



UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS SOBRAL
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA E DE
COMPUTAÇÃO (PPGEEC)

MICHEL BERNARDO DE PAIVA

ALOCAÇÃO DE RECURSOS PARA REDES DE COMUNICAÇÕES NÃO
TERRESTRES EMPREGANDO RSMA

SOBRAL

2025

MICHEL BERNARDO DE PAIVA

ALOCAÇÃO DE RECURSOS PARA REDES DE COMUNICAÇÕES NÃO TERRESTRES
EMPREGANDO RSMA

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e de Computação (PPGEEC) da Universidade Federal do Ceará, Campus de Sobral, como requisito parcial à obtenção do título de mestre em Engenharia Elétrica e de Computação. Área de Concentração: Sistemas de Informação.

Orientador: Prof. Dr. Francisco Rafael Marques Lima

SOBRAL

2025

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

- P169a Paiva, Michel Bernardo de.
Alocação de Recursos para Redes de Comunicações Não Terrestres Empregando RSMA / Michel Bernardo de Paiva. – 2025.
96 f. : il. color.
- Dissertação (mestrado) – Universidade Federal do Ceará, Campus de Sobral, Programa de Pós-Graduação em Engenharia Elétrica e de Computação, Sobral, 2025.
Orientação: Prof. Dr. Francisco Rafael Marques Lima.
1. Redes Não Terrestres. 2. Satélites LEO. 3. RSMA. 4. TDD. 5. Alocação de Potência. I. Título.
CDD 621.3
-

MICHEL BERNARDO DE PAIVA

ALOCAÇÃO DE RECURSOS PARA REDES DE COMUNICAÇÕES NÃO TERRESTRES
EMPREGANDO RSMA

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e de Computação (PPGEEC) da Universidade Federal do Ceará, Campus de Sobral, como requisito parcial à obtenção do título de mestre em Engenharia Elétrica e de Computação. Área de Concentração: Sistemas de Informação.

Aprovado em: 25 de Agosto de 2025.

BANCA EXAMINADORA

Prof. Dr. Francisco Rafael Marques
Lima (Orientador)
Universidade Federal do Ceará (UFC)

Profa. Dra. Brena Kelly Sousa Lima
Universidade Lusófona de Lisboa (ULUSOFONA)

Prof. Dr. Weskley Vinicius Fernandes Mauricio
Universidade Federal do Cariri (UFCA)

A Jesus Cristo, meu salvador, e à minha família, em especial à minha esposa, Israelly Alves Magalhães Bernardo; à minha mãe, Renata do Nascimento Bernardo; à minha avó, Maria Celeste do Nascimento Bernardo; e ao meu irmão, Carlos Renan Bernardo Gomes, que foram minhas maiores motivações para superar cada desafio deste mestrado.

AGRADECIMENTOS

A Jesus Cristo, por sua multiforme graça, que se manifestou em recursos, pessoas e motivação para percorrer todas as etapas deste programa de mestrado.

À minha esposa, Israelly, por ser minha base nos momentos mais desafiadores deste mestrado. Sou grato por sua compreensão, apoio, carinho e suas demais expressões de amor, que são como um bálsamo em meio a momentos de caos e incerteza.

À minha mãe, à minha avó e ao meu irmão, por terem sido minha base durante toda a minha vida e, em especial, nos momentos iniciais do mestrado. Se não fosse por vocês, eu não teria chegado aqui.

Ao Prof. Rafael, por todos os anos de orientação, desde a iniciação científica, ainda na graduação, até os mínimos detalhes desta dissertação. Sou grato por cada aula, reunião e aprendizado compartilhado, bem como pela oportunidade de aprender com um profissional de referência.

Ao CPQD - Centro de Pesquisa e Desenvolvimento em Telecomunicações, por seu programa de apoio à pós-graduação, que me permitiu conciliar as atividades profissionais e acadêmicas ao longo deste mestrado.

A todos os demais familiares, amigos e colegas que me incentivaram e cooperaram para a minha trajetória neste curso de mestrado.

Aos demais professores que fizeram parte da minha formação durante este curso de mestrado.

Ao Programa de Pós-Graduação em Engenharia Elétrica e de Computação (PP-GEEC), pela oportunidade de cursar o mestrado.

À Universidade Federal do Ceará, em especial ao Campus de Sobral, pelo investimento e pela dedicação de recursos na educação superior pública brasileira.

“O problema fundamental da comunicação é o de reproduzir em um ponto, exata ou aproximadamente, uma mensagem selecionada em outro ponto.”

(Claude Shannon.)

RESUMO

A crescente demanda por conectividade global, impulsionada por aplicações emergentes como internet das coisas, do inglês *Internet of Things*, veículos autônomos e realidade imersiva, evidencia não apenas a limitação das redes puramente terrestres, mas também a necessidade de técnicas de acesso múltiplo mais robustas e flexíveis. Nesse contexto, as redes não terrestres, do inglês *Non-Terrestrial Network* (NTN), especialmente baseadas em satélites de baixa órbita terrestre, do inglês *Low Earth Orbit* (LEO), surgem como solução promissora para ampliar a cobertura e complementar as infraestruturas terrestres. Entretanto, a operação dessas redes impõe desafios complexos, incluindo grandes atrasos de propagação, perdas de percurso e efeitos Doppler severos, os quais demandam estratégias avançadas de gerenciamento de interferência e alocação de recursos. Esta dissertação explora o uso do acesso múltiplo por divisão de taxa, do inglês *Rate-Splitting Multiple Access* (RSMA), como alternativa para aumentar a eficiência espectral em cenários NTN empregando satélites LEO. Inicialmente, é proposta uma solução para alocação de potência entre mensagens comuns e privadas em RSMA, com o objetivo de maximizar a soma das taxas dos usuários. Para isso, são empregadas soluções utilizando expressões de forma fechada, adaptadas para o cenário NTN, de baixa complexidade computacional. No segundo momento, é apresentado um problema de escalonamento de usuários em sistemas de duplexação por divisão no tempo, do inglês *Time Division Duplexing* (TDD), NTN, integrando o mecanismo de avanço do tempo, do inglês *Timing Advance* (TA), para minimizar o tempo total de transmissão, respeitando restrições individuais de qualidade de serviço. As soluções propostas demonstram, por meio de simulações, ganhos expressivos na taxa total de dados e maior robustez frente a imperfeições de conhecimento de canal no transmissor, do inglês *Channel State Information at the Transmitter* (CSIT), quando comparadas a abordagens convencionais como o acesso múltiplo ortogonal, do inglês *Orthogonal Multiple Access* (OMA) e o acesso múltiplo por divisão espacial, do inglês *Spatial Division Multiple Access* (SDMA). Os resultados obtidos confirmam que a adoção do RSMA, nos cenários analisados, possibilita ganhos significativos em eficiência espectral, representando um avanço relevante para o desenvolvimento de futuras redes NTN integradas à visão dos sistemas móveis de Sexta geração, do inglês *6th Generation* (6G).

Palavras-chave: redes não terrestres; satélites LEO; RSMA; TDD; alocação de potência.

ABSTRACT

The growing demand for global connectivity, driven by emerging applications such as the Internet of Things (IoT), autonomous vehicles, and immersive reality, highlights not only the limitations of purely terrestrial networks but also the need for more robust and flexible multiple access techniques. In this context, Non-Terrestrial Networks (NTN), especially those based on Low Earth Orbit (LEO) satellites, emerge as a promising solution to extend coverage and complement terrestrial infrastructures. However, the operation of these networks presents complex challenges, including large propagation delays, path losses, and severe Doppler effects, which require advanced interference management and resource allocation strategies. This dissertation explores the use of Rate-Splitting Multiple Access (RSMA) as an alternative to increase spectral efficiency in NTN scenarios employing LEO satellites. Initially, a solution is proposed for power allocation between common and private messages in RSMA, aiming to maximize the users' sum rate. For this purpose, solutions using closed-form expressions are employed, adapted to the NTN scenario, with low computational complexity. Subsequently, a user scheduling problem is presented in Time Division Duplexing (TDD) NTN systems, integrating the Timing Advance (TA) mechanism to minimize the total transmission time while satisfying individual quality of service constraints. The proposed solutions demonstrate, through simulations, significant gains in total data rate and greater robustness against imperfections in Channel State Information at the Transmitter (CSIT) when compared to conventional approaches such as Orthogonal Multiple Access (OMA) and Spatial Division Multiple Access (SDMA). The results confirm that the adoption of RSMA in the analyzed scenarios enables substantial gains in spectral efficiency, representing a relevant advancement for the development of future NTN networks aligned with the vision of 6th Generation (6G) mobile systems.

Keywords: non-terrestrial networks; LEO satellites; RSMA; TDD; power allocation.

LISTA DE FIGURAS

Figura 1 – Plataforma NTN atuando como UE.	28
Figura 2 – Plataforma NTN como retransmissor: cenário de <i>backhaul</i> sem fio.	29
Figura 3 – Plataforma NTN como retransmissor: acesso retransmitido.	29
Figura 4 – Plataforma NTN atuando como BS.	29
Figura 5 – Conexão simultânea a uma NG-RAN terrestre e a uma NG-RAN NTN.	30
Figura 6 – Conexão simultânea a duas NG-RAN baseada em NTN.	31
Figura 7 – Estrutura de <i>frame</i> assíncrona.	39
Figura 8 – Estrutura de <i>frame</i> síncrono com aplicação do TA.	41
Figura 9 – Ilustração do processo de aplicação do RSMA em comunicação MISO em DL.	48
Figura 10 – Ilustração de um sistema de comunicação LEO-Terrestre empregando RSMA, destacando o AoD.	54
Figura 11 – Taxa total de dados, por unidade de banda, alcançável pelos usuários em função de \bar{D} para avaliar o desempenho dos MAs com CSIT perfeito.	63
Figura 12 – Taxa total de dados, por unidade de banda, alcançável pelos usuários em função de \bar{D} para avaliar o desempenho dos MAs com CSIT imperfeito.	64
Figura 13 – Taxa total de dados, por unidade de banda, alcançável pelos usuários em função de ϵ para avaliação do desempenho dos MAs.	65
Figura 14 – Ilustração de um sistema de comunicação LEO-Terrestre empregando RSMA, destacando o ângulo de elevação do usuário j	67
Figura 15 – Taxa de <i>outage</i> em função de R_{\min}^{DL}	80
Figura 16 – Tempo médio total de <i>frame</i> em função de R_{\min}^{DL}	82
Figura 17 – Taxa de dados média total dos usuários no DL em função de R_{\min}^{DL}	83
Figura 18 – Taxa de dados média total dos usuários no UL em função de R_{\min}^{DL}	83
Figura 19 – Taxa de <i>outage</i> em função de N	84
Figura 20 – Tempo médio total de <i>frame</i> em função de N	85
Figura 21 – Taxa de dados média total dos usuários no DL em função de N	86
Figura 22 – Taxa de dados média total dos usuários no UL em função de N	86
Figura 23 – Taxa de <i>outage</i> em função de J	87
Figura 24 – Tempo médio total de <i>frame</i> em função de J	88
Figura 25 – Taxa de dados média total dos usuários no DL em função de J	89
Figura 26 – Taxa de dados média total dos usuários no UL em função de J	90

LISTA DE TABELAS

Tabela 1 – Comparativo entre os modos TDD e FDD para 5G NTN.	38
Tabela 2 – Parâmetros de simulação para avaliar a alocação de potência para mensagens comum e privadas em redes RSMA LEO-Terrestre.	60
Tabela 3 – Parâmetros de simulação para avaliação do escalonamento em redes NTN com diferentes técnicas de múltiplo acesso.	78

LISTA DE ABREVIATURAS E SIGLAS

3GPP	<i>3rd Generation Partnership Project</i>
4G	<i>4th Generation</i>
5G	<i>5th Generation</i>
6G	<i>6th Generation</i>
AoD	<i>Angle of Departure</i>
AWGN	<i>Additive White Gaussian Noise</i>
BS	<i>Base Station</i>
CDMA	<i>Code Division Multiple Access</i>
CSI	<i>Channel State Information</i>
CSIT	<i>Channel State Information at the Transmitter</i>
DL	<i>Downlink</i>
DU	<i>Distributed Unit</i>
ESSA	<i>Enhanced Slot Allocation and Scheduling Algorithm</i>
FDD	<i>Frequency Division Duplexing</i>
FDMA	<i>Frequency Division Multiple Access</i>
GEO	<i>Geostationary Earth Orbiting</i>
GP	<i>Guard Period</i>
HAP	<i>High Altitude Platform</i>
LEO	<i>Low Earth Orbit</i>
LoS	<i>Line-of-Sight</i>
MA	<i>Multiple Access</i>
MEO	<i>Medium Earth Orbiting</i>
MIMO	<i>Multiple-Input Multiple-Output</i>
MIP	<i>Mixed-Integer Programming</i>
MISO	<i>Multiple-Input Single-Output</i>
MMSE	<i>Minimum Mean-Square Error</i>
MRT	<i>Maximum-Ratio Transmission</i>
NG-RAN	<i>Next Generation Radio Access Network</i>
NGSO	<i>Non-GEO</i>
NOMA	<i>Non-Orthogonal Multiple Access</i>

NR	<i>New Radio</i>
NTN	<i>Non-Terrestrial Network</i>
OFDM	<i>Orthogonal Frequency Division Multiplexing</i>
OFDMA	<i>Orthogonal Frequency-Division Multiple Access</i>
OMA	<i>Orthogonal Multiple Access</i>
PD-NOMA	<i>Power-domain Non-Orthogonal Multiple Access</i>
PRB	<i>Physical Resource Block</i>
QoS	<i>Quality of Service</i>
RB	<i>Resource Block</i>
RR	<i>Round Robin</i>
RSMA	<i>Rate-Splitting Multiple Access</i>
RTT	<i>Round Trip Time</i>
SDMA	<i>Spatial Division Multiple Access</i>
SIC	<i>Successive Interference Cancellation</i>
SINR	<i>Signal-to-Interference-Plus-Noise Ratio</i>
SNR	<i>Signal-to-Noise Ratio</i>
TA	<i>Timing Advance</i>
TDD	<i>Time Division Duplexing</i>
TDMA	<i>Time Division Multiple Access</i>
TN	<i>Terrestrial Network</i>
TR	<i>Technical Report</i>
UAS	<i>Unmanned Aircraft System</i>
UAV	<i>Unmanned Aerial Vehicle</i>
UE	<i>User Equipment</i>
UL	<i>Uplink</i>
ULA	<i>Uniform Linear Array</i>
URLLC	<i>Ultra-Reliable and Low Latency Communications</i>
WF	<i>Water-Filling</i>
ZF	<i>Zero-Forcing</i>

SUMÁRIO

1	INTRODUÇÃO	17
1.1	Revisão bibliográfica	19
1.2	Contribuições científicas	21
1.3	Organização do documento	22
1.4	Produções científicas	23
2	FUNDAMENTAÇÃO TEÓRICA	25
2.1	NTN	25
2.1.1	<i>Definição</i>	26
2.1.2	<i>Arquiteturas de integração e modelos operacionais</i>	27
2.1.2.1	<i>Plataforma NTN como UE</i>	27
2.1.2.2	<i>Plataforma NTN como retransmissor</i>	28
2.1.2.3	<i>Plataforma NTN como BS</i>	28
2.1.2.4	<i>Continuidade de serviço e multi-conectividade</i>	30
2.1.3	<i>Desafios em comunicações NTN</i>	31
2.1.3.1	<i>Atraso de propagação e perda de percurso</i>	31
2.1.3.2	<i>Efeito Doppler</i>	32
2.1.3.3	<i>Cobertura, vazão e handover</i>	32
2.1.3.4	<i>Implantação</i>	33
2.1.3.5	<i>Efeitos do clima espacial</i>	33
2.2	Duplexação e sincronização em NTN satelital	34
2.2.1	<i>Modos de duplexação</i>	35
2.2.1.1	<i>FDD</i>	36
2.2.1.2	<i>TDD</i>	36
2.2.1.3	<i>Comparativo entre FDD e TDD para NTN</i>	37
2.2.2	<i>Sincronização em TDD para NTN</i>	37
2.2.3	<i>TA</i>	40
2.3	Introdução ao MA	41
2.3.1	<i>Esquemas de MA de referência</i>	43
2.3.1.1	<i>OMA</i>	43
2.3.1.2	<i>SDMA</i>	44

2.3.1.3	<i>NOMA</i>	45
2.3.2	Fundamentos do RSMA	46
2.3.2.1	<i>RSMA no DL</i>	47
2.4	Pré-codificação linear	49
2.4.1	Pré-codificação em sistemas MISO	49
2.4.1.1	<i>MRT</i>	50
2.4.1.2	<i>ZF</i>	51
2.4.1.3	<i>MMSE</i>	51
3	ALOCAÇÃO DE POTÊNCIA PARA MENSAGENS COMUM E PRIVADAS EM REDES RSMA LEO-TERRESTRE	53
3.1	Modelo de sistema	53
3.1.1	<i>Modelo de canal</i>	54
3.1.2	<i>Modelo de transmissão</i>	55
3.2	Formulação do problema de otimização	56
3.3	Solução proposta para o problema de otimização	56
3.3.1	<i>Análise dos pré-codificadores</i>	57
3.3.2	<i>Divisão de potência</i>	59
3.3.2.1	<i>OMA/NOMA/Multicasting</i>	59
3.3.2.2	<i>SDMA/RSMA</i>	59
3.4	Avaliação de desempenho	60
3.4.1	<i>Detalhes da implementação</i>	60
3.4.2	<i>Soluções de referência</i>	61
3.4.3	<i>Resultados</i>	62
3.4.3.1	<i>Impacto da distância do usuário (\bar{D})</i>	62
3.4.3.2	<i>Impacto do nível de imperfeição do CSIT</i>	63
3.4.3.3	<i>Nota sobre a complexidade computacional</i>	64
3.5	Considerações parciais	65
4	ESCALONAMENTO DE UES EM SISTEMAS TDD EM REDES RSMA LEO-TERRESTRE	66
4.1	Modelo de sistema	66
4.1.1	<i>Modelo de canal</i>	68
4.1.2	<i>Modelo de transmissão</i>	70

4.1.3	<i>Modelo de frame TDD</i>	71
4.2	Formulação do problema de otimização	72
4.2.1	<i>Problema linearizado</i>	73
4.3	Soluções de <i>baseline</i>	74
4.3.1	<i>Solução ótima com CPLEX</i>	74
4.3.2	<i>RR</i>	75
4.3.3	<i>Cenário de Referência TDMA</i>	76
4.3.4	<i>Nota sobre a complexidade computacional das soluções</i>	76
4.4	Avaliação de desempenho	77
4.4.1	<i>Detalhes da implementação</i>	77
4.4.2	<i>Esquemas avaliados</i>	79
4.4.3	Resultados	79
4.4.3.1	<i>Impacto do requisito mínimo de QoS no DL</i>	79
4.4.3.2	<i>Impacto do número total de slots</i>	83
4.4.3.3	<i>Impacto do número total de usuários</i>	87
4.5	Considerações parciais	89
5	CONSIDERAÇÕES FINAIS	92
	REFERÊNCIAS	94

1 INTRODUÇÃO

A sociedade contemporânea vivencia uma era de transformação digital sem precedentes, impulsionada pela crescente demanda por conectividade ubíqua, de alta capacidade e baixa latência. De acordo com Ericsson (2025), projeta-se que o tráfego global de dados móveis cresça em torno de um fator de 2,3 vezes, alcançando 280 EB por mês em 2030. Este cenário pavimentará o caminho para uma miríade de aplicações emergentes e disruptivas, como a internet das coisas, do inglês *Internet of Things*, a indústria 4.0, as comunicações imersivas e os veículos autônomos (DEEPENDER *et al.*, 2021). Para habilitar plenamente essas aplicações, as futuras redes de sexta geração, do inglês *6th Generation* (6G), surgem como uma evolução natural, visando integrar de maneira sinérgica diferentes domínios e oferecer cobertura global com qualidade consistente (GIORDANI *et al.*, 2020).

No entanto, o paradigma convencional das redes terrestres, do inglês *Terrestrial Networks* (TNs), revela-se limitado diante de barreiras geográficas, econômicas e logísticas, impossibilitando a universalização da conectividade em áreas remotas ou de difícil acesso (HOSSEINIAN *et al.*, 2021). Para superar essas restrições, as redes não terrestres, do inglês *Non-Terrestrial Networks* (NTNs), especialmente as constelações de satélites em baixa órbita terrestre, do inglês *Low Earth Orbit* (LEO), emergem como uma solução tecnológica indispensável (DARWISH *et al.*, 2022). Ao complementar as TNs, as NTNs possibilitam ampliar a cobertura global e aumentar a resiliência das comunicações, configurando uma arquitetura integrada fundamental para a visão do 6G (HAROUNABADI; HEYN, 2023).

De acordo com Geraci *et al.* (2023), a integração vertical tridimensional, que unifica os segmentos espacial, aéreo e terrestre, proporciona benefícios significativos em termos de flexibilidade e robustez. Entretanto, esta integração também introduz desafios técnicos consideráveis, sobretudo no gerenciamento eficiente e inteligente dos recursos de rádio em um ambiente altamente heterogêneo e dinâmico (SHAHID *et al.*, 2024). Nesse contexto, a otimização conjunta da eficiência espectral e da eficiência energética emerge como prioridade central para viabilizar sistemas NTN práticos e sustentáveis (LAKEW *et al.*, 2023). O desempenho de uma rede integrada TN-NTN está diretamente associado ao grau de sofisticação de suas tecnologias de acesso ao rádio, do inglês *Radio Access Technologies*. Entre os diversos aspectos críticos, destacam-se duas frentes de pesquisa intensamente exploradas na literatura recente em sistemas NTN: o desenvolvimento de esquemas de acesso múltiplo, do inglês *Multiple Access* (MA), e a gestão precisa da sincronização em cenários que utilizam a duplexação por divisão no tempo, do

inglês *Time Division Duplexing* (TDD) (SHAH *et al.*, 2024; KANG *et al.*, 2023).

O primeiro grande desafio reside na definição de um esquema de MA capaz de lidar eficientemente com interferências e garantir alta qualidade de serviço em cenários multiusuário. O acesso múltiplo por divisão de taxa, do inglês *Rate-Splitting Multiple Access* (RSMA), surge como uma estrutura promissora por sua flexibilidade e capacidade de generalizar esquemas tradicionais, como o acesso múltiplo ortogonal, do inglês *Orthogonal Multiple Access* (OMA), o acesso múltiplo não-ortogonal, do inglês *Non-Orthogonal Multiple Access* (NOMA), e o acesso múltiplo por divisão espacial, do inglês *Spatial Division Multiple Access* (SDMA) (CLERCKX *et al.*, 2020). O RSMA permite ajustar dinamicamente a divisão de mensagens em feixes, do inglês *streams*, comuns e privados, otimizando a utilização do espectro em cenários heterogêneos. Entretanto, a alocação ótima de potência entre esses *streams* é formulada como um problema não convexo e de alta complexidade. Pesquisas recentes evidenciam os ganhos do RSMA em enlaces LEO, mas recorrem a métodos de busca exaustiva para determinar a alocação de potência ideal (SCHRÖDER *et al.*, 2023). Embora eficientes em termos analíticos, esses métodos são inviáveis para aplicações práticas em tempo real, indicando a necessidade urgente de abordagens de menor complexidade computacional.

O segundo desafio crítico está relacionado à sincronização temporal em sistemas NTN integrados às redes de quinta geração, do inglês *5th Generation* (5G), e 6G, que em grande parte empregam o TDD. A aplicação do TDD em enlaces satelitais é severamente impactada pelos atrasos de propagação elevados e não uniformes, resultando em dificuldades significativas para a sincronização do *Uplink* (UL) (KANG *et al.*, 2023). O mecanismo clássico de compensação de atraso de propagação, o avanço no tempo, do inglês *Timing Advance* (TA), apesar de funcional, induz a necessidade de períodos de guarda, do inglês *Guard Periods* (GPs), excessivamente longos em sistemas NTN. Esses períodos, durante os quais os recursos de rádio permanecem ociosos, podem levar a desperdícios superiores a 50% da capacidade do sistema, conforme demonstrado em estudos recentes (TRASPADINI *et al.*, 2024). Embora existam propostas de agendamento aprimorado para reduzir essa ineficiência, a otimização conjunta do agendamento temporal com esquemas de MA avançados, como o RSMA, permanece uma linha de pesquisa pouco explorada e de grande potencial.

Diante desse panorama, esta dissertação é motivada pela necessidade de avançar na otimização integrada do uso de recursos em redes satelitais. O objetivo central é abordar de forma conjunta a alocação de potência no domínio do RSMA e o gerenciamento temporal

em sistemas TDD com TA, contribuindo para o desenvolvimento de soluções mais eficientes e viáveis para as futuras redes 6G.

1.1 Revisão bibliográfica

O crescente interesse na técnica de RSMA, e na integração das NTN com os sistemas de comunicação 5G e além, tem impulsionado uma extensa gama de pesquisas (CLERCKX *et al.*, 2020; SCHRÖDER *et al.*, 2023; LI *et al.*, 2023; TRASPADINI *et al.*, 2024). Tais investigações visam não apenas otimizar a eficiência espectral e energética através de técnicas avançadas de acesso múltiplo, mas também superar desafios operacionais fundamentais, como o gerenciamento de interferência e a sincronização em enlaces de longa distância.

Neste âmbito, o RSMA trouxe uma abordagem inovadora para lidar com interferências e otimizar o desempenho de sistemas multiusuário. No trabalho de Clerckx *et al.* (2020), foi demonstrado que o RSMA pode ser interpretado como uma estrutura unificadora que generaliza técnicas clássicas como o OMA, NOMA, SDMA, e até mesmo *multicasting*. Essa flexibilidade advém da possibilidade de dividir a mensagem de cada usuário em partes comum e privada, otimizando a alocação de potência entre elas. A partir dessa concepção, diversos estudos passaram a explorar a viabilidade e os ganhos do RSMA em cenários mais complexos, incluindo sistemas via satélite, onde os desafios relacionados à propagação e à obtenção de informações perfeitas de estado de canal no transmissor, do inglês *Channel State Information at the Transmitter* (CSIT), são acentuados. A robustez do RSMA frente a CSIT imperfeito é uma das principais razões para seu potencial em ambientes satelitais, que apresentam alta mobilidade relativa e grandes variações dinâmicas nos canais.

Antes de adentrar nos desafios específicos de redes não terrestres, é importante notar que a maior parte da pesquisa em RSMA se concentrou, inicialmente, em cenários de redes terrestres. Nesses ambientes, a técnica demonstrou grande potencial para otimizar o uso de recursos em redes densas e com múltiplos pontos de acesso. Por exemplo, o trabalho de Tegos *et al.* (2025) introduz o conceito de RSMA de UL distribuído, do inglês *Distributed UL RSMA*, onde múltiplas unidades remotas de rádio, do inglês *Remote Radio Heads*, cooperam para decodificar os sinais dos usuários, expandindo a região de capacidade alcançável. Outra aplicação relevante em redes terrestres é a garantia de segurança na camada física. O estudo de Xia *et al.* (2024) investiga o uso do RSMA para maximizar a taxa de transmissão segura em cenários com interceptadores internos, onde os próprios usuários legítimos podem tentar

decodificar mensagens uns dos outros. Ambos os trabalhos reforçam a flexibilidade do RSMA para gerenciar a interferência de forma sofisticada, seja ela proveniente de múltiplos acessos em células vizinhas ou de ameaças de segurança.

Neste viés, destaca-se o estudo de Schröder *et al.* (2023), que analisou o uso do RSMA em sistemas multi-feixe, do inglês *multibeam*, baseados em satélites LEO. O trabalho compara o desempenho do RSMA com o SDMA e o OMA em cenários com diferentes graus de imperfeição do CSIT, confirmando a superioridade do RSMA em termos de taxa total de dados e robustez a erros de estimativa. Entretanto, a estratégia utilizada pelos autores para otimizar a divisão de potência entre *streams* comum e privados baseia-se em uma busca exaustiva, o que implica uma complexidade computacional elevada e inviável para aplicações em tempo real ou cenários com maior número de usuários. Esta limitação abre espaço para abordagens de menor complexidade, que mantenham os ganhos do RSMA e sejam factíveis em implementações práticas. A primeira contribuição desta dissertação parte diretamente desse desafio, ao propor uma solução analítica de baixa complexidade para a alocação de potência em sistemas RSMA-LEO, mantendo a eficiência espectral mesmo em condições de CSIT imperfeito, mas reduzindo drasticamente a necessidade de processamento.

Outro trabalho que se destaca nesse contexto é o de Schröder *et al.* (2024), que propõe uma abordagem baseada em aprendizado por reforço, do inglês *Reinforcement Learning*, para formação de feixe flexível e robusto em sistemas multi-feixe via satélite. A proposta busca adaptar dinamicamente os padrões de irradiação para maximizar o desempenho sob imperfeições de CSIT. Embora não aplique diretamente o RSMA, o estudo demonstra o crescente interesse em estratégias de alocação flexíveis e resilientes em ambientes satelitais, corroborando a relevância de abordagens que combinem robustez e eficiência espectral.

Além dos estudos focados na camada física e na alocação de potência, outra linha de pesquisa relevante está relacionada ao gerenciamento temporal em redes via satélite, especialmente em sistemas que utilizam o TDD (LI *et al.*, 2023; TRASPADINI *et al.*, 2024). O trabalho de Li *et al.* (2023) analisou detalhadamente o impacto do uso do mecanismo de TA, em redes NTN integradas ao padrão do *3rd Generation Partnership Project* (3GPP). O estudo evidencia que o TA, embora essencial para compensar os grandes atrasos de propagação, força a introdução de GPs, significativamente longos, o que pode levar a uma grave degradação da eficiência espectral. Em cenários realistas, a eficiência do sistema pode cair para menos de 40%, representando uma subutilização drástica dos recursos de rádio. Esta constatação destaca a urgência de se

repensar os mecanismos de agendamento e alocação temporal em sistemas NTN, principalmente em satélites LEO, que apresentam elevada variação de distância e, conseqüentemente, alta diversidade de atrasos entre usuários.

