



UNIVERSIDADE FEDERAL DE GOIÁS
ESCOLA DE ENGENHARIA ELÉTRICA, MECÂNICA E DE
COMPUTAÇÃO

PAULO TEALDI

**Otimização de pré-codificação conjunta
com formação de clusters em um
sistema cell-free auxiliado por RIS**

Goiânia
2025



UNIVERSIDADE FEDERAL DE GOIÁS
ESCOLA DE ENGENHARIA ELÉTRICA, MECÂNICA E DE COMPUTAÇÃO

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TRABALHO DE CONCLUSÃO DE CURSO DE GRADUAÇÃO NO REPOSITÓRIO INSTITUCIONAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio do Repositório Institucional (RI/UFG), regulamentado pela Resolução CEPEC no 1240/2014, sem ressarcimento dos direitos autorais, de acordo com a Lei no 9.610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

O conteúdo dos Trabalhos de Conclusão dos Cursos de Graduação disponibilizado no RI/UFG é de responsabilidade exclusiva dos autores. Ao encaminhar(em) o produto final, o(s) autor(a)(es)(as) e o(a) orientador(a) firmam o compromisso de que o trabalho não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

1. Identificação do Trabalho de Conclusão de Curso de Graduação (TCCG)

Nome(s) completo(s) do(a)(s) autor(a)(es)(as): Paulo Tealdi

Título do trabalho: Otimização de pré-codificação conjunta com formação de clusters em um sistema cell-free auxiliado por RIS

2. Informações de acesso ao documento (este campo deve ser preenchido pelo orientador) Concorda com a liberação total do documento SIM NÃO¹

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. Após esse período, a possível disponibilização ocorrerá apenas mediante: a) consulta ao(à)(s) autor(a)(es)(as) e ao(à) orientador(a); b) novo Termo de Ciência e de Autorização (TECA) assinado e inserido no arquivo do TCCG. O documento não será disponibilizado durante o período de embargo.

Casos de embargo:

- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro.

Obs.: Este termo deve ser assinado no SEI pelo orientador e pelo autor.



Documento assinado eletronicamente por **Rodrigo Pinto Lemos, Professor do Magistério Superior**, em 30/06/2025, às 14:21, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Paulo Tealdi, Discente**, em 30/06/2025, às 14:24, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **5470442** e o código CRC **9B0FB7FC**.

Referência: Processo nº 23070.013482/2025-04

SEI nº 5470442

PAULO TEALDI

Otimização de pré-codificação conjunta com formação de clusters em um sistema cell-free auxiliado por RIS

Trabalho de Conclusão apresentado à Coordenação do Curso de Engenharia Elétrica da Escola de Engenharia Elétrica, Mecânica e de Computação da Universidade Federal de Goiás, como requisito parcial para obtenção do título de Bacharel em Engenharia Elétrica.

Orientador: Prof. Dr. Rodrigo Pinto Lemos

Co-Orientador: Prof. Me. André Almeida Souza Coelho

Goiânia
2025

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

Tealdi, Paulo

Otimização de pré-codificação conjunta com formação de clusters em um sistema cell-free auxiliado por RIS [manuscrito] / Paulo Tealdi. - 2025.

15 f.

Orientador: Prof. Dr. Rodrigo Pinto Lemos; co-orientador André Almeida Souza Coelho.

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de Goiás, Escola de Engenharia Elétrica, Mecânica e de Computação (EMC), Engenharia Elétrica, Goiânia, 2025.

1. Pré-codificação. 2. Otimização. 3. Cluster. 4. Cell-free. 5. Superfície Refleitora Inteligente. I. Lemos, Rodrigo Pinto, orient. II. Título.

CDU 621.3



UNIVERSIDADE FEDERAL DE GOIÁS
ESCOLA DE ENGENHARIA ELÉTRICA, MECÂNICA E DE COMPUTAÇÃO

ATA DE DEFESA DE TRABALHO DE CONCLUSÃO DE CURSO

Ao(s) vinte e três dia(s) do mês de junho do ano de 2025 iniciou-se a sessão pública de defesa do Projeto Final de Curso (PFC) intitulado “Otimização de pré-codificação conjunta com formação de clusters em um sistema Cell-Free auxiliado por RIS”, de autoria de Paulo Tealdi, do curso de Engenharia Elétrica, do(a) Escola de Engenharia Elétrica, Mecânica e de Computação - EMC da UFG. Os trabalhos foram instalados pelo(a) Prof. Dr. Rodrigo Pinto Lemos (EMC/UFG) com a participação dos demais membros da Banca Examinadora: Prof. Dr. Marcelo Stehling de Castro (EMC/UFG) e Prof. Dr. Álisson Assis Cardoso (EMC/UFG), além de contar com a presença do coorientador, Engo. Me. André Almeida Souza Coelho (CERISE/UFG). Após a apresentação, a banca examinadora realizou a arguição do(a) estudante. Posteriormente, de forma reservada, a Banca Examinadora atribuiu a nota final de 10,0 (dez), tendo sido o TCC considerado aprovado.

Proclamados os resultados, os trabalhos foram encerrados e, para constar, lavrou-se a presente ata que segue assinada pelos Membros da Banca Examinadora.



Documento assinado eletronicamente por **Rodrigo Pinto Lemos, Professor do Magistério Superior**, em 23/06/2025, às 16:01, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Marcelo Stehling De Castro, Professor do Magistério Superior**, em 23/06/2025, às 16:04, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Alisson Assis Cardoso, Professor do Magistério Superior**, em 23/06/2025, às 16:04, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **André Almeida Souza Coelho, Discente**, em 30/06/2025, às 14:26, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **5452840** e o código CRC **1AB0112A**.

Otimização de pré-codificação conjunta com formação de clusters em um sistema cell-free auxiliado por RIS

Paulo Tealdi

Escola de Engenharia Elétrica, Mecânica e de Computação, Universidade Federal de Goiás
Goiânia, Goiás, Brasil. E-mail: paulotealdi@discente.ufg.br

Resumo—Este trabalho final de curso busca apresentar os fundamentos teóricos relacionados às Superfícies Refletoras Inteligentes (RIS) e à arquitetura *cell-free*, incluindo aspectos construtivos, restrições de defasagens das RIS, análise de performance e comparações. A partir disso, um algoritmo de formação de clusters voltado para redes *cell-free* auxiliadas por RIS é proposto, no qual, para cada equipamento de usuário (UE), são selecionados dinamicamente determinados pontos de acesso (APs) responsáveis pela transmissão, bem como um subconjunto de RISs que atuarão como auxiliares na comunicação. Dessa maneira, a formação dos clusters ocorre centralizada no UE, buscando otimizar o WSR (*weighted sum-rate*, ou soma ponderada das taxas) e diminuir o *overhead* de sinalização. A otimização do WSR em função do vetor de pré-codificação ativa (pré-codificação das APs) e passiva (matriz de defasagens das RIS) é feita seguindo algoritmos disponíveis na literatura que utilizam otimização alternada e, por isso, com o problema subdividido, a parte dos clusters é adicionada e resolvida aqui via algoritmo guloso.

Palavras-chave—Pré-codificação, Otimização, Cluster, Cell-free, superfície refletora inteligente (RIS).

Abstract—This final course project aims to present the theoretical foundations related to Reconfigurable Intelligent Surfaces (RIS) and the cell-free architecture, including construction aspects, phase-shift constraints of the RIS, performance analysis, and comparisons. Based on this, a clustering algorithm for RIS-aided cell-free networks is proposed, in which, for each user equipment (UE), a set of access points (APs) responsible for transmission is dynamically selected, as well as a subset of RISs that will assist in communication. In this way, cluster formation is carried out in a UE-centric manner, aiming to optimize the WSR (*weighted sum-rate*) and reduce signaling overhead. The optimization of the WSR with respect to the active precoding vector (precoding at the APs) and the passive precoding (RIS phase-shift matrix) is performed based on existing alternating optimization algorithms in the literature. Therefore, with the subdivided problem, the clustering part is introduced and solved here using a greedy algorithm.

Index Terms—Precoding, Optimization, Cluster, Cell-free, Reconfigurable Intelligent Surface (RIS).

I. INTRODUÇÃO

NOS sistemas de transmissão sem fio convencionais, com arquitetura celular, os usuários localizados nas bordas das células geralmente sofrem com interferências inter-celulares severas, causadas pelos sinais de células adjacentes. Uma ideia proposta para o 5G de melhora da capacidade

(bit/s/Hz) das redes é a implementação de redes ultra-densas (UDN), em que o tamanho das células é reduzido (*small cells*), ao mesmo tempo em que há o aumento do número de estações rádio base (BS). O problema da abordagem das UDNs está nas interferências inter-celulares que, ao aumentar a quantidade de células em busca de uma capacidade de rede maior, aumentam-se também as interferências, impedindo assim um maior avanço da performance [8]. Em vista disso, para habilitar a densificação, o *cell-free* é proposto como uma arquitetura em que não existem células e todas as estações transmissoras atuam em cooperação. Com a densificação, o *cell-free* consegue elevar a capacidade da rede através da alta quantidade de APs servindo os UEs sob controle de uma unidade de processamento centralizada (CPU), o que torna a transmissão diversa e resistente a bloqueios [1], [2].

Entretanto, apesar das vantagens citadas, em um cenário com quantidades elevadas de UEs e APs, a escalabilidade do sistema passa a ser uma grande preocupação, já que existirão muitos canais a serem estimados e processados nas CPUs, levando a altas latências e a uma baixa eficiência em geral do sistema, ou a uma baixa qualidade de serviço. Também, com a adoção das UDNs no *cell-free*, a instalação de mais APs representa um alto custo e consumo de potência [2].

Especialmente em cenários *indoor*, alguns desafios aparecem devido às altas densidades de bloqueios, que produzem efeitos negativos causados por reflexões e atenuações nos canais de transmissão. O *cell-free* é um bom caminho para solucionar os problemas citados, uma vez que as múltiplas APs instaladas conseguem servir os UEs por diversas localizações, e isso pode mitigar o desvanecimento profundo, causado por bloqueios, e manter uma alta eficiência espectral [1].

As RIS oferecem uma solução adicional aos problemas citados por conseguirem manter a qualidade de transmissão quando os canais diretos (canais entre APs e UEs) estão severamente atenuados, servindo como um canal direto virtual sem bloqueios, caso a RIS esteja estrategicamente posicionada, e também como uma substituição para um determinado número de APs, já que as RIS são uma opção de baixo custo e consumo energético para melhoria da capacidade da rede [2]. Para tal, as RIS precisam tornar a reflexão dos sinais flexível e eficiente, o que é feito através do controle de defasagens individuais de seus elementos, que habilitam a conformação de feixe (*beamforming*) dos sinais incidentes em direções arbitrárias [5].

Cada elemento é constituído por uma antena e uma impedância controlável via tensão, podendo ser, por exemplo, uma antena patch e um diodo PIN. Com a impedância conectada na antena, podem ser utilizados microcontroladores programados para ajustarem os níveis de tensão conforme os ângulos necessários para o *beamforming* em uma certa direção.

Aqui, será proposta uma solução para o problema de pré-codificação conjunta baseada no algoritmo guloso, visando resolver parte do problema de escalabilidade em sistemas *cell-free* auxiliados por RIS. Também, será desenvolvida uma investigação dos impactos do algoritmo em cenários *indoor*, conforme a modelagem de canal que será apresentada.

Serão apresentados os fundamentos teóricos sobre modelagem de canal, *phased-arrays* e *beamforming*, RIS, *cell-free* e sobre a arquitetura *two-timescale*, que são necessários para a contextualização do algoritmo proposto.

Notações: \mathbb{C} é o conjunto dos números complexos, $[\cdot]^{-1}$, $[\cdot]^T$ e $[\cdot]^H$ denotam as operações de inversa, transposta e transposta-conjugada (operador hermitiano), respectivamente; $w_{i,j}$ é o elemento na linha i e coluna j da matriz w ; $\|\cdot\|$ é a norma euclidiana de seu argumento; $\text{diag}()$ é a operação diagonal; $\mathbf{0}_A$ é a matriz quadrada de zeros com dimensões $A \times A$; \mathbf{I}_A é a matriz identidade com dimensões $A \times A$.

II. TRABALHOS RELACIONADOS

Em [2], os autores desenvolvem um esquema *two-timescale* de pré-codificação conjunta para sistemas *cell-free* auxiliados por RIS em que, no primeiro *timescale*, todos os canais são estimados (*full CSI*) para escolher um conjunto de RIS para cada usuário e, nos próximos *timescales*, apenas a estimação parcial dos canais é realizada em função das RIS selecionadas. Para resolver o problema de otimização, é desenvolvido um algoritmo baseado em relaxações nas restrições do problema para transformá-lo em um equivalente convexo. No artigo, os autores incluem um sistema de seleção apenas das RISs, e a primeira contribuição do presente trabalho é a adição de seleção de APs em conjunto com uma seleção de RISs.

Em [3], os autores introduzem um algoritmo de pré-codificação conjunta para sistemas *cell-free* distribuídos auxiliados por RIS, semelhante ao L-MMSE (*local minimum mean square error*), e reduzem significativamente o *overhead* de sinalização em comparação com a abordagem centralizada do *cell-free*. Os autores desenvolvem a análise de sinalização, que será utilizada aqui como base para as comparações entre o algoritmo proposto e as abordagens centralizada de [2] e distribuída de [3], sendo a segunda contribuição do trabalho.

