o serviço uRLLC, o sistema deve combinar alta confiabilidade e baixa latência. Assim, os resultados numéricos revelam que é possível encontrar a melhor matriz padrão que atenda a cada um dos requisitos de cada um dos serviços separadamente ou em conjunto; Das quinze matrizes padrão PDMA avaliadas, o Modelo 1 é adequado para eMBB, enquanto o Modelo 3 para mMTC e o Modelo 10 para uRLLC. Para transmissão híbrida que atende simultaneamente usuários de eMBB e uRLLC, a matriz padrão que melhor atende a ambos os modos de uso é o Modelo 10.

## 7.1 Sugestões de Trabalhos Futuros

Muitos aspectos dos sistemas NOMA, NOMA-MIMO e mMIMO PDMA ainda podem ser abordados nessa linha de pesquisa. Algumas sugestões para trabalhos futuros são listadas a seguir:

1. Mais parâmetros podem ser acrescentados para avaliar as matrizes padrão em sistemas mMIMO PDMA que atendem aos serviços 5G como análise de desempenho e complexidade dos detectores em ambientes com ondas milimétricas. A análise de desempenho pode validar os resultados dos parâmetros que adotamos em nossa análise e os ambientes com ondas milimétricas são importantes para os sistemas 5G que trabalham entre as frequências de 24 e 60 GHz correspondem a ondas milimétricas.
2. Aprofundar a análise das matrizes padrão em sistemas mMIMO PDMA aplicáveis aos serviços mMTC, avaliando performance, complexidade e adequabilidade do esquema de transmissão resultante. Os serviços mMTC trabalham com conexões densas e o sistema PDMA é capaz de lidar muito bem com essas conexões.
3. Obter a curva da máxima EE *versus* a taxa de bit na condição de máxima equidade para sistemas MIMO NOMA com ondas milimétricas. Essa é uma extensão do trabalho adotando ondas milimétricas característica do 5G.

# Referências

- AL-OBIEDOLLAH, H.; CUMANAN, K.; THIYAGALINGAM, J.; BURR, A. G.; DING, Z.; DOBRE, O. A. Sum Rate Fairness Trade-Off-Based Resource Allocation Technique for MISO NOMA Systems. In: *2019 IEEE Wireless Communications and Networking Conference (WCNC)*. [S.l.]: IEEE, 2019. p. 1–6. ISBN 9781538676462. ISSN 23318422. Cited on page 7.
- AL-WANI, M. M.; SALI, A.; ALI, B. M.; SALAH, A. A.; NAVAIE, K.; LEOW, C. Y.; NOORDIN, N. K.; HASHIM, S. J. On Short Term Fairness and Throughput of User Clustering for Downlink Non-Orthogonal Multiple Access System. In: *IEEE 89th Vehicular Technology Conference*. Kuala Lumpur, Malaysia: IEEE, 2019. p. 1–6. ISBN 9781728112176. ISSN 15502252. Cited on page 7.
- ALI, M. S.; HOSSAIN, E.; KIM, D. I. Non-Orthogonal Multiple Access (NOMA) for Downlink Multiuser MIMO Systems: User Clustering, Beamforming, and Power Allocation. *IEEE Access*, v. 5, p. 565–577, mar 2017. ISSN 21693536. Cited 2 times on page(s) 6 e 7.
- ALI, M. S.; TABASSUM, H.; HOSSAIN, E. Dynamic User Clustering and Power Allocation for Uplink and Downlink Non-Orthogonal Multiple Access (NOMA) Systems. *IEEE Access*, v. 4, p. 6325–6343, 2016. Cited on page 28.
- ALVES, T. A. B.; ABRÃO, T. Massive MIMO and NOMA Bits-Per-Antenna Efficiency Under Power Allocation Policies. *Physical Communication*, v. 51, p. 101588, 2022. ISSN 1874-4907. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1874490721002858>>. Cited on page 5.
- BENJEBBOUR, A.; SAITO, Y.; KISHIYAMA, Y.; LI, A.; HARADA, A.; NAKAMURA, T. Concept and Practical Considerations of Non-Orthogonal Multiple Access (NOMA) for Future Radio Access. In: *Int. Symp. Intell. Signal Process. Commun. Syst. (ISPACS)*. Naha, Japan: [s.n.], 2013. p. 770–774. Cited on page 3.
- BOYD, S.; VANDENBERGHE, L. *Convex Optimization*. [S.l.]: Cambridge University Press, 2004. Cited 2 times on page(s) 20 e 30.
- BUI, V.-p.; NGUYEN, P. X.; NGUYEN, H. V.; NGUYEN, V.-d.; SHIN, O.-s. Optimal User Pairing for Achieving Rate Fairness in Downlink NOMA Networks. *2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)*, IEEE, p. 575–578, 2019. Cited on page 8.
- CAO, D.; CHEN, M. Using Penalty Function and Tabu Search to Solve Cell Formation Problems with Fixed Cell Cost. *Computers & Operations Research*, v. 31, n. 1, p. 21 – 37, 2004. ISSN 0305-0548. Cited on page 42.
- CARVALHO, M.; CARDOSO, A.; VIEIRA, F. H. Adaptive Power Allocation Based Beamforming Algorithm for Multiple Input and Multiple Output Non-Orthogonal Multiple Access System Throughput Optimization. *Transactions on Emerging Telecommunications Technologies*, v. 31, p. 1–18, 07 2020. Cited on page 6.

CHEN, S.; REN, B.; GAO, Q.; KANG, S.; SUN, S.; NIU, K. Pattern Division Multiple Access A Novel Non-Orthogonal Multiple Access for Fifth-Generation Radio Networks. *IEEE Transactions on Vehicular Technology*, v. 66, n. 4, p. 3185–3196, 2017. Cited 3 times on page(s) 4, 47 e 49.

CHEN, W.-E.; FAN, X.-Y.; CHEN, L.-X. A CNN-Based Packet Classification of eMBB, mMTC and URLLC Applications for 5G. In: *2019 International Conference on Intelligent Computing and its Emerging Applications (ICEA)*. [S.l.: s.n.], 2019. p. 140–145. Cited on page 11.

CHEN, X.; BENJEBBOUR, A.; LAN, Y.; LI, A.; JIANG, H. Impact of Rank Optimization on Downlink Non-Orthogonal Multiple Access (NOMA) with SU-MIMO. In: *2014 IEEE International Conference on Communication Systems, IEEE ICCS 2014*. [S.l.: s.n.], 2014. p. 233–237. ISBN 9781479958320. Cited on page 6.

CHOI, J. Non-Orthogonal Multiple Access in Downlink Coordinated Two-point Systems. *IEEE Communications Letters*, v. 18, n. 2, p. 313–316, 2014. ISSN 10897798. Cited on page 6.

DAI, L.; WANG, B.; YUAN, Y.; HAN, S.; I, C.-L.; WANG, Z. Non-Orthogonal Multiple Access for 5G: Solutions, Challenges, Opportunities, and Future Research Trends. *IEEE Communications Magazine*, v. 53, n. 9, p. 74–81, September 2015. Cited 2 times on page(s) 2 e 5.

DAI, X.; CHEN, S.; SUN, S.; KANG, S.; WANG, Y.; SHEN, Z.; XU, J. Successive Interference Cancellation Amenable Multiple Access (SAMA) for Future Wireless Communications. In: *2014 IEEE International Conference on Communication Systems*. [S.l.: s.n.], 2014. p. 222–226. Cited 4 times on page(s) 2, 3, 4 e 10.

DENG, R.; DI, B.; ZHANG, H.; SONG, L. HDMA: Holographic-Pattern Division Multiple Access. *IEEE Journal on Selected Areas in Communications*, v. 40, n. 4, p. 1317–1332, 2022. Cited on page 10.

DEVROYE, L. *Non-Uniform Random Variate Generation*. [S.l.]: Springer Science+Business Media, LLC, 1986. Cited 2 times on page(s) 64 e 75.

DHAKAL, S.; MARTIN, P. A.; SMITH, P. J. NOMA with Guaranteed Weak User QoS: Design and Analysis. *IEEE Access*, IEEE, v. 7, p. 32884–32896, 2019. ISSN 21693536. Cited on page 7.

DING, J.; CAI, J. Efficient MIMO-NOMA Clustering Integrating Joint Beamforming and Power Allocation. *2017 IEEE Global Communications Conference, GLOBECOM 2017 - Proceedings*, p. 1–6, 2018. Cited on page 7.

DING, Z.; ADACHI, F.; POOR, H. V. The Application of MIMO to Non-Orthogonal Multiple Access. *IEEE Transactions on Wireless Communications*, v. 15, n. 1, p. 537–552, 2016. ISSN 15361276. Cited 3 times on page(s) 5, 7 e 27.

DING, Z.; SCHÖBER, R.; POOR, H. V. A General MIMO Framework for NOMA Downlink and Uplink Transmission Based on Signal Alignment. *IEEE Transactions on Wireless Communications*, v. 15, n. 6, p. 4438–4454, 2016. ISSN 15361276. Cited on page 6.

- DING, Z.; YANG, Z.; FAN, P.; POOR, H. V. On the Performance of Non-orthogonal Multiple Access in 5G Systems with Randomly Deployed Users. *IEEE Signal Processing Letters*, v. 21, n. 12, p. 1501–1505, 2014. ISSN 10709908. Cited 2 times on page(s) 17 e 18.
- EHRGOTT, M.; RUZIKA, S. Improved  $\epsilon$ -Constraint Method for Multiobjective Programming. *Journal of Optimization Theory and Applications*, Springer New York, v. 138, n. 3, p. 375–396, 2008. Cited on page 37.
- EL-SAYED, M. M.; IBRAHIM, A. S.; KHAIRY, M. M. Power Allocation Strategies for Non-Orthogonal Multiple Access. In: *2016 International Conference on Selected Topics in Mobile and Wireless Networking, MoWNeT 2016*. [S.l.: s.n.], 2016. p. 1–6. ISBN 9781509017430. Cited 2 times on page(s) 29 e 58.
- ELGHARIANI, A.; ZOLTOWSKI, M. Low Complexity Detection Algorithms in Large-Scale MIMO Systems. *IEEE Transactions on Wireless Communications*, v. 15, n. 3, p. 1689–1702, 2016. ISSN 15361276. Cited on page 75.
- GUI, G.; SARI, H.; SARI, H.; BIGLIERI, E. A New Definition of Fairness for Non-Orthogonal Multiple Access. *IEEE Communications Letters*, IEEE, v. 23, n. 7, p. 1267–1271, 2019. ISSN 1089-7798. Disponível em: <<https://ieeexplore.ieee.org/document/8713476/>>. Cited on page 24.
- HIGUCHI, K.; KISHIYAMA, Y. Non-Orthogonal Access with Random Beamforming and Intra-Beam SIC for Cellular MIMO Downlink. In: *2013 IEEE 78th Vehicular Technology Conference (VTC Fall)*. [S.l.: s.n.], 2013. p. 1–5. Cited 3 times on page(s) 2, 54 e 55.
- HOSHYAR, R.; WATHAN, F. P.; TAFAZOLLI, R. Novel low-density signature for synchronous cdma systems over awgn channel. *IEEE Transactions on Signal Processing*, v. 56, n. 4, p. 1616–1626, 2008. Cited on page 2.
- HUANG, J.; PENG, K.; PAN, C.; YANG, F.; JIN, H. Scalable Video Broadcasting Using Bit Division Multiplexing. *IEEE Transactions on Broadcasting*, v. 60, n. 4, p. 701–706, 2014. Cited on page 2.
- HUANG, Y.; HE, S.; WANG, J.; ZHU, J. Spectral and Energy Efficiency Tradeoff for Massive MIMO. *IEEE Transactions on Vehicular Technology*, IEEE, v. 67, n. 8, p. 6991–7002, 2018. ISSN 00189545. Cited on page 23.
- ISLAM, S. M.; AVAZOV, N.; DOBRE, O. A.; KWAK, K. S. Power-Domain Non-Orthogonal Multiple Access (NOMA) in 5G Systems: Potentials and Challenges. *IEEE Communications Surveys and Tutorials*, v. 19, n. 2, p. 721–742, 2017. ISSN 1553877X. Cited 2 times on page(s) 1 e 3.
- JACOB, J. L.; ABRÃO, T. Nonorthogonal Multiple Access Systems Optimization to Ensure Maximum Fairness to Users. *Transactions on Emerging Telecommunications Technologies*, v. 31, n. 4, p. 1–14, 2020. ISSN 21613915. Cited on page 37.
- JACOB, J. L.; MARTINEZ, C. A. P.; MARTINEZ, A. L. M.; ABRÃO, T. Non-Linear Biobjective EE-SE Optimization for NOMA-MIMO Systems Under User-Rate Fairness and Variable Number of Users Per Cluster. *AEU - International Journal of Electronics and Communications*, v. 138, p. 153870, 2021. ISSN 1434-8411. Cited on page 57.

- JACOB, J. L.; PANAZIO, C. M.; ABRÃO, T. Energy and Spectral Efficiencies Trade-Off in MIMO-NOMA System Under User-Rate Fairness and Variable User Per Cluster. *Physical Communication*, v. 47, p. 1–14, 05 2021. Cited on page 57.
- JAIN, R. K.; CHIU, D.-M. W.; HAWKES, W. R. A Quantitative Measure of Fairness and Discrimination for Resource Allocation in Shared Computer System. *ACM Transaction on Computer Systems*, p. 1–38, 1984. Cited 2 times on page(s) 23 e 29.
- JI, B.; SONG, K.; LI, C.; ZHU, W. ping; YANG, L. Energy Harvest and Information Transmission Design in Internet-of-Things Wireless Communication Systems. *AEU - International Journal of Electronics and Communications*, v. 87, p. 124–127, 2018. ISSN 1434-8411. Cited on page 8.
- JI, Z.; JI, Y.; WANG, B.; GAO, F.; WANG, H.; ZHANG, C. A New Uplink Channel Estimation Architecture for Massive MIMO Systems with PDMA. In: *2019 IEEE 13th International Conference on ASIC (ASICON)*. [S.l.: s.n.], 2019. p. 1–4. Cited on page 10.
- JIANG, Y.; LI, P.; DING, Z.; ZHENG, F.-C.; MA, M.; YOU, X. Joint Transmitter and Receiver Design for Pattern Division Multiple Access. *IEEE Transactions on Mobile Computing*, v. 18, n. 4, p. 885–895, 2019. Cited 6 times on page(s) 4, 5, 12, 44, 45 e 55.
- JOHNSON, A.; JACOBSON, S. On the Convergence of Generalized Hill Climbing Algorithms. *Discrete Applied Mathematics*, v. 119, n. 1, p. 37 – 57, 2002. Cited on page 42.
- KANG, J.-M.; KIM, I.-M.; CHUN, C.-J. Short Papers Deep Learning-Based MIMO-NOMA with Imperfect SIC Decoding. *IEEE Systems Journal*, IEEE, v. 14, n. 3, p. 3414–3417, 2020. Cited on page 6.
- KIM, B.; LIM, S.; KIM, H.; SUH, S.; KWUN, J.; CHOI, S.; LEE, C.; LEE, S.; HONG, D. Non-Orthogonal Multiple Access in a Downlink Multiuser Beamforming System. In: *Proceedings - IEEE Military Communications Conference MILCOM*. [S.l.: s.n.], 2013. p. 1278–1283. ISBN 9780769551241. Cited on page 6.
- KONG, D.; ZENG, J.; SU, X.; RONG, L.; XU, X. Multiuser Detection Algorithm for PDMA Uplink System Based on SIC and MMSE. In: . [S.l.: s.n.], 2016. p. 1–5. Cited on page 47.
- LI, L.; MA, H.; REN, H.; CHENG, Q.; WANG, D.; BAI, T.; HAN, Z. Learning-Aided Resource Allocation for Pattern Division Multiple Access-Based SWIPT Systems. *IEEE Wireless Communications Letters*, v. 10, n. 1, p. 131–135, 2021. Cited on page 11.
- LI, M.; ZHANG, S.; GAO, F.; FAN, P.; DOBRE, O. A. A New Path Division Multiple Access for the Massive MIMO-OTFS Networks. *IEEE Journal on Selected Areas in Communications*, v. 39, n. 4, p. 903–918, 2021. Cited on page 10.
- LI, P.; JIANG, Y.; KANG, S.; ZHENG, F.; YOU, X. Pattern Division Multiple Access with Large-Scale Antenna Array. In: *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*. [S.l.: s.n.], 2017. p. 1–6. Cited 4 times on page(s) 4, 5, 12 e 44.
- LI, S.; SU, B.; JIN, L. Research on PDMA Communication System Based on Complete Complementary Sequence. *EURASIP Journal on Wireless Communications and Networking*, p. 1–18, 2020. Cited on page 4.

- LI, S.; SU, B.; SUN, Y.; SUN, S.; JIN, L. Uplink Grant-Free Pattern Division Multiple Access Transmission Scheme by Exploiting Poly Complementary Sequence. *Transactions on Emerging Telecommunications Technologies*, n/a, n. n/a, p. 1–18, 2021. Cited 7 times on page(s) [xi](#), [3](#), [10](#), [53](#), [54](#), [90](#) e [91](#).
- LIU, F.; YANG, Q.; HE, Q.; KWAK, K. S. Energy Efficiency and Spectral Efficiency Tradeoff in Downlink OFDMA Systems with Imperfect CSI. *AEU - International Journal of Electronics and Communications*, v. 85, p. 54 – 58, 2018. Cited on page [8](#).
- LIU, Q.; TAN, F.; LV, T.; GAO, H. Energy Efficiency and Spectral-Efficiency Tradeoff in Downlink NOMA Systems. In: *2017 IEEE International Conference on Communications Workshops, ICC Workshops 2017*. [S.l.]: IEEE, 2017. p. 247–252. ISBN 9781509015252. Cited on page [8](#).
- LIU, Y.; PAN, G.; ZHANG, H.; SONG, M. On the Capacity Comparison Between MIMO-NOMA and MIMO-OMA. *IEEE Access*, IEEE, v. 4, p. 2123–2129, 2016. ISSN 21693536. Cited on page [6](#).
- MARTINEZ, C. A. P.; ABRÃO, T. Spectral and Energy Efficiency Tradeoff in Optical Code Division Multiple Access Networks. *Transactions on Emerging Telecommunications Technologies*, v. 32, p. 1–20, 2020. Cited 9 times on page(s) [8](#), [9](#), [36](#), [37](#), [38](#), [40](#), [42](#), [43](#) e [80](#).
- MOLTAFET, M.; MOKARI, N.; JAVAN, M. R.; SAEEDI, H.; PISHRO-NIK, H. A New Multiple Access Technique for 5G: Power Domain Sparse Code Multiple Access (PSMA). *IEEE Access*, v. 6, p. 747–759, 2018. Cited on page [2](#).
- NI, D.; HAO, L.; QIAN, X.; TRAN, Q. T. Energy-Spectral Efficiency Tradeoff of Downlink NOMA System with Fairness Consideration. In: *IEEE Vehicular Technology Conference*. [S.l.]: IEEE, 2018. v. 2018-June, p. 1–5. ISBN 9781538663554. ISSN 15502252. Cited on page [8](#).
- NIKOPOUR, H.; BALIGH, H. Sparse Code Multiple Access. In: *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. [S.l.: s.n.], 2013. p. 332–336. Cited on page [2](#).
- NOCEDAL, J.; WRIGHT, S. *Numerical Optimization*. Second Edition. [S.l.]: Springer Science+Business Media, LLC, 1996. v. 17. 685 p. ISSN 978-0387-30303-1. ISBN 0-387-30303-0. Cited 3 times on page(s) [41](#), [66](#) e [80](#).
- POPOVSKI, P.; TRILLINGSGAARD, K. F.; SIMEONE, O.; DURISI, G. 5G Wireless Network Slicing for eMBB, URLLC, and mMTC: A Communication-Theoretic View. *IEEE Access*, v. 6, p. 55765–55779, 2018. Cited on page [12](#).
- REBHI, M.; HASSAN, K.; RAOOF, K.; CHARGÉ, P. Sparse Code Multiple Access: Potentials and Challenges. *IEEE Open Journal of the Communications Society*, v. 2, p. 1205–1238, 2021. Cited on page [1](#).
- REN, B.; WANG, Y.; DAI, X.; NIU, K.; TANG, W. Pattern Matrix Design of PDMA for 5G UL Applications. *China Communications*, v. 13, n. Supplement2, p. 159–173, 2016. Cited 5 times on page(s) [xi](#), [51](#), [53](#), [90](#) e [91](#).

RUAN, Y.; ZHANG, R.; LI, Y.; WANG, C.; ZHANG, H. Spectral-Energy Efficiency Tradeoff in Cognitive Satellite-Vehicular Networks Towards Beyond 5G. p. 1–6, 2019. Cited 2 times on page(s) 8 e 9.

RUSSELL, S.; NORVIG, P. Artificial Intelligence - A Modern Approach. *NJ: Prentice Hall*, 2012. Cited on page 42.

SAITO, Y.; KISHIYAMA, Y.; BENJEBBOUR, A.; NAKAMURA, T.; LI, A.; HIGUCHI, K. Non-Orthogonal Multiple Access (NOMA) for Cellular Future Radio Access. In: *2013 IEEE 77th Vehicular Technology Conference (VTC Spring)*. [S.l.: s.n.], 2013. p. 1–5. Cited on page 2.

SANTOS, H. L. dos; MARINELLO, J. C.; PANAZIO, C. M.; AO, T. A. Machine Learning-Aided Pilot and Power Allocation in Multi-Cellular Massive MIMO Networks. *Physical Communication*, v. 52, p. 101646, 2022. ISSN 1874-4907. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1874490722000295>>. Cited on page 5.

SASIKUMAR, S.; JAYAKUMARI, J. A Novel Method for the Optimization of Spectral-Energy Efficiency Tradeoff in 5G Heterogeneous Cognitive Radio Network. *Computer Networks*, v. 180, 2020. Cited on page 8.

SAYED-AHMED, A.; ELSABROUTY, M.; EL-MALEK, A. H. A.; ABO-ZAHHAD, M. Energy Efficient Framework for Multiuser Downlink MIMO-NOMA Systems. In: *2018 Eleventh International Workshop on Selected Topics in Mobile and Wireless Computing*. [S.l.]: Networking and Communications (WiMob), 2018. p. 48–54. Cited 2 times on page(s) 5 e 7.

SHIN, W.; VAEZI, M.; LEE, B.; LOVE, D. J.; LEE, J.; POOR, H. V. Coordinated Beamforming for Multi-Cell MIMO-NOMA. *IEEE Communications Letters*, v. 21, n. 1, p. 84–87, 2017. ISSN 10897798. Cited on page 6.

SONG, Z.; NI, Q.; SUN, X. Spectrum and Energy Efficient Resource Allocation with QoS Requirements for Hybrid MC-NOMA 5G systems. *IEEE Access*, v. 6, p. 37055–37069, 2018. Cited 2 times on page(s) 8 e 9.

SOUZA, J. a. H. I. de; AMIRI, A.; ABRÃO, T.; CARVALHO, E. de; POPOVSKI, P. Quasi-Distributed Antenna Selection for Spectral Efficiency Maximization in Subarray Switching XL-MIMO Systems. *IEEE Transactions on Vehicular Technology*, v. 70, n. 7, p. 6713–6725, 2021. Cited on page 5.

STOREY, C. Applications of a Hill Climbing Method of Optimization. *Chemical Engineering Science*, v. 17, n. 1, p. 45 – 52, 1962. ISSN 0009-2509. Cited on page 42.

SUN, J.; WANG, C.; ZENG, J.; SU, X.; LV, T. Design of PDMA Pattern Matrix in 5G Scenarios. In: *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*. [S.l.: s.n.], 2020. p. 1–6. Cited 9 times on page(s) xi, 4, 12, 51, 52, 53, 55, 90 e 91.

SUN, Q.; HAN, S.; I, C.-L.; PAN, Z. On the Ergodic Capacity of MIMO NOMA Systems. *IEEE Wireless Commun. Lett.*, v. 4, n. 4, p. 405–408, 2015. Cited on page 7.

TANG, J.; LUO, J.; LIU, M.; SO, D. K.; ALSUSA, E.; CHEN, G.; WONG, K. K.; CHAMBERS, J. A. Energy Efficiency Optimization for NOMA with SWIPT. *IEEE Journal on Selected Topics in Signal Processing*, IEEE, v. 13, n. 3, p. 452–466, 2019. ISSN 19324553. Cited 2 times on page(s) 20 e 32.

- TANG, S.; MA, Z.; XIAO, M.; HAO, L. Hybrid Transceiver Design for BeamSpace MIMO-NOMA in Code-Domain for MmWave Communication Using Lens Antenna Array. *IEEE Journal on Selected Areas in Communications*, v. 38, n. 9, p. 2118–2127, 2020. Cited on page 10.
- TSE, D.; PRAMOD, V. *Fundamentals of Wireless Communication*. [S.l.]: Cambridge University Press, 2005. v. 9780521845. 1-564 p. ISBN 9780511807213. Cited on page 5.
- UBIALI, G. A.; MARINELLO, J. C.; ABRÃO, T. Energy-Efficient Flexible and Fixed Antenna Selection Methods for XL-MIMO Systems. *AEU - International Journal of Electronics and Communications*, v. 130, p. 153568, 2021. ISSN 1434-8411. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1434841120327722>>. Cited on page 5.
- UDALCOVS, A.; SCHATZ, R.; WOSINSKA, L.; MONTI, P. Analysis of Spectral and Energy Efficiency Tradeoff in Single-Line Rate WDM Links. *J. Lightwave Technol., OSA*, v. 35, n. 10, p. 1847–1857, 2017. Cited on page 8.
- VAEZI, M.; DING, Z.; POOR, H. V. *Multiple Access Techniques for 5G Wireless Networks and Beyond*. [S.l.]: Springer International Publishing AG, 2019. Cited 5 times on page(s) 11, 47, 48, 49 e 50.
- WANG, C. H.; LIN, J. Y.; WU, J. M. Joint Fairness and Sum Rate Resource Allocation for NOMA Communications. In: *2017 IEEE Conference on Standards for Communications and Networking, CSCN 2017*. [S.l.: s.n.], 2017. p. 269–274. ISBN 9781538630709. Cited 2 times on page(s) 23 e 29.
- WANG, D.; ZHANG, Y.; WEI, H.; YOU, X.; WANG, J. An Overview of Transmission Theory and Techniques of Large-Scale Antenna Systems for 5G Wireless Communications. In: *Science CHINA Information Sciences*. [S.l.: s.n.], 2016. v. 59, n. 8, p. 1–23. Cited 3 times on page(s) 5, 25 e 44.
- WANG, H.; ZHANG, R.; SONG, R.; LEUNG, S. H. A Novel Power Minimization Precoding Scheme for MIMO-NOMA Uplink Systems. *IEEE Communications Letters*, v. 22, n. 5, p. 1106–1109, 2018. ISSN 10897798. Cited 2 times on page(s) 6 e 10.
- YUAN, Z.; YU, G.; LI, W.; YUAN, Y.; WANG, X.; XU, J. Multi-user shared access for internet of things. In: *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*. [S.l.: s.n.], 2016. p. 1–5. Cited on page 2.
- ZENG, J.; FAN, X.; ZHONG, Y.; PENG, M.; LV, T.; JIANG, T. Downlink Power Allocation Optimization in Pattern Division Multiple Access. *IEEE Access*, v. 9, p. 14649–14659, 2021. Cited on page 10.
- ZENG, J.; LI, B.; SU, X.; RONG, L.; XING, R. Pattern Division Multiple Access (PDMA) for Cellular Future Radio Access. In: *2015 International Conference on Wireless Communications Signal Processing (WCSP)*. [S.l.: s.n.], 2015. p. 1–5. Cited 2 times on page(s) 3 e 10.
- ZENG, M.; YADAV, A.; DOBRE, O. A.; POOR, H. V. A Fair Individual Rate Comparison Between MIMO-NOMA and MIMO-OMA. In: *2017 IEEE Globecom Workshops, GC Wkshps 2017 - Proceedings*. [S.l.: s.n.], 2017. p. 1–5. ISBN 9781538639207. Cited on page 7.

- ZENG, M.; YADAV, A.; DOBRE, O. A.; TSIROPOULOS, G. I.; POOR, H. V. Capacity Comparison Between MIMO-NOMA and MIMO-OMA with Multiple Users in a Cluster. *IEEE Journal on Selected Areas in Communications*, v. 35, n. 10, p. 2413–2424, 2017. ISSN 07338716. Cited on page 26.
- ZENG, M.; YADAV, A.; DOBRE, O. A.; POOR, H. V. Energy-Efficient Power Allocation for MIMO-NOMA with Multiple Users in a Cluster. *IEEE Access*, v. 6, p. 5170–5181, 2018. ISSN 21693536. Cited 2 times on page(s) 25 e 26.
- ZENG, Y.; ZHANG, R. Millimeter Wave MIMO with Lens Antenna Array: A New Path Division Multiplexing Paradigm. *IEEE Transactions on Communications*, v. 64, n. 4, p. 1557–1571, 2016. Cited on page 10.
- ZHANG, S.; ZHANG, N.; KANG, G.; LIU, Z. Energy and Spectrum Efficient Power Allocation with NOMA in Downlink HetNets. *Physical Communication*, v. 31, p. 121 – 132, 2018. Cited 2 times on page(s) 8 e 9.
- ZHANG, Y.; WANG, H. M.; ZHENG, T. X.; YANG, Q. Energy-Efficient Transmission Design in Non-Orthogonal Multiple Access. *IEEE Transactions on Vehicular Technology*, IEEE, v. 66, n. 3, p. 2852–2857, 2017. ISSN 00189545. Cited 3 times on page(s) 7, 20 e 32.
- ZUO, Y.; ZHU, X.; JIANG, Y.; WEI, Z.; ZENG, H.; WANG, T. Energy Efficiency and Spectral Efficiency Tradeoff for Multicarrier NOMA Systems with User Fairness. *2018 IEEE/CIC International Conference on Communications in China (ICCC)*, p. 666–670, 2018. Cited 2 times on page(s) 8 e 9.

# **Anexos**

**ANEXO A – Nonorthogonal  
Multiple Access Systems  
Optimization to Ensure Maximum  
Fairness to Users**

(P1 – artigo publicado em 2019)

# Nonorthogonal multiple access systems optimization to ensure maximum fairness to users

Jaime L. Jacob  | Taufik Abrão 

Department of Electrical Engineering,  
State University of Londrina, Londrina,  
Brazil

## Correspondence

Taufik Abrão, Department of Electrical  
Engineering, State University of Londrina,  
Londrina, Parana 86057-970, Brazil.  
Email: taufik@uel.br

## Abstract

In this contribution, the optimization of power proportion allocated for each user in the downlink (DL) nonorthogonal multiple access (NOMA) systems have been developed. Successive interference cancellation technique recovers users' signal with high difference between channel gains, to find the lowest optimum power proportion required to guarantee an equal data rate for all active users. Following the same approach, the optimum power proportion for each user and also the minimum total power to achieve the same rate for all users (maximum fairness) were obtained as a design goal. Moreover, the same design methodology was developed seeking to maximize the NOMA system energy efficiency (EE). It was possible to find the maximum EE point and the respective power distribution among the users for a certain circuitry power consumption. For all NOMA users, in which the optimal operation point for EE maximization was parameterized, it was possible to find the total power, power ratio, and the equal rate values. As a result, one can find the trade-off point between EE and sum rate for each system scenario, as well as the best resource efficiency operation point. By considering the same rate for all users, the system attains maximum fairness among the users.

## 1 | INTRODUCTION

Nonorthogonal multiple access (NOMA) system is a promising technique for 5G technology by optimizing bandwidth usage where users can share it, as they have different power ratings. As a result, users at the edge of the cell will be treated under the same conditions as cell users themselves in wireless cellular systems.<sup>1</sup> To reach this goal, some studies have been done analyzing the necessary conditions for the use of NOMA system. For instance, one can maximize the symbol transmission rate of two users in a cluster<sup>2</sup> aiming at maximizing the sum capacity of them by deploying beamforming technique in the base station (BS). However, is two users per cluster the ideal configuration? In addition, the feedback rank is used to determine the NOMA application in cases where users with the best rank are selected while other users continue using the orthogonal multiple access (OMA).<sup>3</sup> Note that those authors do not consider the optimal power distribution. Instead of setting a single BS, one can use two BS and apply the Alamouti code combined with the successive interference cancellation (SIC) technique to recover the signal at reception.<sup>4</sup>

Abbreviations: BS, base station; DL, downlink; EE, energy efficiency; NOMA, nonorthogonal multiple access; OMA, orthogonal multiple access; QoS, quality of service; RF, radio frequency; SE, spectral efficiency; SIC, successive cancellation of interference; SINR, signal-to-interference-plus-noise power ratio; SQP, sequential quadratic programming.

Another approach to improve NOMA capacity consists in separate the users into three categories (singletons, weak users, and strong users) and to analyze the transmission rate and signal-to-interference-plus-noise power ratio (SINR) for the three categories, emphasizing the quality of service (QoS) of weak users.<sup>5</sup> Such work analyzed only two users per cluster. Furthermore, other authors<sup>6</sup> developed energy efficiency (EE) optimization and got the green point, that is, the optimum system operation point that maximizes the overall EE, but they did not consider fairness among users. The trade-off between EE and spectral efficiency (SE) was studied in Reference 7, but the authors did not analyze different numbers of users. Moreover, the maximization of the minimum rate for each user was analyzed in Reference 8, but despite mentioning fairness such concept was not systematically explored. Besides, the trade-off between EE and SE considering fairness has been analyzed in Reference 9, but it did not guarantee maximum fairness among the users, and the calculation used the traditional concept of Jain's fairness.<sup>10</sup> Currently, the concept of fairness based on the considerations of information theory<sup>11</sup> has become popular.

A research aimed to compare the uplink (UL) NOMA system with OMA using resource allocation fairness as a reference and considering only Jain's index was reported in Reference 12. The authors analyzed fairness for only two users. In another research<sup>13</sup> an optimal power allocation scheme in the system was investigated from the concept of  $\alpha$ -fairness with the solution of the ergodic rate maximization and the sum of capacity problem in a downlink (DL) NOMA system. A waterfilling based power allocation technique has been applied to the DL NOMA combined with proportional fairness scheduler to maximize the average rate with near-optimal power distribution between subbands and high fairness.<sup>14</sup> In other analysis, the trade-off between EE and SE from the combination of orthogonal frequency division multiplexing access (OFDMA) and NOMA systems were made from the user's fairness restriction.<sup>15</sup> Resource efficiency (RE) for different weights between EE and SE and different maximum transmission powers were analyzed. Another article<sup>16</sup> analyzed the average sum rates, together or not, with the maximization of the sum of the delay-limited rates. This analysis is subject to user fairness to obtain optimal power control. This is all done for a two-user DL NOMA system over a fading channel. Research in Reference 17 has addressed the power allocation problem that maximizes the minimum rate obtained among users. In addition, by simulation, the minimum user rate was compared in terms of the Jain's fairness index for different power allocation schemes in NOMA vs OMA systems. Authors 18 have analyzed the jointly optimization of EE and fairness among users with respect to subcarriers and the power allocation of a DL multicarrier (MC) NOMA system. The authors compare the EE for MC-NOMA with equal power, difference of convex programming<sup>19</sup> and their proposed convex programming sequence for the best link and the worst link and combination. Furthermore, a power allocation strategy for two users in a UL NOMA system with proportional fairness guarantees has been proposed in Reference 20 where two scenarios analyzed: basic scenario users are distributed inside the cell, and in the complex scenario where the interfering users are placed outside the cell. Contribution: This work establishes the best power allocation policy in NOMA system in terms of sum-power allocation coefficients, which guarantees of equal rate (fairness approach) for all users. For that, we deploy convex optimization techniques,<sup>21,22</sup> to solve the optimization problem. Differently of Reference 23, in which the power proportion allocated for each user is found, herein we have further considered the EE maximization and the fairness among users in terms of equal data rate guarantee. Indeed, this article also considers the EE-NOMA optimization problem by finding the minimum total power needed to reach a specified equal rate per user. From the jointly EE and sum rate optimization with maximum fairness guarantee optimization problem, we have find the best EE-SE trade-off and thereby obtain get the best NOMA system configuration in terms of RE.

## 2 | NOMA SYSTEM

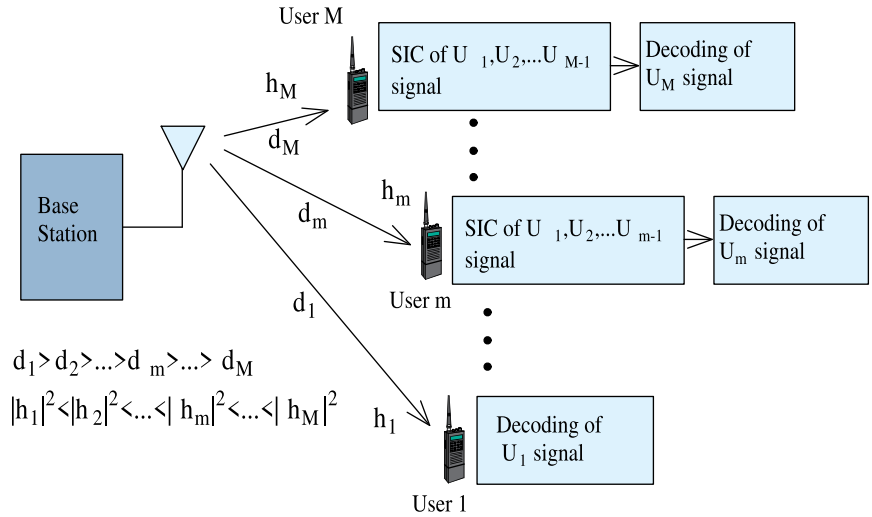
Consider a cellular system with downlink transmission scenario, in which one BS located in the center of the disc with radius  $\mathfrak{R}_D$  and  $M$  users aligned, Figure 1. All terminals are equipped with a single antenna.

The channel between the users and the BS<sup>24</sup> is given by\*

$$h_i = \frac{g_i}{\sqrt{1 + d_i^\alpha}} \text{ for } i = 1, 2, \dots, M, \quad (1)$$

\*It is adding 1 in the channel gain denominator, Equation (1), it is warranted that at the 0 distance, the transmitted signal is equal to the received signal. Indeed, this is one of the possible pathloss models found in the literature, for instance, in Reference 25.

**FIGURE 1** Nonorthogonal multiple access with  $M$  users



where  $g_i$  represents the channel gain with Rayleigh fading,  $\alpha$  is the path loss factor, and  $d_i$  represents the distance from the user to the BS.<sup>25</sup> Without loss of generality, the channels are sorted as  $|h_1|^2 \leq |h_2|^2 \leq \dots \leq |h_M|^2$ . The transmission signal in the BS with application of the NOMA technique is given by

$$x = \sum_{i=1}^M \sqrt{\beta_i P} s_i, \quad (2)$$

where  $s_i$  is message from  $i$ th user,  $P$  is the transmission power, and  $\beta_i$  is the power allocation coefficient, that is,  $\beta_1 \geq \beta_2 \geq \dots \geq \beta_M$ . There are the superposition of  $M$  signals to its  $M$  users via power-domain division. The signal received at the  $m$ th subscriber is given by

$$y_m = h_m \sum_{i=1}^M \sqrt{\beta_i P} s_i + z_m, \quad (3)$$

where  $z_i$  is i.i.d. additive white complex Gaussian noise with zero mean and variance  $\sigma_n^2$ . In the system model presented here, only user 1 does not use the SIC technique in signal detection. The SIC technique is initially used in user 2, where user 1 has its interference fully extracted and the other users are treated as noise. The privilege of the user closest to BS, that is, the  $M$ th user is to have the interfering signals fully extracted. To calculate the capacity, it is first necessary to calculate the received SINR. From Equation (3), we can obtain the generic SINR for the  $1 \leq m \leq (M - 1)$  given by

$$\gamma_m = \frac{P|h_m|^2\beta_m}{P|h_m|^2 \sum_{i=m+1}^M \beta_i + \sigma_n^2}, \quad m = 1, \dots, M - 1, \quad (4)$$

and for the last (near-BS) user results:

$$\gamma_M = \frac{P|h_M|^2\beta_M}{\sigma_n^2}. \quad (5)$$

The channel capacity is given by  $\log_2(1 + SINR)$ , so the capacity of users are given by Reference 24

$$R_m = \log_2(1 + \gamma_m) \quad m = 1, \dots, M \quad (\text{bits/s/Hz}). \quad (6)$$

Herein, it is assumed that the interference from the first user has been completely canceled when using the SIC technique.

EE is given by the ratio between the sum of rates achievable by all users in system and the total consumed power, given by

$$EE = \frac{R_s}{P_t + MP_c} = \frac{MR}{P_t + MP_c}, \quad \left( \frac{\text{bits}}{J \cdot \text{Hz}} \right), \quad (7)$$

where  $R_s = \sum_{m=1}^M R_m = MR$  is the sum of the rates transmitted by all the users given by Equation (6) where all users have the same capacity, that is,  $R_m = R$ ; the total RF power consumed is given by:

$$P_t = \rho \cdot \sum_{i=1}^M \beta_i P_c \quad (W), \quad (8)$$

where  $\rho \geq 1$  is the RF power amplifier inefficiency, assumed equal for all of them feeding each Tx antenna;  $\beta_i$  is the power allocation coefficient, and  $P_c$  is the circuitry power required to signal processing at the transmitter side, assumed be fixed value per user.

### 3 | DL NOMA OPTIMIZATION

In this section, the optimization of the NOMA system is done in two ways. In the same optimization process, the goal is to get the minimum proportion of power for each user needed to attain a specific target equal rate (maximum fairness), and still lower total transmit power to obtain such same minimum fair data rate. This goal is achieved by simultaneously maximizing the EE via power transmitting optimization and the proportion of power among users.

In the sequel, we discuss the present the fairness index, discuss the solution for the ratio of power optimization problem, as well as the power allocation solution under the perspective of EE optimization. Finally, the concept of RE is revisited.

#### 3.1 | Fairness indexes

Fairness index can be obtained by the formula based on the Jain's fairness index,<sup>10,26</sup> as well as making considerations based on information theory. Hence, from the conventional definition of fairness, the Jain's fairness index can be written as

$$\mathcal{F}_J = \frac{\left(\sum_{k=1}^M R_k\right)^2}{M \sum_{k=1}^M R_k^2}. \quad (9)$$

In the optimization design context of this article the same rate has been adopted for all the users; hence, it can be verified by Equation (9) that under this condition, the calculated fairness results in  $\mathcal{F}_J = 1$  being its maximum value. For illustration purpose, let us consider the following toy example: assuming  $M = 4$  users, with  $R_1 = R_2 = R_3 = 1$  and  $R_4 = 0.7$ , resulting in  $\left(\sum_{k=1}^4 R_k\right)^2 = 13.69$  and  $4 \cdot \sum_{k=1}^4 R_k^2 = 13.96$ ; so,  $\mathcal{F}_J = 0.98$  results very close to 1, as only one user presents data rate different than one.

One alternative way to define fairness is based purely on information theory considerations, as discussed in Reference 11, being expressed by:

$$\mathcal{F}_{IT} = 1 - \frac{\hat{\Delta}^2}{\hat{R}^2}, \quad (10)$$

where  $\hat{\Delta}^2 = \frac{1}{M-1} \sum_{m=1}^M \Delta_m^2$  is the unbiased estimate of the mean-square rate deviations and  $\hat{R}^2 = \frac{R_s^2}{M}$  is average square sum rate. The  $\Delta_m^2$  is the squared deviations of the actual user rates from their fair rate ( $R^f$ ) as follows:

$$\Delta_m^2 = \left(R_m^f - R_m\right)^2, \quad m = 1, 2, \dots, M. \quad (11)$$

The fair rate for the  $m$ th user is defined as can be represented by

$$R_m^f = \log_2 \left( 1 + \frac{P |h_m| \beta_m}{\sigma_n^2} \right) \cdot \frac{R_s}{R_c}, \quad (12)$$

where the cumulative rate is given by

$$R_c = \sum_{m=1}^M \log_2 \left( 1 + \frac{P |h_m| \beta_m}{\sigma_n^2} \right). \quad (13)$$

The authors in Reference 11 justify the cumulative rate as the sum rates of an OMA system that allows a separate channel for each user. Calculating fairness in this way, the index of Equation (10) reaches the maximum value of 1 when both condition are achieved, a uniform power distribution and the equal rate for all users. For instance, when a user makes full use of the network resources, the  $\mathcal{F}_{IT}$  index reaches its minimum value of 0.

### 3.2 | Power optimization problem

The relative power ratio  $\beta_i$  as well as the minimum total power  $P$  to be considered for each user to attain a minimum equal data rate for all users (maximum fairness) can be obtained solving the follows optimization problem:

$$\begin{aligned} & \min_{(P, \beta_i) \in \mathbb{R}^+} \sum_{i=1}^M \beta_i P \\ \text{s.t. (C.1)} \quad & \sum_{j=1}^M \beta_j = 1 \\ \text{(C.2)} \quad & \beta_j > 0; P > 0; R > 0 \\ \text{(C.3)} \quad & \beta_m = (2^R - 1) \sum_{i=m+1}^M \beta_i + (2^R - 1) \frac{\sigma_n^2}{P |h_m|^2} \text{ to } m = 1, \dots, M - 1 \\ \text{(C.4)} \quad & \beta_M = (2^R - 1) \frac{\sigma_n^2}{P |h_M|^2}, \end{aligned} \quad (14)$$

where the expression for the proportions of total power can be derived from Equations (4) to (6).

