

Brasil 6G

Projeto Brasil 6G

Atividade 3.4 - Técnicas MIMO



UNIVERSIDADE FEDERAL
DO RIO DE JANEIRO



Histórico de Atualizações:

Versão	Data	Autor(es)	Notas
1	04/08/2023	Davi da Silva Brilhante Jonathan Gois Juliano Silveira Ferreira Pedro Henrique Carneiro de Souza Luciano Leonel Mendes	Elaboração de conteúdo
2	17/08/2023	Luciano Leonel Mendes	Revisão de texto

Lista de Figuras

1	Diagrama em blocos de um sistema empregando MIMO.	2
2	Diagrama em blocos simplificado do modelo para MIMO.	3
3	Comparação da BER dos detectores ZF, LMMSE e ZF-SIC(V-BLAST) para as modulações BPSK (azul), 4QAM (verde) e 16QAM (vermelho), com diferentes configurações de antenas no receptor e transmissor.	10
4	Ideia do algoritmo SD.	11
5	Fluxograma de detecção utilizando <i>Lattice Reduction Aided</i> (LRA).	14
6	Desempenho dos detectores por máxima verossimilhança (5), <i>Sphere Detector</i> (SD), <i>Zero Forcing</i> (ZF) (11) e <i>Approximate Message-Passing</i> (AMP) (47), considerando a probabilidade de erro do vetor transmitido para uma faixa de valores de <i>Signal-to-Noise Ratio</i> (SNR).	17
7	Arquitetura de redes neurais baseada nos princípios do <i>deep unfolding</i>	19
8	Desempenho dos detectores por máxima verossimilhança (5), <i>Deep Unfolded Probability Data Association</i> (DU-PDA), <i>Probability Data Association</i> (PDA), ZF (11) e AMP (47), considerando a probabilidade de erro de bit para uma faixa de valores de SNR.	22
9	Diagrama simplificado da rede de acesso Brasil 6G.	23
10	Ilustração da troca de informações do canal entre BS e UE.	25
11	<i>Set-ups</i> de testes.	26
12	MIMO com diversidade espacial: taxa de dados alcançada pela UE.	28
13	MIMO com multiplexação espacial: taxa de dados alcançada pela UE.	28
14	<i>Set-up</i> empregado para a captura da resposta do canal com enlace via antena.	30
15	Comparação da BER dos detectores MF, ML, LMMSE-SC e ZF-SIC(V-BLAST) para as modulações 4QAM (azul), 16QAM (verde) e 64QAM (vermelho), com resposta do canal real capturada a partir do transceptor com enlace via cabo de RF.	31
16	Comparação da BER dos detectores MF, ML, LMMSE-SC e ZF-SIC(V-BLAST) para as modulações 4QAM (azul), 16QAM (verde) e 64QAM (vermelho), com resposta do canal real capturada a partir do transceptor com enlace via antenas.	32

Lista de Tabelas

1	Relação de multiplicação entre as unidades <i>quaterniônicas</i>	16
2	Hiperparâmetros da arquitetura DU-PDA.	22
3	Tempo de processamento dos algoritmos.	27
4	Detalhes do <i>set-up</i> de captura considerando comunicação via cabo de RF.	30
5	Detalhes do set-up de captura considerando comunicação via antenas.	30