Como resposta parcial a esses desafios temporais, o estudo de Traspadini *et al.* (2024) propôs o algoritmo de alocação e agendamento de *slots* aprimorado, do inglês *Enhanced Slot Allocation and Scheduling Algorithm* (ESSA), uma solução que explora os longos períodos de guarda para alocar transmissões oportunísticas. O ESSA divide os usuários em grupos baseados em seus atrasos de propagação e utiliza o tempo ocioso do GP para agendar transmissões adicionais, recuperando parte da capacidade perdida. Os resultados apresentados mostram ganhos significativos em *throughput*, particularmente no UL, e destacam a viabilidade de estratégias baseadas em aproveitamento oportunístico do GP. Contudo, o ESSA atua exclusivamente no domínio temporal, sem explorar outras dimensões de otimização, como o domínio de potência ou o de código, e tampouco emprega esquemas de acesso múltiplo mais robustos, como o RSMA, o que limita seu potencial de ganhos adicionais em sistemas mais complexos ou densos.

Neste contexto, a segunda contribuição desta dissertação surge como uma extensão natural e complementar às soluções existentes. Ao combinar um agendamento temporal inteligente com a flexibilidade proporcionada pelo RSMA, é possível otimizar não apenas o uso do tempo, mas também o compartilhamento espectral e de potência entre usuários. A proposta desta dissertação integra um modelo de agendamento que considera a seleção otimizada de grupos de usuários por *slot*, incorporando requisitos de qualidade de serviço, do inglês *Quality of Service* (QoS), no *Downlink* (DL) e UL. Além disso, utiliza o RSMA no DL para potencializar ainda mais a eficiência espectral. Esse enfoque integrado permite minimizar os períodos de guarda, melhorar a utilização do canal e garantir maior equidade de serviço entre os usuários. Ao avançar além das abordagens tradicionais, esta solução evidencia a importância de estratégias conjuntas de alocação temporal e espacial, consolidando uma linha de pesquisa inovadora para sistemas satelitais de próxima geração.

1.2 Contribuições científicas

Esta dissertação visa estudar e propor soluções para otimização de recursos em NTNs com suporte à RSMA, considerando cenários de comunicação via satélite LEO, e integração com sistemas terrestres. No primeiro estudo, abordamos o problema de maximização da taxa total de dados em enlaces DL por meio da otimização da alocação de potência e do *design* de

pré-codificadores para mensagens comum e privadas em sistemas RSMA-LEO, considerando cenários com CSIT imperfeito e reduzindo significativamente a complexidade computacional ao empregar soluções analíticas, adaptadas ao cenário NTN, em substituição à busca exaustiva. Já no segundo estudo, investigamos a alocação otimizada de *slots* TDD com suporte a TA, integrando requisitos de QoS no DL e UL para múltiplos usuários em cenários LEO, com o objetivo de minimizar o tempo total de transmissão e melhorar a eficiência do uso espectral e temporal. Para isso, foi formulado um problema de otimização misto contínuo-binário, considerando a seleção de grupos de usuários, e proposto um modelo que incorpora pré-cálculo das taxas no DL usando RSMA e no UL usando OMA. Entre as principais contribuições, destacam-se: (i) a adaptação de uma abordagem de baixa complexidade para a divisão de potência em RSMA com CSIT imperfeito (Capítulo 3); (ii) o estudo pioneiro da interação entre RSMA e TA em ambientes NTN; e (iii) a formulação e análise de estratégias conjuntas de alocação temporal e espacial, visando reduzir períodos de guarda e ampliar o desempenho global do sistema (Capítulo 4).

1.3 Organização do documento

O Capítulo 1 apresenta a contextualização do problema estudado, a motivação para a integração de NTN com técnicas avançadas de acesso múltiplo, as lacunas na literatura, bem como os principais objetivos e contribuições desta dissertação. Além disso, a presente seção descreve o conteúdo abordado nos demais capítulos.

No Capítulo 2, são abordados os conceitos teóricos essenciais para a compreensão deste trabalho. A seção detalha os fundamentos de NTNs, com foco em sistemas de satélites LEO; os desafios de duplexação e sincronização em redes TDD via satélite, incluindo o mecanismo de TA; os princípios de esquemas de MA, com ênfase em RSMA; e as bases de pré-codificação e modelagem de canal.

O Capítulo 3 desenvolve a primeira contribuição principal desta dissertação. São apresentados o modelo de sistema, a formulação do problema de maximização da taxa de transmissão em redes RSMA via satélite LEO e a proposta de uma solução adaptada de alocação de potência de baixa complexidade. O capítulo inclui ainda a avaliação de desempenho da solução proposta em comparação com *benchmarks* da literatura, bem como uma conclusão parcial que sintetiza os principais achados e destaca as implicações práticas dos resultados.

No Capítulo 4, é desenvolvida a segunda contribuição principal. São apresentados o modelo de sistema, a formulação do problema de otimização de agendamento que visa minimizar

o tempo total de transmissão em sistemas TDD-NTN, integrando o mecanismo de TA com o esquema de acesso múltiplo RSMA. O capítulo descreve detalhadamente a metodologia adotada para resolver o problema, as considerações técnicas para a implementação e apresenta uma conclusão parcial que resume as contribuições obtidas, apontando as principais vantagens da abordagem proposta.

Por fim, o Capítulo 5 apresenta as considerações finais da dissertação, sintetizando as principais conclusões extraídas a partir das contribuições apresentadas. Ademais, discute as implicações dos resultados para o avanço das redes satelitais de próxima geração e sugere direções promissoras para trabalhos futuros.

1.4 Produções científicas

O seguinte trabalho está diretamente relacionado com o conteúdo desta dissertação e foi aceito para publicação:

M. B. de Paiva and F. R. M. Lima, "Power Allocation for Common and Private Messages in LEO-Terrestrial RSMA Networks," accepted for publication in *XLIII Brazilian Symposium on Telecommunications and Signal Processing (SBrT 2025)*, Natal, RN, Brazil, 2025, pp. 1–5.

Ao longo do período de mestrado, o autor desta dissertação teve oportunidade de cooperar com outros trabalhos científicos que estão listados abaixo:

J. V. B. Labres, P. S. S. de Souza, D. V. L. Moreira, P. R. B. da Silva, M. B. de Paiva, E. C. Alves, and L. M. Azevedo, "Order, Unite, and Conquer: A Group Formulation for Multi-Armed Bandits in Microservice Provisioning," in *Proc. 33rd Euromicro Int. Conf. on Parallel, Distributed, and Network-Based Processing (PDP)*, 2025, pp. 58–61. doi:10.1109/PDP66500.2025.00017.

D. R. de Luna, F. M. Abinader, P. R. B. da Silva, M. B. de Paiva, V. A. de Sousa, and A. J. Venâncio Neto, "Multi-Objective Optimization of Open RAN Data Center Placement for Enhanced 5G Deployments," in *Proc. IEEE Conf. on Network Function Virtualization and Software Defined Networks (NFV-SDN)*, 2024, pp. 1–4. doi:10.1109/NFV-SDN61811.2024.10807497.

J. Z. Correia, F. M. Abinader, P. R. B. da Silva, and M. B. de Paiva, "KPI-Based Predictive

Computational Resource Consumption in 5G Open RAN Networks,” in *Proc. IEEE Conf. on Network Function Virtualization and Software Defined Networks (NFV-SDN)*, 2024, pp. 1–6. doi:10.1109/NFV-SDN61811.2024.10807488.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, são apresentados os conceitos teóricos fundamentais que servem de base para o desenvolvimento das contribuições desta dissertação. Serão explorados aspectos essenciais relacionados ao contexto de NTN, às técnicas avançadas de MA e às tecnologias complementares necessárias para a operação eficiente do sistema proposto. Na Seção 2.1, são discutidos os conceitos gerais sobre NTN, incluindo suas arquiteturas, classificações orbitais e os principais desafios associados à operação de enlaces via satélite. Em seguida, a Seção 2.2 aborda os aspectos de duplexação e sincronização em sistemas NTN via satélite, com ênfase no TDD e no mecanismo de TA, fundamentais para viabilizar a comunicação bidirecional. A Seção 2.3 introduz as principais técnicas de MA, contextualizando a evolução desde os esquemas ortogonais até o RSMA, destacando suas vantagens em cenários multiusuário. Por fim, a Seção 2.4 trata dos fundamentos de pré-codificação linear, apresentando conceitos básicos e técnicas utilizadas para mitigar interferências e melhorar o desempenho dos enlaces de comunicação.

2.1 NTN

Com a evolução dos sistemas de comunicação móvel surgem diversos novos requisitos de rede, como alta capacidade de conectividade, latência ultra baixa, comunicação ubíqua e de alta confiabilidade, dentre outras (3GPP, 2020). No entanto, as TNs tradicionais não são suficientes para alcançar esse conjunto de requisitos, principalmente quando pensamos em situações extremas, como regiões onde a implementação de infraestrutura terrestre é dispendiosa ou impraticável, por exemplo, em áreas rurais, embarcações e aeronaves. Ou ainda, em situações de desastres, naturais ou causados pelo homem, que facilmente comprometem as instalações e interrompem a capacidade de conectividade móvel (AZARI *et al.*, 2022).

Neste contexto, as NTNs são consideradas como tecnologia potencial para garantir os requisitos supracitados, tanto pela capacidade de alcançar regiões remotas e maior resiliência a desastres, quanto por garantir alta performance de comunicação (LAKEW *et al.*, 2023). As NTNs podem ser implantadas em coordenação com redes terrestres tradicionais, o que demonstra o potencial não só de prover conectividade em localidades remotas, mas também de melhorar a capacidade das redes terrestres, gerando implicações tecnológicas e sociais de grande significado a longo prazo (AZARI *et al.*, 2022). Esta convergência das comunicações terrestres, aéreas e espaciais representa um esforço considerável tanto para a indústria de comunicações móveis

quanto para a de satélites, pois implica na definição e orquestração inteligente de uma nova arquitetura de rede tridimensional (GERACI *et al.*, 2023).

2.1.1 Definição

De acordo com o 3GPP, NTN é um sistema que opera parcialmente, ou completamente, para fins de comunicação através de um veículo espacial ou aéreo (AZARI *et al.*, 2022). Os veículos espaciais são classificados de acordo com três principais parâmetros: altitude, tamanho de feixe e órbita. Assim, as plataformas espaciais são classificadas em (RINALDI *et al.*, 2020; LAKEW *et al.*, 2023):

- Órbita terrestre geoestacionária, do inglês *Geostationary Earth Orbiting* (GEO): altitude de aproximadamente 35.786 km, com tamanho de feixe entre 200 e 3.500 km, e órbita circular equatorial com período de órbita igual ao período de rotação da Terra, de modo que o satélite aparenta estar fixo no céu para observadores terrestres.
- Órbita terrestre média, do inglês *Medium Earth Orbiting* (MEO): altitude entre 7.000 e 25.000 km, com tamanho de feixe entre 100 e 1.000 km, e órbita circular, com movimento relativo moderado em relação à superfície terrestre.
- LEO: altitude entre 300 e 1.500 km, com tamanho de feixe entre 100 e 1.000 km, e órbita circular, com alta mobilidade relativa e menores atrasos de propagação.

Vale destacar que LEO e MEO também são categorizados como satélites não geoestacionários, do inglês *Non-GEO* (NGSO), porque seus períodos de órbita são menores que período de rotação da Terra, variando entre 1.5 a 10 horas (RINALDI *et al.*, 2020).

No caso dos veículos aéreos, são utilizados plataformas de sistemas de aeronaves não tripuladas, do inglês *Unmanned Aircraft Systems* (UASs), destacando-se as plataformas de altas altitudes, *High Altitude Platforms* (HAPs), e os veículos aéreos não tripulados, *Unmanned Aerial Vehicles* (UAVs). Os HAPs operam a uma altitude de 20 km enquanto os UAVs a algumas centenas de metros. Em geral, o tamanho de feixe dos UASs fica entre 5 a 200 km no solo (RINALDI *et al.*, 2020).

As plataformas NTN, sejam espaciais ou aéreas, podem ser classificadas conforme o tipo de carga útil embarcada. Existem, basicamente, duas configurações: carga útil transparente, ou *bent-pipe*, e carga útil regenerativa. Na configuração transparente, a plataforma realiza apenas funções básicas, como filtragem em radiofrequência, conversão de frequência e amplificação de sinal, sem processamento adicional do tráfego. Já na configuração regenerativa, a plataforma

integra funções completas da estação rádio base, em inglês *Base Station* (BS), permitindo processamento avançado a bordo (RINALDI *et al.*, 2020; AZARI *et al.*, 2022).

Por fim, além das plataformas espaciais ou aéreas, a arquitetura de acesso NTN é composta por outros elementos importantes. O terminal NTN pode ser um equipamento de usuário, do inglês *User Equipment* (UE), conforme definido pelo 3GPP, ou terminais específicos para satélites. Terminais com antena de abertura muito pequena costumam operar na banda Ka (cerca de 30 GHz no UL e 20 GHz no DL), enquanto terminais portáteis operam na banda S (aproximadamente 2 GHz). Outro elemento essencial é o *gateway* NTN, que funciona como um nó lógico conectando a plataforma NTN ao núcleo de rede 5G. Já o *link* de serviço corresponde ao enlace de rádio entre o terminal NTN e a plataforma não-terrestre, enquanto o *link* de alimentação, ou *feeder link*, conecta o *gateway* NTN à plataforma (RINALDI *et al.*, 2020).

Vale ressaltar que, apesar da discussão geral sobre diferentes órbitas e plataformas aéreas, este trabalho foca exclusivamente no contexto de satélites LEO, em função de suas características de baixa latência, maior proximidade com os usuários e potencial para oferecer altas taxas de dados (RINALDI *et al.*, 2020; LAKEW *et al.*, 2023).

Após descrever as características das plataformas, discute-se agora como estas podem ser integradas funcionalmente às redes terrestres.

2.1.2 Arquiteturas de integração e modelos operacionais

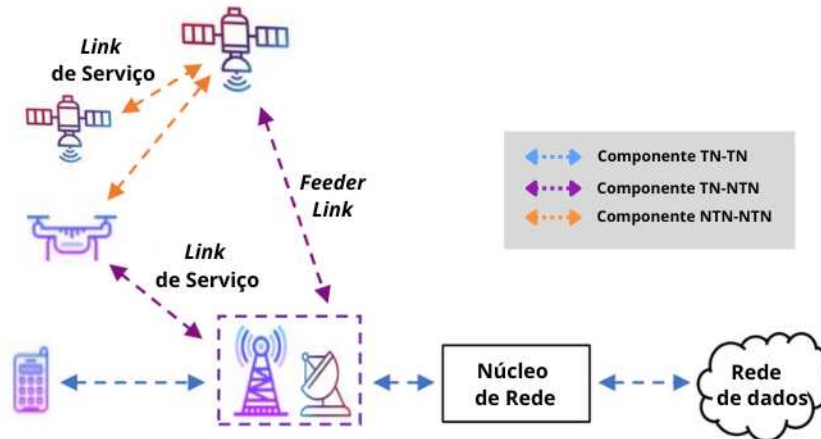
A integração das NTN com as redes terrestres pode ocorrer na camada física (PHY), exigindo que a plataforma NTN utilize a mesma tecnologia de acesso por rádio da rede terrestre, ou na camada de rede (NET), o que permite o uso de tecnologias distintas (RINALDI *et al.*, 2020; AZARI *et al.*, 2022). Independentemente da camada de integração, a plataforma NTN pode assumir diferentes papéis funcionais, definindo a arquitetura operacional da rede, como será descrito a seguir.

2.1.2.1 Plataforma NTN como UE

Neste modelo, a plataforma NTN (tipicamente um UAV ou satélite LEO) atua como um UE, sendo atendida pela infraestrutura de uma TN ou por satélites em altitudes superiores (RINALDI *et al.*, 2020), conforme ilustra a Figura 1. O 3GPP já estudou os desafios e oportunidades de servir UAVs como um novo tipo de UE (3GPP, 2020). Esta arquitetura é interessante pois pode dispensar a necessidade de uma vasta rede de estações terrestres para

coletar dados ou enviar comandos para a plataforma (RINALDI *et al.*, 2020).

Figura 1 – Plataforma NTN atuando como UE.



Fonte: Adaptado de Azari *et al.* (2022).

2.1.2.2 Plataforma NTN como retransmissor

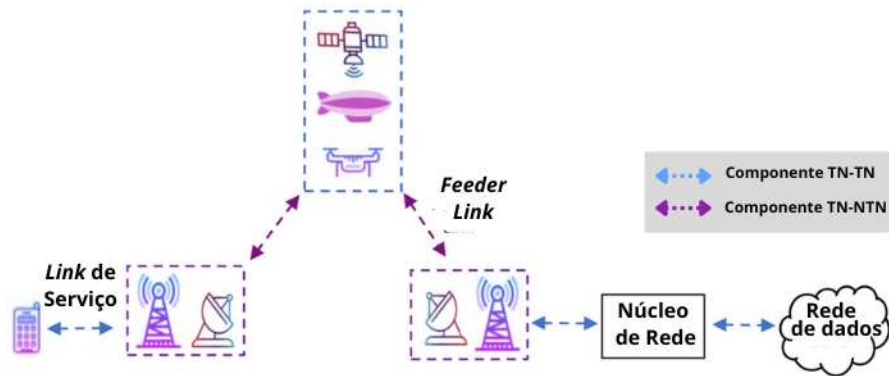
A plataforma NTN pode funcionar como um repetidor, utilizando uma carga útil transparente para retransmitir sinais (AZARI *et al.*, 2022). Este modelo se desdobra em dois cenários principais, conforme ilustra as Figuras 2 e 3:

- *Backhaul Sem Fio*: A plataforma fornece o *link de backhaul* entre uma BS terrestre e o núcleo da rede, substituindo ou complementando as conexões de fibra óptica.
- *Acesso Retransmitido*, do inglês *Relay-Like*: A plataforma retransmite o sinal entre usuários finais e a BS, estendendo a cobertura. Este conceito está alinhado com as arquiteturas de acesso e *backhaul* integrado, do inglês *Integrated Access and Backhaul*, descritas pelo 3GPP (3GPP, 2018).

2.1.2.3 Plataforma NTN como BS

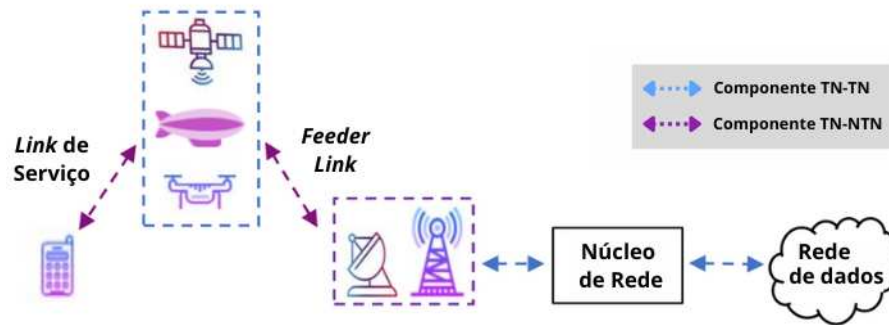
Equipada com uma carga útil regenerativa e capacidade de processamento a bordo, a plataforma NTN pode incorporar as funcionalidades de uma BS, conforme ilustra a Figura 4 (RINALDI *et al.*, 2020; AZARI *et al.*, 2022). Esta arquitetura pode seguir o modelo de divisão funcional da rede de acesso de próxima geração, do inglês *Next Generation Radio Access Network* (NG-RAN), onde a plataforma hospeda a unidade distribuída, do inglês *Distributed Unit* (DU), e se conecta via interface F1 a uma unidade central, do inglês *Central Unit*, em solo (RINALDI *et al.*, 2020).

Figura 2 – Plataforma NTN como retransmissor: cenário de *backhaul* sem fio.



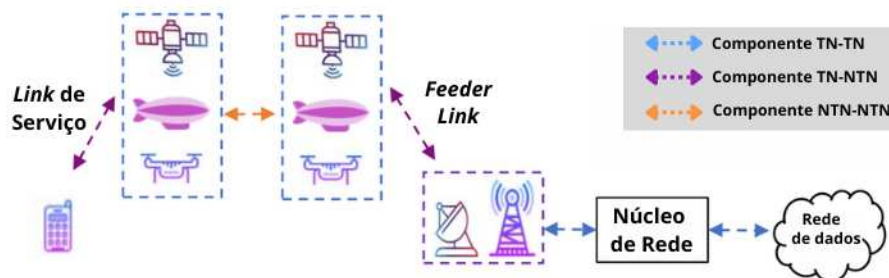
Fonte: Adaptado de Azari *et al.* (2022).

Figura 3 – Plataforma NTN como retransmissor: acesso retransmitido.



Fonte: Adaptado de Azari *et al.* (2022).

Figura 4 – Plataforma NTN atuando como BS.



Fonte: Adaptado de Azari *et al.* (2022).

2.1.2.4 Continuidade de serviço e multi-conectividade

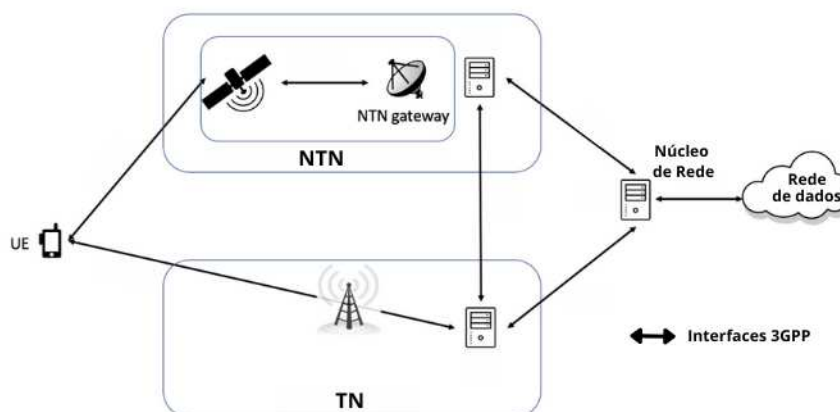
As arquiteturas mencionadas são fundamentais para habilitar dois dos casos de uso mais importantes do 5G e além: a continuidade de serviço e a multi-conectividade (RINALDI *et al.*, 2020). A integração NTN-TN é essencial para garantir serviços contínuos não apenas entre diferentes tipos de redes (TN e NTN), mas também entre duas redes NTN distintas (RINALDI *et al.*, 2020).

O 3GPP estuda o recurso de multi-conectividade, que permite a um terminal acessar simultaneamente a uma TN e a uma NTN, ou até mesmo duas redes NTN diferentes, por exemplo, GEO e LEO (RINALDI *et al.*, 2020). Esta capacidade oferece enormes benefícios em termos de resiliência, balanceamento de carga e otimização de desempenho. Cenários de multi-conectividade incluem (RINALDI *et al.*, 2020):

- Conexão simultânea a uma NG-RAN terrestre e a uma NG-RAN baseada em NTN transparente (RINALDI *et al.*, 2020).
- Combinação de duas NG-RANs baseadas em NTN, como uma GEO para fornecer largura de banda adicional e, conseqüentemente, maior throughput, e uma LEO para baixa latência (RINALDI *et al.*, 2020; AZARI *et al.*, 2022).
- Conexão a uma NG-RAN terrestre e a uma NG-RAN com DU regenerativo a bordo do satélite (RINALDI *et al.*, 2020).

As Figuras 5 e 6 ilustram simplificada os modelos de continuidade de serviço.

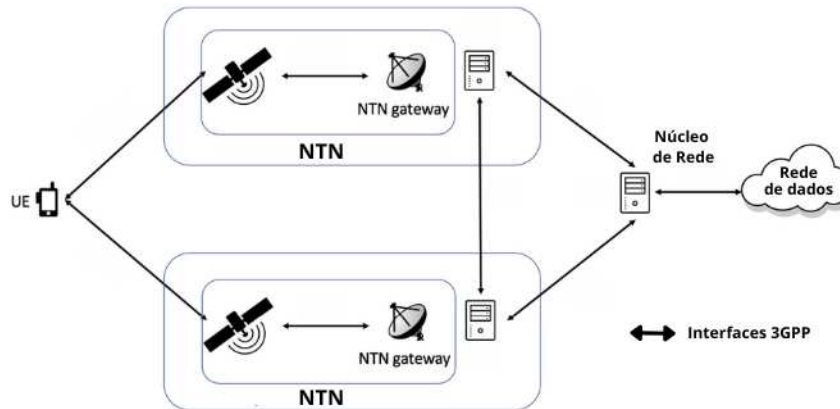
Figura 5 – Conexão simultânea a uma NG-RAN terrestre e a uma NG-RAN NTN.



Fonte: Adaptado de Rinaldi *et al.* (2020).

Estas arquiteturas, muitas vezes interligadas por enlaces inter-satélite, são cruciais para a visão de redes 5G e futuras redes 6G, visando conectividade global e onipresente (RI-

Figura 6 – Conexão simultânea a duas NG-RAN baseada em NTN.



Fonte: Adaptado de Rinaldi *et al.* (2020).

NALDI *et al.*, 2020; LAKEW *et al.*, 2023; AZARI *et al.*, 2022). Neste sentido, esta dissertação explora a arquitetura operacional simplificada com o satélite LEO atuando com BS, com enlace direto aos UEs.

Apesar de seu vasto potencial, a integração de plataformas NTN em redes de comunicação impõe um conjunto significativo de desafios técnicos, que serão discutidos na próxima seção.

2.1.3 Desafios em comunicações NTN

Os principais desafios da implantação e *design* de NTNs decorrem diretamente das características intrínsecas ao ambiente NTN, como as grandes altitudes, alta mobilidade das plataformas e vastas áreas de cobertura. A seguir, são detalhados os desafios mais proeminentes.

2.1.3.1 Atraso de propagação e perda de percurso

O atraso de propagação é um dos desafios mais impactantes nas comunicações NTN, especialmente para aplicações que exigem baixa latência, como as de comunicações ultra-confiáveis e de baixa latência, do inglês *Ultra-Reliable and Low Latency Communications* (URLLC) (AZARI *et al.*, 2022). A altitude das plataformas NTN introduz atrasos consideráveis no enlace de comunicação. Para satélites GEO, por exemplo, a latência de ida e volta, do inglês *Round Trip Time* (RTT), pode chegar a 270 ms (AZARI *et al.*, 2022). Este atraso prolongado afeta negativamente os protocolos das camadas superiores, como o controle de retransmissões e o tempo de resposta, e pode tornar as medições da qualidade do canal desatualizadas no momento em que são utilizadas (AZARI *et al.*, 2022; LAKEW *et al.*, 2023).

Além do atraso absoluto, as grandes áreas de cobertura dos satélites geram um atraso diferencial significativo, onde usuários no centro e na borda da célula recebem o sinal em instantes distintos (LAKEW *et al.*, 2023). Em um sistema LEO a 800 km de altitude, por exemplo, o atraso de propagação unidirecional pode chegar a 2,82 ms, o que, em sistemas TDD, pode exigir um período de guarda de até 5,64 ms — equivalente a mais de cinco subframes do novo rádio 5G, do inglês *5G New Radio* (NR), levando a um uso ineficiente dos recursos de rádio (TRASPADINI *et al.*, 2024).

A perda de percurso, do inglês *Path Loss*, também é substancialmente maior em sistemas NTN devido às longas distâncias que o sinal precisa percorrer (LAKEW *et al.*, 2023). Embora plataformas de menor altitude como UAVs e HAPs apresentem atrasos e perdas mais próximos aos de sistemas terrestres (LAKEW *et al.*, 2023), a modelagem do canal NTN permanece complexa, necessitando considerar fatores como condições atmosféricas e a presença de obstáculos (AZARI *et al.*, 2022).

2.1.3.2 Efeito Doppler

A mobilidade das plataformas NTN, especialmente a dos satélites LEO, introduz um efeito Doppler muito mais severo do que o encontrado em TN (AZARI *et al.*, 2022). Um satélite LEO a 600 km de altitude, operando a 2 GHz, pode gerar um desvio de frequência de até 48 kHz, o que é aproximadamente dez vezes maior que o experimentado por um usuário em um trem de alta velocidade (AZARI *et al.*, 2022).

Este grande desvio de frequência representa um desafio crítico para a camada física, exigindo adaptações no projeto de sistemas baseados em acesso múltiplos por divisão de frequência, do inglês *Frequency Division Multiple Access* (FDMA), como o ajuste do espaçamento entre subportadoras para evitar interferência inter-portadoras (AZARI *et al.*, 2022). Os sinais de referência, utilizados para monitoramento do canal, também precisam ser adaptados para compensar o Doppler e os efeitos específicos de multi-percurso.

2.1.3.3 Cobertura, vazão e handover

A cobertura de uma plataforma NTN é diretamente proporcional à sua altitude; um satélite GEO pode cobrir uma área com raio de até 3.500 km, enquanto LEOs e MEOs cobrem áreas com raios de até 1.000 km. No entanto, a alta velocidade dos satélites NGSO faz com que essa cobertura seja dinâmica, havendo dois principais cenários: (i) as células se movem

sobre a superfície da Terra (AZARI *et al.*, 2022), (ii) ou satélites se revezando no provimento de cobertura para células fixas.

Essa natureza dinâmica tem duas implicações principais:

1. Gerenciamento de *handover*: A mobilidade do satélite, somada à mobilidade dos usuários, exige procedimentos de *handover* frequentes e complexos para manter a continuidade do serviço. O *handover* pode ocorrer entre feixes de um mesmo satélite (intra-satélite), entre satélites diferentes (inter-satélite) ou entre a rede NTN e uma rede terrestre (inter-rede ou vertical) (AZARI *et al.*, 2022).
2. Gerenciamento da área de rastreamento, do inglês *Paging*: A mobilidade das células NTN dificulta o gerenciamento da área de rastreamento dos usuários. O terminal pode não conseguir determinar com precisão em qual área está localizado, resultando em altas cargas de *paging* para a rede (AZARI *et al.*, 2022).