One need to prove that the set of constraints in Equation (14) is convex. For this, the restriction (C.1) can be developed by replacing the restrictions (C.3) and (C.4) in the constraint (C.1) subjected to the constraints of (C.2); hence, the following expression is obtained.

$$f = \sum_{m=1}^{M-1} \left( (2^R - 1) \sum_{i=m+1}^M \beta_i + (2^R - 1) \frac{\sigma_n^2}{P |h_m|^2} \right) + (2^R - 1) \frac{\sigma_n^2}{P |h_M|^2} - 1. \quad (15)$$

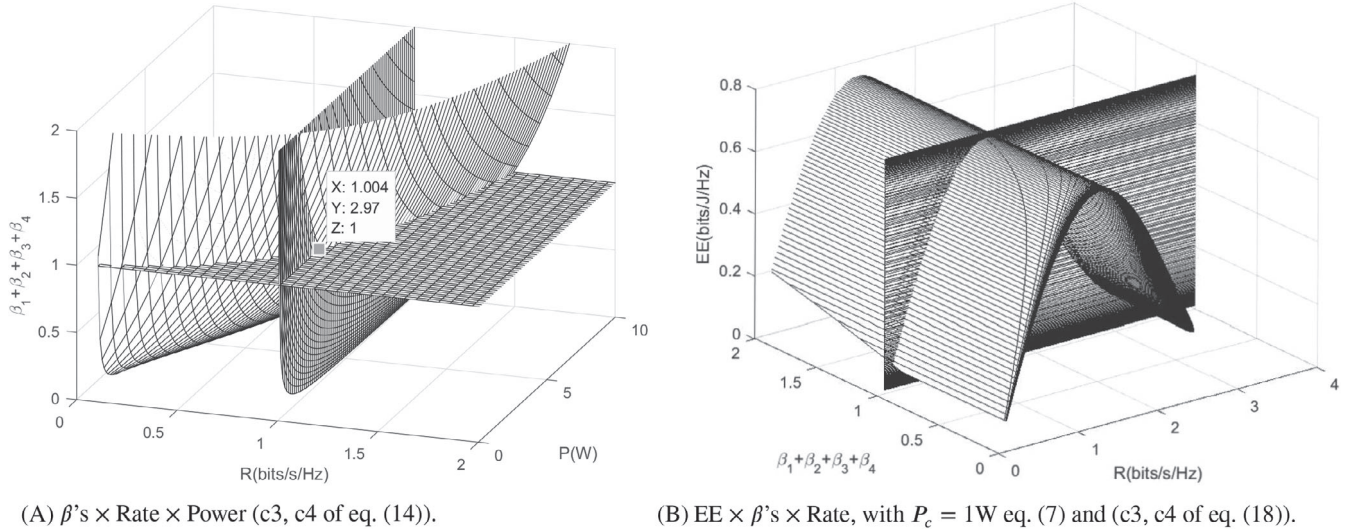
The Hessian of Equation (15) must be semidefinite positive, that is,  $\nabla^2 f \geq 0$  to be convex.<sup>21</sup> Therefore, the eigenvalues of the Hessian must be greater than or equal to zero. After applying the Hessian and eigenvalues definition in Equation (15), one can obtain the following polynomial characteristic for the eigenvalues determination:

$$-\lambda^M + \frac{2(2^R - 1)\sigma_n^2 \lambda^{M-1}}{\sum_{m=1}^{M-1} |h_m|^2 P^3} + \frac{2(2^R - 1)\sigma_n^2 \lambda^{M-1}}{P^3 |h_M|^2} = 0. \quad (16)$$

Hence, taking  $\lambda^{M-1}$  in evidence, so that there are  $M - 1$  eigenvalues equal to zero, and, therefore, one can obtain:

$$\lambda = \frac{2(2^R - 1)\sigma_n^2}{P^3} \left( \frac{1}{\sum_{m=1}^{M-1} |h_m|^2} + \frac{1}{|h_M|^2} \right). \quad (17)$$

As the variable  $P$  is positive, as well as the constants  $R$ ,  $\sigma_n^2$ , and  $|h_i|^2$ , the eigenvalues associated to Equation (15) are positive or equal to zero and the equation in question is semidefinite positive, then the constraints (C.1) to (C.4) constitute a convex set.



**FIGURE 2** Closed set for  $\sigma_n^2 = 1 \mu\text{W}$  and  $M = 4$  users with distances  $d_1 = 1.0 \text{ km}$ ,  $d_2 = 0.7 \text{ km}$ ,  $d_3 = 0.5 \text{ km}$ , and  $d_4 = 0.3 \text{ km}$ . (A)  $\beta$ 's  $\times$  Rate  $\times$  Power (c3, c4 of Equation (14)) and (B) EE  $\times$   $\beta$ 's  $\times$  Rate, with  $P_c = 1 \text{ W}$  Equation ((7)) and (c3, c4 of Equation (18))

The constraints form a collection of convex sets and the intersection of these sets is convex. According to Figure 2A the constraints defined in Equation (14) form a closed set; in this way, the presented optimization problem has a solution with optimal global. In Figure 2A one can observe the subspace given by the condition of  $\beta_1 + \beta_2 + \beta_3 + \beta_4 = 1$  and by the rate of each user  $R = 1 \text{ bits/s/Hz}$ . In this case, the solution is obtained for a total power  $P \approx 2.97 \text{ W}$ . However, the power proportion values cannot be obtained so easily in real practical interest system and channel scenarios. Moreover, Figure 2B depicts the set of solutions for the EE optimization problem. This set is formed by the sum of the power proportions equal to 1, that is,  $\beta_1 + \beta_2 + \beta_3 + \beta_4 = 1$  and the intersection of the other plane defined by the EE and user's rate.

### 3.3 | Optimization of user power allocation from the EE perspective

The Equation (7) is quasiconcave.<sup>6</sup> We use the Dinkelbach method to turn a quasiconcave objective function into a concave one,<sup>27</sup> as indicated in Table 1 to solve problem Equation (18), where  $U_R = MR$  and  $U_T = \rho \sum_{j=1}^M \beta_j P + MP_c$ .

The EE optimization problem to obtain the optimal power allocation  $P$ , data rate  $R$ , and power proportion  $\beta_i$  for all the users can be formulated as:

$$\begin{aligned}
 \min_{(P, \beta_j, R) \in \mathbb{R}^+} & - \left( MR - q^n \left( \rho \sum_{j=1}^M \beta_j P + MP_c \right) \right) \\
 \text{s.t. (C.1)} & \sum_{j=1}^M \beta_j = 1 \\
 \text{(C.2)} & \beta_j > 0; P > 0; R > 0 \\
 \text{(C.3)} & \beta_m = (2^R - 1) \sum_{i=m+1}^M \beta_i + (2^R - 1) \frac{\sigma_n^2}{P|h_m|^2}, \quad \text{for } m = 1, \dots, M-1 \\
 \text{(C.1)} & \beta_M = (2^R - 1) \frac{\sigma_n^2}{P|h_M|^2}.
 \end{aligned} \tag{18}$$

By analyzing the convexity of the constraints regarding the total power  $P$ , proportion of the total power among users  $\beta_i$  and data rate  $R$ , one can conclude that the set (C.1) to (C.4) in Equation (18) are nonconvex, as the eigenvalues of the Hessian with these variables result complex values.

However, by setting the data rate  $R$  while maintaining the total power  $P$  and the proportion of the total power  $\beta_i$  among users as optimization variables, the problem in Equation (18) becomes convex as Equation (15) is obtained. As

**TABLE 1** Iterative power allocation and data rate and/or power allocation algorithm based on Dinkelbach method

Dinkelbach's algorithm
1) Let $n = 0$ and $q^{(n)} = 0$ ;
Set $\epsilon > 0$ as the stopping criterion;
2) <b>REPEAT</b>
For a given $q^{(n)}$ , solve Equations (18) and/or (19) to obtain first the power allocation and data rate $P^{(n)}$ and $R^{(n)}$ and/or second the power allocation $P^{(n)}$ ;
<b>IF</b> $U_R(P^{(n)}, R^{(n)}) - q^{(n)}U_T(P^{(n)}, R^{(n)}) \leq \epsilon$ and/or $U_R(P^{(n)}) - q^{(n)}U_T(P^{(n)}) \leq \epsilon$
5) Convergence = <b>TRUE</b> ;
6) <b>RETURN</b> $P^* = P^{(n)}, R^* = R^{(n)}, q^* = q^{(n)}$ and/or $P^* = P^{(n)}, q^* = q^{(n)}$ ;
7) <b>ELSE</b>
8) Convergence = <b>FALSE</b> ;
9) Set $n = n + 1$ and $q^{(n)} = \frac{U_R(P^{(n-1)}, R^{(n-1)})}{U_T(P^{(n-1)}, R^{(n-1)})}$ or $q^{(n)} = \frac{U_R(P^{(n-1)})}{U_T(P^{(n-1)})}$ ;
10) <b>END IF</b>
11) <b>UNTIL</b> Convergence = <b>TRUE</b> .

proved, earlier this set of constraints with such variables is convex. Hence, the problem Equation (18) is reformulated as Equation (19). We use the Dinkelbach method to turn a quasiconvex objective function into a convex one, as Table 1 to the problem Equation (19).

$$\begin{aligned} \min_{(P, \beta) \in \mathfrak{R}^+} \quad & MR - q^n \left( \rho \sum_{j=1}^M \beta_j P + MP_c \right) \\ \text{s.t.} \quad & \text{(C.1), (C.2), (C.3), (C.4).} \end{aligned} \tag{19}$$

### 3.4 | Resource efficiency

In conventional system designs instead of focusing on the SE or the EE separately, it is much more effective balancing the attainable system SE and EE by adopting RE metric.<sup>6</sup> The SE is defined for a single cell system with  $M$  single-antenna users as  $SE = \sum_{i=1}^M R_i$ . The RE is expressed as a weighted sum of the EE and the SE can be formulated as:

$$RE = \xi_0 EE + SE, \tag{20}$$

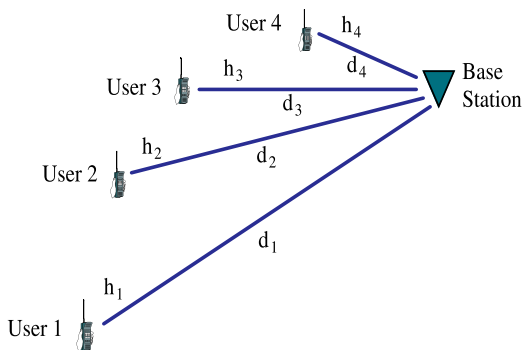
where  $\xi_0$  is the weighting factor in (W) controlling the weights of EE and the SE on the design. Hence, when  $\xi_0 = 0$  the expression (20) reduces to the SE.

## 4 | NUMERICAL RESULTS

The parameter values used in the simulations are described in Table 2. The tool used for the optimization was the MatLab<sup>®</sup> fmincon function with the technique sequential quadratic programming (SQP). The SQP algorithm<sup>28</sup> is a powerful iterative method for nonlinear optimization and nonconvex systems.<sup>29</sup> This algorithm can be considered as one of the best nonlinear programming method for constrained optimization problems. It outperforms every other nonlinear programming method in terms of efficiency, accuracy and percentage of successful solutions over a large number of test problems.<sup>30</sup> The SQP algorithm converges to the local minimum,<sup>30</sup> but very close to the global minimum. It was considered a range of values from the lowest to the highest to obtain the behavior of the system with different number of users, path loss exponent values, target equal user's rate and  $\mathfrak{R}_D$ . In this way, one can trace a system's behavior considering different practical scenarios of interest. Unlike Scenario 4, where users have a uniform distribution for the respective  $\mathfrak{R}_D$ , in the Scenario 1, 2, and 3 users have a fixed distance from BS.

Parameter	Value
Users number	$M = 2$ to 12
Path-loss exponent	$\alpha = [2.0; 3.0; 4.5]$
Power of white noise	$\sigma_n^2 = 0.1 \mu W$
Power budget	$P = 120 W$
Target equal use's rate	$R \in [0.1; 0.5; 1; 1.5; 2; 2.5; 3]$ (bit/s/Hz)
Circuits power	$P_c = 250 \left( \frac{mW}{\text{user}} \right)$
RF power amplifier inefficiency	$\rho = 1.4$
Number of samples	$10^4$
Monte-Carlo simulation trials	5
Scenarios: user distribution inside the cell	
Scenario 1— $\sum d_i = 1.0$ (km)	$d_1 = 0.34, d_2 = 0.29, d_3 = 0.22, d_4 = 0.15$
Scenario 2— $\sum d_i = 2.0$ (km)	$d_1 = 0.8, d_2 = 0.6, d_3 = 0.4, d_4 = 0.2$
Scenario 3— $\sum d_i = 4.0$ (km)	$d_1 = 1.8, d_2 = 1.2, d_3 = 0.7, d_4 = 0.3$
Scenario 4— $\mathcal{R}_D$ (m) (uniform distribution)	50, 100, 200, 300, 400
Power allocation strategies	
Inverse power allocation	$P_i$ defined by Equation (21)
Equal-rate optimal power allocation	Subsection 3.1

**TABLE 2** Adopted parameters for the simulation scenarios



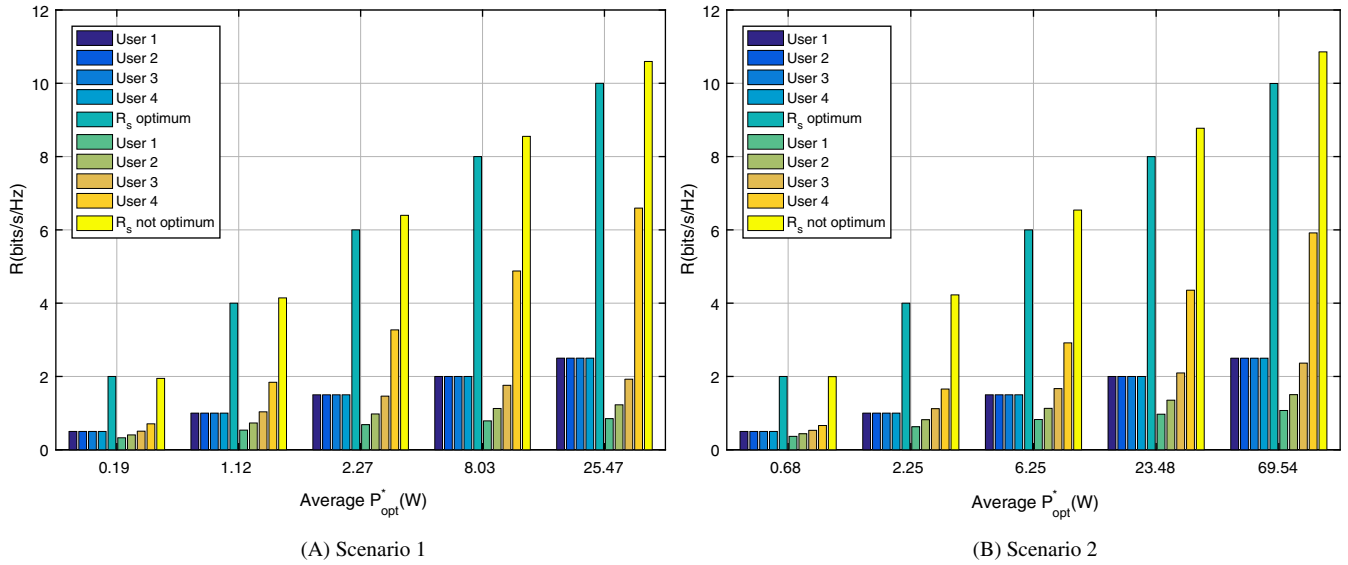
**FIGURE 3** Distribution of users in a nonorthogonal multiple access system with four users

Figure 3 shows the distribution of users relative to the BS in a cell. We used the power assignment technique among users proportional to the inverse of the channel response (ICA)<sup>31</sup> to compare the sum rates with egalitarian rates among all users, given by:

$$P_i = \frac{\sum_{k=1}^M |h_k|^2}{|h_i|^2} P, \quad i = 1, \dots, M. \quad (21)$$

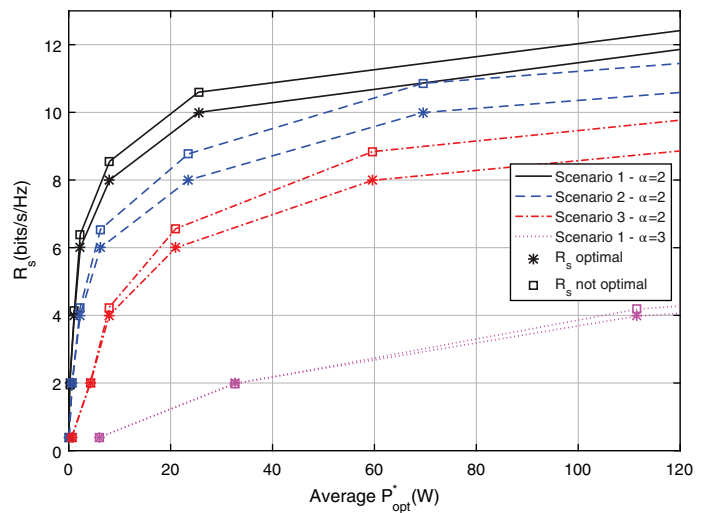
#### 4.1 | ERPA × ICA performance

To demonstrate the effectiveness of the proposed equal-rate optimal power allocation (ERPA) optimization following the rate fairness principle (same rate for all users), Figure 4 depicts the respective optimum and nonoptimum sum rate. It was considered four users with fixed rates equal to 0.5, 1, 1.5, 2, and 2.5 bits/s/Hz and optimum average powers for Scenario 1 with  $\alpha = 2$  (LoS—line of sight) for each user. The distributed ERPA strategy is compared with the one obtained by Equation (21) with the same total power. It was observed that only the optimal system kept the same rates for all users at the cost of marginally reduced sum rate at each total power scenario. In the nonoptimal system, the user 4 is closer to the BS and receives a greater proportion of power. This generates greater interference in other users, reducing their rates and



**FIGURE 4** Data rates vs allocated power considered line of sight pathloss ( $\alpha = 2$ ), problem (14). The nonoptimal rates were obtained with the power proportion given by Equation (21)

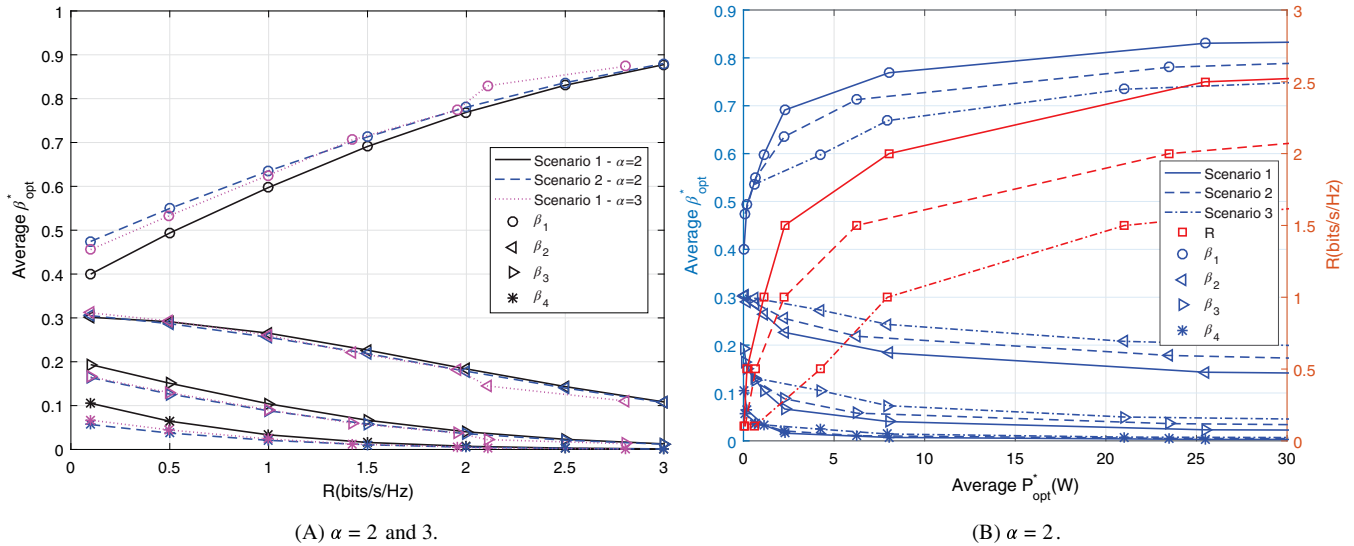
**FIGURE 5**  $R_s$  vs  $P$  to  $\alpha = 2$  and 3. The nonoptimal rates were obtained with the power proportion given by Equation (21), problem (14)



increasing its own rate, which results in a bit higher sum rates. Notice that to ensure fairness, the strongest user (nearest) partially loses rate, while far away users gain rate by allocating more power for them; this mechanism happens for all the power-rate scenarios analyzed. This is the price paid to have fairness among users. Indeed, to maintain a rate of  $R = 2$  bits/s/Hz for all users in Scenario 1 and  $\alpha = 2$ , the total power required is 8.03 W, Figure 4 ; for the same case in Scenario 2 the total power required is 23.48 W. In Scenario 2, the total power has to be 2.9 times greater because users are more distant.

Figure 5 shows the sum rate for optimal and a nonoptimal distribution of total powers, for scenarios with different distances between users, and for  $\alpha = 2$  and 3. With the same total power, the sums of nonoptimal rates are somewhat larger for the nonideal energy distribution because the closest users, with distances  $d_2, d_3,$  and  $d_4$  have a higher power proportion and consequently higher rate than the farthest user, with distance  $d_1$ . For instance, in the environment with  $\alpha = 2$  and Scenario 3, the sum of nonoptimal rates with an average power budget of  $P_{opt}^* \approx 59.7$  W is  $R_s = 8.8$  bits/s/Hz while for the sum of optimal rates is  $R_s = 8.0$  bits/s/Hz for the same power budget of 59.7 W.

In Figure 6A, we have the variation of the power ratios of users  $\beta_{opt}^*$  as a function of the rate of each user given by  $R$  in bits/s/Hz. The power proportions for users of Scenario 2 with path loss coefficient  $\alpha = 2$  and Scenario 1 with  $\alpha = 3$  have a marginal difference, because the increase of distances of users in the first one is compensated by the increased path loss



**FIGURE 6** Optimal power proportion and sum-rate behavior, problem (14), for the three scenarios and different path loss coefficient: (A)  $\beta_{opt}^* \times R$ ; (B)  $\beta_{opt}^* \times P_{opt}^*$  and  $R \times P_{opt}^*$

R (bits/s/Hz)	$\alpha = 2$			$\alpha = 3$
	Scenario 1	Scenario 2	Scenario 3	Scenario 1
1.0	0.947	0.968	<u>0.976</u>	0.953
1.5	0.943	0.962	0.970	0.948
2.0	0.941	0.958	0.966	0.945
2.5	0.940	0.955	0.962	0.943
3.0	<u>0.939</u>	0.952	0.960	0.942

Note: The underlined values are the maximum and minimum value in the table, respectively.

**TABLE 3** Fairness  $\mathcal{F}_{IT}$  obtained from the Equation (10)

coefficient of the second one. For users of Scenario 1 and the path loss coefficient  $\alpha = 2$ , the power ratio of users is smaller because of decrease in the path loss effect.

Figure 6B demonstrates how the power proportions  $\beta_i$  behaves concerning the total power  $P$  and the fixed rate  $R$  of each user. For a fixed rate of  $R = 1.5$  bits/s/Hz per user, the optimal total power is around 2.3, 6.2, and 21.0 W for the Scenarios 1, 2 and 3, respectively, while the power proportion for users 1 are around  $\beta_1 = 0.7, 0.71$ , and 0.73. The same power proportion, the different total power, and the same data rate occur because of the distance differences among users for each of the three scenarios.

## 4.2 | Fairness

Fairness index  $\mathcal{F}_{IT}$  defined in Equation (10) has been applied to the proposed ERPA-NOMA strategy for the four different scenarios. As shown in Table 3, fairness index results are very close to 1 demonstrating the effectiveness of the proposed optimization method. The maximum value of  $\mathcal{F}_{IT} = 0.976$  is in Scenario 3 for  $R = 1$  bits/s/Hz and the minimum value of  $\mathcal{F}_{IT} = 0.939$  is in Scenario 1 for  $R = 3$  bits/s/Hz.

Comparing the Scenarios 1 to 3 for  $\alpha = 2$ , one can conclude that the increase in fairness index is gradual, approaching to 1 when the sum of users' distance increases (Scenario 3). Moreover, the fairness index  $\mathcal{F}_{IT}$  increases when the nonline of sight occurs, that is, when  $\alpha > 2$  happens. Finally,  $\mathcal{F}_{IT}$  decreases slowly when the equal rate per user  $R$  increases, and it demonstrates how difficult it is to attain more restrictive tied QoS goals.

Jain's fairness index  $\mathcal{F}_j$  Equation (9) resulted in 1 once it was ensured that all users had the same rate in optimization process. This was sufficient to ensure maximum fairness when calculated for the index. However, fairness index  $\mathcal{F}_{IT}$

based on the information theory<sup>11</sup> resulted close to 1. For index  $\mathcal{F}_{IT}$  to be equal to 1, powers must also be uniformly distributed.

### 4.3 | EE optimal operation points

The optimal point for  $EE^*$  is depicted in Figure 7A,B for the three fixed-value scenarios and for LoS channel  $\alpha = 2$ . For Scenario 1,  $EE^* = 1.13$  bits/J/Hz operation point results in a rate of  $R = 6.35$  bits/s/Hz for a total power consumption of  $P = 4.60$  W. Although for Scenarios 2 and 3, we have rates of  $R = 5.99$  bits/s/Hz and  $R = 4.56$  bits/s/Hz for a total power of  $P = 6.30$  W and  $P = 13.57$  W, which resulted in  $EE^* = 0.99$  bits/J/Hz and  $EE^* = 0.63$  bits/J/Hz, respectively. Of course, Scenario 1 (closer users), less power is required to reach the specified rate, so that very low power reaches the highest rate and consequently higher  $EE^*$  as it can be observed in graph of Figure 7A,B in the points in asterisk. Indeed, the  $EE^*$  point for Scenario 1 is higher than users of Scenarios 2 and 3, as users of the Scenario 1 are closer each other, which can be corroborated by the behavior of Figure 7B.

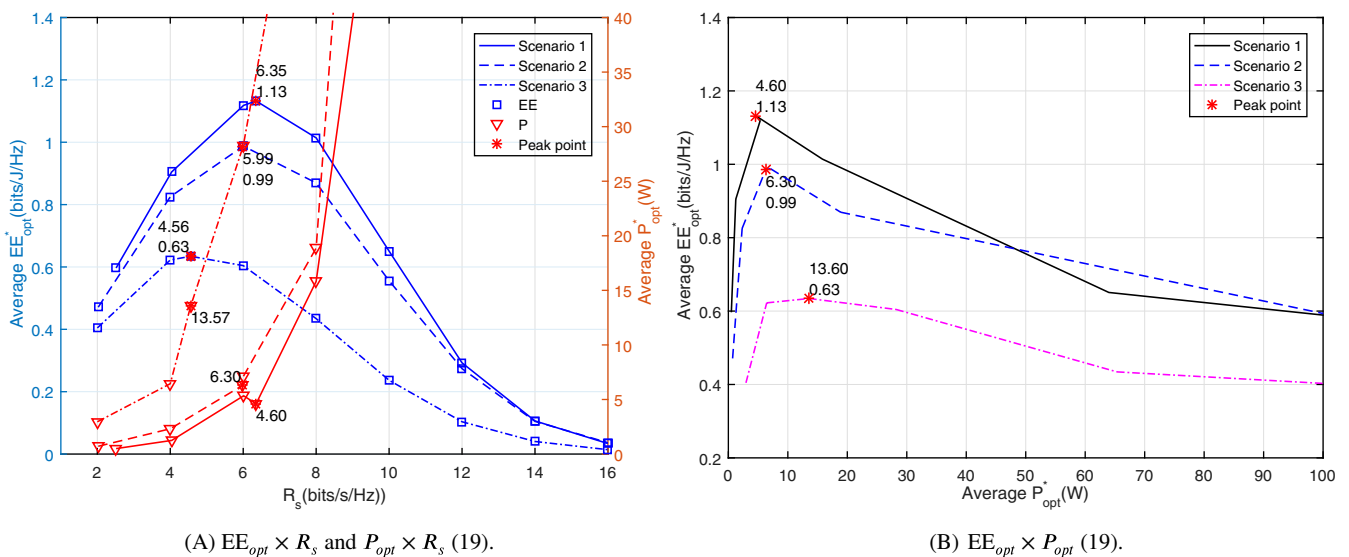


FIGURE 7 Optimal points considering  $\alpha = 2$ . Peak point problem (18)

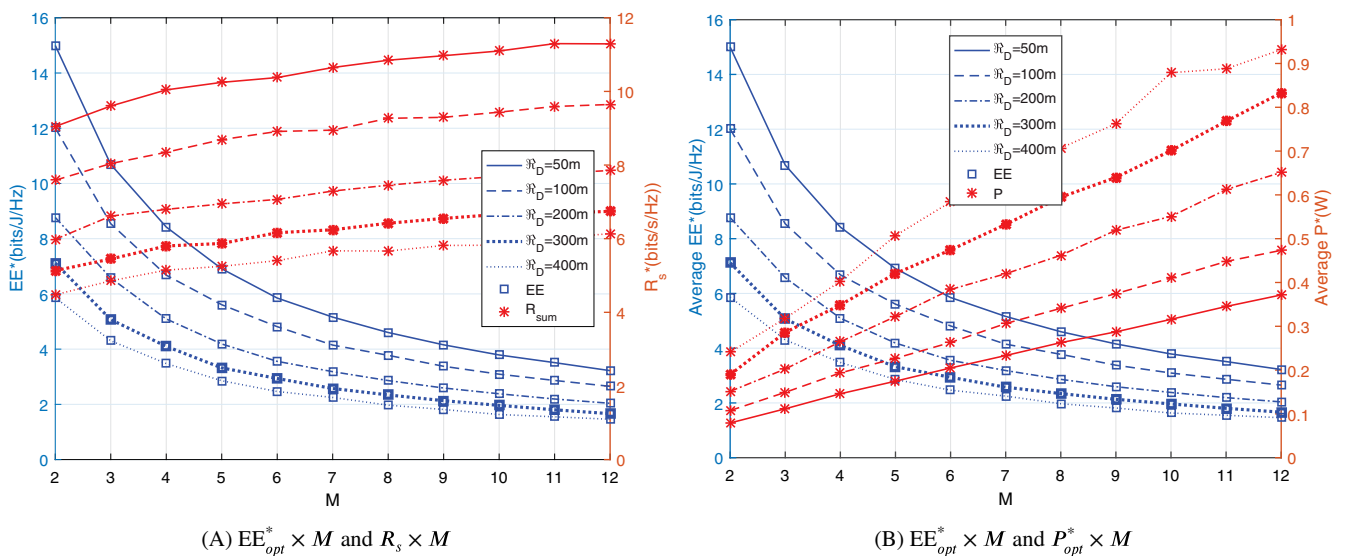
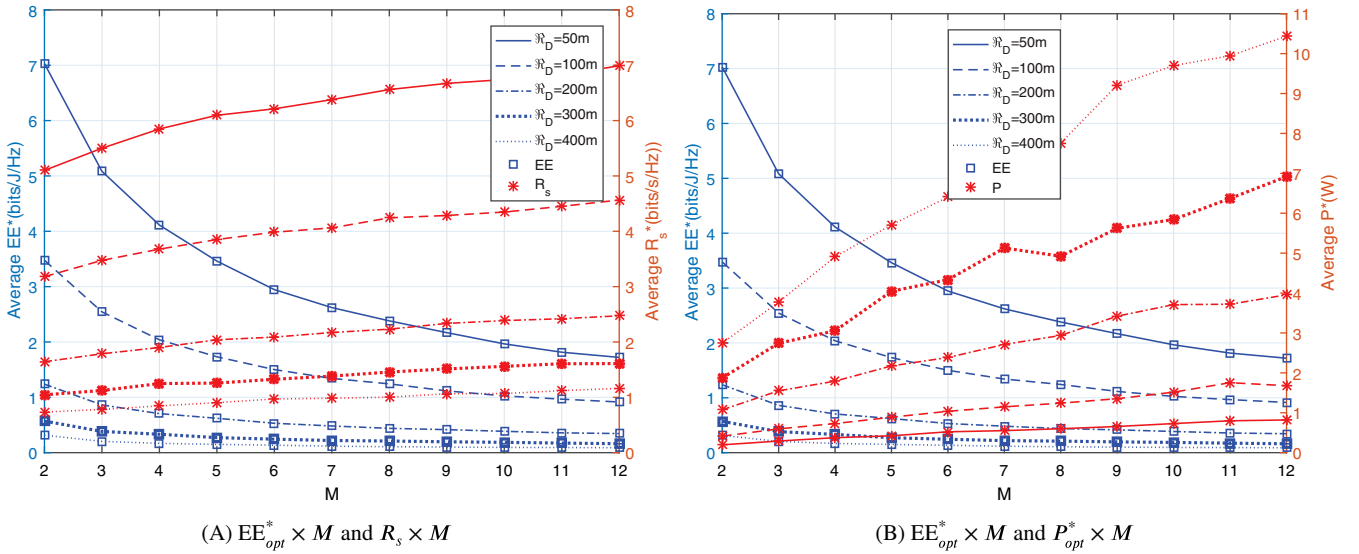
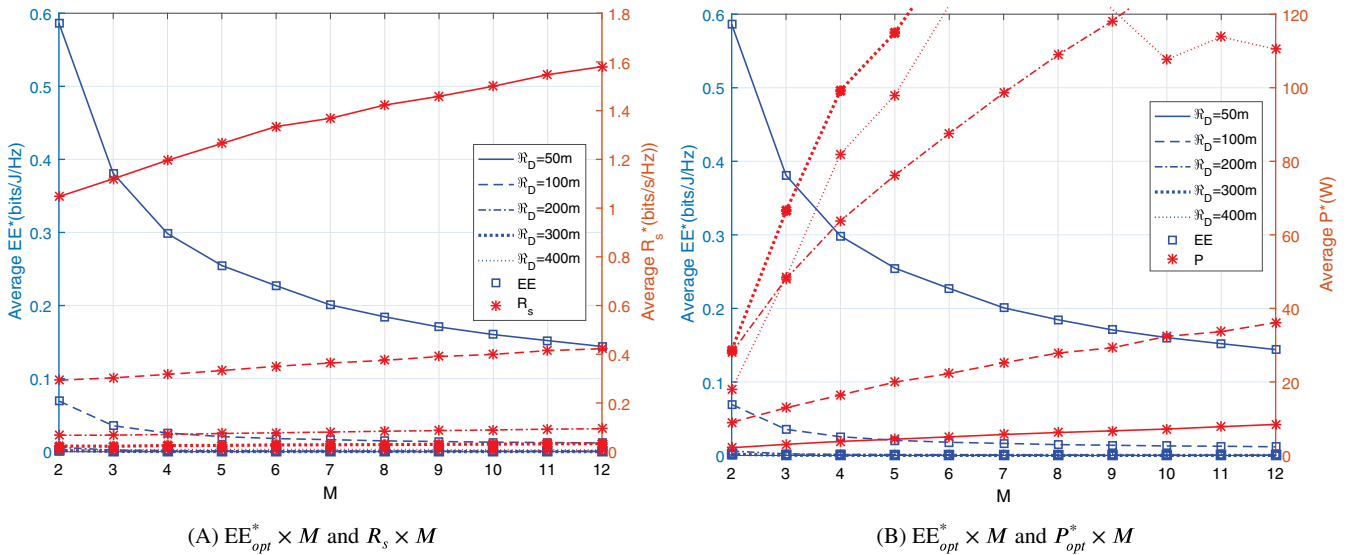


FIGURE 8 Trade-off between energy efficiency and spectral efficiency for some radius  $\mathfrak{R}_D$  and  $\alpha = 2$ , problem (18)



**FIGURE 9** Trade-off between energy efficiency and spectral efficiency for some radius  $\mathfrak{R}_D$  and  $\alpha = 3$ , problem (18)



**FIGURE 10** Trade-off between energy efficiency and spectral efficiency for some radius  $\mathfrak{R}_D$  and  $\alpha = 4.5$ , problem (18)

#### 4.4 | RE (EE $\times$ SE) optimal point as a function of the number of users

Problem (18) was used to obtain RE from EE. As it can be observed from the graph in Figure 7, the peak point is very close to maximum point to fixed  $R$  in problem (19). In Figures 8 to 10, one can see the optimal values for EE, sum rates and total power as a function of the number of users  $M$  from the perspective of the trade-off EE  $\times$  SE. In this case, the sum rate  $R_s$  and total power  $P$  are considered as variables in the optimization process of problem (18). To obtain the curves, it was considered that users have uniform distribution within the cell of radius  $\mathfrak{R}_D$  and  $P_c = 250$  mW per user.<sup>32</sup> We considered  $10^4$  samples for each number of users  $M$ . Table 4 summarized the relevant points of numerical results obtained from Figures 8 to 10, where the best trade-off between EE and  $R_s$  were selected for each scenario. This trade-off represented the number of users that obtained the best system RE to  $\xi_0 = 1.8$  in Equation (20).

For different rays  $\mathfrak{R}_D$  and  $\alpha$ , two users/cluster configuration showed the best RE as observed in Table 4. EE and SE reduced with the increase of rays of cells and consequently the increase of optimum powers obtained due to larger distances of users. These values were the best configuration to obtain the highest RE in NOMA system in the presented scenarios.

**TABLE 4** Trade-off point between energy efficiency (EE) and  $R_s$  for some  $\mathfrak{R}_D$  and users with uniform distributions

$\mathfrak{R}_D$	$\alpha = 2$				$\alpha = 3$				$\alpha = 4.5$			
	User	EE	$R_s$	$P$	User	EE	$R_s$	$P$	User	EE	$R_s$	$P$
50	2	15.00	9.05	0.08	2	7.03	5.10	0.20	2	0.59	1.05	2.11
100	2	12.01	7.59	0.11	2	3.48	3.18	0.41	2	0.07	0.29	9.03
200	2	8.77	5.98	0.15	2	1.24	1.64	1.07	2	0.006	0.07	27.94
300	2	7.12	5.13	0.19	2	1.04	0.57	1.87	2	0.0005	0.02	28.65
400	2	4.48	5.86	0.24	2	0.32	0.73	2.74	2	0.0005	0.008	17.92

## 5 | CONCLUSIONS

In this work, we defined the proportion of energy distributed optimally for each user to a given transmission rate in a NOMA system of two users or more. In our design, it was also possible to optimize the total power for the same rate maximum fairness among all users. We used the techniques of SQP, which presents fast convergence feature. The numerical results corroborate the minimum total power allocation for two or more users need to reach the defined capacity. The design and analysis were developed aiming to find the minimum equal rate for each user in each scenario, with the maximum EE. Maximum fairness was guaranteed in our project. We also found the trade-off point between EE and the sum rate, that is, the system operating in the point of best RE.

### ORCID

Jaime L. Jacob  <https://orcid.org/0000-0003-4943-0397>

Taufik Abrão  <https://orcid.org/0000-0001-8678-2805>

### REFERENCES

1. Tse D, Viswanath P. *Fundamentals of Wireless Communication*. Cambridge, MA: Cambridge University Press; 2005.
2. Kimy B, Lim S, Kim H, et al. Non-orthogonal multiple access in a downlink multiuser beamforming system. Paper presented at: MILCOM 2013 - IEEE Military Communications Conference; 2013, 1278-1283.
3. Chen X, Benjebbour A, Lan Y, Li A, Jiang H. Impact of rank optimization on downlink non-orthogonal multiple access (NOMA) with SU-MIMO. Paper presented at: IEEE International Conference on Communication Systems, IEEE ICCS 2014; 2014, 233-237.
4. Choi J. Non-orthogonal multiple access in downlink coordinated two-point systems. *IEEE Commun Lett*. 2014;18(2):313-316.
5. Dhakal S, Martin PA, Smith PJ. NOMA with guaranteed weak user QoS: design and analysis. *IEEE Access*. 2019;7:32884-32896.
6. Zhang Y, Wang HM, Zheng TX, Yang Q. Energy-efficient transmission design in non-orthogonal multiple access. *IEEE Trans Veh Technol*. 2017;66(3):2852-2857.
7. Liu Q, Tan F, Lv T, Gao H. Energy efficiency and spectral-efficiency tradeoff in downlink NOMA systems. Paper presented at: IEEE International Conference on Communications Workshops; 2017, 247-252.
8. Bui VP, Nguyen PX, Nguyen HV, Nguyen VD, Shin OS. Optimal user pairing for achieving rate fairness in downlink NOMA networks. Paper presented at: International Conference on Artificial Intelligence in Information and Communication; 2019, 575-578.
9. Ni D, Hao L, Qian X, Tran QT. Energy-spectral efficiency tradeoff of downlink NOMA system with fairness consideration. Paper presented at: IEEE Vehicular Technology Conference; 2018, 1-5.
10. Jain RK, Chiu D-MW, Hawe WR. A quantitative measure of fairness and discrimination for resource allocation in shared computer system. *Technical report DEC-TR-301*. 1984;1-38. Available at: <https://www.cse.wustl.edu/~jain/papers/ftp/fairness.pdf>. Accessed January 7, 2020.
11. Gui G, Sari H, Sari H, Biglieri E. A new definition of fairness for non-orthogonal multiple access. *IEEE Commun Lett*. 2019;23(7):1267-1271.
12. Wei Z, Guo J, Ng DWK, Yuan J. Fairness comparison of uplink NOMA and OMA. Paper presented at: IEEE 85th Vehicular Technology Conference (VTC Spring); 2017, 1-6.
13. Xu P, Cumanan K. Optimal power allocation scheme for non-orthogonal multiple access with  $\alpha$ -fairness. *IEEE J Selected Areas Commun*. 2017;35(10):2357-2369.
14. Hojeij M-R, Nour CA, Farah J, Douillard C. Waterfilling-based proportional fairness scheduler for downlink non-orthogonal multiple access. *IEEE Wirel Commun Lett*. 2017;6(2):230-233.
15. Zuo Y, Zhu X, Jiang Y, Wei Z, Zeng H, Wang T. Energy efficiency and spectral efficiency tradeoff for multicarrier NOMA systems with user fairness. Paper presented at: IEEE/CIC International Conference on Communications in China (ICCC); 2018, 666-670.
16. Xing H, Liu Y, Nallanathan A, Ding Z, Poor HV. Optimal throughput fairness tradeoffs for downlink non-orthogonal multiple access over fading channels. *IEEE Trans Wirel Commun*. 2018;17(6):3556-3571.
17. Qi T, Feng W, Chen Y, Wang Y. Optimum fairness for non-orthogonal multiple access. Paper presented at: IEEE/CIC International Conference on Communications in China (ICCC); 2018, 116-120.

18. Muhammed AJ, Ma Z, Diamantoulakis PD, Li L, Karagiannidis George K. Energy-efficient resource allocation in multicarrier NOMA systems with fairness. *IEEE Trans Commun.* 2019;67(12):8639-8654.
19. Fang F, Zhang H, Cheng J, Leung VCM. Energy efficiency of resource scheduling for non-orthogonal multiple access (NOMA) wireless network. Paper presented at: IEEE ICC 2016 - Next-Generation Networking and Internet Symposium; 2016, 1-5.
20. Chen L, Ma L, Xu Y. Proportional fairness-based user pairing and power allocation algorithm for non-orthogonal multiple access system. *IEEE Access.* 2019;7:19602-19615.
21. Boyd S, Vandenberghe L. *Convex Optimization.* Cambridge, UK: Cambridge University Press; 2004.
22. Bertsekas DP. *Convex Optimization Theory.* Belmont, MA: Athena Scientific; 2009.
23. Timotheou S, Krikidis I. Fairness for non-orthogonal multiple access in 5G systems. *IEEE Signal Process Lett.* 2015;22(10):1647-1651.
24. Ding Z, Yang Z, Fan P, Poor HV. On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users. *IEEE Signal Process Lett.* 2014;21(12):1501-1505.
25. Inaltekin H, Chiang M, Poor HV, Wicker SB. On unbounded path-loss models: effects of singularity on wireless network performance. *IEEE J Selected Areas Commun.* 2009;27(7):1078-1092.
26. Wang CH, Lin JY, Wu JM. Joint fairness and sum rate resource allocation for noma communications. Paper presented at: IEEE Conference on Standards for Communications and Networking, CSCN 2017; 2017, 269-274.
27. Tang J, Luo J, Liu M, et al. Energy efficiency optimization for NOMA with SWIPT. *IEEE J Select Topics Signal Process.* 2019;13(3):452-466.
28. Nocedal J, Wright SJ. *Numerical Optimization.* Springer Series in Operations Research and Financial Engineering. 2nd ed. New York, NY: Springer Science, LLC; 2006.
29. Ding J, Dutkiewicz E, Huang X, Fang G. Energy-efficient distributed beamforming in uwb based implant body area networks. Paper presented at: IEEE Vehicular Technology Conference; 2015, 1-5.
30. Elaiw AM, Xia X, Shehata AM. Solving dynamic economic emission dispatch problem with valve-point effects using hybrid DE-SQP. Paper presented at: IEEE Power and Energy Society Conference and Exposition in Africa: Intelligent Grid Integration of Renewable Energy Resources; 2012, 1-6.
31. El-Sayed MM, Ibrahim AS, Khairy MM. Power allocation strategies for non-orthogonal multiple access. Paper presented at: International Conference on Selected Topics in Mobile & Wireless Networking (MoWNeT); 2016, 1-6.
32. Devroye L. *Non-Uniform Random Variate Generation.* New York, NY: Springer Science, LLC; 1986.

**How to cite this article:** Jacob JL, Abrão T. Nonorthogonal multiple access systems optimization to ensure maximum fairness to users. *Trans Emerging Tel Tech.* 2020;e3875. <https://doi.org/10.1002/ett.3875>

**ANEXO B – Energy and Spectral  
Efficiencies Trade-off in  
MIMO-NOMA System Under  
User-Rate Fairness and Variable User  
Per Cluster**

(P2 – artigo publicado em 2021)



## Full length article

## Energy and spectral efficiencies trade-off in MIMO-NOMA system under user-rate fairness and variable user per cluster

Jaime L. Jacob<sup>a</sup>, Cristiano Magalhaes Panazio<sup>b</sup>, Taufik Abrão<sup>a,\*</sup><sup>a</sup> Department of Electrical Engineering, State University of Londrina, Londrina 86057-970, Brazil<sup>b</sup> Department of Telecommunication and Control (PTC) of Polytechnic School of São Paulo University, Brazil

## ARTICLE INFO

## Article history:

Received 23 November 2020

Received in revised form 30 March 2021

Accepted 22 April 2021

Available online 1 May 2021

## Keywords:

Non-orthogonal multiple access (NOMA)

Multiple-input and multiple-output

(MIMO)

Optimization

Successive interference cancellation (SIC)

Energy efficiency (EE)

Fairness

## ABSTRACT

The power proportion allocation for each user in down-link (DL) multiple-input and multiple-output (MIMO) non-orthogonal multiple access (NOMA) systems is analyzed. Successive interference cancellation (SIC) technique is deployed to recover users' signal with highest differences between channel gains. In proposed optimization design, initially, the minimum total power and the power proportion factor for each active user are obtained. After that, the maximum system energy efficiency (EE), the equal data rate, total power and power-proportion distribution along the users and clusters are determined under a certain circuitry power consumption. Both optimization procedures have deployed the same constraints and all users were subjected to the equal data rate. The numerical results show that the total number of users reach the maximum EE in each analyzed clusters-users configurations. The number of total users in the cell for specific number of users-per-cluster (2, 3 or 4) that simultaneously maximize EE and the equal data rate per user (representing the maximum fairness) among all active MIMO-NOMA users depends on the path-loss channel coefficients, the number of users-per-cluster and the cell coverage (radius).

© 2021 Elsevier B.V. All rights reserved.

## 1. Introduction

Non-Orthogonal multiple Access (NOMA) is considered a promising technique for fifth generation (5G) technology. The power domain NOMA uses different power levels between users to transmit on the same channel simultaneously [1]. The NOMA receiver extracts its signals using successive interference cancellation (SIC) to eliminate interference between users [2]. It explores the different path loss between users to separate signals, thereby increasing spectral efficiency (SE) and user fairness by applying a scheme to achieve the desired capacity on the downlink (DL) [3].