Acrônimos

6G Sexta Geração de Rede Móvel Celular

AMP *Approximate Message-Passing*

AWGN *Additive White Gaussian Noise*

BER *Bit Error Rate*

BPSK *Binary Phase-Shift Keying*

BS *Base Station*

CLLL *Complex Lenstra-Lenstra-Lovász*

DU-PDA *Deep Unfolded Probability Data Association*

FER *Frame Error Rate*

GFDM *Generalized Frequency Division Multiplexing*

IAI *Inter Antenna Interference*

LC-RBD-LR *Low Complexity Regularized Block Diagonalization Lattice Reduction*

LLL *Lenstra-Lenstra-Lovász*

LMMSE *Linear Minimum Mean Square Error*

LRA *Lattice Reduction Aided*

MCS *Modulation Coding Scheme*

MF *Matched Filter*

MIMO *Multiple-Input Multiple-Output*

ML *Machine Learning*

MMSE *Minimum Mean Square Error*

MRC *Maximum Ratio Combining*

OSIC *Ordered Successive Interference Cancellation*

PDA *Probability Data Association*

QAM *Quadrature Amplitude Modulation*

QPSK *Quadrature Phase Shift Keying*

RAN *Radio Access Network*

RBD *Regularized Block Diagonalization*

SD *Sphere Detector*

SDR *Software Defined Radio*

SIC *Successive Interference Cancellation*

SIMD *Single Instruction Multiple Data*

SNR *Signal-to-Noise Ratio*

UE *User Equipment*

ZF *Zero Forcing*

BER *Bit Error Rate*

ML *Maximum Likelihood*

MMSE *Minimum Mean Square Error*

MIMO *Multiple-Input Multiple-Output*

ZF *Zero Forcing*

MRC *Maximal Ratio Combining*

SIC *Successive Interference Cancellation*

Sumário

1	Introdução	1
2	Detectores Clássicos para MIMO	2
2.1	Princípios das Técnicas Clássicas para Detecção MIMO	3
2.1.1	Modelo do Sistema de Comunicação MIMO	3
2.1.2	Detector por Máxima Verossimilhança	3
2.1.3	<i>Matching Filter</i>	4
2.1.4	<i>Zero Forcing</i>	4
2.1.5	<i>Linear Minimal Mean Square Error</i>	5
3	Técnicas avançadas e de alto desempenho	6
3.1	<i>Successive Interference Cancellation</i>	6
3.2	<i>Sphere Detection</i>	11
3.3	<i>Lattice Reduction Aided</i>	13
3.4	<i>Approximate Message-passing</i>	16
3.5	<i>Deep Unfolded MIMO</i>	18
4	Implementação de MIMO no transceptor Brasil 6G	23
4.1	Desafios da implementação	23
4.2	Implementação de detectores MIMO com multiplexação espacial	24
4.3	Análise do condicionamento da matriz de canal	24
4.4	Ajustes na definição dos parâmetros do enlace de comunicação	25
4.5	Resultados de testes práticos	26
5	Avaliações complementares envolvendo simulações computacionais com condições reais do canal de comunicação	29
5.1	<i>Set-up</i> de captura das condições reais do canal	29
5.2	Resultado inicial das simulações considerando os canais reais	30
6	Conclusão	33

1 Introdução

O projeto Brasil 6G tem como um de seus objetivos principais conceber e implementar uma plataforma de testes e avaliação de diferentes tecnologias que possam vir a compôr a futura rede de comunicação móvel, a Sexta Geração de Rede Móvel Celular (6G). Dentre as tecnologias consideradas como habilitadoras, o *Multiple-Input Multiple-Output* (MIMO) merece especial destaque devido ao aumento da eficiência espectral e aumento de robustez.

A rede 6G deve prover suporte à diferentes aplicações que requerem enlaces de alta vazão e confiabilidade e as técnicas MIMO podem ser empregadas para atender a estes requisitos. As tecnologias MIMO envolvem um vasto campo de pesquisas com diferentes objetivos, que levam a algoritmos com variada complexidade de implementação e que proporcionam distintos benefícios. Dentre as pesquisas recentes pode-se destacar o emprego de aprendizado de máquina e de redes neurais em algoritmos de MIMO visando a redução de complexidade e maior robustez às variações das condições de operação do enlace.

O presente relatório, que tem por objetivo abordar tecnologias MIMO com multiplexação espacial, apresenta uma breve introdução aos conceitos e técnicas MIMO de reduzida complexidade, assim como técnicas mais complexas e de alto desempenho que tem sido alvo de pesquisas recentes. Além disso, o relatório apresenta os principais avanços relacionados à implementação de MIMO na plataforma Brasil 6G, mais especificamente no transceptor Brasil 6G. Estão apresentadas também o resultado de avaliação de alguns detectores MIMO a partir de uma metodologia que associa simulação computacional com informações capturadas a partir de condições reais do enlace de comunicação.