A taxa de vazão de dados, do inglês *throughput*, é afetada por todos esses fatores, além do risco aumentado de interferência co-canal devido às grandes áreas de cobertura.

2.1.3.4 Implantação

Os diferentes tipos de plataformas NTN oferecem flexibilidades e cronogramas de implantação muito distintos. Plataformas aéreas como UAVs e HAPs oferecem uma solução de implantação rápida, flexível e de menor custo, podendo ser colocadas em operação em questão de horas. Essa agilidade as torna ideais para cenários temporários, como resposta a desastres ou cobertura de eventos especiais (AZARI *et al.*, 2022).

Em contrapartida, os satélites representam uma solução de longo prazo, cujo processo de projeto, construção, lançamento e comissionamento pode levar meses ou até anos. Embora o custo inicial seja significativamente maior, uma vez em órbita, um satélite oferece cobertura sobre uma vasta área de forma contínua e estável (AZARI *et al.*, 2022).

2.1.3.5 Efeitos do clima espacial

As operações de satélite, especialmente em órbitas mais altas, são vulneráveis aos efeitos do clima espacial. O ambiente espacial é permeado por um fluxo variável de radiação de alta energia, proveniente de raios cósmicos, radiação solar e dos cinturões de radiação de Van Allen (AZARI *et al.*, 2022). Esses cinturões, compostos por partículas de baixa energia, representam um risco particular para satélites que passam longos períodos em órbita MEO.

Satélites GEO operam acima dos cinturões, enquanto os LEOs, em sua maioria, orbitam abaixo deles, estando mais protegidos. A radiação pode danificar componentes eletrônicos sensíveis, representando um perigo para a integridade e o tempo de vida útil das espaçonaves.

Nesta dissertação explora-se cenários simplificados, cujos modelos de canal utilizados consideram apenas os efeitos das perdas de propagação, perdas de percursos, e do efeito Doppler no CSIT.

A seguir, são discutidos os conceitos de duplexação e sincronização, fundamentais para o contexto deste trabalho.

2.2 Duplexação e sincronização em NTN satelital

A comunicação bidirecional, que permite a transmissão e recepção de dados, é uma função elementar de qualquer sistema de telecomunicações. O método pelo qual os fluxos de DL e UL compartilham o meio físico, conhecido como duplexação, representa uma escolha de projeto fundamental com profundas implicações no desempenho e na complexidade da rede (LI *et al.*, 2024). Historicamente, as redes de satélite têm operado predominantemente em duplexação por divisão de frequência, do inglês *Frequency Division Duplexing* (FDD), onde o UL e o DL utilizam bandas de frequência distintas. Essa abordagem simplifica o *design* do sistema, pois elimina a necessidade de uma sincronização temporal rigorosa entre os enlaces, um desafio particular em redes com longos atrasos de propagação (TRASPADINI *et al.*, 2024).

Em contrapartida, as TNs modernas, incluindo o 5G NR, favorecem o TDD (TRASPADINI *et al.*, 2024; KANG *et al.*, 2023). O TDD oferece vantagens significativas, como a flexibilidade na alocação de recursos para tráfego assimétrico e um projeto de *hardware* mais simples, que pode exigir apenas um conjunto de antenas para o transceptor (LI *et al.*, 2024). Talvez o benefício mais relevante para as NTN seja a eficiência no uso do espectro, pois o TDD não requer bandas de frequência simétricas, cuja obtenção é particularmente difícil em faixas de frequência mais baixas, adequadas para terminais portáteis (TRASPADINI *et al.*, 2024; KANG *et al.*, 2023; SHAHID *et al.*, 2024). Para garantir uma integração transparente entre os ecossistemas terrestre e não terrestre — permitindo que um usuário utilize um terminal unificado para se comunicar tanto com uma BS terrestre quanto com um satélite — o 3GPP sugere que as NTN também sejam projetadas para operar em TDD (3GPP, 2020).

No entanto, a aplicação do TDD em redes NTN via satélite não é trivial e introduz um conjunto único de desafios (KANG *et al.*, 2023). A eficácia do TDD depende de uma

sincronização temporal precisa para evitar que as transmissões de UL e DL interfiram entre si. Em redes via satélite, essa sincronização é extremamente complexa devido à longa distância de transmissão, que resulta em atrasos de propagação muito grandes. Mesmo para satélites LEO, o RTT pode se estender da ordem de símbolos de multiplexação por divisão de frequência ortogonal, do inglês *Orthogonal Frequency Division Multiplexing* (OFDM), para a ordem de múltiplos *slots* (MEDINA-ACOSTA *et al.*, 2024).

Este grande atraso de propagação impacta diretamente a estrutura de quadros TDD, exigindo a inserção de um GP para evitar a colisão entre os sinais de UL e DL. No entanto, o parâmetro crítico para a definição do GP não é o atraso absoluto de ida e volta (RTT), mas sim o atraso diferencial entre os usuários atendidos simultaneamente. Em cenários com grande dispersão geográfica, especialmente em NTN, esse atraso diferencial pode atingir valores da ordem de milissegundos, conforme discutido em (TRASPADINI *et al.*, 2024). Isso demanda períodos de guarda mais longos, comprometendo a eficiência espectral do sistema. Para mitigar esse efeito, mecanismos de sincronização como o TA são empregados. Contudo, como será detalhado, a aplicação do TA em NTN, embora necessária, introduz novas complexidades e não elimina por completo a ineficiência dos GPs, o que motiva a busca por estratégias de otimização (TRASPADINI *et al.*, 2024).

Esta seção tem como objetivo apresentar esses desafios. Primeiramente, serão apresentados os modos de duplexação FDD e TDD. Em seguida, aprofundaremos nas dificuldades específicas da sincronização em TDD para NTN. Posteriormente, o mecanismo de TA será detalhado como a solução padrão para este problema. Finalmente, analisaremos a ineficiência gerada pelos GPs, estabelecendo assim o alicerce técnico e a motivação para o desenvolvimento de esquemas de alocação de *slots* e agendamento avançados, que constituem uma das contribuições centrais desta dissertação.

2.2.1 Modos de duplexação

As duas principais técnicas de duplexação empregadas em redes de comunicação sem fio são o FDD e o TDD (JAAFARI *et al.*, 2023). A escolha entre FDD e TDD é uma decisão de projeto fundamental que impacta a alocação de espectro, a complexidade do *hardware*, a latência e a gestão de interferência do sistema.

2.2.1.1 FDD

No FDD, o UL e o DL operam simultaneamente, mas em bandas de frequência distintas e separadas. Tradicionalmente, o FDD tem sido o modo de operação predominante e mais popular em redes de satélite. Essa preferência é motivada por várias vantagens inerentes ao FDD em cenários de longa distância (JAAFARI *et al.*, 2023):

- Transmissão Contínua: O FDD permite a comunicação *full-duplex*, com transmissões de UL e DL ocorrendo de forma simultânea e contínua.
- Sincronização Simplificada: Como os canais de UL e DL são separados por frequência, o FDD elimina a necessidade de uma sincronização temporal rigorosa entre o transmissor e o receptor, o que é uma vantagem significativa para enlaces de satélite com longos atrasos de propagação.
- Gerenciamento de Interferência: A separação em frequência mitiga naturalmente a interferência entre os sinais de UL e DL. Além disso, o FDD não requer a inserção de um GP entre as transmissões de UL e DL, otimizando o tempo de transmissão.

A principal desvantagem do FDD é a necessidade de alocação de bandas de frequência simétricas e pareadas para o UL e o DL, o que pode ser ineficiente se o tráfego for assimétrico e representa um desafio devido à escassez de espectro. Adicionalmente, o FDD geralmente exige dois conjuntos de antenas e filtros de radiofrequência para separar os sinais de transmissão e recepção, o que pode aumentar a complexidade e o custo do *hardware*, tanto no satélite quanto no terminal (JAAFARI *et al.*, 2023).

2.2.1.2 TDD

No TDD, os enlaces de UL e DL compartilham a mesma banda de frequência, mas transmitem em instantes de tempo diferentes e não sobrepostos. O TDD é o modo de operação típico em redes de comunicação terrestre, incluindo os padrões desde a terceira geração, do inglês *3th Generation*, até o 5G. Embora raramente aplicado em sistemas de satélite comerciais, o TDD está sendo fortemente considerado para futuras NTN's pelos seguintes motivos (JAAFARI *et al.*, 2023; KANG *et al.*, 2023; ALLIANCE, 2019; SAAD *et al.*, 2024):

- Flexibilidade Espectral: O TDD não exige bandas de frequência pareadas, permitindo uma alocação de espectro mais flexível e eficiente, o que é especialmente vantajoso nas faixas de frequência mais baixas, onde o espectro é mais escasso.

- Flexibilidade de Tráfego: A proporção de tempo alocada para UL e DL pode ser ajustada dinamicamente para acomodar padrões de tráfego assimétricos de forma eficiente.
- *Hardware* Simplificado: O uso de uma única frequência para UL e DL permite que o transmissor e o receptor compartilhem o mesmo conjunto de antenas e cadeia de radio-frequência, do inglês *Radio Frequency*, simplificando o projeto e reduzindo o custo dos terminais e dos satélites.
- Reciprocidade do Canal: Como UL e DL operam na mesma frequência, eles experimentam condições de canal muito semelhantes (reciprocidade do canal). Isso pode ser explorado para simplificar a estimação de canal e otimizar técnicas de pré-codificação.

Apesar de suas vantagens, a aplicação do TDD em NTN enfrenta desafios significativos. O principal deles é a necessidade de um GP entre os *slots* de UL e DL para evitar interferência mútua. Devido aos longos atrasos de propagação em enlaces de satélite, este GP precisa ser correspondentemente longo, o que leva a uma sobrecarga, do inglês *overhead*, de sistema severa e a uma utilização ineficiente dos recursos de rádio. Além disso, a sincronização temporal rigorosa exigida pelo TDD é difícil de alcançar em NTN devido aos grandes atrasos diferenciais entre os usuários (KANG *et al.*, 2023).

2.2.1.3 Comparativo entre FDD e TDD para NTN

A Tabela 1 resume as principais diferenças entre os modos FDD e TDD no contexto de aplicação em NTN.

Em suma, enquanto o FDD oferece uma solução robusta e comprovada para a comunicação via satélite tradicional, o TDD se apresenta como uma tecnologia essencial para a futura integração das NTN com as redes 5G terrestres, desde que seus desafios intrínsecos de sincronização e *overhead* de guarda sejam adequadamente resolvidos (SAAD *et al.*, 2024; KANG *et al.*, 2023).

2.2.2 Sincronização em TDD para NTN

A viabilidade do TDD depende de uma premissa fundamental: a sincronização temporal precisa entre todos os participantes da rede. Em sistemas terrestres, onde as distâncias entre a BS e os terminais são relativamente curtas e homogêneas, alcançar essa sincronia é um problema bem estabelecido e solucionado. No entanto, o cenário NTN introduz uma complexidade drasticamente maior, tornando a sincronização um dos desafios técnicos mais

Tabela 1 – Comparativo entre os modos TDD e FDD para 5G NTN.

Fator Chave	Características do TDD	Características do FDD
Alocação de Frequência	Não necessita de bandas simétricas, o que economiza recursos de frequência.	Requer bandas de frequência simétricas e pareadas.
Arquitetura de Antena	Transmissor e receptor podem compartilhar a mesma antena, simplificando o <i>hardware</i> .	Geralmente requer conjuntos de antenas separados para transmissão e recepção.
Estrutura de Quadro	Exige um GP para separar UL e DL, o que pode desperdiçar recursos, especialmente com longos atrasos.	Não há necessidade de GP entre UL e DL.
Sincronização e Latência	A sincronização é rigorosa e a latência pode ser maior devido ao compartilhamento de tempo e à necessidade de GP.	A sincronização é mais simples e a transmissão contínua pode resultar em menor latência.
Gerenciamento de Interferência	A interferência pode ocorrer entre os sinais de UL e DL, tornando o gerenciamento mais complexo.	A interferência entre UL e DL é inerentemente mitigada pela separação de frequência.

Fonte: Adaptado de Jaafari *et al.* (2023).

críticos para a implementação do TDD via satélite (TRASPADINI *et al.*, 2024; JAAFARI *et al.*, 2023).

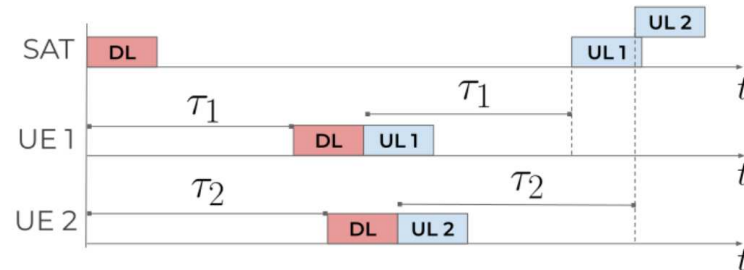
O principal obstáculo reside nos longos e variáveis atrasos de propagação inerentes aos enlaces de satélite. A grande altitude das plataformas NTN, mesmo no caso de satélites LEO, resulta em um RTT que pode se estender por múltiplos *slots* de transmissão, diferentemente do ambiente terrestre, onde o atraso é da ordem de símbolos de OFDM (TRASPADINI *et al.*, 2024). Esse longo atraso, por si só, já impacta toda a temporização do sistema, incluindo os procedimentos de acesso aleatório, agendamento e os mecanismos de retransmissão (KANG *et al.*, 2023).

Contudo, o desafio mais complexo não é apenas o atraso absoluto, mas o atraso diferencial entre os múltiplos usuários servidos por um único feixe de satélite. Devido à vasta área de cobertura, um usuário localizado diretamente abaixo do satélite, perpendicularmente à sua plataforma, isto é, no nadir, experimentará um atraso de propagação, τ_1 , significativamente menor do que um usuário na borda da célula, que terá um atraso τ_2 , com $\tau_2 \gg \tau_1$. Em uma célula com 1.000 km de raio, por exemplo, o atraso de propagação diferencial pode chegar a 10 ms (TRASPADINI *et al.*, 2024).

Essa disparidade temporal tem uma consequência direta e prejudicial para a sincronização do UL. Considere o cenário ilustrado na Figura 7 (TRASPADINI *et al.*, 2024):

1. A estação base do satélite transmite um *slot* de DL em um instante t_0 .
2. O usuário 1 recebe este *slot* no instante $t_0 + \tau_1$, enquanto o usuário 2 o recebe no instante $t_0 + \tau_2$.
3. Se ambos os usuários transmitirem seus slots de UL imediatamente após a recepção do DL, seus sinais chegarão de volta ao satélite em instantes de tempo diferentes: o sinal do usuário 1 chegará em $t_0 + 2\tau_1$ e o do usuário 2 em $t_0 + 2\tau_2$.

Figura 7 – Estrutura de *frame* assíncrona.



Fonte: Adaptado de Traspadini *et al.* (2024).

Essa chegada assíncrona dos sinais de UL ao satélite resulta em transmissões não coordenadas que causam severa interferência mútua, tornando impossível a correta decodificação dos dados. Um sinal de UL de um usuário pode invadir o recurso de tempo alocado para outro, um problema conhecido como colisão de UL (JAAFARI *et al.*, 2023).

Para evitar essa interferência, o sistema TDD deve garantir que os sinais de todos os usuários cheguem ao satélite de forma sincronizada dentro dos seus respectivos recursos de tempo alocados. Isso exige a inserção de um GP suficientemente longo entre as transmissões de DL e UL. Este GP deve ser, no mínimo, igual ao RTT do usuário mais distante da célula para garantir que a transmissão de DL não interfira com a recepção de UL no satélite. A necessidade desse GP, cuja duração é ditada pelos longos atrasos de propagação, é a principal causa da ineficiência do TDD em NTN, pois representa um tempo em que nenhum dado útil pode ser transmitido (TRASPADINI *et al.*, 2024).

Em suma, a sincronização em TDD para redes NTN é um problema multifacetado, impulsionado pelos longos atrasos de propagação e, mais criticamente, pelos atrasos diferenciais entre usuários. A falha em gerenciar essa assincronia compromete a integridade do sistema, enquanto as soluções mais simples, como a inserção de longos períodos de guarda, levam a uma drástica perda de eficiência. Isso estabelece a necessidade de mecanismos de compensação mais

inteligentes, como o TA, que será explorado na próxima seção (TRASPADINI *et al.*, 2024).

2.2.3 TA

Para solucionar o problema da chegada assíncrona dos sinais de UL em redes TDD via satélite, o 3GPP introduziu um mecanismo de compensação de atraso conhecido como TA (JAAFARI *et al.*, 2023; KANG *et al.*, 2023; TRASPADINI *et al.*, 2024). O objetivo do TA é garantir que os sinais transmitidos por todos os UEs, independentemente de suas localizações e respectivos atrasos de propagação, cheguem à estação base do satélite de forma perfeitamente sincronizada, cada um em seu recurso de tempo designado.

O funcionamento do mecanismo de TA em um cenário NTN é um processo cooperativo entre o UE e a rede, que pode ser descrito nos seguintes passos (TRASPADINI *et al.*, 2024):

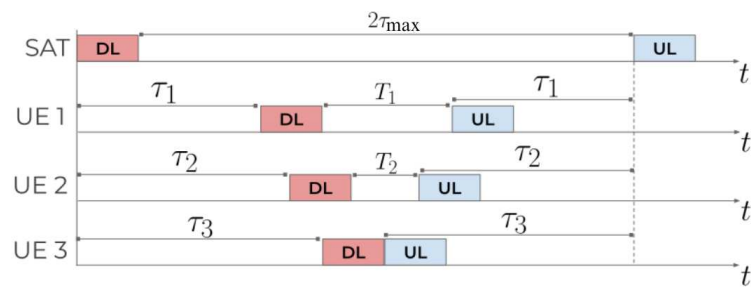
1. **Estimação do Atraso de Propagação:** O processo requer que cada UE tenha conhecimento tanto de sua própria posição quanto da posição do satélite. Esta informação é tipicamente obtida através de sistemas de navegação, como o *Global Navigation Satellite System*. Com esses dados, cada UE j é capaz de estimar seu próprio atraso de propagação unidirecional, τ_j .
2. **Determinação do Atraso de Referência:** Os UEs podem reportar suas estimativas de atraso para o satélite. A rede, por sua vez, calcula o maior atraso de propagação entre todos os usuários dentro de sua área de cobertura. Este valor máximo, denotado como τ_{\max} , serve como um ponto de referência temporal para toda a célula. A informação sobre τ_{\max} é então comunicada a todos os UEs, por exemplo, através do canal físico de controle de DL, do inglês *Physical Downlink Control Channel*.
3. **Aplicação do Avanço no Tempo:** Com base no seu próprio atraso τ_j e no atraso máximo de referência τ_{\max} , cada UE j calcula um valor de avanço no tempo, T_j , dado pela seguinte equação:

$$T_j = 2(\tau_{\max} - \tau_j). \quad (2.1)$$

Este valor T_j representa o quanto o UE j deve atrasar a transmissão do seu sinal de UL em relação a um ponto de referência comum. Na prática, o UE não compensa todo o atraso de propagação, mas sim a diferença de tempo de ida e volta entre ele e o usuário mais distante da célula, em um método conhecido como TA diferencial.

O efeito da aplicação do TA é uma transformação fundamental na estrutura temporal do canal de UL. Ao fazer com que cada UE atrase sua transmissão pelo valor exato de T_j , os sinais de todos os usuários, que partem de pontos diferentes e em instantes físicos distintos, convergem para o receptor do satélite de forma perfeitamente alinhada, como ilustrado na Figura 8 (TRASPADINI *et al.*, 2024). Essa sincronização permite que a BS do satélite aloque diferentes recursos, como símbolos OFDM, para múltiplos usuários dentro do mesmo *slot* de UL sem que haja interferência mútua (KANG *et al.*, 2023).

Figura 8 – Estrutura de *frame* síncrono com aplicação do TA.



Fonte: Adaptado de Traspadini *et al.* (2024).

Apesar de resolver eficazmente o problema de desalinhamento e colisão no UL, a implementação do TA impõe uma nova e rígida restrição ao sistema. Para que o mecanismo funcione e para evitar que o sinal de DL interfira com um sinal de UL transmitido antecipadamente, um GP entre um *slot* de DL e o seu correspondente *slot* de UL deve agora ser igual a $2\tau_{\max}$, ou seja, o RTT do usuário mais distante da célula (TRASPADINI *et al.*, 2024), essa consequência direta do TA é a principal fonte de ineficiência em sistemas TDD para NTN (KANG *et al.*, 2023).

A seguir, são apresentados os principais conceitos de MA pertinentes a esta dissertação.

2.3 Introdução ao MA

O desenvolvimento de técnicas de MA sempre foi um elemento central na evolução das diferentes gerações das tecnologias de comunicação celular (MATHUR H., 2021). As técnicas de MA determinam como os escassos recursos do espectro de rádio são compartilhados entre múltiplos usuários. As gerações anteriores de redes celulares adotaram predominantemente esquemas OMA, como o FDMA, acesso múltiplo por divisão de tempo, do inglês *Time Division Multiple Access* (TDMA), acesso múltiplo por divisão de frequência ortogonal, do inglês

Orthogonal Frequency-Division Multiple Access (OFDMA), e o acesso múltiplo por divisão de código, do inglês *Code Division Multiple Access* (CDMA) (CAI *et al.*, 2018). A característica unificadora dessas técnicas é a alocação de recursos em domínios ortogonais — seja no tempo, na frequência ou em códigos — de modo que os sinais de diferentes usuários sejam ortogonais no receptor. Essa ortogonalidade simplifica significativamente a detecção do sinal, pois, idealmente, minimiza ou elimina a interferência entre usuários adjacentes, um princípio que foi a base para o projeto de redes sem fio até a quarta geração, do inglês *4th Generation* (4G) (SHAH, 2022).

De acordo com Shah (2022), o OFDMA, em particular, foi amplamente utilizado em sistemas de banda larga e continua a ser um componente relevante para as futuras gerações de sistemas celulares, oferecendo flexibilidade no espaçamento de subportadoras e na duração dos *slots* de tempo para acomodar diversos casos de uso. Contudo, o paradigma da ortogonalidade, apesar de sua simplicidade e robustez, apresenta limitações intrínsecas que se tornam críticas diante das novas e exigentes demandas das redes 5G e além. A principal restrição do OMA reside no número limitado de usuários que podem ser suportados simultaneamente, uma vez que este número é diretamente atrelado à quantidade de blocos de recursos, do inglês *Resource Blocks* (RBs), ortogonais disponíveis. Essa limitação impõe um teto à eficiência espectral e à capacidade geral da rede, tornando o OMA insuficiente para o cenário futuro das comunicações.

A transição para o 5G foi marcada por uma mudança de paradigma nos requisitos de serviço, e a visão para o 6G expande ainda mais esses horizontes. As redes futuras devem suportar um ecossistema heterogêneo de aplicações, cada uma com necessidades distintas e, por vezes, conflitantes. No contexto do 5G, o 3GPP delineou cenários de uso chave, como largura de banda móvel melhorada, do inglês *Enhanced Mobile Broadband*, comunicações massivas do tipo máquina, do inglês *Massive Machine Type Communications*, e URLLC (MATHUR H., 2021). Atender a essas demandas significa alcançar metas ambiciosas, como picos de taxa de dados da ordem de 10 Gbps, latência inferior a 1 ms e uma densidade de conexão de até um milhão de dispositivos por quilômetro quadrado. Fica evidente que um sistema que aloca um recurso ortogonal exclusivo para cada usuário não consegue, de forma eficiente, escalar para suportar tal densidade de conexões massivas ou prover a flexibilidade necessária para os diferentes requisitos de QoS.

É neste contexto que a pesquisa e o desenvolvimento de novos esquemas de MA se tornam não apenas importantes, mas imperativos (MATHUR H., 2021). Para superar as limitações do OMA, a comunidade de pesquisa voltou-se para duas direções principais: o domínio

da potência e o domínio espacial. No domínio da potência, o NOMA emergiu como um candidato promissor. O NOMA permite que múltiplos usuários compartilhem os mesmos recursos de tempo e frequência, distinguindo-os pela alocação de diferentes níveis de potência. Através da Codificação por Superposição no transmissor e do *Successive Interference Cancellation* (SIC) no receptor, o NOMA pode melhorar significativamente a eficiência espectral e a conectividade massiva. Paralelamente, no domínio espacial, o SDMA, implementado principalmente via sistemas de múltiplas antenas na entrada e na saída, do inglês *Multiple-Input Multiple-Output* (MIMO), explora a separabilidade espacial dos usuários para servi-los simultaneamente nos mesmos recursos de tempo-frequência, utilizando pré-codificação para direcionar os sinais e mitigar a interferência (SHAH, 2022).

Apesar de seus méritos, tanto o NOMA quanto o SDMA apresentam um desempenho ótimo apenas sob condições específicas (CLERCKX *et al.*, 2020). O NOMA, por exemplo, é mais eficaz quando os usuários apresentam uma grande disparidade em suas condições de canal, enquanto o SDMA funciona melhor quando os canais dos usuários são ortogonais ou fracamente correlacionados. Diante dessas limitações, o RSMA surge como uma estrutura mais poderosa, geral e flexível (MISHRA *et al.*, 2022). O RSMA revoluciona o gerenciamento de interferência ao dividir a mensagem de cada usuário em uma parte comum, que é decodificada por todos, e uma parte privada, destinada apenas ao usuário específico. Essa abordagem permite que o RSMA trate a interferência de forma híbrida e adaptativa: parte dela é decodificada e parte é tratada como ruído. Essa flexibilidade intrínseca permite ao RSMA englobar o OMA, NOMA e o SDMA como casos particulares de sua operação, adaptando-se dinamicamente a uma gama muito mais ampla de cenários de canal e topologias de rede. Por essa capacidade de unificar e superar os esquemas de MA existentes, o RSMA representa uma fronteira de pesquisa fundamental e é, portanto, o foco central desta dissertação.

A seguir são descritas as principais técnicas de referência que formam a base para o desenvolvimento de esquemas mais avançados, como OMA, SDMA e NOMA.

2.3.1 Esquemas de MA de referência

2.3.1.1 OMA

O OMA foi o paradigma dominante nas gerações de redes celulares até o 4G. O princípio central do OMA é a alocação de recursos ortogonais e exclusivos para cada usuário,

garantindo que, teoricamente, não haja interferência mútua entre eles durante o compartilhamento do meio sem fio. Essa separação pode ocorrer em diferentes domínios (SHAH, 2022):

- FDMA: O espectro de frequência disponível é dividido em canais de frequência menores e não sobrepostos. Cada usuário recebe um canal de frequência exclusivo para sua comunicação, eliminando a interferência co-canal.
- TDMA: Neste esquema, os usuários compartilham a mesma banda de frequência, mas transmitem em instantes de tempo diferentes. O tempo é dividido em *slots*, e cada usuário tem acesso ao canal completo durante seu *slot* de tempo designado.
- CDMA: O CDMA permite que múltiplos usuários transmitam simultaneamente na mesma frequência e ao mesmo tempo. A ortogonalidade é alcançada no domínio do código, onde cada usuário recebe uma sequência de espalhamento (ou código) única e ortogonal. No receptor, o sinal desejado é recuperado correlacionando o sinal recebido com o código do usuário específico, enquanto os sinais dos outros usuários, por serem ortogonais, são tratados como ruído.
- OFDMA: É a técnica de acesso utilizada no DL do 4G *Long Term Evolution* e constitui a base para o 5G NR. O OFDMA combina o OFDM com uma estratégia de MA. O espectro é dividido em um grande número de subportadoras ortogonais, que são agrupadas em RBs. Esses RBs são então alocados dinamicamente a diferentes usuários em diferentes *slots* de tempo, proporcionando alta granularidade e flexibilidade na alocação de recursos.

A principal vantagem do OMA é a simplicidade na detecção do sinal no receptor, uma vez que a interferência entre usuários é minimizada pela alocação de recursos ortogonais. No entanto, sua maior desvantagem reside na limitação do número de usuários que podem ser servidos simultaneamente, que é restrito ao número total de RBs ortogonais disponíveis. Essa limitação intrínseca restringe a eficiência espectral e a capacidade da rede, tornando o OMA inadequado para atender às demandas de conectividade massiva das redes 5G e além (SHAH, 2022).

2.3.1.2 SDMA

O SDMA explora uma nova dimensão para o acesso múltiplo: o espaço. Diferente do OMA, o SDMA permite que vários usuários compartilhem o mesmo recurso de tempo-frequência, separando-os espacialmente. A implementação do SDMA depende fundamentalmente de sistemas MIMO, especialmente em sua configuração para múltiplos usuários (MATHUR H., 2021).

O princípio de funcionamento do SDMA envolve o uso de um conjunto de antenas na estação base para criar feixes de transmissão altamente direcionados. Cada feixe pode ser apontado para um usuário ou um grupo de usuários geograficamente distintos, permitindo transmissões simultâneas e paralelas. Para que isso seja eficaz, a BS emprega técnicas avançadas de processamento de sinal, como pré-codificadores, para conformar os feixes e anular a interferência entre os usuários. O SDMA é particularmente eficaz quando os canais dos usuários são espacialmente distintos e apresentam baixa correlação, permitindo que a estação base os separe de forma eficiente. Embora aumente drasticamente a eficiência espectral, o SDMA exige conhecimento preciso do CSIT no transmissor e aumenta a complexidade computacional (MATHUR H., 2021).