Em contraste com [2] e [3], que utilizam otimização alternada, os autores em [10] realizam a mesma otimização da pré-codificação conjunta para maximização da soma das taxas dos UEs, mas utilizando aprendizado por reforço multiagente (MARL) em duas camadas com lógica *fuzzy* (FL-MARL), realizando também a seleção de APs. Os autores demonstram uma redução na capacidade de *backhaul* e ainda uma redução na complexidade computacional em comparação com sistemas MARL convencionais, mantendo uma performance semelhante.

Aqui, será adotada a abordagem centralizada do *cell-free* como em [2], mas com um algoritmo aprimorado de seleção

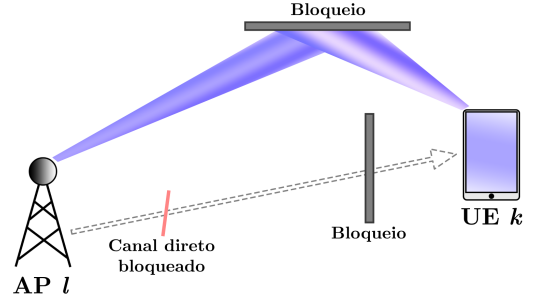


Figura 1: Propagação via reflexões em cenário com bloqueios. Fonte: Próprio autor.

que incorpora também a seleção de APs, ou seja, será realizada a seleção conjunta de APs e RISs. Tal abordagem reduz o *overhead* de sinalização, como será demonstrado.

III. FUNDAMENTOS TEÓRICOS

A. Modelo de canal

Em transmissões de sinais sem fio (*wireless*), os canais de transmissão possuem comportamento altamente aleatório devido à grande quantidade de reflexões nos obstáculos do ambiente, que podem variar em posicionamento, tamanho, material e outras características que interferem no perfil de propagação. A Figura 1 mostra um exemplo de uma transmissão com canal direto bloqueado, mas que o UE ainda consegue receber o sinal via reflexões decorrentes dos objetos no ambiente que interferem no padrão de propagação sem fio.

Em banda estreita, a resposta do canal de comunicação $h \in \mathbb{C}$ com L caminhos pode ser escrita por:

$$h = \sum_{i=1}^L \alpha_i e^{-j2\pi \frac{d_i - d}{\lambda}} \quad (1)$$

onde $d = \frac{1}{L} \sum_{i=1}^L d_i$ é a média das distâncias dos canais e d_i é a distância de um caminho de propagação i em metros. Seja $\theta_1, \dots, \theta_L$ amostras independentes de uma variável aleatória com distribuição contínua uniforme entre $-\pi$ e π e, portanto, $\theta_i \sim U[-\pi, \pi]$, e sua função densidade de probabilidade correspondente é:

$$f_{\Theta}(\theta) = \begin{cases} \frac{1}{2\pi}, & \text{se } -\pi \leq \theta < \pi \\ 0, & \text{caso contrário} \end{cases} \quad (2)$$

Também, o canal h pode ser separado em parte real e parte imaginária:

$$h = \sum_{i=1}^L \alpha_i \cos(\theta_i) - j \sum_{i=1}^L \alpha_i \sin(\theta_i) \quad (3)$$

A parte real e a parte imaginária possuem médias nulas, uma vez que trata-se da média de funções cosseno e seno. Também, elas são descorrelacionadas:

$$\begin{aligned} \mathbb{E} \left\{ \left(\sum_{i=1}^L \alpha_i \cos(\theta_i) \right) \left(\sum_{j=1}^L \alpha_j \sin(\theta_j) \right) \right\} \\ = \sum_{i=1}^L \mathbb{E} \{ \alpha_i^2 \} \mathbb{E} \{ \cos(\theta_i) \sin(\theta_i) \} = 0 \quad (4) \end{aligned}$$

Então, pelo teorema central do limite, a soma de L variáveis reais aleatórias independentes e identicamente distribuídas resulta em uma distribuição Gaussiana, para L grande. Aplicando o teorema na parte real e na parte imaginária de h , que são descorrelacionadas, resulta que a resposta de canal h em um espaço rico em reflexões é aproximadamente distribuída por uma Gaussiana complexa:

$$h \sim \mathcal{N}_{\mathbb{C}}(0, \beta) \quad (5)$$

sendo β o ganho médio de canal $\mathbb{E}\{|h|^2\} = \beta$, que também representa a variância de h , que possui média nula.

O tipo de distribuição de canal como em (5) é também conhecida como desvanecimento Rayleigh, ou *Rayleigh fading*, e o motivo é que o módulo do canal $|h|$ segue a distribuição de Rayleigh, ou seja, $|h| \sim \text{Rayleigh}(\sqrt{\beta/2})$, sendo a distribuição densidade de probabilidade dada por [5], [11]:

$$f_{|h|}(x) = \frac{2x}{\beta} e^{-\frac{x^2}{\beta}}, \quad x \geq 0 \quad (6)$$

Para $\beta = 1$, o gráfico clássico da distribuição de Rayleigh surge, como na Figura 2 [5]. Percebe-se que, para valores de módulo de canal acima de 3, apenas olhando a distribuição, é possível dizer que a densidade é nula; os valores de $|h|$ estarão mais próximos do intervalo entre 0.5 e 1.

É interessante também observar o comportamento da distribuição Rayleigh com o gráfico em escala log no eixo x , conforme está na Figura 3 [5]. Observa-se que existem duas regiões de desvanecimento profundo, ou *deep fade*, que

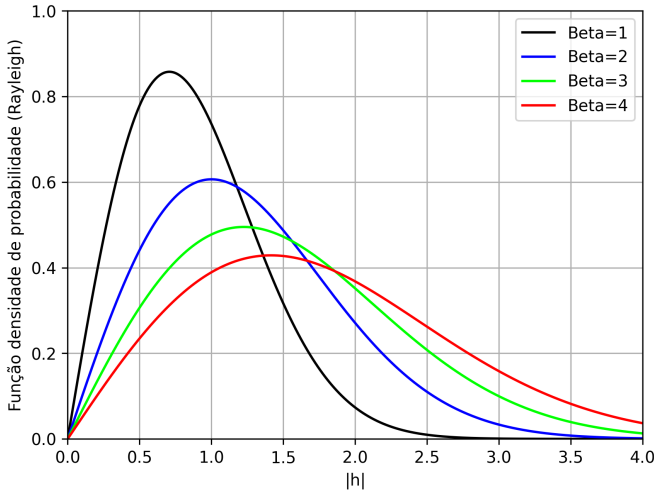


Figura 2: Função densidade de probabilidade Rayleigh. Fonte: Próprio autor.

estão aproximadamente abaixo de 0.1 e, também, acima de 3, conforme visto no outro gráfico. As regiões de *deep fade* são as regiões em que o canal entrou em obstrução completa, ou quase completa, e, portanto, são regiões com baixa qualidade de sinal.

Como visto, a derivação do canal distribuído por Rayleigh é via um somatório de reflexões, ou seja, um canal rico em caminhos NLoS (sem visada direta), mas não foi discutido o que aconteceria com a distribuição do canal se o mesmo fosse acompanhado de uma parcela LoS (visada direta). Logicamente, a componente LoS é muito mais forte que qualquer componente NLoS, uma vez que é a transmissão de canal direta, que não passa por bloqueios e reflexões. Por isso, não é possível utilizar a mesma estratégia de aplicar o teorema do limite central. Seja α_0^2 o ganho de canal LoS e $\theta_0 \sim U[-\pi, \pi)$ a defasagem, a resposta de canal é dada pela soma das parcelas LoS e NLoS:

$$h = \alpha_0 e^{-j\theta_0} + \sum_{i=1}^L \alpha_i e^{-j\theta_i} \quad (7)$$

e, para $L \rightarrow \infty$, $h \rightarrow \alpha_0 e^{-j\theta_0} + h_{\text{NLoS}}$, sendo $h_{\text{NLoS}} \sim \mathcal{N}_{\mathbb{C}}(0, \beta_{\text{NLoS}})$, que é o mesmo de antes e representa, portanto, a parcela do canal gerada pelas múltiplas reflexões. O modelo de canal descrito é chamado de desvanecimento Riciano, ou *Rician fading*, já que $|h| \sim \text{Rice}(\alpha_0, \sqrt{\beta_{\text{NLoS}}/2})$, que é a distribuição Rician [5], [11]. Finalmente, seja $\beta = \mathbb{E}\{|h|^2\} = \alpha_0^2 + \beta_{\text{NLoS}}$ o ganho médio do canal, o conhecido fator κ da distribuição de Rice é definido:

$$\kappa = \frac{\alpha_0^2}{\beta_{\text{NLoS}}} \quad (8)$$

que determina o quanto o ganho de canal LoS é mais forte do que o ganho do canal NLoS.

Na literatura existem padrões reconhecidos que definem parâmetros a serem utilizados nos modelos de desvanecimento de canal em larga escala (*large-scale fading* ou *pathloss*) e de

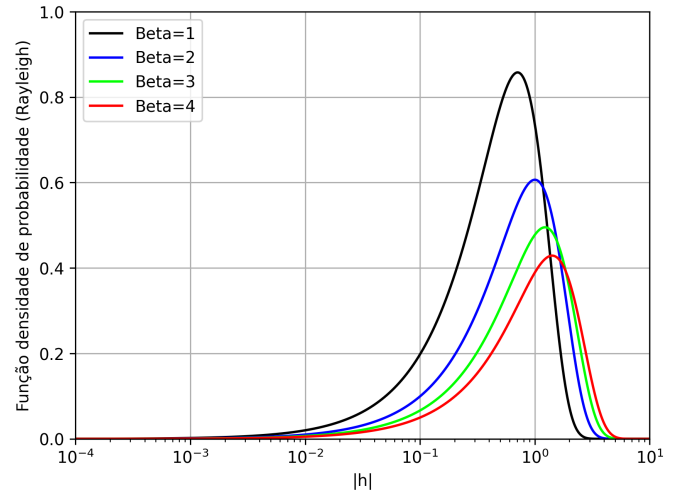


Figura 3: Função densidade de probabilidade Rayleigh com eixo x em escala log. Fonte: Próprio autor.

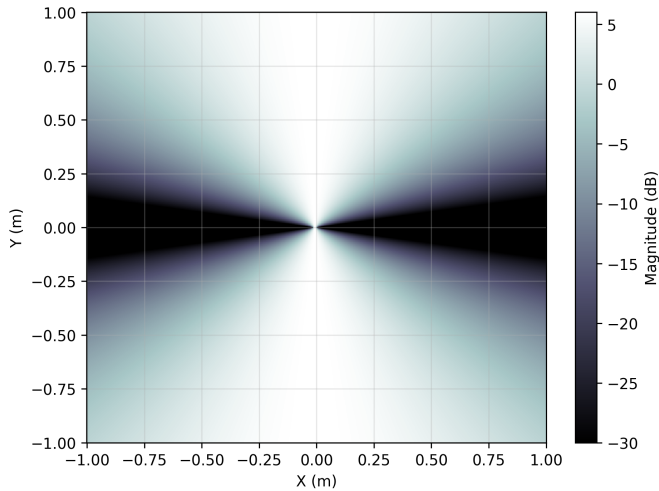


Figura 4: Transmissão em 2 antenas com defasagens iguais, o *beamforming* fica centralizado, perpendicular à linha que contém as antenas do arranjo linear. Fonte: Próprio autor.

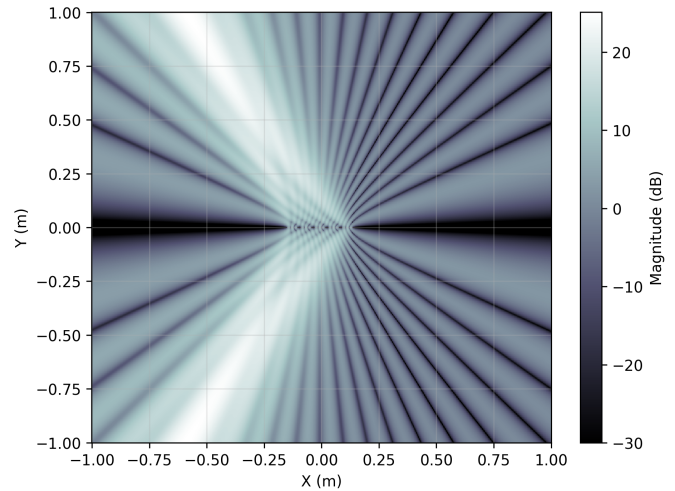


Figura 5: Transmissão em 20 antenas ULA com direcionamento em 30° . Fonte: Próprio autor.

desvanecimento de canal em pequena escala (*small-scale fading*), como o 3GPP [12]. Neste trabalho, como desenvolvido anteriormente, será utilizado o modelo denominado *Rician fading* como modelo de *small-scale fading* [5]:

$$\mathbf{H} = \sqrt{\frac{\kappa}{1+\kappa}} \mathbf{H}^{\text{LoS}} + \sqrt{\frac{1}{1+\kappa}} \mathbf{H}^{\text{NLoS}} \quad (9)$$

onde \mathbf{H}^{LoS} é a componente do canal arbitrário \mathbf{H} com visada direta (*Line of Sight*, ou LoS), \mathbf{H}^{NLoS} é a componente do canal arbitrário \mathbf{H} sem visada direta (*Non Line of Sight*, ou NLoS) e κ é o fator de Rice, que controla o quanto o canal LoS é mais forte que o NLoS. Os parâmetros do *InH-Shopping Malls-LOS* e do *InH-Shopping Malls-NLOS single slope* (FFS) de [4] são utilizados no modelo de *pathloss*:

$$\text{PL}(f, d) = 20 \log_{10} \left(\frac{4\pi f}{c} \right) + 10n \log_{10} \left(\frac{d}{1 \text{ m}} \right) + X_\sigma \quad (10)$$

onde f é a frequência, d é a distância e X_σ é o termo de *shadow fading*, seguindo distribuição log-normal com desvio padrão σ , que modela o efeito de obstáculos no ambiente. O modelo de *pathloss* descrito é utilizado em simuladores famosos da área, como o simulador especializado em RIS, o SimRIS [13].