Sendo assim, visando alcançar seus objetivos, o presente relatório foi organizado da seguinte forma: a Seção 2, apresenta uma introdução aos conceitos de MIMO com foco em técnicas de detecção lineares. A Seção 3 apresenta os conceitos de técnicas mais avançadas e de alto desempenho, baseadas em técnicas não lineares destacando-se, dentre elas, *Sphere Detection*, *LRA*, *Approximate Message-Passing* e *Deep Unfolded MIMO*. A Seção 4 apresenta avanços com relação a implementação de MIMO no transceptor Brasil 6G. A Seção 5 mostra o resultado de simulações computacionais de alguns detectores MIMO considerando a resposta real do canal de comunicação. Finalmente, a Seção 6 apresenta a conclusão do relatório.

2 Detectores Clássicos para MIMO

Jonathan Gois, Juliano Silveira Ferreira, Luciano Mendes
jonathan.gois@cefet-rj.br, silveira@inatel.br, luciano@inatel.br

As técnicas MIMO se referem a um conjunto de tecnologias de processamento de sinais que proporcionam melhoria do desempenho de sistemas de comunicação sem fio e que empregam como base múltiplas antenas de transmissão e recepção [1, 2]. A melhoria de desempenho mencionada é alcançada devido à exploração do domínio espacial como complemento aos domínios do tempo e da frequência. Duas categorias importantes de técnicas MIMO que pode-se destacar são aquelas denominadas de diversidade espacial e de multiplexação espacial.

As técnicas MIMO associadas à diversidade espacial permitem melhorar o desempenho de comunicação reduzindo a taxa de *Bit Error Rate* (BER) do enlace de comunicação, o que impacta no aumento da área do cobertura graças a redução dos efeitos de desvanecimento causados pelo canal de comunicação [1]. Um sistema de comunicação que emprega M antenas de transmissão e N antenas de recepção, por exemplo, possui ordem de diversidade máxima MN [2]. Quanto maior a ordem de diversidade, menores são os efeitos do desvanecimento no sistema de comunicação e, conseqüentemente, menor é a probabilidade de erros no enlace de comunicação em uma determinada localidade.

As técnicas MIMO associadas à multiplexação espacial, por sua vez, permitem aumentar a vazão de dados de um sistema de comunicação, pois, empregam as múltiplas antenas para a transmissão de múltiplos fluxos de dados de maneira simultânea. Um sistema que utiliza M antenas de transmissão, nesse caso, pode teoricamente ampliar a vazão de comunicação do sistema em que é empregado em até M vezes quando comparado com o sistema com antena única [2], desde que o número de antenas receptoras seja $N \geq M$. O ganho de multiplexação espacial é alcançado sem demandar o aumento da banda de transmissão [1]. Os dados transmitidos pelas múltiplas antenas transmissoras se sobrepõem nas antenas receptoras resultando na ocorrência de *Inter Antenna Interference* (IAI). Para separar os diferentes fluxos de dados recebidos e mitigar a IAI, emprega-se técnicas baseadas em algoritmos lineares ou não-lineares. A Figura 1 ilustra a arquitetura básica de um sistema MIMO com multiplexação espacial.

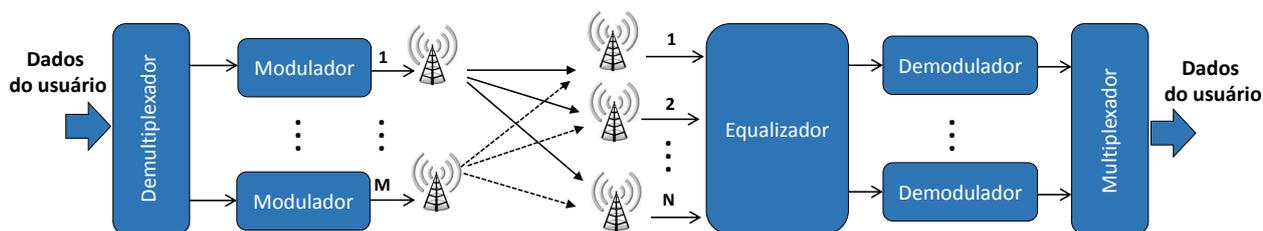


Figura 1: Diagrama em blocos de um sistema empregando MIMO.