2.3.1.3 NOMA

O NOMA também representa uma mudança de paradigma fundamental em relação ao OMA (MATHUR H., 2021). Semelhantemente ao SDMA, o NOMA permite que múltiplos usuários compartilhem os mesmos elementos de recurso (tempo, frequência e código) de forma não ortogonal. A separação dos sinais dos usuários é então realizada no receptor através de técnicas avançadas. O NOMA é classificado principalmente em duas categorias (SHAH, 2022):

- NOMA no domínio da potência, do inglês *Power-domain Non-Orthogonal Multiple Access* (PD-NOMA): É a forma mais proeminente de NOMA. Nela, os usuários são multiplexados no mesmo recurso de tempo-frequência, mas com diferentes níveis de potência alocados. O transmissor utiliza a Codificação por Superposição para combinar os sinais de múltiplos usuários em um único sinal composto. No receptor, o SIC é empregado para decodificar os sinais. Uma característica chave é a alocação de potência: usuários com piores condições de canal recebem mais potência, enquanto usuários com melhores canais recebem menos potência. Isso permite que o usuário com melhor canal decodifique primeiro o sinal do usuário com pior canal (que tem mais potência), subtraia-o do sinal recebido e, em seguida, decodifique seu próprio sinal. Esse método melhora a equidade entre os usuários e aumenta a capacidade geral da célula (SHAH, 2022).
- NOMA no domínio do código, do inglês *Code-domain Non-Orthogonal Multiple Access*: Nesta variante, a sobreposição de múltiplos usuários no mesmo recurso de tempo-frequência é gerenciada através da atribuição de sequências de espalhamento ou códigos não-ortogonais de baixa correlação cruzada. Exemplos incluem Acesso Múltiplo por código esparsos, do inglês *Sparse Code Multiple Access*, CDMA de baixa densidade de

espalhamento, do inglês *Low-Density Spreading CDMA*, e acesso múltiplo por divisão de padrão, do inglês *Pattern Division Multiple Access*. Essas técnicas permitem um sistema sobrecarregado, onde o número de usuários ativos pode exceder o número de recursos ortogonais disponíveis, o que é ideal para cenários de conectividade massiva. A decodificação é geralmente realizada por receptores de baixa complexidade baseados em algoritmos de passagem de mensagens (SHAH, 2022).

O NOMA e o SDMA são, portanto, as principais técnicas de referência que buscam superar as limitações do OMA tradicional, servindo como base para o desenvolvimento de esquemas ainda mais avançados e flexíveis, como o RSMA.

2.3.2 Fundamentos do RSMA

O RSMA é uma estratégia de acesso múltiplo e gerenciamento de interferência que generaliza e supera os esquemas clássicos, como SDMA e PD-NOMA (CLERCKX *et al.*, 2020). O princípio de projeto do RSMA é decodificar parcialmente a interferência e tratar parcialmente a interferência como ruído. Essa abordagem híbrida permite ao RSMA uma flexibilidade sem precedentes para se adaptar às condições do canal e à carga da rede, preenchendo a lacuna entre as duas abordagens extremas de gerenciamento de interferência, isto é, tratar toda a interferência como ruído (característico do SDMA) e decodificar completamente a interferência (característico do PD-NOMA) (MISHRA *et al.*, 2022).

Antes de prosseguir, é importante compreender que o RSMA pode ser implementado em diferentes formas, notavelmente como RSMA de uma camada, do inglês *1-Layer RSMA*, ou de múltiplas camadas, do inglês *Multi-Layer RSMA* (MISHRA *et al.*, 2022). O *1-layer RSMA*, que é a base para a maioria das discussões e a estrutura mais fundamental, consiste em dividir as mensagens dos usuários e combinar as partes comuns em um único *stream* comum, que é transmitido juntamente com os *streams* privados. Estratégias de *multi-layer RSMA* também foram extensivamente investigadas e demonstraram superar o SDMA e o NOMA em termos de eficiência espectral (AHMAD *et al.*, 2020). Essas estratégias mais avançadas geram múltiplos *streams* comuns hierárquicos, cada um destinado a um subconjunto diferente de usuários, o que exige múltiplas camadas de SIC no receptor. Embora promissor, este aumento de desempenho vem à custa de uma maior complexidade no transmissor e no receptor (MISHRA *et al.*, 2022). Este trabalho foca na implementação do *1-layer RSMA*.

Além disso, vale ressaltar que o RSMA se baseia em princípios semelhantes tanto no

DL quanto no UL. No entanto, as motivações diferem: no DL, o objetivo principal é aprimorar a eficiência espectral, enquanto no UL, busca-se atingir a região de capacidade do canal de acesso múltiplo sem a necessidade de compartilhamento de tempo, do inglês *time sharing*. Os estudos desenvolvidos neste trabalho focam na aplicação do RSMA no DL. A seguir, apresenta-se o detalhamento matemático do RSMA no DL.

2.3.2.1 RSMA no DL

No cenário de DL, considerando-se um canal de *broadcast* com múltiplas antenas na entrada e antena única na saída, do inglês *Multiple-Input Single-Output* (MISO), onde uma BS equipada com M antenas atende a J usuários de antena única, indexados pelo conjunto $\mathcal{J} = \{1, \dots, J\}$, o processo de transmissão no RSMA de uma camada pode ser descrito da seguinte forma:

1. Divisão da Mensagem: A mensagem W_j destinada ao usuário $j \in \mathcal{J}$ é dividida em uma parte comum, $W_{c,j}$, e uma parte privada, $W_{p,j}$.
2. Codificação dos *Streams*: As partes comuns de todos os usuários, $\{W_{c,1}, \dots, W_{c,J}\}$, são combinadas para formar uma única mensagem comum W_c , que é então codificada em um *stream* de dados comum $s_c \in \mathbb{C}$. Por outro lado, a parte privada $W_{p,j}$ de cada usuário é codificada de forma independente em um *stream* de dados privado $s_{p,j} \in \mathbb{C}$.
3. Pré-codificação Linear: Pré-codificadores lineares são aplicados aos *streams*. O *stream* comum s_c é multiplicado pelo vetor de pré-codificação $\mathbf{p}_c \in \mathbb{C}^{M \times 1}$, e cada *stream* privado $s_{p,j}$ é multiplicado por seu respectivo pré-codificador $\mathbf{p}_{p,j} \in \mathbb{C}^{M \times 1}$. O sinal transmitido $\mathbf{x} \in \mathbb{C}^{M \times 1}$ é a superposição de todos os *streams* pré-codificados:

$$\mathbf{x} = \mathbf{p}_c s_c + \sum_{j=1}^J \mathbf{p}_{p,j} s_{p,j}. \quad (2.2)$$

Este sinal está sujeito a uma restrição de potência total, $\mathbb{E}\{\|\mathbf{x}\|^2\} \leq P$.

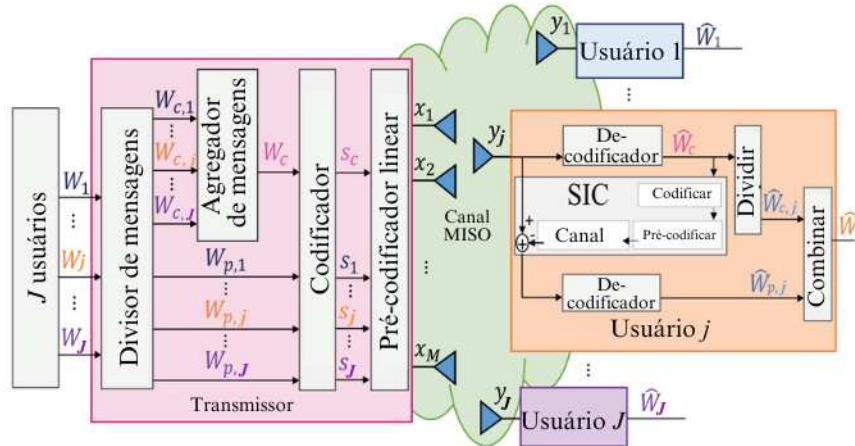
4. Recepção e Decodificação: O sinal recebido pelo usuário j , denotado por $y_j \in \mathbb{C}$, é dado por:

$$y_j = \mathbf{h}_j^H \mathbf{x} + n_j, \quad (2.3)$$

onde $\mathbf{h}_j \in \mathbb{C}^{M \times 1}$ é o vetor de canal entre a BS e o usuário j , e $n_j \sim \mathcal{CN}(0, \sigma^2)$ é o ruído branco gaussiano aditivo, do inglês *Additive White Gaussian Noise* (AWGN). A

Figura 9 ilustra de forma sintetizada o processo de aplicação do RSMA em sistemas de comunicação MISO em DL.

Figura 9 – Ilustração do processo de aplicação do RSMA em comunicação MISO em DL.



Fonte: Adaptado de Mishra *et al.* (2022).

No receptor, o usuário j primeiro decodifica o *stream* comum s_c , tratando todos os *streams* privados como interferência. A razão sinal-interferência mais ruído, do inglês *Signal-to-Interference-Plus-Noise Ratio* (SINR), para o *stream* comum no usuário j é:

$$\gamma_{c,j} = \frac{|\mathbf{h}_j^H \mathbf{p}_c|^2}{\sum_{i=1}^J |\mathbf{h}_j^H \mathbf{p}_{p,i}|^2 + \sigma^2}. \quad (2.4)$$

Após decodificar s_c com sucesso, o usuário j aplica o SIC para remover a contribuição do *stream* comum do sinal recebido. Em seguida, o usuário j decodifica seu próprio *stream* privado $s_{p,j}$, tratando os *streams* privados dos outros usuários como interferência residual. A SINR para o *stream* privado do usuário j é:

$$\gamma_{p,j} = \frac{|\mathbf{h}_j^H \mathbf{p}_{p,j}|^2}{\sum_{i=1, i \neq j}^J |\mathbf{h}_j^H \mathbf{p}_{p,i}|^2 + \sigma^2}. \quad (2.5)$$

Como o *stream* comum deve ser decodificado por todos os usuários, sua taxa máxima alcançável é limitada pelo usuário com a pior SINR, sendo $R_c = \log_2(1 + \min_{j \in \mathcal{J}} \{\gamma_{c,j}\})$. Esta taxa comum é então dividida entre os usuários ($R_c = \sum_{j=1}^J C_j$), onde C_j é a porção da taxa destinada ao usuário j . A taxa do *stream* privado é $R_{p,j} = \log_2(1 + \gamma_{p,j})$. Portanto, a taxa total para o usuário j é a soma da sua porção da taxa comum com sua taxa privada: $R_j = C_j + R_{p,j}$ (MISHRA *et al.*, 2022).

Compreendidos os princípios básicos do RSMA no DL, a seguir são apresentados os principais conceitos de pré-codificação pertinentes a esta dissertação.

2.4 Pré-codificação linear

O desempenho efetivo de qualquer esquema de MA, especialmente em sistemas MISO/MIMO, não depende apenas da alocação de recursos em alto nível, mas fundamentalmente de como os sinais são moldados, transmitidos e gerenciados no domínio físico (NGO, 2015). A maximização de métricas de desempenho chave, como a taxa de transmissão ou a eficiência energética, se traduz em um complexo problema de otimização, cujas variáveis são precisamente as estratégias de transmissão.

Para tanto, esta seção aborda o conceito de Pré-codificação Linear, que é o processamento de sinal realizado na BS antes da transmissão. Utilizando o *Channel State Information* (CSI), a pré-codificação linear molda espacialmente os sinais para direcionar a energia aos usuários desejados e, ao mesmo tempo, mitigar a interferência causada aos outros. Enquanto técnicas ótimas podem ser complexas, a pré-codificação linear oferece um balanço entre desempenho e complexidade computacional, sendo quase ótima em muitos cenários (NGO, 2015; ARTIGA; VÁZQUEZ, 2022).

A seguir, serão explorados os principais esquemas de pré-codificação linear, como o de transmissão de razão máxima, do inglês *Maximum-Ratio Transmission* (MRT), de forçamento a zero, do inglês *Zero-Forcing* (ZF), e o de erro quadrático médio mínimo, do inglês *Minimum Mean-Square Error* (MMSE), que são centrais para o projeto dos sistemas analisados nesta dissertação.

2.4.1 Pré-codificação em sistemas MISO

De acordo com Ngo (2015), em sistemas de comunicação MISO, a pré-codificação é uma técnica de processamento de sinal fundamental, empregada no enlace de DL. Seu principal objetivo é gerenciar a interferência multiusuário que surge quando a BS atende a múltiplos usuários simultaneamente sobre o mesmo recurso de tempo-frequência. A pré-codificação utiliza o CSI, disponível no transmissor, para processar os sinais de dados antes da transmissão, de forma a maximizar a energia do sinal no usuário desejado e, idealmente, minimizar a interferência gerada para os outros usuários.

Em um sistema MISO com M antenas na BS servindo J usuários, o vetor de símbolos a ser transmitido, $\mathbf{s} \in \mathbb{C}^{J \times 1}$, onde s_j é o símbolo destinado ao usuário j , é multiplicado por uma matriz de pré-codificação $\mathbf{P} \in \mathbb{C}^{M \times J}$. O sinal vetorial resultante $\mathbf{x} \in \mathbb{C}^{M \times 1}$ a ser transmitido pelas M antenas é dado por:

$$\mathbf{x} = \sqrt{\alpha} \mathbf{P} \mathbf{s}, \quad (2.6)$$

onde α é uma constante de normalização para satisfazer a restrição de potência total de transmissão. O sinal recebido no usuário j é, então, uma superposição do sinal desejado, da interferência dos outros usuários e do ruído:

$$y_j = \underbrace{\sqrt{\alpha P_{\max}} \mathbf{h}_j^H \mathbf{p}_j s_j}_{\text{Sinal Desejado}} + \underbrace{\sqrt{\alpha P_{\max}} \sum_{j' \neq j}^J \mathbf{h}_j^H \mathbf{p}_{j'} s_{j'}}_{\text{Interferência Multiusuário}} + \underbrace{n_j}_{\text{Ruído}}, \quad (2.7)$$

onde P_{\max} é a potência de transmissão, $\mathbf{h}_j \in \mathbb{C}^{M \times 1}$ é o vetor de canal entre a BS e o usuário j , \mathbf{p}_j é a j -ésima coluna da matriz \mathbf{P} , e n_j é o ruído aditivo. A SINR no usuário j define a qualidade da comunicação:

$$\text{SINR}_j = \frac{\alpha P_{\max} |\mathbf{h}_j^H \mathbf{p}_j|^2}{\alpha P_{\max} \sum_{i=1, i \neq j}^J |\mathbf{h}_j^H \mathbf{p}_i|^2 + 1}. \quad (2.8)$$

A escolha da matriz de pré-codificação \mathbf{P} determina como o sistema lida com o balanço entre maximizar o termo do sinal desejado e minimizar o termo de interferência. A seguir, são descritos os principais esquemas de pré-codificação linear.

2.4.1.1 MRT

O MRT, também conhecida como *conjugate beamforming*, é a abordagem de pré-codificação linear mais simples. Seu único objetivo é maximizar a potência do sinal recebido no usuário alvo, ignorando completamente a interferência que isso possa causar a outros usuários. Para alcançar isso, o pré-codificador simplesmente alinha o sinal transmitido com o conjugado do canal do usuário. A matriz de pré-codificação MRT é, portanto, a transposta Hermitiana (conjugada transposta) da matriz de canal $\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_J]$ (NGO, 2015):

$$\mathbf{P}_{\text{MRT}} = \mathbf{H}^{\text{H}}. \quad (2.9)$$

A principal vantagem do MRT é sua baixíssima complexidade computacional, pois não requer inversões de matriz e pode ser implementado de forma distribuída. Funciona bem em cenários de baixa razão sinal-ruído, do inglês *Signal-to-Noise Ratio* (SNR), onde o ruído é o fator limitante e a interferência multiusuário é desprezível. Contudo, seu desempenho é severamente limitado pela interferência em regimes de alta SNR, pois a SINR satura e não aumenta com o aumento da potência de transmissão.

2.4.1.2 ZF

Em contraste com o MRT, o ZF tem como objetivo principal eliminar completamente a interferência multiusuário. O ZF projeta o sinal de cada usuário de forma a ser ortogonal aos canais de todos os outros usuários, garantindo que o termo de interferência na equação da SINR seja nulo. A matriz de pré-codificação ZF é a pseudo-inversa da matriz de canal:

$$\mathbf{P}_{\text{ZF}} = \mathbf{H}(\mathbf{H}^{\text{H}}\mathbf{H})^{-1}. \quad (2.10)$$

Para que a matriz $\mathbf{H}^{\text{H}}\mathbf{H}$ seja invertível, o número de antenas na BS deve ser maior ou igual ao número de usuários ($M \geq J$). A grande vantagem do ZF é seu excelente desempenho em cenários limitados por interferência (alto SNR), onde a SINR pode crescer linearmente com a potência de transmissão. Sua principal desvantagem é a inversão de matriz, que aumenta a complexidade computacional. Além disso, ao focar exclusivamente em anular a interferência, o ZF ignora o impacto do ruído. Se os canais dos usuários forem pouco condicionados (próximos de serem colineares), a inversão da matriz pode amplificar significativamente o ruído, degradando o desempenho em regimes de baixo SNR (NGO, 2015).

2.4.1.3 MMSE

O MMSE representa um balanço otimizado entre os objetivos do MRT e do ZF. O MMSE busca minimizar o erro quadrático médio entre os símbolos transmitidos e os sinais estimados nos receptores, o que, na prática, corresponde a maximizar a SINR de cada usuário,

considerando simultaneamente o sinal desejado, a interferência e o ruído. A matriz de pré-codificação MMSE é uma forma regularizada da pseudo-inversa:

$$\mathbf{P}_{\text{MMSE}} = \left[\mathbf{H}^H \mathbf{H} + \frac{J}{P_{\max}} \mathbf{I}_M \right]^{-1} \mathbf{H}^H. \quad (2.11)$$

O termo de regularização, $\frac{J}{P_{\max}} \mathbf{I}_M$, onde \mathbf{I}_M é a matriz identidade de dimensão M , representa a contribuição do ruído. Esse termo permite que o pré-codificador MMSE se comporte de maneira adaptativa:

- Em baixo SNR, quando P_{\max} é pequeno, o termo de regularização domina, e o pré-codificador MMSE se aproxima do MRT, focando em maximizar a potência do sinal contra o ruído.
- Em alto SNR, quando P_{\max} é grande, o termo de regularização se torna desprezível, e o pré-codificador MMSE converge para o ZF, focando em eliminar a interferência.

Devido a essa adaptabilidade, o MMSE é considerado o pré-codificador linear ótimo, oferecendo o melhor desempenho em toda a faixa de SNR em comparação com MRT e ZF. Sua desvantagem é a maior complexidade computacional, que envolve a inversão de matriz e o conhecimento da potência de ruído, isto é, SNR (NGO, 2015).

3 ALOCAÇÃO DE POTÊNCIA PARA MENSAGENS COMUM E PRIVADAS EM REDES RSMA LEO-TERRESTRE

Este capítulo apresenta a primeira contribuição principal desta dissertação, dedicada à análise e à otimização da alocação de potência entre as mensagens comum e privadas em sistemas RSMA LEO-terrestre. Inicialmente, na Seção 3.1 são detalhados o modelo de sistema, canal e transmissão adotados, destacando as particularidades do ambiente LEO e a modelagem de CSIT imperfeito. Em seguida, na Seção 3.2 é formulado o problema de otimização para maximizar a soma das taxas dos usuários, sob restrição de potência total. Para solucionar o problema, na Seção 3.3 propõe-se uma abordagem de baixa complexidade, adaptada de soluções da literatura para o cenário NTN, que combina a pré-codificação ZF para as mensagens privadas com uma solução max-min para o *stream* comum, além de apresentar uma expressão em forma fechada para a divisão de potência. Em seguida, na Seção 3.4 são apresentados os resultados numéricos que validam a eficácia da abordagem proposta em diferentes cenários, incluindo variação de distância e imperfeição de CSIT, destacando a superioridade em relação às soluções de referência existentes na literatura. Por fim, na Seção 3.5 são apresentadas as considerações parciais deste capítulo.

3.1 Modelo de sistema

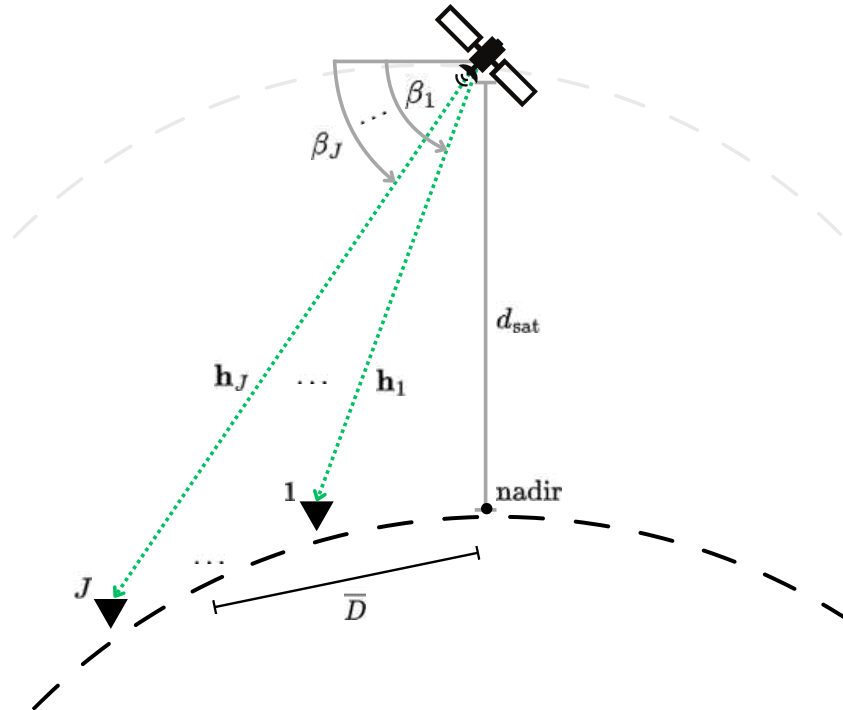
Considera-se um sistema de comunicação via enlace de DL utilizando a técnica RSMA em um canal com linha de visada, do inglês *Line-of-Sight* (LoS), operando a partir de um satélite LEO. O satélite é equipado com um arranjo linear uniforme, do inglês *Uniform Linear Array* (ULA), composto por M antenas, cada uma com ganho G_{sat} , atendendo a J usuários com antena única e ganho G_{ue} , conforme ilustrado na Figura 10. O ângulo de saída, do inglês *Angle of Departure* (AoD), do arranjo para o usuário $j \in J$ é representado por β_j . A distância mínima entre o satélite e a superfície terrestre é denotada por d_{sat} , e os usuários estão uniformemente distribuídos ao redor de um centro de célula localizado a uma distância \bar{D} do nadir do satélite. O canal de LoS entre o satélite e o usuário j é representado por $\mathbf{h}_j \in \mathbb{C}^{M \times 1}$. O sinal recebido pelo usuário j pode ser expresso como:

$$y_j = \mathbf{h}_j^H \mathbf{x} + n_j = \sum_{m=1}^M h_{j,m} x_m + n_j, \quad (3.1)$$

onde $\mathbf{x} \in \mathbb{C}^{M \times 1}$ representa o sinal pré-codificado, x_m é o seu m -ésimo elemento, $h_{j,m}$ corresponde

à resposta do canal entre a antena m e o usuário j , e $n_j \sim \mathcal{CN}(0, \sigma^2)$ representa o AWGN.

Figura 10 – Ilustração de um sistema de comunicação LEO-Terrestre empregando RSMA, destacando o AoD.



Fonte: Autoral.

3.1.1 Modelo de canal

Adota-se o modelo de canal com LoS apresentado em (SCHRÖDER *et al.*, 2023), no qual a resposta entre a m -ésima antena do satélite e o usuário j é dada por:

$$h_{j,m}(\cos(\beta_j)) = \sqrt{G_{ue}G_{sat}} \frac{\lambda_{port}}{4\pi d_j} e^{-i\omega_j} a_{j,m}(\cos(\beta_j)), \quad (3.2)$$

onde d_j representa a distância entre o satélite e o usuário j , λ_{port} é o comprimento de onda da portadora, $\omega_j \in [0, 2\pi]$ corresponde ao defasamento total, e $a_{j,m}(\cos(\beta_j))$ representa o defasamento relativo, definido como:

$$a_{j,m}(\cos(\beta_j)) = \exp^{-i\pi \frac{d_a}{\lambda_{port}} (M+1-2m) \cos(\beta_j)}, \quad (3.3)$$

sendo d_a a distância entre os elementos da ULA. Definindo o vetor de direcionamento como $\mathbf{a}_j(\cos(\beta_j)) = [a_{j,1}(\cos(\beta_j)), \dots, a_{j,M}(\cos(\beta_j))]^T$, o vetor de resposta do canal pode ser expresso por:

$$\mathbf{h}_j(\cos(\beta_j)) = \sqrt{G_{\text{ue}}G_{\text{sat}}}\frac{\lambda_{\text{port}}}{4\pi d_j}e^{-i\omega_j}\mathbf{a}_j(\cos(\beta_j)). \quad (3.4)$$

Devido aos altos deslocamentos Doppler característicos de sistemas LEO, considera-se um canal com CSIT imperfeito, modelada como uma perturbação no cosseno do AoD, dada por uma variável aleatória $\varepsilon_j \sim \mathcal{U}(-\varepsilon, +\varepsilon)$ (SCHRÖDER *et al.*, 2023). Assim, o canal estimado pode ser representado como:

$$\mathbf{h}'_j(\cos(\beta_j)) = \mathbf{h}_j(\cos(\beta_j)) \circ \mathbf{a}_j(\varepsilon_j), \quad (3.5)$$

onde \circ denota o produto de Hadamard. Ressalta-se que a qualidade do CSIT impacta diretamente o *design* dos pré-codificadores, influenciando, portanto, o desempenho do esquema RSMA, o qual é abordado na próxima seção.

3.1.2 Modelo de transmissão

No enlace de DL considerado, adota-se o RSMA em sua forma *1-layer*, conforme detalhado na Seção 2.3.2. Nesse esquema, cada mensagem de usuário é dividida em uma parte comum e uma parte privada, possibilitando maior flexibilidade no gerenciamento da interferência multiusuário, especialmente em cenários com CSIT imperfeito.

O sinal transmitido pelo satélite corresponde à superposição linear do *stream* comum e dos *streams* privados, sendo cada um associado a seu respectivo vetor de pré-codificação. O procedimento de decodificação no receptor segue a estratégia do SIC, em que o *stream* comum é decodificado primeiro, seguido da decodificação dos *streams* privados.

As equações gerais que regem o cálculo das taxas, os modelos de SINR e a composição final das taxas totais dos usuários estão descritas em detalhe na Seção 2.3.2, e são diretamente utilizadas neste trabalho. A adaptação central no presente modelo consiste na consideração de um canal com CSIT imperfeito, o que influencia o *design* dos pré-codificadores e a estratégia de alocação de potência entre as mensagens comum e privadas, conforme será aprofundado nas seções subsequentes.

Na próxima seção, apresenta-se a formulação do problema de maximização da taxa total no sistema baseado em RSMA.

3.2 Formulação do problema de otimização

O objetivo deste trabalho é maximizar a soma das taxas de dados no sistema de comunicação por enlace de DL baseado em RSMA, conforme descrito nas seções anteriores. As variáveis de otimização consideradas são os vetores de pré-codificação e a alocação de potência entre os *streams* comum e privados. A única restrição considerada no problema é à potência total de transmissão.

A formulação matemática do problema é dada por:

$$\max_{\mathbf{p}, s} \sum_{j=1}^J R_j, \quad (3.6)$$

sujeito a:

$$\|\mathbf{p}\|_2^2 \leq P_{\max}, \quad (3.7)$$

onde $\mathbf{p} \in \mathbb{C}^{(J+1)M \times 1} = [(\mathbf{p}_1)^T, \dots, (\mathbf{p}_J)^T, (\mathbf{p}_c)^T]^T$ representa o vetor agregado de pré-codificação dos *streams* privados e do *stream* comum, e $s \in [0, 1]$ é o parâmetro que define a divisão de potência entre as mensagens privadas e comuns.

Na próxima seção, apresenta-se a estratégia proposta para solução do problema acima.

3.3 Solução proposta para o problema de otimização

Observa-se que o problema formulado é muito complexo, pois a função objetivo e a restrição envolvida são funções não convexas nas variáveis de otimização. Neste capítulo, recorreremos a uma solução de baixa complexidade, adaptando as soluções apresentadas em (CLERCKX *et al.*, 2020) e (HSIAO *et al.*, 2015), onde o ZF é empregado para as mensagens privadas e um pré-codificador de taxa de dados max-min é aplicado para a mensagem comum. Também limitamos a análise a um cenário simplificado com dois usuários.

3.3.1 Análise dos pré-codificadores

A aplicação do ZF consiste em fixar a direção dos pré-codificadores e ajustar a potência dos *streams*, de tal forma que os usuários recebam o sinal desejado com a máxima intensidade e os outros sinais cheguem com intensidade desprezível. Podemos expressar o vetor de pré-codificação ZF \mathbf{p}_j para o usuário j como:

$$\mathbf{p}_j = \sqrt{P_j} \frac{\mathbf{c}_j}{\|\mathbf{c}_j\|}, \quad (3.8)$$

onde \mathbf{c}_j é obtido a partir da pseudo-inversa da matriz de canal \mathbf{H} :

$$\mathbf{P}_{\text{ZF}} = \mathbf{H}(\mathbf{H}^H \mathbf{H})^{-1}, \quad (3.9)$$

onde $\mathbf{H} = [\mathbf{h}_1 \dots \mathbf{h}_j]$ e \mathbf{c}_j corresponde à j -ésima coluna de \mathbf{P}_{ZF} . Além disso, P_j é a potência alocada ao j -ésimo *stream* privado.