B. Phased-arrays e beamforming

Uma antena isotrópica propaga o sinal transmitido com potência uniformemente distribuída em todas as direções, formando um perfil de propagação como uma casca esférica em expansão. Se duas antenas isotrópicas forem posicionadas em proximidade, propagando o mesmo sinal ao mesmo tempo, formam-se interferências construtivas e destrutivas no espaço, alterando o perfil de propagação do sinal, como mostra a Figura 4.

Apesar de a propagação não possuir perfil plano, é comum na literatura a adoção da hipótese do campo distante, que consiste em assumir uma distância entre transmissor e receptor

grande o suficiente para que a curvatura da frente de onda torne-se negligenciável nos cálculos. Assumindo-se a hipótese do campo distante, na recepção com um arranjo de duas antenas, tem-se a situação da Figura 6.

Na Figura 6, L representa a distância que a onda plana precisa percorrer para atingir a segunda antena após atingir a primeira e, por isso, L é a medida que representa a defasagem de recepção $\Delta\phi$. Segue que:

$$\Delta\phi = \frac{L}{\lambda} \cdot 2\pi \quad (11)$$

onde L/λ representa quantos comprimentos de onda existem de defasagem entre as antenas, e multiplicar por 2π significa converter tal valor em radianos. Como $L = d_x \sin(\varphi)$, por substituição, está demonstrada a seguinte equação:

$$\Delta\phi = \frac{2\pi d_x \sin(\varphi)}{\lambda} \quad (12)$$

onde d_x é a distância entre as duas antenas, φ é o ângulo de incidência da onda plana em relação à perpendicular do arranjo e λ é o comprimento de onda. A equação (12) fornece uma maneira de calcular quais são as defasagens necessárias para um certo arranjo de antenas de forma que ele esteja configurado em uma direção específica. O vetor de combinação das

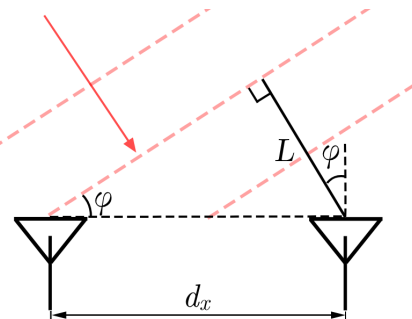


Figura 6: Diagrama de recepção em campo distante com 2 antenas. Fonte: Próprio autor.

defasagens para um arranjo linear uniforme (*Uniform Linear Array*, ULA) com N antenas, para um determinado φ de diretividade, seria dado por $[0, \Delta\phi, 2\Delta\phi, \dots, (N - 1)\Delta\phi]$. Com a diretividade em 30° , 20 antenas em arranjo linear uniforme, espaçamento de $d_x = \lambda/2$ sendo $f = 10$ GHz, o padrão de transmissão é como na Figura 5, em que já observa-se o *beamforming* na direção desejada.

A configuração de defasagens para transmissão pode ser entendida como uma pré-codificação, ou *precoding*, em que os conjuntos de antenas são configurados para direcionar a potência de transmissão aos usuários. Entretanto, não é suficiente que as antenas transmissoras estejam devidamente configuradas para direcionarem a potência ao usuário se o arranjo de antenas do usuário estiver configurado para algum outro ângulo que não o do arranjo de transmissão. Nesse sentido, também é importante configurar as defasagens das antenas receptoras para combinarem coerentemente os sinais recebidos. A configuração de defasagens para recepção é o vetor *combining*, ou *receive combining*.

Diante disso, a configuração da pré-codificação conjunta em sistemas auxiliados por RIS é fundamental, já que agora é necessário configurar as defasagens das antenas transmissoras e dos elementos da RIS, além de configurar o *receive combining* nas antenas receptoras para receberem sinais tanto das RIS quanto das APs. Apesar disso, no algoritmo de otimização dos *precodings* utilizado, desenvolvido em [2], a otimização do *receive combining* não é considerada, por simplicidade, o que reduz uma etapa do problema de otimização via abordagem alternada (*Alternating Optimization*, AO).

Estudando agora a importância do *receive combining* em conjunto com o *precoding*, configurando-se o cenário da Figura 5 com uma outra ULA receptora de 20 antenas em um plano paralelo ao da transmissora, localizada conforme a Figura 7, altera-se a configuração do *receive combining* em vários ângulos, de -90° até 90° , e o comportamento esperado é observar um pico de recepção no ângulo que corresponde à direção de localização da transmissora. O resultado da

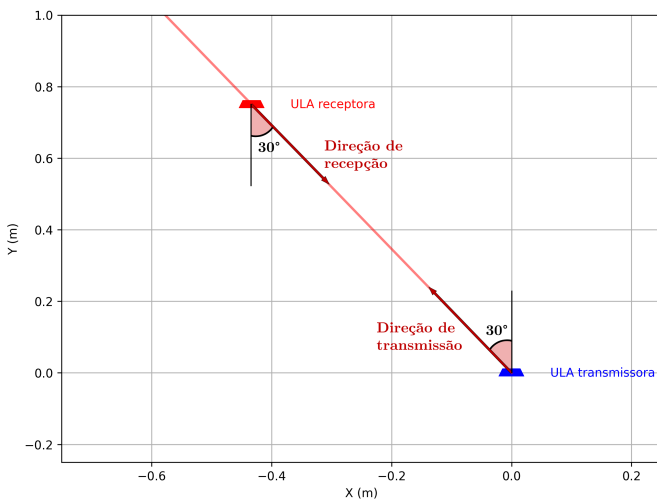


Figura 7: Localizações para experimento de estimação de ângulo de chegada (*Angle of Arrival*, AoA), com transmissão em 30° . Fonte: Próprio autor.

varredura está na Figura 8.

Percebe-se na 8 o resultado esperado: como o *precoding* está configurado na ULA transmissora para direcionar a potência no ângulo de 30° em sentido anti-horário partindo da perpendicular do arranjo, o arranjo ULA de recepção também precisa configurar suas antenas na direção dos mesmos 30° , que é de onde está chegando o sinal.

Sobre a polaridade do resultado, no código das simulações, a convenção utilizada é que as defasagens são aplicadas da antena na extremidade esquerda até a da extremidade direita e, por isso, para gerar um *beamforming* na ULA transmissora no sentido dado pela linha vermelha da Figura 7, configura-se um conjunto de defasagens para direcionar em um ângulo de -30° e, para receber o sinal, configuram-se as defasagens para o ângulo de $+30^\circ$, conforme o resultado obtido.

Também, é importante lembrar que o cenário é contido em duas dimensões e, por isso, a estimação do ângulo de chegada torna-se bastante simples, já que é suficiente realizar a varredura de apenas 1 ângulo. Em cenários tridimensionais, caso fosse necessário identificar a localização de uma antena transmissora de arranjo planar uniforme (UPA), por exemplo, via recepção de seu *beamforming*, seria necessário instalar uma quantidade maior de antenas receptoras para fornecer uma diferença de defasagem no outro eixo, uma vez que, com uma ULA, não seria possível estimar um ângulo que não esteja contido em seu próprio plano.

Naturalmente, para as UPAs, tem-se um caso bastante semelhante ao equacionamento das ULAs, mas agora em duas dimensões. Pode-se considerar que, em uma das dimensões do plano da UPA, tem-se uma ULA, e assim constrói-se uma matriz de defasagens correspondente a essa dimensão. Para a outra dimensão, tem-se outra matriz de defasagens, novamente como se fosse a ULA, e tem-se outra matriz de defasagens. As duas matrizes de defasagem são então combinadas através de uma soma simples, conforme a Figura 9, ou caso estejam na forma $e^{j\Delta\phi_n}$, a combinação pode ser feita através de um produto *element-wise*, ou produto *Hadamard*, em que as fases seriam somadas da mesma forma $e^{j(\Delta\phi_n + \Delta\phi_m)}$. Na Figura 9,

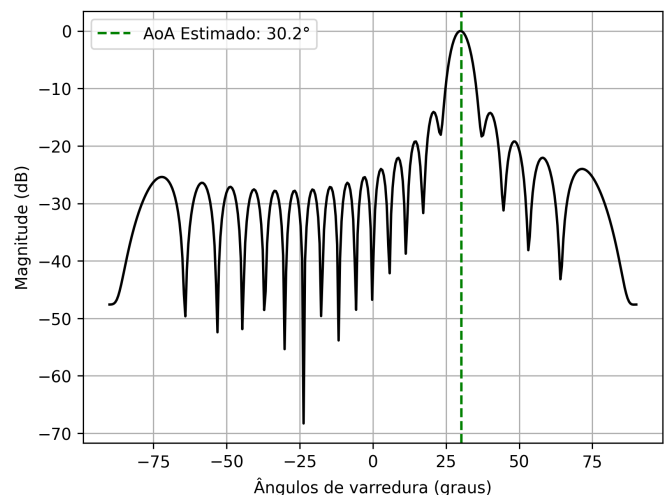


Figura 8: Resultado da estimação de ângulo de chegada (*Angle of Arrival*, AoA), com transmissão em 30° . Fonte: Próprio autor.

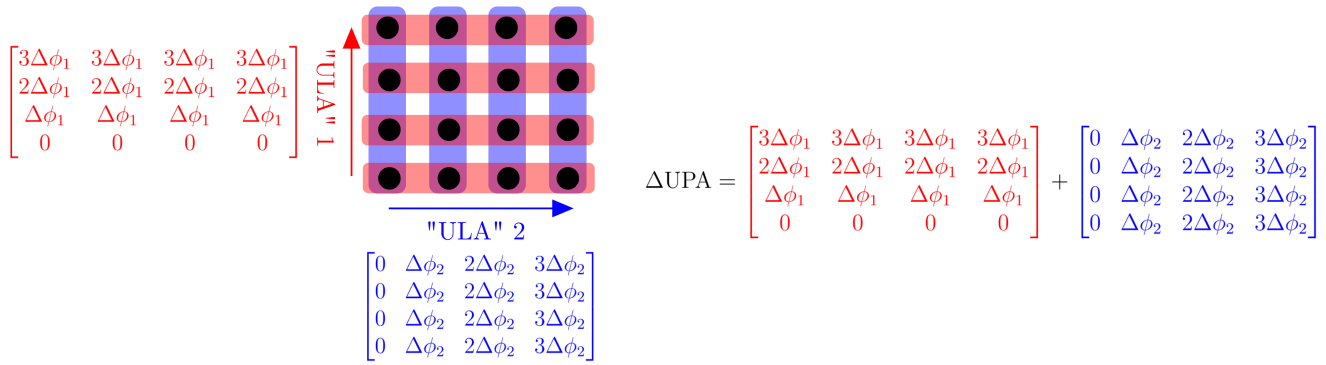


Figura 9: Defasagens de uma UPA com 16 antenas a partir de defasagens de ULAs. Fonte: Próprio autor.

os elementos com fase nula, de referência, são utilizados como os primeiros de baixo para cima ou os primeiros da esquerda para a direita, mas é uma convenção flexível, é possível adotar como referência os primeiros da direita para a esquerda e os primeiros de cima para baixo.

Em outras palavras, a ideia é que, com uma UPA, existem dois graus de liberdade de defasagens que formam a rotação do feixe, um em cada dimensão do arranjo. Assim, para formar uma base de rotações, são necessários dois vetores ortogonais para suprir as duas dimensões do arranjo. O equacionamento a seguir esclarece a ideia.