Os diferentes detectores ou equalizadores existentes, que são os algoritmos de cancelamento de IAI, apresentam diferentes desempenhos e diferentes níveis de complexidade de implementação. A classe de algoritmos lineares, como ZF e *Minimum Mean Square Error* (MMSE), empregam alguma transformação linear do sinal recebido para estimar os símbolos transmitidos. Embora estes esquemas apresentem baixa complexidade de implementação, esta classe de algoritmos tipicamente apresenta um desempenho limitado em termos de BER por não serem capazes de colher a diversidade presente nas antenas de recepção. Por outro lado, técnicas não-lineares de estimação, como o *Maximum Likelihood* (ML), podem proporcionar desempenho

ótimo, atingindo o máximo ganho de diversidade, porém à um custo de complexidade de implementação que muitas vezes impede a construção de um detector para uso prático. A seguir estão apresentados maiores detalhes sobre as principais técnicas lineares de detectores MIMO. Algumas das técnicas não lineares, assim como técnicas avançadas e de alto desempenho, por sua vez, estão apresentadas na Seção 3.

2.1 Princípios das Técnicas Clássicas para Detecção MIMO

A seguir, a modelagem matemática de um sistema de comunicação que emprega MIMO é apresentada, assim como uma introdução às principais técnicas de detecção.

2.1.1 Modelo do Sistema de Comunicação MIMO

Dado um sinal de mensagem $\mathbf{q} \in \mathbb{C}^{N \times 1}$ a ser transmitido por um canal linear MIMO $\mathbf{G} \in \mathbb{C}^{M \times N}$, então o sinal presente nas antenas receptoras é dado por

$$\mathbf{y} = \mathbf{G}\mathbf{q} + \boldsymbol{\omega}, \quad (1)$$

onde $\boldsymbol{\omega}$ é um ruído complexo aditivo gaussiano $\mathcal{CN}(0, \sigma^2)$. A Fig. 2 mostra o diagrama em blocos simplificado deste modelo.

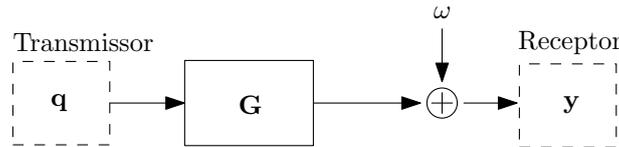


Figura 2: Diagrama em blocos simplificado do modelo para MIMO.

Os detectores lineares tem por objetivo inverter a influência do canal no sinal recebido, visando mitigar a IAI. O detector linear pode ser implementado através de uma matriz de equalização $\mathbf{D} \in \mathbb{C}^{M \times N}$ para a separação dos dados transmitidos. Logo, a estimativa do vetor de dados obtida no receptor é dada por

$$\hat{\mathbf{q}} = \mathbf{D}^T \mathbf{y}, \quad (2)$$

onde $(\hat{\cdot})$ representa uma grandeza estimada.

2.1.2 Detector por Máxima Verossimilhança

O objetivo de todo detector é estimar qual símbolo foi enviado com maior probabilidade, dada a observação do sinal recebido \mathbf{y} , tendo-se conhecimento sobre a matriz de canal \mathbf{G} . Existem diversas maneiras de realizar tal estimação, sendo a minimização da probabilidade de erro [3] a primeira a ser analisada neste relatório. Neste caso, o objetivo é encontrar $\hat{\mathbf{q}}$ que maximize $P(\mathbf{q} = \hat{\mathbf{q}} | \mathbf{y}, \mathbf{G})$, que pode ser expressa como

$$P(\mathbf{q} = \hat{\mathbf{q}} | \mathbf{y}, \mathbf{G}) = \frac{P(\mathbf{q} = \hat{\mathbf{q}}) f_{\mathbf{y} | \mathbf{q}, \mathbf{G}}(\mathbf{y} | \mathbf{q} = \hat{\mathbf{q}}, \mathbf{G})}{f_{\mathbf{y} | \mathbf{G}}(\mathbf{y} | \mathbf{G})}, \quad (3)$$

onde $f_{\mathbf{y} | \mathbf{q}, \mathbf{G}}$ e $f_{\mathbf{y} | \mathbf{G}}$ são as funções de densidade de probabilidade condicional de \mathbf{y} dado (\mathbf{q}, \mathbf{G}) e \mathbf{y} dado \mathbf{G} respectivamente. Como nem $P(\mathbf{q} = \hat{\mathbf{q}})$ nem $f_{\mathbf{y} | \mathbf{G}}(\mathbf{y} | \mathbf{G})$ dependem de $\hat{\mathbf{q}}$ em (3), o