Dessa forma, temos $|\mathbf{h}_j^H \mathbf{p}_i| = 0, \forall i \neq j$, e $|\mathbf{h}_j^H \mathbf{p}_j|^2 = \|\mathbf{h}_j\|^2 \rho P_j$, onde $\rho \in [0, 1]$ expressa a correlação entre os canais, de modo que 0 e 1 correspondem a canais completamente alinhados e ortogonais, respectivamente. O fator de correlação é dado pelo determinante da matriz de Gram, que contém todos os produtos internos dos vetores de canal normalizados. Para $J = 2$, $\rho = 1 - \left| \bar{\mathbf{h}}_1^H \bar{\mathbf{h}}_2 \right|^2$, onde $\bar{\mathbf{h}}_j = \frac{\mathbf{h}_j}{\|\mathbf{h}_j\|}$.

Por outro lado, o pré-codificador do *stream* comum é projetado como o vetor de *beamforming* que maximiza a taxa comum. Dessa forma, temos:

$$\max_{\mathbf{p}_c} \min \left(\frac{|\mathbf{h}_1^H \mathbf{p}_c|^2}{\sigma^2 + |\mathbf{h}_1^H \mathbf{p}_1|^2}, \frac{|\mathbf{h}_2^H \mathbf{p}_c|^2}{\sigma^2 + |\mathbf{h}_2^H \mathbf{p}_1|^2} \right), \quad (3.10)$$

sujeito a:

$$\|\mathbf{p}_c\|^2 = P_c. \quad (3.11)$$

Seja $\psi_j^2 = \sigma^2 + |\mathbf{h}_j^H \mathbf{p}_j|^2 = \sigma^2 + \|\mathbf{h}_j\|^2 \rho P_j$, e $\tilde{\mathbf{h}}_j = \mathbf{h}_j / \psi_j$, assim, reformulamos o problema da seguinte forma:

$$\max_{\mathbf{p}_c} \min \left(|\tilde{\mathbf{h}}_1^H \mathbf{p}_c|^2, |\tilde{\mathbf{h}}_2^H \mathbf{p}_c|^2 \right), \quad (3.12)$$

sujeito a:

$$\|\mathbf{p}_c\|^2 = P_c, \quad (3.13)$$

de modo que a solução para (3.12) é $\mathbf{p}_c = \sqrt{P_c}\mathbf{f}_c$, onde $\mathbf{f}_c(\|\mathbf{f}_c\|^2 = 1)$ é a direção do pré-codificador, obtida por (HSIAO *et al.*, 2015):

$$\mathbf{f}_c = \frac{1}{\sqrt{\lambda}} \left(\mu_1 \tilde{\mathbf{h}}_1 + \mu_2 \tilde{\mathbf{h}}_2 e^{-j\angle\alpha_{12}} \right), \quad (3.14)$$

com

$$\lambda = \frac{\alpha_{11}\alpha_{22} - |\alpha_{12}|^2}{\alpha_{11} + \alpha_{22} - 2|\alpha_{12}|}, \quad (3.15)$$

$$\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \frac{1}{\alpha_{11} + \alpha_{22} - 2|\alpha_{12}|} \begin{bmatrix} \alpha_{22} - |\alpha_{12}| \\ \alpha_{11} - |\alpha_{12}| \end{bmatrix}, \quad (3.16)$$

$$\begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{12}^* & \alpha_{22} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{h}}_1^H \\ \tilde{\mathbf{h}}_2^H \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{h}}_1 & \tilde{\mathbf{h}}_2 \end{bmatrix}. \quad (3.17)$$

Também definimos $r_j^2 = 1 + \frac{|\mathbf{h}_j^H \mathbf{p}_j|^2}{\sigma^2} = 1 + \frac{\|\mathbf{h}_j\|^2 \rho P_j}{\sigma^2}$. Assim, a taxa total alcançada pelos usuários, empregando RSMA com os pré-codificadores mencionados, pode ser escrita como $R_{\text{sum}} = R_c + \log_2(r_1^2) + \log_2(r_2^2)$, onde R_c satisfaz (3.10). Observa-se que, dado \mathbf{p}_c por (3.14), e como $|\tilde{\mathbf{h}}_1^H \mathbf{p}_c| = |\tilde{\mathbf{h}}_2^H \mathbf{p}_c|$, então $R_c = \log_2 \left(1 + |\tilde{\mathbf{h}}_2^H \mathbf{p}_c|^2 \right)$. Além disso, observamos que $\psi_j^2 = \sigma^2 r_j^2$. Portanto, a soma total de taxa é obtida por

$$R_{\text{tot}} = \log_2(r_1^2) + \log_2 \left(r_2^2 + \frac{|\mathbf{h}_2^H \mathbf{p}_c|^2}{\sigma^2} \right). \quad (3.18)$$

A seguir, apresentamos a formulação da alocação de potência entre os *streams* comum e privados, e sua otimização.

3.3.2 Divisão de potência

Sejam P_{\max} , P_p e P_c a potência de transmissão total disponível, a potência total alocada aos *streams* privados e a potência total alocada ao *stream* comum, respectivamente, tal que, $P_p = sP_{\max}$ e $P_c = (1 - s)P_{\max}$, otimizamos a alocação de potência entre os *streams* privados aplicando a solução *Water-Filling* (WF). Desta forma, a potência alocada ao j -ésimo *stream* privado é dada por

$$P_j = \frac{1}{J} \left(\sum_{\forall i} \frac{\sigma^2}{\rho \|\mathbf{h}_i\|^2} + P_p \right) - \frac{\sigma^2}{\rho \|\mathbf{h}_j\|^2}. \quad (3.19)$$

Seguindo (CLERCKX *et al.*, 2020), definimos $\Gamma = \frac{1}{\rho} \left[\frac{1}{\|\mathbf{h}_2\|^2} - \frac{1}{\|\mathbf{h}_1\|^2} \right]$, que reflete a correlação do canal e a disparidade da intensidade dos canais. Desta forma, consideramos dois regimes principais de operação do RSMA.

3.3.2.1 OMA/NOMA/Multicasting

Se $sP_{\max} \leq \Gamma$, alocamos $P_2 = 0$ e $P_1 = sP_{\max}$. De acordo com (CLERCKX *et al.*, 2020), observamos que o RSMA opera como *multicasting* para $s = 0$, NOMA para $0 \leq s \leq 1$, e OMA para $s = 1$. Assim, ajustamos s para que o RSMA execute o melhor esquema de MA para este regime de operação.

3.3.2.2 SDMA/RSMA

Por outro lado, se $sP_{\max} > \Gamma$, temos que P_1 e P_2 são dados pelo WF e são maiores que zero. Além disso, o RSMA opera como SDMA se $s = 1$, enquanto não executa nenhum outro esquema se $0 < s < 1$. Note que, substituindo r_j^2 , ψ_j^2 , e P_j em (3.18), de modo que

$$R_{\text{tot}} = \log_2 \left[\frac{ac + (ad + bc)s + bds^2}{(\sigma^2)^2} \right], \quad (3.20)$$

onde $b = \frac{\|\mathbf{h}_1\|^2 \rho P_{\max}}{2}$, $a = \sigma^2 + \frac{\Gamma}{P_{\max}} b$, $d = \frac{\|\mathbf{h}_2\|^2 \rho P_{\max}}{2} - |\mathbf{h}_2^H \mathbf{f}_c|^2 P_{\max}$, e $c = \sigma^2 - \frac{\Gamma}{P_{\max}} d + |\mathbf{h}_2^H \mathbf{f}_c|^2 (P_{\max} - \Gamma)$. Adicionalmente, se $P_1 > 0$ e $P_2 > 0$, então (3.20) não é uma função de s nem da diferença nas intensidades do canal, mas apenas das direções do canal (CLERCKX *et al.*, 2020). Assim, a solução de $\frac{\partial R_{\text{tot}}}{\partial s} = 0$ fornece o s ótimo para $0 < s < 1$, cuja solução é $s = -\frac{a}{2b} - \frac{c}{2d}$. Portanto, o s ótimo para o regime é

$$s = \min \left(-\frac{a}{2b} - \frac{c}{2d}, 1 \right). \quad (3.21)$$

Na próxima seção, o desempenho da abordagem proposta é avaliada usando simulações numéricas e análise comparativa.

3.4 Avaliação de desempenho

Nesta seção são apresentados os detalhes da implementação, as soluções de referência e os resultados, discutindo as vantagens e desvantagens da abordagem proposta na Seção 3.3. A Tabela 2 resume os principais parâmetros para a implementação do sistema proposto neste capítulo, cujos valores foram baseados no sistema de referência de Schröder *et al.* (2023).

Tabela 2 – Parâmetros de simulação para avaliar a alocação de potência para mensagens comum e privadas em redes RSMA LEO-Terrestre.

Parâmetro	Variável	Valor
Altitude do satélite	d_{sat}	600 km
Frequência da portadora	f	2 GHz
Tamanho do arranjo do satélite	M	6
Espaçamento entre antenas	d_a	7,5 cm
Ganho da antena do satélite	G_{sat}	16 dBi
Ganho da antena do usuário	G_{ue}	0 dBi
Potência de ruído	σ^2	-122 dBW
Potência de transmissão do satélite	P_{max}	25 dBW

3.4.1 Detalhes da implementação

Como mencionado anteriormente, foi considerado um cenário simplificado com dois usuários, distribuídos uniformemente ao redor de um centro de célula localizado a uma distância \bar{D} do nadir do satélite. Deste modo, a distância do usuário j ao nadir do satélite é modelado por uma variável aleatória $\bar{D}_j \sim \mathcal{U}(\bar{D} - \Delta\bar{D}, \bar{D} + \Delta\bar{D})$. Além disso, é garantido uma separação mínima entre os usuário, também de $\Delta\bar{D}$.

Foram considerados dois cenários para avaliação do desempenho dos esquemas de MAs: (i) variando \bar{D} com $\Delta\bar{D} = 10$ km, e $\varepsilon = 0$ (CSIT perfeito) e $\varepsilon = 0,05$ (CSIT imperfeito); (ii) fixando $\bar{D} = 100$ km e $\Delta\bar{D} = 10$ km, enquanto se aumenta ε . Cada caso utiliza 1.000 repetições de Monte Carlo para garantir a confiabilidade estatística. Outros parâmetros seguem Schröder *et al.* (2023) e Clerckx *et al.* (2020).

3.4.2 Soluções de referência

Para fins de comparação, também analisamos algumas soluções de referência. Em Schröder *et al.* (2023), os autores comparam o RSMA, OMA e o SDMA em um sistema de comunicação via satélite LEO, com dois usuários no DL. Os autores propõem uma divisão de potência no RSMA baseada no parâmetro α , que é otimizado por meio de busca exaustiva, de alta complexidade computacional. Especificamente, $P_c = P_{\max} - P_{\max}^\alpha$ e $P_p = P_{\max}^\alpha$, onde P_c e P_p são as potências dos *streams* comum e privados, respectivamente, e P_{\max} é a potência total disponível para o sistema. Adicionalmente, um pré-codificador fixo para o *stream* comum é utilizado, modelado por $\mathbf{p}_c = \sqrt{\frac{P_c}{N}} \begin{bmatrix} 1 & 1 \end{bmatrix}^T$. Para os *streams* privados, é aplicado o MMSE, cuja matriz de pré-codificação é modelada como $\mathbf{P}_p = \sqrt{\frac{P_p}{\text{tr}(\mathbf{W}^H \mathbf{W})}} \mathbf{W}$, onde $\mathbf{W} = \left[\mathbf{H}^H \mathbf{H} + \sigma^2 \frac{J}{P_{\max}} \mathbf{I}_M \right]^{-1} \mathbf{H}^H$, sendo \mathbf{H} a matriz de canal, e J o número de usuários. Assim, a taxa comum e a taxa privada do j -ésimo usuário são dadas, respectivamente, por:

$$R_{c,j} = \log_2 (1 + \gamma_{c,j}), \quad (3.22)$$

$$R_{p,j} = \log_2 (1 + \gamma_{p,j}), \quad (3.23)$$

com $\gamma_{c,j}$ e $\gamma_{p,j}$ definidos em (2.4) e (2.5), respectivamente. Assim, a taxa total alcançável empregando o RSMA é dada por:

$$R^{\text{RSMA}} = \min_j R_{c,j} + \sum_{j=1}^J R_{p,j}. \quad (3.24)$$

O MMSE também foi empregado na implementação do SDMA, mas com a matriz de pré-codificação normalizada com a potência total, uma vez que não há divisão de mensagem. Assim, a taxa total alcançável aplicando o SDMA é dada por:

$$R^{\text{SDMA}} = \sum_{j=1}^J \log_2 \left(1 + \frac{|\mathbf{h}_j^H \mathbf{p}_j^{\text{MMSE}}|^2}{\sigma^2 + \sum_{i=1, i \neq j}^J |\mathbf{h}_j^H \mathbf{p}_i^{\text{MMSE}}|^2} \right), \quad (3.25)$$

onde $\mathbf{p}_j^{\text{MMSE}}$ é o vetor de precodificação do j -ésimo usuário aplicando o MMSE. Por fim, os autores utilizaram o pré-codificador MRT na implementação do OMA, cuja matriz de pré-codificação

para o j -ésimo usuário é $\mathbf{p}_j^{\text{MRT}} = \sqrt{\frac{P_{\text{max}}}{J}} \frac{\mathbf{h}_j}{\|\mathbf{h}_j\|}$. Assim, a taxa total alcançável aplicando o OMA é dada por:

$$R^{\text{OMA}} = \frac{1}{J} \sum_{j=1}^J \log_2 \left(1 + \frac{|\mathbf{h}_j^H \mathbf{p}_j^{\text{MRT}}|^2}{\sigma^2} \right). \quad (3.26)$$

Na próxima seção, são apresentados os resultados das campanhas de simulação e as análises do comportamento do sistema.

3.4.3 Resultados

Nesta seção, são apresentados e analisados os resultados das campanhas de simulação, comparando o desempenho da abordagem descrita na Seção 3.3 (RSMA Proposto) com o OMA, SDMA e com o RSMA com pré-codificador do *stream* comum fixo Schröder *et al.* (2023) (RSMA *baseline*). O objetivo é avaliar o impacto da variação da distância do usuário ao nadir do satélite e o impacto da variação do nível de imperfeição do CSIT, em termos de taxa total alcançável normalizada por banda (bps/Hz).

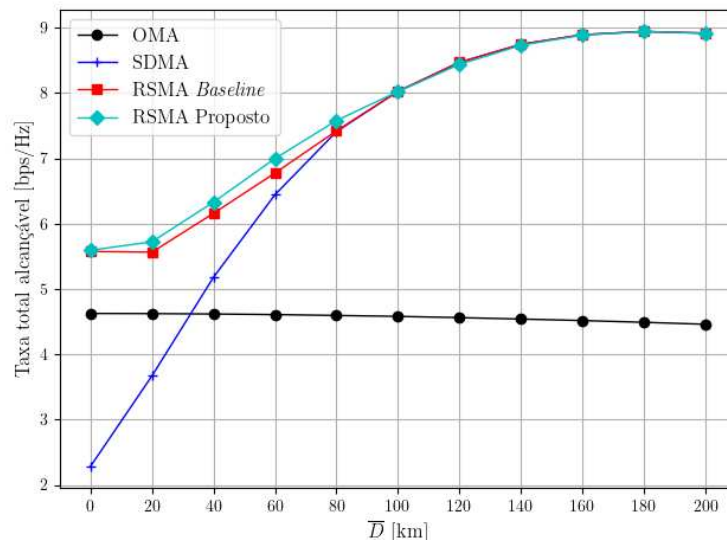
3.4.3.1 Impacto da distância do usuário (\bar{D})

Para avaliar o cenário (i), as Figuras 11 e 12 mostram a taxa total alcançável normalizada por banda (bps/Hz), em função de \bar{D} , considerando CSIT perfeito e imperfeito, respectivamente. Os resultados apresentados nas figuras revelam tendências de desempenho distintas entre as técnicas avaliadas.

O OMA mantém um desempenho relativamente estável devido à sua alocação de recursos ortogonais, que elimina eficazmente a interferência entre usuários, mas ao custo da eficiência espectral. Em contraste, o SDMA e o RSMA exibem ganhos significativos à medida que \bar{D} aumenta, particularmente sob condições de CSIT perfeito. Esse comportamento é atribuído à sua capacidade de aproveitar a diversidade espacial e otimizar as estratégias de pré-codificação. Notavelmente, o RSMA proposto atinge taxas maiores, ou iguais, à solução RSMA *baseline* em todas as distâncias, demonstrando melhor alocação de potência e gerenciamento de interferência sob condições de maior confiabilidade do canal. Essa vantagem torna-se mais pronunciada em regiões mais próximas ao satélite, onde valores de taxa alcançada mais altos são observados para o RSMA proposto.

Sob CSIT imperfeito, representado na Figura 12, todas as técnicas sofrem degradação de desempenho devido à reduzida precisão do estado do canal. Os ganhos de desempenho do SDMA são mais modestos em comparação com o cenário de CSIT perfeito, refletindo sua dependência do conhecimento preciso do canal para um *beamforming* eficaz. O RSMA *baseline* ainda mantém um desempenho superior ao OMA e SDMA, capitalizando sua robustez contra imperfeições do CSIT. O RSMA proposto apresenta desempenho superior ao RSMA *baseline* para um \bar{D} de até 120 km, após o qual o desempenho começa a apresentar uma queda na taxa total alcançada em comparação com as soluções de referência. Embora a queda na taxa não seja muito significativa, isso indica que a combinação da distância ao satélite e as condições de canal imperfeitas causam uma degradação maior na solução proposta. No entanto, a faixa de operação apropriada para a solução proposta é considerável, o que torna a solução útil para o cenário estudado.

Figura 11 – Taxa total de dados, por unidade de banda, alcançável pelos usuários em função de \bar{D} para avaliar o desempenho dos MAs com CSIT perfeito.

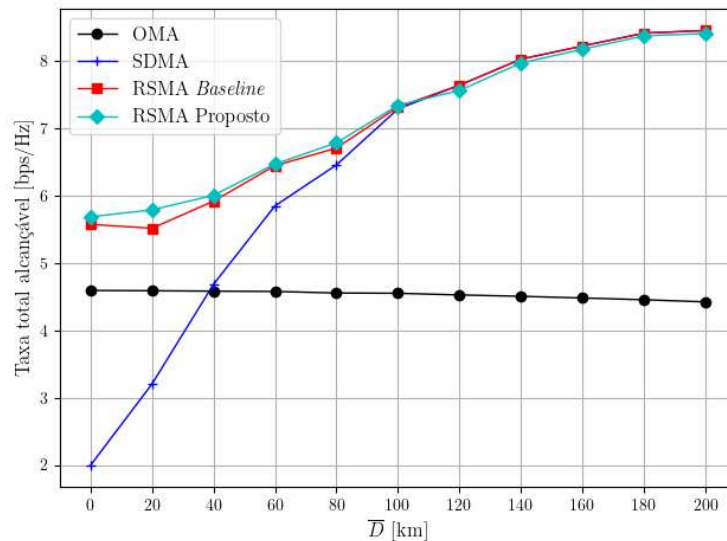


Fonte: Autoral.

3.4.3.2 Impacto do nível de imperfeição do CSIT

Para avaliar o cenário (ii), o valor de \bar{D} é fixado em 100 km, que é uma distância média capaz de proporcionar uma boa separação entre os usuários. Isso nos permite avaliar o impacto das imperfeições do canal em um cenário com uma distância representativa do satélite e com uma correlação de canal que também reflete essa distribuição, conforme aplicado no

Figura 12 – Taxa total de dados, por unidade de banda, alcançável pelos usuários em função de \bar{D} para avaliar o desempenho dos MAs com CSIT imperfeito.



Fonte: Autoral.

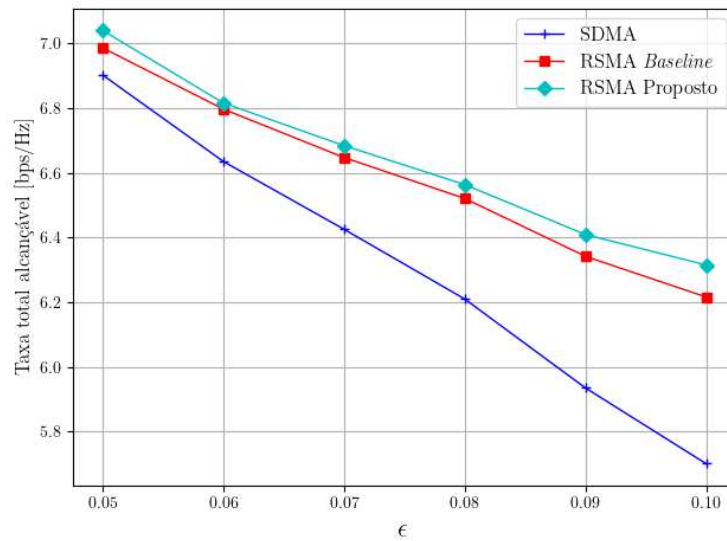
cenário de Schröder *et al.* (2024). A Fig. 13 mostra a taxa total alcançável normalizada por banda (bps/Hz), em função de ε . É possível destacar a resiliência do RSMA proposto ao nível de imperfeição do CSIT.

O resultado do OMA foi omitido devido ao seu baixo desempenho, o que explica suas limitações em termos de adaptabilidade às variações de interferência devido à sua natureza ortogonal. O SDMA também sofre com o aumento de ε , reforçando sua sensibilidade às imprecisões do CSIT. Por outro lado, o RSMA mitiga esses efeitos ao equilibrar eficazmente as alocações de *streams* comum e privados, levando a uma degradação de desempenho mais gradual. O RSMA proposto aprimora ainda mais essa estratégia, superando consistentemente a implementação do RSMA *baseline* em todos os valores de ε , demonstrando uma estrutura de alocação de potência superior que aumenta a robustez contra imperfeições do CSIT para o cenário considerado.

3.4.3.3 Nota sobre a complexidade computacional

É importante notar que na solução RSMA avaliada em (SCHRÖDER *et al.*, 2023), a divisão da potência de transmissão entre as mensagens comum e privadas é otimizada empregando uma solução de busca exaustiva. Em contraste, a solução empregada neste trabalho aplica uma expressão de forma fechada para a alocação de potência entre essas mensagens, resultando em menor complexidade computacional.

Figura 13 – Taxa total de dados, por unidade de banda, alcançável pelos usuários em função de ϵ para avaliação do desempenho dos MAs.



Fonte: Autoral.

3.5 Considerações parciais

Este capítulo apresentou uma expressão de forma fechada como solução para o problema de alocação de potência entre as mensagens comum e privadas em sistemas RSMA LEO-terrestre, considerando as limitações práticas de CSIT imperfeito e restrição de potência total. A estratégia proposta, fundamentada na utilização de pré-codificação ZF para os *streams* privados e uma solução max-min para o *stream* comum, demonstrou ser eficiente ao evitar a necessidade de busca exaustiva, reduzindo a complexidade computacional do sistema.

Os resultados numéricos obtidos confirmaram que a solução proposta atinge soma de taxas superiores ou equivalentes em comparação às abordagens de referência, sobretudo em cenários com boa separação espacial entre os usuários e níveis moderados de imperfeição no CSIT. Adicionalmente, a expressão analítica para otimização da divisão de potência entre os *streams* privados e comum mostrou-se robusta e de fácil implementação.

Esses avanços reforçam o potencial do RSMA como técnica viável para sistemas NTN, contribuindo para maior eficiência espectral e melhor gerenciamento da interferência multiusuário. No capítulo seguinte, são explorados os aspectos de agendamento de usuários em sistemas TDD em cenário LEO-terrestre, complementando o estudo iniciado neste capítulo.

4 ESCALONAMENTO DE UES EM SISTEMAS TDD EM REDES RSMA LEO-TERRESTRE

Este capítulo apresenta um estudo de caso sobre o escalonamento de usuários em um sistema TDD NTN LEO-terrestre, empregando RSMA no DL e OMA no UL. O objetivo central consiste em escalonar G usuários em cada *slot* de transmissão, de modo a minimizar o tempo total necessário para a transmissão de N pares de *slots* DL/UL, garantindo, ao mesmo tempo, os requisitos de QoS individuais de cada usuário em termos de taxa mínima. Na Seção 4.1 são apresentados o modelo de canal, o modelo de transmissão e o *frame* TDD considerados neste capítulo. Na Seção 4.2, é formulado matematicamente um problema de otimização que leva em conta as particularidades do ambiente satelital, incluindo o impacto dos atrasos de propagação e a necessidade de GPs para sincronização, além das restrições de QoS. Em seguida, na Seção 4.3, são detalhadas as soluções implementadas para resolução do problema. Para resolver o problema de forma ótima, utilizou-se o solver CPLEX. Também foi implementado um escalonamento cíclico, do inglês *Round Robin* (RR), como solução de *benchmark*. Considerou-se, ainda, o TDMA como cenário de referência, no qual o problema consiste em escalonar um único usuário por *slot*. Essas soluções de referência permitem compreender os limiares entre complexidade computacional, simplicidade de implementação e desempenho em cenários práticos. Além disso, as soluções de escalonamento são avaliadas empregando diferentes esquemas de MA no DL. Dessa forma, na Seção 4.4 são apresentados os resultados numéricos que mostram o comportamento das soluções implementadas em diferentes cenários, incluindo variação do número de usuários e *slots*, bem como diferentes requisitos de taxa de dados por usuário no DL, possibilitando destacar *insights* relevantes para o entendimento do modelo avaliado. Por fim, a Seção 4.5 apresenta as considerações parciais deste capítulo.

4.1 Modelo de sistema

Semelhante ao Capítulo 3, considera-se um sistema de comunicação NTN integrado a redes terrestres, composto por um único satélite LEO, posicionado a uma altitude d_{sat} e equipado com uma ULA de M antenas, cada uma com ganho G_{sat} . O satélite atua como BS e fornece conectividade direta ao conjunto de usuários $\mathcal{J} = \{1, 2, \dots, J\}$, com J usuários distribuídos em sua área de cobertura. Cada usuário é equipado com uma única antena, de ganho G_{ue} .

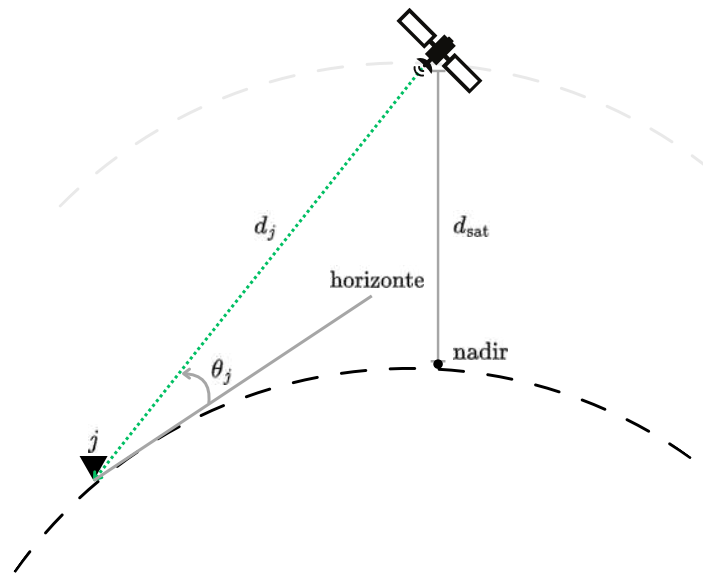
Define-se o ângulo de elevação do usuário j , θ_j , como o ângulo entre a linha que conecta o usuário ao satélite e a linha do horizonte, conforme ilustrado na Figura 14. Assim, a

distância entre o j -ésimo usuário e o satélite é dada por:

$$d_j = \sqrt{R_E^2 \sin^2(\theta_j) + d_{\text{sat}}^2 + 2d_{\text{sat}}R_E - R_E \sin(\theta_j)}, \quad (4.1)$$

onde R_E é o raio da Terra.

Figura 14 – Ilustração de um sistema de comunicação LEO-Terrestre empregando RSMA, destacando o ângulo de elevação do usuário j .



Fonte: Autoral.

Assume-se ainda que o satélite possua mobilidade desprezível em relação às transmissões de DL e UL, permitindo considerar atrasos de propagação constantes durante a análise em nível de *frame*, como também adotado em (TRASPADINI *et al.*, 2024).

Considera-se que o horizonte temporal é representado pelo conjunto $\mathcal{T} = \{1, 2, \dots, N\}$, onde N é o número total de pares de *slots* de transmissão. Em cada instante $t \in \mathcal{T}$, um subconjunto de G usuários é selecionado para transmissão, com $G < J$. Denotando por \mathcal{G} o conjunto de todos os subconjuntos possíveis de \mathcal{J} com cardinalidade G , tem-se:

$$\mathcal{G} = \{g \subset \mathcal{J} : |g| = G\}. \quad (4.2)$$

Assim, o escalonamento em cada instante t consiste na escolha de um elemento $g \in \mathcal{G}$ que define o grupo de usuários servidos naquele *slot*.

A seguir, são descritos o modelo de canal adotado neste capítulo, o modelo de transmissão e o modelo de estrutura de *frame* TDD adotado nesta dissertação.

4.1.1 Modelo de canal

Neste capítulo é considerado o modelo de canal 3GPP NTN com sombreamento, em cenário urbano, do inglês *Urban*, conforme descrito no relatório técnico, do inglês *Technical Report* (TR), 38.811 (3GPP, 2020). A geração do canal ocorre em três etapas: (i) definição da topologia e dos parâmetros de larga escala; (ii) geração dos parâmetros de pequena escala; e (iii) síntese dos coeficientes de canal.