Seja cada elemento da UPA localizado em $(x_m, y_n) = (m \cdot d_x, n \cdot d_y)$, $m = 0, 1, \dots, M$, φ o ângulo azimutal e θ o ângulo de elevação. O vetor unitário que aponta na direção de propagação é dado por:

$$\vec{u} = \begin{bmatrix} \sin \theta \cos \varphi \\ \sin \theta \sin \varphi \\ \cos \theta \end{bmatrix} \quad (13)$$

As componentes de \vec{u} que estão no plano da UPA são então utilizadas para cálculo da fase do direcionamento, resultando em:

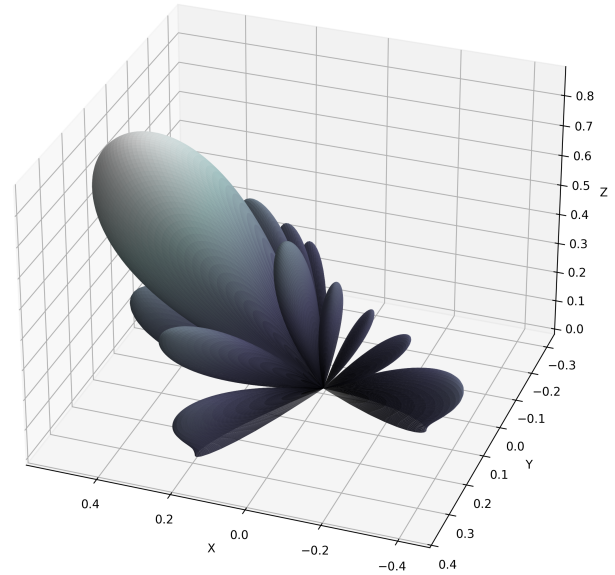
$$\psi_{m,n} = -k(x_m \sin \theta \cos \varphi + y_n \sin \theta \sin \varphi) \quad (14)$$

onde $k = 2\pi/\lambda$ é o número de onda. Por fim, a fase é computada por:

$$e^{j\psi_{m,n}} = e^{jk(x_m \sin \theta \cos \varphi + y_n \sin \theta \sin \varphi)} \quad (15)$$

Utilizando o que foi desenvolvido, a Figura 10 demonstra um *beamforming* realizado por uma UPA de 64 antenas (8x8) na frequência de 10 GHz, espaçamento $\lambda/2$ e ângulo de propagação de 30° de elevação e 20° de azimute.

Agora, com o desenvolvimento das UPAs finalizado, segue naturalmente o desenvolvimento das RIS, que também são UPAs (geralmente; aqui serão UPAs em todos os casos), mas que não transmitem sinais, e sim refletem de maneira controlada. As RIS podem também assumir diferentes formatos, como nas recentes BD-RIS [15], em que há a interconexão dos elementos da RIS e, por isso, é bastante justificada e vantajosa


 Figura 10: *beamforming* utilizando uma UPA com o plano do arranjo em xy. Fonte: Próprio autor.

a fabricação de RIS com diferentes formatos, ou também na utilização de antenas fluidas [14], que também não possuem arranjo planar, e, por último, as RIS sobrepostas em várias camadas [16], que podem atuar como redes neurais físicas.

C. Superfícies Refletoras Inteligentes

A tecnologia MIMO (*Multiple-Input Multiple-Output*) consegue elevar consideravelmente a força de transmissão dos sinais, empregando várias antenas nos transmissores e receptores, utilizadas para o *precoding* e *combining*. Entretanto, no caso em que a condição dos canais é desfavorável, como em situações de bloqueio, a transmissão torna-se muito dependente de reflexões, em grande parte severamente atenuadas devido ao *pathloss* e perdas por absorção nos materiais.

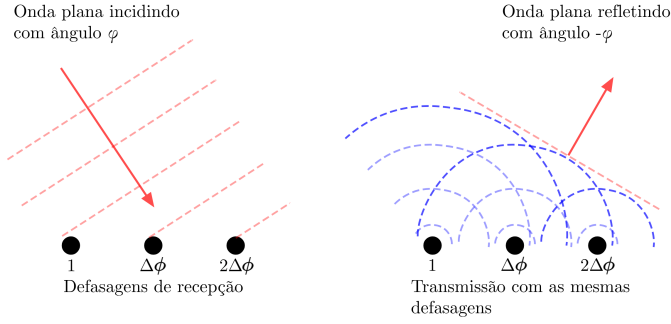


Figura 11: Onda plana incidindo em uma superfície plana na direção φ , as defasagens podem ser encontradas visualizando a superfície como uma ULA. Com as mesmas defasagens, a onda é transmitida na direção $-\varphi$. Fonte: Próprio autor.

Na reflexão em uma superfície plana qualquer, de acordo com o princípio de Huygens-Fresnel, cada ponto do objeto de incidência de uma frente de onda atua como um novo transmissor, emitindo ondas esféricas (transmissão isotrópica). No caso da Figura 11, como a onda plana incide anteriormente na extremidade esquerda, ela retransmite primeiro o sinal, e a extremidade direita retransmite por último, provocando assim a troca de direção, ou seja, as retransmissões nos pontos ocorrem com as mesmas defasagens das recepções.

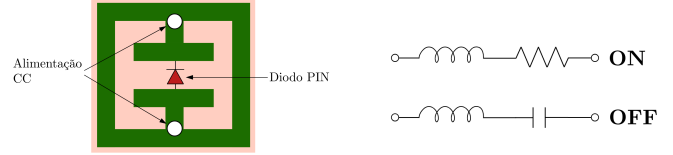
As RIS atuam contrariando o comportamento descrito de que as retransmissões ocorrem com as mesmas defasagens, os elementos que recebem o sinal possuem impedâncias variáveis que, se devidamente configuradas, podem alterar o perfil de propagação da reflexão, controlando, por exemplo, o ângulo do feixe para um novo ângulo arbitrário. Dessa forma, as RIS disponibilizam um caminho de reflexão alternativo para a transmissão dos sinais que auxilia os sistemas de comunicação *wireless* em situações de NLOS, tornando as transmissões via reflexões mais eficientes. Portanto, idealmente, a RIS atua como um espelho dinamicamente controlável e sem perdas, retransmitindo os sinais na direção desejada.

O primeiro passo para investigar o comportamento das RIS é estudar o coeficiente de reflexão de cada um dos elementos da superfície, que podem ser dinamicamente ajustados através de uma tensão de entrada que controla a reatância do elemento. Será visto como a reatância afeta as defasagens de reflexão e, por consequência, como a RIS reconfigura o *beamforming* [5]. Seja Z_0 a impedância do primeiro meio, o espaço livre, e Z_n a impedância do segundo meio, o elemento n da RIS. Então:

$$\Gamma_{0n} = \frac{Z_n - Z_0}{Z_n + Z_0} \quad (16)$$

e, portanto, se o elemento da RIS possuir impedância de característica puramente reativa (resistência nula) $Z_n = jX_n$, para algum $X_n \in \mathbb{R}$, o módulo do coeficiente de reflexão é unitário, conforme segue:

$$|\Gamma_{0n}| = \left| \frac{jX_n - Z_0}{jX_n + Z_0} \right| = \frac{\sqrt{X_n^2 + Z_0^2}}{\sqrt{X_n^2 + Z_0^2}} = 1 \quad (17)$$



(a) Circuito de uma RIS fabricada com diodos PIN (b) Circuito equivalente do diodo PIN

Figura 12: Elemento de RIS fabricada com diodos PIN. Fonte: Próprio autor.

sendo o ângulo do coeficiente de reflexão dado por:

$$\begin{aligned} \arg(\Gamma_{0n}) &= \arg\left(\frac{jX_n - Z_0}{jX_n + Z_0}\right) \\ &= \begin{cases} \pi - 2 \arctan\left(\frac{X_n}{Z_0}\right), & \text{se } X_n \geq 0 \\ -\pi - 2 \arctan\left(\frac{X_n}{Z_0}\right), & \text{se } X_n < 0 \end{cases} \quad (18) \end{aligned}$$

Um elemento da forma descrita não insere perdas na reflexão, já que o módulo do coeficiente é unitário, mas insere uma defasagem em função da reatância X_n e, com um X_n controlável, a defasagem também torna-se controlável. Tal controle pode ser feito por um diodo varicap, que é um dispositivo eletrônico que possui camada de depleção variável em função da tensão aplicada, ou seja, pode ser utilizado como uma capacitância variável. O controle de defasagens depende da fabricação da RIS, é possível que os níveis de defasagem não sejam contínuos, mas discretos, como em uma RIS fabricada para funcionar com 2 bits de defasagem, contendo assim 4 ângulos disponíveis, como $0, \pi/2, \pi$ e $3\pi/2$, ou com apenas 1 bit (*On/Off*), contendo assim 2 ângulos disponíveis, como 0 e π . Uma RIS fabricada com diodos PIN, por exemplo, seria como mostra a Figura 12a, onde a impedância equivalente do diodo PIN utilizado é como mostra a Figura 12b.

Com tais configurações, o canal completo com as reflexões pela RIS inclusas fica da seguinte forma:

$$h = h_s + \sum_{n=1}^N h_{r,n} e^{j\theta_n} h_{t,n} \quad (20)$$

onde $h_s \in \mathbb{C}$ é o canal estático, ou canal direto entre o AP e o UE, $h_{r,n} \in \mathbb{C}$ é o canal entre o elemento da RIS e o UE, e $h_{t,n} \in \mathbb{C}$ é o canal entre o elemento da RIS e o AP. Utilizando notação matricial, tem-se

$$h = h_s + \mathbf{h}_r^T \Theta \mathbf{h}_t \quad (21)$$

sendo $\mathbf{h}_t = [h_{t,1} \cdots h_{t,N}]^T$, $\mathbf{h}_r = [h_{r,1} \cdots h_{r,N}]^T$ e $\Theta = \text{diag}(e^{j\theta_1}, \dots, e^{j\theta_N})$.

A ideia principal é maximizar o módulo do canal completo h , uma vez que a matriz Θ da RIS fornece a possibilidade de ajustar as reflexões para interferirem construtivamente no receptor juntamente ao sinal enviado via canal direto. Ou seja, é necessário maximizar como na equação (19), em que a desigualdade é obtida pelo teorema de Cauchy-Schwarz [5].

O máximo é obtido observando-se que o ângulo do primeiro elemento é $\arg(h_s)$, que é a defasagem do canal estático, e os

$$|h_s + \mathbf{h}_r^T \Theta \mathbf{h}_t|^2 \leq \left\| \begin{bmatrix} \sqrt{|h_s|} \\ \sqrt{|h_{r,1} h_{t,1}|} \\ \vdots \\ \sqrt{|h_{r,N} h_{t,N}|} \end{bmatrix} \right\|^2 \left\| \begin{bmatrix} \sqrt{|h_s|} e^{j \arg(h_s)} \\ \sqrt{|h_{r,1} h_{t,1}|} e^{j(\arg(h_{r,1} h_{t,1}) + \theta_1)} \\ \vdots \\ \sqrt{|h_{r,N} h_{t,N}|} e^{j(\arg(h_{r,N} h_{t,N}) + \theta_N)} \end{bmatrix} \right\|^2 = \left(|h_s| + \sum_{n=1}^N |h_{r,n} h_{t,n}| \right)^2 \quad (19)$$

demais são dependentes dos ângulos dos canais em cascata que passam pela RIS, juntamente ao ângulo configurável da RIS θ_n . O módulo do vetor de canal completo resultante da soma entre o canal estático e as várias reflexões pela RIS é máximo quando todos esses vetores estão alinhados (Cauchy-Schwarz), então configura-se θ_n para que

$$\begin{aligned} e^{j(\arg(h_{r,n} h_{t,n}) + \theta_n)} &= e^{j \arg(h_s)} \\ \arg(h_{r,n} h_{t,n}) + \theta_n &= \arg(h_s) \\ \theta_n &= \arg(h_s) - \arg(h_{r,n} h_{t,n}) + 2\pi k_n \end{aligned} \quad (22)$$

Essa é a solução para o problema de otimização do ganho de canal por meio da configuração de defasagens da RIS. Se $|h_{t,n}| = \sqrt{\beta_t}$, $|h_{r,n}| = \sqrt{\beta_r}$ e $|h_s| = \sqrt{\beta_s}$, tem-se que o ganho de canal é dado por [5]:

$$(\sqrt{\beta_s} + N \sqrt{\beta_r \beta_t})^2 = \begin{cases} N^2 \beta_r \beta_t, & \text{se } \beta_s \ll N^2 \beta_r \beta_t \\ \beta_s, & \text{se } \beta_s \gg N^2 \beta_r \beta_t \end{cases} \quad (23)$$

Pelo resultado, é possível perceber que o ganho fornecido pela RIS depende de N^2 , em que N é a quantidade de elementos da RIS, que depende de seu tamanho. O N aparece ao quadrado devido à sua contribuição tanto no ganho de *beamforming* na transmissão, que depende da quantidade de elementos, quanto no ganho de abertura na recepção, que depende de seu tamanho. A conclusão da análise é que, com canais estáticos fortes, a RIS apresenta pouca influência positiva no ganho de canal do sistema, e com um canal estático fraco, ela pode elevar substancialmente o ganho, a depender de seu tamanho e quantidade de elementos.

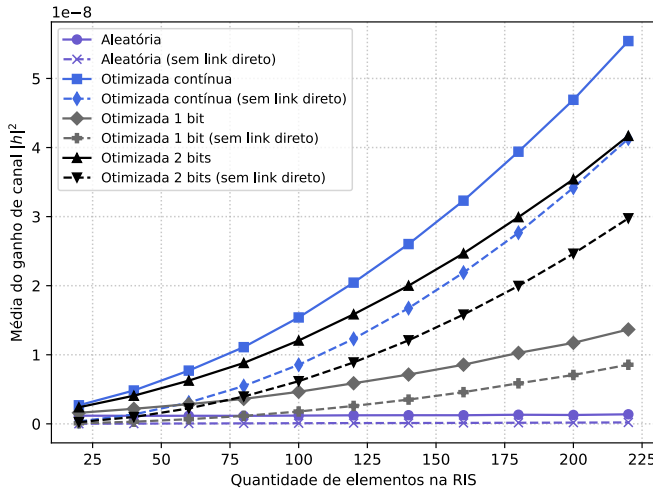


Figura 13: Análise da influência da quantidade de elementos da RIS no ganho de canal de sistemas SISO *indoor*. Fonte: Próprio autor.