The application of multiple input and multiple output (MIMO) technology in NOMA is promising because the use of MIMO technique provides high data rate and low error rate through multiplexing gain and spatial diversity, respectively [4,5]. A novel MIMO-NOMA framework for DL and uplink (UL) transmission applying the concept of signal alignment has been proposed [5]. A new MIMO-NOMA UL implementation scheme with group interference cancellation is shown in the literature to be a better scheme than orthogonal multiple access transmission and signal alignment [6]. In [7], it is shown analytically and through

simulations the superiority of MIMO-NOMA over conventional orthogonal multiple access (OMA)-based MIMO system in terms of the sum of channel capacity and also with regard to the sum of ergodic capacity. Besides, the maximization of ergodic capacity of MIMO-NOMA system with the restriction of the maximum transmitted power and the minimum rate for the weaker user has been proposed showing its superiority over the traditional MIMO-OMA [8]. To cancel inter cluster interference in MIMO-NOMA systems, a linear beamforming technique has been suggested, and for intra cluster and inter cluster power allocation (PA) to maximize the overall capacity of the cell, a dynamic PA solution was proposed [2]. In order to minimize the total power consumed by grouping of mobile users it has been investigated by means of a linear optimization approach to jointly optimize beam-forming vectors and energy allocation coefficients for a MIMO-NOMA cluster [9]. An important way of applying the MIMO technique in NOMA system has been considered so that a new project of precoding and detection of matrices for MIMO-NOMA systems and its performance were analyzed [10]. For a generic number of users per cluster in a DL MIMO-NOMA system, the energy efficiency (EE) and PA schemes were proposed to convert a difficult EE-PA problem into an equivalent SE-PA problem [1]. A PA strategy to ensure that the MIMO-NOMA system achieves a high individual rate for each user was compared to the arbitrary PA in MIMO-OMA system [11]. It was shown that MIMO-NOMA

\* Corresponding author.

E-mail addresses: [jaimel.jacob@uel.br](mailto:jaimel.jacob@uel.br) (J.L. Jacob), [cpanazio@lcs.poli.usp.br](mailto:cpanazio@lcs.poli.usp.br) (C.M. Panazio), [taufik@uel.br](mailto:taufik@uel.br) (T. Abrão).

could achieve better fairness through appropriate PA. The paper [12] made an analysis of the trade-off between sum rate and fairness using single and multi-objective optimization based on beamforming design for a SIMO-NOMA system. To achieve a better trade-off, an algorithm for selecting both strong and weak users based on fairness ratio was developed in [13] for a NOMA downlink system with zero-forcing beamforming in order to improve throughput-fairness trade-off in NOMA system.

In [14], it establishes the EE-NOMA optimization problem by finding the minimum total power needed to reach a specified equal rate per user. Based on the tradeoff that relates EE optimization and SE optimization in resource efficiency (RE) problem with maximum fairness, [14] calculates these values for each number of users and presents one with the highest RE in the cell. The current contribution differs from [14] in the sense that herein we consider a variable number of user per cluster. Unlike this work, in [14] the MIMO configuration was not considered. The objective is to obtain the best number of clusters for two, three and four users per cluster that maximizes the energy efficiency (EE) and the average equal-rate per user, guaranteeing maximum fairness among all active MIMO-NOMA users.

Fairness is an important concept when it comes aggregated to 5G systems. There are many works analyzing the behavior of fairness in 5G systems. However, differently what exists in the literature, in our work we reveal the behavior of the MIMO-NOMA systems simultaneously operating under the maximum user fairness (including edge users) and maximum system energy efficiency.

**Contribution:** The contribution of this work is to establish the best PA policy in MIMO-NOMA system in terms of sum-PA coefficients, which guarantees the same rate (fairness approach) for all users in MIMO-NOMA system. For that, we deploy optimization techniques [15,16] to solve the optimization problem. This paper also considers the EE-MIMO-NOMA optimization problem by finding the minimum total power needed to reach a specified equal rate per user. From the jointly EE and sum rate optimization with maximum fairness, i.e., the same rate among users, one can find the best EE-SE trade-off and thereby get the cluster number in the cell which results in maximum EE in the analyzed configurations.

This paper is organized as follows. Section 2 describes the MIMO-NOMA system model. Section 3 formulates the MIMO-NOMA EE-SE trade-off optimization problem from the perspective of RE with equal user-rate and specific number of user per cluster. Then, numerical simulation is presented in Section 4, and Section 5 concludes the paper.

## 2. MIMO-NOMA system

Consider a DL transmission in a multiple-input and multiple-output MIMO cell system combined with a non-orthogonal multiple-access (NOMA) system in which the base station (BS) is equipped with  $A$  antennas, as depicted in Fig. 1.

In order to have a NOMA scenario, the cellular system is divided into  $A$  clusters randomly grouped. There are  $K \geq 2$  users for each cluster, where users have  $U$  antennas each one. Herein, we assume  $U \geq A$  to make the system feasible. All clusters have the same number of users and all users belong to the same cell (single-cell system); hence, without loss of generality, the total number of users in the cell is selected to be multiple of the number of users per cluster, i.e. equal to  $K \cdot A$ . The signal transmitted by the BS to the user can be described by [17]

$$\mathbf{x} = \mathbf{W}\hat{\mathbf{s}} \quad (1)$$

where the data vector  $\hat{\mathbf{s}} \in \mathcal{C}^{A \times 1}$  can be written as:

$$\hat{\mathbf{s}} = \begin{bmatrix} \sqrt{P\beta_{1,1}}s_{1,1} + \dots + \sqrt{P\beta_{1,k}}s_{1,k} + \dots + \sqrt{P\beta_{1,K}}s_{1,K} \\ \vdots \\ \sqrt{P\beta_{a,1}}s_{a,1} + \dots + \sqrt{P\beta_{a,k}}s_{a,k} + \dots + \sqrt{P\beta_{a,K}}s_{a,K} \\ \vdots \\ \sqrt{P\beta_{A,1}}s_{A,1} + \dots + \sqrt{P\beta_{A,k}}s_{A,k} + \dots + \sqrt{P\beta_{A,K}}s_{A,K} \end{bmatrix} = \begin{bmatrix} \hat{s}_1 \\ \vdots \\ \hat{s}_a \\ \vdots \\ \hat{s}_A \end{bmatrix}. \quad (2)$$

In Eq. (1),  $\mathbf{W}$  represents the precoding matrix given by  $\mathbf{W} = \mathbf{I}_A$ , where  $\mathbf{I}_A$  is  $A \times A$  identity matrix and  $\mathbf{w}_a$  is the  $a$ th column of matrix  $\mathbf{W}$ . It allows to properly identify the cluster to be considered at the transmitter, as well as to describe such cluster signal at receiver side. This choice prevents the transmitter from sending channel state information (CSI), thereby reducing system overhead [10]. In Eq. (2),  $s_{a,k}$  and  $\beta_{a,k}$  represent the signal and the power proportion coefficients for the  $a$ th cluster and the  $k$ th user, respectively, and  $P$  indicates the total power transmitted by the BS. Besides, the *proportion power coefficients* must satisfy

$$\sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} = 1. \quad (3)$$

Hence, to attain data rate fairness across all NOMA users, those users with lower channel gain in BS will receive larger  $\beta_{a,k}$ . Precisely, such values will be obtained from optimization processes.

Consequently, in the DL, the signal received by a certain user  $k$  belonging to the  $a$ th cluster is given by

$$\mathbf{y}_{a,k} = \mathbf{H}_{a,k}\mathbf{W}\hat{\mathbf{s}} + \mathbf{z}_{a,k}, \quad \text{where } \mathbf{H}_{a,k} = \frac{\mathbf{Q}_{a,k}}{\sqrt{1+d_{a,k}^\alpha}}. \quad (4)$$

In Eq. (4),  $\mathbf{Q}_{a,k} \in \mathcal{C}^{U \times A}$  denotes a Rayleigh fading channel matrix,  $\alpha$  is the path loss coefficient,  $d_{a,k}$  represents the distance from the  $k$ th user that belongs to the  $a$ th cluster of the BS, and  $\mathbf{z}_{a,k}$  is an additive Gaussian noise (AWGN) vector,  $\mathcal{CN}(0, \sigma^2 \mathbf{I})$ . The received signal must be multiplied by a detection vector  $\mathbf{v}_{a,k}$  which overrides the interfering signal [18]. This way, one can identify the interfering cluster term:

$$\mathbf{v}_{a,k}^H \mathbf{y}_{a,k} = \mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_a \sum_{k=1}^K \sqrt{P\beta_{a,k}} s_{a,k} + \underbrace{\sum_{i=1, i \neq a}^A \mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_i \hat{s}_i}_{\text{interfering clusters}} + \mathbf{v}_{a,k}^H \mathbf{z}_{a,k} \quad (5)$$

where  $\hat{s}_i$  represents the  $i$ th element of  $\hat{\mathbf{s}}$  in Eq. (2), and  $\mathbf{w}_\ell$  is the  $\ell$ th column of matrix  $\mathbf{W}$ . Hence, the product  $\mathbf{H}_{a,k} \mathbf{w}_i$  results in the  $i$ th column of  $\mathbf{H}_{a,k}$ . This column represents the antenna that serves the  $i$ th cluster.

To remove interference from other clusters, the detection matrix must obey the condition  $\mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{p}_i = 0$  for any  $i \neq a$ . This constraint is reduced to  $\mathbf{v}_{a,k}^H \mathbf{H}_{a,k} = 0$ , where  $\mathbf{H}_{a,k} = [\mathbf{h}_{1,ak} \dots \mathbf{h}_{a-1,ak} \mathbf{h}_{a+1,ak} \dots \mathbf{h}_{A,ak}]$  and  $\mathbf{h}_{j,ak}$  is the  $j$ th column of  $\mathbf{H}_{a,k}$ . We have [10]

$$\mathbf{v}_{a,k} = \mathbf{U}_{a,k} \mathbf{g}_{a,k}, \quad (6)$$

where  $\mathbf{U}_{a,k}$  is the matrix of the left singular vectors of  $\overline{\mathbf{H}}_{a,k}$  after decomposing on non-zero singular values and the maximal ratio

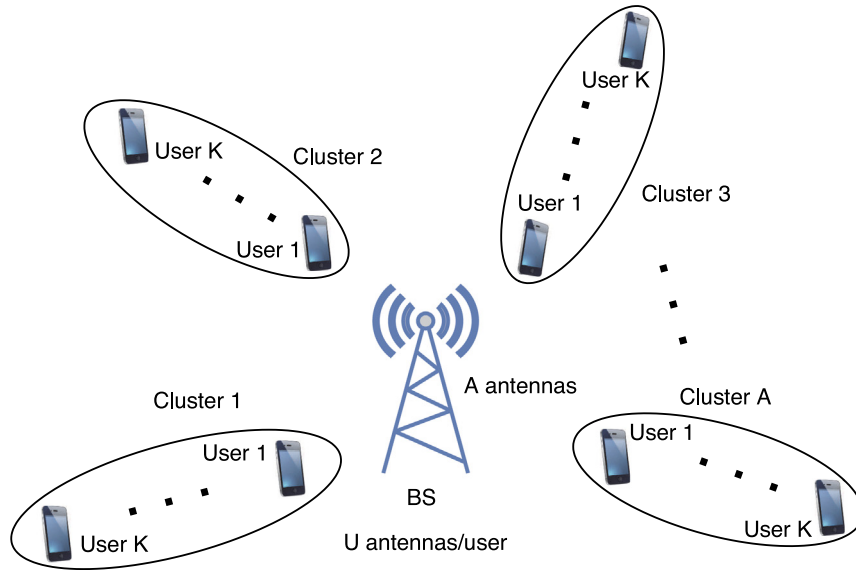


Fig. 1. Cellular system with a BS with  $A$  antennas,  $A \cdot K$  users with  $A$  clusters,  $K$  users/clusters and each user has  $U$  antennas ( $U \geq A$ ).

combining vector expressed as

$$\mathbf{g}_{a,k} = \frac{\mathbf{U}_{a,k}^H \mathbf{h}_{a,ak}}{|\mathbf{U}_{a,k}^H \mathbf{h}_{a,ak}|}. \quad (7)$$

After removing the interference from the other clusters, Eq. (5) can be rewritten as

$$\mathbf{v}_{a,k}^H \mathbf{y}_{a,k} = \mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_a \sum_{k=1}^K \sqrt{P \beta_{a,k}} \mathbf{s}_{a,k} + \mathbf{v}_{a,k}^H \mathbf{z}_{a,k} \quad (8)$$

Without loss of generality, the effective channel gains are ordered as [10]

$$|\mathbf{v}_{a,1}^H \mathbf{H}_{a,1} \mathbf{w}_a|^2 \geq \dots \geq |\mathbf{v}_{a,K}^H \mathbf{H}_{a,K} \mathbf{w}_a|^2 \quad (9)$$

and the power proportion coefficients are ordered as follows

$$\beta_{1,1} \leq \dots \leq \beta_{1,K}. \quad (10)$$

The BS still needs to be informed about the scalar effective channel gain order to implement the NOMA as shown in (9).

At the reception, each user applies SIC to remove interference from other users with worse channel gain. In a generic way, interfering users  $(a, k+1), \dots, (a, K)$  are removed by user  $(a, k)$ . We consider that all interference is removed because we use a perfect SIC. In this way, the signal-to-interference-plus-noise ratio (SINR) for the  $k$ th user in the  $a$ th cluster is given by [10]

$$\gamma_{a,k} = \frac{P |\mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_a|^2 \beta_{a,k}}{P \sum_{j=1}^{k-1} \beta_{a,j} |\mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_a|^2 + |\mathbf{v}_{a,k}|^2 \sigma^2} \quad (11)$$

for  $1 \leq a \leq A$  and  $2 \leq k \leq K$ . The first user results in the following SINR:

$$\gamma_{a,1} = \frac{P |\mathbf{v}_{a,1}^H \mathbf{H}_{a,1} \mathbf{w}_a|^2 \beta_{a,1}}{|\mathbf{v}_{a,1}|^2 \sigma^2} \quad (12)$$

while the normalized data rate or spectral efficiency for the  $k$ th user of the  $a$ th NOMA cluster reached by a user in a given cluster results:

$$R_{a,k} = \log_2(1 + \gamma_{a,k}) \quad \left[ \frac{\text{bits}}{\text{s Hz}} \right] \quad (13)$$

The energy efficiency is given by the ratio between the sum-rate achievable by the users to the total power consumed in the system:

$$EE = \frac{R_\Sigma}{P_{RF} + P_c} = \frac{K \cdot A \cdot R}{P_{RF} + K \cdot A \cdot P_k + A \cdot P_{BS}} \quad \left[ \frac{\text{bits}}{\text{J Hz}} \right]. \quad (14)$$

By assuming normalized bandwidth, the normalized sum-rate is computed as the total of data rates transmitted by all users under maximum fairness. In this case all users attain the same capacity ( $R_{ak} = R$ ) and consequently:

$$R_\Sigma = \sum_{a=1}^A \sum_{k=1}^K R_{a,k} = K \cdot A \cdot R. \quad (15)$$

In Eq. (14),  $P_c$  is the circuitry power consumption required to signal processing, cooling etc, at the transmitters and receivers. For simplicity, only the  $P_k$  and  $P_{BS}$  parts were explicitly considered, where the power required to operate the circuitry of each user antenna is  $P_k$ , and to operate the circuit components of each BS antenna is  $P_{BS}$ . Besides, the total RF power consumption is given by:

$$P_{RF} = \varrho \cdot \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} P = \varrho \cdot P \quad [\text{W}] \quad (16)$$

where  $\varrho \geq 1$  is the RF power amplifier inefficiency, assumed equal for all of them feeding each Tx antenna. Finally,  $P_c$  is the circuitry power consumption required to signal processing at the transmitter side, assumed be a fixed value per user.

### 3. Optimization of downlink MIMO-NOMA system

In this section, we perform the optimization of MIMO-NOMA system in two different ways. First, we obtained the minimum total power and equal rate for all users, using distributed equal-rate optimal power allocation (ERPA) strategy. Second, we computed the EE optimization to find the equal data rate  $R$  that give maximum EE. The concept of fairness is presented in the next subsection to better understand the process of obtaining its maximum value, followed by the problem of power optimization and EE optimization.

### 3.1. Fairness

Fairness can be defined by the **Jain's fairness index** [19,20]:

$$\mathcal{F}_{\text{Jain}} = \frac{\left(\sum_{a=1}^A \sum_{k=1}^K R_{a,k}\right)^2}{A \cdot K \sum_{a=1}^A \sum_{k=1}^K R_{a,k}^2}. \quad (17)$$

In the optimization design of this work, the equal data rate is adopted for all users; hence, from Eq. (17), the full fairness results  $\mathcal{F}_{\text{Jain}} = 1$ , being its maximum value.

### 3.2. Power allocation based on inverse of the channel response (ICR)

The technique used to define the proportional power to each user in [21] is given by the *Inverse of the Channel Response* (ICR). We use this technique to compare the sum-rate with egalitarian rates among all users obtained in next subsection 3.3. The power assignment for the  $j$ th user belonging to the  $i$ th cluster is defined as:

$$P_{ij} = \underbrace{\left(\sum_{a=1}^A \sum_{k=1}^K \frac{1}{\frac{\sigma^2 |\mathbf{v}_{a,k}|^2}{|\mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_a|^2}}\right)^{-1}}_{\beta'_{ij}} \cdot \left(\frac{|\mathbf{v}_{i,j}|^2}{|\mathbf{v}_{i,j}^H \mathbf{H}_{i,j} \mathbf{w}_i|^2}\right)^{-1} \cdot P. \quad (18)$$

In Eq. (18), the term  $\beta'_{i,j}$  is constituted by the product of the sum of the inverse of channel gain of all users of the cell in first parentheses and the inverse of the channel gain of the user of interest in the second parentheses. This equation represents the power proportion of each user given by the inverse of the normalized channel gain  $\beta'_{ij}$  multiplied by the total power. In this case, the power proportion will be higher for users with lower channel gain and lower for users with higher channel gain.

### 3.3. Equal-rate optimal power allocation (ERPA) in NOMA systems

Power proportions to be considered for each user and cluster to attain the minimum total power with equal-rate for all NOMA users with the distributed *equal-rate optimal power allocation* (ERPA) strategy is given by

$$\begin{aligned} & \min_{(P, \beta_{a,k})} \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} P \\ \text{s.t. } & \text{c.1 } \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} = 1 \\ & \text{c.2 } \beta_{a,k} > 0; P > 0 \\ & \text{c.3 } \beta_{a,k} = (2^R - 1) \frac{\sigma^2 |\mathbf{v}_{a,k}|^2}{P |\mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_a|^2} + (2^R - 1) \sum_{i=1}^{k-1} \beta_{a,i} \\ & \text{where } \begin{cases} 2 \leq k \leq K \\ 1 \leq a \leq A \end{cases} \\ & \text{c.4 } \beta_{a,1} = \frac{(2^R - 1) \sigma^2 |\mathbf{v}_{a,1}|^2}{P |\mathbf{v}_{a,1}^H \mathbf{H}_{a,1} \mathbf{w}_a|^2} \text{ to } 1 \leq a \leq A \end{aligned} \quad (19)$$

where the set of constraints is convex. To deal with this, let us show that the constraints c.1-c.4. can be incorporated into a convex cost function. The c.1 can be developed by replacing the restrictions c.3 and c.4 into the constraint c.1 subject to the constraints of c.2, obtaining the following equation:

$$f = \sum_{a=1}^A \sum_{k=2}^K \left[ (2^R - 1) \sum_{i=1}^{k-1} \beta_{a,i} + (2^R - 1) \frac{\sigma^2 |\mathbf{v}_{a,k}|^2}{P |\mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_a|^2} \right] +$$

$$\sum_{a=1}^A (2^R - 1) \frac{\sigma^2 |\mathbf{v}_{a,1}|^2}{P |\mathbf{v}_{a,1}^H \mathbf{H}_{a,1} \mathbf{w}_a|^2} - 1 \quad (20)$$

Hence, to be convex, the Hessian of Eq. (20) must be positive semi-definite [15], that is,  $\Delta^2 f > 0$ . Therefore all the eigenvalues of the Hessian must be greater than or equal to zero. After obtaining the Hessian and applying the definition of eigenvalues in Eq. (20), the following characteristic polynomial equally can be obtained:

$$\begin{aligned} -\lambda^M + \frac{2(2^R - 1)\sigma^2}{P^3} \sum_{a=1}^A \sum_{k=2}^K \frac{|\mathbf{v}_{a,k}|^2}{|\mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_a|^2} \lambda^{M-1} + \\ \frac{2(2^R - 1)\sigma^2}{P^3} \sum_{a=1}^A \frac{|\mathbf{v}_{a,1}|^2}{|\mathbf{v}_{a,1}^H \mathbf{H}_{a,1} \mathbf{w}_a|^2} \lambda^{M-1} = 0 \end{aligned} \quad (21)$$

Taking  $\lambda^{M-1}$  in evidence, there are  $M - 1$  eigenvalues equal to zero, and therefore the value obtained for the last one is:

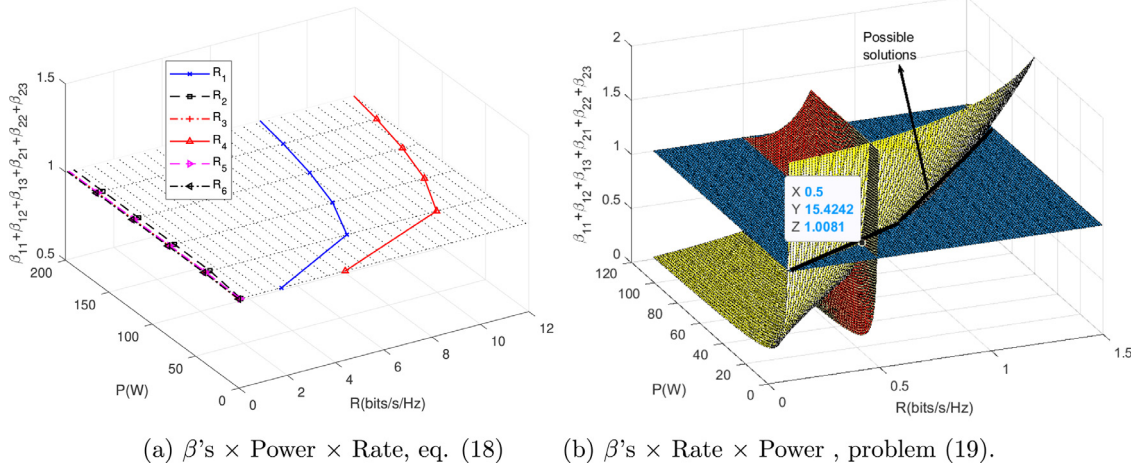
$$\lambda = \frac{2(2^R - 1)\sigma^2}{P^3} \left( \sum_{a=1}^A \sum_{k=2}^K \frac{|\mathbf{v}_{a,k}|^2}{|\mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_a|^2} + \sum_{a=1}^A \frac{|\mathbf{v}_{a,1}|^2}{|\mathbf{v}_{a,1}^H \mathbf{H}_{a,1} \mathbf{w}_a|^2} \right) \quad (22)$$

Since the variable  $P$  is positive as well as the constants  $R, \sigma^2, \frac{|\mathbf{v}_{a,k}|^2}{|\mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_a|^2}$  and  $\frac{|\mathbf{v}_{a,1}|^2}{|\mathbf{v}_{a,1}^H \mathbf{H}_{a,1} \mathbf{w}_a|^2}$ , the eigenvalues are positive or equal to zero. Hence, Eq. (20) is semi-definite positive.

In Fig. 2(a), the values of  $R_{a,k}$  for each user were obtained by applying the ICR principle, Eq. (18). Notice that in this sample scenario, the two clusters have similar distances between users, so the values of the respective  $R_{a,k}$  are close. Hence, applying (18), the data rate  $R_{a,k}$  values do not result in maximum fairness. By applying the ERPA formulation and solving the optimization problem in (19), the optimal values of  $\beta_{a,k}$  and  $P$  that minimizes the overall transmit power according to Fig. 2(b) while holding the equal data rate for all users are obtained ("Possible solutions" curve). Indeed, for a fixed data rate  $R$ , the total power is minimized by the objective function without violating the four constraints of the problem. The constraints form a collection of convex sets and therefore the intersection of these sets is convex. According to Fig. 2(b), the constraints defined in (19) form a closed set. In this way, the presented optimization problem has a true global solution. Fig. 2(b) depicts the subspace given by the condition of  $\beta_{11} + \beta_{12} + \beta_{13} + \beta_{21} + \beta_{22} + \beta_{23} = 1$  in the Z axis identified by the blue color plane, and in this example, the X axis has the value of  $R = 0.5$  bits/s/Hz for each user and is represented by the vertical plane. The feasible solution is the intersection by the three planes, resulting in a value of  $P \approx 15.42$  W for total power  $P$  in Y axis. However, the power proportion  $\beta_{a,k}$  values in question cannot be obtained so easily in practical MIMO-NOMA system operation conditions. There is a need to use an optimization tool. The trace in Fig. 2(b) shows the feasible solutions that can be given by the problem (19). The marker in the figure identifies the optimized values for the data rate in x-axis, the total power in the y-axis, and  $\sum_{a=1}^A \sum_{k=1}^K \beta_{a,k}$  in the z-axis.

### 3.4. Energy efficiency optimization problem

The energy efficiency optimization variables considered are data rate, power and the power proportion between users in which the objective function is given by Eq. (14). Such equation is quasi-concave [22]. We use the Dinkelbach method to turn the quasi-concave objective function into a concave one [23], as expressed by the problem (23). Hence, the EE optimization design in MIMO-NOMA system to obtain the optimal power allocation



**Fig. 2.** Closed set for  $\alpha = 3$ ,  $\sigma_n^2 = 1 \mu\text{W}$ ,  $A = 2$  clusters and  $K = 3$  users/cluster with distances  $d_{11} = 0.08$  km,  $d_{12} = 0.15$  km,  $d_{13} = 0.28$  km,  $d_{21} = 0.07$  km,  $d_{22} = 0.15$  km and  $d_{23} = 0.27$  km.

$P$ , under equal data rate  $R$  (maximum rate fairness) and power proportion  $\beta_{a,k}$  values for all users and different clusters can be defined by an equivalent problem as follows:

$$\begin{aligned}
 & \max_{(P, \beta_{a,k}, R)} K \cdot A \cdot R - q^n \left( \varrho \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} P + K \cdot A \cdot P_K + A \cdot P_{BS} \right) \\
 \text{s.t. } & \text{c.1 } \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} = 1 \\
 & \text{c.2 } \beta_{a,k} > 0; P > 0; R > 0 \\
 & \text{c.3 } \beta_{a,k} = (2^R - 1) \frac{\sigma^2 |\mathbf{v}_{a,k}|^2}{P |\mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_a|^2} + (2^R - 1) \sum_{i=1}^{k-1} \beta_{a,i} \\
 & \text{to } \begin{cases} 2 \leq k \leq K \\ 1 \leq a \leq A \end{cases} \\
 & \text{c.4 } \beta_{a,1} = \frac{(2^R - 1) \sigma^2 |\mathbf{v}_{a,1}|^2}{P |\mathbf{v}_{a,1}^H \mathbf{H}_{a,1} \mathbf{w}_a|^2} \text{ to } 1 \leq a \leq A
 \end{aligned} \quad (23)$$

where  $q^n$  is an EE fixed value. A pseudo-code for the iterative MIMO-NOMA energy efficiency maximization based on Dinkelbach method described in (23) is depicted in Algorithm 1, where EE value at the  $n$ th iteration is determined by the previous power and rate allocation as:

$$\begin{aligned}
 q^{(n)} &= \frac{U_R(P^{(n-1)}, R^{(n-1)}, \beta_{a,k}^{(n-1)})}{U_T(P^{(n-1)}, R^{(n-1)}, \beta_{a,k}^{(n-1)})} \\
 &= \frac{K \cdot A \cdot R^{(n-1)}}{\varrho \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k}^{(n-1)} P^{(n-1)} + K \cdot A \cdot P_K + A \cdot P_{BS}}
 \end{aligned}$$

**Algorithm 1 EE-NOMA – Iterative power-rate allocation based on Dinkelbach method**

**Initialization:**  $n = 0$  and  $q^{(n)} = 0$   
**while**  $\varepsilon > 0$  **do**  
 For a given  $q^{(n)}$ , solve (23) to obtain  $P^{(n)}$  and  $R^{(n)}$  and  $\beta_{a,k}$   
**if**  $U_R(P^{(n)}, R^{(n)}, \beta_{a,k}^{(n)}) - q^{(n)} U_T(P^{(n)}, R^{(n)}, \beta_{a,k}^{(n)}) \leq \varepsilon$  **then**  
 $P^* = P^{(n)}, R^* = R^{(n)}, q^* = q^{(n)}, \beta_{a,k}^* = \beta_{a,k}^{(n)}$   
 to  $a = 1, \dots, A$  and  $k = 1, \dots, K$   
**else**  
 Set  $n = n + 1$  and  $q^{(n)} = \frac{U_R(P^{(n-1)}, R^{(n-1)}, \beta_{a,k}^{(n-1)})}{U_T(P^{(n-1)}, R^{(n-1)}, \beta_{a,k}^{(n-1)})}$   
**end if**  
**end while**

By analyzing the convexity of constraints with the variables total power, proportion of total power among users and data rate, it was found to be non-convex, since the Hessian eigenvalues with these variables result in a complex value. However, by setting data rate  $R$  (fixing the equal rate for all NOMA users, to attain maximum fairness), while maintaining total power  $P$  and the proportion of total power  $\beta_{a,k}$  among users as optimization variables, the convexity of constraints set is hold, similarly as posed by Eq. (20). Hence, in a same way as pointed out earlier, the set of constraints under these modified conditions results convex. As a consequence, the EE-NOMA maximization under user-rate fairness problem using Dinkelbach method can be re-stated as follows. In problem Eq. (24), again, the Dinkelbach method was deployed to turn the quasi-concave objective function into a concave one. Algorithm 2 describes the EE maximization adopting fixed data rate.

$$\begin{aligned}
 & \max_{(P, \beta_{a,k}) \in \mathfrak{R}^+} K \cdot A \cdot R^* - q^n \left( \varrho \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} P + K \cdot A \cdot P_K + A \cdot P_{BS} \right) \\
 \text{s.t. } & \text{c.1, c.2, c.3 and c.4}
 \end{aligned} \quad (24)$$

In this reduced problem, the fixed EE at the  $n$ th iteration is defined as:

$$\begin{aligned}
 q^{(n)} &= \frac{U_R(P^{(n-1)}, \beta_{a,k}^{(n-1)})}{U_T(P^{(n-1)}, \beta_{a,k}^{(n-1)})} \\
 &= \frac{K \cdot A \cdot R^*}{\varrho \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k}^{(n-1)} P^{(n-1)} + K \cdot A \cdot P_K + A \cdot P_{BS}}
 \end{aligned}$$

while the equal user-rate  $R$  is a pre-defined value. Algorithm 2 describes the EE maximization adopting fixed data rate  $R^*$ .

Algorithm 2 obtains the maximal EE for each fixed data rate constraint ( $R = R^*$ ). Thus, to find the optimal data rate that provides the optimal EE, i.e., the optimal solution, we need to scan a whole range of fixed data rates and then use a finer grid around the most promising solution. This is (obviously) a costly and inefficient computational procedure. On the other hand, although Algorithm 1 does not guarantee convergence to the global optimum because the problem is non-convex, our numerical results point out that it can obtain a performance as good as that obtained by the search using Algorithm 2, with a considerably

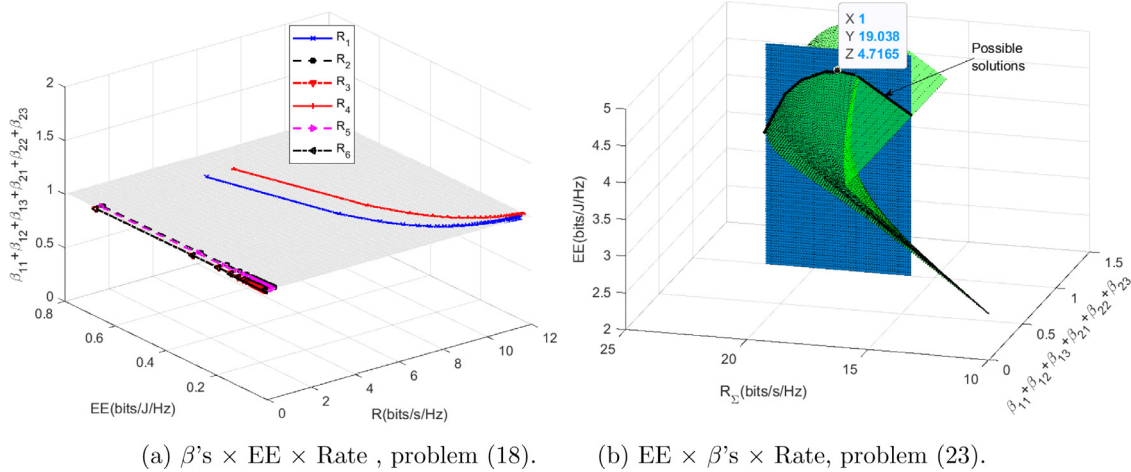


Fig. 3. Closed set for power  $P_{BS} = 1$  W/antenna BS and  $P_K = 100$  mW/antenna user,  $\alpha = 2$ ,  $\sigma_n^2 = 1 \mu\text{W}$ ,  $A = 2$  clusters and  $K = 3$  users/cluster with distances  $d_{11} = 0.08$  km,  $d_{12} = 0.15$  km,  $d_{13} = 0.28$  km,  $d_{21} = 0.07$  km,  $d_{22} = 0.15$  km and  $d_{23} = 0.27$  km.

**Algorithm 2 EE-NOMA – Iterative power allocation based on Dinkelback, equal data rate**

**Initialization:**  $n = 0$ ,  $q^{(n)} = 0$ , and  $R = R^*$  (Target equal data rate)  
**while**  $\varepsilon > 0$  **do**  
 For a given  $q^{(n)}$ , solve (24) to obtain  $P^{(n)}$  and  $\beta_{a,k}$   
**if**  $U_R(P^{(n)}, \beta_{a,k}^{(n)}) - q^{(n)}U_T(P^{(n)}, \beta_{a,k}^{(n)}) \leq \varepsilon$  **then**  
 $P^* = P^{(n)}$ ,  $q^* = q^{(n)}$ ,  $\beta_{a,k}^* = \beta_{a,k}^{(n)}$  to  $a = 1, \dots, A$  and  $k = 1, \dots, K$   
**else**  
 Set  $n = n + 1$  and  $q^{(n)} = \frac{U_R(P^{(n-1)}, \beta_{a,k}^{(n-1)})}{U_T(P^{(n-1)}, \beta_{a,k}^{(n-1)})}$   
**end if**  
**end while**

smaller computational burden. Since the solution provided by Algorithm 1 is so close to that obtained through extensive search of the optimal data rate and EE using Algorithm 2, for convenience, herein we will call such answer as the optimal point solution.

The  $R_{a,k}$  values were obtained for each user in Fig. 3(a) from the respective proportions of power given by Eq. (18), and they were used to calculate the EE values as in Eq. (14). As expected, the data rates for each user resulted in different values. On the other hand, optimization problem posed by Eq. (23) computes the optimal value of total power  $P$  and the data rate  $R$  that maximizes the EE, which generates Fig. 3(b). EE is maximized by the objective function without violating the four constraints of the problem. Elaborating further, Fig. 3(b) reveals the set of solutions of the optimization problem pointed out by the intersection of the two surfaces in the graph, as indicated by the “Possible solutions” curve: the sum of the power proportions equal to 1, i.e.,  $\beta_{11} + \beta_{12} + \beta_{13} + \beta_{21} + \beta_{22} + \beta_{23} = 1$ , in  $x$ -axis, and the other plane defined by EE in  $z$ -axis versus sum-rate users in  $y$ -axis for each value of the sum of  $\beta$ 's. The intersection between the planes is the possible solution to the problem given by Eq. (23). The values for maximum EE and maximum fairness for this example are approximately  $EE = 4.72$  bits/J/Hz and  $R = 3.17$  bits/s/Hz or  $R_\Sigma = 19.04$  bits/s/Hz. The curve with feasible solutions and example solutions are highlighted in the graph in Fig. 3(b), identified as:  $\sum_{a=1}^A \sum_{k=1}^K \beta_{a,k}$  in the  $x$ -axis, the sum users-rate in  $y$ -axis and EE in  $z$ -axis.

**4. Numerical results**

The parameter values used in the simulations are described in Table 1. The tool used for optimization was MatLab® *fmincon* function with the technique sequential quadratic programming (SQP). SQP algorithm [24] is a powerful iterative method for nonlinear optimization. It was considered a range of values from the lowest to the highest in order to obtain the behavior of the system with different number of users, path loss exponent values, target equal user's rate and  $\mathfrak{H}_D$ . In this way, one can trace a system's behavior considering different practical scenarios of interest, as shown in Fig. 4. In Scenarios 1, 2 and 3, users have a fixed distance from BS; however, the users in Scenario 4 have a uniform distribution for the respective  $\mathfrak{H}_D$ . The sum of distances between users and BS increases progressively from Scenarios 1, 2 to 3. The Fig. 4(a), (c) and (e) show the distribution of users in 3 clusters with 2 users/cluster, while Fig. 4(b), (d) and (f) present users in 2 clusters with 3 users/cluster.

**4.1. Proposed ERPA versus ICR policies in realistic MIMO-NOMA scenarios**

We demonstrated that there is actually an optimization problem in order to get the same rate for all users. We assumed the Scenarios 1 and 2 with  $\alpha = 2$  (LoS – line of sight). Fig. 5(a) and (c) show the ERPA (left bars) and ICR sum-rate (right bars), considering six users with fixed rates equal to 1, 2, 3 and 4 bits/s/Hz for 3 clusters with 2 users/cluster. Fig. 5(b) and (d) show the rates 0.5, 1, 1.5, 2 and 2.5 bits/s/Hz for 2 clusters with 3 users/cluster. The distributed ERPA strategy is compared with the one obtained by Eq. (18) with the same total power. We observe that only ERPA system kept the same rates for all users at the cost of worst sum rates at each rate. In ICR strategy, the further users receive a greater proportion of power that causes greater interference in other users and consequently reduces their rates and increases its own rate. The strongest user (closest) loses rate partially to ensure fairness, while distant users gain rate by allocating more power for them. This mechanism happened for all rates analyzed and this is the price paid to have fairness among users. Indeed, to maintain a rate of  $R = 2$  bits/s/Hz for all users in Scenario 1 with 3 clusters and 2 users/cluster and  $\alpha = 2$ , the total power required is 0.08 W, as shown in Fig. 5(a). For the same case, Fig. 5(b) shows total power required of 1.42 W for 2 clusters and 3 users/cluster. So, the total power needs to be 17.75 times greater for 2 clusters and 3 users/cluster, because there is greater

**Table 1**  
Adopted parameters for the simulation scenarios.

Parameter	Value
#Users	$M \in \{4; 24\}$
Path-loss exponent	$\alpha = [2.0; 3.0; 3.5]$
Power of white noise	$\sigma_n^2 = 0.1 \mu\text{W}$
Power budget	$P = 100 \text{ W}$
# clusters	$A \in \{2; 12\}$
Users/clusters	$K \in \{2, 3, 4\}$
# antennas per user	$U \geq A$ , with $U \in \{2; 12\}$ .
Target equal normalized user's rate	$R^* \in [0.1; 0.5; 1; 1.5; 2; 2.5; 3; 3.5; 4]$ [bits/s/Hz]
Circuits power	$P_{BS} = 1 \text{ W/antenna BS}$ $P_K = 100 \text{ mW/antenna user}$
RF power amplifier inefficiency	$\varrho = 1.4$
# MCS realizations	$10^4$
<b>Scenarios:</b> User distribution inside the cell: 6 users, 2 or (3) clusters	
Scen. 1 - $\sum_{a=1}^A \sum_{k=1}^K d_{ak} = 1.0$	$d_{11(11)} = 0.08, d_{12(12)} = 0.15,$ $d_{22(13)} = 0.28, d_{21(21)} = 0.07,$ $d_{31(22)} = 0.15, d_{32(23)} = 0.27$ [km]
Scen. 2 - $\sum_{a=1}^A \sum_{k=1}^K d_{ak} = 2.0$	$d_{11(11)} = 0.20, d_{12(12)} = 0.30,$ $d_{22(13)} = 0.50, d_{21(21)} = 0.10,$ $d_{31(22)} = 0.30, d_{32(23)} = 0.60$ [km]
Scen. 3 - $\sum_{a=1}^A \sum_{k=1}^K d_{ak} = 3.0$	$d_{11(11)} = 0.20, d_{12(12)} = 0.50,$ $d_{22(13)} = 0.80, d_{21(21)} = 0.20,$ $d_{31(22)} = 0.40, d_{32(23)} = 0.90$ [km]
Scen. 4 - $\mathfrak{N}_D$ (uniform. distrib.)	0.05, 0.10, 0.20 [km]
<b>Power Allocation Strategies</b>	
Inverse of the channel response (ICR)	$P_{ij}$ in Eq. (18)
Equal-rate optimal power allocation (ERPA)	Section 3.3

diversity or multiplexing capacity in the case of 3 clusters and 2 users/cluster, apart from the fact that the interference in each cluster is smaller.

Fig. 6 shows the sum-rate for ERPA and ICR distribution of total powers, in Scenarios 1, 2 and 3, for  $\alpha = 2$  and 3. The sums of ICR rates  $R_\Sigma$  are higher than ERPA distributions for the same total power. It is justified by the fact that ERPA distributions are more restrictive than ICR ones. This difference reduces as power increases, because power is no longer a restrictive factor. There is greater interference within the cluster for systems with 3 users/cluster, Fig. 6(a), so it requires high power to achieve the same rate as systems with 2 users/cluster, Fig. 6(c). For instance, in the environment with  $\alpha = 3$  and Scenario 1, the sum of ICR distribution rates with an average power budget of  $P_{opt}^* \approx 1.90 \text{ W}$  is  $R_\Sigma = 10.32 \text{ bits/s/Hz}$ , while the sum of ERPA distribution rates is  $R_\Sigma = 3.00 \text{ bits/s/Hz}$  for the same power budget of 1.90 W for 2 users/cluster, Fig. 6(c). The ratio between the ERPA and ICR sum rates increases as the sum rate increases, as shown in Fig. 6(b) and (d). The slopes of the curves are similar. The ratio between ERPA and ICR sum rates is close to ninety percent for high powers. The most restrictive scenarios need more power to reach this level. For 3 users/cluster, the power required for stabilization is greater because there is greater interference between users in the cluster.

Fig. 7 depicts the variation of power ratios of users  $\beta_{opt}^*$  as a function of the same rate to the six users given by  $R$  in bits/s/Hz. The Fig. 7(a) and (c) show the results for  $\alpha = 2$  in the configuration of 3 users/cluster and 2 users/cluster, respectively, in Scenarios 1, 2 and 3. Fig. 7(b) and (d) show the results of Scenario 2 for  $\alpha = 2, 3$  and 3.5 in the configuration of 3 users/cluster and 2 users/cluster, respectively.

In Scenarios 1 and 3, the power distribution among users assumes close values because users have a distribution with similar proportion of distances in the cell, Fig. 7(a). In cluster 1 of Scenario 2, users have marginal values of path loss and consequently nearby channel gain. As the gains are close, the user with the lowest channel gain has an increase in the proportion

of power, data rate and total power. Near values of total power proportion cause more interference between users. In cluster 2 of Scenario 2, the proportions of distances between users are greater than in Scenarios 1 and 3 and cluster 1 of Scenario 2. This leads to an increase in the proportion of power for the user with the greatest path loss as the data rate and total power of users increases. It can be justified by the need for the most critical user to have more power to maintain the same data rate, due to high interference from users of the same cluster in signal detection.

In Fig. 7(a) and (b), there is greater interference between users within the cluster, for 3 users/cluster situation, which results in a more pronounced proportion of power for users with less channel gain. This proportion of power increases the data rate. This process occurs in reverse for users with higher channel gain. Therefore, there is little variation in power proportions with the increase in the rate of users in this cluster. Users of cluster 2 in the Scenario 2 have proportionally greater distances between them.

The proportion of power for the user with the greatest path loss  $\beta_{23}$  results in a greater value of the proportion of power for low values of data rate  $R$ , when compared with the user that have the greatest path loss of cluster 1 in Scenario 2 given by  $\beta_{13}$ , Fig. 7(b). The greater proportion of power for the most distant user predominates in this case. The user with the greatest path loss in cluster 1 given by  $\beta_{13}$  will have a greater proportion of power associated with the increase of data rate  $R$ . The predominance of interference suffered by the user with the greatest path loss is caused by users with shorter distances. The total power proportion  $\beta_{23}$  values are greater for low  $R$  values due to the predominance of users with greater path loss coefficient  $\alpha$ . The values of  $\beta_{13}$  increase with the increment of  $R$  due to the predominance of user interference with channel gain of close values.

In Fig. 7(c), total power proportion  $\beta_{22}$  decreases when data rate  $R$  increases in clusters with greater distances. In contrast, total power proportions  $\beta_{12}$  and  $\beta_{32}$  increase when data rate  $R$  increases in clusters with shorter distances. There is a predominance of interference between users of the same cluster

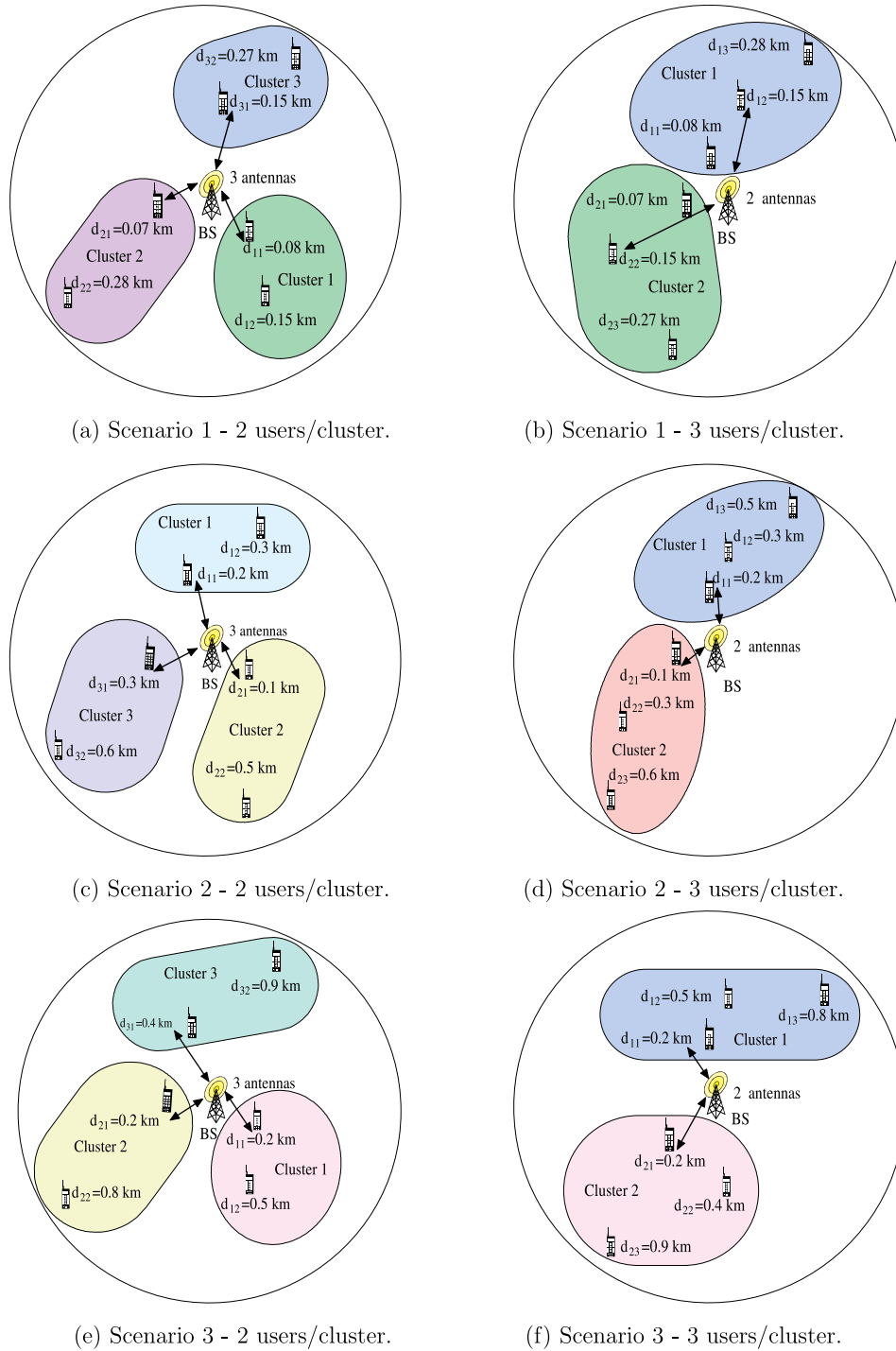


Fig. 4. Distribution of users in a NOMA system with 6 users as Table 1.

with shorter distance. In Fig. 7(d), we increased the path loss coefficient  $\alpha$ , and the behavior was similar to Fig. 7(b) and (c). The curve moved up because of the increase in data rate and the higher total power proportion values.