O canal entre a antena m do satélite e o usuário j , no instante t , é modelado por uma resposta ao impulso composta por múltiplos percursos agrupados em Z_{cl} *clusters*, cada um contendo Z_{ray} raios. A resposta ao impulso é expressa como:

$$h_{j,m}(\mathbf{v}, t) = \sum_{c=1}^{Z_{\text{cl}}} \sum_{z=1}^{Z_{\text{ray}}} \alpha_{c,z,j,m}(t) \cdot \delta(\mathbf{v} - \mathbf{v}_{c,z}), \quad (4.3)$$

onde $\mathbf{v}_{c,z}$ é o atraso do raio z do *cluster* c , e $\alpha_{c,z,j,m}(t)$ representa seu ganho complexo no instante t . Este ganho modela os efeitos de desvanecimento de pequena escala e é definido como:

$$\alpha_{c,z,j,m}(t) = \sqrt{P_{c,z}} \cdot G_{j,c,z}^{(\text{Rx})} \cdot G_{m,c,z}^{(\text{Tx})} \cdot \Phi_{c,z} \cdot e^{i\omega_{j,c,z}^{(\text{Rx})}(t)} \cdot e^{i\omega_{m,c,z}^{(\text{Tx})}(t)} \cdot e^{i\omega_{c,z}^{(\text{D})}(t)}, \quad (4.4)$$

sendo:

- $P_{c,z}$ a potência do raio z do *cluster* c ;
- $G_{m,c,z}^{(\text{Tx})}$ o ganho da m -ésima antena transmissora do raio z do *cluster* c , na direção de partida;
- $G_{j,c,z}^{(\text{Rx})}$ o ganho da j -ésima antena receptora do raio z do *cluster* c , na direção de chegada;
- $\Phi_{c,z}$ a fase inicial aleatória do raio z do *cluster* c ;
- $\omega^{(\text{Rx})}$, $\omega^{(\text{Tx})}$ e $\omega^{(\text{D})}$ as fases associadas às posições relativas e ao efeito Doppler.

Assim, a resposta em frequência do canal entre a antena m do satélite e o usuário j , obtida via Transformada de Fourier da resposta ao impulso, para uma portadora com frequência f no instante t , é:

$$H_{j,m}(f,t) = \sum_{c=1}^{Z_{cl}} \sum_{z=1}^{Z_{ray}} \alpha_{c,z,j,m}(t) \cdot e^{-i2\pi f v_{c,z}}. \quad (4.5)$$

Agrupando as respostas de todos os M elementos da antena do satélite ao usuário j , obtém-se o vetor de canal:

$$\mathbf{h}_j(f,t) = \begin{bmatrix} H_{j,1}(f,t) \\ H_{j,2}(f,t) \\ \vdots \\ H_{j,M}(f,t) \end{bmatrix} \in \mathbb{C}^{M \times 1}. \quad (4.6)$$

Este vetor é posteriormente ajustado pelas perdas de larga escala, aplicadas como um fator escalar multiplicativo:

$$\mathbf{h}_j(f,t) \leftarrow \mathbf{h}_j(f,t) \cdot 10^{-\frac{PL_j}{20}}, \quad (4.7)$$

onde a perda total PL_j , em dB, é composta por:

$$PL_j = FSPL_j + A_g + A_s + SF_j, \quad (4.8)$$

onde $FSPL_j$ é a perda de percurso no espaço livre, A_g representa perdas atmosféricas (gás, chuva, nuvens), A_s a perda por cintilação, e SF_j o desvanecimento por sombreamento, do inglês *shadow fading*, conforme modelagem estatística do TR 38.811 (3GPP, 2020).

O emulador de canal utilizado, implementado conforme M. V. Tim Düe (2025), permite a geração dos vetores $\mathbf{h}_j(f,t)$ em diversos cenários, como *Urban*, *Dense Urban* e *Suburban*, com variação de parâmetros como frequência portadora, ângulo de elevação e altura orbital. Estes vetores são utilizados diretamente no projeto dos pré-codificadores e na avaliação do desempenho da transmissão via satélite.

Neste estudo, assume-se um único bloco de recursos físicos, do inglês *Physical Resource Block* (PRB), de forma que o canal é tratado como *flat fading* dentro da banda de interesse. A consideração de múltiplos PRBs, bem como o impacto do seu escalonamento entre diferentes usuários, é deixada como perspectiva para trabalhos futuros, pois implicaria em

aumento significativo da complexidade computacional. Além disso, assume-se o conhecimento prévio dos canais em todos os *slots* de tempo.

4.1.2 Modelo de transmissão

Os modelos de transmissão adotados neste estudo seguem os modelos apresentados no Capítulo 3, nas Seções 3.3.1 e 3.4.2. Primeiramente, para o UL, adota-se o OMA, dividindo a banda disponível de forma equitativa entre os usuários ativos. Deste modo, a partir de (3.26), obtêm-se que a taxa para o j -ésimo usuário, quando pertencente ao grupo $g \in \mathcal{G}$, com g transmitindo no instante t , empregando o OMA, é dada por:

$$R_{j,g,t}^{\text{OMA}} = \frac{B}{G} \log_2 \left(1 + \frac{|\mathbf{h}_j^H(f,t) \mathbf{p}_j^{\text{MRT}}(f,t)|^2}{\sigma_{UL}^2} \right), \quad (4.9)$$

onde B é a banda total disponível, σ_{UL}^2 é a potência de ruído do UL, e $\mathbf{p}_j^{\text{MRT}}(f,t)$ é o vetor de pré-codificação do j -ésimo usuário, para uma portadora com frequência f no instante t , aplicando o MRT. O pré-codificador MRT aplicado neste estudo segue a descrição da Seção 3.4.2.

Por outro lado, no DL, considerou-se quatro modelos de transmissão, conforme os esquemas de MA adotados no DL: (i) RSMA com divisão de potência otimizada, conforme proposto no capítulo 3 (RSMA Proposto). Assim, a partir de (3.18), obtêm-se que a taxa para o j -ésimo usuário, quando pertencente ao grupo $g \in \mathcal{G}$, com g transmitindo no instante t , empregando o RSMA Proposto, é dada por:

$$R_{j,g,t}^{\text{RSMA Proposto}} = \log_2 \left(1 + \frac{\|\mathbf{h}_j(f,t)\|^2 \rho P_j}{\sigma_{DL}^2} \right) + \frac{1}{G} \log_2 \left(1 + \min_{j' \in g} \frac{|\mathbf{h}_{j'}^H(f,t) \mathbf{p}_c(f,t)|^2}{\sigma_{DL}^2} \right), \quad (4.10)$$

onde σ_{DL}^2 é a potência de ruído no DL, e $\mathbf{p}_c(f,t)$ é o pré-codificador comum, dado por (3.14), para uma portadora com frequência f no instante t . Além disto, destaca-se que foi considerado que cada usuário j tenha a mesma contribuição na taxa comum, por isso a taxa comum é dividida por G . Otimizações da divisão de taxa comum também são deixadas como perspectiva para trabalhos futuros, pois implicaria em complexidades não consideradas nesta dissertação; (ii) também foi considerado o RSMA *Baseline*, com pré-codificador do *stream* comum fixo, conforme Schröder *et al.* (2023). Assim, a partir de (4.11), obtêm-se que a taxa para o j -ésimo usuário, quando pertencente ao grupo $g \in \mathcal{G}$, com g transmitindo no instante t , empregando o RSMA *Baseline*, é dada por:

$$R_{j,g,t}^{\text{RSMA}_{\text{Baseline}}} = R_{p,j} + \frac{1}{G} \min_{j' \in g} R_{c,j'}, \quad (4.11)$$

onde $R_{p,j}$ e $R_{c,j'}$ são dados por (3.23) e (3.22), respectivamente. Também foi considerado contribuição igualitária dos usuários na taxa comum; o terceiro esquema de MA é o (iii) OMA, conforme (4.9), porém com σ_{DL}^2 no denominador; por fim, o (iv) SDMA, aplicando (3.25). Assim, obtêm-se que a taxa para o j -ésimo usuário, quando pertencente ao grupo $g \in \mathcal{G}$, com g transmitindo no instante t , empregando o SDMA, é dada por:

$$R_{j,g,t}^{\text{SDMA}} = \log_2 \left(1 + \frac{|\mathbf{h}_j^H(f,t)\mathbf{p}_j^{\text{MMSE}}(f,t)|^2}{\sigma_{DL}^2 + \sum_{i=1, i \neq j}^J |\mathbf{h}_j^H(f,t)\mathbf{p}_i^{\text{MMSE}}(f,t)|^2} \right), \quad (4.12)$$

onde $\mathbf{p}_j^{\text{MMSE}}(f,t)$ é o vetor de pré-codificação do j -ésimo usuário, para uma portadora com frequência f no instante t , aplicando o MMSE. O pré-codificador MMSE aplicado neste estudo segue a descrição da Seção 3.4.2.

4.1.3 Modelo de frame TDD

O *frame* de transmissão considerado neste trabalho adota o modo TDD, em que os recursos de tempo são divididos dinamicamente entre enlaces de DL e UL. Esta escolha visa explorar a reciprocidade do canal, facilitando a estimação de canal no DL a partir de sondagens no UL e vice-versa, além de permitir uma alocação flexível da largura de banda entre as direções de tráfego, característica essencial para cenários heterogêneos e dinâmicos como os dos sistemas NTN.

Em cada *frame* TDD, assume-se que a duração total é dividida em N pares de slots DL/UL, cada um com duração T . Para cada par de *slots*, são escalonados exatamente G usuários para transmissão simultânea no DL e no UL. Esta estratégia permite explorar as vantagens do RSMA no DL ao multiplexar *streams* comum e privados entre usuários selecionados, enquanto no UL adota-se o OMA, dividindo a banda de forma equitativa entre os usuários ativos naquele *slot*.

Conforme apresentado na Seção 2.2, um dos principais desafios na estruturação do *frame* TDD em redes via satélite é a necessidade de GPs, introduzidos para evitar colisões e interferências decorrentes dos diferenciais de atraso de propagação. Seja τ_j o atraso de propagação do j -ésimo usuário, dado por:

$$\tau_j = \frac{d_j}{c}, \quad (4.13)$$

onde c é a velocidade da luz. O tempo de guarda necessário entre os enlaces de DL e UL depende diretamente da diferença entre os atrasos dos usuários escalonados. Em particular, para cada par de *slots*, o valor de guarda é dimensionado com base no maior atraso de propagação entre os usuários escalonados naquele instante, representado por $\tau_{\max} = \max_{j \in \mathcal{G}_t} \tau_j$, onde \mathcal{G}_t denota o grupo de usuários escalonados no par de *slots* t .

Essa característica torna o tempo de guarda uma variável dinâmica e dependente das escolhas de escalonamento, impactando diretamente na eficiência espectral do sistema. Para mitigar esse impacto, o mecanismo de TA é empregado: cada usuário j aplica um ajuste de tempo $T_j = 2(\tau_{\max} - \tau_j)$, garantindo que todos os sinais de UL cheguem sincronizados ao satélite, independentemente da distância. Ainda assim, o tempo de guarda permanece necessário como proteção, e sua duração cresce à medida que aumenta a diversidade de atrasos entre os usuários escalonados.

Esse fenômeno motiva a necessidade de estratégias inteligentes de escalonamento, capazes de selecionar grupos de usuários que não apenas maximizem as taxas de transmissão, mas também reduzam os tempos de guarda, aumentando assim a eficiência temporal e espectral do sistema.

4.2 Formulação do problema de otimização

Neste capítulo, o problema central consiste em minimizar o tempo total necessário para a transmissão de N pares de *slots* DL/UL, atendendo aos requisitos de taxa de cada usuário em um sistema TDD-NTN que emprega RSMA no DL e OMA no UL. Assume-se que, em cada par de *slots*, são escalonados G usuários.

Define-se $x_{g,t}$ como a variável binária de escalonamento, de forma que:

$$x_{g,t} = \begin{cases} 1, & \text{se o grupo } g \in \mathcal{G} \text{ é escalonado no instante } t \in \mathcal{T}, \\ 0, & \text{caso contrário.} \end{cases} \quad (4.14)$$

Para indicar a associação entre usuários e grupos, utiliza-se $a_{g,j}$, onde $a_{g,j} = 1$ se o usuário $j \in \mathcal{J}$ pertence ao grupo $g \in \mathcal{G}$, e $a_{g,j} = 0$ caso contrário.

O problema é formulado como um problema min-max, cujo objetivo é determinar a matriz de escalonamento que minimize o pior atraso de propagação total, garantindo simultaneamente os requisitos individuais de QoS. A formulação matemática é expressa como:

$$\min_{x_{g,t}} \sum_{t \in \mathcal{T}} \left(\max_{j \in \mathcal{J}} \left(\sum_{g \in \mathcal{G}} x_{g,t} \cdot a_{g,j} \cdot \tau_j \right) \right), \quad (4.15)$$

sujeito a:

$$\frac{\sum_{t \in \mathcal{T}} \sum_{g \in \mathcal{G}} T \cdot x_{g,t} \cdot a_{g,j} \cdot R_{j,g,t}^{\text{DL}}}{N \cdot 2T + 2 \sum_{t \in \mathcal{T}} \max_{k \in \mathcal{J}} \left(\sum_{g \in \mathcal{G}} x_{g,t} \cdot a_{g,k} \cdot \tau_k \right)} \geq R_j^{\text{DL}}, \quad \forall j \in \mathcal{J}, \quad (4.16)$$

$$\frac{\sum_{t \in \mathcal{T}} \sum_{g \in \mathcal{G}} T \cdot x_{g,t} \cdot a_{g,j} \cdot R_{j,g,t}^{\text{UL}}}{N \cdot 2T + 2 \sum_{t \in \mathcal{T}} \max_{k \in \mathcal{J}} \left(\sum_{g \in \mathcal{G}} x_{g,t} \cdot a_{g,k} \cdot \tau_k \right)} \geq R_j^{\text{UL}}, \quad \forall j \in \mathcal{J}, \quad (4.17)$$

$$\sum_{g \in \mathcal{G}} x_{g,t} = 1, \quad \forall t \in \mathcal{T}, \quad (4.18)$$

onde $R_{j,g,t}^{\text{DL}}$ e $R_{j,g,t}^{\text{UL}}$ são as taxas efetivas alcançadas pelo usuário j ao ser escalonado no grupo g no instante t . Os valores R_j^{DL} e R_j^{UL} são os requisitos mínimos de taxa para garantir o QoS no DL e no UL, respectivamente. Assim, (4.16) e (4.17) são as restrições de QoS, e implicam que cada usuário deve receber, em média, uma taxa superior ou igual ao requisito mínimo, considerando o tempo total (incluindo *slots* e atrasos de propagação). Além disso, (4.18) garante que, em cada par de *slots*, apenas um grupo seja escalonado.

Essa formulação, no entanto, constitui um problema de programação inteira não linear, do inglês *Mixed-Integer Non-Linear Problem*, cuja complexidade decorre da presença da função max e das variáveis no denominador das restrições. A seguir, é detalhada a reformulação do problema, de forma a obter um problema computacionalmente menos complexo.

4.2.1 Problema linearizado

Para possibilitar a resolução prática, recorre-se à linearização do problema, convertendo-o em um problema de programação inteira mista, do inglês *Mixed-Integer Programming* (MIP).

Primeiro, introduz-se a variável auxiliar contínua $u_t \geq 0$, representando o atraso máximo (dentro do grupo escalonado) no instante t . Essa variável é definida por:

$$u_t \geq \sum_{g \in \mathcal{G}} x_{g,t} \cdot a_{g,j} \cdot \tau_j, \quad \forall j \in \mathcal{J}, \forall t \in \mathcal{T}. \quad (4.19)$$

Com isso, o problema pode ser reescrito da seguinte forma:

$$\min_{x_{g,t}, u_t} \sum_{t \in \mathcal{T}} u_t, \quad (4.20)$$

sujeito a:

$$\sum_{t \in \mathcal{T}} \sum_{g \in \mathcal{G}} T \cdot x_{g,t} \cdot a_{g,j} \cdot R_{j,g,t}^{\text{DL}} \geq R_j^{\text{DL}} (N \cdot 2T + 2 \sum_{t \in \mathcal{T}} u_t), \quad \forall j \in \mathcal{J}, \quad (4.21)$$

$$\sum_{t \in \mathcal{T}} \sum_{g \in \mathcal{G}} T \cdot x_{g,t} \cdot a_{g,j} \cdot R_{j,g,t}^{\text{UL}} \geq R_j^{\text{UL}} (N \cdot 2T + 2 \sum_{t \in \mathcal{T}} u_t), \quad \forall j \in \mathcal{J}, \quad (4.22)$$

$$\sum_{g \in \mathcal{G}} x_{g,t} = 1, \quad \forall t \in \mathcal{T}, \quad (4.23)$$

$$x_{g,t} \in \{0, 1\}, \quad \forall g \in \mathcal{G}, \forall t \in \mathcal{T}, \quad (4.24)$$

$$u_t \geq \sum_{g \in \mathcal{G}} x_{g,t} \cdot a_{g,j} \cdot \tau_j, \quad \forall j \in \mathcal{J}, \forall t \in \mathcal{T}, \quad (4.25)$$

$$u_t \geq 0, \quad \forall t \in \mathcal{T}, \quad (4.26)$$

onde (4.21) e (4.22) são as restrições de QoS linearizadas. Além disso, (4.25) e (4.26) impõem que u_t capture o maior atraso de propagação entre os usuários escalonados no instante t . Isso substitui o operador max original, obtendo assim a linearidade. Analogamente ao problema original, (4.23) garante a unicidade da escolha do grupo em cada *slot* e (4.24) garante que cada decisão seja binária (escalar ou não o grupo).

Na seção seguinte, são apresentadas as soluções de *benchmark* utilizadas para avaliar o sistema.

4.3 Soluções de *baseline*

Nesta seção são apresentadas as soluções utilizadas como referência para o estudo do problema de escalonamento. As soluções implementadas abrangem desde uma solução ótima via CPLEX até heurísticas simples, permitindo analisar os limites entre desempenho e complexidade.

4.3.1 Solução ótima com CPLEX

A solução ótima é obtida resolvendo o problema MIP com o *solver* CPLEX, conforme a modelagem linearizada apresentada na Seção 4.2. Esta abordagem considera explicita-

mente todas as restrições de QoS, os atrasos de propagação e a restrição de escalonamento. A solução encontrada serve como *benchmark* para avaliar a eficiência das heurísticas, estabelecendo um limite inferior para a função objetivo, que é o tempo total de transmissão. O algoritmo 1 sumariza a implementação do problema com CPLEX.

Algorithm 1 Solução ótima via CPLEX

Require: $\mathcal{J}, \mathcal{G}, N, R_j^{DL}, R_j^{UL}, R_{j,g,t}^{DL}, R_{j,g,t}^{UL}, a_{g,j}, \tau_j, T$

Ensure: $x_{g,t}, u_t$

- 1: Inicializar modelo de otimização (CPLEX)
- 2: Definir variáveis binárias $x_{g,t}$ para cada $g \in \mathcal{G}, t \in \mathcal{T}$
- 3: Definir variáveis contínuas $u_t \geq 0$ para cada $t \in \mathcal{T}$
- 4: Definir objetivo: minimizar $\sum_t u_t$
- 5: **for all** usuário $j \in \mathcal{J}$ **do**
- 6: Adicionar restrição de QoS no DL:

$$\sum_t \sum_g T \cdot x_{g,t} \cdot a_{g,j} \cdot R_{j,g,t}^{DL} \geq R_j^{DL} (2NT + 2 \sum_t u_t)$$

- 7: Adicionar restrição de QoS no UL:

$$\sum_t \sum_g T \cdot x_{g,t} \cdot a_{g,j} \cdot R_{j,g,t}^{UL} \geq R_j^{UL} (2NT + 2 \sum_t u_t)$$

- 8: **end for**
- 9: **for all** slot $t \in \mathcal{T}$ **do**
- 10: Adicionar restrição de unicidade:

$$\sum_g x_{g,t} = 1$$

- 11: **for all** usuário $j \in \mathcal{J}$ **do**
- 12: Adicionar restrição de atraso:

$$u_t \geq \sum_g x_{g,t} \cdot a_{g,j} \cdot \tau_j$$

- 13: **end for**
 - 14: **end for**
 - 15: Resolver modelo com CPLEX
 - 16: Retornar variáveis $x_{g,t}$ e u_t
-

4.3.2 RR

A heurística RR é baseada em um escalonamento cíclico fixo. Os grupos de usuários são previamente ordenados e alocados sequencialmente ao longo dos *slots*, reiniciando o ciclo caso o número de *slots* exceda o número de grupos. É importante notar que, ao contrário da

solução ótima, a heurística RR não incorpora as restrições de QoS em seu processo de decisão. O atendimento (ou não) dos requisitos de taxa é uma consequência a ser avaliada *a posteriori*, e não uma garantia do algoritmo. O algoritmo 2 sumariza a implementação.

Algorithm 2 Heurística RR

Require: N, \mathcal{G}

Ensure: $x_{g,t}$

- 1: **for** $t = 1$ até N **do**
 - 2: Seja g o grupo na posição $((t - 1) \bmod |\mathcal{G}|)$ da lista ordenada de grupos
 - 3: Escalonar grupo g no slot t , definindo $x_{g,t} = 1$
 - 4: **end for**
 - 5: Retornar $x_{g,t}$
-

4.3.3 Cenário de Referência TDMA

Para avaliar os ganhos da multiplexação espacial, um cenário de referência baseado em TDMA é considerado. Neste cenário, o tamanho do grupo é fixado em $G = 1$, ou seja, apenas um único usuário é atendido por *slot*, utilizando toda a banda disponível sem sofrer interferência de outros usuários. A SNR do usuário j no instante t é dada por:

$$\gamma_{j,t}^{\text{TDMA}} = \frac{|\mathbf{h}_j^H(f,t)\mathbf{p}_j(f,t)|^2}{\sigma^2}, \quad (4.27)$$

onde $\mathbf{h}_j(f,t)$ é o vetor de canal, $\mathbf{p}_j(f,t)$ é o vetor de pré-codificação aplicando o MRT, e σ^2 corresponde à potência do ruído. A taxa alcançável é então obtida pela expressão de Shannon:

$$R_{j,t}^{\text{TDMA}} = B \cdot \log_2 \left(1 + \gamma_{j,t}^{\text{TDMA}} \right). \quad (4.28)$$

O problema de escalonamento, que consiste em definir qual usuário será servido em cada *slot*, é então resolvido com duas abordagens para fins de comparação: (i) a solução ótima via CPLEX, adaptada para o caso com $G = 1$, e (ii) a heurística RR, que simplesmente cicla entre os J usuários disponíveis.

4.3.4 Nota sobre a complexidade computacional das soluções

A solução ótima via CPLEX apresenta elevada complexidade computacional, especialmente para grandes valores de J e N . Um fator central é que o número de grupos possíveis

de usuários cresce de forma combinatorial com J , dado por $|\mathcal{G}| = \binom{J}{G}$. Mesmo para valores moderados de J e G , o número de variáveis binárias $x_{g,t}$ torna o problema intratável, pois este se enquadra na classe de problemas NP-difíceis. Como consequência, embora o CPLEX ofereça a solução globalmente ótima, seu tempo de execução cresce exponencialmente, tornando-o impraticável para sistemas em tempo real ou para cenários de grande escala.

As heurísticas, por outro lado, possuem baixa complexidade computacional. O RR tem complexidade linear em relação ao número de *slots* ($\mathcal{O}(N)$), pois a decisão de escalonamento em cada *slot* é computacionalmente simples.

Em resumo, embora as heurísticas não garantam a otimalidade, elas permitem obter soluções com baixo custo computacional. Assim, são adequadas para aplicações práticas e servem como referência para quantificar o quão próximo de uma solução ótima uma proposta mais sofisticada pode chegar.

A seguir, é apresentada a análise comparativa de desempenho entre as soluções, discutindo detalhadamente os impactos em termos de tempo total, eficiência espectral e atendimento de QoS.

4.4 Avaliação de desempenho

Esta seção apresenta a metodologia utilizada para avaliar o desempenho das diferentes abordagens de escalonamento discutidas na Seção 4.3. A Tabela 3 resume os principais parâmetros para implementação do sistema proposto neste capítulo, seguindo os valores adotados no Capítulo 3, e valores de transmissão e ruído adotados na literatura, de modo a obter taxas de dados compatíveis com valores práticos (FILIPPO *et al.*, 2025; M. V. Tim Düe, 2025). Além disso, os parâmetro de canal seguem os parâmetros do cenário *Urban*, conforme o TR 38.811 (3GPP, 2020).

4.4.1 Detalhes da implementação

As simulações foram implementadas em *Python*, utilizando como base o emulador de canal *Open NTN* (M. V. Tim Düe, 2025), baseado no TR 38.811, considerando-se o cenário de propagação *Urban*, conforme as diretrizes do modelo 3GPP. Foram geradas 5.000 instâncias por meio de simulações de Monte Carlo. Para cada instância de Monte Carlo, são geradas N amostras do canal, representando a evolução temporal do mesmo ao longo dos N pares de slots

Tabela 3 – Parâmetros de simulação para avaliação do escalonamento em redes NTN com diferentes técnicas de múltiplo acesso.

Parâmetro	Variável	Valor
Altitude do satélite	d_{sat}	600 km
Frequência da portadora	f	2 GHz
Número de antenas no satélite	M	6
Ganho da antena do satélite	G_{sat}	16 dBi
Ganho da antena do usuário	G_{ue}	0 dBi
Numerologia	μ	0
Espaçamento entre subportadoras	–	15 kHz
Bloco de recurso físico	–	1
Banda total	B	180 kHz
Duração de slot	T	1 ms
Potência de transmissão no DL	P_{DL}	51,76 dBm
Potência de transmissão no UL	P_{UL}	20 dBm
Temperatura do ruído (DL)	T_{DL}	290 K
Temperatura do ruído (UL)	T_{UL}	1 K
Fator de ruído (DL)	F_{DL}	5,5 dB
Fator de ruído (UL)	F_{UL}	5 dB
Potência de ruído (DL)	σ_{DL}^2	–116 dBm
Potência de ruído (UL)	σ_{UL}^2	–140 dBm

de transmissão.

Além disso, considera-se um cenário simplificado com $G = 2$, isto é, para cada par de *slots* é escalonada uma dupla de usuários. Os usuários são posicionados com base nos seus respectivos ângulos de elevação, gerados a partir de uma variável aleatória uniforme, $\theta_j \sim \mathcal{U}(\theta_{\min}, \theta_{\max})$. Os cenários de simulação considerados neste capítulo utilizam valores fixos de $\theta_{\min} = 35^\circ$ e $\theta_{\max} = 90^\circ$.

Adicionalmente, as taxas utilizadas no problema de escalonamento são previamente calculadas para cada instante de tempo, considerando a matriz de canal gerada por Monte Carlo e a técnica de MA aplicada em cada enlace. Essas taxas são então armazenadas e utilizadas como entradas fixas no problema de escalonamento, permitindo separar o cálculo físico das capacidades do canal da lógica de alocação temporal. Os requisitos mínimos de taxa para cada usuário nos enlaces de DL e UL, denotados por R_{\min}^{DL} e R_{\min}^{UL} , foram definidos, para o enlace de UL, como $R_{\min}^{\text{UL}} = 350$ bps. R_{\min}^{DL} é utilizada como variável, conforme discutido nas seções posteriores.

Foram considerados três cenários para avaliação do desempenho das soluções de escalonamento: (i) variando R_{\min}^{DL} , com $J = 6$ e $N = 6$; (ii) variando N , com $J = 6$ e $R_{\min}^{\text{DL}} = 800$ bps; (iii) variando J , com $N = 8$ e $R_{\min}^{\text{DL}} = 800$ bps;

4.4.2 Esquemas avaliados

Conforme dito anteriormente, foram implementadas quatro combinações distintas de MA nos enlaces, todas empregando OMA no UL. Deste modo, na próxima seção são avaliadas as seguintes curvas diante dos cenários de simulação: (i) RSMA Proposto + OMA - CPLEX; (ii) RSMA Proposto + OMA - RR; (iii) RSMA *Baseline* + OMA - CPLEX; (iv) RSMA *Baseline* + OMA - RR; (v) OMA + OMA - CPLEX; (vi) OMA + OMA - RR; (vii) SDMA + OMA - CPLEX; (viii) SDMA + OMA - RR; (ix) TDMA + TDMA - CPLEX; e (x) TDMA + TDMA - RR.

4.4.3 Resultados

Nesta seção são apresentados e analisados os resultados das campanhas de simulação, comparando as soluções de *benchmark* (TDMA e RR) com a solução ótima, obtida via *solver* CPLEX, nos diferentes cenários de avaliação, permitindo compreender os compromissos entre simplicidade computacional e desempenho. O objetivo é avaliar o impacto da variação dos requisitos mínimos de taxa no DL, do número total de usuários, e do número total de *slots*, em termos de taxas de *outage*, tempo total de transmissão e eficiência espectral.

4.4.3.1 Impacto do requisito mínimo de QoS no DL

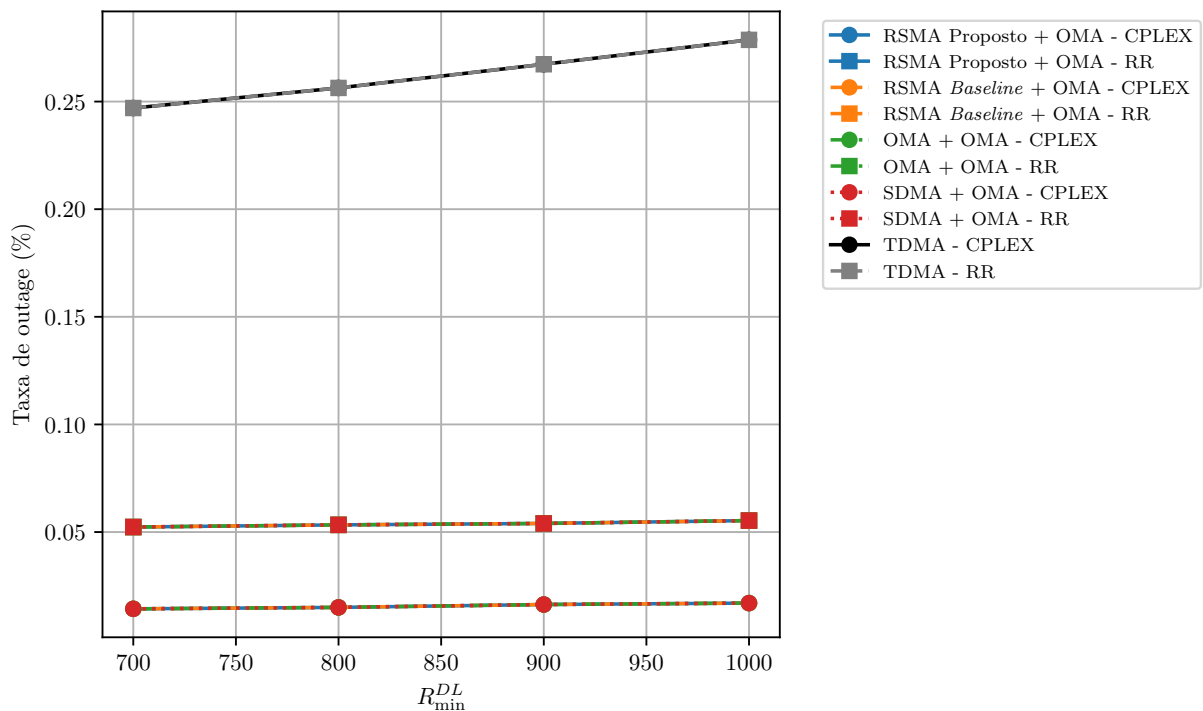
O primeiro conjunto de análises investiga o comportamento do sistema ao variar o requisito mínimo de taxa no DL, R_{\min}^{DL} , mantendo o número de usuários $J = 6$ e o número de pares de *slots* $N = 6$.