máximo de $P(\mathbf{q} = \hat{\mathbf{q}}|y, \mathbf{G})$ é dado por $\hat{\mathbf{q}}$ tal que $f_{y|\mathbf{q}, \mathbf{G}}(y|\mathbf{q} = \hat{\mathbf{q}}, \mathbf{G})$. Esse é o dado critério de máxima verossimilhança e o detector é dado por

$$\hat{\mathbf{q}} = \operatorname{argmax} f_{y|\mathbf{q}, \mathbf{G}}(y|\mathbf{q} = \hat{\mathbf{q}}, \mathbf{G}). \quad (4)$$

Dado que o ruído seja do tipo *Additive White Gaussian Noise* (AWGN), a estimativa que maximiza a verossimilhança é aquela que minimiza a distância entre o vetor de símbolos recebido y e a estimativa na saída do canal \mathbf{G} , ou seja

$$\hat{\mathbf{q}}_{\text{ML}} = \operatorname{argmin} \|y - \mathbf{G}\hat{\mathbf{q}}\|^2. \quad (5)$$

O detector de máxima verossimilhança atinge o desempenho ótimo, pois obtém o ganho máximo de diversidade presente no canal de comunicações. Porém, a complexidade computacional da detecção é proporcional ao número de símbolos da constelação empregada elevado ao número de antenas transmissoras, tornando impraticável a sua implementação. Os próximos detectores lineares a serem retratados nesta seção visam reduzir a complexidade de implementação em detrimento do desempenho em termos de BER.

2.1.3 Matching Filter

O detector *Matched Filter* (MF) é obtido fazendo-se $\mathbf{D} = \mathbf{G}^H$, de modo que o sinal estimado é dado por

$$\hat{\mathbf{q}}_{\text{MF}} = \mathbf{G}^H y \quad (6)$$

Esta detecção funciona de forma apropriada quando $M \ll N$ [4] e possui desempenho inferior aos demais detectores mais complexos. Também conhecido como *Maximal Ratio Combining* (MRC), o sinal recebido pode ser decomposto como [5]:

$$\hat{\mathbf{q}}_{\text{MF}} = \mathbf{G}^H (\mathbf{G}\mathbf{q} + \boldsymbol{\omega}) \quad (7)$$

$$= \mathbf{G}^H \mathbf{G}\mathbf{q} + \mathbf{G}^H \boldsymbol{\omega} \quad (8)$$

$$(9)$$

Para um dado terminal d , tem-se que

$$\hat{\mathbf{q}}_{\text{MF}, d} = \underbrace{\mathbf{g}_d^H \mathbf{g}_d y_d}_{\text{Sinal}} + \underbrace{\sum_{j=1, j \neq d}^K \mathbf{g}_d^H \mathbf{g}_j y_j}_{\text{IAI}} + \underbrace{\mathbf{g}_d^H \boldsymbol{\omega}}_{\text{Ruído}}, \quad (10)$$

onde considerou que o receptor tem pleno conhecimento do canal entre o transmissor e o terminal d .

2.1.4 Zero Forcing

O detector por *Zero Forcing* [5] é obtido fazendo-se $\mathbf{D} = \mathbf{G}^{-1}$, levando a

$$\hat{\mathbf{q}} = \mathbf{G}^{-1} y. \quad (11)$$

Caso o sistema MIMO seja não quadrado, ou seja, caso $M < N$, então deve-se usar a matriz pseudo-inversa de Moore-Penrose, definida como

$$\mathbf{D} = \mathbf{G}^\dagger = \mathbf{G}^H (\mathbf{G}\mathbf{G}^H)^{-1}. \quad (12)$$

O desempenho deste detector é prejudicado por canais mal-condicionados, onde os autovalores de $(\mathbf{G}\mathbf{G}^H)^{-1}$ são próximos de zero.

No caso de terminais com pleno conhecimento de \mathbf{G} , tem-se que $\mathbf{G}^\dagger\mathbf{G} = \mathbf{I}$.