#### 4.2. Fairness

The fairness was calculated by Eq. (17). Once we have guaranteed that all users have the same rate  $R$ , the maximum fairness resulted in  $\mathcal{F}_j = 1$ . In Fig. 5, the total power  $P$  values, as well as the power proportion  $\beta$  for each user in Fig. 7 have been

obtained in order to guarantee the same rate for all users and therefore the maximum fairness. It can be seen that the values of the user rates are really the same, which shows the precision of the optimization process.

#### 4.3. Energy efficiency optimal operation points

The ERPA points, Eq. (24), for EE are shown in Fig. 8 for three fixed-value scenarios to  $\alpha = 2$ . The ERPA points, Eq. (23) of Fig. 8 are summarized in Table 2. For Scenario 2, a rate of  $R = 19.30$  and  $R = 11.30$  bits/s/Hz was obtained for a total power of  $P = 1.21$

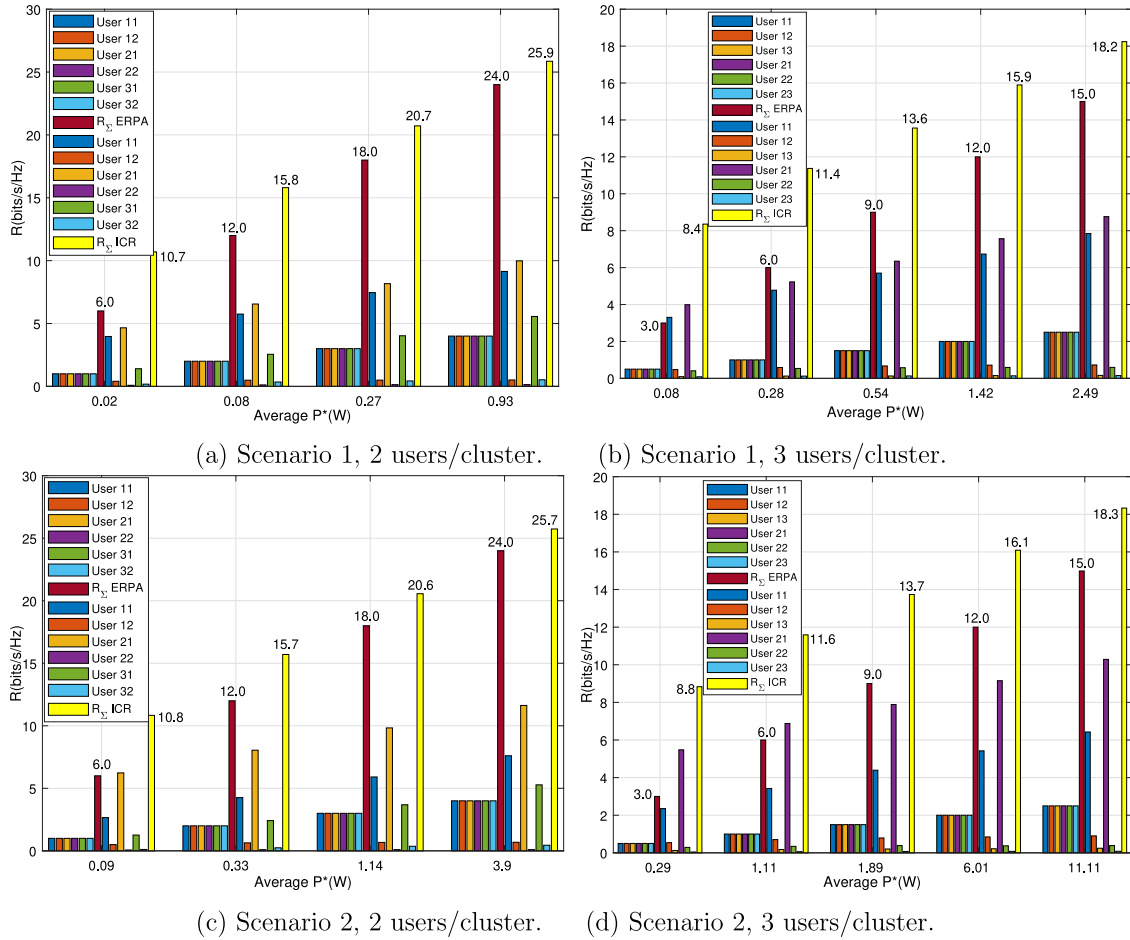


Fig. 5. Data rates versus allocated power considered LoS pathloss ( $\alpha = 2$ ) to 6 users. The ICR rates were obtained with the power proportion given by Eq. (18) and ERPA rates by problem (19). (a) Scenario 1 as in Fig. 4(a). (b) Scenario 1 as in Fig. 4(b). (c) Scenario 2 as in Fig. 4(c). (d) Scenario 2 as in Fig. 4(d).

and  $P = 1.92$  W, which resulted  $EE = 2.99$  and  $EE = 1.91$  bits/J/Hz with 2 users/cluster and 3 users/cluster, respectively. That is approximately three times less power to 1.56 times more rate and EE for 2 users/cluster. In the same analysis, the results for Scenario 3 show a reduction in approximately twenty per cent in all parameters, but with the same proportionality when comparing for 2 and 3 users/cluster. In Scenario 1, we obtain rates of  $R = 24.50$  bits/s/Hz and  $R = 14.90$  bits/s/Hz and total power of  $P = 0.85$  W and  $P = 1.41$  W, which resulted in  $EE = 4.10$  and  $EE = 2.82$  bits/J/Hz with 2 users/cluster and 3 users/cluster, respectively. That is approximately sixty percent less power to double the rate and EE for 2 users/cluster. For closest users in Scenario 1, less power is required to reach the specified rate, so that very low power reaches the highest rate and consequently higher EE, as can be seen in the points in asterisk of Fig. 8. The EE ERPA point Eq. (23) of users in Scenario 1 is higher than users of Scenarios 2 and 3 because users of the Scenario 1 are closer or less path loss and with greater evidence for total power according to Fig. 8(b) and (d). There are more interfering users in the same cluster for the 3 users/cluster situation, therefore, the data rate and energy efficiency are lower in these cases, as shown in Table 2. EE and data rate decrease when we move from Scenarios 1 to 2 and 3 because in the last scenarios users have less channel gain. By increasing the number of users per cluster from 2 to 3 there is a reduction in the ERPA distribution of EE and  $R$  and an increase in  $P$ . The interference between the users increase with the increase in the number of users per cluster, according to Table 2. The points generated by ERPA Eq. (23)

Table 2

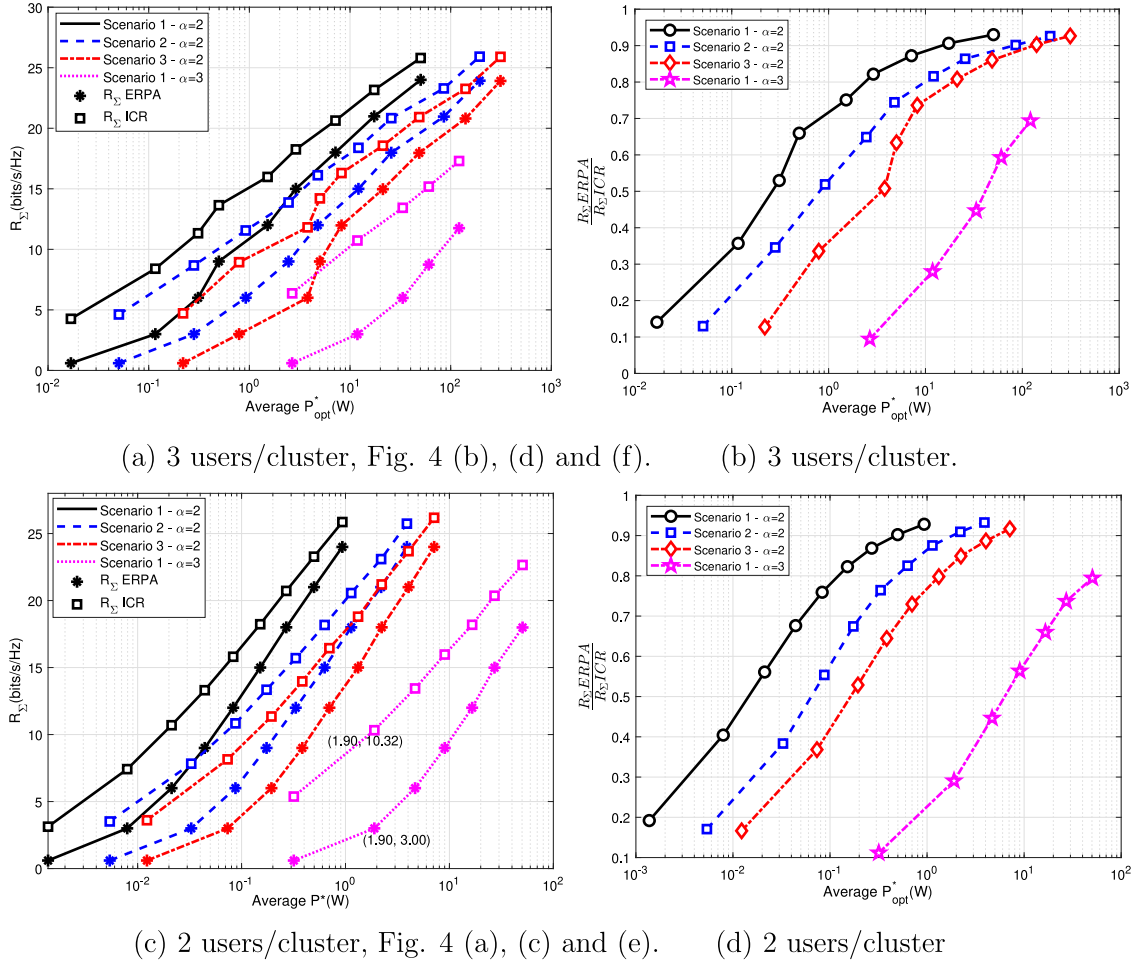
ERPA distribution to  $\alpha = 2$  and six users obtained according to graphs of Fig. 8.

Scenario	2 Users per cluster			3 Users per cluster		
	$EE_{opt}^*$	$P_{opt}^*$	$R_{\Sigma}^*$	$EE_{opt}^*$	$P_{opt}^*$	$R_{\Sigma}^*$
1	4.10	0.85	24.50	2.82	1.41	14.90
2	2.99	1.21	19.30	1.90	1.92	11.30
3	2.46	1.58	17.10	1.45	2.60	9.57

are consistent with the curve generated by ERPA Eq. (24). This shows the robustness of the results. The maximum EE results of the Eq. (23) are coinciding with the maximum of Eq. (24). This allows us to use Eq. (23) to obtain the maximum EE points in the following graphs.

#### 4.4. Resource efficiency optimal point as a function of the number of clusters with users uniformly distributed

In this section, the resource efficiency (RE), i.e., the optimal  $EE \times R_{\Sigma}$  trade-off is analyzed in terms of number of clusters  $A$ . The optimization problem was solved to obtain each point on the graph and the points were interconnected to identify the points with the same simulation characteristics. We adopted  $U = A$ , i.e., the number of antennas of each user and in the BS is equal to the number of clusters. We considered  $10^4$  MCS (Monte Carlo Simulations) realizations for each number of users



**Fig. 6.**  $R_\Sigma$  versus  $P$  and  $\frac{R_\Sigma ERPA}{R_\Sigma ICR}$  versus  $P_{opt}^*$  to  $\alpha = 2$  and 3. The ICR distribution were obtained with the power proportion, Eq. (18), and ERPA distribution by problem (19).

**Table 3**  
Combination of  $K$  users per cluster for  $A$  clusters to  $KA$  users in decreasing order of channel gain.

k	Cluster				
	# 1	# 2	# 3	...	# A
1	1	2	3	...	A
2	A + 1	A + 2	A + 3	...	2A
⋮	⋮	⋮	⋮	⋮	⋮
k	(k - 1)A + 1	(k - 1)A + 2	(k - 1)A + 3	...	kA
⋮	⋮	⋮	⋮	⋮	⋮
K	(K - 1)A + 1	(K - 1)A + 2	(K - 1)A + 3	...	KA

M. To define the combination of users for NOMA, the channel gain of  $KA$  users was ordered decreasingly. We considered the scenario in which the distances change at random with uniform distribution obtained by Scenario 4. The parameters and scenario considered are in Table 1. Users belonging to each cluster were defined according to Table 3, based on Ref. [25]. For example, for  $K = 2$  users/cluster and  $A = 3$  clusters, it results in 6 users ordered in descending order of channel gain, so that we have the combination of users (1,4), (2,5), and (3,6) for clusters 1, 2, and 3, respectively.

In Figs. 9, 10, 11, we considered the cell radius  $\mathfrak{R}_D = 50$  m,  $\mathfrak{R}_D = 100$  m and  $\mathfrak{R}_D = 200$  m, respectively. The optimal values for EE, sum-rates and total power were obtained by solving the

optimization problem posed by Eq. (23). In this case, EE, sum-rate  $R_\Sigma$  and total power  $P$  were considered as variables in the optimization process to ensure the equal rate for all NOMA users. In this setup, all users were uniformly distributed [26] within the cell of radius  $\mathfrak{R}_D$ , while the adopted circuitry power consumption was  $P_{BS} = 1$ W per antenna of BS and  $P_K = 100$  mW per antenna of user.

The optimization problem (23) was simulated to a fixed total number of users in the cell, number of users by clusters  $K$ , cell radius  $\mathfrak{R}_D$  and path loss coefficient  $\alpha$  for different values according to Scenario 4 and parameters of Table 1. Thereby, the maximum EE and the corresponding  $R_\Sigma$  were obtained for the maximum fairness at each point of the graphs in Figs. 9(a), 10(a) and 11(a). The respective powers  $P$  and the corresponding  $R$  of fairness or the data rate for each user at each point of the graphs are in Figs. 9(b), 10(b) and 11(b).

The maximum value of EE for  $\alpha = 2$  and 2 users/cluster resulted in 6 users in the cell, for 3 users/cluster in 9 users in the cell, and for 4 users/cluster in 12 users in the cell, that is equivalent to 3 clusters in the cell in the three cases. The curves obtained for  $\alpha = 2$  and 2, 3 and 4 users/cluster, in Figs. 9(b), 10(b) and 11(b), corroborate this result, as the power values increase exponentially, while  $R$  fairness increases marginally for more than 6, 9 and 12 users in the cell in the three cases.

Similarly for  $\alpha = 3$  and  $\mathfrak{R}_D = 50$  m, the maximum value of EE resulted 8 users in 4 clusters for 2 users/cluster, and 9 and 12 users in 3 clusters for 3 and 4 users/cluster respectively, as shown in Fig. 9(a). As we can see in Fig. 9(b), power increased

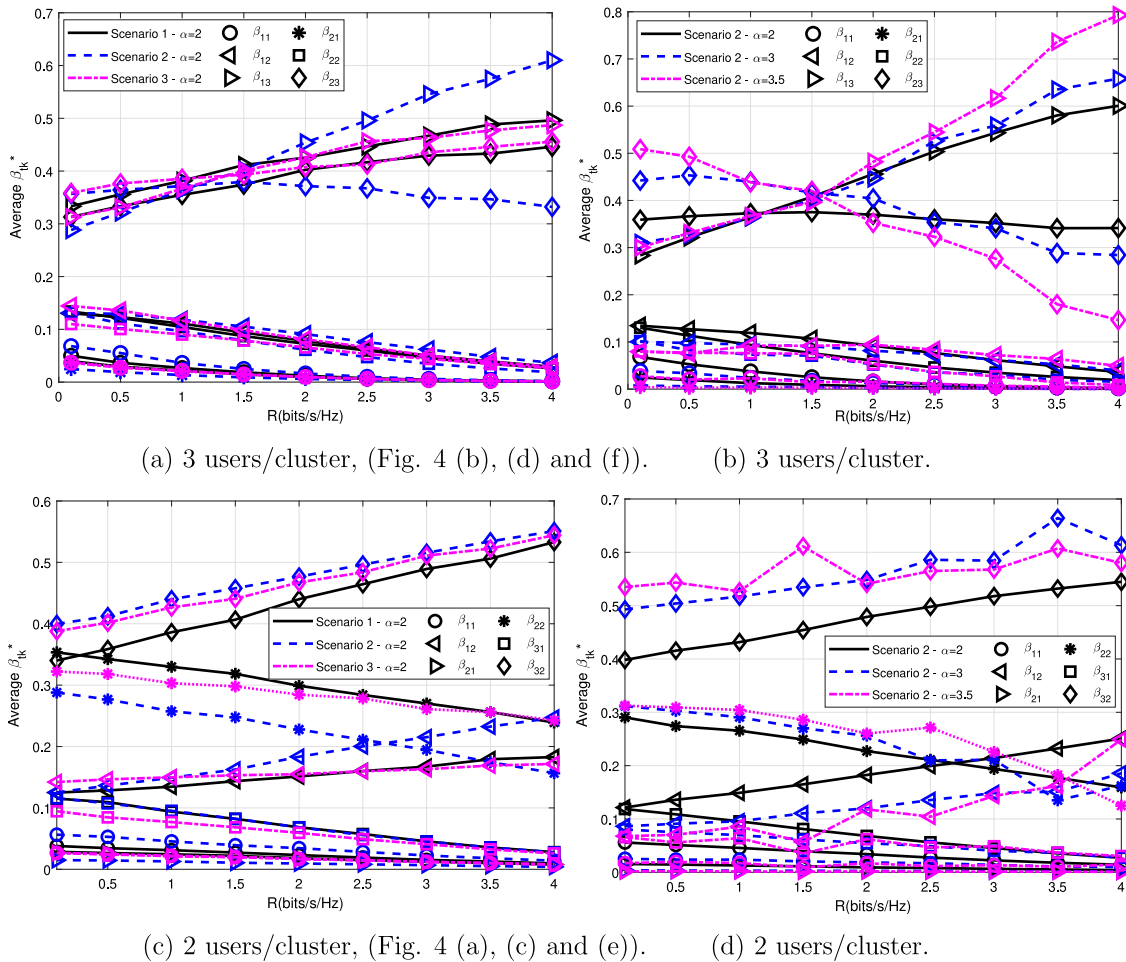


Fig. 7.  $\beta_{ik}^*$  versus data rate, problem (19).

exponentially for more than 8 users in cell with 2 users/cluster, for more than 9 users in cell with 3 users/cluster and for more than 12 users in cell with 4 users/cluster, while  $R$  fairness increased marginally. Therefore, EE resulted in lower values after these values.

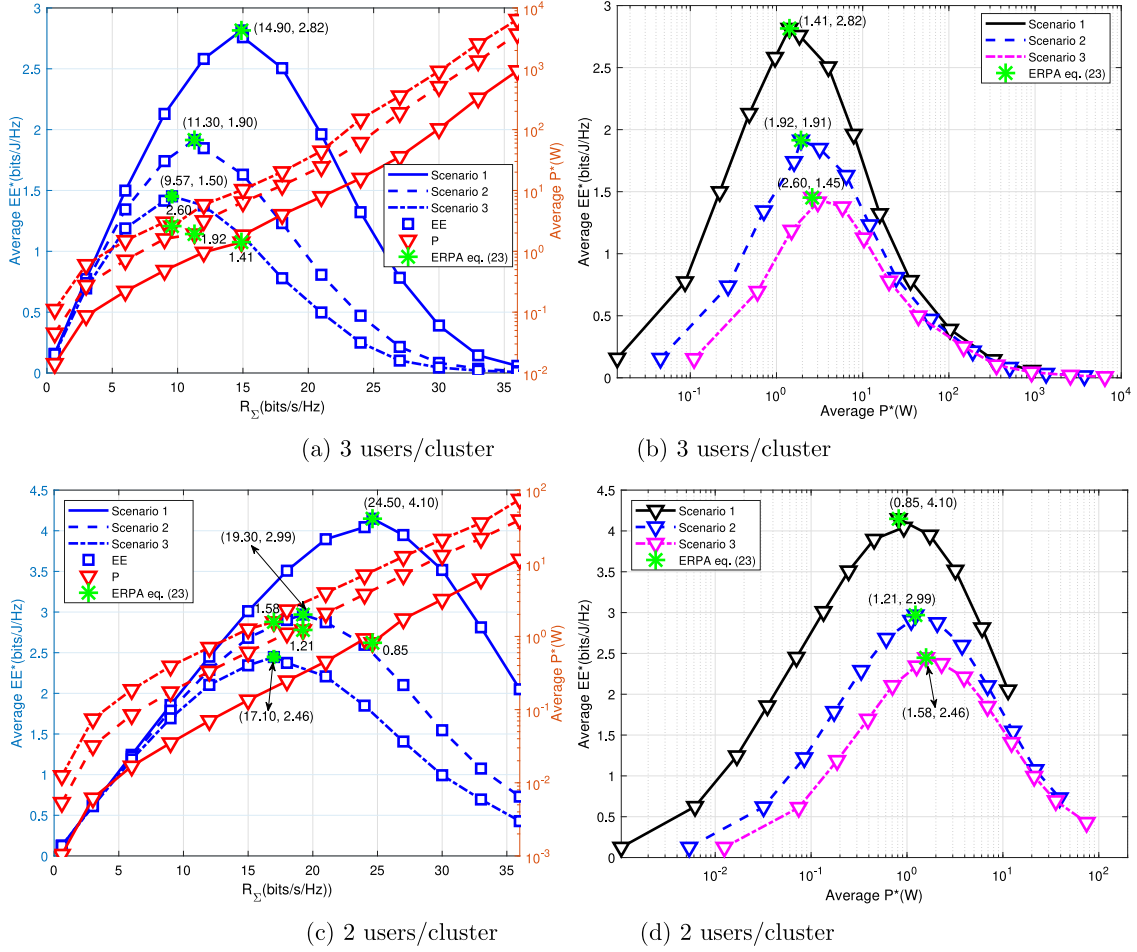
In Fig. 10(a), maximum EE and  $\mathfrak{R}_D = 100$  m resulted in 12 users in cell corresponding to 6 clusters for 2 users/cluster, 12 users in cell corresponding to 4 clusters for 3 users/cluster and 16 users in cell corresponding to 4 clusters for 4 users/cluster. Moreover, in Fig. 11(a), maximum EE and  $\mathfrak{R}_D = 200$  m also resulted in 12 users in cell corresponding to 6 clusters for 2 users/cluster, but 18 users in cell corresponding to 6 clusters for 3 users/cluster and 24 users in cell corresponding to 6 clusters for 4 users/cluster. In Figs. 10(b) and 11(b), we observed exponential growth for total power and marginal increase for the  $R$  fairness starting from the number of users that reached maximum EE. This is consistent with results presented previously in Figs. 10(a) and 11(a), and it shows that the maximum EE with maximum fairness for 2, 3 and 4 users/cluster depends on the cell radius and the path loss coefficient.

Finally, the number of users in the cell for maximum EE and  $\alpha = 3.5$  resulted in 12 users or 6 clusters in cell for 2 users/cluster for the three cell radius under analysis in Figs. 9(a), 10(a) and 11(a). In Fig. 9(a), for 3 and 4 users/cluster, it resulted in 12 and 16 users in 4 clusters in cell in both cases, respectively. For 3 users/cluster the maximum EE obtained was 21 users in 7 clusters with radius of  $\mathfrak{R}_D = 100$  and  $\mathfrak{R}_D = 200$  m according to 10(a)

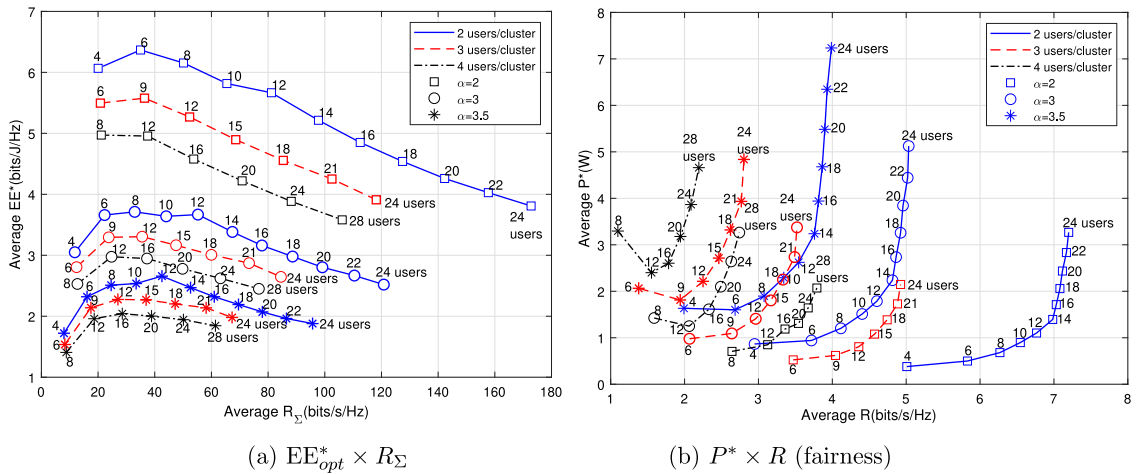
and 11(a) under analysis. The results for 4 users/cluster were 24 users in 6 clusters in cell with radius  $\mathfrak{R}_D = 100$  m according to Fig. 10(a). For cell rays  $\mathfrak{R}_D = 200$  m, it resulted in 28 users in 7 clusters in the cell according to Fig. 11(a). The results shown in Figs. 9(b), 10(b) and 11(b) confirm the displayed values.

## 5. Conclusions

The resource optimization problem in down-link of MIMO-NOMA system was solved in terms of proportion of energy distributed optimally among each user for an obtained equal transmission rate or maximum fairness condition that guarantees minimum total power and maximum EE. The power optimization problem proved to be a simple way to obtain the minimum total power and the appropriate distribution of power proportion across the users. The formulated system EE optimization problem was analyzed in two parts. First, the maximum EE and maximum fairness were obtained for fixed distances and data rate for each user. The optimum set points for the near-optimized variable data rate behaved very similarly to a fixed data rate analysis. In the second part, for a given cell radius, the analysis process considered the distances obtained randomly with uniform distribution of the users in two, three and four users per cluster. The data rate was an optimization variable obtained for different number of clusters in the cell. The goal was to find the approximate number of clusters in the cell that resulted in maximum EE or



**Fig. 8.**  $EE_{opt}^* \times R_{\Sigma}$ ,  $P_{opt}^* \times R_{\Sigma}$  and  $EE_{opt}^* \times P_{opt}^*$ , problem (24) to  $\alpha = 2$  and 6 users. The ERPA distribution (solution for the problem (23)) have been highlighted as green stars markers.



**Fig. 9.**  $EE \times R_{\Sigma}$  and  $P \times R$  (fairness) for 2, 3 and 4 users/cluster,  $\alpha = 2, 3$  and 3.5 and  $\mathfrak{R}_D = 50$  m, problem (23).

alternatively maximum fairness. For the adopted values of users-per-cluster ( $K$ ), number of clusters ( $A$ ), number of BS antennas per user ( $U \geq A$ ), path-loss coefficient ( $\alpha$ ), and cell coverage scenarios ( $\mathfrak{R}_D$ ), the numerical results indicated that the number of clusters-per-cell that maximizes the EE and the number of users with equal data rate depend on the number of users per cluster,

the path loss coefficient factor and the cell radius. Notably, with regard to the channel path-loss coefficient, a general value of  $\alpha$  impacts (increases) the number of clusters needed to achieve the highest system energy efficiency, as well as to attain maximum fairness.

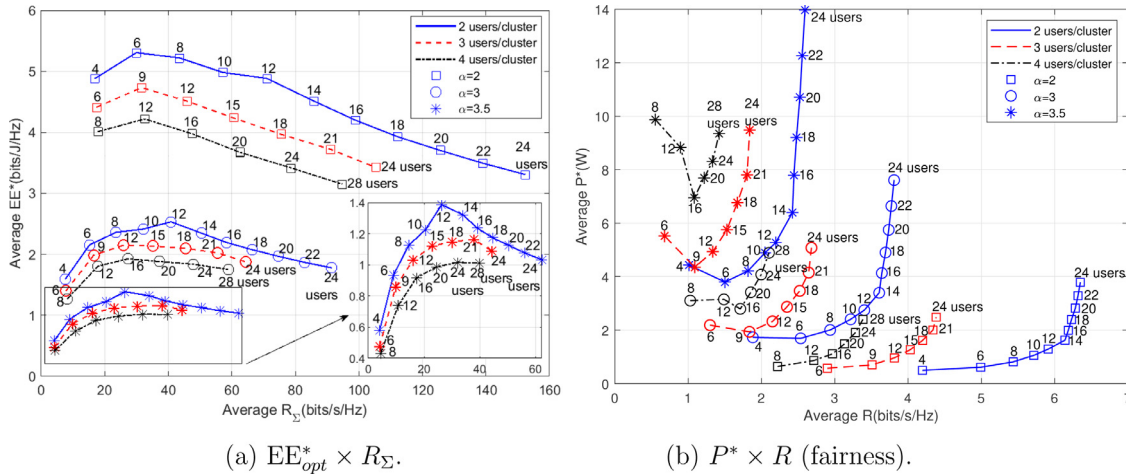


Fig. 10.  $EE \times R_{\Sigma}$  and  $P \times R$  (fairness) for 2, 3 and 4 users/cluster,  $\alpha = 2, 3$  and 3.5 and  $\mathfrak{H}_D = 100$  m, problem (23).

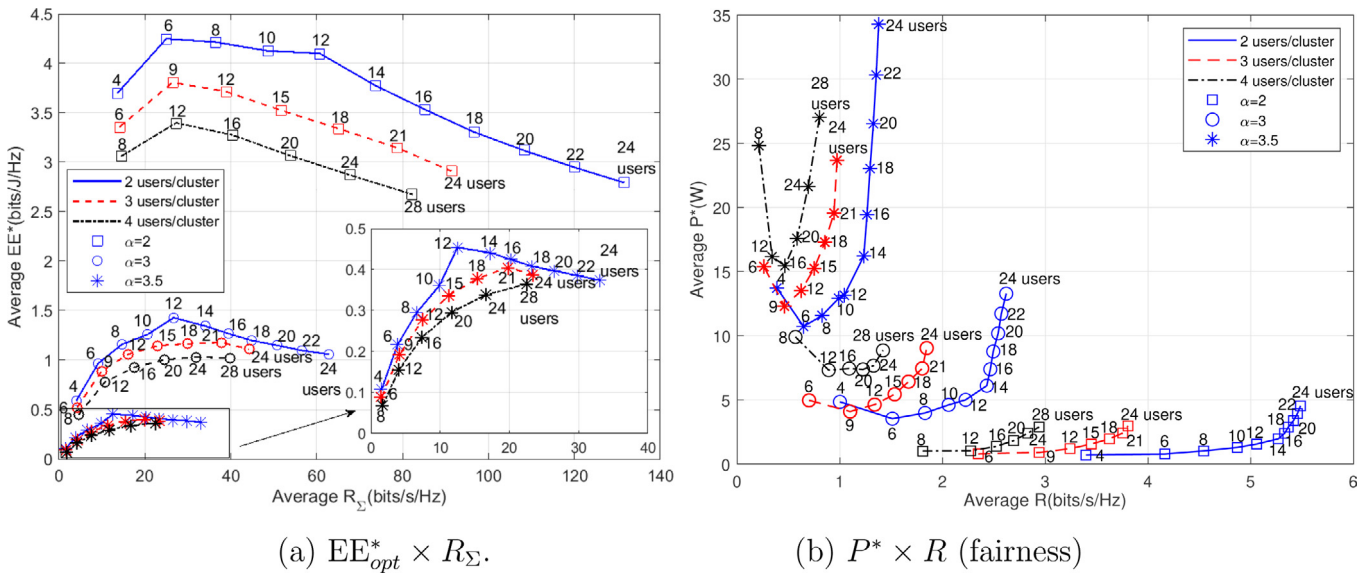


Fig. 11.  $EE \times R_{\Sigma}$  and  $P \times R$  (fairness) for 2, 3 and 4 users/cluster,  $\alpha = 2, 3$  and 3.5 and  $\mathfrak{H}_D = 200$  m, problem (23).

**Declaration of competing interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Acknowledgments**

This work was supported in part by the National Council for Scientific and Technological Development (CNPq) of Brazil under Grant 10681/2019-7; in part by Londrina State University, Parana State Government (UEL); and in part by the Coordenacao de Aperfeicoamento de Pessoal de Nivel Superior (CAPES), Brasil.

**References**

[1] A. Sayed-Ahmed, M. Elsabrouty, A.H.A. El-Malek, M. Abo-Zahhad, Energy efficient framework for multiuser downlink MIMO-NOMA systems, in: Eleventh International Workshop on Selected Topics in Mobile and Wireless Computing, 2018, pp. 48–54.  
 [2] M.D.S. Ali, E. Hossain, D.I. Kim, Non-orthogonal multiple access (NOMA) for downlink multiuser MIMO systems: User clustering, beamforming, and power allocation, IEEE Access 5 (2017) 565–577.

[3] W. Shin, M. Vaezi, B. Lee, D.J. Love, J. Lee, H.V. Poor, Coordinated beamforming for multi-cell MIMO-NOMA, IEEE Commun. Lett. 21 (1) (2017) 84–87.  
 [4] J.-M. Kang, I.-M. Kim, C.-J. Chun, Short papers deep learning-based MIMO-NOMA with imperfect SIC decoding, IEEE Syst. J. (2019) 1–4.  
 [5] Z. Ding, R. Schober, H.V. Poor, A general MIMO framework for NOMA downlink and uplink transmission based on signal alignment, IEEE Trans. Wireless Commun. 15 (6) (2016) 4438–4454.  
 [6] H. Wang, R. Zhang, R. Song, S.H. Leung, A novel power minimization precoding scheme for MIMO-NOMA uplink systems, IEEE Commun. Lett. 22 (5) (2018) 1106–1109.  
 [7] Y. Liu, G. Pan, H. Zhang, M. Song, On the capacity comparison between MIMO-NOMA and MIMO-OMA, IEEE Access 4 (2016) 2123–2129.  
 [8] Q. Sun, S. Han, C.-L. I, Z. Pan, On the ergodic capacity of MIMO NOMA systems, IEEE Wirel. Commun. Lett. 4 (4) (2015) 405–408.  
 [9] J. Ding, J. Cai, Efficient MIMO-NOMA clustering integrating joint beamforming and power allocation, in: IEEE Global Communications Conference, GLOBECOM, 2018, pp. 1–6.  
 [10] Z. Ding, F. Adachi, H.V. Poor, The application of MIMO to non-orthogonal multiple access, IEEE Trans. Wireless Commun. 15 (1) (2016) 537–552.  
 [11] M. Zeng, A. Yadav, O.A. Dobre, H.V. Poor, A fair individual rate comparison between MIMO-NOMA and MIMO-OMA, in: IEEE Globecom Workshops, IEEE Globecom Workshops, 2017, pp. 1–5.  
 [12] H. Al-Obiedollah, K. Cumanan, J. Thiyagalingamy, A.G. Burr, Z. Dingz, O.A. Dobrex, Sum rate fairness trade-off-based resource allocation technique for MISO NOMA systems, in: IEEE Wireless Communications and Networking Conference (WCNC), 2019, pp. 1–6.

- [13] M.M. Al-Wani, A. Sali, B.M. Ali, A.A. Salah, K. Navaie, C.Y. Leow, N.K. Noordin, S.J. Hashim, On short term fairness and throughput of user clustering for downlink non-orthogonal multiple access system, in: *IEEE Conference on Vehicular Technology (VTC)*, 2019, pp. 1–6.
- [14] J.L. Jacob, T. Abrao, Nonorthogonal multiple access systems optimization to ensure maximum fairness to users, *Trans. Emerg. Telecommun. Technol. - Wiley* 31 (2020) 1–14.
- [15] S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.
- [16] D.P. Bertsekas, *Convex Optimization Theory*, Athena Scientific, Belmont, MA USA, 2009.
- [17] M. Zeng, A. Yadav, O.A. Dobre, H.V. Poor, Energy-efficient power allocation for MIMO-NOMA with multiple users in a cluster, *IEEE Access* 6 (2018) 5170–5181.
- [18] M. Zeng, A. Yadav, O.A. Dobre, G.I. Tsiropoulos, H.V. Poor, Capacity comparison between MIMO-NOMA and MIMO-OMA with multiple users in a cluster, *IEEE J. Sel. Areas Commun.* 35 (10) (2017) 2413–2424.
- [19] R.K. Jain, D.-M.W. Chiu, W.R. Hawe, A quantitative measure of fairness and discrimination for resource allocation in shared computer system, *ACM Trans. Comput. Syst.* (1984) 1–38.
- [20] C.H. Wang, J.Y. Lin, J.M. Wu, Joint fairness and sum rate resource allocation for NOMA communications, in: *IEEE Conference on Standards for Communications and Networking, CSCN*, 2017, pp. 269–274.
- [21] M.M. El-Sayed, A.S. Ibrahim, M.M. Khairy, Power allocation strategies for non-orthogonal multiple access, in: *International Conference on Selected Topics in Mobile and Wireless Networking, MoWNeT*, 2016, pp. 1–6.
- [22] Y. Zhang, H.M. Wang, T.X. Zheng, Q. Yang, Energy-efficient transmission design in non-orthogonal multiple access, *IEEE Trans. Veh. Technol.* 66 (3) (2017) 2852–2857.
- [23] J. Tang, J. Luo, M. Liu, D.K. So, E. Alsusa, G. Chen, K.K. Wong, J.A. Chambers, Energy efficiency optimization for NOMA with SWIPT, *IEEE J. Sel. Top. Signal Process.* 13 (3) (2019) 452–466.
- [24] J. Nocedal, S.J. Wright, *Numerical Optimization*, second ed., Springer Science, LLC, 2006.
- [25] M.S. Ali, H. Tabassum, E. Hossain, Dynamic user clustering and power allocation for uplink and downlink non-orthogonal multiple access (NOMA) systems, *IEEE Access* 4 (2016) 6325–6343.
- [26] L. Devroye, *Non-Uniform Random Variate Generation*, Springer Science, LLC, 1986.



**Jaime L. Jacob** received his B.S. and M.Sc. degree in Electrical Engineering from CEFET - Centro Federal de Educação Tecnológica, Curitiba, Paraná, Brazil, in 1992 and Londrina State University - UEL, Londrina, Paraná, Brazil, in 2006, respectively. He worked at Telecommunications Company (TELEPAR) from 1992 to 2008. Since August 2008, he has been Professor of Telecommunications in Department of Electrical Engineering, UEL, Paraná, Brazil. He is a member of SBrT. His current research interests include communications and signal processing, especially NOMA, massive MIMO, detection and estimation, multicarrier systems, resource allocation and convex optimization aspects of 5G wireless systems.



**Cristiano Panazio** received the engineer's and master's degrees in electrical engineering in 1999 and 2001, respectively, from the State University of Campinas, Campinas, Brazil, and the Ph.D. degree in telecommunications from the Conservatoire National des Arts et Métiers, Paris, France, in 2005.

He is currently an Assistant Professor with the Laboratory of Communications and Signals, Escola Politécnica of the University of São Paulo, São Paulo, Brazil. His research interests include MIMO, multicarrier modulation, signal processing and optimization for

wireless communications and synchronization techniques.

Prof. Panazio served as an Associated Editor for the *Transactions on Circuits and Systems I – Regular Papers* from 2016–2018 and also he has been an ad hoc Editor for the *Journal of Communication and Information Systems (JCIS)* since 2016. He has been a TPC member in several symposiums and conferences. He is a senior member of the Brazilian Communication Society (SBrT) and is currently serving as the President of SBrT from 2020 to 2022.



**Taufik Abrão** (SM'12, SM-SBrT) received the B.S., M.Sc., and Ph.D. degrees in electrical engineering from the Polytechnic School of the University of São Paulo, São Paulo, Brazil, in 1992, 1996, and 2001, respectively. Since March 1997, he has been with the Communications Group, Department of Electrical Engineering, Londrina State University, Paraná, Brazil, where he is currently an Associate Professor in Telecommunications and the Head of the Telecomm. & Signal Processing Lab. He is a Productivity Researcher from the CNPq Brazilian Agency (Pq-1D).

From July–October 2018 he was with the Connectivity section, Aalborg University as a Guest Researcher. In 2012, he was an Academic Visitor with the Southampton Wireless Research Group, University of Southampton, Southampton, U.K. From 2007 to 2008, he was a Post-doctoral Researcher with the Department of Signal Theory and Communications, Polytechnic University of Catalonia (TSC/UPC), Barcelona, Spain. He has participated in several projects funded by government agencies and industrial companies. He is involved in editorial board activities of several journals in the telecommunications area and has served as TPC member in several symposiums and conferences. He has also served as an Associate Editor for the *IEEE ACCESS* since 2016, the *IET Journal of Engineering* since 2014, the *IET Signal Processing* since Dec-2018, and *JCIS-SBrT journal* since 2018. Previously, he served as AE of the *IEEE Communication Surveys & Tutorials* (2013–2017). Moreover, Prof. Abrao has been served as Executive Editor of the *ETT-Wiley journal* since 2016. He is a member of SBrT and a senior member of IEEE. His current research interests include communications and signal processing, especially massive MIMO, ultra-reliable low latency communications, detection and estimation, multicarrier systems, cooperative communication and relaying, resource allocation, as well as heuristic and convex optimization aspects of 5G wireless systems. He has supervised 27 M.Sc. and 4 Ph.D. students, as well as 3 postdocs, co-authored twelve book chapters on mobile radio communications and +280 research papers published in international journals and conferences.

**ANEXO C – Non-Linear  
Biobjective EE-SE Optimization for  
NOMA-MIMO Systems under  
User-rate Fairness and Variable  
Number of Users per Cluster**

(P3 – artigo publicado em 2021)



Contents lists available at ScienceDirect

## International Journal of Electronics and Communications

journal homepage: [www.elsevier.com/locate/aeue](http://www.elsevier.com/locate/aeue)

Regular paper

## Non-linear biobjective EE-SE optimization for NOMA-MIMO systems under user-rate fairness and variable number of users per cluster

Jaime L. Jacob<sup>a</sup>, Cristiane A. Pendeza Martinez<sup>b</sup>, André Luís Machado Martinez<sup>b</sup>, Taufik Abrão<sup>a,\*</sup><sup>a</sup> Department of Electrical Engineering, Londrina State University, Rod. Celso Garcia Cid - PR445, Po.Box 10.011, CEP: 86057-970, Londrina, PR, Brazil<sup>b</sup> Department of Mathematics, Universidade Tecnológica Federal do Paraná – Campus Cornélio Procopio Avenida Alberto Carazzai, 1640 - CEP 86300-000, Cornélio Procopio, PR, Brazil

## ARTICLE INFO

## Keywords:

Non-orthogonal multiple access (NOMA)  
 Multiple-input and multiple-output (MIMO)  
 Optimization  
 Spectrum efficiency  
 Energy efficiency (EE)  
 Fairness  
 EE-SE trade-off  
 Sequential quadratic programming (SQP)  
 Dinkelback method (DK)  
 $\epsilon$ -Constraint scalarization method ( $\epsilon$ -C)  
 Hill Climbing (HC) greedy local search  
 Weighted sum (WS) method

## ABSTRACT

The trade-off between energy efficiency (EE) and spectrum efficiency (SE) in downlink (DL) multiple-input and multiple-output (MIMO) non-orthogonal multiple access (NOMA) systems under equal-rate constraint has been investigated. The maximum fairness of the NOMA-MIMO system is attained when it is assured the same data rate for all active users in the system. The SE-EE trade-off was formulated as a multi-objective (MOO) problem. The optimization problem was solved by using primarily  $\epsilon$ -Constraint ( $\epsilon$ -C) scalarization method combined with nonlinear programming (NLP) techniques, as sequential quadratic programming (SQP) and Dinkelback method (DK). The proposed iterative-analytical  $\epsilon$ -C-DK and  $\epsilon$ -C-SQP methods were able to achieve optimal EE-SE solutions at the Pareto frontier. Besides, both proposed iterative-analytical optimization methods are compared with a greedy local search heuristic method, by combining the  $\epsilon$ -C with the Hill Climbing (HC) heuristic and weighted sum (WS) scalarization method. Numerical results have demonstrated the ability of the  $\epsilon$ -C method in finding a diversity of EE-SE NOMA-MIMO solutions in the Pareto frontier. Furthermore, the proposed  $\epsilon$ -C-SQP method is able to attain a much more variety of solutions in the Pareto frontier in a remarkably reduced computational time.

## 1. Introduction

Non-Orthogonal Multiple Access (NOMA) is one of the key techniques for 5G systems since NOMA can significantly improve the spectral efficiency (SE) and user fairness in mobile wireless communications. NOMA allows multiple users to share time and frequency resources in the same spatial layer via power domain or code domain multiplexing [1]. The users with poorer channel conditions are destined to more transmission power. These users decode their own messages by treating others' information as noise. The users with better channel conditions use the successive interference cancellation (SIC) technique to cancel the respective interference generated by the other paired users. Such users first detect messages from other paired users with low channel state information (CSI) conditions and then detect their own message by removing information from the other users [2].

Multiple-input multiple-output (MIMO) communications have the potential to achieve significant gains in the transmission rate of the

system [3]. Hence, in a downlink (DL) NOMA-MIMO system, multiple receiving antennas from different users with different channel gains are grouped into different NOMA clusters based on the disparity of received powers. Hence, the user equipments (UEs) in each cluster works based on SIC detection.