A Figura 15 ilustra a taxa de *outage* do sistema. Neste contexto, o *outage* é definido como a fração de instâncias de Monte Carlo em que os requisitos mínimos de QoS não são atendidos por todos os usuários. É importante notar como essa métrica é apurada para cada tipo de solução: para os esquemas baseados em CPLEX, o *outage* ocorre quando o otimizador não encontra uma solução viável, declarando a instância como infactível; já para a heurística RR, que sempre produz um escalonamento, o *outage* é verificado *a posteriori*, contabilizando os casos em que a alocação resultante não satisfaz as restrições de taxa. Feita essa distinção, a primeira observação notável na figura é que, para todos os esquemas, a taxa de *outage* aumenta ligeiramente com o aumento de R_{\min}^{DL} . Este comportamento é esperado, pois um requisito de QoS mais rigoroso restringe o espaço de soluções viáveis, tornando mais provável que, para uma dada

realização do canal, nenhuma combinação de escalonamento consiga satisfazer a demanda.

A análise comparativa entre os esquemas revela duas divisões de desempenho fundamentais. A primeira é a enorme diferença entre os esquemas que empregam multiplexação espacial ($G = 2$) e o TDMA ($G = 1$). O TDMA apresenta uma taxa de *outage* significativamente superior (em torno de 25-28%), evidenciando sua ineficiência espectral ao não explorar o domínio espacial para servir múltiplos usuários simultaneamente. A segunda divisão ocorre entre os escalonadores. As soluções baseadas em CPLEX apresentam uma taxa de *outage* muito baixa (abaixo de 2%), enquanto a heurística RR, por ser agnóstica às condições de canal e aos requisitos de QoS, falha em encontrar uma solução viável em cerca de 5,5% dos casos. É interessante notar que, para um dado escalonador (CPLEX ou RR), todos os esquemas de múltiplo acesso com $G = 2$ (RSMA Proposto, RSMA *Baseline*, SDMA e OMA) apresentam uma taxa de *outage* muito próxima. Isso sugere que, para este cenário, a viabilidade do sistema é limitada pela capacidade do escalonador em alocar os recursos no tempo, e não pela capacidade do esquema de múltiplo acesso em si. No entanto, vale ressaltar que a abordagem de RSMA Proposta é computacionalmente menos complexa que o RSMA *Baseline*.

Figura 15 – Taxa de *outage* em função de R_{\min}^{DL} .



Fonte: Autoral.

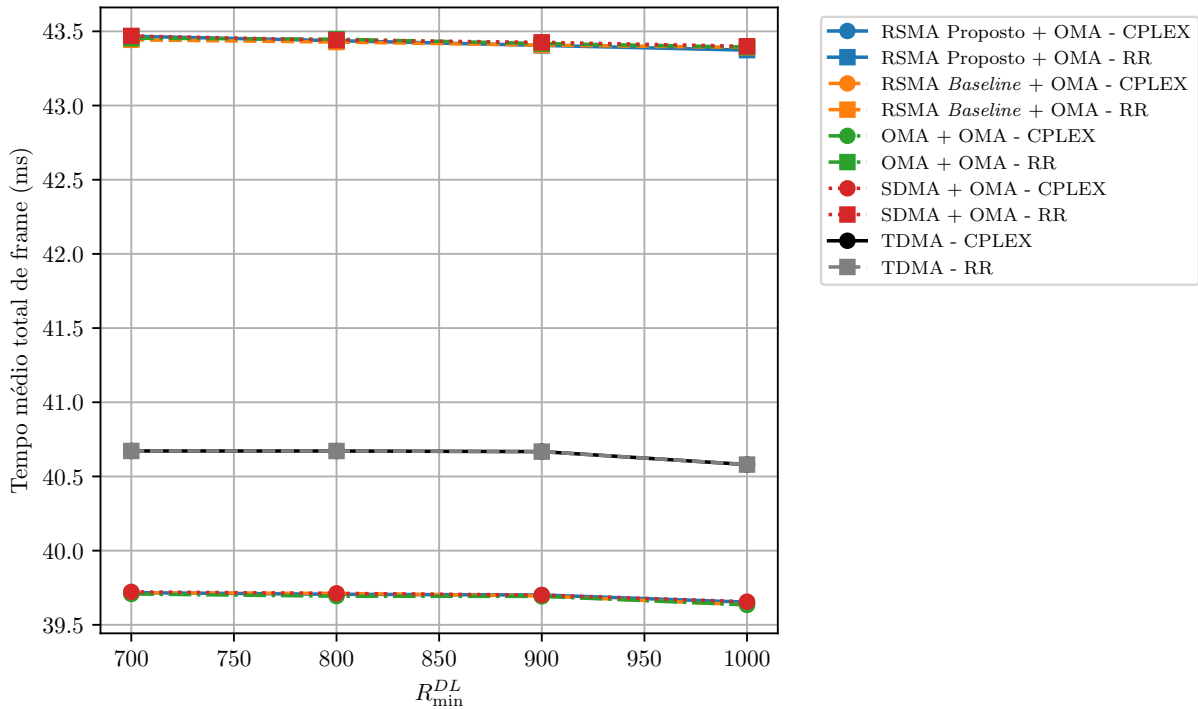
A Figura 16 apresenta o tempo médio total do *frame*, que é a soma da duração dos *slots* e dos tempos de guarda necessários, sendo diretamente proporcional à função objetivo

minimizada. Observa-se que o escalonador CPLEX alcança um tempo total consistentemente menor (em torno de 39,7 ms) em comparação com o RR (43,4 ms). Isso ocorre porque o CPLEX, ao minimizar $\sum u_t$, seleciona ativamente grupos de usuários com menor dispersão de atraso de propagação para reduzir os tempos de guarda. A heurística RR, por outro lado, cicla por todos os grupos sem critério, resultando em uma média de tempos de guarda mais elevada. O TDMA, embora escalonado com CPLEX, apresenta um tempo ligeiramente superior (40,6 ms) aos demais esquemas com CPLEX, pois a otimização se restringe a ordenar usuários individuais, e não a formar pares eficientes. Um comportamento notável é a leve tendência de queda do tempo total com o aumento de R_{\min}^{DL} para as soluções com CPLEX. A explicação para este fenômeno, que é aparentemente contraintuitivo, pode ser atribuída a uma combinação de fatores. A primeira hipótese, ligada à estrutura do problema de otimização, é que para satisfazer uma restrição de QoS mais exigente (conforme a Equação (4.21)), o otimizador é forçado a encontrar soluções mais eficientes. Isso pode envolver não apenas a seleção de grupos com canais de alta taxa, mas também daqueles com menor dispersão de atraso, de modo a não "inflar" o denominador da restrição e, conseqüentemente, reduzir o tempo total de guarda.

Uma segunda hipótese, de natureza estatística, sugere que o aumento do requisito de QoS torna o problema mais difícil de ser resolvido, elevando a taxa de *outage*. Como o tempo médio é calculado apenas sobre as instâncias factíveis, um aumento no *outage* poderia enviesar a amostragem, fazendo com que a média seja computada sobre um subconjunto de cenários inerentemente "mais fáceis" (e.g., com usuários geograficamente mais próximos e menor atraso). Contudo, deve-se ponderar que a variação na taxa de *outage* para as soluções com CPLEX, conforme visto na Figura 15, é relativamente pequena. Isso indica que este viés estatístico, embora potencialmente presente, pode não ser o fator dominante. Portanto, é plausível que a tendência de queda observada seja resultado da combinação de ambos os efeitos: uma pressão otimizatória mais forte por soluções eficientes e um leve viés na seleção de cenários factíveis.

As Figuras 17 e 18 mostram a taxa de dados média total nos enlaces de DL e UL, respectivamente. Na Figura 17, observa-se uma clara hierarquia de desempenho que reflete a eficiência espectral de cada esquema de múltiplo acesso. As variantes de RSMA e SDMA apresentam o melhor desempenho, superando o OMA no DL e, com larga margem, o TDMA. Isso demonstra o benefício da multiplexação espacial e da gestão de interferência mais sofisticada do RSMA e SDMA. Em todos os casos, o escalonador CPLEX alcança uma taxa superior à do RR, pois seleciona de forma inteligente os grupos com melhores condições de canal para

Figura 16 – Tempo médio total de *frame* em função de R_{\min}^{DL} .

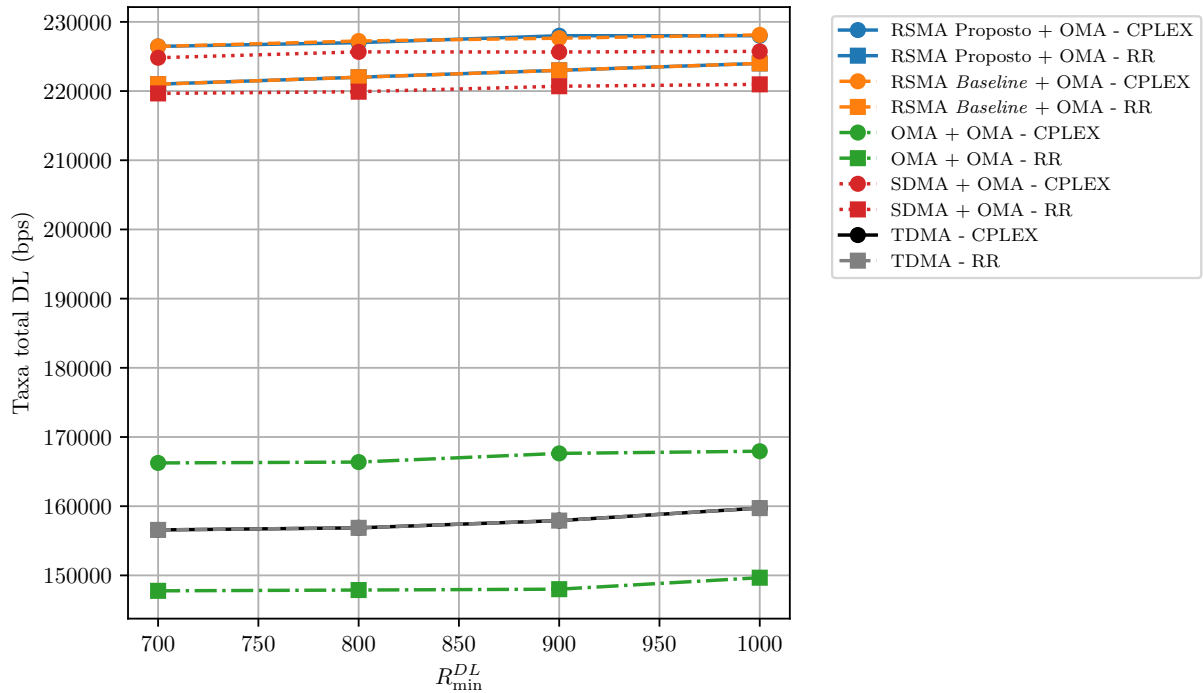


Fonte: Autoral.

maximizar a transmissão de dados, sempre respeitando as restrições.

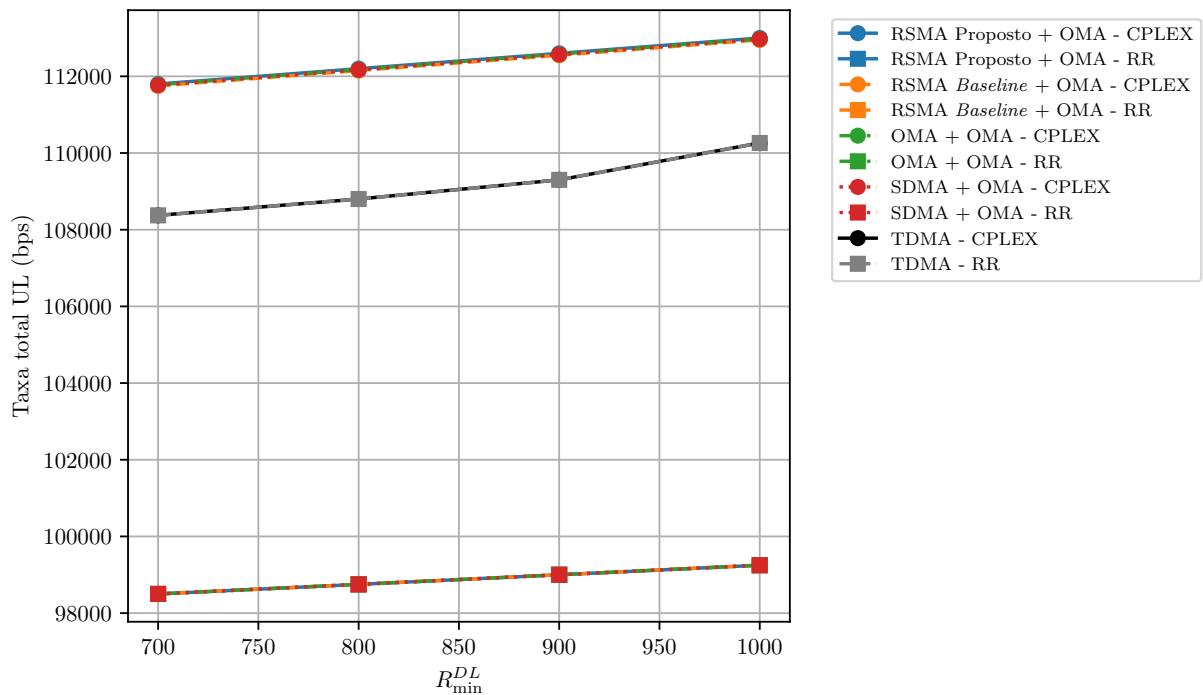
No enlace de UL (Figura 18), onde o OMA é majoritariamente utilizado, o comportamento dos diferentes esquemas com $G = 2$ é quase idêntico quando se utiliza o mesmo escalonador. Isso era esperado, pois a estratégia de transmissão no UL é a mesma para esses casos, e as pequenas variações de taxa decorrem apenas das diferentes decisões de escalonamento, que foram primariamente motivadas pelos requisitos do DL. O leve aumento na taxa de dados em ambos os enlaces com o crescimento de R_{\min}^{DL} também pode ser explicado por dois fatores complementares. Primeiramente, ao ser forçado a encontrar soluções mais eficientes para o DL, o otimizador pode, como efeito secundário, selecionar combinações de usuários que também resultam em um desempenho de taxa de dados ligeiramente superior. Em segundo lugar, e em linha com a análise anterior, este comportamento pode ser parcialmente atribuído a um viés estatístico: como um R_{\min}^{DL} mais alto aumenta a probabilidade de *outage*, a média de taxa é calculada sobre um conjunto de cenários que tendem a ser mais favorável, ou seja, com condições de canal intrinsecamente melhores para ambos os enlaces.

Figura 17 – Taxa de dados média total dos usuários no DL em função de R_{\min}^{DL} .



Fonte: Autoral.

Figura 18 – Taxa de dados média total dos usuários no UL em função de R_{\min}^{DL} .



Fonte: Autoral.

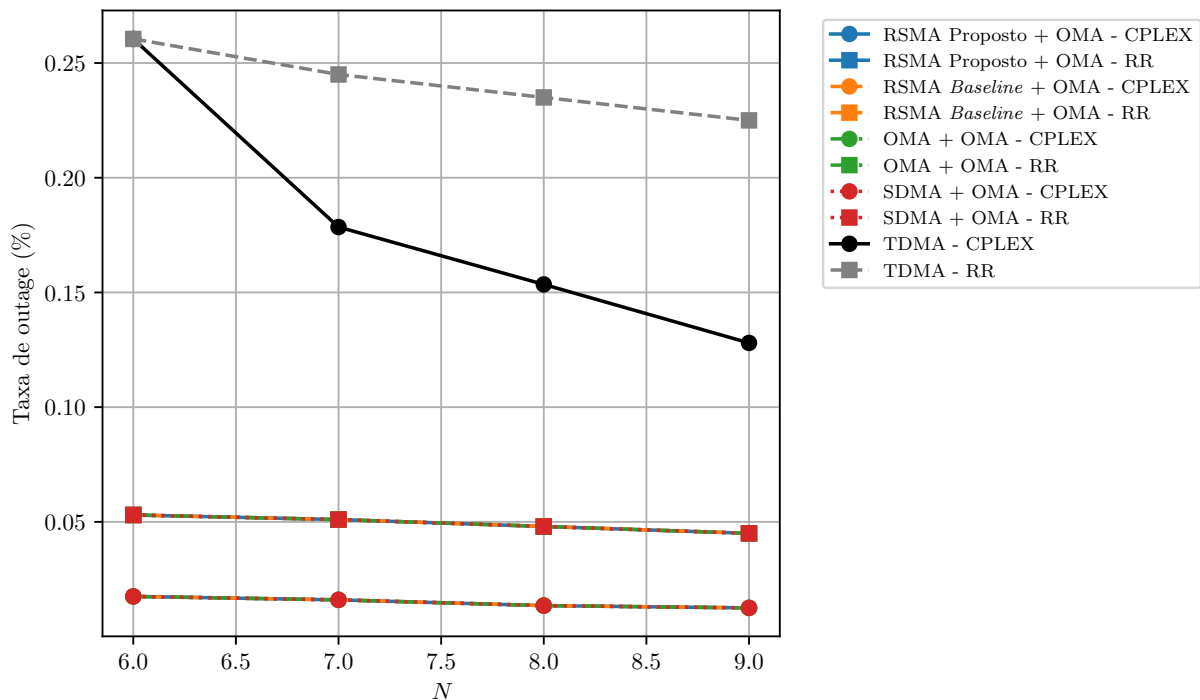
4.4.3.2 Impacto do número total de slots

Nesta segunda análise, avalia-se o impacto do número de pares de *slots* de transmissão, N , que representa o horizonte de tempo disponível para o escalonador. O número de usuários

foi fixado em $J = 6$ e o requisito de taxa no DL em $R_{\min}^{DL} = 800$ bps. O objetivo é compreender como a maior flexibilidade temporal afeta o desempenho do sistema.

A Figura 19 apresenta a taxa de *outage* em função de N . A observação mais importante é que, para todos os esquemas, a taxa de *outage* diminui consistentemente à medida que N aumenta. Teoricamente, um maior número de *slots* concede ao escalonador mais graus de liberdade e diversidade temporal. Se um usuário possui condições de canal desfavoráveis em um determinado instante, o escalonador tem mais oportunidades futuras para alocá-lo em um *slot* com melhores condições, facilitando o cumprimento da restrição de QoS acumulada. Este efeito é particularmente acentuado para as soluções ótimas com CPLEX, que exploram essa flexibilidade de forma eficiente, reduzindo o *outage* a valores próximos de 1% para $N = 9$. É notável a divergência entre as curvas do TDMA-CPLEX e TDMA-RR. Enquanto o escalonador ótimo consegue reduzir drasticamente o *outage* do TDMA ao ter mais *slots* para escolher os melhores instantes para cada usuário, a heurística RR melhora de forma muito mais modesta, pois permanece agnóstica ao canal.

Figura 19 – Taxa de *outage* em função de N .

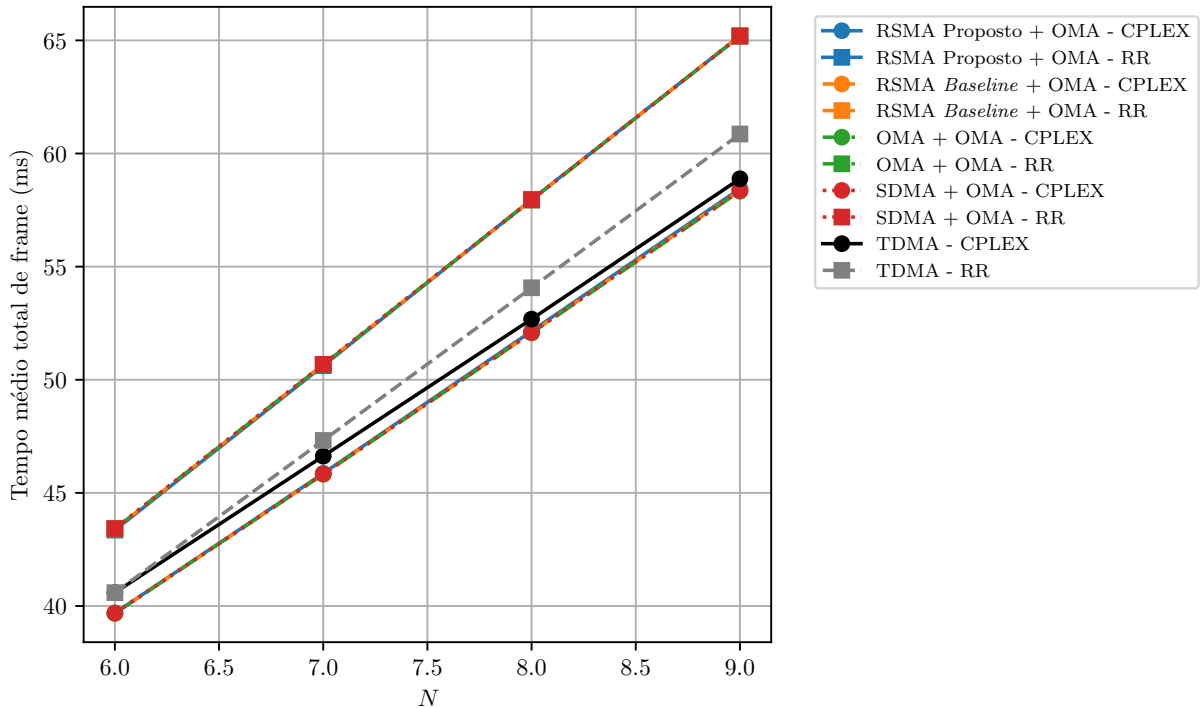


Fonte: Autoral.

A Figura 20 ilustra o tempo médio total do *frame*. Como esperado, o tempo total cresce de forma aproximadamente linear com N para todos os esquemas. Isso ocorre porque o tempo total é dominado pelo termo $2NT$ na formulação da taxa média, que é inerentemente

linear em N . As diferenças de desempenho entre os escalonadores e os esquemas de múltiplo acesso se mantêm consistentes com a análise anterior: o CPLEX otimiza os tempos de guarda, resultando em um tempo total menor em comparação com a heurística RR.

Figura 20 – Tempo médio total de *frame* em função de N .



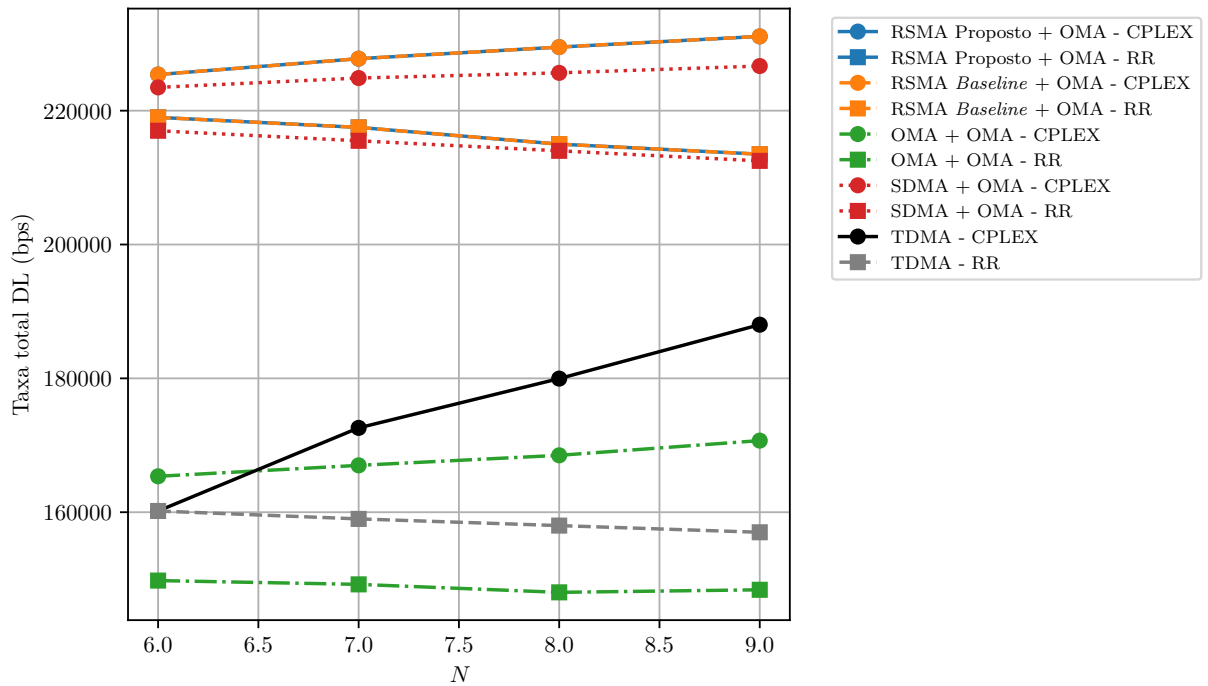
Fonte: Autoral.

As taxas de dados médias totais, exibidas nas Figuras 21 e 22, revelam que, para o DL, as soluções com CPLEX demonstram uma leve tendência de aumento na taxa de dados com N , enquanto as soluções com RR apresentam uma leve queda. O escalonador CPLEX, ao dispor de mais opções, pode não apenas garantir a QoS, mas também selecionar os grupos com as melhores condições de canal, aumentando a taxa de dados total. Por outro lado, a heurística RR, ao percorrer um horizonte de tempo maior, tende a uma média de desempenho que inclui mais instantes de canal de baixa qualidade, resultando em uma taxa total ligeiramente menor.

O resultado mais expressivo, contudo, é observado na taxa de dados do UL para o TDMA-CPLEX, que apresenta um crescimento significativo com N . Com poucos *slots*, o escalonador está fortemente restringido a atender os requisitos de DL. À medida que N cresce e a restrição de DL se torna mais fácil de cumprir, o escalonador ganha liberdade para otimizar objetivos secundários, como a taxa no UL, alocando os usuários em instantes em que seus canais de UL são mais fortes. Este resultado evidencia a capacidade de um escalonador ótimo em explorar os recursos de forma holística, otimizando o sistema em múltiplas dimensões quando

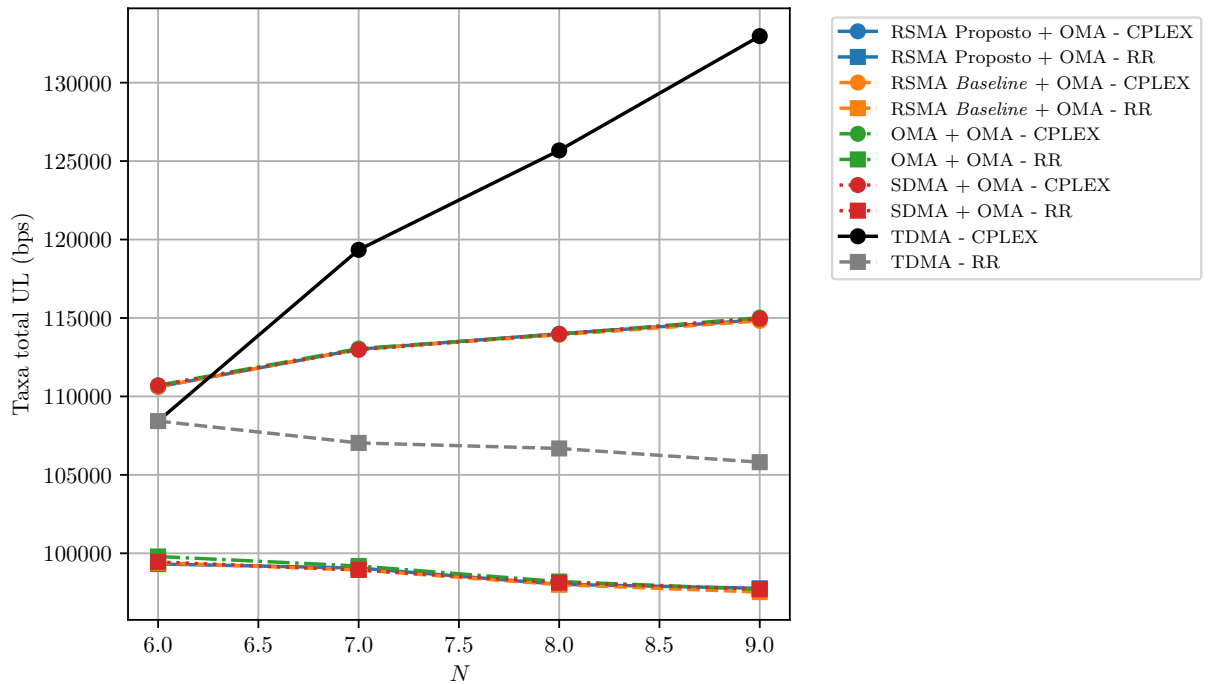
dispõe de flexibilidade suficiente.