2.1.5 *Linear Minimal Mean Square Error*

O detector LMMSE é obtido fazendo-se $\mathbf{D} = (\mathbf{G} + \sigma\mathbf{I})^\dagger$, onde o objetivo é adicionar um valor inversamente proporcional à SNR na diagonal principal da matriz do canal, minimizando o problema de mal-condicionamento quando a SNR for baixa. Assim, a estimativa do vetor de símbolos transmitidos é dada por

$$\hat{\mathbf{q}} = (\mathbf{G} + \sigma\mathbf{I})^\dagger y \quad (13)$$

onde σ é a variância do ruído. Neste caso, assumiu-se que a energia do sinal transmitido está normalizada para o valor unitário.

3 Técnicas avançadas e de alto desempenho

Davi da Silva Brilhante, Pedro Henrique Carneiro de Souza, Jonathan Nogueira Gois,
 dbrilhante@land.ufrj.br, pedro.carneiro@dtel.inatel.br, jonathan.gois@cefet-rj.br,

O projeto de técnicas de detecção MIMO exige um balanço entre a complexidade do algoritmo e seu desempenho quanto à detecção, usualmente medida em BER. Técnicas de detecção lineares têm implementação simples e baixa complexidade. Contudo, a robustez das técnicas de detecção lineares é muitas vezes comprometida.

Por exemplo, o detector *Zero Forcing* tem, no pior caso, complexidade computacional $\mathcal{O}(M^3)$. Contudo, se o valor dos autovalores de $(\mathbf{G}\mathbf{G}^H)^{-1}$ é pequeno, o processo de detecção introduz uma amplificação no ruído, reduzindo o desempenho. Além disso, os detectores lineares não conseguem colher a diversidade presente nos sinais recebidos. O efeito da amplificação do ruído na saída do detector ZF é dado por

$$\begin{aligned} P_\omega &= [\mathbf{G}^\dagger \omega]^2 = [(\mathbf{G}\mathbf{G}^H)^{-1} \mathbf{G}^H]^2 \omega \omega^H \\ &= (\mathbf{G}\mathbf{G}^H)^{-1} \omega \omega^H \end{aligned} \quad (14)$$

O baixo condicionamento da matriz do canal é causada por uma alta correlação entre os sinais transmitidos, o que pode acontecer em função da proximidade entre as antenas transmissoras e receptoras ou pela existência de linha de visada direta entre todas as antenas. Outros detectores, tipicamente não-lineares, podem apresentar um desempenho melhor neste cenário, sem um custo proibitivo para a complexidade de implementação. Esta seção irá explorar alguns destes detectores.

3.1 *Successive Interference Cancellation*

Os algoritmos baseados em *Successive Interference Cancellation* (SIC) partem do princípio que um dos símbolos transmitidos pode ser detectado com elevada probabilidade de acerto e que a matriz do canal está disponível na recepção. Ao receber este símbolo com alta confiabilidade, é possível remover a sua influência nos demais símbolos ponderando-o pelo ganho do canal e subtraindo-o do sinal recebido. Desta forma, a detecção do próximo símbolo se dará sem a interferência do primeiro. Isso é feito de forma sucessiva até que todos os símbolos tenham sido recebidos. Este procedimento consegue colher alguma ordem de diversidade presente no canal a partir do segundo símbolo detectado, sendo que o último símbolo é aquele que terá o maior ganho de diversidade. Para que este algoritmo funcione adequadamente, é importante garantir que sempre os símbolos mais confiáveis sejam detectados primeiro, pois um erro de detecção no início do processo irá resultar em uma propagação de erros, influenciando o desempenho do sistema. Isto levou à busca por uma ordem de detecção ótima, originando a variante chamada de *Ordered Successive Interference Cancellation* (OSIC).