The spectral efficiency (SE) optimization problem under user-rate fairness and NOMA configuration present limitations and challenges. Improving the SE unilaterally leads to high energy consumption and, thus, increases the overall cost of operating the network. By using as much power as possible to maximize SE, the overall energy efficiency (EE) is drastically reduced [4]. SE and EE are two important metrics for improving the resource efficiency (RE) in wireless communication systems with lower energy costs per bit transferred and user data rate increasing simultaneously. However, both EE and SE metrics are often conflicting, *i.e.*, increasing one metric lead to a reduction in the other, and vice versa [5]. Therefore, a better optimizing approach consists in deal with the trade-off between SE and EE metrics, by adopting a multi-

\* Corresponding author.

E-mail addresses: [jaimel.jacob@uel.br](mailto:jaimel.jacob@uel.br) (J.L. Jacob), [crismartinez@utfpr.edu.br](mailto:crismartinez@utfpr.edu.br) (C.A. Pendeza Martinez), [martinez@utfpr.edu.br](mailto:martinez@utfpr.edu.br) (A.L. Machado Martinez), [taufik@uel.br](mailto:taufik@uel.br) (T. Abrão).<https://doi.org/10.1016/j.aeue.2021.153870>

Received 26 January 2021; Accepted 18 June 2021

Available online 26 June 2021

1434-8411/© 2021 Elsevier GmbH. All rights reserved.

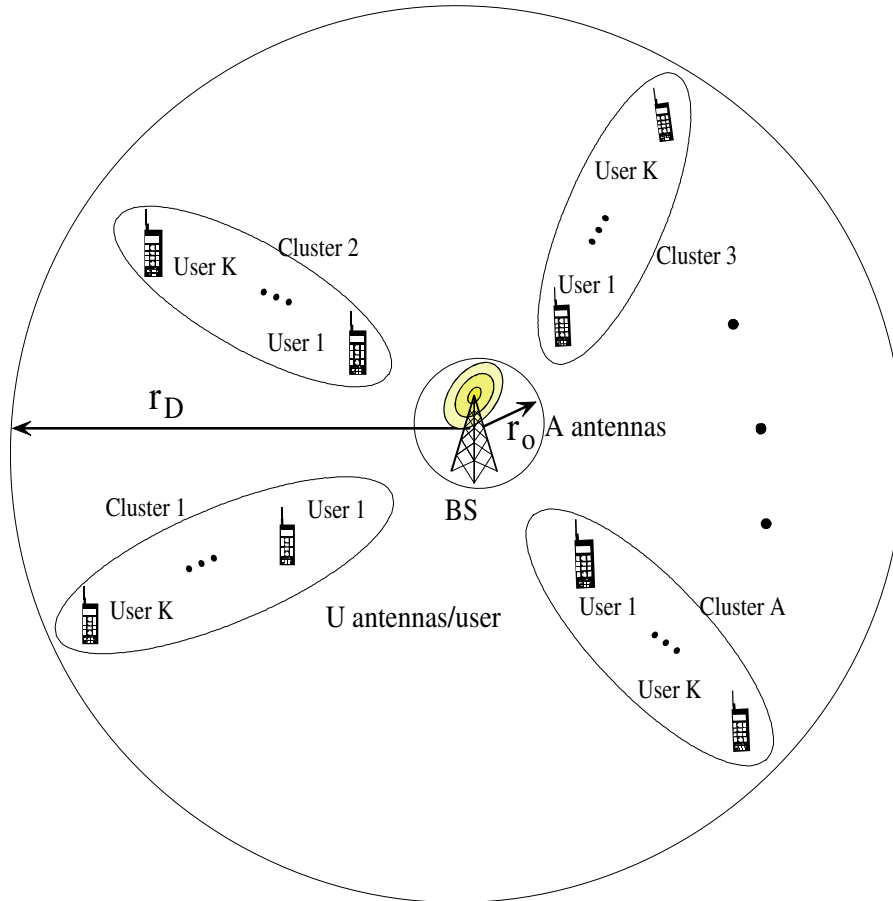


Fig. 1. Cellular system with a BS with  $A$  antennas,  $A \times K$  users with  $A$  clusters,  $K$  users/clusters and each user has  $U$  antennas ( $U \geq A$ ).

objective optimization description. Indeed, recently, several works in the literature have been dedicated to solving multi-objective optimization problems related to communication networks.

The EE-NOMA optimization problem is discussed in [6] by finding the minimum total power needed to reach a specified equal rate per user. Based on the EE-SE trade-off optimization problem with maximum fairness, [6] calculates these values for each number of users and presents one with the highest resource efficiency (RE) in the cell. The current contribution differs from [6] in the sense that herein we consider a variable number of users per cluster while adopting a multi-objective optimization approach.

Other works adopting multi-objective optimization approaches include [7–13]. In [7], a novel technique, namely superior population generation algorithm (SPGA), modifies the regular genetic algorithm (GA) to incorporate the capabilities of conventional techniques. The algorithm was developed to solve the SE-EE trade-off optimization problem for fifth-generation (5G) heterogeneous networks, which uses key future technologies such as carrier aggregation and cognitive radio.

The EE-SE trade-off in DL orthogonal frequency division multiplexing access (OFDMA) systems are investigated in [8], whilst considering the channel estimation cost and the corresponding effect of imperfect CSI on the SE and EE. The problem is formulated as a multi-objective optimization (MOO) to determine the optimal pilot transmission power, data transmission power, and subcarrier assignment. The original MOO problem is transformed into a single-objective optimization (SOO) problem, which characterizes a non-convex mixed-integer nonlinear programming (NMINP) and NP-hard. To address it, the authors propose an efficient algorithm by adopting alternating convex optimization sub-problem approach in the lower-power region, as well as an approximate conversion and branch-and-bound methods in

the high-power region. Simulation results validate the attainable performance in terms of EE-SE trade-off optimization.

Authors in [9] investigate the EE-SE trade-off in DL NOMA-based heterogeneous networks (HetNets). The EE-SE trade-off problem is formulated as a bi-objective optimization (BOO) problem. Under the constraints of maximum transmit power of small cell users (SUEs) and minimum rate requirements, both EE and SE are maximized. In order to make the MOO tractable, a weighted sum method is used to transform it into an SOO problem, which enables the network providers to dynamically tune the weight factor to adapt to different EE and SE design requirements. To solve the transformed SOO, an SUE allocation algorithm is proposed to determine the multiple SUEs for each sub-channel, based on which a power proportional factor determination algorithm for SUEs and a novel power allocation algorithm for sub-channels are proposed. Numerical results show the fast convergence feature, and demonstrate that the EE and SE values attained by the proposed algorithms are Pareto optimal solutions and superior to that obtained under conventional orthogonal multiple access (OMA) schemes.

The EE-SE trade-off in a downlink of a single-cell massive MIMO transmission with statistical CSI, aiming to optimize the system RE is investigated in [10]. Moreover, in [11] the authors propose a novel low-complexity resource allocation scheme to attain a superior EE-SE trade-off in multicarrier NOMA systems with the constraint of user fairness. The objective is to maximize both EE and SE metrics through the number of users and power allocation among them. The network, in turn, has different communication demands in different scenarios. Hence, it is more important to increase SE at peak traffic, but on the other hand, outside the peak hours, it becomes more important to prioritize EE. As a result, it is paramount to describe the RE problem as a multi-objective optimization problem.

In [12] the EE-SE trade-off is discussed by formulating a bi-objective optimization problem in optical code division multiple access (OCDMA) networks, while solutions are compared by deploying the weighted sum (WS) and  $\epsilon$ -Constraint ( $\epsilon$ -C) scalarization procedures combined with nonlinear programming as augmented lagrangian method (ALM) and sequential quadratic programming (SQP) method in order to solve the generated sub-problems. In [13], the authors investigate the resource allocation problem for achieving superior SE-EE trade-offs under minimum user-rate requirements in hybrid multi-carrier non-orthogonal multiple access (MC-NOMA) systems; such systems incorporate both NOMA and orthogonal multiple access (OMA) modes into a unified framework analysis. First, the SE-EE trade-off is formulated as a MOO problem with minimum rate requirement constraints. Considering the non-convexity nature of the MOO problem, it is converted into an SOO problem by utilizing the weighted Tchebycheff method. Lagrangian dual decomposition and sequential convex programming are applied to solve the equivalent SOO problem.

**Contribution:** The contribution of this work is twofold. First, the SE-EE trade-off in NOMA-MIMO systems operating under maximum equal-rate fairness and a different number of users per cluster has been formulated and comprehensively developed as a MOO problem. Second, for solving such knotty optimization problem, we have selected and compared distinct approaches by using primarily  $\epsilon$ -C scalarization method combined with two NLP techniques, as well as a greedy local search approach. The goal is to investigate which NLP technique, SQP, and DK, combined with  $\epsilon$ -C scalarization, attains the best EE-SE performance. Besides, the analyses and comparisons are extended to include a hybrid optimization method that combines the  $\epsilon$ -C and the HC greedy heuristic methods, as well as a comparison with the WS scalarization method, is performed.

The rest of this paper is organized as follows. In Section 2, the MIMO-NOMA system model is described. Section 3, formulates the MIMO-NOMA EE-SE trade-off optimization problem from the perspective of RE with equal user rate and a specific number of users per cluster. Section 4, formulates the solutions for the EE  $\times$  SE trade-off problem, describing the main feature of each approach adopted based on the  $\epsilon$ -Constraint approach. Comprehensive numerical results are presented in Section 5, while Section 6 concludes the paper.

## 2. MIMO-NOMA system

Consider a DL transmission in a multiple-input and multiple-output MIMO cell system combined with a non-orthogonal multiple-access (NOMA) system in which the base station (BS) is equipped with  $A$  antennas, as depicted in Fig. 1.

In order to have a NOMA scenario, the cellular system is divided into  $A$  clusters randomly grouped. There are  $K \geq 2$  users<sup>1</sup> for each cluster, where users have  $U$  antennas each one. Herein, we assume  $U \geq A$  to make the system feasible. As a trend, small cells have becoming ultra-densely deployed in 5G networks, in which low-cost, low-power small cell BSs have being more frequently used; hence, it is very likely that such BSs possess the same number of antennas as users' devices ( $A = U$ ), or even less, mainly due to the rapidly growing capabilities of smartphones and tablets<sup>2</sup> [2]. It worth noting that the adopted MIMO-NOMA scheme can also be applicable to the IoT, such as smart homes, in which the capability of a home BS is similar to that of laptops and other digital devices.

<sup>1</sup> All clusters have the same number of users and all users belong to the same cell (single-system); hence, without loss of generality, the total number of users in the cell is selected to be multiple of the number of users per cluster.

<sup>2</sup> In the case that the BS has more antennas than the users, different approaches need to be deployed to implement the MIMO-NOMA scheme explored herein. As pointed out in [14], one possible approach consists in allocating different beamforming vector to different users individually, and the precoding matrices at the BS can be optimized by taking user fairness into consideration.

The signal transmitted by the BS to the user can be described by [15]

$$\mathbf{x} = \mathbf{W}\hat{\mathbf{s}} \quad (1)$$

where the data vector  $\hat{\mathbf{s}} \in \mathcal{C}^{A \times 1}$  can be written as:

$$\hat{\mathbf{s}} = \begin{bmatrix} \sqrt{P\beta_{1,1}}s_{1,1} + \dots + \sqrt{P\beta_{1,k}}s_{1,k} + \dots + \sqrt{P\beta_{1,K}}s_{1,K} \\ \vdots \\ \sqrt{P\beta_{a,1}}s_{a,1} + \dots + \sqrt{P\beta_{a,k}}s_{a,k} + \dots + \sqrt{P\beta_{a,K}}s_{a,K} \\ \vdots \\ \sqrt{P\beta_{A,1}}s_{A,1} + \dots + \sqrt{P\beta_{A,k}}s_{A,k} + \dots + \sqrt{P\beta_{A,K}}s_{A,K} \end{bmatrix} = \begin{bmatrix} \hat{s}_1 \\ \vdots \\ \hat{s}_a \\ \vdots \\ \hat{s}_A \end{bmatrix} \quad (2)$$

with  $s_{a,k}$  and  $\beta_{a,k}$  representing the signal and the power proportion coefficients for the  $a$ th cluster and the  $k$ th user, respectively,  $P$  indicates the total power transmitted by the BS and  $\mathbf{W}$  the precoding matrix given by the diagonal matrix  $\mathbf{W} = \mathbf{I}_A$  an identity matrix and  $\mathbf{w}_a$  is the  $a$ th column of matrix  $\mathbf{W}$ . It allows to properly identify the cluster to be considered at the transmitter, as well as describe such cluster signal at the receiver side. This choice prevents the transmitter from sending channel state information (CSI), thereby reducing system overhead [2]. Besides, the *proportion power coefficients* must satisfy

$$\sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} = 1. \quad (3)$$

Hence, to attain data-rate fairness across all NOMA users, those users with lower channel gain in BS will receive larger  $\beta_{a,k}$ . Precisely, such values will be obtained from optimization processes.

Consequently, in the DL, the signal received by a certain user  $k$  belonging to the  $a$ th cluster is given by

$$\mathbf{y}_{a,k} = \mathbf{H}_{a,k}\mathbf{W}\hat{\mathbf{s}} + \mathbf{z}_{a,k}, \quad \text{where } \mathbf{H}_{a,k} = \frac{\mathbf{Q}_{a,k}}{\sqrt{1 + d_{a,k}^\alpha}} \quad (4)$$

and  $\mathbf{Q}_{a,k} \in \mathcal{C}^{U \times A}$  denoting a Rayleigh fading channel matrix, and  $\alpha$  is the path loss coefficient;  $d_{a,k}$  represents the distance from the  $k$ th user belongs to the  $a$ th cluster to the BS, and  $\mathbf{z}_{a,k}$  is an additive Gaussian noise (AWGN) vector,  $\mathcal{CN}(0, \sigma^2 \mathbf{I})$ . The received signal must be multiplied by a detection vector  $\mathbf{v}_{a,k}$  which overrides the interfering signal [16]. This way, one can identify the interfering cluster term:

$$\mathbf{v}_{a,k}^H \mathbf{y}_{a,k} = \mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_a \sum_{k=1}^K \sqrt{P\beta_{a,k}}s_{a,k} + \underbrace{\sum_{i=1, i \neq a}^A \mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_i \hat{s}_i}_{\text{interfering clusters}} + \mathbf{v}_{a,k}^H \mathbf{z}_{a,k} \quad (5)$$

where  $\hat{s}_i$  represents the  $i$ th element of  $\hat{\mathbf{s}}$  in Eq. (2), and  $\mathbf{w}_\ell$  is the  $\ell$ th column of matrix  $\mathbf{W}$ . Hence, the product  $\mathbf{H}_{a,k}\mathbf{w}_i$  results in the  $i$ -th column of  $\mathbf{H}_{a,k}$ . This column represents the antenna that serves the  $i$ th cluster.

To remove interference from other clusters, the detection matrix must obey the condition  $\mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_i = 0$  for any  $i \neq a$ . In order to ensure the feasibility of the  $\mathbf{v}_{a,k}$ ,  $U \geq A$  is assumed. This constraint is reduced to

$$\mathbf{v}_{a,k}^H \bar{\mathbf{H}}_{a,k} = 0, \quad (6)$$

where  $\bar{\mathbf{H}}_{a,k} = [\mathbf{h}_{1,ak} \dots \mathbf{h}_{a-1,ak} \mathbf{h}_{a+1,ak} \dots \mathbf{h}_{A,ak}]$  and  $\mathbf{h}_{j,ak}$  is the  $j$ th column of  $\mathbf{H}_{a,k}$ . We have [2]

$$\mathbf{v}_{a,k} = \mathbf{U}_{a,k} \mathbf{g}_{a,k}, \quad (7)$$

where  $\mathbf{U}_{a,k}$  is the matrix of the left singular vectors of  $\bar{\mathbf{H}}_{a,k}$  after decomposing on non-zero singular values and the maximal ratio combining vector expressed as

$$\mathbf{g}_{a,k} = \begin{bmatrix} \mathbf{U}_{a,k}^H \mathbf{h}_{a,ak} \\ \mathbf{U}_{a,k}^H \mathbf{h}_{a,ak} \end{bmatrix}. \quad (8)$$

After removing the interference from the other clusters, the Eq. (5) can be rewritten as

$$\mathbf{v}_{a,k}^H \mathbf{y}_{a,k} = \mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_a \sum_{k=1}^K \sqrt{P\beta_{a,k}} \mathbf{s}_{a,k} + \mathbf{v}_{a,k}^H \mathbf{z}_{a,k} \quad (9)$$

Without loss of generality, the effective channel gains are ordered as [2]

$$\left| \mathbf{v}_{a,1}^H \mathbf{H}_{a,1} \mathbf{w}_a \right|^2 \geq \dots \geq \left| \mathbf{v}_{a,K}^H \mathbf{H}_{a,K} \mathbf{w}_a \right|^2 \quad (10)$$

and the power proportion coefficients are ordered as follows

$$\beta_{1,1} \leq \dots \leq \beta_{1,K}. \quad (11)$$

The BS still needs to be informed about the scalar effective channel gain in order to implement the NOMA as shown in (10).

At the reception, each user uses SIC to remove interference from other users with worse channel gain, in a generic way, interfering users  $(a, k+1), \dots, (a, K)$  are removed by user  $(a, k)$ . It is considered that all the interference has been removed, that is, a perfect SIC is used. In this way, the signal-to-interference-plus-noise ratio (SINR) for the  $k$ th user in the  $a$ th cluster is given by [2]

$$\gamma_{a,k} = \frac{P \left| \mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_a \right|^2 \beta_{a,k}}{P \sum_{j=1}^{k-1} \beta_{a,j} \left| \mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_a \right|^2 + \left| \mathbf{v}_{a,k} \right|^2 \sigma^2} \quad (12)$$

for  $1 \leq a \leq A$  and  $2 \leq k \leq K$ . For the first user results in the following SINR:

$$\gamma_{a,1} = \frac{P \left| \mathbf{v}_{a,1}^H \mathbf{H}_{a,1} \mathbf{w}_a \right|^2 \beta_{a,1}}{\left| \mathbf{v}_{a,1} \right|^2 \sigma^2} \quad (13)$$

while the normalized data rate<sup>3</sup> reached by a user in a given cluster results:

$$R_{a,k} = \log_2(1 + \gamma_{a,k}) \quad \left[ \frac{\text{bits}}{\text{s}\cdot\text{Hz}} \right] \quad (14)$$

The energy efficiency is given by the ratio between the sum-rate achievable by the users to the total power consumed in the system:

$$\text{EE} = \frac{R_\Sigma}{P_{rf} + P_c} = \frac{K \cdot A \cdot R}{P_{rf} + A \cdot (K \cdot P_k + P_{bs})}, \quad \left[ \frac{\text{bits}}{\text{J}\cdot\text{Hz}} \right] \quad (15)$$

where the *sum of rates*<sup>4</sup> transmitted by all the users under maximum fairness, i.e. when all users attain the same capacity ( $R_{ak} = R$ ) is given by:

$$R_\Sigma = \sum_{a=1}^A \sum_{k=1}^K R_{a,k} = K \cdot A \cdot R. \quad (16)$$

In Eq. (15), we have assumed  $U = A$  antennas; moreover,  $P_c$  is the circuitry power consumption required to signal processing, cooling etc, at the transmitters and receivers. For simplicity, only the two more relevant terms  $P_k$  and  $P_{bs}$  were explicitly considered, where the power required to operate the circuitry of each user antenna is  $P_k$ , and to operate the circuit components of each BS antenna is  $P_{bs}$ . Besides, the *total RF power consumption* is given by:

$$P_{rf} = \varrho \cdot \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} P \quad \left[ \text{W} \right] \quad (17)$$

where  $\varrho \geq 1$  is the RF power amplifier inefficiency, assumed equal for all of them feeding each Tx antenna. Finally,  $P_c$  is the circuitry power

consumption required to signal processing at the transmitter side, assumed to be a fixed value per user.

### 3. EE-SE optimization in downlink MIMO-NOMA system

In this section, the optimization of the MIMO-NOMA system is done in two ways. First, to use distributed equal-rate optimal power allocation (ERPA) strategy to obtain a minimum equal rate for all users. Second, to compute the EE optimization to find the minimum equal data rate  $R$  that gives maximum EE. The concept of fairness is presented in the next subsection to better understand the process of obtaining its maximum value, followed by the problem of power optimization and EE optimization.

#### 3.1. Fairness

Fairness can be defined by the **Jain's fairness index** [17,18]:

$$\mathcal{F}_{\text{Jain}} = \frac{\left( \sum_{a=1}^A \sum_{k=1}^K R_{a,k} \right)^2}{A \cdot K \sum_{a=1}^A \sum_{k=1}^K R_{a,k}^2}. \quad (18)$$

In the optimization design of this work, the equal data rate has been adopted for all users; hence, from eq. (18), the full fairness results  $\mathcal{F}_{\text{Jain}} = 1$ , being its maximum value.

#### 3.2. Equal-rate Optimal Power Allocation (ERPA) in MIMO-NOMA systems

The formulation for obtaining power proportions to be considered for each user and cluster aiming at guaranteeing a minimum equal-rate for all MIMO-NOMA users with the distributed *equal-rate optimal power allocation* (ERPA) strategy is given by:

$$\text{minimize}_{P, \beta_{a,k}} \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} P$$

$$\text{s.t. c.1} \quad \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} = 1$$

$$\text{c.2} \quad \beta_{a,k} = (2^R - 1) \frac{\sigma^2 \left| \mathbf{v}_{a,k} \right|^2}{P \left| \mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_a \right|^2} + (2^R - 1) \sum_{i=1}^{k-1} \beta_{a,i}, \quad \text{where} \quad \begin{cases} 2 \leq k \leq K \\ 1 \leq a \leq A \end{cases}$$

$$\text{c.3} \quad \beta_{a,1} = \frac{(2^R - 1) \sigma^2 \left| \mathbf{v}_{a,1} \right|^2}{P \left| \mathbf{v}_{a,1}^H \mathbf{H}_{a,1} \mathbf{w}_a \right|^2} \text{ to } 1 \leq a \leq A$$

(19)

where the set of constraints is convex, and  $P, \beta_{a,k} \in \mathfrak{R}^+$ . Next, we show that the constraints c.1-c.3. can be incorporated into a convex cost function. The c.1 can be developed by replacing the restrictions c.2 and c.3 into the constraint c.1 subject to the constraints  $P, \beta_{a,k} \in \mathfrak{R}^+$ , obtaining the following cost function:

$$f = \sum_{a=1}^A \sum_{k=2}^K \left[ (2^R - 1) \sum_{i=1}^{k-1} \beta_{a,i} + (2^R - 1) \frac{\sigma^2 \left| \mathbf{v}_{a,k} \right|^2}{P \left| \mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_a \right|^2} \right] + \sum_{a=1}^A (2^R - 1) \frac{\sigma^2 \left| \mathbf{v}_{a,1} \right|^2}{P \left| \mathbf{v}_{a,1}^H \mathbf{H}_{a,1} \mathbf{w}_a \right|^2} - 1 \quad (20)$$

Hence, to be convex, the Hessian of Eq. (20) must be positive semi-

<sup>3</sup> Or spectral efficiency for the  $k$ th user of the  $a$ th NOMA cluster.

<sup>4</sup> Assuming normalized bandwidth.

definite [19], that is,  $\nabla^2 f \succ 0$ . Therefore all the eigenvalues of the Hessian must be greater than or equal to zero. After obtaining the Hessian and applying the definition of eigenvalues in Eq. (20), the following characteristic polynomial equally can be obtained:

$$-\lambda^M + \frac{2(2^R - 1)\sigma^2}{P^3} \sum_{a=1}^A \sum_{k=2}^K k = 2^k \frac{|\mathbf{v}_{a,k}|^2}{|\mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_a|^2} \lambda^{M-1} + \frac{2(2^R - 1)\sigma^2}{P^3} \sum_{a=1}^A \frac{|\mathbf{v}_{a,1}|^2}{|\mathbf{v}_{a,1}^H \mathbf{H}_{a,1} \mathbf{w}_a|^2} \lambda^{M-1} = 0 \quad (21)$$

Taking  $\lambda^{M-1}$  in evidence, there are  $M-1$  eigenvalues equal to zero. and therefore the value obtained for the last one is:

$$\lambda = \frac{2(2^R - 1)\sigma^2}{P^3} \left( \sum_{a=1}^A \sum_{k=2}^K \frac{|\mathbf{v}_{a,k}|^2}{|\mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_a|^2} + \sum_{a=1}^A \frac{|\mathbf{v}_{a,1}|^2}{|\mathbf{v}_{a,1}^H \mathbf{H}_{a,1} \mathbf{w}_a|^2} \right) \quad (22)$$

Since the variable  $P$  is positive as well as the constants  $R, \sigma^2, \frac{|\mathbf{v}_{a,k}|^2}{|\mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_a|^2}$  and  $\frac{|\mathbf{v}_{a,1}|^2}{|\mathbf{v}_{a,1}^H \mathbf{H}_{a,1} \mathbf{w}_a|^2}$ , the eigenvalues are positive or equal to zero; hence, eq. (20) is semi-definite positive.

### 3.3. Energy efficiency optimization problem

The energy efficiency optimization variables considered are the data rate, power, and the power proportion coefficients along with the users in which the objective function is given by Eq. (15). Such equation is quasi-concave [20]. We use the Dinkelbach (DK) method to turn the quasi-concave objective function into a concave one [21], as expressed by the problem (23). Hence, the EE optimization design in the MIMO-NOMA system to attain the optimal power allocation  $P$ , under equal data-rate  $R$  (maximum rate fairness) and power proportion  $\beta_{a,k}$  values for all users and different clusters can be defined by an equivalent problem as follows:

$$\begin{aligned} & \underset{P, \beta_{a,k}, R}{\text{maximize}} \quad K \cdot A \cdot R - q^n \left( q \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} P + K \cdot A \cdot P_K + A \cdot P_{BS} \right) \\ & \text{s.t. } c.1 \quad \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} = 1 \\ & \quad c.2 \quad \beta_{a,k} = (2^R - 1) \frac{\sigma^2 |\mathbf{v}_{a,k}|^2}{P |\mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_a|^2} + (2^R - 1) \sum_{i=1}^{k-1} \beta_{a,i}, \quad \text{where } \begin{cases} 2 \leq k \leq K \\ 1 \leq a \leq A \end{cases} \\ & \quad c.3 \quad \beta_{a,1} = \frac{(2^R - 1)\sigma^2 |\mathbf{v}_{a,1}|^2}{P |\mathbf{v}_{a,1}^H \mathbf{H}_{a,1} \mathbf{w}_a|^2} \quad \text{to } 1 \leq a \leq A \\ & \quad c.4 \quad \beta_{a,k} > 0; \quad P > 0; \quad R > 0 \end{aligned} \quad (23)$$

where  $q^n$  is a EE fixed value. A pseudo-code for the iterative MIMO-NOMA energy efficiency maximization based on Dinkelbach method described in (23) is depicted in Algorithm 1, where EE value at the  $n$ th iteration is determined by the previous power and rate allocation as:

$$q^{(n)} = \frac{U_R(P^{(n-1)}, R^{(n-1)}, \beta_{a,k}^{(n-1)})}{U_T(P^{(n-1)}, R^{(n-1)}, \beta_{a,k}^{(n-1)})} = \frac{K \cdot A \cdot R^{(n-1)}}{q \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k}^{(n-1)} P^{(n-1)} + K \cdot A \cdot P_K + A \cdot P_{BS}}$$

### Algorithm 1. EE-MIMO-NOMA – Iterative power-rate allocation based on Dinkelbach method

**Initialization:**  $n = 0$  and  $q^{(n)} = 0$   
**while**  $\varepsilon > 0$  **do**  
    For a given  $q^{(n)}$ , solve (23) to obtain  $P^{(n)}$  and  $R^{(n)}$  and  $\beta_{a,k}^{(n)}$   
**if**  $U_R(P^{(n)}, R^{(n)}, \beta_{a,k}^{(n)}) - q^{(n)} U_T(P^{(n)}, R^{(n)}, \beta_{a,k}^{(n)}) \leq \varepsilon$  **then**  
     $P^* = P^{(n)}, R^* = R^{(n)}, q^* = q^{(n)}, \beta_{a,k}^* = \beta_{a,k}^{(n)}$   
**else**  
    Set  $n = n + 1$  and  $q^{(n)} = \frac{U_R(P^{(n-1)}, R^{(n-1)}, \beta_{a,k}^{(n-1)})}{U_T(P^{(n-1)}, R^{(n-1)}, \beta_{a,k}^{(n-1)})}$   
**end if**  
**end while**

By analyzing the convexity of constraints with the variables total power, the proportion of total power among users, and data rate, it was found to be non-convex, since the Hessian eigenvalues with these variables result in a complex value. However, by setting an equal data rate  $R$  for all MIMO-NOMA users aiming to attain maximum fairness while consuming the same total power  $P$  and the proportion of total power  $\beta_{a,k}$  among users as optimization variables, the convexity of constraints set is holding, similarly as posed in eq. (20). Hence, in the same way as pointed out earlier, the set of constraints under these modified conditions results convex. As a consequence, the EE-MIMO-NOMA maximization under user-rate fairness problem using the Dinkelbach method can be re-stated as follows.

$$\begin{aligned} & \underset{P, \beta_{a,k}}{\text{maximize}} \quad K \cdot A \cdot R^* - q^n \left( q \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} P + K \cdot A \cdot P_K + A \cdot P_{BS} \right) \\ & \text{s.t. } c.1, \quad c.2, \quad c.3 \quad \text{and} \quad c.4 \end{aligned} \quad (24)$$

where in this reduced problem, the fixed EE at the DK  $n$ th iteration is defined as:

$$q^{(n)} = \frac{U_R(P^{(n-1)}, \beta_{a,k}^{(n-1)})}{U_T(P^{(n-1)}, \beta_{a,k}^{(n-1)})} = \frac{K \cdot A \cdot R^*}{q \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k}^{(n-1)} P^{(n-1)} + K \cdot A \cdot P_K + A \cdot P_{BS}}$$

while the equal user-rate  $R$  is a pre-defined value. Notice that in the problem (24), again, the Dinkelbach method was deployed to turn the quasi-concave objective function into a concave one. Algorithm 2 describes the EE maximization adopting fixed data rate. Algorithm 2 describes the EE maximization adopting fixed data rate  $R^*$ .

### Algorithm 2. EE-MIMO-NOMA – Iterative power allocation based on Dinkelbach, equal data-rate

**Initialization:**  $n = 0, q^{(n)} = 0$   
**while**  $\varepsilon > 0$  **do**  
    For a given  $q^{(n)}$ , solve (24) to obtain  $P^{(n)}$  and  $\beta_{a,k}^{(n)}$   
**if**  $U_R(P^{(n)}, \beta_{a,k}^{(n)}) - q^{(n)} U_T(P^{(n)}, \beta_{a,k}^{(n)}) \leq \varepsilon$  **then**  
     $P^* = P^{(n)}, q^* = q^{(n)}, \beta_{a,k}^* = \beta_{a,k}^{(n)}$   
**else**  
    Set  $n = n + 1$  and  $q^{(n)} = \frac{U_R(P^{(n-1)}, \beta_{a,k}^{(n-1)})}{U_T(P^{(n-1)}, \beta_{a,k}^{(n-1)})}$   
**end if**  
**end while**

## 4. EE-SE Bi-objective optimization approach

In this section, MOO (BOO) methods and different strategies to solve the MIMO-NOMA resource efficiency, i.e., the EE-SE trade-off problem

**Table 1**  
Adopted Parameters for the Simulation Scenarios.

Parameter	Value
#Users	$M \in \{4; 40\}$
Path-loss exponent	$\alpha = [2.0; 3.0; 3.5]$
Power of white noise	$\sigma_n^2 = 0.1\mu\text{W}$
Power Budget	$P = 100\text{ W}$
# clusters	$A \in \{2; 12\}$
users/clusters	$K \in \{2, 3, 4\}$
# antennas per user	$U \geq A$ , with $U \in \{2; 12\}$ .
Target equal normalized user's rate	$R^* \in [0.1; 0.5 : 0.5 : 4.0]$ [bits/s/Hz]
Circuits Power	$P_{BS} = 1\text{W/antenna (BS)}$ $P_K = 100\text{mW/antenna per user}$
RF power amplifier inefficiency	$\rho = 1.4$
<b>Optimization Methods: Parameters</b>	
$\epsilon$ -C limits	$SE(\bar{P}, \bar{\beta}_{a,k}, \bar{R}) \leq \epsilon \leq \max SE$
Hill Climbing Greedy heuristic	$\rho = 10^8$ ; Max-Iteration=10 <sup>4</sup>
<b>Scenarios: User distribution inside the cell <math>\sim \mathcal{U}(r_0, r_D)</math></b>	
Scenario 1	$r_0 = 10$ [m] and $r_D = 50$ [m]
Scenario 2	$r_0 = 10$ [m] and $r_D = 100$ [m]
Scenario 3	$r_0 = 10$ [m] and $r_D = 200$ [m]
# MCS realizations	10 <sup>4</sup>
<b>Power Allocation Strategies</b>	
Equal-rate power allocation (ERPA)	Section 3.2

under maximum equal user-rate fairness and variable user per cluster is developed using primarily the  $\epsilon$ -Constraint ( $\epsilon$ -C) scalarization method combined with nonlinear programming (NLP) optimization techniques, namely sequential quadratic programming (SQP), as well as Dinkelback (DK) method.

#### 4.1. Resource allocation for MIMO-NOMA problem

Resource allocation problems, mainly considering EE and SE optimization simultaneously, result in bi-objective optimization (BOO) goals instead of just one. These objectives are often in conflict with each other, and there is rarely a solution that is optimal for all objectives simultaneously. The characteristics of the *resource efficiency* (RE), i.e., the EE-SE trade-off multi-objective optimization (MOO) have been explored recently in [12] considering the optical communication context and in the wireless communication [6]. Hence, the RE metric can be expressed as a weighted sum of the EE and the SE as:

$$RE = \xi \cdot EE + SE \quad \left[ \frac{\text{bits}}{\text{s} \cdot \text{Hz}} \right], \quad (25)$$

where  $\xi$  is the weighting factor controlling the EE weight regarding the SE in a BOO design. Moreover, the  $\xi$  factor in [J/s] acts as a physically consistent factor in Eq. (25), while balancing the importance among both metrics.

Let  $\mathbb{R}^n$  and  $\mathbb{R}^m$  be Euclidean vector spaces referred to the decision space and the objective space. Let  $\mathcal{X} \subset \mathbb{R}^n$  be a non-empty feasible set and let  $f$  be a vector-valued objective function  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  composed of  $m$  real-valued objective functions,  $\mathbf{f} = (f_1, \dots, f_m)$ , where  $f_k: \mathbb{R}^n \rightarrow \mathbb{R}$  for  $k = 1, \dots, m$ . A MOO problem can be formulated as [22,12]:

$$\underset{\mathbf{x} \in \mathcal{X}}{\text{maximize}} \quad \{f_1(\mathbf{x}), \dots, f_m(\mathbf{x})\}; \quad \mathbf{x} \in \mathcal{X} \quad (26)$$

We assume that the set  $\mathcal{X}$  is given implicitly in the form of constraints, where:

- $\mathbf{x} = \{x_1, x_2, \dots, x_n\} \in \mathbb{R}^n$ : vector of decision variables;
- $z_i = f_i(\mathbf{x})$ :  $k$ -th objective function to be minimized,  $i = 1, \dots, m$ ;
- $\emptyset \neq \mathcal{X} \subseteq \mathbb{R}^n$ : decision or feasible space.

**Definition 1. (Domination)** One solution  $\mathbf{x}^{(1)}$  dominates another solution  $\mathbf{x}^{(2)}$ , where  $\mathbf{x}^{(1)}, \mathbf{x}^{(2)} \in \mathcal{X}$ , if the following conditions are met:

1. The solution  $\mathbf{x}^{(1)}$  no worse than the solution  $\mathbf{x}^{(2)}$  in all objectives, i.e.,  $f_i(\mathbf{x}^{(1)}) \geq f_i(\mathbf{x}^{(2)})$  for all  $i = 1, \dots, m$ .
2. The solution  $\mathbf{x}^{(1)}$  is strictly better than the  $\mathbf{x}^{(2)}$  solution on at least one component  $i \in \{1, \dots, m\}$ , i.e.,  $f_i(\mathbf{x}^{(1)}) > f_i(\mathbf{x}^{(2)})$

The dominance of the  $\mathbf{x}^{(1)}$  solution over  $\mathbf{x}^{(2)}$  solution is denoted by:  $\mathbf{x}^{(1)} \preceq \mathbf{x}^{(2)}$ .

**Definition 2. (Efficient solution)** A  $\mathbf{x}^* \in \mathcal{X}$  solution is called efficient or optimal in the Pareto sense, if there is no other solution  $\mathbf{x} \in \mathcal{X}$ , so that  $\mathbf{x} \preceq \mathbf{x}^*$ .

In this work, we are interested in the resource efficiency problem in terms of EE-SE trade-off optimization in NOMA-MIMO networks. The EE-SE trade-off can be formulated aiming at maximizing the EE and SE simultaneously. One can formulate it as a BOO problem as follows:

$$\begin{aligned} & \underset{P, \beta_{a,k}, R}{\text{maximize}} \quad SE \left[ \frac{\text{bit}}{\text{s} \cdot \text{Hz}} \right] \quad \text{and} \quad \underset{(P, \beta_{a,k}, R)}{\text{maximize}} EE \left[ \frac{\text{bit}}{\text{J} \cdot \text{Hz}} \right] \\ & \text{s.t. c.1} \quad \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} = 1 \\ & \text{c.2} \quad \beta_{a,k} = (2^R - 1) \frac{\sigma^2 |\mathbf{v}_{a,k}|^2}{P |\mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_a|^2} + (2^R - 1) \sum_{i=1}^{k-1} \beta_{a,i}, \quad \text{where} \quad \begin{cases} 2 \leq k \leq K \\ 1 \leq a \leq A \end{cases} \\ & \text{c.3} \quad \beta_{a,1} = \frac{(2^R - 1) \sigma^2 |\mathbf{v}_{a,1}|^2}{P |\mathbf{v}_{a,1}^H \mathbf{H}_{a,1} \mathbf{w}_a|^2} \quad \text{to} \quad 1 \leq a \leq A. \\ & \text{c.4} \quad \beta_{a,k} > 0; P > 0; R > 0 \end{aligned} \quad (27)$$

The traditional approach to solving MOO is by scalarization methods, which involve formulating a single-objective problem related to MOO, by using a real value scalar function. Sometimes the viable set of the MOO is further restricted by new restriction functions related to the MOO objective functions, besides the new variables introduced. Moreover, in scalarization approaches, it is paramount to consider the necessary and sufficient conditions to attain the optimal solutions in terms of weakly, strictly, or adequately efficient MOO solutions.

#### 4.2. Solving EE-SE trade-off in MIMO-NOMA systems under user-rate maximum fairness and variable user per cluster

The central idea in the  $\epsilon$ -C multi-objective optimization method is that one of the objectives is chosen as the only objective to be optimized, while the other conflicting goals being treated as inequality constraints, allowing the deployment of the classic single-objective optimization approaches. In this way, the  $\epsilon$ -C method does not directly address an appropriate weight between the EE and SE metrics, as occurs in the WS scaling method, to determine solutions in the Pareto frontier [12]. Thus, applying  $\epsilon$ -C technique, the original RE MIMO-NOMA problem presented in (27) can be reformulated as follows:

$$\begin{aligned}
 & \underset{P, \beta_{a,k}, R}{\text{maximize}} && \sum_{a=1}^A \sum_{k=1}^K R_{a,k} Q + K \cdot A \cdot P_K + A \cdot P_{BS} \\
 & \text{s.t. } c.1 && \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} = 1 \\
 & c.2 && \beta_{a,k} = (2^R - 1) \frac{\sigma^2 |\mathbf{v}_{a,k}|^2}{P |\mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_a|^2} + (2^R - 1) \sum_{i=1}^{k-1} \beta_{a,i}, \quad \text{where } \begin{cases} 2 \leq k \leq K \\ 1 \leq a \leq A \end{cases} \\
 & c.3 && \beta_{a,1} = \frac{(2^R - 1) \sigma^2 |\mathbf{v}_{a,1}|^2}{P |\mathbf{v}_{a,1}^H \mathbf{H}_{a,1} \mathbf{w}_a|^2} \quad \text{to } 1 \leq a \leq A \\
 & c.4 && \beta_{a,k} > 0; P > 0; R > 0 \\
 & c.5 && \sum_{a=1}^A \sum_{k=1}^K R_{a,k} \geq \epsilon
 \end{aligned} \tag{28}$$

where  $\epsilon \in [\overline{\text{SE}}(P, \beta_{a,k}, R), \text{maxSE}]$ . The  $\epsilon$  values range from the lower limit of constraint (C. $\epsilon$ ) to the values that SE can take on the Pareto front. The minimum value  $\overline{\text{SE}}(P, \beta_{a,k}, R)$  is determined by obtaining  $\overline{(P, \beta_{a,k}, R)}$  solution of the EE problem. This choice for the minimum value is made because any other vector  $(P, \beta_{a,k}, R)$  whose value of SE is less than that will also have a value less than EE; therefore, this vector will be dominated by the solution vector  $\overline{(P, \beta_{a,k}, R)}$ , therefore, it would not be a Pareto solution. Thus, we consider  $\epsilon$  varying in the range  $\epsilon \in [\overline{\text{SE}}(P, \beta_{a,k}, R), \text{maxSE}]$  indicates the vector where the maximum EE occurs.

The Theorem 1 presents the necessary and sufficient condition for global solutions of (28) to be on the Pareto boundary of the EE-SE trade-off problem.

**Theorem 1.** A global solution to the problem defined in (28) is a solution on the Pareto frontier to the EE-SE trade-off problem if, and only if, the following equality holds in the solution vector.

$$\sum_{a=1}^A \sum_{k=1}^K R_{a,k} = \epsilon \tag{29}$$

*Proof:* The proof is similar to that presented in [12].

Eq. (29) in Theorem 1 indicates that the (C. $\epsilon$ ) inequality constraint in the problem (28) is satisfied as an equality in a solution belonging to Pareto front.

The  $\epsilon$ -C method is summarized in Algorithm 3, where the vector  $\mathbf{p}^k$  denotes the vector solution of problem (28) at the  $k$ -th iteration. As difficulty in applying the  $\epsilon$ -C method is that any Pareto solution found by the method must satisfy the (C. $\epsilon$ ) inequality constraint as equality. A drawback of the WS method consists in different weights may imply equal solutions as reported in [12]. Unlike the WS method, the  $\epsilon$ -C method does not directly consider RE in its objective, but by setting appropriate values to  $\epsilon$ , one is imposing the exact weight that  $\eta_{SE}$  will have on the RE metric; thus, solving the problem in (27) one also can get solutions for RE. Moreover, one motivation for using the  $\epsilon$ -C method is that different values for  $\epsilon$  imply different solutions. So, in the  $\epsilon$ -C method operating under a uniform distribution of  $\epsilon$  values, one can expect a uniform distribution of solutions in the Pareto frontier.

**Algorithm 3.**  $\epsilon$ -C-MIMO-NOMA –  $\epsilon$ -Constraint for the EE-SE Trade-off in MIMO-NOMA

- 1: **Choose:**  $\delta \in (0, 1)$ .
- 2: **Calculate:**  $\text{maxSE}$  and  $\mathbf{p}^0 = \overline{(P, \beta_{a,k}, R)}$ ;
- 3: **Define:**  $k = 1$  and  $\epsilon = \overline{\text{SE}}(P, \beta_{a,k}, R)$ ;
- 4: **while**  $\epsilon \leq \text{maxSE}$  **do**
- 5:    $\mathbf{p}^k \leftarrow$  Solve problem (28) by HC, DK or SQP method;
- 6:    $\epsilon \leftarrow \epsilon + \delta(\text{maxSE} - \overline{\text{SE}}(P, \beta_{a,k}, R))$ ;
- 7:   Save  $\mathbf{p}^k$ ;
- 8:    $k = k + 1$ ;
- 9: **end while**
- 10: **Output:**  $\mathbf{p}^0, \dots, \mathbf{p}^k$ ;

### 4.3. Deployed optimization methods

In the sequel, we present the adopted non-linear optimization methods for solving Step 5 of Algorithm 3. Besides, we discuss how the DK, HC, and SQP methods were implemented to solve such an intricate optimization problem.

#### 4.3.1. ( $\epsilon$ -C)-DK method

We apply the DK method as discussed in Section 3.3 and detailed in Algorithm 2, to solve the specific power allocation in the EE problem inside the  $\epsilon$ -C-MIMO-NOMA Algorithm 3. Notice that the  $c.\epsilon$  constraint coupled with the problem (15) as it appears in the problem (28) is a convex constraint. This fact allows the use of Algorithm 2 in the resolution of Step 5 of Algorithm 3.

#### 4.3.2. ( $\epsilon$ -C)-SQP method

The sequential quadratic programming (SQP) method determines an approximation for the solution of (28) through the limit of the solutions of a sequence of predefined quadratic problems. In the context of the EE-SE trade-off in MIMO-NOMA, to simplify the notation we consider in this subsection  $\mathbf{x} = (P, \beta_{a,k}, R)$ ,  $f$  represents the EE function,  $\mathbf{h}(\mathbf{x}) = 0$  represents restrictions  $c.1, c.2$  e  $c.3$ , and  $g(\mathbf{x}) \geq 0$  represents the constraint  $c.\epsilon$ . In summary, the SQP method at each iteration replaces EE with a quadratic approximation of the  $\mathcal{L}$  Lagrangian function, defined by:

Table 2

Figure of merit for the EE-SE MIMO-NOMA algorithms performance evaluation; path loss exponent  $\alpha = 3$ .