Agora, em acordo com o que foi desenvolvido, serão feitas simulações explorando a otimização de 3 RISs diferentes: uma com defasagens contínuas, outra com 1 bit de configuração e outra com 2 bits de configuração. As simulações serão em um sistema SISO, com 1 AP, 1 RIS e 1 UE, sendo a RIS com $M = 64$ elementos, $N_t = 1$ antenas de transmissão e $N_r = 1$ antenas de recepção.

Para a primeira simulação, o UE está posicionado em (10, 0) m com altura $h_{UE} = 1.5$ m, o AP está posicionado em (0, 0) m com altura $h_{AP} = 2$ m e a RIS está posicionada em (0, 5) m com altura $h_{RIS} = 3$ m. O que está variando é a quantidade de elementos da RIS, entre 20 e 240 elementos, e calcula-se o ganho de canal $|h|^2$, com h dado como na equação (21). As estatísticas de canal são as estatísticas para cenários *indoor* que foram descritas na seção de modelo de canal.

A Figura 13 permite visualizar o aumento no ganho do canal completo para diferentes tipos de RIS com diversas quantidades de elementos. Percebe-se que o crescimento na quantidade de elementos exibe padrão quadrático, que é um resultado já descrito matematicamente aqui, na equação (23), e, portanto, é o comportamento esperado.

Na Figura 14, é investigado o cenário em que a RIS se move de 0 a 50 m no cenário, agora fixando o número de elementos da RIS em $M = 81$, o posicionamento do usuário em (0, 0) m, e o posicionamento do AP em (50, 0) m, todos com as mesmas alturas da simulação anterior. O resultado da simulação permite visualizar um efeito conhecido e apresentado na literatura sobre RIS, que é a respeito do bom posicionamento médio da superfície [5]. Em termos médios, os melhores posicionamentos das RIS são em máxima

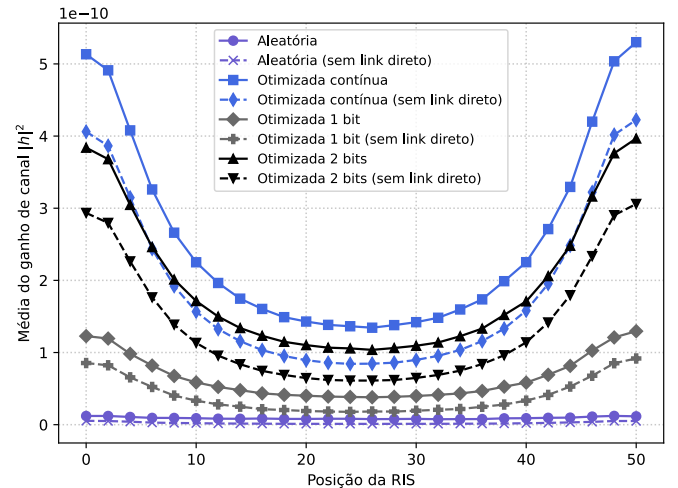


Figura 14: Análise de influência da proximidade da RIS ao UE e à AP. Fonte: Próprio autor.

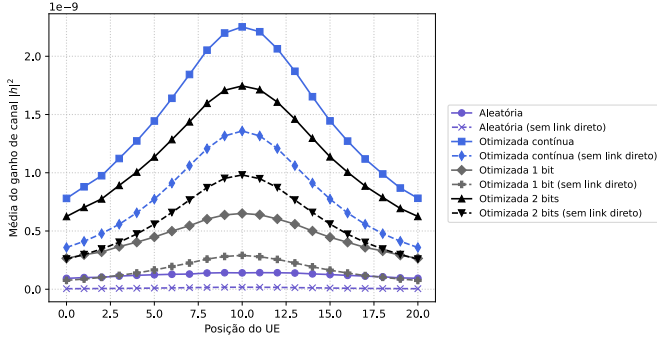


Figura 15: Análise de movimento do UE pelo cenário SISO *indoor*. Fonte: Próprio autor.

proximidade aos UEs ou aos APs, conforme demonstra o gráfico, que apresenta seus picos em 0 e 50 metros, que são os pontos em que estão localizados, respectivamente, o UE e o AP.

Por último, será investigado, na Figura 15, o cenário em que o UE se move de 0 a 20 m no cenário, mesmos parâmetros da simulação anterior, mas o AP agora está em (10, -20) m e a RIS em (10, 5) m. O motivo de o AP ter sido colocado a 20 metros do centro no eixo y é para forçar o destaque do efeito da RIS, onde é gerada uma situação de enfraquecimento do canal direto pela distância maior, uma vez que, como discutido anteriormente, as RIS conseguem ajudar mais quando o canal direto está prejudicado. A simulação é, novamente, uma visualização do efeito da proximidade da RIS ao UE, como na simulação anterior. Percebe-se que a RIS aleatória em todas as simulações não apresenta qualquer influência positiva perceptível na performance do sistema, o que destaca a importância de bons algoritmos de otimização para realizarem as configurações dinâmicas desses equipamentos.

Agora, como em [5], será apresentado um modelo básico para a resposta ao impulso de um elemento da RIS com impedância controlada via diodo varicap, que é um dispositivo comumente utilizado em circuitos de telecomunicações, como sintonizadores e osciladores. O circuito utilizado é da Figura 16 e a impedância equivalente do elemento é dada por:

$$Z_n(f, V_n) = \frac{j2\pi f L_1(j2\pi f L_2 + R + \frac{1}{j2\pi f C(V_n)})}{j2\pi f L_1 + (j2\pi f L_2 + R + \frac{1}{j2\pi f C(V_n)})} \quad (24)$$

sendo V_n a tensão de bias do varicap. Utilizando a expressão (16), do coeficiente de reflexão, com $Z_0 = 377 \Omega$ (impedância característica do vácuo) [17], $R = 1 \Omega$, $L_1 = 2.5 \text{ nH}$, $L_2 = 0.7 \text{ nH}$ e frequências em torno de uma portadora com $f_c = 3 \text{ GHz}$, obtém-se o resultado da Figura 17.

D. Cell-free

Os sistemas de comunicação atuais são construídos com base na abordagem celular, em que os usuários (UEs) são servidos pelas APs com os melhores canais, que geralmente são as APs mais próximas, e a região de proximidade dessas APs determina uma célula, como será visto a seguir. No sistema celular, as APs são equipadas com uma quantidade

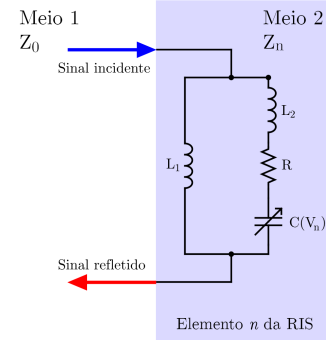


Figura 16: Elemento de RIS fabricada com diodos varicap [5]. Fonte: Próprio autor.

elevada de antenas que atuam em conjunto para gerar feixes de propagação, cada um com o sinal direcionado ao usuário adequado. Já no *cell-free*, as APs possuem uma quantidade menor de antenas, mas existe uma quantidade elevada de APs que trabalham em conjunto para gerar interferências construtivas nos UEs. Nesse sentido, não existem células e todas as APs servem a todos os UEs em cooperação.

Um sistema celular com células pequenas, chamado *small cell* na literatura, na fase de transmissão *uplink* com L APs e 1 UE, teria sinal recebido $y_l^{\text{small-cell}} \in \mathbb{C}$ dado por

$$y_l^{\text{small-cell}} = h_l s + n_l \quad (25)$$

sendo $s \in \mathbb{C}$ a informação, $n_l \sim \mathcal{N}_{\mathbb{C}}(0, \sigma_{\text{ul}}^2)$ o ruído de recepção e $\mathbb{E}\{|s|^2\} = p$ a potência do sinal. Então, a SNR, dada pela razão entre a potência recebida do sinal e a potência do ruído, resulta em:

$$\text{SNR}_l^{\text{small-cell}} = \frac{\mathbb{E}\{|h_l s|^2\}}{\mathbb{E}\{|n_l|^2\}} = \frac{p}{\sigma_{\text{ul}}^2} |h_l|^2 \quad (26)$$

e o UE então é configurado para receber a transmissão de uma das APs, uma vez que, no sistema celular, não há cooperação entre APs. Naturalmente, o AP escolhido para cada usuário é o que fornece o maior valor de SNR, então segue que a SNR do UE é:

$$\text{SNR}^{\text{small-cell}} = \frac{p}{\sigma_{\text{ul}}^2} \max_{l \in \{1, \dots, L\}} |h_l|^2 \quad (27)$$

Portanto, fica explícito que a maior SNR é dada pelo maior ganho de canal $|h_l|^2$, que é também o mesmo que escolher as APs com as menores distâncias aos usuários (em média) e, por isso, em sistemas celulares, as células de cada AP são regiões de proximidade.

Já no sistema *cell-free*, com as mesmas L APs, agora há a cooperação entre as APs e, então, os sinais recebidos da equação (25) são agora dados por:

$$\mathbf{y}^{\text{cell-free}} = \sum_{l=1}^L (h_l s + n_l) = \mathbf{h} s + \mathbf{n} \quad (28)$$

onde $\mathbf{h} = [h_1 \dots h_L]^T$ e $\mathbf{n} = [n_1 \dots n_L]^T$.

Para estimar o sinal s , um vetor *receive combining* $\mathbf{v} \in \mathbb{C}^L$ é aplicado no sinal recebido, resultando em:

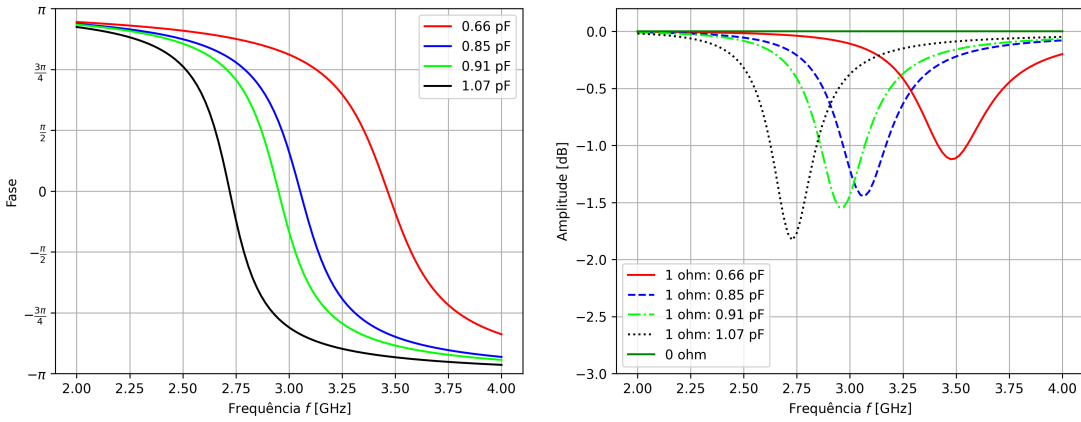


Figura 17: Resposta de amplitude e fase em função da frequência e capacitância do varicap [5]. Fonte: Próprio autor.

$$\hat{\mathbf{s}}^{\text{cell-free}} = \mathbf{v}^H \mathbf{y}^{\text{cell-free}} = \mathbf{v}^H \mathbf{h}_s + \mathbf{v}^H \mathbf{n} \quad (29)$$

e, para maximizar a SNR, utiliza-se o *combining* MR (*maximum-ratio*) [1] dado por $\mathbf{v} = \mathbf{h}/\|\mathbf{h}\|$:

$$\text{SNR}^{\text{cell-free}} = \frac{p}{\sigma_{ul}^2} \|\mathbf{h}\|^2 = \frac{p}{\sigma_{ul}^2} \sum_{l=1}^L |h_l|^2 \quad (30)$$

Agora, com $\text{SNR}^{\text{small-cell}}$ e $\text{SNR}^{\text{cell-free}}$, é possível comparar as duas abordagens. Considerando-se também as interferências, no denominador da expressão de SNR, adicionam-se as potências de interferência:

$$\begin{aligned} \text{SINR}^{\text{small-cell}} &= \frac{p}{\sigma_{ul}^2 + \sum_{\substack{i=1 \\ i \neq k}}^L |h_i|^2} \max_{l \in \{1, \dots, L\}} |h_l|^2 \\ \text{SINR}^{\text{cell-free}} &= \frac{p}{\sigma_{ul}^2} \sum_{l=1}^L |h_l|^2 \end{aligned} \quad (31)$$

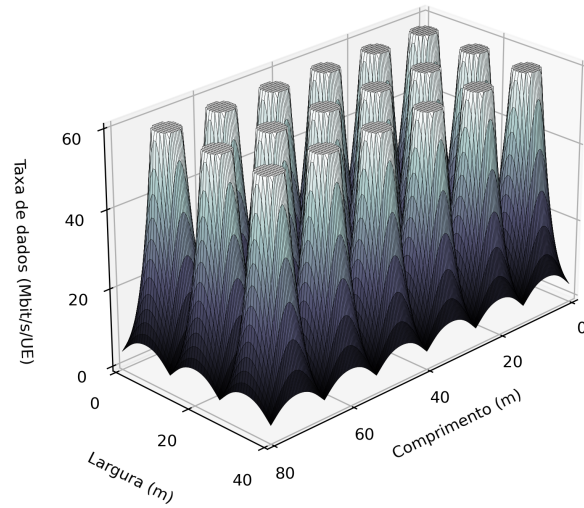
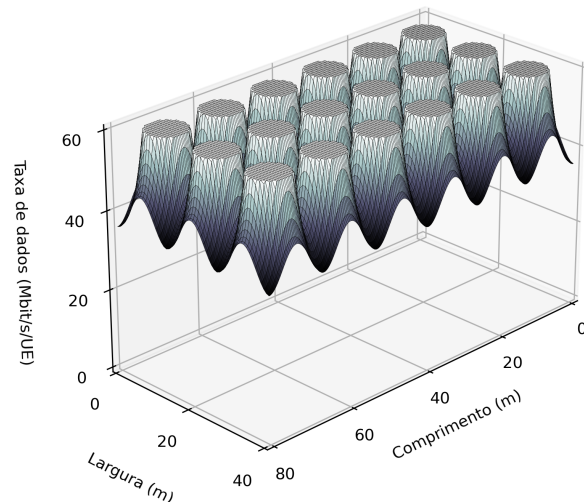
e, por fim, a taxa de dados pode ser calculada por:

$$R = B \log_2(1 + \text{SINR}) \quad (32)$$

sendo B a largura de banda.