Figura 21 – Taxa de dados média total dos usuários no DL em função de N .



Fonte: Autoral.

Figura 22 – Taxa de dados média total dos usuários no UL em função de N .



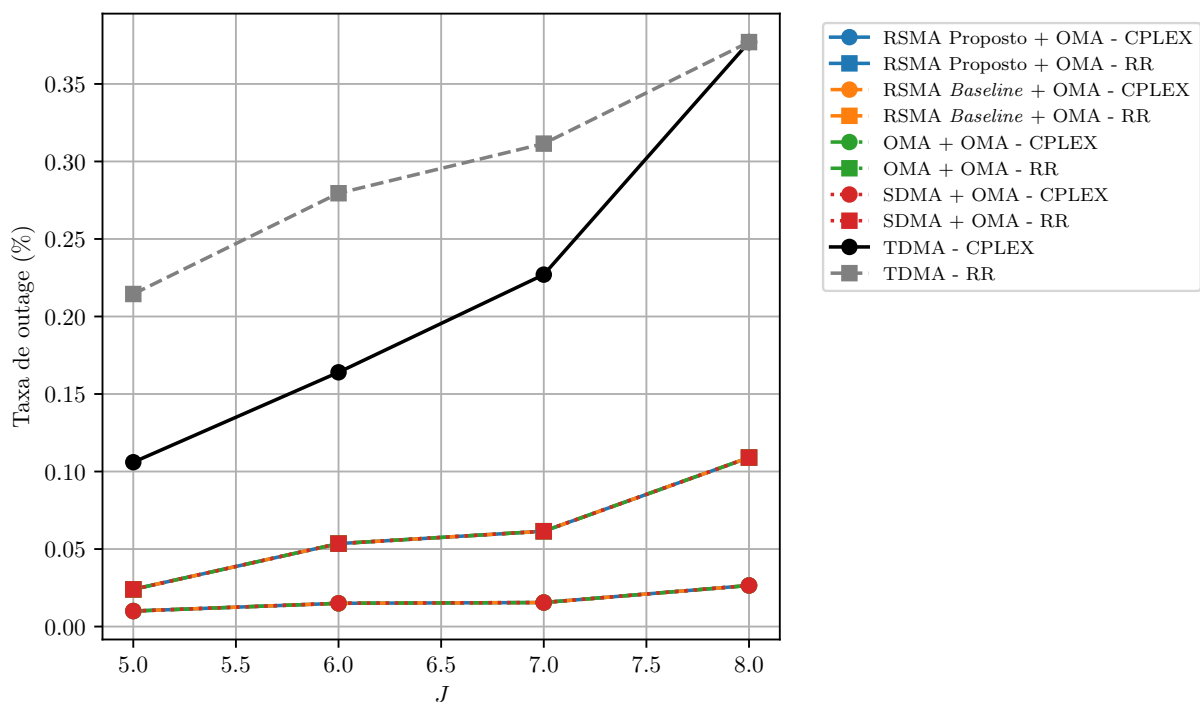
Fonte: Autoral.

4.4.3.3 Impacto do número total de usuários

No cenário final, o número de usuários no sistema, J , é variado, mantendo-se o número de *slots* fixo em $N = 8$ e o requisito de taxa em $R_{\min}^{DL} = 800$ bps. Esta análise visa compreender os efeitos da diversidade de multiusuário e do aumento da carga no sistema.

A Figura 23 demonstra que a taxa de *outage* cresce para todos os esquemas à medida que J aumenta. Este resultado é esperado, pois com um número fixo de recursos temporais ($N = 8$), o problema de satisfazer os requisitos de QoS para um número crescente de usuários torna-se inerentemente mais difícil. É de especial interesse o alargamento da lacuna de desempenho entre o CPLEX e o RR. Com mais usuários, o número de grupos possíveis ($\binom{J}{G}$) cresce significativamente, oferecendo ao escalonador CPLEX um ganho de diversidade de multiusuário. O CPLEX pode explorar este vasto espaço de escolha para encontrar combinações que satisfaçam as restrições. A heurística RR, por outro lado, sofre com o aumento de J , pois seu ciclo de escalonamento se torna mais longo, aumentando a probabilidade de alocar grupos desfavoráveis. Para $J = 8$, as curvas do TDMA convergem para um *outage* elevado, pois com 8 usuários e 8 *slots*, quase não há flexibilidade de escalonamento, forçando tanto o CPLEX quanto o RR a uma solução trivial de alocar cada usuário a um único *slot*.

Figura 23 – Taxa de *outage* em função de J .

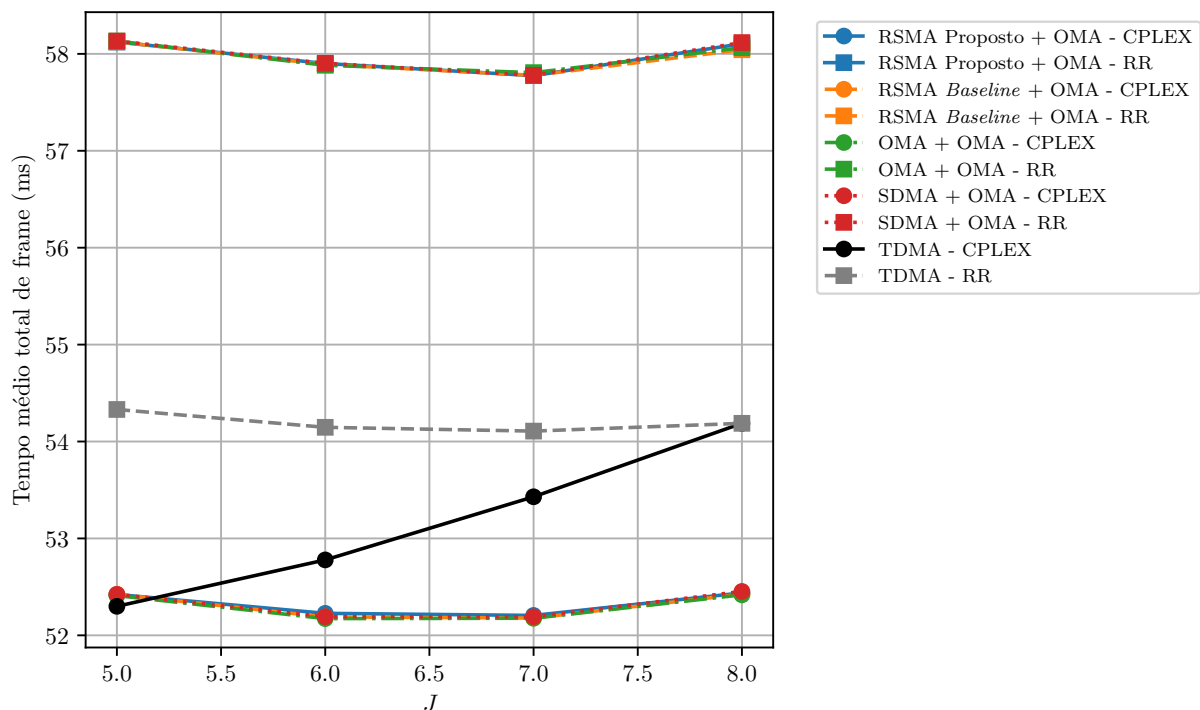


Fonte: Autoral.

O comportamento do tempo médio total, na Figura 24, revela um *trade-off* interessante. Para os esquemas com CPLEX, observa-se uma ligeira curva em "U", com o tempo mínimo ocorrendo para $J = 6$ ou $J = 7$. A explicação para esta forma reside na interação entre a diversidade de multiusuário e a rigidez das restrições. Inicialmente, ao passar de $J = 5$ para $J = 7$, o ganho de diversidade de multiusuário permite ao otimizador explorar um conjunto maior de grupos de usuários ($\binom{J}{G}$). Essa maior flexibilidade aumenta a probabilidade de encontrar grupos que são eficientes tanto em taxa de transmissão quanto em dispersão de atraso, resultando na redução do tempo de guarda médio.

No entanto, ao aumentar J para 8, a tendência se inverte. Neste ponto, o problema se torna significativamente mais restrito, pois há 8 usuários a serem atendidos em apenas 8 *slots*. Como aponta a Figura 23, este aumento na dificuldade eleva a taxa de *outage*. Consequentemente, a prioridade do escalonador muda de "otimizar o tempo" para "garantir a viabilidade a qualquer custo". Isso significa que, para evitar um *outage*, o otimizador pode ser forçado a aceitar combinações de usuários com maior dispersão de atraso — soluções que seriam descartadas como sub-ótimas em cenários menos restritos. O aumento no tempo total para $J = 8$ é, portanto, uma consequência direta da drástica redução do espaço de soluções viáveis, onde as poucas soluções restantes que cumprem a QoS exigem um tempo de guarda maior.

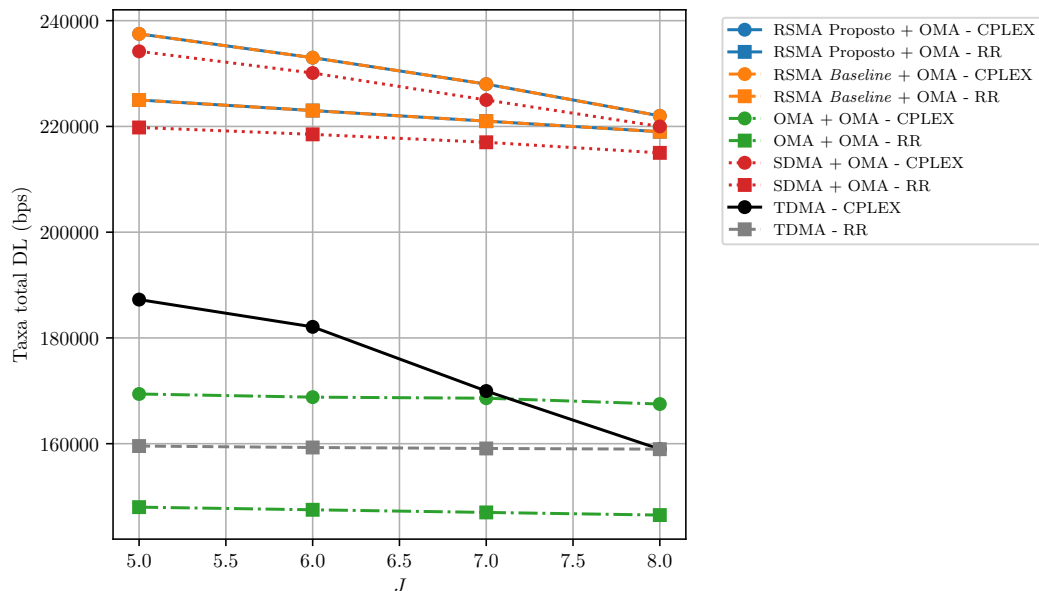
Figura 24 – Tempo médio total de *frame* em função de J .



Fonte: Autoral.

As taxas de dados totais, apresentadas nas Figuras 25 e 26, mostram uma tendência de queda com o aumento de J . Este comportamento evidencia que, sob uma carga de usuários mais elevada, o escalonador ótimo precisa se concentrar em garantir a viabilidade do sistema (atender a QoS de todos), em detrimento da maximização da taxa total. Ele não pode mais se dar ao luxo de escolher apenas os melhores usuários, precisando alocar recursos para aqueles com canais mais fracos para evitar o *outage*, o que naturalmente reduz a taxa de dados agregada do sistema. A diferença de desempenho entre o CPLEX e o RR também diminui para $J = 8$, pois o espaço de soluções viáveis se torna tão pequeno que até mesmo o escalonador ótimo tem pouca margem para manobra.

Figura 25 – Taxa de dados média total dos usuários no DL em função de J .



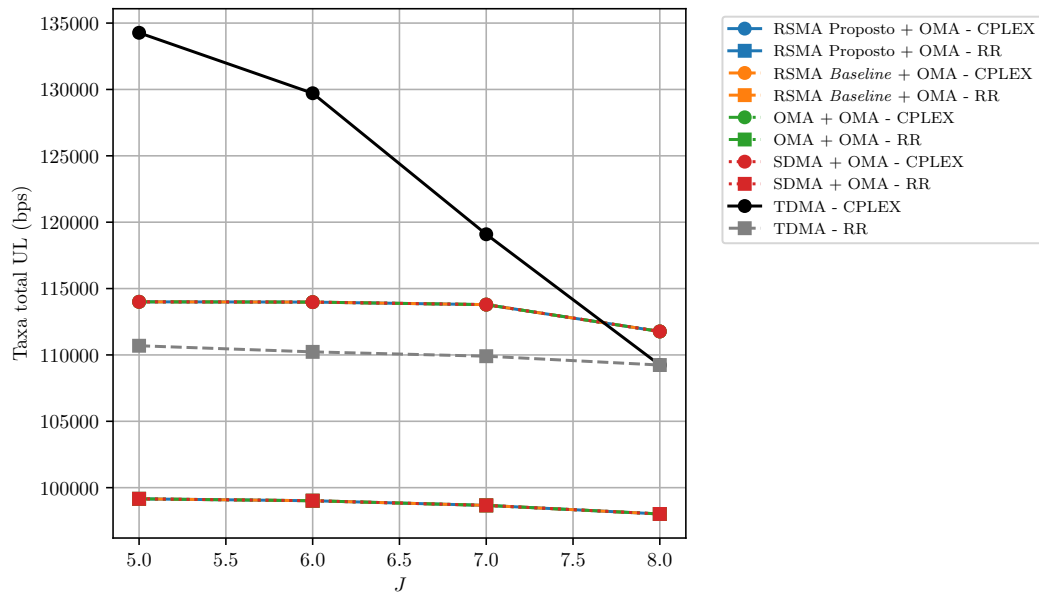
Fonte: Autoral.

4.5 Considerações parciais

Neste capítulo, foi apresentado e avaliado um problema de otimização para o escalonamento de usuários em um sistema TDD-NTN, que emprega RSMA no DL. A formulação do problema se destacou por integrar, de forma inédita, o impacto físico dos tempos de guarda, decorrentes dos atrasos de propagação, diretamente nas restrições de QoS e na função objetivo. A avaliação de desempenho, conduzida por meio de uma solução ótima via CPLEX e heurísticas de referência, permitiu extrair conclusões importantes sobre a dinâmica do sistema.

Os resultados confirmaram consistentemente a superioridade dos esquemas de múltiplo

Figura 26 – Taxa de dados média total dos usuários no UL em função de J .



Fonte: Autoral.

tiplo acesso que exploram o domínio espacial (RSMA, SDMA, OMA com $G = 2$) sobre o TDMA. A capacidade de servir múltiplos usuários simultaneamente resultou em taxas de *outage* drasticamente menores e taxas de dados agregadas significativamente maiores, reforçando a importância da multiplexação espacial para a eficiência de sistemas NTN.

Ficou evidente o valor de um escalonamento inteligente e ciente do canal. O escalonador ótimo (CPLEX) não apenas superou a heurística agnóstica RR em todas as métricas, mas também demonstrou a capacidade de explorar os *trade-offs* do sistema. A análise mostrou que o CPLEX consegue minimizar ativamente os tempos de guarda, adaptar-se a requisitos de QoS mais restritos e explorar tanto a diversidade temporal (maior N) quanto a de multiusuário (maior J) para maximizar o desempenho.

As análises dos três cenários revelaram os compromissos fundamentais do sistema. Verificou-se que (i) aumentar os requisitos de QoS degrada a confiabilidade, mas força o sistema a ser mais eficiente temporalmente; (ii) aumentar o horizonte de tempo (N) melhora a confiabilidade e permite que o escalonador otimize objetivos secundários, como a taxa de UL; e (iii) aumentar o número de usuários (J) introduz um balanço entre o ganho de diversidade de multiusuário e a maior dificuldade em satisfazer todas as restrições.

Em suma, este capítulo estabeleceu um limite superior de desempenho para o escalonamento em sistemas TDD-NTN e demonstrou que um projeto de escalonador eficiente deve, obrigatoriamente, considerar os efeitos da camada física, como os atrasos de propagação. A elevada complexidade computacional da solução ótima, no entanto, a torna impraticável para

implementações em tempo real, motivando a necessidade de desenvolver heurísticas avançadas e de baixa complexidade, capazes de se aproximar do desempenho ótimo aqui estabelecido.

5 CONSIDERAÇÕES FINAIS

Nesta dissertação, investigou-se o problema de alocação de potência em sistemas RSMA para redes LEO-terrestres, visando maximizar a soma das taxas dos usuários sob restrição de potência total e CSIT imperfeito. O problema de otimização resultante é não-convexo, e soluções existentes dependem de busca exaustiva de alta complexidade. Para superar essa limitação, esta dissertação adaptou e validou a estratégia de baixa complexidade proposta em Clerckx *et al.* (2020) para o contexto específico de redes LEO-terrestres. A contribuição consistiu em aplicar esta abordagem, que combina pré-codificação ZF e um projeto max-min, a um modelo de canal com linha de visada e CSIT imperfeito, resultando na derivação de uma expressão analítica fechada para a divisão de potência nesse cenário. Em paralelo, abordou-se a ineficiência temporal em sistemas TDD para NTN, formulando um problema misto contínuo-binário que captura o acoplamento entre escalonamento e períodos de guarda, posteriormente linearizado para permitir solução ótima.

A solução proposta de alocação de potência demonstrou desempenho de soma de taxas superior ou equivalente às abordagens de referência, com baixa complexidade computacional, especialmente em cenários com boa separação espacial. A análise do escalonamento TDD mostrou que uma otimização ciente do atraso de propagação reduz significativamente a taxa de *outage* e o tempo total de transmissão frente a heurísticas agnósticas, como o RR. O escalonador ótimo explora a diversidade multiusuário e temporal para mitigar a ineficiência dos períodos de guarda, provando que a otimização inteligente de recursos é fundamental para a viabilidade do TDD em NTN. A integração dessas contribuições evidencia que uma otimização adaptada às particularidades do ambiente satelital, tanto na camada física quanto na de acesso ao meio, é essencial para sistemas práticos e eficientes.

Para trabalhos futuros, a generalização da solução de alocação de potência para mais de dois usuários é uma direção natural, possivelmente exigindo algoritmos iterativos ou baseados em aprendizado de máquina. No que tange ao escalonamento, o alto custo computacional da solução ótima motiva o desenvolvimento de heurísticas quase-ótimas e de baixa complexidade. Estratégias para isso poderiam incluir um tratamento prévio dos grupos de usuários, onde combinações desfavoráveis (e.g., com alta dispersão de atraso ou canais fortemente correlacionados) são descartadas para reduzir o espaço de busca, ou mesmo a otimização explícita da seleção e pareamento de usuários antes do escalonamento temporal. Adicionalmente, o estudo de escalonamento assumiu o conhecimento prévio dos canais em todos os *slots* de tempo. Uma extensão

valiosa seria investigar abordagens de escalonamento dinâmico que operem com conhecimento causal, ou seja, sem prever o futuro. Nesse contexto, poderiam ser integradas estratégias de predição de canal, que explorem a trajetória previsível do satélite para estimar as condições futuras e informar decisões de escalonamento mais realistas e proativas. Além disso, investigar um *framework* para otimização conjunta de escalonamento e alocação de potência, incorporar modelos de canal mais complexos, como cenários multi-feixe com interferência inter-feixe, analisar o impacto de imperfeições práticas da camada física, como SIC imperfeito, e avaliar o impacto da justiça entre usuários no desempenho do sistema são perspectivas que tornariam o estudo mais abrangente.

REFERÊNCIAS

- 3GPP. *NR; Study on integrated access and backhaul*. [S. l.], 2018. Release 15.
- 3GPP. *Study on New Radio (NR) to support non-terrestrial networks*. [S. l.], 2020. Version 15.4.0.
- AHMAD, A. A.; MAO, Y.; SEZGIN, A.; CLERCKX, B. Rate splitting multiple access in c-ran. In: IEEE. **IEEE 31st PIMRC'20 - Workshop on Rate-Splitting and Robust Interference Management for Beyond 5G**. [S. l.], 2020.
- ALLIANCE, N. Non-terrestrial networks position paper. **project: Extreme Long-Range Communications for Deep Rural Coverage**, v. 11, 2019.
- ARTIGA, X.; VÁZQUEZ, M. 5g new radio non-terrestrial networks multiuser precoding. In: **2022 11th Advanced Satellite Multimedia Systems Conference and the 17th Signal Processing for Space Communications Workshop (ASMS/SPSC)**. [S. l.: s. n.], 2022. p. 1–8.
- AZARI, M. M.; SOLANKI, S.; CHATZINOTAS, S.; KODHELI, O.; SALLOUHA, H.; COLPAERT, A.; MONTOYA, J. F. M.; POLLIN, S.; HAQIQATNEJAD, A.; MOSTAANI, A.; LAGUNAS, E.; OTTERSTEN, B. Evolution of non-terrestrial networks from 5g to 6g: A survey. **IEEE Communications Surveys Tutorials**, v. 24, n. 4, p. 2633–2672, 2022.
- CAI, Y.; QIN, Z.; CUI, F.; LI, G. Y.; MCCANN, J. A. Modulation and multiple access for 5g networks. **IEEE Communications Surveys Tutorials**, v. 20, n. 1, p. 629–646, 2018.
- CLERCKX, B.; MAO, Y.; SCHOBER, R.; POOR, H. V. Rate-splitting unifying sdma, oma, noma, and multicasting in miso broadcast channel: A simple two-user rate analysis. **IEEE Wireless Communications Letters**, v. 9, n. 3, p. 349–353, 2020.
- DARWISH, T.; KURT, G. K.; YANIKOMEROGLU, H.; BELLEMARE, M.; LAMONTAGNE, G. Leo satellites in 5g and beyond networks: A review from a standardization perspective. **IEEE Access**, v. 10, p. 35040–35060, 2022.
- DEEPENDER; MANOJ; SHRIVASTAVA, U.; VERMA, J. K. A study on 5g technology and its applications in telecommunications. In: **2021 International Conference on Computational Performance Evaluation (ComPE)**. [S. l.: s. n.], 2021. p. 365–371.
- Ericsson. **Ericsson Mobility Report**. [S. l.], 2025. Uen Rev B.
- FILIPPO, B. D.; AMATETTI, C.; CAMPANA, R.; GUIDOTTI, A.; VANELLI-CORALLI, A. **An SCMA Receiver for 6G NTN based on Multi-Task Learning**. 2025. Disponível em: <https://arxiv.org/abs/2506.05111>.
- GERACI, G.; LÓPEZ-PÉREZ, D.; BENZAGHTA, M.; CHATZINOTAS, S. Integrating terrestrial and non-terrestrial networks: 3d opportunities and challenges. **IEEE Communications Magazine**, v. 61, n. 4, p. 42–48, 2023.
- GIORDANI, M.; POLESE, M.; MEZZAVILLA, M.; RANGAN, S.; ZORZI, M. Toward 6g networks: Use cases and technologies. **IEEE Communications Magazine**, v. 58, n. 3, p. 55–61, 2020.

HAROUNABADI, M.; HEYN, T. Toward integration of 6g-ntn to terrestrial mobile networks: Research and standardization aspects. **IEEE Wireless Communications**, v. 30, n. 6, p. 20–26, 2023.

HOSSEINIAN, M.; CHOI, J. P.; CHANG, S.-H.; LEE, J. Review of 5g ntn standards development and technical challenges for satellite integration with the 5g network. **IEEE Aerospace and Electronic Systems Magazine**, v. 36, n. 8, p. 22–31, 2021.

HSIAO, C.-L.; GUEY, J.-C.; SHEEN, W.-H.; CHEN, R.-J. A two-user approximation-based transmit beamforming for physical-layer multicasting in mobile cellular downlink systems. **Journal of the Chinese Institute of Engineers**, Taylor & Francis, v. 38, n. 6, p. 742–750, 2015.

JAAFARI, M. E.; CHUBERRE, N.; ANJUERE, S.; COMBELLES, L. Introduction to the 3gpp-defined ntn standard: A comprehensive view on the 3gpp work on ntn. **International Journal of Satellite Communications and Networking**, Wiley Online Library, v. 41, n. 3, p. 220–238, 2023.

KANG, S.; MIAO, D.; SUN, S.; CHEN, S. Tdd mode for ntn direct to satellite service. **IEEE Future Networks**, 2023.

LAKIEW, D. S.; TRAN, A.-T.; MASOOD, A.; DAO, N.-N.; CHO, S. A review on satellite-terrestrial integrated wireless networks: Challenges and open research issues. In: **2023 International Conference on Information Networking (ICOIN)**. [S. l.: s. n.], 2023. p. 638–641.

LI, H.; SUN, C.; WANG, S.; CUI, T.; WANG, X.; GONG, Y.; ZHANG, W. 5g sub-band full duplex: 3gpp standardization progress and performance analysis. In: **2024 IEEE/CIC International Conference on Communications in China (ICCC)**. [S. l.: s. n.], 2024. p. 1940–1945.

LI, H.; SUN, C.; WANG, X.; BAI, Y.; CUI, T.; WANG, S. Performance analysis of a non-terrestrial network with timing advance compensation in 3gpp. In: **2023 International Conference on Wireless Communications and Signal Processing (WCSP)**. [S. l.: s. n.], 2023. p. 851–855.

M. V. Tim Düe. **OpenNTN, open source channel models for NTN simulations**. 2025. <https://github.com/ant-uni-bremen/OpenNTN>. Disponível em 21 de janeiro de 2025, conforme citado na referência [7] do artigo.

MATHUR H., D. T. A survey on advanced multiple access techniques for 5g and beyond wireless communications. **Wireless Personal Communications**, p. 1775–1792, 2021.

MEDINA-ACOSTA, G. A.; MUNGARA, R. K.; ERIKSSON, S. G.; KHAN, T. 3gpp release-18 physical layer enhancements for iot-ntn. **IEEE Communications Standards Magazine**, v. 8, n. 3, p. 18–24, 2024.

MISHRA, A.; MAO, Y.; DIZDAR, O.; CLERCKX, B. Rate-splitting multiple access for 6g—part i: Principles, applications and future works. **IEEE Communications Letters**, v. 26, n. 10, p. 2232–2236, 2022.

NGO, H. Q. **Massive MIMO: Fundamentals and System Designs**. Tese (Doutorado) – Linköping University, Linköping, Sweden, 2015. Dissertations, No. 1642.

- RINALDI, F.; MAATTANEN, H.-L.; TORSNER, J.; PIZZI, S.; ANDREEV, S.; IERA, A.; KOUCHERYAVY, Y.; ARANITI, G. Non-terrestrial networks in 5g beyond: A survey. **IEEE Access**, v. 8, p. 165178–165200, 2020.
- SAAD, M. M.; TARIQ, M. A.; KHAN, M. T. R.; KIM, D. Non-terrestrial networks: An overview of 3gpp release 17–18. **IEEE Internet of Things Magazine**, v. 7, n. 1, p. 20–26, 2024.
- SCHRÖDER, A.; GRACLA, S.; RÖPER, M.; WÜBBEN, D.; BOCKELMANN, C.; DEKORSY, A. Flexible robust beamforming for multibeam satellite downlink using reinforcement learning. In: **ICC 2024 - IEEE International Conference on Communications**. [S. l.: s. n.], 2024. p. 3809–3814.
- SCHRÖDER, A.; ROEPER, M.; WUEBBEN, D.; MATTHIESEN, B.; POPOVSKI, P.; DEKORSY, A. A comparison between rsma, sdma, and oma in multibeam leo satellite systems. In: **WSA SCC 2023; 26th International ITG Workshop on Smart Antennas and 13th Conference on Systems, Communications, and Coding**. [S. l.: s. n.], 2023. p. 1–6.
- SHAH, A. F. M. S. A survey from 1g to 5g including the advent of 6g: Architectures, multiple access techniques, and emerging technologies. In: **2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC)**. [S. l.: s. n.], 2022. p. 1117–1123.
- SHAH, A. F. M. S.; KARABULUT, M. A.; RABIE, K. Multiple access schemes for 6g enabled ntn-assisted iot technologies: Recent developments, prospects and challenges. **IEEE Internet of Things Magazine**, v. 7, n. 1, p. 48–54, 2024.
- SHAHID, H.; AMATETTI, C.; CAMPANA, R.; TONG, S.; PANAITOPOL, D.; VANELLI-CORALLI, A.; MOHAMED, A.; ZHANG, C.; KHALIFA, E.; MEDEIROS, E.; RECAYTE, E.; GHASEMIFARD, F.; LIANGHAI, J.; BUCHELI, J.; CAUS, M.; GURELLI, M.; VAZQUEZ, M. A.; SHAAT, M.; BORIOS, N.; ERIKSSON, P.-E.; EULER, S.; LI, Z.; FU, X. Emerging advancements in 6g ntn radio access technologies: An overview. In: **2024 Joint European Conference on Networks and Communications 6G Summit (EuCNC/6G Summit)**. [S. l.: s. n.], 2024. p. 593–598.
- TEGOS, A. A.; XIAO, Y.; TEGOS, S. A.; KARAGIANNIDIS, G. K.; DIAMANTOULAKIS, P. D. Distributed uplink rate splitting multiple access (du-rsma): Principles and performance analysis. **IEEE Transactions on Wireless Communications**, p. 1–1, 2025.
- TRASPADINI, A.; GIORDANI, M.; ZORZI, M. Enhanced time division duplexing slot allocation and scheduling in non-terrestrial networks. In: **2024 58th Asilomar Conference on Signals, Systems, and Computers**. [S. l.: s. n.], 2024. p. 835–841.
- XIA, H.; MAO, Y.; ZHOU, X.; CLERCKX, B.; HAN, S.; LI, C. Weighted sum-rate maximization of rate-splitting multiple access with confidential messages. **IEEE Transactions on Wireless Communications**, v. 23, n. 10, p. 13738–13751, 2024.