Method	$r_D$ [m]	SN [%]	Time [sec.]	NMSE	$R^*$ [bps/Hz]	FunC	S-PF
<b>1 cluster and 2 users</b>							
$\epsilon$ -C-DK	50	100	160.043556	9.0350e-07	2.8982	1.4011e+6	11
$\epsilon$ -C-SQP		100	1.2163	1.1234e-17	2.9740	8928	11
$\epsilon$ -C-HC		100	127.464792	0.0070	<b>3.4612</b>	4.0396e+6	11
$\epsilon$ -C-DK	100	100	109.003807	1.0804e-06	1.4479	6.7499e+5	11
$\epsilon$ -C-SQP		100	1.1529	2.9737e-17	1.8296	9046	11
$\epsilon$ -C-HC		100	132.431144	0.0132	<b>2.4335</b>	4.6132e+6	11
$\epsilon$ -C-DK	200	100	79.314762	4.7637e-06	0.5808	5.4086e+5	11
$\epsilon$ -C-SQP		100	1.1211	4.4042e-17	1.2096	9246	11
$\epsilon$ -C-HC		100	252.629958	0.0077	<b>2.0623</b>	6.5318e+6	11
<b>2 cluster and 4 users</b>							
$\epsilon$ -C-DK	50	100	367.431154	7.4973e-06	<b>3.7934</b>	2.9768e+6	11
$\epsilon$ -C-SQP		100	1.3553	1.3395e-17	3.7915	12174	11
$\epsilon$ -C-HC		100	630.490926	0.0101	3.5929	1.3880e+7	11
$\epsilon$ -C-DK	100	100	260.128037	1.4427e-06	2.0865	1.5796e+6	11
$\epsilon$ -C-SQP		100	1.3469	3.9909e-17	<b>2.4385</b>	12620	11
$\epsilon$ -C-HC		100	789.395572	0.0234	2.2856	2.0171e+7	11
$\epsilon$ -C-DK	200	100	146.104273	1.5602e-06	<b>1.6552</b>	1.2693e+6	11
$\epsilon$ -C-SQP		100	1.3934	7.2783e-17	1.4033	12898	11
$\epsilon$ -C-HC		100	1150.971378	0.0110	1.3273	2.4700e+7	11
<b>3 cluster and 6 users</b>							
$\epsilon$ -C-DK	50	100	545.3777	8.0740e-7	4.0584	1.9246e+6	11
$\epsilon$ -C-SQP		100	2.3179	5.4295e-17	4.3099	18058	11
$\epsilon$ -C-HC		100	1037.2966	0.0080	<b>4.7923</b> <sup>†</sup>	1.2149e+7	11
$\epsilon$ -C-DK	100	100	285.8444	6.738e-7	3.0794	1.4711e+6	11
$\epsilon$ -C-SQP		100	2.3476	1.1659e-16	3.3178	20002	11
$\epsilon$ -C-HC		100	1661.6796	0.0145	<b>3.8713</b>	1.9571e+7	11
$\epsilon$ -C-DK	200	100	264.7762	9.0229e-6	1.7936	1.4498e+6	11
$\epsilon$ -C-SQP		100	2.3907	9.7081e-16	1.9476	20080	11
$\epsilon$ -C-HC		100	1383.3902	0.0140	<b>2.5699</b>	1.7240e+7	11
$\epsilon$ -C-DK	300	100	287.2231	0.0011	1.2122	1.5383e+6	11
$\epsilon$ -C-SQP		100	2.2659	3.9103e-16	1.2734	19172	11
$\epsilon$ -C-HC		100	1652.4219	0.0145	<b>1.8702</b>	1.2149e+7	11
$\epsilon$ -C-DK	400	100	532.9553	0.0334	1.0171	2.7164e+6	11
$\epsilon$ -C-SQP		100	2.4897	6.7735e-16	0.9118	19274	11
$\epsilon$ -C-HC		100	1123.8642	0.0140	<b>1.5050</b>	1.2417e+7	11

<sup>†</sup>Highest target equal normalized user's rate.

$$\mathcal{L}(\mathbf{x}, \lambda, \mu) = f(\mathbf{x}) + \sum_i \lambda_i h_i(\mathbf{x}) - \mu g(\mathbf{x}). \quad (30)$$

Hence, non-linear constraints are replaced by linear approximations. Thus, at the  $k$ -th iteration, the SQP method solves the following quadratic programming problem:

$$\begin{aligned} & \underset{\mathbf{d} \in \mathbb{R}^k}{\text{maximize}} && f(\mathbf{x}^k) + \nabla f(\mathbf{x}^k)^T \mathbf{d} + \frac{1}{2} \mathbf{d}^T \nabla^2 \mathcal{L}(\mathbf{x}^k, \lambda^k, \mu^k) \mathbf{d} \\ & \text{s.t.} && -\nabla g(\mathbf{x}^k)^T \mathbf{d} - g(\mathbf{x}^k) \geq 0, \end{aligned} \quad (31)$$

$$\nabla h_i(\mathbf{x}^k)^T \mathbf{d} + h_i(\mathbf{x}^k) = 0, \quad \forall i.$$

Updating to the next solution,  $\mathbf{x}^{k+1}$  is obtained by designing  $\mathbf{x}^k + \bar{\alpha}_k \mathbf{d}^k$  in the set  $\Omega$  defined by constraint c.2,  $\bar{\alpha}_k \in (0, 1]$  is the size of the step determined to guarantee the convergence properties of the method,  $\mathbf{d}$  is the search direction. A more detailed description of the SQP method can

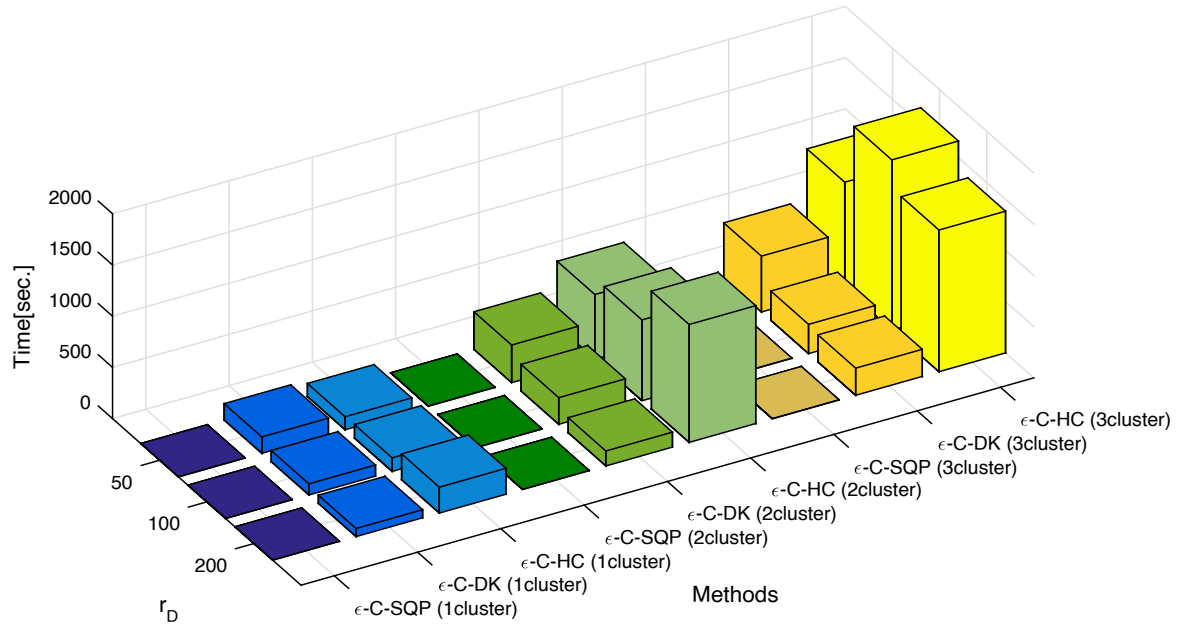
be found at [23]. Pseudocode for the SQP-based method solving the EE-SE trade-off in MIMO-NOMA is described in [Algorithm 4](#).

**Algorithm 4.** SQP-MIMO-NOMA – Sequential quadratic programming Trade-off in MIMO-NOMA

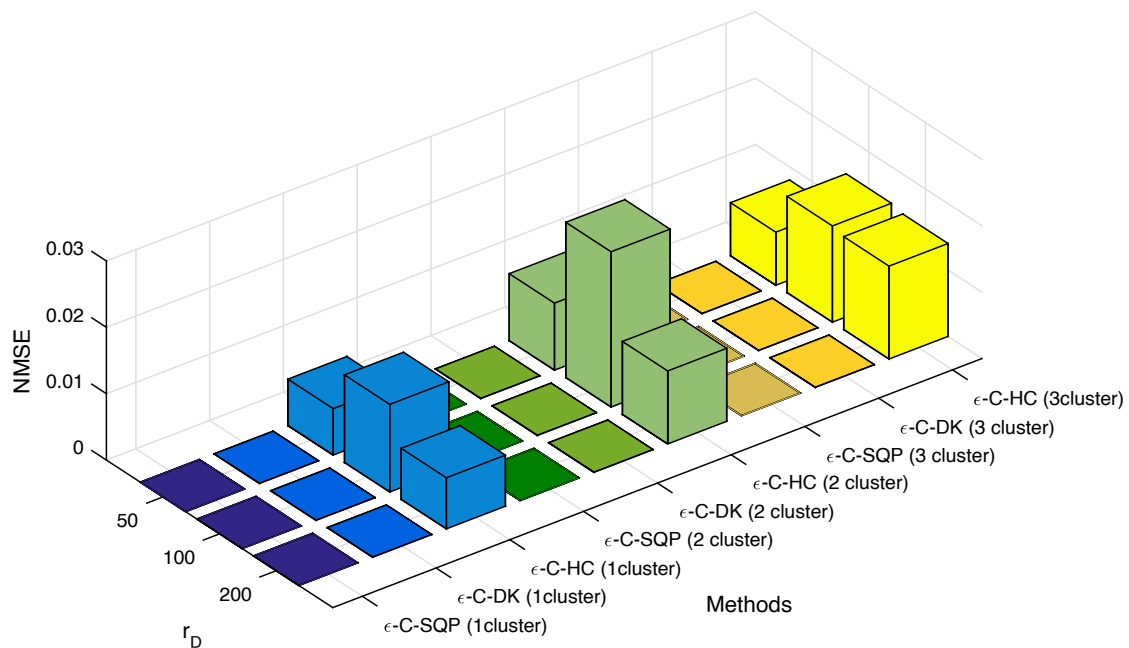
- 1: **Choose:** initial point  $(\mathbf{x}^0, \mu_0, \lambda_0)$  and tolerance  $\zeta > 0$ ;
- 2: **Define:**  $k = 0$ ; 3: **while**  $\|\mathbf{p}^k - \mathbf{p}^{k-1}\| \leq \zeta$  **do**
- 4: **Compute:**  $f(\mathbf{x}^k)$ ,  $\nabla f(\mathbf{x}^k)$ ,  $\nabla^2 \mathcal{L}(\mathbf{x}^k)$ ,  $\mathbf{h}(\mathbf{x}^k)$ ,  $\nabla \mathbf{h}(\mathbf{x}^k)$ ,  $g(\mathbf{x}^k)$  and  $\nabla g(\mathbf{x}^k)$ ;
- 5: **Solve:** (31) to get  $\mathbf{d}^k$  and  $\mu^{k+1}$ ;
- 6:  $\mathbf{x}^{k+1} \leftarrow \mathcal{P}(\mathbf{x}^k + \bar{\alpha}_k \mathbf{d}^k)$ ; where  $\mathcal{P}$  is the orthogonal projection operator in the set  $\Omega$ ;
- 7:  $\mathbf{x}^* = \mathbf{x}^k$ ; 8: **end while**
- 9: **Output:**  $\mathbf{x}^* = \mathbf{x}^k$ .

#### 4.4. ( $\epsilon$ -C)-HC heuristic Greedy method

The Hill Climbing (HC) algorithm is a greedy local search optimization method that starts with a solution called the current node and



(a) Time elapsed for the three BOO proposed methods.



(b) NMSE for the three BOO proposed methods.

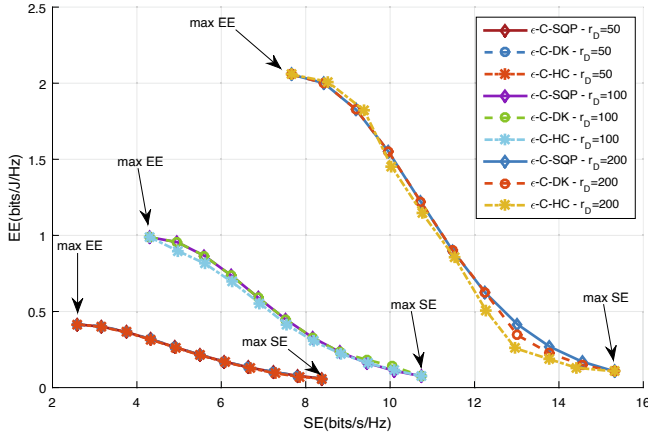
Fig. 2. Comparison of the elapsed time and NMSE attained by the three BOO proposed methods  $\epsilon$ -C-SQP,  $\epsilon$ -C-DK, and  $\epsilon$ -C-HC, under full convergence considering three different Scenarios ( $r_D$  values).

looks in nearby neighbours, where the distance is controlled by the step size, in order to find a better node; if there is no better node associated, one has attained a solution to the problem. Hence, if the algorithm detects a neighboring node better than the current node, it transforms the neighbor as the new approximation to the current solution and then iterates until reaches a much better location (based on distances, following a termination criterion) or it has achieved the maximum number of pre-established iterations [24–26]. A pseudo-code for the HC algorithm adapted to the EE-SE trade-off in MIMO-NOMA communication systems is described in Algorithm 5. The cost function  $F(\mathbf{x})$  deployed

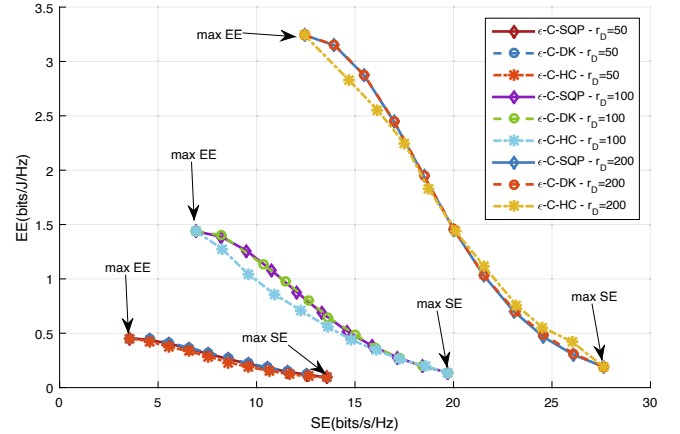
in the Algorithm 5 can be defined as:

$$F(\mathbf{x}) = EE(\mathbf{x}) + \rho \left\{ \sum [h_i(\mathbf{x})]^2 + \sum [\max(0, g_i(\mathbf{x}))]^2 \right\}. \quad (32)$$

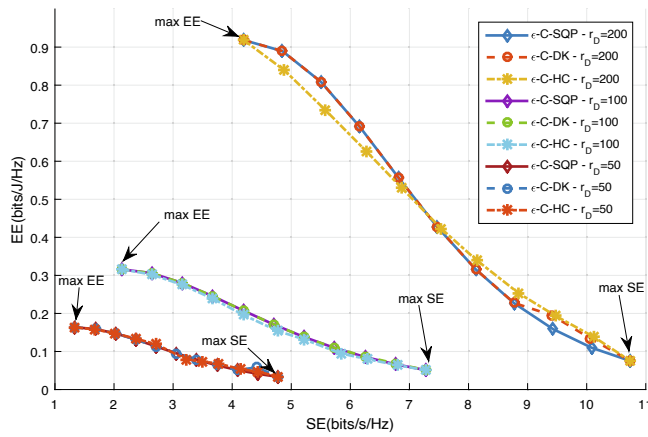
where  $\rho$  is the penalty parameter associated to the violation of the problem restrictions, while functions  $h$  and  $g$  in the EE-SE MIMO-NOMA problem were discussed in Section 4.3.2. The cost function (32) with constant penalty parameter requires a high but limited value to inhibit restrictions' violation that makes up the problem [27].



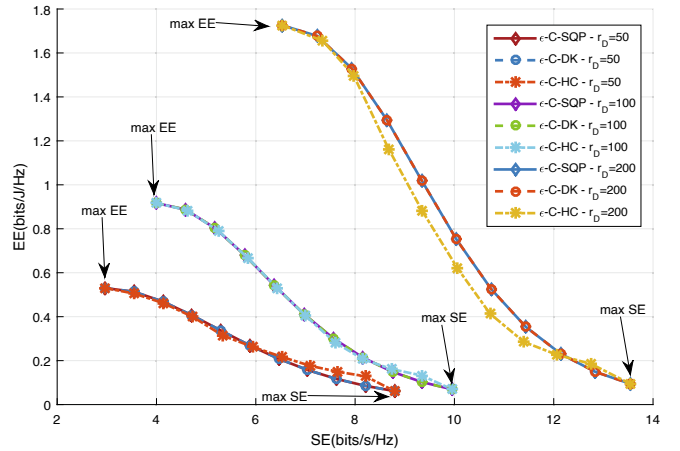
(a) Pareto front for 1 clusters and 2 users,  $\alpha = 3$ .



(b) Pareto front for 2 clusters and 4 users,  $\alpha = 3$ .



(c) Pareto front for 1 clusters and 2 users,  $\alpha = 3.5$ .



(d) Pareto front for 2 clusters and 4 users,  $\alpha = 3.5$ .

Fig. 3. Pareto front obtained by the three BOO methods  $\epsilon$ -C-SQP,  $\epsilon$ -C-DK, and  $\epsilon$ -C-HC for different Scenarios ( $r_D$  values); path-loss coefficients  $\alpha = 3$  and  $\alpha = 3.5$ .

**Algorithm 5.** HC-MIMO-NOMA – Hill Climbing EE-SE Trade-off in MIMO-NOMA

- 1: **Choose:**  $\mathbf{x} = (P^0, \beta_{a,k}^0, R^0)$ .
- 2: **Define:** Solution = 0; Iteration = 0; Max-Iteration > 0.
- 3: **while** Solution = 0 and Iteration  $\leq$  Max-Iteration **do**
- 4: Iteration = Iteration + 1;
- 5: Generate neighbours;
- 6: **if** any neighbor has a better value of  $F$  than  $x$  **then**
- 7:  $\mathbf{x}$  is updated as the neighbor with best cost for  $F$ ;
- 8: **else**
- 9: Solution = 1;
- 10: **end if**
- 11: **end while**
- 12: **Output:**  $\mathbf{x}$ ;

**5. Numerical results**

The parameter values used in the simulations are defined in Table 1. The tool used for optimization was MatLab® *fmincon* function. It was considered a range of values from the lowest to the highest in order to obtain the behavior of the system with a different number of users, path loss exponent values. In this way, one can trace a system's behavior considering different practical scenarios of interest. We have considered Scenario 1, 2, and 3 in which users are uniformly distributed inside the ring area of radius  $r_0$  and  $r_D$  [meters], Fig. 1.

**5.1. Implementation details**

Since the problem  $EE$  is fractional and the objective function of  $SE$  is concave (Section 4.2), in addition to the suitable results of the SQP method obtained in a similar problem [12], it can be expected that the solution obtained in Step 5 of Algorithm 3 be global. However, just a good performance of DK or SQP in the subproblems generated by the  $\epsilon$ -C method is not enough to guarantee the Pareto front solutions. Therefore, to complete the numerical analysis obtained with the  $\epsilon$ -C method, we have evaluated the normalized mean squared error (NMSE).

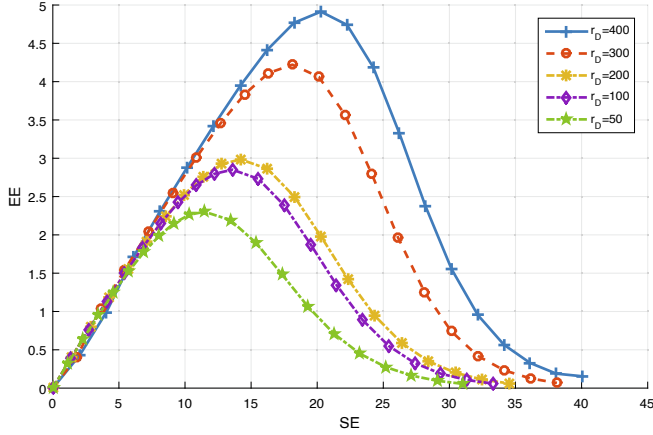
From Theorem 1, one can measure the quality of the attained solution using the Algorithm 3, in the  $k$ -th iteration, by expressing how far the current solution ( $\epsilon$ ) is from the final optimum solution  $SE(\mathbf{x}^*)$ , using the solution  $\mathbf{x}^*$  determined in Step 4 of the algorithm. This quality of the solution is quantified by the NMSE when equilibrium is attained. Numerically, the quality of the solution obtained by an EE optimization algorithm is ready defined by

$$nmse = \mathbb{E} \left[ \frac{\|se(\mathbf{x}^*) - \epsilon\|^2}{\|\epsilon\|^2} \right], \tag{33}$$

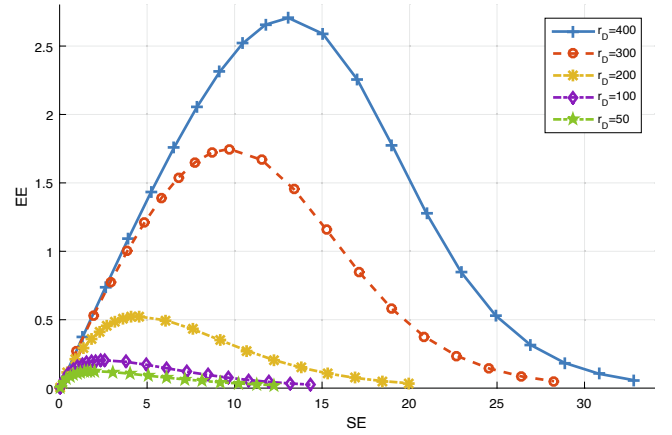
where  $\|\cdot\|^2$  denotes the Euclidean quadratic distance and  $\mathbb{E}[\cdot]$  is the statistical expectation operator.

**5.1.1. SQP vs DK in Step 5 of Algorithm 3**

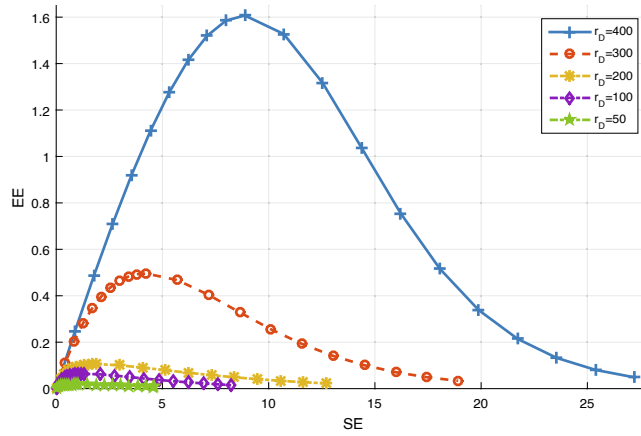
We have applied the SQP and DK methods to solve the optimization



(a) EE×SE for 2 clusters and 6 users,  $\alpha = 2$ .



(b) EE×SE for 2 clusters and 6 users,  $\alpha = 3$ .



(c) EE×SE for 2 clusters and 6 users,  $\alpha = 3.5$ .

Fig. 4. EE × SE obtained for the  $\epsilon$ -C-SQP method under different cell radius  $r_D$  values, considering path-loss coefficients  $\alpha = 2, 3$  and  $3.5$ .

problem (28) in Step 5 of Algorithm 3. The DK method as presented in Algorithm 1 uses a quadratic penalty method to solve the problem in (23). For implementing the SQP method, we have deployed Matlab's `fmincon` function. Besides, the gradient calculation was implemented by finite differences, while the quadratic subproblem was solved using a method of interior points [23] applied to convex quadratic subprograms.

**Stopping criterion:** For the DK and SQP methods, the following condition was considered as a stopping criterion.

$$\zeta = \|\mathbf{x}^k - \mathbf{x}^{k-1}\| < 10^{-6}, \quad \text{with } \mathbf{x}^k \text{ a feasible point.}$$

**DK × SQP sub-steps:** The choice of DK and SQP methods for comparison purposes is due to the simplicity of definition and the good theoretical properties of both methods, specifically for DK in fractional optimization problems where the method converges to the overall solution. Such methods require solving a simpler sequence of NLP problems than the original. Considering the sequential resolution of the problem (28), we apply DK and SQP, so that the first problem uses null multipliers (SQP). For the other sequential subproblems resolution, the starting point is obtained by solving the previous problem; moreover, in the SQP, the initial multipliers and penalty parameters are assumed as the multipliers and penalty parameters obtained with the solution of the previous subproblems. Finally, in the  $\epsilon$ -C method, as mentioned earlier, a challenge is to determine an appropriate range for the  $\epsilon$  parameter. Thus, the target EE was maintained, while SE was defined as a constraint, as described in Algorithm 3.

### 5.2. EE-SE trade-off comparison in realistic MIMO-NOMA scenarios

The comparison of the  $\epsilon$ -C-DK and  $\epsilon$ -C-SQP and  $\epsilon$ -C-HC methods are extensively discussed considering Scenario 1, 2, and 3 with a focus on the EE-SE trade-off performance analysis in MIMO-NOMA systems with data rate fairness and a variable number of users per cluster.

Let us consider the parameters values previously described in Table 1 for different  $r_D \in [50; 100; 200]$  [m] values associated to Scenario 1, and 3, respectively, and different path-loss coefficients. The three EE-SE-MIMO-NOMA algorithms performance is compared in Table 2 in terms of execution time (Time), percentage average number of convergence success (SN), and NMSE for the  $\epsilon$ -C algorithm, number of function evaluations (FunC), and number of the distinct solutions in Pareto front (S-PF) considering a different number of users and number of users per cluster.

Hence, for the MIMO-NOMA EE-SE trade-off, it can be seen that the DK method generates sub-problems with a degree of complexity similar to the problem addressed directly by the SQP method; thus, one can expect a higher elapsed time performance of  $\epsilon$ -C-DK regarding the  $\epsilon$ -C-SQP. This computational complexity increasing ranges from 100 times (1 cluster and 2 users) to 235 times (3 clusters and 6 users MIMO-NOMA configuration). In addition, it was observed that when the MIMO-NOMA scenarios become more complex, e.g., 3 clusters and 6 users, the  $\epsilon$ -C-DK method did not achieve full convergence in all executions. However, in the other soft configurations, when the optimization problem dimension is reduced, the  $\epsilon$ -C-DK method always attains full convergence in all

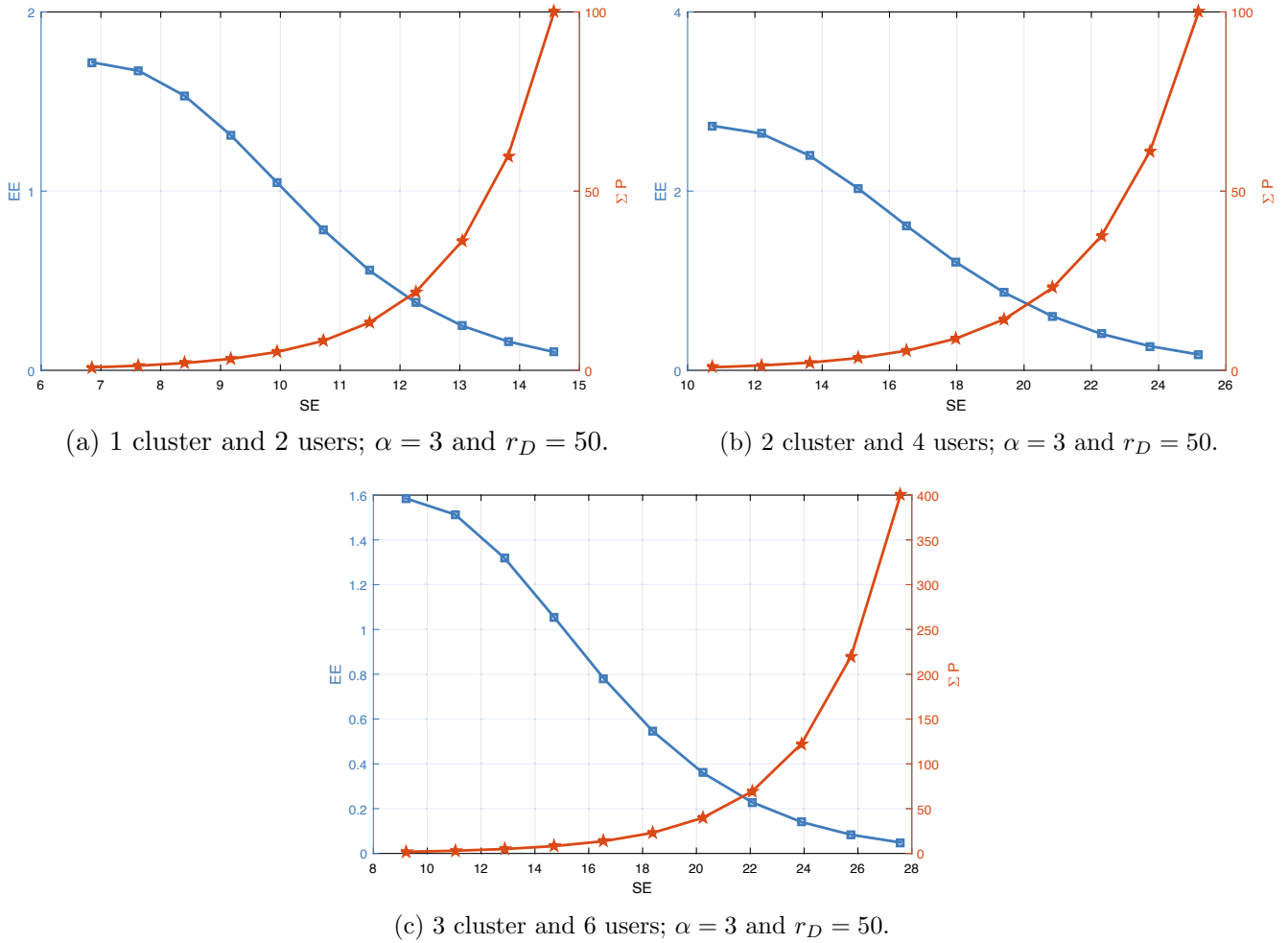


Fig. 5. EE-SE trade-off and power budget considering  $\epsilon$ -C-SQP method. 3 users/cluster and 6 users;  $\alpha = 3$  and  $r_D = 50$  m.

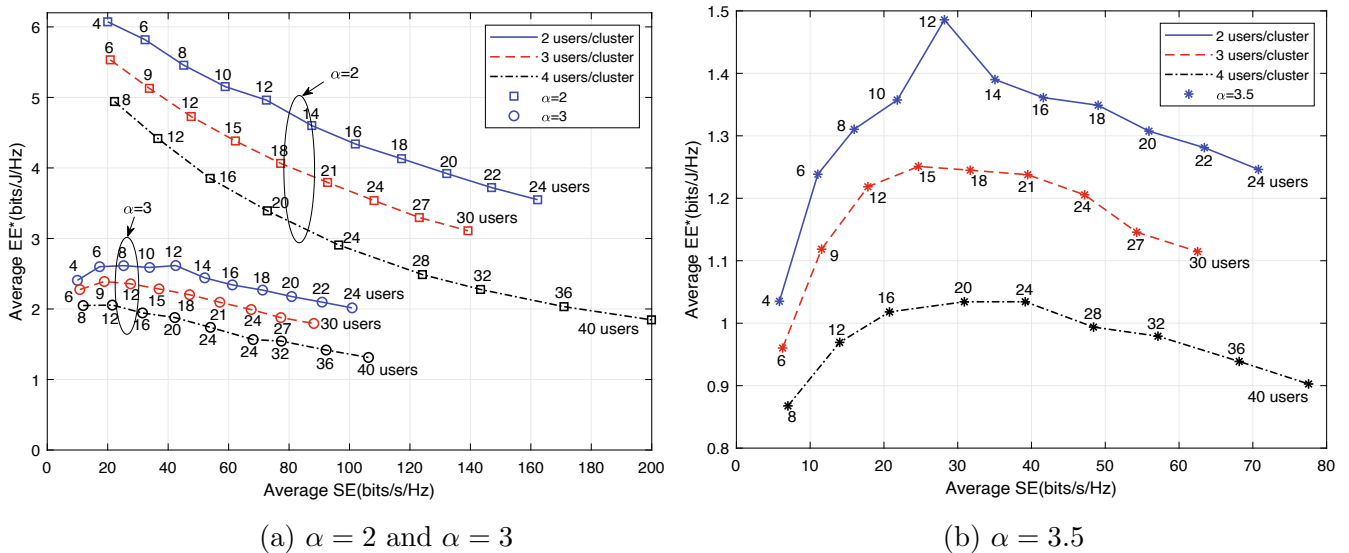


Fig. 6. Maximum EE extracted from Pareto front against SE for  $\epsilon$ -C-SQP method. Cell radius  $r_D = 100$ m; different distinct number of users for 2, 3 and 4 users/cluster, and path loss coefficients  $\alpha$ .

executions such as the  $\epsilon$ -C-SQP method. As expected by Theorem 1, the constraint  $c.\epsilon$  is satisfied as equality in all evaluated cases, as  $\epsilon$ -C-SQP. It is worth notice that the HC greedy local search-based method resulted

inefficient in solving the BOO EE-SE trade-off problem for all the three cell user distribution scenarios and the number of users and MIMO-NOMA clusters configurations: for more complicated (three clusters

**Table 3**

Optimal number of users  $M^*$  to attain optimum EE, considering cell radius  $r_D = 100$  m, different number of users/cluster  $K$ , and users per cell  $M$ .

Pathloss coef.	$K = 2$	$K = 3$	$K = 4$
$\alpha = 2.0$	4	6	8
$\alpha = 3.0$	8	9	12
$\alpha = 3.5$	12	15	24

**Table 4**

Figure of merit attained with WS-SQP method for the EE-SE MIMO-NOMA problem; path loss exponent  $\alpha = 3$ .

Method	$r_D$ [m]	SN [%]	Time [sec.]	$R^*$ [bps/Hz]	FunC	S-PF
<b>1 cluster and 2 users</b>						
WS-SQP	50	100	1.4610	3.5113	14077	6
WS-SQP	100	100	1.3420	2.2626	13032	7
WS-SQP	200	100	1.5191	1.3445	13711	8
<b>2 cluster and 4 users</b>						
WS-SQP	50	100	2.5017	2.9938	22374	5
WS-SQP	100	100	1.5666	1.6799	21050	6
WS-SQP	200	100	1.5398	1.1583	20392	7
<b>3 cluster and 6 users</b>						
WS-SQP	50	100	2.5882	3.9422	32346	4
WS-SQP	100	100	2.2437	2.8636	31798	6
WS-SQP	200	100	2.1345	1.9545	30772	6

and six users) scenario, the increase in the elapsed time required by the  $\epsilon$ -C-HC ranges from twice to six times regarding the  $\epsilon$ -C-DK method, and  $\approx 500$  times higher than the run time needed by the  $\epsilon$ -C-SQP method, while the NMSE accuracy of the  $\epsilon$ -C-HC is substantially worst regarding the  $\epsilon$ -C-SQP and  $\epsilon$ -C-DK methods. Indeed, the *superiority* of the  $\epsilon$ -C-SQP method can be confirmed comparing the complexity (Time elapsed) and NMSE of the three proposed methods in Fig. 2. (a) and 2. (b), respectively. Thus, SQP demands a smaller number of FunC, the iterative  $\epsilon$ -C-SQP optimization method have attained the best choice in terms of performance-complexity.

Fig. 3 depicts the Pareto front obtained with the three BOO methods considered in the EE-SE trade-off solutions in the MIMO-NOMA system. The solutions obtained were close to those determined by the methods  $\epsilon$ -C-DK and  $\epsilon$ -C-SQP, which reinforces the good performance of both proposed methods. It is worth noticing that for  $r_D = 50$ m, the  $\epsilon$ -C-HC method performed similarly to the other two proposed iterative-

analytical methods; however, for  $r_D = 100$ m and  $r_D = 200$ m a small degradation in the Pareto front solutions can be clearly observed. This degradation is reported in Table 2, through the NMSE accuracy values obtained by  $\epsilon$ -C-HC.

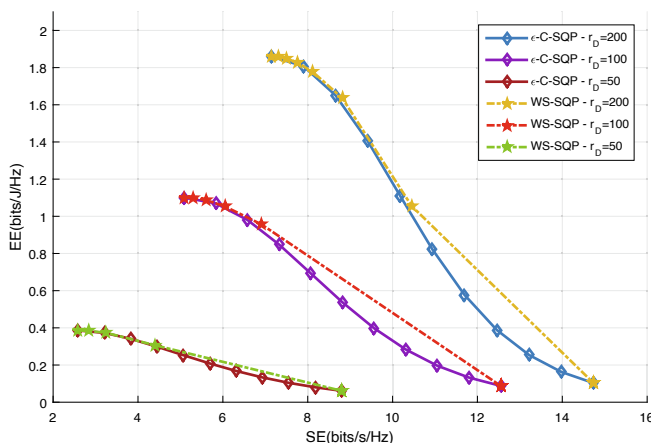
Fig. 4 evidences the EE and SE interdependence, considering the adopted system parameters. It can be noted that the curve obtained is bell-shaped in the considered domain. Also, the curves reveal the conflicting nature of the simultaneous SE and EE optimization, *i.e.*, the maximum SE is attainable after the EE reaches its maximum level and decreases monotonically. In fact, the points that satisfy the bi-objective EE-SE trade-off solution can be identified in the segments of the curves between the points where  $\max EE$  and  $\max SE$  occur, representing the solutions in the Pareto front.

Fig. 5 plots the SE vs. EE and the total power allocated ( $\Sigma P$ ) for the most effective iterative-analytical EE-SE optimization method, the  $\epsilon$ -C-SQP. The associated EE in Fig. 5 decreases with the increasing of SE, while the corresponding power increases to attain the equal user-rate  $R^*$  for all users across the entire cell area. This is because an increase in SE impacts negatively the total power requirement, which grows exponentially in the system, and therefore more energy resources must be allocated to satisfy the equal user rate fairness constraint.

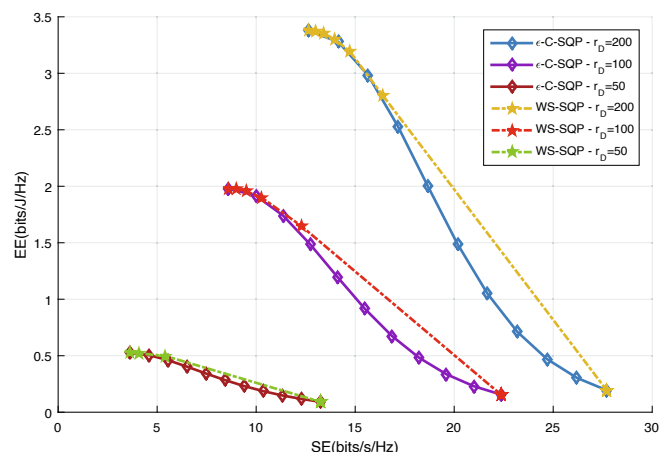
From the Pareto front, average EE  $\times$  SE curves have been obtained in Fig. 6 for  $\epsilon$ -C-SQP method, considering  $r_D = 100$ m. This graph shows the maximum (optimum) EE is attained with different number of users in the range of  $M \in \{4; 40\}$  users,  $K \in [2; 3; 4]$  users/cluster and considering three different path loss coefficients  $\alpha \in [2; 3; 3.5]$  and cell radius of  $r_D = 100$  m. Remembering that for each point in the same curve, all users have the same equal rate and therefore maximum fairness. Indeed, from these curves, it is possible to obtain the optimum number of users in the cell that results in the maximum EE for different path loss channel coefficients and the number of users/cluster in NOMA configurations. For instance, for  $\alpha = 2$  the  $EE^* = 6$ [bps/Hz] is attained under  $M^* = 4$  users when  $K = 2$  users/cluster. Such maximum EE decreases to  $EE^* \approx 1.03$ [bps/Hz] in  $K = 4$  users/cluster, which is attained with  $M^* = 20$  users in the cell. Table 3 summarizes the optimum number of users  $M^*$  to attain maximum EE under a distinct number of users per cluster and path loss coefficients. In general, the greater the number of users per cluster and the path loss coefficient, the lower the optimum EE.

### 5.3. Comparison with other scalarization method

In this subsection, we compare our scaling method  $\epsilon$ -C with the *weighted sum* (WS) optimization method proposed in [12]. Indeed, in the previous Section 5.2 the method  $\epsilon$ -C was extensively evaluated by



(a) Pareto front for 1 clusters and 2 users,  $\alpha = 3$ .



(b) Pareto front for 2 clusters and 4 users,  $\alpha = 3$ .

**Fig. 7.** Pareto front obtained by the two BOO methods  $\epsilon$ -C-SQP and WS-SQP for different Scenarios ( $r_D$  values); path loss coefficients  $\alpha = 3$ .

combining itself with the SQP, DK, and HC methods. Herein, we evaluate the WS scaling method combined with the SQP method in the same system and channel scenarios as those performed with  $\epsilon$ -C in Section 5.2.

The WS method re-scales the objective functions EE and SE of the original BOO problem in (27) as a single-objective problem, particularly as a convex combination of such objectives. Hence, given  $\tau \in [0, 1]$  one can consider the convex combination:

$$\left(1 - \tau\right) \cdot \text{EE} + \frac{\tau}{\bar{P}_{rf} + P_c} \text{SE},$$

where  $\bar{P}_{rf}$  is calculated in the solution vector  $(\bar{P}, \bar{\beta}_{a,k}, \bar{R})$  of the EE problem, as discussed in Section 4.2. Thus, we divide SE by  $\bar{P}_{rf} + P_c$  so that the possible combinations of EE and SE to be maximized have the same order of magnitude. The original bi-objective EE-SE trade-off problem can be described by the WS method as the following SOO problem:

$$\begin{aligned} & \underset{P, \beta_{a,k}, R}{\text{maximize}} \quad \left(1 - \tau\right) \frac{\sum_{a=1}^A \sum_{k=1}^K R_{a,k}}{Q \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} P + K \cdot A \cdot P_K + A \cdot P_{BS}} + \tau \frac{\sum_{a=1}^A \sum_{k=1}^K R_{a,k}}{\bar{P}_{rf} + P_c} \\ & \text{s.t. c.1} \quad \sum_{a=1}^A \sum_{k=1}^K \beta_{a,k} = 1 \\ & \text{c.2} \quad \beta_{a,k} = (2^R - 1) \frac{\sigma^2 |\mathbf{v}_{a,k}|^2}{P |\mathbf{v}_{a,k}^H \mathbf{H}_{a,k} \mathbf{w}_a|^2} + (2^R - 1) \sum_{i=1}^{k-1} \beta_{a,i}, \text{ where } \begin{cases} 2 \leq k \leq K \\ 1 \leq a \leq A \end{cases} \\ & \text{c.3} \quad \beta_{a,1} = \frac{(2^R - 1) \sigma^2 |\mathbf{v}_{a,1}|^2}{P |\mathbf{v}_{a,1}^H \mathbf{H}_{a,1} \mathbf{w}_a|^2} \text{ to } 1 \leq a \leq A. \\ & \text{c.4} \quad \beta_{a,k} > 0; P > 0; R > 0 \end{aligned} \quad (34)$$

The WS scalarization method to solve the EE-SE trade-off optimization problem can be summarized in Algorithm 6. Global solutions of (34), assuming  $\tau \in [0, 1]$ , belong to the Pareto boundary of the original BOO problem in (27). Details on the WS method applied to EE-SE trade-off in optical networks can be found in [12].

A pseudo-code for the MIMO-NOMA EE-SE trade-off solved by the WS method is presented in Algorithm 6. The vector  $\mathbf{p}^k$  denotes the vector solution for the problem (34) at the  $k$ -th iteration.

**Algorithm 6.** WS-MIMO-NOMA – Weighted Sum for the EE-SE Trade-off in MIMO-NOMA

- 1: **Choose:**  $\delta \in (0, 1)$ .
- 2: **Calculate:**  $\mathbf{p}^0 = (\bar{P}, \bar{\beta}_{a,k}, \bar{R})$ ;
- 3: **Define:**  $k = 1$  and  $\tau = \delta$ ;
- 4: **while**  $\tau \leq 1$  **do**
- 5:  $\mathbf{p}^k \leftarrow$  Solve problem (34) by SQP method;
- 6:  $\tau \leftarrow \tau + \delta$ ;
- 7: Save  $\mathbf{p}^k$ ;
- 8:  $k = k + 1$ ;
- 9: **end while**
- 10: **Output:**  $\mathbf{p}^0, \dots, \mathbf{p}^k$ ;

Table 4 presents the associated numerical results by applying the WS-SQP method to solve the EE-SE trade-off in MIMO-NOMA. It can be observed that the WS method generates more complex problems to be solved by the SQP method compared to the  $\epsilon$ -C method. The  $\epsilon$ -C-SQP method achieved convergence in all executions, for which the method attains distinct solutions, and, as expected by the Theorem 1, the (C. $\epsilon$ ) constraint in the problem (28) is equally satisfied in all cases evaluated, by checking the NMSE values achieved in Section 5.2. Hence, for the WS-SQP and  $\epsilon$ -C-SQP methods, one can conclude by the same behavior regarding the convergence success (SN), but the number of the distinct

solutions in Pareto front (S-PF) regarding those obtained by WS-SQP are less than  $\epsilon$ -C-SQP. Specifically, WS-SQP took more time and have consumed more FunC evaluations to achieve convergence compared with the  $\epsilon$ -C-SQP method.

Fig. 7 depicts the Pareto front obtained for both scalarization methods considered in the EE-SE trade-off MIMO-NOMA problem. It is also noted that the WS-SQP methods repeated solutions and, as expected due to their construction, the solutions are not evenly distributed on the curve. Besides, the solutions obtained by the WS-SQP were close to the  $\epsilon$ -C-SQP method, which reinforces the good performance of the proposed approaches.

Our extensive numerical results indicate that for the WS scaling method to be considered as an option for the EE-SE trade-off problem in MIMO-NOMA systems, it is necessary to consider a greater number of variations for the parameter  $\tau$  in order to determine more points in Pareto frontier. However, this strategy implies the further computational cost of the WS-SQP method; so, the method would have an even lower performance in this regard than the  $\epsilon$ -C-SQP method.

## 6. Conclusions

The numerical results have demonstrated that the EE-SE trade-off optimization design can be handled properly through the  $\epsilon$ -C method associated with the SQP within a bi-objective approach given the conflicting nature between EE and SE metrics from the point of view of resource efficiency in the MIMO-NOMA system. In addition, the  $\epsilon$ -C-DK and  $\epsilon$ -C-SQP methods were able to find EE-SE solutions at the Pareto front. As expected, it can be seen through extensive numerical tests carried out that the  $\epsilon$ -C method was able to find the different solutions at the Pareto frontier.

The remarkable performance-complexity of the  $\epsilon$ -C-SQP method, revealed by our extensive numerical results comes from the characteristics of the component functions of the EE-SE formulated as a bi-objective optimization trade-off. The method proved to be more suitable for this kind of optimization problem. Moreover, the three methods  $\epsilon$ -C-SQP,  $\epsilon$ -C-DK, and  $\epsilon$ -C-HC derived from  $\epsilon$ -C have validated the Theorem 1. Thus, if the constraint C. $\epsilon$  is satisfied as equality, then the solutions found are on the Pareto border, meaning that such solutions are efficient solutions. Additionally, we have compared the  $\epsilon$ -C-SQP method with the WS-SQP method; the sum-weighted method revealed a significantly lower diversity of points in the Pareto frontier than the  $\epsilon$ -C method while requiring a slightly higher computational cost as well. This reinforces the clear superiority of the  $\epsilon$ -C-SQP method for solving the EE-SE trade-off optimization problem in MIMO-NOMA systems.

The numerical results for the EE-SE trade-off optimization in MIMO-NOMA systems have demonstrated that the system is more efficient when there are major differences between users channel gain in the same cluster meaning that the MIMO-NOMA power domain principle can be explored advantageously. Besides, a high path-loss coefficient representing NLoS channels is a way to obtain it. As a result, for the maximum EE operation point, higher path-loss coefficient value implies in a higher number of users with same data-rate fairness in the cell. In contrast, for a smaller number of users per cluster, the maximum attainable EE increases.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgment

This work was supported in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil (CAPES) - Finance Code

001, in part by the National Council for Scientific and Technological Development (CNPq) of Brazil under Grant 310681/ 2019-7, and in part by the Universidade Tecnológica Federal do Paraná, Brazil (UTFPR), and State University of Londrina-Paraná State Government (UEL).