Percebe-se que, no *small cell*, a SNR está em função apenas do ganho máximo de canal dentre todas as APs em relação ao UE, que corresponde à AP que maximiza a SNR e, portanto, foi a escolhida para a transmissão. Já no *cell-free*, como há cooperação entre as APs, a SNR não está em função apenas do ganho de canal máximo, mas sim em função de todos os ganhos somados, já que todas as APs estão transmitindo para o UE.

Diante do exposto, movendo-se o UE pelo cenário que será utilizado na seção de resultados e calculando a taxa em cada ponto, a taxa por usuário em um sistema celular convencional com 18 APs possui o padrão da Figura 18, e em um sistema *cell-free*, com as mesmas configurações, possui o padrão da Figura 19. As simulações foram inspiradas nas simulações disponíveis em [1].


 Figura 18: Padrão de taxa de dados em um sistema celular convencional com $L = 18$ num galpão 80x40m. Fonte: Próprio autor.

 Figura 19: Padrão de taxa de dados em um sistema *cell-free* com $L = 18$ num galpão 80x40m. Fonte: Próprio autor.

Surge então o primeiro problema fundamental do *cell-free*, que é a escalabilidade do sistema. A abordagem principal de solução é a utilização de *clusters*, como ilustrado na Figura 20, inspirada em [1], que são conjuntos compostos por múltiplas APs e, geralmente, por um UE, quando o *cluster* é centrado no usuário (para o restante do texto, assume-se que os *clusters* são centrados nos usuários). Assim, um algoritmo de otimização pode criar associações entre os usuários com as melhores APs disponíveis, como, por exemplo, as N APs com os melhores ganhos de canal. Tal processo é realizado para cada *cluster* (para cada UE, se os *clusters* são centrados no usuário).

Há no sistema *cell-free* uma central de processamento, a CPU, na qual todas as APs estão conectadas por enlaces físicos para que os cálculos de estimativa de canal, pré-codificação e de vetores *combining* sejam realizados. Para os processos de cálculo, algumas abordagens são possíveis, como a abordagem centralizada, em que todos os cálculos são concentrados na CPU, e a abordagem distribuída, em que alguns cálculos são feitos pelas próprias APs [3]. Neste trabalho, será adotado o sistema centralizado, como em [2], mas será incorporada a formação de *clusters* com APs e RISs.

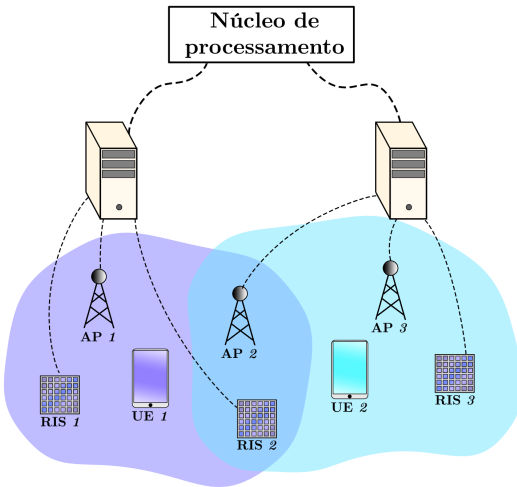


Figura 20: *Clusters* em um sistema cell-free com RIS. Fonte: Próprio autor.

E. Sistema two-timescale com clusters

O sistema *two-timescale* é um meio para atingir um equilíbrio entre *overhead* de sinalização e performance do sistema. A ideia é utilizar o primeiro *timescale* para realizar a estimação completa dos canais, formar os *clusters* com essas informações e, nos próximos *timescales*, realizar apenas a estimação parcial dos canais que estão ativos pelos *clusters*. Dessa forma, os cálculos de pré-codificação e de vetores *combining* podem ser reduzidos pela menor quantidade de canais presentes nos cálculos, além de que o *overhead* de sinalização também pode ser reduzido pelo mesmo motivo. Intuitivamente, o sistema *two-timescale* pode ser observado pela Figura 21, inspirada em [2].

Neste trabalho, será utilizada a variável T para representar a quantidade total de *timescales*, ou um *large timescale*, e X para representar a quantidade de *small timescales*, com CSI

parcial, ou seja, $T = X + 1$ uma vez que apenas o primeiro *timescale* é destinado à realização do CSI completo.

IV. MODELO DO SISTEMA

O canal completo entre o usuário k e o AP l inclui o canal direto e os canais indiretos, ou canais em cascata, que passam pela RIS, e pode ser descrito pelo seguinte (Figura 22):

$$\mathbf{h}_{l,k}^H = \mathbf{H}_{l,k}^H + \sum_{r=1}^R \mathbf{F}_{r,k}^H \Theta_r^H \mathbf{G}_{l,r} \quad (33)$$

onde $\mathbf{H}_{l,k} \in \mathbb{C}^{N_t \times N_r}$ é o canal direto entre o AP l e o UE k , $\mathbf{F}_{r,k} \in \mathbb{C}^{M \times N_r}$ é o canal entre a RIS e o UE, $\mathbf{G}_{l,r} \in \mathbb{C}^{N_t \times M}$ é o canal entre AP e RIS e $\Theta_r = \text{diag}(\theta_{r,1}, \dots, \theta_{r,M})$ representa a matriz de defasagens da RIS r , sendo $\theta_{r,m}$ a defasagem individual do elemento m da RIS r onde $|\theta_{r,m}| \leq 1$. Neste trabalho, assume-se que os canais são conhecidos perfeitamente.

Denota-se N_t , N_r e M como o número de antenas de cada AP, o número de antenas de cada UE e o número de elementos de cada RIS, respectivamente. Então, $\mathcal{K} = \{1, \dots, K\}$, $\mathcal{R} = \{1, \dots, R\}$, $\mathcal{L} = \{1, \dots, L\}$ e $\mathcal{M} = \{1, \dots, M\}$ denotam o índice de usuários, RISs, APs e elementos de RIS, respectivamente.

Uma maior simplificação pode ser feita definindo $\mathbf{F}_k = [\mathbf{F}_{1,k}^T, \dots, \mathbf{F}_{R,k}^T]^T \in \mathbb{C}^{MR \times N_r}$, $\Theta = \text{diag}(\Theta_1, \dots, \Theta_R) \in \mathbb{C}^{MR \times MR}$ e $\mathbf{G}_l = [\mathbf{G}_{l,1}^T, \dots, \mathbf{G}_{l,R}^T]^T \in \mathbb{C}^{MR \times N_t}$; então, o canal completo (33) pode ser reescrito como:

$$\mathbf{h}_{l,k}^H = \mathbf{H}_{l,k}^H + \mathbf{F}_k^H \Theta^H \mathbf{G}_l \quad (34)$$

Um símbolo transmitido para o usuário k denotado por $s_k \in \mathbb{C}$ é inicialmente pré-codificado com um vetor *precoding* $\mathbf{w}_{l,k} \in \mathbb{C}^{N_t \times 1}$ na transmissão *downlink*, então os símbolos pré-codificados $\mathbf{w}_{l,k} s_k$ são somados para todos os k no AP l para gerar o símbolo pré-codificado \mathbf{x}_l do AP. Então, o usuário k recebe todos os símbolos transmitidos \mathbf{x}_l que foram destinados a todos os usuários, mais um ruído branco aditivo gaussiano (AWGN) distribuído como $\mathcal{CN}(\mathbf{0}, \sigma_k^2 \mathbf{I}_{N_r})$ e denotado por $\mathbf{z}_k \in \mathbb{C}^{N_r \times 1}$. Portanto, o sinal recebido pelo usuário k pode ser escrito como segue:

$$\mathbf{y}_{l,k} = \mathbf{h}_k^H \mathbf{w}_k s_k + \sum_{\substack{i=1 \\ i \neq k}}^K \mathbf{h}_k^H \mathbf{w}_i s_i + \mathbf{z}_k \quad (35)$$

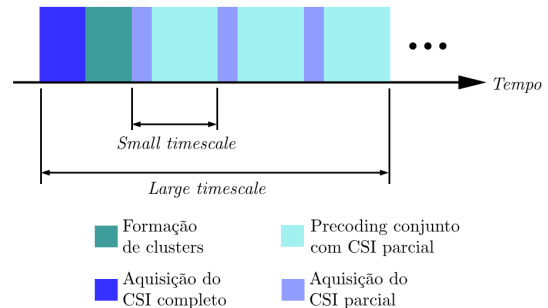


Figura 21: Visualização do esquema two-timescale com clusters. Fonte: Próprio autor.

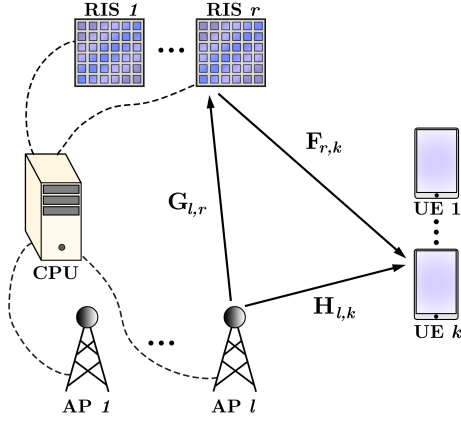


Figura 22: Diagrama de um sistema *cell-free* auxiliado por RIS. Fonte: Próprio autor.

em que a igualdade se mantém definindo $\mathbf{h}_k = [\mathbf{h}_{1,k}^T, \dots, \mathbf{h}_{L,k}^T]^T \in \mathbb{C}^{LN_t \times N_r}$ e $\mathbf{w}_k = [\mathbf{w}_{1,k}^T, \dots, \mathbf{w}_{L,k}^T]^T \in \mathbb{C}^{LN_t \times 1}$. O primeiro termo é o sinal desejado, correspondente ao símbolo destinado de fato ao usuário k junto ao seu vetor *precoding* compatível \mathbf{w}_k . O segundo termo corresponde à soma dos símbolos recebidos que foram destinados aos outros usuários que não o k , com seus respectivos vetores *precoding* e, portanto, o termo representa a interferência inter-usuário. Segue que a razão sinal-ruído mais interferência (SINR) pode ser calculada como:

$$\text{SINR}_k = \frac{|\mathbf{h}_k^H \mathbf{w}_k|^2}{\sum_{i=1, i \neq k}^K |\mathbf{h}_k^H \mathbf{w}_i|^2 + \sigma_k^2}, \forall k \in \mathcal{K} \quad (36)$$

de tal forma que $\text{SINR}_{l,k}$ são os valores de SINR com somente os canais do AP l , ou seja, calcula-se (36) com $\mathbf{h}_k = [0, \dots, 0, \mathbf{h}_{l,k}^T, 0, \dots, 0]^T$, e a ideia é retirar qualquer atuação de outros APs que não o l para que a contribuição na SINR por parte do mesmo seja calculada. De forma análoga, $\text{SINR}_{r,k}$ são os valores de SINR com somente os canais da RIS r , então calcula-se (36) com $\mathbf{F}_k = [0, \dots, 0, \mathbf{F}_{r,k}^T, 0, \dots, 0]^T$ e $\mathbf{G}_l = [0, \dots, 0, \mathbf{G}_{l,r}^T, 0, \dots, 0]^T$.