## References

- [1] Dai L, Wang B, Yuan Y, Han S, I C-L, Wang Z. Non-orthogonal multiple access for 5g: Solutions, challenges, opportunities, and future research trends. *IEEE Commun Mag* 2015;53(9):74–81.
- [2] Ding Z, Adachi F, Poor HV. The application of mimo to non-orthogonal multiple access. *IEEE Trans Wireless Commun* 2016;15(1):537–52.
- [3] Ali MS, Hossain E, Kim DI. Non-orthogonal multiple access (noma) for downlink multiuser mimo systems: User clustering, beamforming, and power allocation. *IEEE Access* 2017;5:565–77.
- [4] Ji B, Song K, Li C, ping Zhu W, Yang L. Energy harvest and information transmission design in internet-of-things wireless communication systems. *AEU Int J Electron Commun* 2018;87:124–7. <https://doi.org/10.1016/j.aeue.2018.01.038>.
- [5] Udalcovs A, Schatz R, Wosinska L, Monti P. Analysis of spectral and energy efficiency tradeoff in single-line rate wdm links. *J Lightwave Technol* 2017;35(10):1847–57.
- [6] Jacob JL, Abrão T. Nonorthogonal multiple access systems optimization to ensure maximum fairness to users. *Trans Emerg Telecommun Technol* 2020;31:1–14.
- [7] Sasikumar S, Jayakumari J. A novel method for the optimization of spectral-energy efficiency tradeoff in 5 g heterogeneous cognitive radio network. *Comput Netw* 2020;180:107389.
- [8] Liu F, Yang Q, He Q, Kwak KS. Energy efficiency and spectral efficiency tradeoff in downlink ofdma systems with imperfect csi. *AEU Int J Electron Commun* 2018;85:54–8.
- [9] Zhang S, Zhang N, Kang G, Liu Z. Energy and spectrum efficient power allocation with noma in downlink hetnets. *Phys Commun* 2018;31:121–32.
- [10] Ruan Y, Zhang R, Li Y, Wang C-X, Zhang H. Spectral-energy efficiency tradeoff in cognitive satellite-vehicular networks towards beyond 5g. In: 2019 IEEE Wireless Communications and Networking Conference (WCNC); 2019. p. 1–6. <https://doi.org/10.1109/WCNC.2019.8885680>.
- [11] Zuo Y, Zhu X, Jiang Y, Wei Z, Zeng H, Wang T. Energy efficiency and spectral efficiency tradeoff for multicarrier noma systems with user fairness. In: 2018 IEEE/CIC International Conference on Communications in China (ICCC); 2018. p. 666–70.
- [12] Pendeza Martinez CA, Abrão T. Spectral and energy efficiency tradeoff in optical code division multiple access networks. *Trans Emerg Telecommun Technol* 2021; 32:e4107. <https://doi.org/10.1002/ett.4107>.
- [13] Song Z, Ni Q, Sun X. Spectrum and energy efficient resource allocation with qos requirements for hybrid mc-noma 5g systems. *IEEE Access* 2018;6:37055–69.
- [14] Hanif MF, Ding Z, Ratnarajah T, Karagiannidis GK. A minorization-maximization method for optimizing sum rate in the downlink of non-orthogonal multiple access systems. *IEEE Trans Signal Process* 2016;64(1):76–88. <https://doi.org/10.1109/TSP.2015.2480042>.
- [15] Zeng M, Yadav A, Dobre OA, Poor HV. Energy-efficient power allocation for mimo-noma with multiple users in a cluster. *IEEE Access* 2018;6:5170–81.
- [16] Zeng M, Yadav A, Dobre OA, Tsiropoulos GI, Poor HV. Capacity comparison between mimo-noma and mimo-oma with multiple users in a cluster. *IEEE J Sel Areas Commun* 2017;35(10):2413–24.
- [17] Jain RK, Chiu D-MW, Hawe WR. A quantitative measure of fairness and discrimination for resource allocation in shared computer system. *ACM Trans Comput Syst* 1984:1–38.
- [18] Wang CH, Lin JY, Wu JM. Joint fairness and sum rate resource allocation for noma communications. In: *IEEE Conference on Standards for Communications and Networking. CSCN*; 2017. p. 269–74.
- [19] Boyd S, Vandenberghe L. *Convex Optimization*. Cambridge University Press; 2004.
- [20] Zhang Y, Wang HM, Zheng TX, Yang Q. Energy-efficient transmission design in non-orthogonal multiple access. *IEEE Trans Veh Technol* 2017;66(3):2852–7.
- [21] Tang J, Luo J, Liu M, So DK, Alsusa E, Chen G, et al. Energy efficiency optimization for noma with swipt. *IEEE J Sel Top Sign Proces* 2019;13(3):452–66.
- [22] Ehrhoff M, Ruzika S. Improved  $\epsilon$ -constraint method for multiobjective programming. *J Optim Theory Appl* 2008;138(3):375–96.
- [23] Nocedal J, Wright SJ. *Numerical Optimization*. 2nd ed. Springer Science, LLC; 2006.
- [24] Storey C. Applications of a hill climbing method of optimization. *Chem Eng Sci* 1962;17(1):45–52. [https://doi.org/10.1016/0009-2509\(62\)80005-0](https://doi.org/10.1016/0009-2509(62)80005-0).
- [25] Russell S, Norvig P. *Artificial intelligence: a modern approach*. 3rd ed. NJ: Prentice Hall; 2016.
- [26] Johnson A, Jacobson S. On the convergence of generalized hill climbing algorithms. *Discrete Appl Math* 2002;119(1):37–57.
- [27] Cao D, Chen M. Using penalty function and tabu search to solve cell formation problems with fixed cell cost. *Comput Oper Res* 2004;31(1):21–37. [https://doi.org/10.1016/S0305-0548\(02\)00144-2](https://doi.org/10.1016/S0305-0548(02)00144-2).

**ANEXO D – PDMA Large-Scale  
MIMO Energy Efficiency  
Maximization for eMBB, mMTC and  
uRLLC**

(P4 – artigo em processo de re-submissão, outubro de  
2022)

# PDMA Large-Scale MIMO Energy Efficiency Maximization for eMBB, mMTC and uRLLC

Jaime L. Jacob<sup>a,\*</sup>, Taufik Abrão<sup>a</sup>

<sup>a</sup>*Department of Electrical Engineering, State University of Londrina, Londrina 86057-970, Brazil.*

---

## Abstract

Recently, pattern division multiple access (PDMA) has been introduced in the fifth generation (5G) and beyond (B5G) scenarios to solve the insufficient spectrum resources problem. Unlike non-orthogonal multiple access (NOMA), PDMA schemes adopt pattern segmentation to separate user signals at the transmitter and SIC at the receiver. In this work, we combine the pattern division multiple access (PDMA) principle with Large-Scale Antenna MIMO (LSA-MIMO) to efficiently support the enhanced mobile broadband (eMBB), massive machine type communication (mMTC) and ultra-reliable and low latency communications (uRLLC) use modes in 5G systems. The addressed optimization problems also take into account the system energy efficiency (EE) metric. In addition, different parameters related to the pattern matrix design, including the overload factor (OF), quadratic sum squared correlation (SSC), diversity order (DIV), maximum row weight (MRW) and performance evaluation factor (PEF) have been considered to find the suitable, yet feasible solutions. We select pattern matrices that best meet the heterogeneous eMBB, mMTC, and uRLLC use mode requirements. Due to the wide range of possible combinations of pattern matrices, several PDMA matrices available in the literature with favorable features in terms of fitted requirements imposed by the different use modes have been analyzed and compared using EE, SSC, DIV, MRW PEF metrics; one that best meets such requirements has been selected.

*Keywords:* Pattern Division Multiple Access (PDMA), Large-Scale Antenna MIMO (LSA-MIMO), Energy Efficiency (EE), Fifth generation (5G), Enhanced Mobile Broadband (eMBB), Massive Machine Type Communication (mMTC), Ultra-Reliable and Low Latency Communications (uRLLC)

---

## 1. Introduction

5G wireless networks have been deploying to meet the growing demands for *internet of things* (IoT) networks. It is difficult for *orthogonal multiple access* (OMA) systems to simultaneously handle the needs of so many machine-type devices. To solve the current problem of insufficient spectrum

---

\*Corresponding author

*Email addresses:* `jaime.jacob@uel.br` (Jaime L. Jacob), `taufik@uel.br` (Taufik Abrão)

resources, 5G systems should adopt new multiple access schemes to increase spectrum efficiency [1]. In theory, *non-orthogonal multiple access* (NOMA) is ideal for reaching the limit of multi-user capacity, and therefore can be considered a promising candidate to meet the massive connectivity and super-system capacity requirements of 5G and B5G systems [2]. Recently, *pattern division multiple access* (PDMA) has been introduced in the 5G-B5G scenarios [3], [4], [5]. Unlike NOMA schemes, PDMA adopts pattern segmentation to separate user signals at the transmitter and applies SIC at the receiver side [3], [6], [7]. Indeed, PDMA is a newly proposed NOMA scheme to meet the numerous requirements of 5G-B5G by adopting patterns to separate user signals at the transmitter, where such patterns can be realized in multiple resource domains [1], [8]. By deploying *multi-domain multiplexing*, PDMA can take full advantage of wireless resources to increase the sum-rate with affordable computational complexity. Recently, increasingly advanced receptors have been investigated for operating in PDMA schemes. Although PDMA is a promising candidate for future communication systems, comprehensive and thorough research into the transceiver pattern and design is still needed [8].

We cannot confuse *path-division multiple access* [9, 10, 11, 12, 13] with *pattern division multiple access* (PDMA), [3], [4], [5]. Both have the same acronym, but with different meaning and systems application. The first is a multiplexing scheme that creates small radio frequency chains with high gains in a MIMO system under millimeter wave (mmWave) lenses whose response follows a *sync* function. This spatial multiplexing is generated for narrow band and broadband in millimeter wave communications, where parallel data streams are transmitted over different propagation paths with single-path processing. In second context of interest herein, the PDMA means design of the pattern matrix that defines the users and the resources deployed to transmit the users' signals, generating temporal, spatial and coding domains. In this paper we work with the latter concept of PDMA. Another recently published variation of PDMA is *holographic-pattern division multiple access* (HDMA), [14]. The main feature of HDMA is to use the holographic pattern obtained from the *reconfigurable holographic surface* (RHS). The RHS uses the metamaterial radiation elements to build a holographic pattern based on the principle of holographic interference. The purpose of this holographic pattern is to control the radiation amplitude of the leaking wave in each element effectively to achieve the desired signal transmission directions.

Considering the PDMA context treated in our work, the authors in [15] propose the use of PDMA with optimized power allocation (PA) strategy. They study the PA scheme to maximize the sum-rate in a downlink (DL) PDMA system with perfect and imperfect channel state information (CSI). The authors applied pattern matrices existing in the literature to develop the numerical results and corroborate the proposed PDMA with PA scheme. Besides, a resource allocation scheme that uses the Markov constrained decision process (MCDP) based on Deep Q-Network (DQN) is proposed in [16] to maximize the average network EE of a *simultaneous wireless information and power transfer* (SWIPT) based on PDMA with fast convergence. However, such analysis was carried out on a certain

pattern matrix without taking into account its characteristics and possible optimization.

PDMA technology has many applications in 5G and B5G services and use modes, such as eMBB, mMTC and uRLLC. The eMBB application services require high bandwidth to support 4K and 8K movies that can contain work or entertainment in the cloud. In the eMBB scenario, PDMA increase the multi-user capacity, providing more fairness against the near-far effect while improving the user experience in ultra-dense networks [17]. For small eMBB packets, PDMA can increase sum throughput, reduce latency and overhead when combining with grant-free protocols. Besides, the uRLLC services have extremely strict requirements regarding data throughput, delay, and reliability. Examples of uRLLC services are industrial automation, telemedicine surgery, smart grid distribution automation, transportation security, and unmanned aerial vehicle (UAV). Moreover, in the uRLLC use mode, grant-free transmission can be applied to achieve ultra-low latency. Although dedicated physical resources can be allocated to uRLLC user equipment (UE) aiming to achieve high reliability, the spectrum efficiency (SE) results poorly low, especially for aperiodic or burst-based services. To attain high reliability, low latency and high spectrum efficiency at the same time, the PDMA transmission method represents an appropriate choice, and can be applied to increase reliability in a communication network subject to collisions [17]. The application of PDMA allows efficient multiplexing of uRLLC and eMBB services to further improve resource utilization. The mMTC services contain a large number of IoT devices with low amount of data sent and high latency tolerance, while require low manufacturing cost and long battery life, such as in Smart City applications [18].

### *1.1. PDMA for Distinct 5G User Mode Scenarios*

In a mMTC scenario, a BS needs to provide a connection to a large number of endpoints. Hence, the main challenge for a mMTC scenario is how to effectively handle massive connections with power constraints [17]. Indeed, PDMA is a very competitive method to solve the problem of massive connectivity along with the requirement of wide system coverage. A succinct characterization of these services can be presented as follows: (a) eMBB supports stable connections with very high peak data rates as well as moderate rates for cellular edge users; (b) mMTC supports a large number of IoT devices, which are active only sporadically and send small data loads; (c) uRLLC supports low latency transmissions of small loads with very high reliability from a limited set of endpoints, which are active according to patterns normally specified by external events such as alarms [19]. For the PDMA pattern to better serve the eMBB use mode, the user's throughput should be maximized. On the other hand, in the mMTC use mode, the overload factor (OF) in the PDMA must be maximized, while in the uRLLC scenario, for simultaneously attaining low latency and high reliability, the inter-patterns interference (IPI) must be mitigated, by deploying patterns with low OFs and minimizing the sum squared correlation (SSC) [1].

## 1.2. PDMA in LSA-MIMO Systems

*Large-scale antenna* (LSA) arrays MIMO are one of the key technologies of the future of mobile communications systems. LSA-MIMO systems have been proposed to significantly improve the sum of system rates with extra degrees of freedom that facilitate spatial multiplexing gains and transmission diversity [8], [20]. Examples of LSA-MIMO systems include *massive* MIMO (M-MIMO) [21, 22] and *extra-large* MIMO (XL-MIMO) [23, 24] systems. Such arrangements can provide enough spatial resources for a large number of simultaneous connected devices in a so called crowded scenarios. Currently, the application of LSA-MIMO to PDMA has received increasing attention aiming to confirm further performance improvement [7], [8].

In [1], the authors establish criteria for obtaining the PDMA pattern matrices in UL and DL for the three 5G use mode scenarios, presenting a numerical performance simulation for each service; the authors consider only one transmission antenna. On the other hand, [7] proposes the LSA-MIMO PDMA system in DL transmission scenario with hybrid detection where the spatial filter (SF) is used at the receiver with the SIC to separate signals from multiple overlapping domains. Furthermore, it formulates the sum-rate maximization problem to obtain the optimal pattern mapping policy. The simulation results show that the LSA-MIMO PDMA scheme presents a substantial sum-rate performance gain w.r.t. the OMA and NOMA systems. The approach presented in [8] is similar to the one presented in [7]; the difference lies in the addition of beam allocation optimization step. The simulations reveal that the combination of the two optimizations steps presents better performance. Importantly, in [8] and [7] the authors do not consider 5G applications, such as neither uRLLC, nor mMTC, eMBB. So far, there is a lack in the literature in analyzing the PDMA pattern matrices that best meet the 5G service requirements. In this work we propose systematically evaluate the pattern matrices that suitably meet the requirements and features of 5G services aiming to optimize either the sum-rate or system energy efficiency metrics.

## 1.3. Contributions

We propose to apply an innovative transmission scheme, the multiuser PDMA, operating under LSA-MIMO hybrid 5G use modes traffic, combining eMBB, mMTC, and uRLLC use modes. Such a transmission scheme is based on the *pattern matrix* design. Hence, we adapt the pattern matrices available in the literature to work on the DL LSA-PDMA system. Hence, we evaluate their parameters by applying five different figures-of-merit to identify the one pattern matrix that better meets the specifications and requirements of combined use mode 5G services.

To meet the combined eMBB, mMTC and uRLLC use mode requirements and constraints, we propose optimizing different PDMA transmission parameters and metrics. Hence, the overload factor (OF) PDMA parameter has been maximized for the mMTC scenario, while for the uRLLC use mode, the *sum squared correlation* (SSC) and the *weighted maximum row weight* (MRW) metrics were minimized. The transfer rate *vs.* energy efficiency (EE) curves with the respective pattern

matrix for each scenario have been comprehensively analysed. To corroborate the superiority of the proposed transmitting methodology, we propose and analyse the EE maximization method taking into account the maximum fairness condition or alternatively assume the equal rate among users. The EE optimization under maximum fairness ensures suitable power distribution among users to attain the same rate. Various parameters of the pattern matrix will also be evaluated for different hybrid eMBB, uRLLC and mMTC scenarios.

#### 1.4. Notation and Organization

Uppercase boldface letters is adopted for matrices and lowercase boldface for vectors. The operation  $\mathbb{E}[\cdot]$  represents the statistical expectation operator,  $(\cdot)^T$  and  $(\cdot)^H$  stands for the transpose and the conjugate transpose operations respectively. The forms  $\|\cdot\|$  and  $|\cdot|$  represents the norm of the vector and the absolute value of a number, respectively. The operator  $\circ$  denotes the Hadamard product;  $!$  represents the factorial of a number, and  $\lceil \cdot \rceil$  rounding to the first higher integer.

The remainder of this work is organized as follows. The system model for the mMIMO PDMA is described in section 2. Section 3 explores the pattern matrix design for hybrid 5G use modes (eMBB, mMTC and uRLLC). Section 4 describes criteria used to select the arrays that best serve 5G services. The energy efficiency (EE) optimization formulation is described in Section 5. Section 6 presents the analysis of the pattern matrices and the PDMA matrix selection methodology that best serve each 5G service. Finally, Section 7 summarizes the paper conclusions and future research directions.

## 2. System Model

A large number of antennas at the BS can be spread along the geographical cell area, namely large-scale distributed MIMO<sup>1</sup>, or implemented centrally at the BS, known as large scale arrays (LSA) [20]. As shown in Fig. 1, in this paper, a base station (BS) equipped with LSA-MIMO transmitting to multiple users (downlink) using the pattern division multiple access (PDMA) principles is considered as in [7], [8]. Each user is equipped with  $N_R$  antennas. It is assumed that there are multiple *antenna clusters* (AC) located at the BS side; each AC is equipped with  $N_T$  antennas that form  $N$  beams, where

$$N_T \geq N. \quad (1)$$

All users are divided into multiple *user groups* (UGs) and each UG contains  $K$  users, where  $N_T \leq KN_R$ . An AC is supposed to cover a UG with

$$N \leq K \leq 2^N - 1. \quad (2)$$

---

<sup>1</sup>Also known as cell-free systems.

Table 1: Acronyms

	Definition		Definition
5G	Fifth Generation	PA	Power Allocation
AC	Antenna Clusters	PDMA	Pattern Division Multiple Access
AWGN	Additive White Complex Gaussian Noise	PEF	Performance Evaluation Factor
B5G	Beyond 5G	QAM	Quadrature amplitude modulation
BPSK	Binary Phase Shift Keying	QoS	Quality of Service
BS	Base Station	QPSK	Quadrature Phase Shift Keying
CSI	Channel State Information	RE	Resource Efficiency
MCDP	Markov Constrained Decision Process	RF	Radio Frequency
DL	Downlink	RHS	Reconfigurable Holographic Surface
DQN	Deep Q-Network	RV	Resource Values
EE	Energy Efficiency	SCUs	Small Cell Users
eMBB	Enhanced Mobile Broadband	SE	Spectral Efficiency
HDMA	Holographic-Pattern Division Multiple Access	SIC	Successive Interference Cancellation
IoT	Internet of Things	SINR	Signal-to-Interference-plus-Noise Ratio
IPI	Inter-Patterns Interference	SSC	Sum Squared Correlation
LSA	Large-Scale Antenna	SQP	Sequential Quadratic Programming
MC	Multi-Carrier	SWIPT	Simultaneous Wireless Information and Power Transfer
MIMO	Multiple Input Multiple Output	UEs	User Equipments
mMTC	Massive Machine Type Communication	UL	Uplink
MRW	Maximum Row Weight	UG	User Group
NOMA	Non-Orthogonal Multiple Access	uRLLC	Ultra-Reliable and Low Latency Communications
OF	Overload Factor	XL	Extra-Large
OMA	Orthogonal Multiple Access	ZFBF	Zero-Forcing Beam-Forming

Without loss of generality, we simplify the scenario in which an AC communicates with a UG. For example, consider a *cluster of antennas* (CA) with  $N_T$  antennas; in such a simplification, the set of antennas serves only a user group of  $K$  users and only a specific UG.

Let  $\mathbf{H}_k \in \mathbb{C}^{N_R \times N_T}$  be the matrix of the channel between the AC and the  $k$ th user of the UG where

$$\mathbf{H}_k = \frac{\mathbf{Q}_k}{\sqrt{1 + d_k^\alpha}} \quad (3)$$

and  $\mathbf{Q}_k \in \mathbb{C}^{N_R \times N_T}$  denoting a Rayleigh fading channel matrix, and  $\alpha$  is the path loss exponent;  $d_k$  represents the distance from the  $k$ th user to the BS. It was adopted that the BS has perfect channel status information (CSI). For convenience, we will use zero-forcing beam-forming (ZFBF) in the transmitter. Let  $\mathbf{b}_n \in \mathbb{C}^{N_T \times 1}$  represent ZFBF vector of the  $n$ th beam, and it is generated based on the user's CSI selected among the users covered by the  $n$ th beam represented with the index  $n_{\text{sel}}$ . And the set of beam pairs and the corresponding selected user is expressed as  $\Omega = \{(1, 1_{\text{sel}}), (2, 2_{\text{sel}}), \dots, (n, n_{\text{sel}}), \dots, (N, N_{\text{sel}})\}$ . Then, the composite channel matrix of the se-

Table 2: Nomenclature

	<b>Description</b>
$N_T$	Number of Transmit Antennas on the BS
$N_R$	Number of Reception Antennas Per User
$N$	Beams Number
$K$	Users Number
$\mathbf{H}_k$	Channel Matrix of the $k$ th UG User
$\mathbf{Q}_k$	Channel Matrix With Rayleigh Fading
$\alpha$	Path Loss Exponent
$d_k$	Distance From the $k$ -th User To the BS
$\mathbf{b}_n$	Vector ZFBF Of the $n$ -th Beam
$\mathbf{H}_C$	Selected User Channel Matrix.
$\mathbf{B}_C$	Matrix Precoding ZFBF
$\mathbf{B}$	Adjusted ZFBF Precoding Array
$\mathbf{s}$	Vector Transmission Symbol
$\mathbf{P}$	Power Allocation Matrix
$\mathbf{G}$	Beam Allocation Matrix
$\mathbf{t}$	Signal Vector Overlay after Pattern Mapping
$\mathbf{x}$	AC Signal Transmission Vector
$\mathbf{y}_k$	Signal Vector Received from $k$ -th User
$\mathbf{w}_k$	Vector Additive White Gaussian Noise for the $k$ -th User
$\mathbf{G}_{\text{PDMA}}^{N \times K}$	Pattern Matrix PDMA of Dimensions $N \times K$
$\iota$	Overload Factor
$v_j$	$j$ -th RV Transmission Signal
$x_k$	$k$ -th User Modulation Symbol
$\mu$	Power Scale Factor
$\varphi$	Phase Change Factor
$\lambda^N$	Number of possible PDMA Patterns
$r$	Repeatability
SSC	Correlation of the Quadratic Sum to the Square
DIV	Diversity Order
$d_{\text{max}}^w$	Maximum Weighted Line Weight
$\Lambda$	Performance Evaluation Factor
$\mathbf{V}_k$	SF Spatial Filter Matrix
$\mathbf{v}_{nk}$	Vector of the $n$ -th Column of $\mathbf{V}_k$
$z_{nk}$	Scalar Result after Suppressing Beam Interference
$u_{nk}$	Normalized Channel Gain Equivalent Between the $k$ -th User and AC
$q_{nk}$	Sum of Interference between Beams and Noise after Normalization
$\mathbf{U}$	Normalized Channel Gain Equivalent Between UG and AC
$\gamma_{nk}$	Signal-Interference+Noise Ratio (SINR) of the $k$ -th User on the $n$ -th Beam

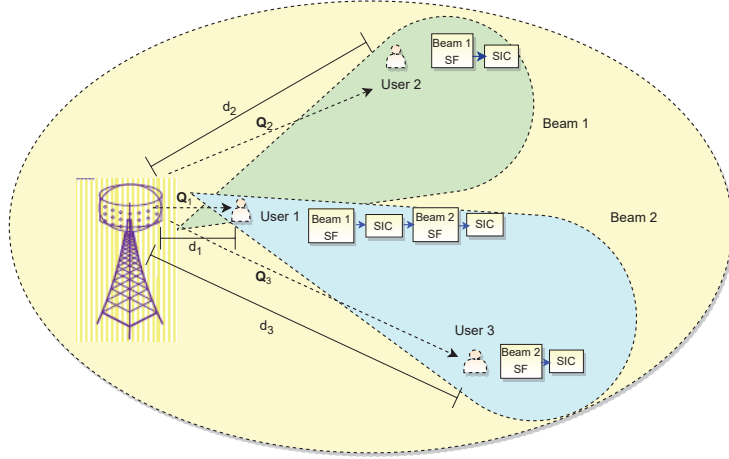


Figure 1: The LSA-MIMO PDMA transmission scheme.

lected users can be expressed as  $\mathbf{H}_C = [\mathbf{H}_{1_{\text{sel}}}^H, \mathbf{H}_{2_{\text{sel}}}^H, \dots, \mathbf{H}_{N_{\text{sel}}}^H]^H$  the ZFBF composite matrix can be expressed as

$$\mathbf{B}_C = \mathbf{H}_C^H (\mathbf{H}_C \mathbf{H}_C^H)^{-1}. \quad (4)$$

We partition the composite ZFBF matrix as  $\mathbf{B}_C = [\mathbf{B}_1, \mathbf{B}_2, \dots, \mathbf{B}_n, \dots, \mathbf{B}_N]$  and then the ZFBF vector of the  $n$ th beam can be expressed as follows

$$\mathbf{b}_n = \mathbf{B}_n \mathbf{1}^{N_R \times 1}, \quad (5)$$

where  $\mathbf{1}^{N_R \times 1}$  denotes an  $N_R$ -dimension column vector with all one elements. Therefore, let  $\mathbf{B} \in \mathbb{C}^{N_T \times N}$  denote the ZFBF matrix and it can be expressed as

$$\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n, \dots, \mathbf{b}_N]. \quad (6)$$

The PDMA system in Figure 1 uses pattern mapping at the transmitter with power allocation and beam allocation domains to overlay the signals. Consider that  $\mathbf{s} \in \mathbb{C}^{K \times 1}$  denote the transmit symbol vector for the UG with  $\mathbf{s} \sim \mathcal{CN}(0, \mathbf{I}_K)$ ,  $[\mathbf{s}]_k = s_k$  the transmit symbol for the  $k$ th user. Let  $\mathbf{P}^{1/2} \in \mathbb{R}^{N \times K}$  denote the power allocation matrix,

$$[\mathbf{P}^{1/2}]_{nk} = \sqrt{p_{nk}} \quad (7)$$

the transmit power allocated to the  $k$ th user in the  $n$ th beam. Let us consider the binary matrix  $\mathbf{G} \in [0, 1]^{N \times K}$  denoting the *beam allocation matrix*:

$$[\mathbf{G}]_{nk} = g_{nk} \quad (8)$$

where the beam allocation factor for the  $k$ th user in the  $n$ th beam assumes the values  $g_{nk} \in \{0, 1\}$ . When  $g_{nk} = 1$  it means that the  $k$ th user is covered by the  $n$ th beam and for  $g_{nk} = 0$  this user is

not covered by this beam. The result of the *superimposed signal vector after pattern mapping* can be expressed by

$$\mathbf{t} = \left( \mathbf{G} \circ \mathbf{P}^{1/2} \right) \mathbf{s}, \quad (9)$$

where  $\circ$  denotes the operation of Hadamard product, and the  $n$ th element of the superimposed signal vector after pattern mapping is defined as:

$$[\mathbf{t}]_n = t_n = \sum_{k=1}^K b_{nk} \sqrt{p_{nk}} s_k, \quad n = 1, 2, \dots, N. \quad (10)$$

For two matrices of the same dimensions  $I \times J$ , the Hadamard product is a matrix of the same dimensions as the operands, with the  $(i, j)$ th element given by  $[\mathbf{G} \circ \mathbf{P}]_{i,j} = [\mathbf{G}]_{i,j} [\mathbf{P}]_{i,j}$ .

At the transmitter, we have  $\mathbf{t} \in \mathbb{C}^{N \times 1}$  representing the superimposed signal vector after the pattern mapping process. Let  $\mathbf{x} \in \mathbb{C}^{N_T \times 1}$  be the vector of the AC transmission signal which can be expressed as follows

$$\mathbf{x} = \mathbf{B}\mathbf{t}. \quad (11)$$

At the receiver, we have  $\mathbf{y}_k \in \mathbb{C}^{N_T \times 1}$  which represents the vector of the received signal vector for the  $k$ th user which can be expressed as follows

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x} + \mathbf{w}_k \quad (12)$$

where  $\mathbf{w}_k \sim \mathcal{CN}(0, \sigma_k^2 \mathbf{I}_{N_R})$  is the additive white Gaussian noise vector whose elements have zero mean and variance  $\sigma_k^2$ .

### 3. Pattern Matrix for Hybrid 5G Use Modes

#### 3.1. Basic Pattern Matrix Rules

The PDMA pattern matrix, dimension  $N \times K$ , consists of 0s and 1s elements, being represented by matrix  $\mathbf{G}_{\text{PDMA}}^{N \times K}$ . The 0s elements indicate that the user cannot use the corresponding resource values (RV), while the 1s elements represent the corresponding use of such RV. The rows and columns of  $\mathbf{G}_{\text{PDMA}}$  represent users and RV, respectively. The dimension and sparsity level are properties of the PDMA pattern matrix that determine the complexity and system performance [6]. Therefore, the pattern matrix is designed with the following principles [17], [25]:

- (a) A pattern with more 1 elements in the pattern gives a higher order of diversity.
- (b) The number of users with 1 elements on each row must be equal.
- (c) Selected patterns should have as many different diversity orders as possible.

Suppose that  $K$  users are mapped into  $N$  orthogonal physical resource elements based on the PDMA pattern matrix. Based on the requirements specified above, the maximum number of users supported with  $N$  resource elements can be calculated by

$$K_{\max} = \sum_{n=1}^N \binom{N}{n} = \binom{N}{1} + \binom{N}{2} + \cdots + \binom{N}{N} = 2^N - 1 \quad (13)$$

where  $\binom{N}{K}$  denotes the set of all combinations  $K$  users of a set  $N$  RV. As an example by applying (13), one can draw the PDMA structures for  $\mathbf{G}_{\text{PDMA}}^{N \times K}$  with  $N = 2, 3$  and 4 as follows:

$$\mathbf{G}_{\text{PDMA}}^{2 \times 3} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ \underbrace{\binom{2}{2}}_{=1} & \underbrace{\binom{2}{1}}_{=2} \end{bmatrix} \quad (14)$$

$$\mathbf{G}_{\text{PDMA}}^{3 \times 7} = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ \underbrace{\binom{3}{3}}_{=1} & \underbrace{\binom{3}{2}}_{=3} & \underbrace{\binom{3}{1}}_{=3} \end{bmatrix} \quad (15)$$

$$\mathbf{G}_{\text{PDMA}}^{4 \times 15} = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ \underbrace{\binom{4}{4}}_{=1} & \underbrace{\binom{4}{3}}_{=4} & \underbrace{\binom{4}{2}}_{=6} & \underbrace{\binom{4}{1}}_{=4} \end{bmatrix} \quad (16)$$

Specifically, consider a pattern matrix  $3 \times 5$  used as an example, which is shown as follows

$$\mathbf{G}_{\text{PDMA}}^{3 \times 5} = \begin{array}{c} \text{RV1} \\ \text{RV2} \\ \text{RV3} \end{array} \begin{array}{ccccc} \text{User1} & \text{User2} & \text{User3} & \text{User4} & \text{User5} \\ \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{bmatrix} \end{array} \quad (17)$$

where element 1 indicates that the data is mapped to the specified RV. The matrix in (17) shows the process of mapping resources to PDMA. We assume that each user is assigned a PDMA standard. After encoding, multiple data streams are transmitted simultaneously by different RV according to the standard matrix. As we can see in (17), the data of user 1 (first the column) is mapped to the first, second and third RV. Data from the user 2 is mapped to the first and second RV and so on.

### 3.2. Overload Factor in the Pattern Matrix

The *overload factor* (OF) is defined as the ratio of the number of columns to the number of rows in a PDMA pattern matrix [17]:

$$\iota = \frac{K}{N} \quad (18)$$

It reflects the excessive number of users multiplexed on the same PDMA resource elements, and can be compared directly to the OMA schemes. Hence, assuming the number of resources  $N = 2$  and the number of users  $K = 3$  as an example, the OF  $\iota = \frac{K}{N} = 1.50$  means that the PDMA supports one and a half times the number of users compared to OMA.

To obtain a certain overload factor, just vary the dimensions of the PDMA pattern matrix. As the array size increases, greater diversity gains and an increase in detection complexity are also provided; besides, this is accompanied by an improvement in performance [6]. For each overload factor, the pattern matrix can be dimensioned to obtain the best performance-complexity trade off.

### 3.3. Extended PDMA Pattern Matrix Design

Let us define *ambiguity* in the pattern matrix and then present the extended PDMA pattern matrix; we start with an example where there are three users and two resource values RV, represented by  $\mathbf{G}_{\text{PDMA}}^{2 \times 3}$  [17]. The signals to be transmitted with this RV are given by

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad (19)$$

where  $v_j$  is the transmission signal on the  $j$ th RV, and  $x_k$  is the modulation symbol of the  $k$ th user. From (19), the signal to be transmitted in each RV is given by the linear combination of the modulated symbols:

$$\begin{aligned} v_1 &= x_1 + x_2 \\ v_2 &= x_1 + x_3. \end{aligned} \quad (20)$$

Let us assume that the three users transmit with BPSK modulation. If users 1 and 2, therefore  $x_1$  and  $x_2$ , transmit the symbol  $-1$  and user 3, therefore  $x_3$ , the symbol  $+1$ . For a noise-free channel, the  $v_1$ , from RV 1, receives the combined signal  $-2$  and the  $v_2$ , from RV 2, receives the combined signal  $0$ . In RV 1 the receiver can conclude that the transmitted symbols were  $[+1, +1]$ . After receiving the symbol  $0$  from RV 2, the receiver can retrieve the symbol as  $[+1, -1]$  or  $[-1, +1]$ , therefore it is impossible to retrieve transmitted symbols correctly and accurately.

With this simple example, one can see that the combined constellation has a non-uniform distribution which results in a combination that leads to the ambiguity problem. Such an *ambiguity problem* can be solved by adding a scalar factor and a phase shift to the pattern matrix PDMA.

Before combining the two symbols, the power scaling factor ( $\mu$ ) and phase-shifting factor ( $\varphi$ ) must be entered properly, modifying the elements of PDMA pattern matrix; hence,  $v_1$  in (20) becomes:

$$v_1 = \sqrt{\mu}e^{j\varphi}x_1 + \sqrt{1-\mu}x_2 \quad (21)$$

Consider the example in Fig. 2 that plots  $v_i$  of Eq. (20) (with ambiguity), and Eq. (21) (unambiguous) with  $\varphi = \pi/4$  and  $\mu = 0.5$  for two users combined, for both QPSK and 16-QAM modulations. Fig. 2 a) and c) are unambiguous where it can be seen that there are several points that coincide with the  $x$  and  $y$ -axis. In Fig. 2 b) and d) with the correction of the ambiguity, the same does not occur, so it can be concluded that the ambiguity has been resolved. Additionally, this new distribution of the combined constellation approximates a Gaussian distribution. Consequently, there is a maximization of the capacity under the additive white Gaussian noise (AWGN) channels. This means that introducing a phase shift factor results in a channel capacity gain.

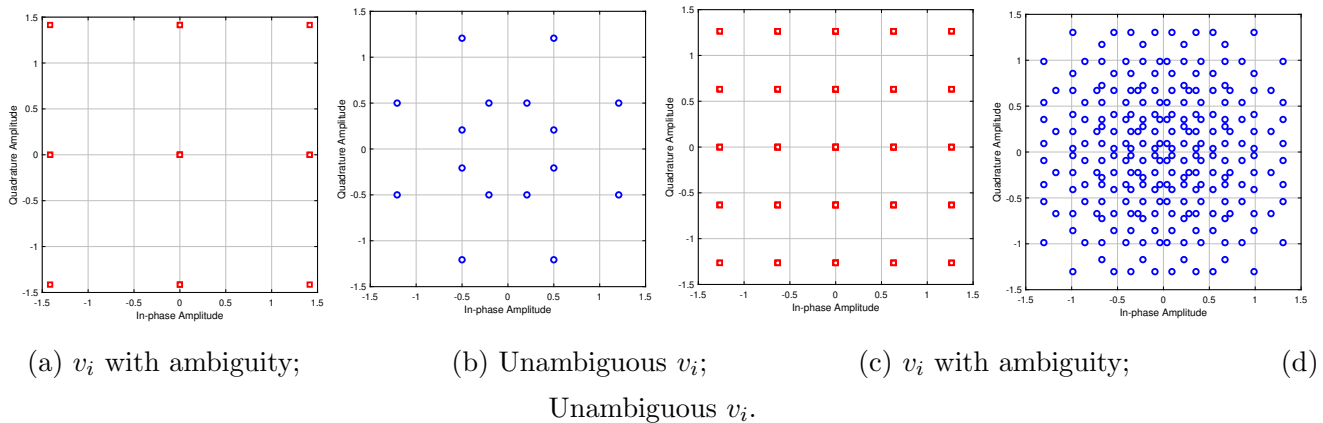


Figure 2: Combined constellation of two users for phase shift  $\varphi = \pi/4$  and power scaling factor  $\mu = 0.5$ . (a) and (b) represent  $v_i$  points from QPSK constellations, while subplots (c) and (d) represent  $v_i$  points from 16-QAM constellations.

As a result, power scaling and phase shift can be incorporated directly into the PDMA pattern matrix. The 1 value is replaced by the complex value that corresponds to the squared root of the power scaling and phase shift in the PDMA pattern matrix. This matrix is now called the *extended PDMA pattern* matrix. For this example, the extended pattern matrix is given by

$$\mathbf{G}_{\text{E-PDMA}}^{2 \times 3} = \begin{bmatrix} \sqrt{\mu_{11}}e^{-j\varphi_{11}} & \sqrt{\mu_{12}}e^{-j\varphi_{12}} & 0 \\ \sqrt{\mu_{21}}e^{-j\varphi_{21}} & 0 & \sqrt{\mu_{23}}e^{-j\varphi_{23}} \end{bmatrix} \quad (22)$$

where  $\mu_{jk}$  and  $\varphi_{jk}$  is the power scaling factor and phase shifting factor on the  $j$ th RV of the  $k$ th user. To obtain the ideal value of the power scale factor and phase shift factor, it is necessary to know the number of users and the adopted constellation.

#### 4. Pattern Matrix Selection and Evaluation Criteria

This section provide a simple methodology for the PDMA matrix selection operating under hybrid uRLLC, mMTC and eMBB use. The pattern matrix used in the DL PDMA transmission should be constrained by the total transmit power available in the BS, which is limited by [1]:

$$\sum_{k=1}^K \sum_{n=1}^N |g_{nk}|^2 = 1 \quad (23)$$

This limitation ensures that the total transmit power of the BS is not greater than the maximum DL transmit power. Besides, let us consider

$$\mathbf{G}^0 = \lceil \mathbf{G}_{\text{PDMA}}^{N \times K} \rceil = [\mathbf{g}_1^0, \mathbf{g}_2^0, \dots, \mathbf{g}_K^0] = [\mathbf{g}_1^{0T}, \mathbf{g}_2^{0T}, \dots, \mathbf{g}_N^{0T}]^T \quad (24)$$

where the symbol  $\lceil \cdot \rceil$  represents rounding to the first higher integer value.

##### 4.0.1. Pattern Matrix Availability

There are  $\lambda^N$  possible PDMA patterns for a PDMA pattern matrix, where  $\lambda$  is the number of possible values that  $g_{nk}$  in  $\mathbf{g}_n = [g_{n1} g_{n2} \dots g_{nK}]$  can assume. Knowing that  $K$  is the number of columns needed for a given value of OF, one can choose  $K$  PDMA patterns among  $\lambda^N$  candidates to build a PDMA matrix of patterns. For this case, the number of possible combinations is given by [26]:

$$C_{\lambda^N}^K = \frac{\lambda^N!}{(\lambda^N - K)!K!}. \quad (25)$$

Assuming  $N$  or  $\lambda$  becomes large,  $C_{\lambda^N}^K$  is too large to obtain via *exhaustive search* (ES) the ideal PDMA pattern matrix.

##### 4.1. Challenging in Optimizing Pattern Matrices for Hybrid Use Modes in 5G

The number of possible combinations to generate the pattern matrix can be calculated by the equation (25). It can be concluded that the range of possible combinations is extremely large, which makes it practically impossible to be generated exhaustively. Hence, we skip such approaches coming with huge efforts, by adopting and analyzing pattern matrices available in the literature. Notice that the pattern matrices available in the literature are suboptimal. In this work, such pattern matrices were analyzed and compared by adopting several figure-of-merit in order to validate our search methodology to find the best yet quasi-optimal pattern matrix that meets the requirements of different use modes (uRLLC, mMTC and eMBB) in 5G services.

The difficulty in defining the pattern matrix that serves more than one 5G use mode lies in the antagonistic characteristics among the use modes. As an example, among the mMTC and the uRLLC use modes, the former mode must serve as many machine-type terminals as possible, which requires a pattern matrix of higher OF, while a more appropriate pattern matrix for the uRLLC service must

result in a lower OF to meet low-latency requirement [1]. It can be concluded that hardly a pattern matrix will fully meet all use modes requirements simultaneously, *i.e.*, when operating in hybrid mode. In such cases, one of the use mode will have to tolerate a certain deterioration in one of its requirements. Hence, our approach is to find the PDMA matrix among those matrices available in the literature that provide the best combined performance considering the 5G system operating under hybrid (two or three) use modes simultaneously.

#### 4.2. Pattern Matrix Evaluation Criteria

There are different yet complementary criteria to directly and indirectly evaluate the PDMA matrix, including: a) *Repeatability*; b) *Correlation*; c) *Diversity*; d) *Performance factor*; e) *Weighted maximum row weight*; f) *Energy efficiency of the system*; g) *Spectral efficiency*; h) *Resource efficiency*. In the sequel we discuss the former five criteria.

##### 4.2.1. Repeatability

In the PDMA context, the parameter related to the interference between users is named *repeatability*,  $r$ . Such a parameter expresses the repeatability of binary parameters corresponding to the standard matrix. By reducing its value, the interference between users is also reduced. The repeatability is given by

$$r = \max[\text{REP}(\mathbf{g}_1^0, \mathbf{g}_2^0, \dots, \mathbf{g}_K^0)] \quad (26)$$

where  $\text{REP}(\cdot)$  is the number of repeated patterns.

##### 4.2.2. Quadratic Sum Squared Correlation

Another parameter to be considered is the *quadratic sum squared correlation* (SSC) [26]. To maximize the number of users accessing the system, the SSC must be minimized. Conceptually, its value is performed as follows

$$\text{SSC} = \sum_{k=1}^K \sum_{k'=k+1}^K |\mathbf{g}_k^H \mathbf{g}_{k'}|^2 \quad (27)$$

for transmission in DL. Besides, the SSC reduction makes sparsity increase [1].

##### 4.2.3. Diversity Order

System *diversity* works in accordance with detector convergence. With increasing diversity, it is possible to increase the speed of convergence of the detector. The system diversity order can be determined as [5]

$$\text{DIV} = \max \left( \sum_{n=1}^N \mathbf{g}_n^{0T} \right) \quad (28)$$

#### 4.2.4. Weighted Maximum Row Weight

The *weighted maximum row weight* (MRW) is denoted by [1]:

$$d_{\max}^w = \sum_{n=1}^N \left[ \frac{1}{2} \sum_{k=1}^K |g_{n,k}|^2 \left( \sum_{k=1}^K |g_{n,k}|^2 - 1 \right) \right] \quad (29)$$

The systematic reduction of  $d_{\max}^w$  minimizes the inter-user interference, consequently, reducing the complexity of multi-user detection on the receive side. The lower the MRW value, the better the receiver performance. To illustrate, let us consider again the pattern matrix of Eq. (19); the obtained pattern matrix for QPSK modulation with ambiguity, Fig. 2 (a) is given by

$$\begin{bmatrix} 1+i & -1-i & 0 \\ 1+i & 0 & -1-i \end{bmatrix}, \quad (30)$$

and the correspondent MRW is  $d_{\max}^w = 18$ . For the same matrix but now adopting unambiguous pattern matrix, given by

$$\begin{bmatrix} 0.7071i & -0.7071 - 0.7071i & 0 \\ 0.7071i & 0 & -0.7071 - 0.7071i \end{bmatrix}, \quad (31)$$

as in Fig. 2.(b) results in a reduced MRW value of  $d_{\max}^w = 1$ .

#### 4.2.5. Performance Evaluation Factor

A comprehensive performance indicator for the PDMA pattern matrix is the *performance evaluation factor* (PEF) [5]:

$$\Lambda = \frac{\text{SSC} \cdot d_{\max}^w}{\text{DIV}}. \quad (32)$$

It is desired to reduce the PEF values of the pattern matrix in order to improve the PDMA system performance. To illustrate, let us consider the same example of subsection 4.2.4. The pattern matrix with ambiguity in Eq. (30), with  $\text{SSC} = 8$ ,  $d_{\max}^w = 18$ ,  $\text{DIV} = 4$  resulted in  $\Lambda = 36$ . Hence, comparing with the unambiguous matrix, eq. (31), with  $\text{SSC} = 1$ ,  $d_{\max}^w = 1$ ,  $\text{DIV} = 2$  resulted in  $\Lambda = 0.5$ , indicating a huge overall PDMA system performance improvement.

### 4.3. Hybrid Detection of the Spatial Filter and SIC at the Receiver

The detector is used to recover received signals. Combining SIC with spatial diversity/beamforming generates hybrid detection schemes, which additionally could explore diversity in the power domain, in the so-called non-orthogonal multiple access (NOMA) systems, as sketched in Fig. 1. The associated hybrid detection scheme is composed by the *spatial filter* (SF) that suppresses the interference between beams caused by the beam domain multiplexing and by the SIC that removes the intra-beam interference caused by the power domain multiplexing.

First, let us consider the SF matrix given by  $\mathbf{V}_k \in \mathbb{C}^{N_R \times N}$  for the  $k$ th user and  $\mathbf{v}_{nk}$  the  $n$ th column vector of  $\mathbf{V}_k$ . In this step, SF is performed on the received signal vector to suppress inter-beam interference. The scalar resultant is represented by  $z_{nk}$ , and can be expressed as follows

$$z_{nk} = \mathbf{v}_{nk}^H \mathbf{y}_k = \underbrace{\mathbf{v}_{nk}^H \mathbf{H}_k \mathbf{b}_n t_n}_{\text{Intra-beam interference}} + \underbrace{\mathbf{v}_{nk}^H \mathbf{H}_k \sum_{\substack{i=1 \\ i \neq n}}^N \mathbf{b}_i t_i}_{\text{Inter-beam interference}} + \mathbf{v}_{nk}^H \mathbf{w}_k, \quad (33)$$

$$n = 1, 2, \dots, N \quad \text{and} \quad k = 1, 2, \dots, K,$$

where the first term on the right-hand side of the equality represents the combination of desired information and intra-beam interference, while the other two terms represent inter-beam interference and noise, respectively.

The SF matrix  $\mathbf{V}_k$  can be constructed based on the MMSE or ZF criteria. Assume that  $\mathbf{V}_k$  calculated based on MMSE criterion [27] as follows:

$$\mathbf{V}_k = \min_{\tilde{\mathbf{V}}_k} \mathbb{E} \left\{ \left\| \mathbf{t} - \tilde{\mathbf{V}}_k^H \mathbf{y}_k \right\|_2^2 \right\} \quad (34)$$

$$= (\mathbf{H}_k \mathbf{B} \mathbf{A} \mathbf{B}^H \mathbf{H}_k^H + \sigma_k^2 \mathbf{I}_{N_R})^{-1} \mathbf{H}_k \mathbf{B} \mathbf{A}, \quad k = 1, 2, \dots, K, \quad (35)$$

where  $\mathbf{A} \triangleq \mathbb{E} \{ \mathbf{t} \mathbf{t}^H \}$  with its elements being expressed as  $a_{i,j} = \sum_{k=1}^K g_{i,k} g_{j,k} \sqrt{p_{i,k} p_{j,k}}$ , and  $\mathbf{t}$  is the *superimposed signal* vector, given in Eq. (9). The expression for  $\mathbf{V}_k$  developed based on the ZF criterion [8] has the following form

$$\mathbf{V}_k = \mathbf{H}_k \mathbf{B} (\mathbf{B}^H \mathbf{H}_k^H \mathbf{H}_k \mathbf{B})^{-1}, \quad k = 1, 2, \dots, K. \quad (36)$$

Second, let the aggregate power of inter-beam interference and noise in  $z_{nk}$  normalized to be one. Normalization is applied to the received scalar signal to meet the SIC implementation condition. And let  $u_{nk}$  be the equivalent normalized channel gain between the  $k$ th user and AC which can be expressed [8] as follows

$$u_{nk} = \sqrt{\frac{|\mathbf{v}_{nk}^H \mathbf{H}_k \mathbf{b}_n|^2}{\sum_{\substack{i=1 \\ i \neq n}}^N |\mathbf{v}_{nk}^H \mathbf{H}_k \mathbf{b}_i|^2 + \sigma_k^2 \|\mathbf{v}_{nk}\|^2}}, \quad n = 1, 2, \dots, N \quad \text{and} \quad k = 1, 2, \dots, K. \quad (37)$$

Consequently, the equation in (34) can be rewritten as follows

$$z_{nk} = u_{nk} \sum_{i=1}^K g_{ni} \sqrt{p_{ni}} s_i + q_{nk}, \quad n = 1, 2, \dots, N \quad \text{and} \quad k = 1, 2, \dots, K, \quad (38)$$

where  $q_{nk}$  represents the sum of interference between beams and noise after normalization such as  $\mathbb{E} [ |q_{nk}|^2 ] = 1$ . The matrix given by  $\mathbf{U} \in \mathbb{R}^{N \times K}$  corresponds to the equivalent normalized channel gain between UG and AC where we can represent  $[\mathbf{U}]_{nk} = u_{nk}$ . Thus, the MIMO channel between

the  $k$ th user and the AC can be degraded into single-input single-output (SISO) channels after normalization [27], which meets the implementation condition of SIC.

The SIC is employed for the received normalized scalar signal to remove intra-beam interference. It is also used to detect symbols by subtracting the detected symbols from users transmitted with a higher power to facilitate the detection of users transmitted with lower power. We assume that users are sorted in ascending order of normalized channel gain, so  $u_{ni} \leq u_{nj}$  is valid if  $1 \leq i \leq j \leq K$ . As a result, the  $j$ th user can correctly detect its signal symbol despite interference from the  $i$ th user. Consequently, the signal-to-interference-to-noise ratio (SINR) of the  $k$ th user in the  $n$ th beam can be expressed [1] as follows

$$\gamma_{nk} = \begin{cases} \frac{|u_{nk}|^2 \cdot |g_{nk}|^2 \cdot p_{nk}}{1 + |u_{nk}|^2 \cdot \sum_{i=k+1}^K |g_{ni}|^2 \cdot p_{ni}}, & n = 1, 2, \dots, N \quad \text{and} \quad k = 1, 2, \dots, K-1, \\ |u_{nk}|^2 \cdot |g_{nk}|^2 \cdot p_{nk}, & n = 1, 2, \dots, N \quad \text{and} \quad k = K \end{cases} \quad (39)$$

Therefore, the *sum rate* of the ACs can be expressed as follows

$$R_{\Sigma} = \sum_{n=1}^N \sum_{k=1}^K \log_2(1 + \gamma_{nk}) \quad \left[ \frac{\text{bits}}{\text{s} \cdot \text{Hz}} \right]. \quad (40)$$

The energy efficiency is given by the ratio between the sum rate achievable by the users and the total power consumed in the system represented by:

$$\text{EE} = \frac{R_{\Sigma}}{P_{RF} + P_c} = \frac{R_{\Sigma}}{P_{RF} + N_R \cdot K \cdot N \cdot P_K + N \cdot N_T \cdot P_{BS}} \quad \left[ \frac{\text{bits}}{\text{J} \cdot \text{Hz}} \right] \quad (41)$$

where  $P_c$  is the circuitry power consumption required to signal processing, cooling, etc, at the transmitters and receivers. For simplicity, only the  $P_K$  and  $P_{BS}$  parts were explicitly considered, where the power required to operate the circuitry of each user antenna is  $P_K$ , and to operate the circuit components of each BS antenna is  $P_{BS}$ . Besides, the total RF power consumption can be written as

$$P_{RF} = \varrho \cdot \sum_{n=1}^N \sum_{k=1}^K \beta_{nk} \cdot P = \varrho \cdot P \quad [W] \quad (42)$$

where  $\varrho \geq 1$  is the RF power amplifier inefficiency, assumed the same for all that feed each transmit antenna and  $\beta_{nk}$  is the power proportion for the  $k$ th user in the  $n$ th beam. The sum of the powers between users is given by  $\sum_{n=1}^N \sum_{k=1}^K \beta_{nk} \cdot P$  where  $\sum_{n=1}^N \sum_{k=1}^K \beta_{nk} = 1$  and  $p_{nk} = \beta_{nk} \cdot P$ .

## 5. Energy Efficiency Optimization Problem

The EE optimization design in MIMO-PDMA system to obtain the optimal power allocation  $P$ , under equal data rate  $R$  (maximum rate fairness) and power proportion  $\beta_{nk}$  values for all users can

be defined by an equivalent problem as follows:

$$\begin{aligned}
& \underset{P, \beta_{nk}}{\text{maximize}} && \frac{R_S}{P_{RF} + N_R \cdot K \cdot N \cdot P_K + N \cdot N_T \cdot P_{BS}} \\
& \text{s.t.} && \text{C.1 } \sum_{n=1}^N \sum_{k=1}^K \beta_{nk} = 1 \\
& && \text{C.2 } \beta_{nk} = \frac{(2^R - 1)}{u_{n,k}^2 |g_{nk}|^{2P}} + \frac{(2^R - 1)}{|g_{nk}|^2} \sum_{i=k+1}^K |g_{ni}|^2 \beta_{ni}, \quad \text{where } \begin{cases} 1 \leq n \leq N \\ 1 \leq k \leq K - 1 \end{cases} \\
& && \text{C.3 } \beta_{nK} = \frac{(2^R - 1)}{u_{nK}^2 |g_{nK}|^{2P}} \text{ to } 1 \leq n \leq N \\
& && \text{C.4 } P > 0 \quad \text{and} \quad \beta_{nk} > 0
\end{aligned} \tag{43}$$

The objective function is not convex in  $P$  and  $\beta_{nk}$  simultaneously, despite the constraints C.1–C.4 form a convex set. The proof of convexity of constraints is similar to the that provided in [28], hence omitted herein for convenience. To solve the EE problem (43) efficiently, we slightly modify the objective function to conveniently deploy the sequential quadratic programming (SQP) technique as in [29], aiming to attain a near-optimal solution under fast convergence despite the non-convexity of the objective function.

## 6. Numerical Results

This section seeks to obtain the pattern matrix that best meets the requirements for each of the eMBB, mMTC and uRLLC services in the LSA-PDMA system. For this, we will consider the calculation of several parameters involving the standard matrices such as repeatability  $r$ , quadratic sum squared correlation SSC, diversity DIV, weighted maximum row weight  $d_{\max}^w$  and the performance evaluation factor  $\Lambda$ . Three different scenarios have been evaluated, Scenario A, B, and C, as sketched in Fig. 3. The optimization problem was solved to obtain the maximum EE under each standard PDMA matrix considering that *all users have the same rate* or *maximum fairness* criterion is selected for the LSA-MIMO PDMA system. By analyzing these data, the pattern matrix that best meets each type of use service (eMBB, mMTC and uRLLC) can be obtained.

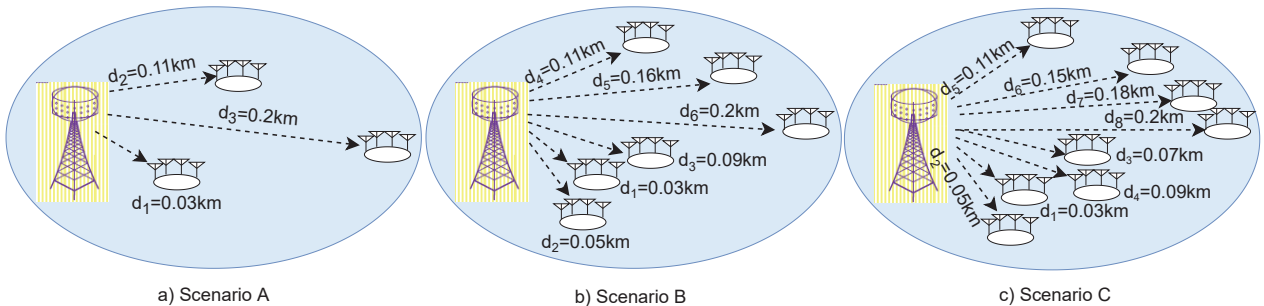


Figure 3: Scenario A, B and C (max-fairness) for the (DL) EE optimization problem, eq. (43).

Table 3: Adopted Parameters for the Simulation Scenarios.

Parameter	Value
#Users	$K \in \{3; 6; 8\}$
Path-loss exponent	$\alpha = 3.0$
Power of white noise	$\sigma_n^2 = -100$ dB
Power Budget	$P = 100$ W
# beamforming users/beamforming antennas/beamforming	$N = 2, 3$ and 4
	$K \in \{2, 3, 4, 5\}$
	$N_T = 16$
# antennas per user	$N_R = 4$
Target equal normalized user's rate	$R \in [4; 45]$ [bits/s/Hz]
Circuits Power	$P_{BS} = 1$ W/antenna (BS) $P_K = 100$ mW/antenna per user
RF power amplifier inefficiency	$\rho = 1.4$
<b>Scenarios:</b> User distribution inside the cell: 3, 6 or 8 users, 2, 3 or 4 beamforming	
Scen. A – 3 users (Fig. 3 (a))	$d_1 = 0.03, d_2 = 0.11, d_3 = 0.20$ [km]
Scen. B – 6 users (Fig. 3 (b))	$d_1 = 0.03, d_2 = 0.05, d_3 = 0.09,$ $d_4 = 0.11, d_5 = 0.16, d_6 = 0.20$ [km]
Scen. C – 8 users (Fig. 3 (c))	$d_1 = 0.03, d_2 = 0.05,$ $d_3 = 0.07, d_4 = 0.09, d_5 = 0.11,$ $d_6 = 0.15, d_7 = 0.18, d_8 = 0.20$ [km]

The tool used for the optimization process of eq. (43) was the *fmincon* function of the MatLab<sup>®</sup> software with the sequential quadratic programming (SQP) technique. The parameters values adopted in the numerical simulation are depicted in Table 3. The number of antennas in the BS that form each beamforming and for each user are defined as  $N_T = 16$  and  $N_R = 4$ , respectively. Users stand around the BS with a radius of up to 200 meters. Scenarios 1, 2 and 3 define the fixed distances from users to the BS according to the number of users by beamforming of the pattern matrix, according to Figure 3.

### 6.1. Selecting PDMA Pattern Matrices

The pattern matrices Type 1 to Type 9 in Table 4 were generated from existing matrices in [1], while the matrices Type 10 to Type 12 have been generated from [5], and Type 13 to Type 15 from [26]. Furthermore, for simplicity of analysis, others PDMA matrices found in the literature in which the values of the figure of merit  $\iota$ , SSC,  $d_{\max}^w$ , DIV,  $\Lambda$  fully coincide with those from Type 1 to Type 15 have not been included in our analyses, since despite the different pattern matrix designs, they

Table 4: PDMA design pattern matrix of different dimensions [1], [26], [5].

Dimension/Type $G_{\text{Type}}^{N \times K}$	Pattern Matrix	$\iota$	SSC	$d_{\max}^W$	DIV	$\Lambda$
$G_{\text{Type1}}^{4 \times 6}$	$\begin{bmatrix} 0.4082 & 0 & 0 & 0 & 0.4082 & 0 \\ 0 & 0.4082i & 0 & 0 & 0 & 0.4082i \\ 0 & 0 & 0.4082 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.4082i & 0 & 0 \end{bmatrix}$	150%	0.06	2	1	0.12
$G_{\text{Type2}}^{4 \times 6}$	$\begin{bmatrix} 0.408 & 0 & 0 & 0.236 & 0.236 & 0 \\ 0 & 0.408i & 0 & 0.236i & 0.236 & 0 \\ 0 & 0 & 0.408 & 0.236 & 0.236i & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.408i \end{bmatrix}$	150%	0.06	9	3	0.18
$G_{\text{Type3}}^{4 \times 8}$	$\begin{bmatrix} 0.25 & 0.25i & 0 & 0 & 0.354 & 0 & 0 & 0 \\ 0.25i & 0 & 0.25 & 0 & 0 & 0.354i & 0 & 0 \\ 0 & 0.25 & 0 & 0.25i & 0 & 0 & 0.354 & 0 \\ 0 & 0 & 0.25i & 0.25 & 0 & 0 & 0 & 0.354i \end{bmatrix}$	200%	0.08	12	2	0.48
$G_{\text{Type4}}^{4 \times 6}$	$\begin{bmatrix} 0.2911 & 0 & 0.2276 & 0.2276i & 0.2633 & 0 \\ 0 & 0 & 0.2276i & 0.2276 & 0.2633i & 0.2633 \\ 0 & 0 & 0.2276 & 0.2276i & 0.2633 & 0.2633i \\ 0 & 0.2911i & 0.2276i & 0.2276 & 0 & 0.2633 \end{bmatrix}$	150%	0.10	24	4	0.60
$G_{\text{Type5}}^{4 \times 6}$	$\begin{bmatrix} 0.2692 & 0 & 0.1860 + 0.1867i & 0.2104i & 0.1999 + 0.2005i & 0 \\ 0 & 0 & 0.2104 & 0.2124 + 0.1603i & 0.2434 & 0.2434i \\ 0 & 0 & 0.1860 + 0.1867i & 0.2104 & 0.2434i & 0.2434 \\ 0 & 0.2692i & 0.2104i & 0.2124 + 0.1603i & 0 & 0.1999 + 0.2005i \end{bmatrix}$	150%	0.13	24	4	0.78
$G_{\text{Type6}}^{4 \times 8}$	$\begin{bmatrix} 0.2911 & 0.2633 & 0.2276 & 0.2633 & 0 & 0 & 0 & 0 \\ 0.2276i & 0.2276i & 0 & 0 & 0.2276i & 0.2633i & 0 & 0 \\ 0.2276 & 0 & 0.2633i & 0 & 0.2633 & 0 & 0.2276 & 0 \\ 0 & 0 & 0 & 0.2276 & 0 & 0.2276i & 0.2633 & 0.2911i \end{bmatrix}$	200%	0.11	24	3	0.88
$G_{\text{Type7}}^{4 \times 6}$	$\begin{bmatrix} 0.4083 & 0 & 0.2042 & 0.2042i & 0.2356 & 0 \\ 0 & 0 & 0.2042i & 0.2042 & 0.2356i & 0.2356 \\ 0 & 0 & 0.2042 & 0.2042i & 0.2356 & 0.2356i \\ 0 & 0.4083i & 0.2042i & 0.2042 & 0 & 0.2356 \end{bmatrix}$	150%	0.09	24	4	0.54
$G_{\text{Type8}}^{4 \times 6}$	$\begin{bmatrix} 0.4082 & 0 & 0.1021 + 0.1768i & 0.1445 + 0.1445i & 0.1180 + 0.2041i & 0 \\ 0 & 0 & 0.1021 + 0.1768i & 0.2041i & 0.2356 & 0.2356 \\ 0 & 0 & 0.2041 & 0.1445 + 0.1445i & 0.2356 & 0.2356 \\ 0 & 0.4082i & 0.1021 + 0.1768i & 0.2041 & 0 & 0.1180 + 0.2041i \end{bmatrix}$	150%	0.12	24	4	0.72
$G_{\text{Type9}}^{4 \times 8}$	$\begin{bmatrix} 0.204 & 0.25 & 0.25i & 0.25 & 0 & 0 & 0 & 0 \\ 0.204i & 0.25i & 0 & 0 & 0.25i & 0.25 & 0 & 0 \\ 0.204 & 0 & 0.25 & 0 & 0.25 & 0 & 0.25i & 0 \\ 0 & 0 & 0 & 0.25i & 0 & 0.25i & 0.25 & 0.3536 \end{bmatrix}$	200%	0.10	24	3	0.80
$G_{\text{Type10}}^{2 \times 3}$	$\begin{bmatrix} -0.4082 & -0.5774i & 0 \\ -0.4082i & 0 & -0.5774 \end{bmatrix}$	150%	0.11	2	2	0.11
$G_{\text{Type11}}^{3 \times 6}$	$\begin{bmatrix} 0.2357 & 0 & 0 & 0 & 0.4082i & -0.4082 \\ 0.2357 & 0.2887 & 0 & 0.4082 & 0 & 0 \\ -0.2357 & 0.2887 & 0.4082 & 0 & 0 & 0 \end{bmatrix}$	200%	0.09	9	3	0.27
$G_{\text{Type12}}^{4 \times 8}$	$\begin{bmatrix} 0.1768i & -0.2041i & 0 & -0.25 & -0.25i & 0 & 0 & -0.3536i \\ -0.1768i & 0 & -0.25i & 0 & 0 & 0 & -0.3536 & 0 \\ -0.1768 & 0.2041i & 0.25 & 0 & -0.25 & 0 & 0 & 0 \\ -0.1768 & 0.2041 & 0 & 0.25i & 0 & 0.3536 & 0 & 0 \end{bmatrix}$	200%	0.08	25	4	0.50
$G_{\text{Type13}}^{4 \times 8}$	$\begin{bmatrix} 0.3536 & 0 & 0 & 0.204 & -0.1022 + 0.1768i & -0.1022 - 0.1768i & -0.1022 - 0.1768i & 0 \\ 0 & 0.3536 & 0 & 0.204 & 0.204 & 0.204 & -0.204 & 0 \\ 0 & 0 & 0.3536 & 0.204 & 0.204 & 0.204 & 0.204 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.3536 \end{bmatrix}$	200%	0.08	30	3	0.80
$G_{\text{Type14}}^{4 \times 6}$	$\begin{bmatrix} 0.4082 & 0 & -0.1021 + 0.1768i & -0.1768 - 0.1021i & -0.1445 + 0.1445i & -0.1768 + 0.1021i \\ 0 & 0.4082 & -0.1021 - 0.1768i & 0.1021 + 0.1768i & -0.2041i & 0.1021 - 0.1768i \\ 0 & 0 & 0.2041 & -0.2041i & 0.1445 + 0.1445i & 0.2041i \\ 0 & 0 & -0.1021 + 0.1768i & -0.1021 + 0.1768i & -0.2041 & -0.1021 - 0.1768i \end{bmatrix}$	150%	0.13	32	4	1.04
$G_{\text{Type15}}^{4 \times 8}$	$\begin{bmatrix} 0.2041 & 0 & 0.2041 & 0.25 & 0.25 & 0 & 0 & 0.3537 \\ 0 & 0.2041 & 0.2041 & 0.25 & 0 & 0.25 & 0 & 0 \\ 0.2041 & 0.2041 & 0.2041 & 0 & 0 & 0.25 & 0.25 & 0 \\ 0.2041 & 0.2041 & 0 & 0 & 0.25 & 0 & 0.25 & 0 \end{bmatrix}$	200%	0.14	32	3	1.49

produce the same system performance. Originally, the Type 7 to Type 15 pattern matrices in Table 4 meet the conditions for operating in UL; however, in this work all pattern matrices were generated to satisfy the condition of *total transmit power available* at the BS, Eq. (23). This condition has been guaranteed in all the three scenarios A, B, and C satisfying system operating in DL.

The operating conditions of PDMA matrices on the DL are different from the those on the UL. Indeed, for the DL scenarios, the existing PDMA matrix designs in the literature are notably scarce yet. Hence, to meet this need, the pattern matrices designed originally to operate in the UL have been adapted herein to operate in the DL direction, by applying the condition expressed in Eq. (23). This adaptation was made simply normalizing the matrix elements left side value of Eq. (23).

The graphs in Fig. 4 reveal the behavior of the optimal energy efficiency  $EE^*$  with the increase of user rate  $R$ . The  $EE^*$  points obtained in such graphs are the result of the fairness EE solution of the problem in (43). Notice that in these curves the results of maximum  $EE^*$  under equal user rate (fairness) condition, shown in Fig. 4 and summarized in Fig. 5, present different resource efficiency, *i.e.*  $RE = EE^* \times R$  performance.

In the next subsections, the pattern matrix that meets the requirements of the respective 5G services described as eMBB, mMTC and uRLLC are analyzed.

## 6.2. PDMA Pattern Matrix Analysis for eMBB

In Table 5 we represent the optimum resource efficiency (RE) as:

$$RE^* = EE_{\max}^* \cdot R \quad \left[ \frac{\text{bits}^2}{\text{watts}} \right] \quad (44)$$

to compose the values of  $EE_{\max}$  and  $R$ . Notice that  $RE^*$  is attained for the maximum of  $EE^*$ , namely  $EE_{\max}^*$ , which comes from the result of  $EE^*$  solving (43) for each value of  $R$  with a fixed number of users  $K$  according to each pattern matrix; the  $EE^*$  behavior is revealed in Fig. 4. Hence, to use PDMA in eMBB services it is necessary to maximize the transmission rate. In this work, we optimized the LSA-MIMO PDMA system to operate, for all the pattern matrices in Scenario A, B and C under maximum EE at the condition of the same rate between users or *maximum fairness*.

The pattern matrices that presented the highest bit rate for the maximum EE are summarized in Table 5 and directly obtained from Fig. 4.(a) and Fig. 5. **Scenario A** (Fig. 3) is adopted in the evaluation of pattern matrix Type 10; **Scenario B** for evaluation of Type 1, Type 2, Type 4, Type 5, Type 7, Type 8, Type 11 and Type 14; and **Scenario C** for evaluation of Type 3, Type 6, Type 9, Type 12, Type 13 and Type 15 pattern matrices.

Although the pattern matrices Type 1 and Type 5 have resulted in higher maximum EE values, the pattern matrix that resulted in the highest user rate for the maximum EE was the Type 1 matrix, despite the  $RE^*$  being smaller. The pattern matrix Type 5 is not preferred for the eMBB service because it has the parameter  $d_{\max}^W = 24$ , therefore it results in high complexity, as inferred from Table 4, and summarized in Table 5. Excluding Type 5, we have to decide between Type 1 and

Table 5: Parameters to be analyzed for the eMBB services.

PDMA Type	$R$ $\left[\frac{\text{bits}}{\text{s}\cdot\text{Hz}}\right]$	$EE_{\max}$ $\left[\frac{\text{bits}}{\text{W}}\right]$	$RE^*$ $\left[\frac{\text{bits}^2}{\text{W}}\right]$	#Users	$\iota$	SSC	$d_{\max}^W$	DIV	$\Lambda$
<b>1<sup>†</sup></b>	34	0.92	31.28	6	150%	0.06	2	1	0.12
5	22	1.94	42.68	6	150%	0.13	24	4	0.78
10	21	2.52	52.92	3	150%	0.11	2	2	0.11

<sup>†</sup> Pattern matrix that best meets the requirements of eMBB services.

Type 10 matrices. Type 1 has a much higher bit rate for maximum EE than Type 10. Additionally, parameter SSC of pattern matrix Type 1 is 0.06 and for Type 10 is SSC = 0.11, which results the highest sparsity attained by the pattern matrix Type 1. Moreover, the Type 1 matrix can serve up to 6 users while the Type 10 matrix serves up to 3 users, increasing advantages to the Type 1 matrix. Hence, we concluded that matrix Type 1 is the pattern matrix that best suits the eMBB service, as highlighted in Table 5.

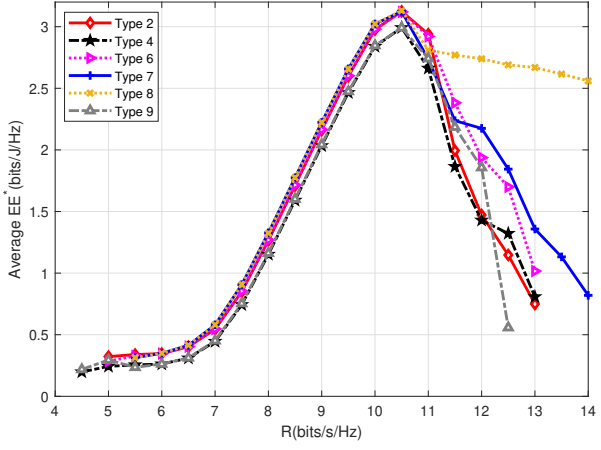
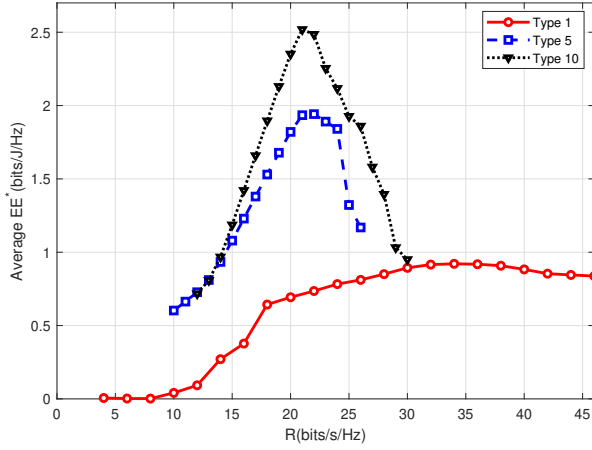
### 6.3. PDMA Pattern Matrix Analysis for mMTC

To evaluate the pattern matrices for mMTC services, a high OF and high matrix dimension must be considered. It is clear that the higher the OF the better the mMTC service meet its demand. The high dimension is essential for the mMTC service because they need to connect with a large number of endpoints while accept high latency. Analyzing the pattern matrix candidates one can select Type 3, 6, and 9 matrices as the highest OFs and dimension with  $\iota = 200\%$  and 8 users according to the Table 4 and, for convenience, also presented in Table 6. Indeed, among them, Type 3 is the pattern matrix with the best EE-SE trade-off, resulting in the highest bit-rate  $\times$  max-EE product (with maximum fairness), *i.e.*, maximum RE according Table 6 and Fig. 5. Type 3 pattern matrix resulted in the highest maximum EE of 3.51 [bits/W]. When considering the composition between  $EE^*$  and  $R$  the correspondent  $RE^*$  for the Type 3 matrix has resulted in 38.61 [bits<sup>2</sup>/W], the highest value among the results of Table 6. One can concluded that the pattern matrix Type 3 is the best among the PDMA pattern matrix candidates from Table 4 for the mMTC service.

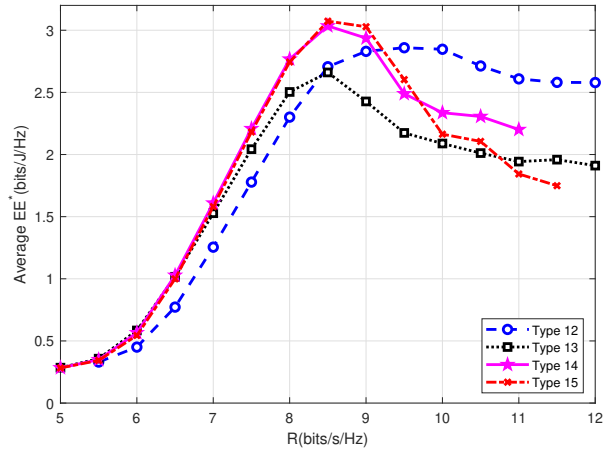
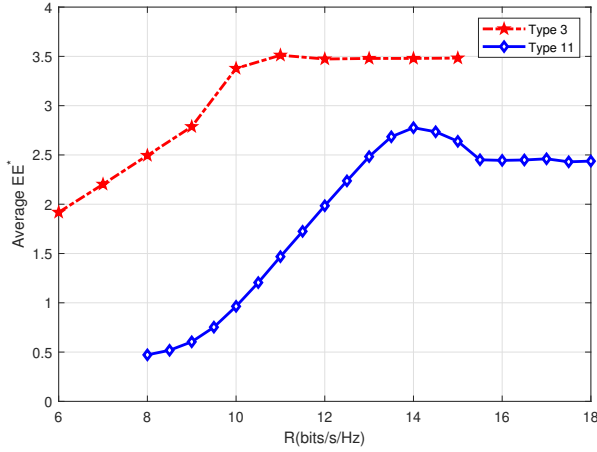
Table 6: Parameters to be analyzed for the mMTC services.

PDMA Type	$R$ $\left[\frac{\text{bits}}{\text{s}\cdot\text{Hz}}\right]$	$EE_{\max}$ $\left[\frac{\text{bits}}{\text{W}}\right]$	$RE^*$ $\left[\frac{\text{bits}^2}{\text{W}}\right]$	#Users	$\iota$	SSC	$d_{\max}^W$	DIV	$\Lambda$
<b>3<sup>†</sup></b>	11.0	3.51	38.61	8	200%	0.08	12	2	0.48
6	10.5	3.12	32.76	8	200%	0.08	12	2	0.48
9	10.5	2.99	31.40	8	200%	0.08	12	2	0.48

<sup>†</sup> Pattern matrix that best meets the requirements of mMTC services.



a) Type 1, 5 ( $K = 6$  users) and Type 10 ( $K = 3$  users)    b) Type 2, 4, 7, 8 ( $K = 6$  users), and Type 6, 9 ( $K = 8$  users).



c) Type 3 ( $K = 8$  users) and Type 11 ( $K = 6$  users)    d) Type 12, 13, 15 ( $K = 8$  users), and Type 14 ( $K = 6$  users).

Figure 4:  $EE^* \times R$  for  $\alpha = 3$ . Optimal EE values were obtained by solving Eq. (43).

#### 6.4. PDMA Pattern Matrix Analysis for uRLLC

For uRLLC services, a pattern matrix PDMA should minimize the interference while maintains sparsity. This is achieved with: a) low OF; b) minimizing sum squared correlation (SSC); c) minimizing  $d_{\max}^W$  complexity; d) consequently reducing performance evaluation factor (PEF). The pattern matrices with the lowest OF are Type 1, Type 2, Type 4, Type 7 and Type 10, as shown in Table 4 and for convenience in Table 7. Among these matrices the ones with the smallest PEF are the standard matrices Type 1 and Type 10. Furthermore, analysing the curves obtained after the optimization process of maximum EE and maximum fairness considering this pattern matrix, one can obtain the  $EE_{\max} = 2.52$  [bits/W] for a bit rate of 21 [bits/s/Hz] for Type 10 and just 0.92 [bits/W] and 34 [bits/s/Hz] for Type 1, as shown in Fig. 4 (a) and summarized in the Figure 5. Although the pattern matrix Type 1 has a higher rate when composing the  $EE^*$  and  $R$ , this matrix results in

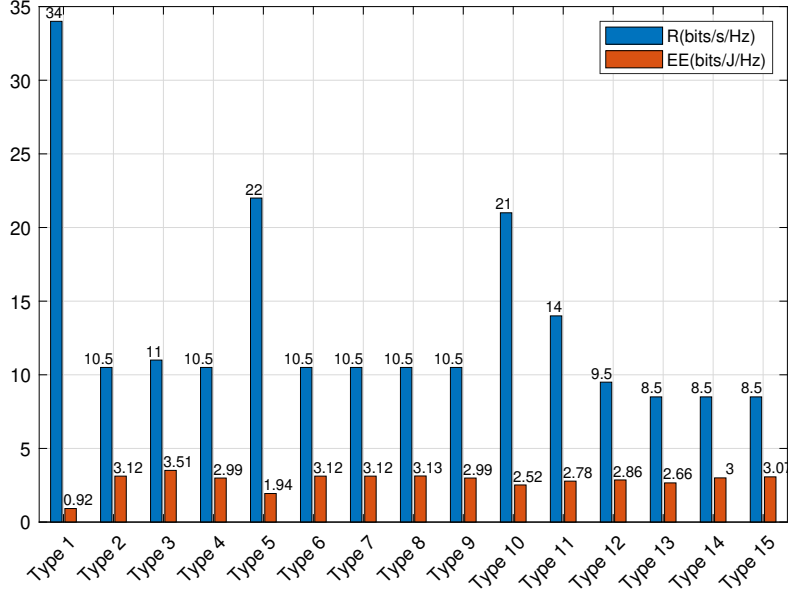


Figure 5: Attained values of maximum  $EE^*$  and the respective equal rate  $R$  obtained by solving optimization problem (43) under a fixed number of users  $K$  according to each pattern matrix. Scenario A was adopted to evaluate the pattern matrix Type 10; Scenario B in Type 1, 2, 4, 5, 7, 8, 11 and 14; Scenario C with the Type 3, 6, 9, 12, 13 and Type 15.

Table 7: Parameters to be analyzed for the uRLLC services.

PDMA Type	$R$ [ $\frac{\text{bits}}{\text{s}\cdot\text{Hz}}$ ]	$EE_{\max}$ [ $\frac{\text{bits}}{\text{W}}$ ]	$RE^*$ [ $\frac{\text{bits}^2}{\text{W}}$ ]	#Users	$\iota$	SSC	$d_{\max}^W$	DIV	$\Lambda$
1	34.0	0.92	31.28	6	150%	0.06	2	1	0.12
2	10.5	3.12	32.76	6	150%	0.06	9	3	0.18
4	10.5	2.99	31.40	6	150%	0.10	24	4	0.60
7	10.5	3.12	32.76	6	150%	0.09	24	4	0.54
<b>10<sup>†</sup></b>	21.0	2.52	52.92	3	150%	0.11	2	2	0.11

<sup>†</sup> Pattern matrix that best meets the requirements of uRLLC services.

31.28 while the pattern matrix Type 10 results in 52.92, so in this regard the pattern matrix Type 10 is much higher. Which represents a high rate and EE for this pattern matrix.

Considering that all pattern matrices in Table 7 have the same OF, the one that simultaneously guarantee the low latency (smallest dimension), and guarantee the best performance (lowest PEF) is the Type 10 pattern matrix. The PEF for pattern matrix Type 1 is 0.12 while for pattern matrix Type 10 results in 0.11 which is smaller mainly because it has greater diversity, Table 7. In conclusion the pattern matrix that best meets the uRLLC service requirements is the Type 10.

### 6.5. PDMA Pattern Matrix Analysis for Hybrid Transmission with eMBB and uRLLC Services

For eMBB services it requires high baud rate, while for the uRLLC service it is highly constrained in terms of delay and reliability. According to Tables 5 and 8 among the pattern matrix candidates

matching such requirements is just Type 10. This is a matrix that fully supports the uRLLC service but partially supports the eMBB service. This is because the rate for maximum EE\* for the pattern matrix Type 10 is 21 [bits/s/Hz], while the pattern matrix selected for the eMBB service was Type 1 with maximum EE\* of 34 [bits/s/Hz]. On the other hand, the Type 10 matrix has a better performance, because its PEF value is smaller. The rate of 21 [bits/s/Hz] for the pattern matrix Type 10 is still high compared to the other options in Table 4 and Fig. 5. This rate is the third highest among the rates obtained for the maximum EE\* considering the all the fifteen pattern matrices. In conclusion, the Type 10 pattern matrix should be selected to work in a hybrid dual eMBB and uRLLC use mode.

Table 8: Parameters to be analyzed for the hibrid eMBB and uRLLC services.

PDMA Type	$R$ $\left[\frac{\text{bits}}{\text{s}\cdot\text{Hz}}\right]$	$EE_{\max}$ $\left[\frac{\text{bits}}{W}\right]$	$RE^*$ $\left[\frac{\text{bits}^2}{W}\right]$	#Users	$\iota$	SSC	$d_{\max}^W$	DIV	$\Lambda$
<b>10<sup>†</sup></b>	21.0	2.52	52.92	3	150%	0.11	2	2	0.11

<sup>†</sup> Pattern matrix that best meets the requirements of hibrid eMBB and uRLLC services.

## 7. Conclusions

In this work we analyze the standard matrices that best meet the requirements of 5G services such as eMBB, mMTC and uRLLC. In addition to the different parameters of the standard matrices, we also analyzed the power distribution among users which results in the equal user rate for all users, attaining maximum fairness and maximum EE, simultaneously. After profiling the standard matrices, we analyzed and defined each matrix that meets each user mode requirements from eMBB, mMTC and uRLLC services.

For the eMBB service one of the essential requirements is the high user rate, while for the mMTC service it is paramount serve a high number of machine-type devices; finally, for the uRLLC use mode, the system should combine high reliability and low latency. Hence, the numerical results reveal that it is possible to find the best pattern matrix that meets each of the requirements of each of the services separately or jointly; From the fifteen PDMA pattern matrices evaluated, the Type 1 is suitable for eMBB, while Type 3 for mMTC and Type 10 for uRLLC. For hybrid transmission serving simultaneously eMBB and uRLLC users, the pattern matrix that best meets both use modes is Type 10.

## References

- [1] J. Sun, C. Wang, J. Zeng, X. Su, T. Lv, Design of pdma pattern matrix in 5g scenarios, in: 2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring), 2020, pp. 1–6. doi:10.1109/VTC2020-Spring48590.2020.9129446.

- [2] Y. Tao, L. Liu, S. Liu, Z. Zhang, A survey: Several technologies of non-orthogonal transmission for 5g, *China Communications* 12 (10) (2015) 1–15. doi:10.1109/CC.2015.7315054.
- [3] X. Dai, S. Chen, S. Sun, S. Kang, Y. Wang, Z. Shen, J. Xu, Successive interference cancelation amenable multiple access (sama) for future wireless communications, in: 2014 IEEE International Conference on Communication Systems, 2014, pp. 222–226. doi:10.1109/ICCS.2014.7024798.
- [4] J. Zeng, B. Li, X. Su, L. Rong, R. Xing, Pattern division multiple access (pdma) for cellular future radio access, in: 2015 International Conference on Wireless Communications Signal Processing (WCSP), 2015, pp. 1–5. doi:10.1109/WCSP.2015.7341229.
- [5] S. Li, B. Su, Y. Sun, S. Sun, L. Jin, Uplink grant-free pattern division multiple access transmission scheme by exploiting poly complementary sequence, *Transactions on Emerging Telecommunications Technologies* n/a (n/a) (2021) 1–18. doi:https://doi.org/10.1002/ett.4416.
- [6] S. Chen, B. Ren, Q. Gao, S. Kang, S. Sun, K. Niu, Pattern division multiple access—a novel non orthogonal multiple access for fifth-generation radio networks, *IEEE Transactions on Vehicular Technology* 66 (4) (2017) 3185–3196. doi:10.1109/TVT.2016.2596438.
- [7] P. Li, Y. Jiang, S. Kang, F. Zheng, X. You, Pattern division multiple access with large-scale antenna array, in: 2017 IEEE 85th Vehicular Technology Conference (VTC Spring), 2017, pp. 1–6. doi:10.1109/VTCSpring.2017.8108482.
- [8] Y. Jiang, P. Li, Z. Ding, F.-C. Zheng, M. Ma, X. You, Joint transmitter and receiver design for pattern division multiple access, *IEEE Transactions on Mobile Computing* 18 (4) (2019) 885–895. doi:10.1109/TMC.2018.2845364.
- [9] Y. Zeng, R. Zhang, Millimeter wave mimo with lens antenna array: A new path division multiplexing paradigm, *IEEE Transactions on Communications* 64 (4) (2016) 1557–1571. doi:10.1109/TCOMM.2016.2533490.
- [10] B. Wang, F. Gao, S. Jin, H. Lin, G. Y. Li, Spatial- and frequency-wideband effects in millimeter-wave massive mimo systems, *IEEE Transactions on Signal Processing* 66 (13) (2018) 3393–3406. doi:10.1109/TSP.2018.2831628.
- [11] Z. Ji, Y. Ji, B. Wang, F. Gao, H. Wang, C. Zhang, A new uplink channel estimation architecture for massive mimo systems with pdma, in: 2019 IEEE 13th International Conference on ASIC (ASICON), 2019, pp. 1–4. doi:10.1109/ASICON47005.2019.8983624.
- [12] S. Tang, Z. Ma, M. Xiao, L. Hao, Hybrid transceiver design for beamspace mimo-noma in code-domain for mmwave communication using lens antenna array, *IEEE Journal on Selected Areas in Communications* 38 (9) (2020) 2118–2127. doi:10.1109/JSAC.2020.3000885.

- [13] M. Li, S. Zhang, F. Gao, P. Fan, O. A. Dobre, A new path division multiple access for the massive mimo-otfs networks, *IEEE Journal on Selected Areas in Communications* 39 (4) (2021) 903–918. doi:10.1109/JSAC.2020.3018826.
- [14] R. Deng, B. Di, H. Zhang, L. Song, Hdma: Holographic-pattern division multiple access, *IEEE Journal on Selected Areas in Communications* 40 (4) (2022) 1317–1332. doi:10.1109/JSAC.2022.3143221.
- [15] J. Zeng, X. Fan, Y. Zhong, M. Peng, T. Lv, T. Jiang, Downlink power allocation optimization in pattern division multiple access, *IEEE Access* 9 (2021) 14649–14659. doi:10.1109/ACCESS.2021.3049493.
- [16] L. Li, H. Ma, H. Ren, Q. Cheng, D. Wang, T. Bai, Z. Han, Learning-aided resource allocation for pattern division multiple access-based swipt systems, *IEEE Wireless Communications Letters* 10 (1) (2021) 131–135. doi:10.1109/LWC.2020.3023108.
- [17] M. Vaezi, Z. Ding, H. V. Poor, *Multiple Access Techniques for 5G Wireless Networks and Beyond*, Springer International Publishing AG, 2019.
- [18] W.-E. Chen, X.-Y. Fan, L.-X. Chen, A cnn-based packet classification of embb, mmhc and urllc applications for 5g, in: *2019 International Conference on Intelligent Computing and its Emerging Applications (ICEA)*, 2019, pp. 140–145. doi:10.1109/ICEA.2019.8858305.
- [19] P. Popovski, K. F. Trillingsgaard, O. Simeone, G. Durisi, 5g wireless network slicing for embb, urllc, and mmhc: A communication-theoretic view, *IEEE Access* 6 (2018) 55765–55779. doi:10.1109/ACCESS.2018.2872781.
- [20] D. Wang, Y. Zhang, H. Wei, X. You, J. Wang, An overview of transmission theory and techniques of large-scale antenna systems for 5g wireless communications, in: *Science CHINA Information Sciences*, Vol. 59, 2016, pp. 1–23. doi:10.1007/s11432-016-0278-5.
- [21] H. L. dos Santos, J. C. Marinello, C. M. Panazio, T. Abrão, Machine learning-aided pilot and power allocation in multi-cellular massive mimo networks, *Physical Communication* 52 (2022) 101646. doi:https://doi.org/10.1016/j.phycom.2022.101646.  
URL <https://www.sciencedirect.com/science/article/pii/S1874490722000295>
- [22] T. A. B. Alves, T. Abrão, Massive mimo and noma bits-per-antenna efficiency under power allocation policies, *Physical Communication* 51 (2022) 101588. doi:https://doi.org/10.1016/j.phycom.2021.101588.  
URL <https://www.sciencedirect.com/science/article/pii/S1874490721002858>
- [23] J. H. I. de Souza, A. Amiri, T. Abrão, E. de Carvalho, P. Popovski, Quasi-distributed antenna selection for spectral efficiency maximization in subarray switching xl-mimo systems, *IEEE Transactions on Vehicular Technology* 70 (7) (2021) 6713–6725. doi:10.1109/TVT.2021.3081462.

- [24] G. A. Ubiali, J. C. Marinello, T. Abrão, Energy-efficient flexible and fixed antenna selection methods for xl-mimo systems, *AEU - International Journal of Electronics and Communications* 130 (2021) 153568. doi:<https://doi.org/10.1016/j.aeue.2020.153568>.  
URL <https://www.sciencedirect.com/science/article/pii/S1434841120327722>
- [25] D. Kong, J. Zeng, X. Su, L. Rong, X. Xu, Multiuser detection algorithm for pdma uplink system based on sic and mmse, in: 2016 IEEE/CIC International Conference on Communications in China (ICCC), 2016, pp. 1–5. doi:[10.1109/ICCChina.2016.7636865](https://doi.org/10.1109/ICCChina.2016.7636865).
- [26] B. Ren, Y. Wang, X. Dai, K. Niu, W. Tang, Pattern matrix design of pdma for 5g ul applications, *China Communications* 13 (Supplement2) (2016) 159–173. doi:[10.1109/CC.2016.7833470](https://doi.org/10.1109/CC.2016.7833470).
- [27] K. Higuchi, Y. Kishiyama, Non-orthogonal access with random beamforming and intra-beam sic for cellular mimo downlink, in: 2013 IEEE 78th Vehicular Technology Conference (VTC Fall), 2013, pp. 1–5. doi:[10.1109/VTCFall.2013.6692307](https://doi.org/10.1109/VTCFall.2013.6692307).
- [28] J. L. Jacob, C. M. Panazio, T. Abrão, Energy and spectral efficiencies trade-off in mimo-noma system under user-rate fairness and variable user per cluster, *Physical Communication* 47 (2021) 1–14. doi:[10.1016/j.phycom.2021.101348](https://doi.org/10.1016/j.phycom.2021.101348).
- [29] J. L. Jacob, C. A. Pendeza Martinez, A. L. Machado Martinez, T. Abrão, Non-linear biobjective ee-se optimization for noma-mimo systems under user-rate fairness and variable number of users per cluster, *AEU - International Journal of Electronics and Communications* 138 (2021) 153870. doi:<https://doi.org/10.1016/j.aeue.2021.153870>.