Ocorre que, na prática, as RIS não possuem capacidade de aquisição de sinais, a não ser que tenham sido fabricadas com sensores e, portanto, não é possível estimar \mathbf{G} e \mathbf{F} separadamente, mas é possível estimar $\mathbf{F}_k^H \Theta^H \mathbf{G}_l$ em conjunto. Por isso, a solução seria zerar todos os $\mathbf{F}_{i,k}^H \Theta_i^H \mathbf{G}_{l,i}$ para $i \neq l$ e, assim, $\mathbf{h}_{l,k}$ de (33) seria composto de todos os canais diretos \mathbf{H} e somente um canal cascata $\mathbf{F}_{r,k}^H \Theta_r^H \mathbf{G}_{l,r}$.

A. Algoritmo proposto

Matrizes de seleção $\mathcal{D}_{r,k}$ e $\mathcal{D}_{l,k}$ são geradas por matrizes $\text{SINR}_{r,k}$ e $\text{SINR}_{l,k}$ através de um algoritmo guloso (*greedy algorithm*), escolhendo-se os pares RIS-UE e AP-UE com os maiores valores de SINR. Define-se:

$$\mathcal{D}_{l,k} = \begin{cases} \mathbf{I}_{N_t}, & \text{se o AP } l \text{ é escolhido para o UE } k \\ \mathbf{0}_{N_t}, & \text{se o AP } l \text{ não é escolhido para o UE } k \end{cases} \quad (37)$$

e, da mesma forma:

$$\mathcal{D}_{r,k} = \begin{cases} \mathbf{I}_{N_t}, & \text{se a RIS } r \text{ é escolhida para o UE } k \\ \mathbf{0}_{N_t}, & \text{se a RIS } r \text{ não é escolhida para o UE } k \end{cases} \quad (38)$$

O canal completo $\mathbf{h}_{l,k}$ de (33) é, então, modificado pelas matrizes de seleção, resultando em:

$$\mathbf{h}_{l,k}^H = \mathcal{D}_{l,k} \left[\mathbf{H}_{l,k}^H + \sum_{r=1}^R \mathcal{D}_{r,k} (\mathbf{F}_{r,k}^H \Theta_r^H \mathbf{G}_{l,r}) \right] \quad (39)$$

B. O problema de otimização

Com (36), a WSR (*weighted sum-rate*, ou soma ponderada das taxas) R_{sum} pode ser calculada por:

$$R_{\text{sum}} = \sum_{k=1}^K \eta_k \log_2(1 + \text{SINR}_k) \quad (40)$$

onde η_k é o peso de cada usuário k , mas será assumido que $\eta_k = 1, \forall k \in \mathcal{K}$. Finalmente, o problema de otimização é o seguinte:

$$\begin{aligned} & \max_{\Theta, \mathbf{W}, \mathcal{D}_{l,k}, \mathcal{D}_{r,k}} R_{\text{sum}}(\Theta, \mathbf{W}, \mathcal{D}_{l,k}, \mathcal{D}_{r,k}) \\ & \text{s.t. } C_1 : \sum_{k=1}^K \|\mathbf{w}_{l,k}\|^2 \leq P_t, \forall l \in \mathcal{L} \\ & C_2 : |\theta_{r,m}| \leq 1, \forall r \in \mathcal{R}, \forall m \in \mathcal{M} \\ & C_3 : \sum_{l \in \mathcal{L}} \mathcal{D}_{l,k} = \mathbf{I}_{N_t} AP_{\text{cluster}}, \forall k \in \mathcal{K} \\ & C_4 : \sum_{r \in \mathcal{R}} \mathcal{D}_{r,k} = \mathbf{I}_{N_t} R_{\text{cluster}}, \forall k \in \mathcal{K} \\ & C_5 : \mathcal{D}_{l,k} \in \{\mathbf{0}_{N_t}, \mathbf{I}_{N_t}\}, \forall l \in \mathcal{L}, \forall k \in \mathcal{K} \\ & C_6 : \mathcal{D}_{r,k} \in \{\mathbf{0}_{N_t}, \mathbf{I}_{N_t}\}, \forall r \in \mathcal{R}, \forall k \in \mathcal{K} \end{aligned} \quad (\text{P1})$$

que é separado em subproblemas, onde Θ e \mathbf{W} são otimizados como em [2], empregando otimização alternada, e a maximização com respeito a $\mathcal{D}_{l,k}$ e $\mathcal{D}_{r,k}$ é feita de acordo com o algoritmo proposto, detalhado no Algoritmo 1. O Algoritmo 1 é executado uma vez, na primeira iteração da otimização durante o primeiro *timescale*, em que, anteriormente, é feito CSI completo e geração das matrizes completas de SINR. Nos demais *timescales*, onde é feita apenas estimativa parcial dos canais, os *clusters* estão formados e permanecem estáticos até que seja finalizado o *large timescale* e o processo de estimativa completa seja feito novamente.

Nos próximos *timescales* o *input* é composto somente pelo resultado da estimativa parcial dos canais e não ocorrem as etapas de cálculo das matrizes de SINR e de *cluster*. O que ocorre são as modificações com as mesmas matrizes de *cluster* anteriormente geradas e a otimização do *precoding* conjunto.

V. SIMULAÇÕES E RESULTADOS

As RISs são instaladas no teto, nas coordenadas xy (20, 0), (60, 0), (100, 0) e (140, 0), com altura (coordenada z) de

Algoritmo 1 Algoritmo proposto (primeiro *timescale*)

Input: Todos os canais $\mathbf{H}_{l,k}$, $\mathbf{G}_{l,r}$ and $\mathbf{F}_{r,k}$ onde $l \in \mathcal{L}$, $k \in \mathcal{K}$, $r \in \mathcal{R}$.

Output: \mathbf{W}_{opt} e Θ_{opt} .

- 1: Inicializar \mathbf{W} e Θ ;
- 2: Calcular $\text{SINR}_{l,k}$ e $\text{SINR}_{r,k}$
- 3: Calcular $\mathcal{D}_{l,k}$ com $\text{SINR}_{l,k}$ via algoritmo guloso
- 4: Calcular $\mathcal{D}_{r,k}$ com $\text{SINR}_{r,k}$ via algoritmo guloso
- 5: Atualizar os canais como em (39)
- 6: **while** R_{sum} não convergir **do**
- 7: Atualizar \mathbf{W} e Θ como em [2];
- 8: **end while**
- 9: **return** \mathbf{W}_{opt} e Θ_{opt} .

$h_R = 4$ m. As APs são instaladas na parede, uniformemente distribuídas de $(0, -20)$ até $(160, -20)$, com altura $h_A = 3$ m. São considerados $K = 4$ UEs ativos aleatoriamente distribuídos em uma área circular com raio de 1 metro, centralizado no meio do cenário $(80, 0)$, $L = 40$ APs, $R = 4$ RIS disponíveis, frequência de operação de $f = 15$ GHz, potência máxima de transmissão por AP de $P_t = 1$ mW, em que as APs e UEs possuem uma antena cada, $N_t = N_r = 1$. O cenário pode ser visualizado na Figura 24. Para todas as simulações, é assumido um nível de ruído de $\sigma_k^2 = -80$ dBm, $\forall k \in \mathcal{K}$.

A. Funcionamento e efeito dos clusters

Como demonstrado em [5], as RIS conseguem promover melhorias significativas na qualidade da transmissão quando os canais diretos AP para UE estão prejudicados. Aqui, serão explorados diferentes valores de R_{cluster} e AP_{cluster} com um canal direto exclusivamente NLoS (sem visada), com fator de Rice $\kappa = 0$. Canais AP para RIS são configurados com $\kappa = 10$, e canais das RIS para os UEs são configurados com $\kappa = 8$, considerando-se RISs estrategicamente posicionadas.

A Figura 23 mostra que a seleção de mais RISs e APs em um *cluster* resulta em uma melhor performance geral do sistema. Também, a diferença de taxa de crescimento entre as curvas de R_{cluster} diferentes é notável, e a explicação está no algoritmo de formação dos *clusters*: um AP só estará transmitindo para um certo usuário por uma RIS específica se tanto a RIS quanto o AP forem selecionados para serem parte do *cluster*. Devido a isso, com um canal direto fraco, aumentar o número de APs no *cluster* com uma única RIS ajudando a transmissão fará o sistema mais suscetível a oscilações dos canais do que se duas ou mais RIS forem utilizadas, e assim por diante.

Além disso, a simulação da Figura 13, que demonstra o resultado de crescimento quadrático do ganho de canal em função da quantidade de elementos da RIS, está presente aqui: aumentar a quantidade de RISs também é aumentar a quantidade de elementos disponíveis. A diferença é que antes havia 1 RIS estática com uma quantidade variável de elementos, e agora são múltiplas RIS distribuídas atuando em ângulos diferentes para auxiliar a transmissão. Por isso, os dois efeitos descritos estão presentes, resultando no aumento citado das taxas de crescimento das curvas.

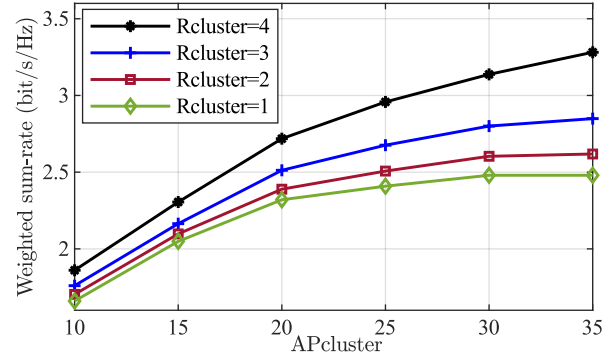


Figura 23: Valores de WSR com variação dos tamanhos de *cluster* e canal direto em NLoS. Fonte: Próprio autor.

B. Análise de sinalização

Considerando sobreposição de *clusters*, onde uma RIS ou AP pode fazer parte de mais de um *cluster* ao mesmo tempo, algumas APs ou RISs podem não ser utilizadas caso o tamanho dos *clusters* seja suficientemente pequeno e elas estejam distantes de todos os usuários. Por isso, seja L_a o número de APs ativas (APs sendo utilizadas por pelo menos um UE) e R_a o número de RISs ativas. O número total de *timescales* é T e X é o número de *timescales* com estimação parcial de canal. Para o primeiro *timescale*, o sistema fará a estimação completa, então o algoritmo proposto possui *overhead* de sinalização depois de I_o iterações de $N_r N_t L K + I_o (K N_r + R M + 2 L K N_t)$ [3], que é a mesma quantidade do algoritmo centralizado convencional [3], e $N_r N_t L K + I_o (K N_r + R M + (2L - 1) K N_t)$ é o *overhead* de sinalização do algoritmo distribuído proposto em [3]. Entretanto, para os próximos X *timescales*, apenas a estimação parcial é realizada devido aos *clusters*, então o algoritmo proposto para C *clusters* possui um *overhead* de $C N_r N_t L_c K + I_o (K N_r + R_a M + 2 L_a K N_t)$, enquanto os algoritmos centralizado convencional e distribuído continuam apresentando o mesmo *overhead*.

Para comparação entre os *overheads* de sinalização dos diferentes *frameworks*, define-se para uma análise com K usuários:

$$\alpha = \frac{\text{Overhead do framework proposto}}{\text{Overhead do framework distribuído}} \quad (41)$$

$$\beta = \frac{\text{Overhead do framework proposto}}{\text{Overhead do framework centralizado}} \quad (42)$$

Os parâmetros utilizados foram $C = K$, $L_c = 0.25 \cdot L$, $I_o = 30$, $R_a = 3$, $R = 4$, $L_a = 0.5 \cdot L$, $T = 11 = X + 1$, $M = 100$ e $N_t = N_r = 1$, com resultado na Figura 25.

Com $K = 4$, para visualizar o *overhead* da situação utilizada na simulação da Figura 23, a Figura 25 mostra que $\alpha = 0.6829$ e $\beta = 0.6792$, que representa um *overhead* do algoritmo proposto 31.71% menor em comparação com o *overhead* do algoritmo distribuído, e 32.08% menor em comparação com o *overhead* do algoritmo centralizado convencional. O resultado apresentado é flexível e depende dos parâmetros de configuração, onde, por exemplo, o número de *timescales* com estimação parcial de canal pode ser ajustado para atender a requisitos de performance, uma vez que não

Figura sem escala

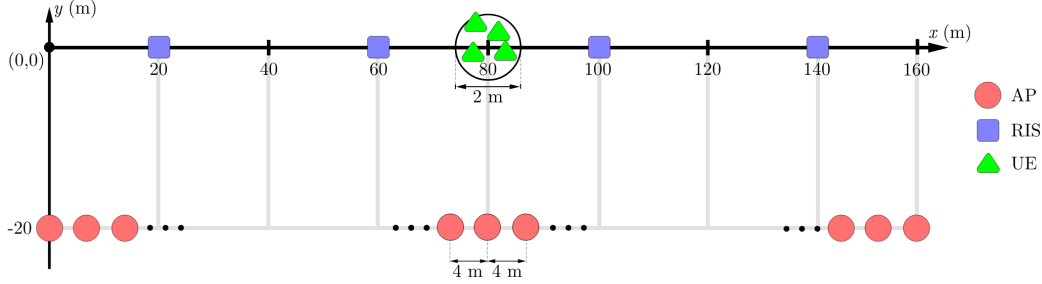
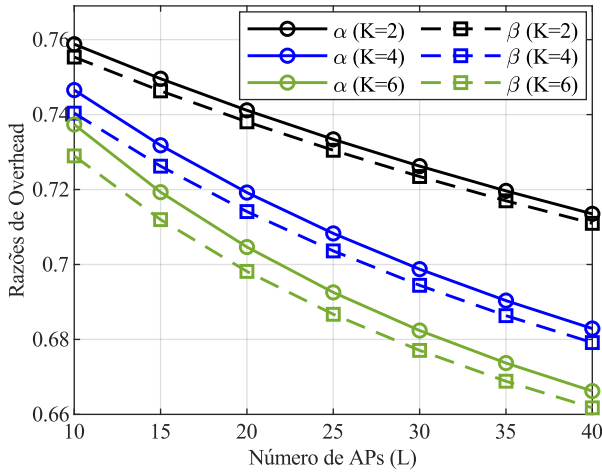
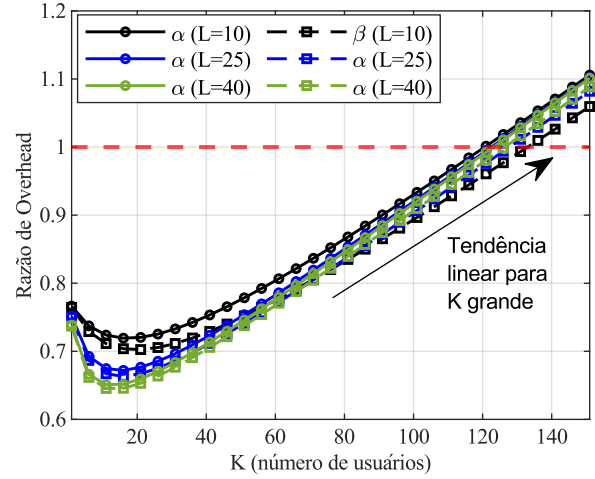


Figura 24: Cenário de simulação. Fonte: Próprio autor.


 Figura 25: Razão de *overhead* de sinalização $\alpha = \text{Proposto/Distribuído}$ e $\beta = \text{Proposto/Centralizado}$ para $K = 2, 4$ ou 6 (número de UEs). Fonte: Próprio autor.

 Figura 26: Razão de *overhead* de sinalização $\alpha = \text{Proposto/Distribuído}$ e $\beta = \text{Proposto/Centralizado}$ em função de K (número de UEs), com $L = 10, 25$ ou 40 (número de APs). Fonte: Próprio autor.

atualizar os *clusters* por um período longo o suficiente pode degradar a performance do sistema além de níveis toleráveis.

É interessante notar a diferença entre a melhora em função do número de usuários, de $K = 2$ para $K = 4$ e de $K = 4$ para $K = 6$, percebe-se uma tendência de estabilização, o que demonstra uma potencial limitação do algoritmo proposto, uma vez que, para quantidades muito elevadas de usuários, o *overhead* de sinalização começa a apresentar tendência de crescimento semelhante à dos outros algoritmos.

Identificada a possível limitação, uma rápida análise das equações evidencia o problema. Inicialmente, desconsiderando o primeiro *timescale*, tem-se, para valores muito grandes de K e, com $C = K$, $L_c = 0.25L$, $I_o = 30$, $R_a = 3$, $L_a = 0.5L$ e $N_r = N_t = 1$, tem-se:

$$\begin{aligned} \alpha &= \frac{CN_r N_t L_c K + I_o(KN_r + R_a M + 2L_a K N_t)}{N_r N_t L K + I_o(KN_r + R M + (2L - 1)K N_t)} \\ &= \frac{0.25LK^2 + 30(K + 300 + LK)}{LK + 30(400 + 2LK)} \\ &\stackrel{(a)}{=} \frac{0.25LK + 30(1 + L)}{61L} \stackrel{(b)}{=} \frac{0.25K + 30}{61} \end{aligned} \quad (43)$$

em que (a) sustenta-se assumindo $K(1 + L) \gg 300$, $2LK \gg$

400 e (b) assumindo $L \gg 1$. Nota-se que o resultado para K grande apresenta tendência linear em K , independente do número de APs L , que é possível observar na Figura 26, para valores maiores de K , as curvas começam a apresentar tendências lineares com coeficientes angulares iguais, independentes do valor de L . Para encontrar o valor de K em que o algoritmo passa a não possuir vantagem, substitui-se $\alpha = 1$ para encontrar $K = 124$. A Figura 26 confirma tal análise, em que, para valores de K próximos de 120, as razões de *overhead* ultrapassam o valor unitário.

É importante perceber que a análise conduzida acima é específica do sistema com os parâmetros listados. Em uma implementação real, com parâmetros diferentes, seria necessário realizar uma nova análise com os novos parâmetros, onde se identificariam as limitações e ajustes a serem feitos. Imaginando uma quantidade muito maior de RIS, ou RISs com quantidades muito maiores de elementos, de tal forma que $RM \rightarrow \infty$, teria-se que $\alpha \rightarrow 1$, situação que faria a análise anterior ser pouco importante. Também, foi assumido que as APs e os UEs possuem quantidades de antenas iguais $N_t = N_r = 1$, mas são dois parâmetros que podem alte-

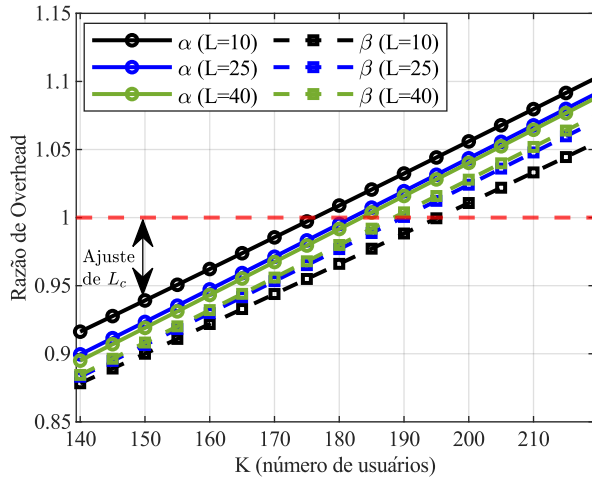


Figura 27: Razão de *overhead* de sinalização $\alpha = \text{Proposto/Distribuído}$ e $\beta = \text{Proposto/Centralizado}$ em função de K (número de UEs), com $L = 10, 25$ ou 40 (número de APs) e parâmetros de *cluster* ajustados para $K = 150$. Fonte: Próprio autor.

rar radicalmente o resultado da Figura 26. Como exemplo, considerando $N_t = N_r = 10$, o resultado de (43) seria de apenas $K = 16$ para $\alpha = 1$, ou seja, com apenas 16 usuários, nessa configuração, o algoritmo proposto já não apresentaria vantagem em comparação com os outros.

Por fim, como exemplo de ajuste de parâmetros de *cluster* para adequação de performance do algoritmo, com os mesmos parâmetros utilizados na simulação da Figura 26, mas diminuindo apenas a quantidade de APs por *cluster* para $L_c = L/6$, tem-se o resultado da Figura 27, em que os valores de razão de *overhead* já estão abaixo de 1 com 150 UEs, indicando uma vantagem na utilização do algoritmo, o que não ocorria com os parâmetros anteriores.

VI. CONCLUSÃO

Nos fundamentos teóricos, foi demonstrada, via diversas simulações, a eficiência das RIS devidamente configuradas em um cenário *indoor* SISO. Foi possível observar efeitos clássicos das RIS, descritos na literatura, como a tendência de ganho com M^2 , sendo M a quantidade de elementos da RIS, e o posicionamento médio ótimo, que é perto do UE ou perto do AP. Também foi demonstrada a eficiência dos sistemas *cell-free* a partir do equacionamento da SNR, onde foi demonstrado que o *cell-free* habilita a cooperação entre as APs e controla as interferências características de sistemas celulares convencionais, efeito perceptível comparando-se a Figura 18, do padrão de taxa de dados do sistema celular, com a Figura 19, do padrão de taxa de dados do *cell-free*.

A abordagem de formação de *clusters* via valores de SINR com algoritmo guloso demonstra-se eficiente em reduzir o *overhead* de sinalização do sistema, em que a redução registrada chegou a 32.08% na comparação do algoritmo proposto com o centralizado convencional. Para sistemas grandes, com muitos UEs e muitas APs e RISs, torna-se fundamental a aplicação de *clusters* para limitar a sinalização, que foi investigada. Também, foi identificada uma limitação do algoritmo

para os parâmetros escolhidos, em que uma quantidade de usuários acima de 120 torna o algoritmo ineficiente na análise de sinalização. A abordagem ideal para solucionar o problema encontrado seria ajustar os parâmetros de *cluster* de forma a equilibrar a performance e o *overhead* em uma situação com tantos usuários, ou, em casos mais extremos, utilizar algoritmos mais robustos e com métricas mais elaboradas.

A partir da diminuição do *overhead* de sinalização, são obtidas menores latências de transmissão e, por consequência, sistemas mais eficientes, abrindo espaço para o estabelecimento de tecnologias novas, como os *wearables*, que necessitam de taxas de dados bastante elevadas [9].

Trabalhos futuros serão focados em investigar os impactos na complexidade computacional e em desenvolver novos algoritmos de seleção capazes de capturar padrões mais profundos.

REFERÊNCIAS

- [1] Özlem Tuğfe Demir, Emil Björnson and Luca Sanguinetti (2021), "Foundations of User-Centric Cell-Free Massive MIMO", Foundations and Trends® in Signal Processing: Vol. 14, No. 3-4, pp 162–472. DOI: 10.1561/2000000109.
- [2] Z. Zhang and L. Dai, "A Joint Precoding Framework for Wideband Reconfigurable Intelligent Surface-Aided Cell-Free Network," in IEEE Transactions on Signal Processing, vol. 69, pp. 4085–4101, 2021, doi: 10.1109/TSP.2021.3088755.
- [3] Zhang, Peng, et al. "Joint distributed precoding and beamforming for RIS-aided cell-free massive MIMO systems." IEEE Transactions on Vehicular Technology (2023).
- [4] K. Haneda et al., "Indoor 5G 3GPP-like channel models for office and shopping mall environments," 2016 IEEE International Conference on Communications Workshops (ICC), Kuala Lumpur, Malaysia, 2016, pp. 694–699, doi: 10.1109/ICCW.2016.7503868.
- [5] Emil Björnson and Özlem Tuğfe Demir. (2024). Introduction to Multiple Antenna Communications and Reconfigurable Surfaces. Boston–Delft: Now Publishers
- [6] E. Björnson and L. Sanguinetti, "Scalable Cell-Free Massive MIMO Systems," in IEEE Transactions on Communications, vol. 68, no. 7, pp. 4247–4261, July 2020, doi: 10.1109/TCOMM.2020.2987311.
- [7] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: A tutorial," IEEE Trans. Commun., vol. 69, no. 5, pp. 3313–3351, May 2021.
- [8] J. G. Andrews, X. Zhang, G. D. Durgin, and A. K. Gupta, "Are we approaching the fundamental limits of wireless network densification?" IEEE Commun. Mag., vol. 54, no. 10, pp. 1558–1896, Oct. 2016.
- [9] H. Pennanen, T. Hänninen, O. Tervo, A. Tölli and M. Latva-Aho, "6G: The Intelligent Network of Everything," in IEEE Access, vol. 13, pp. 1319–1421, 2025, doi: 10.1109/ACCESS.2024.3521579.
- [10] E. Shi, J. Zhang, Z. Liu, Y. Zhu, C. Yuen, D. W. K. Ng, M. Di Renzo, and B. Ai, "Joint precoding and AP selection for energy efficient RIS aided cell-free massive MIMO using multi-agent reinforcement learning," arXiv preprint arXiv:2411.11070, 2024.
- [11] D. Tse and P. Viswanath, Fundamentals of Wireless Communication. Cambridge, United Kingdom: Cambridge University Press, 07 2005.
- [12] ETSI, TR 138 901 V15.0.0 (2018-07)-5G; Study on channel model for frequencies from 0.5 to 100 GHz (3GPP TR 38.901 version 15.0.0 Release 15)," Technical Report, ETSI, Tech. Rep., 2018.
- [13] E. Basar and I. Yildirim, "SimRIS Channel Simulator for Reconfigurable Intelligent Surface-Empowered Communication Systems," 2020 IEEE Latin-American Conference on Communications (LATINCOM), Santo Domingo, Dominican Republic, 2020, pp. 1-6, doi: 10.1109/LATINCOM50620.2020.9282349.
- [14] Wu, Tuo, et al. "Fluid Antenna Systems Enabling 6G: Principles, Applications, and Research Directions." arXiv preprint arXiv:2412.03839 (2024).
- [15] M. Nerini and B. Clerckx, "Dual-Polarized Beyond Diagonal RIS," in IEEE Communications Letters, vol. 29, no. 5, pp. 928–932, May 2025, doi: 10.1109/LCOMM.2025.3547175.
- [16] Liu, Hao, et al. "Stacked intelligent metasurfaces for wireless sensing and communication: Applications and challenges." arXiv preprint arXiv:2407.03566 (2024).
- [17] Orfanidis, Sophocles J. "Electromagnetic waves and antennas." (2002).