Considere novamente que o sinal recebido pelo sistema MIMO seja dado por (1), reescrita a seguir com a notação $(\cdot)^{(i)}$ para representar o sinal a ser empregado para o i -ésimo símbolo a ser detectado.

$$\mathbf{y}^{(0)} = \mathbf{G}\mathbf{q} + \omega, \quad (15)$$

Existem diversas formas para detectar o primeiro símbolo do vetor de dados transmitidos. Uma destas formas consiste em utilizar o detector ZF. Neste caso, o primeiro passo consiste em encontrar a matriz de detecção dada por

$$\mathbf{B}^{(0)} = \mathbf{G}^\dagger. \quad (16)$$

Em seguida, deve-se identificar o símbolo que pode ser recebido com maior probabilidade de sucesso. Isso é feito analisando o fator de amplificação de ruído, que é dado pela soma do módulo quadrado dos elementos presentes em cada linha da matriz \mathbf{B} . Assim, o fator de amplificação para o i -ésimo símbolo é dado por

$$\xi_i^{(0)} = \|\mathbf{B}_i^{(0)}\|_2^2 = \sum_{j=0}^{N-1} |\mathbf{B}_{i,j}^{(0)}|^2, \quad (17)$$

com $i \in \{0, 1, \dots, M-1\}$. Deve-se escolher o símbolo mais confiável aquele com índice associado ao menor valor de $\xi^{(0)}$. Assuma que o índice que leva ao menor valor de $\xi^{(0)}$ seja s_0 . Então, o primeiro símbolo ser detectado é aquele transmitido pela antena s_0 e, para isso, deve-se empregar a linha s_0 da matriz de detecção $\mathbf{B}^{(0)}$, de modo que

$$x_{s_0} = \mathbf{B}_{s_0,:}^{(0)} y^{(0)}, \quad (18)$$

onde a operação $\mathbf{A}_{i,:}$ resulta na i -ésima linha de uma matriz \mathbf{A} qualquer. Comparando o resultado x_{s_0} com os símbolos da constelação utilizada na transmissão, pode-se estimar o símbolo que foi transmitido, \hat{q}_{s_0} , como sendo

$$\hat{q}_{s_0} = \mathcal{Q}_C[x_{s_0}] = \arg \min_{c \in \mathcal{C}} \|c - x_{s_0}\|, \quad (19)$$

onde c é o símbolo candidato da constelação usada na transmissão e \mathcal{C} é o conjunto de todos os símbolos possíveis.

Assumindo que a estimativa foi correta, isto é, $q_{s_0} = \hat{q}_{s_0}$, é possível cancelar a interferência causada pelo sinal detectado \hat{q}_{s_0} sobre o sinal recebido no primeiro estágio de detecção $y^{(0)}$ e assim obter

$$y^{(1)} = y^{(0)} - \mathbf{G}_{:,s_0}^{(0)} \hat{q}_{s_0}, \quad (20)$$

onde a operação $\mathbf{A}_{:,i}$ resulta na i -ésima coluna de uma matriz \mathbf{A} qualquer. Note que para $y^{(1)}$ é como se o símbolo q_{s_0} não tivesse sido transmitido. Assim, o canal equivalente para $y^{(1)}$ corresponde à matriz $\mathbf{G}^{(0)}$ com a s_0 -ésima coluna nula, de modo que

$$\mathbf{G}^{(1)} = \mathbf{G}_{:,s_0}^{(0)} \leftarrow \mathbf{0}_{M,1}, \quad (21)$$

onde a operação \leftarrow denota a substituição de linhas e colunas de uma matriz e $\mathbf{0}_{i,j}$ representa a matriz toda zero de dimensão $i \times j$.

A partir deste ponto, o processo se repete até os M símbolos sejam detectados, tal como mostra o Algoritmo 1.

Este algoritmo tem complexidade de ordem $\mathcal{O}(M^4)$ [6], principalmente em função das sucessivas inversões de matrizes [7], o que pode ser uma complexidade proibitiva em algumas aplicações [8].

Algorithm 1: Algoritmo de detecção *Zero-Forcing SIC*

```

1 Entrada:  $\mathbf{G}$ ,  $\mathbf{y}$ ,  $\sigma^2$ ;
2  $\mathbf{B}^{(0)} \leftarrow (\mathbf{G}^H \mathbf{G})^{-1} \mathbf{G}^H$ ;
3  $\ell \leftarrow 0$ ;
4 while  $\ell < M$  do
5    $\xi_j^{(\ell)} \leftarrow \|\mathbf{B}_j^{(\ell)}\|_2^2$ ;
6    $s_\ell \leftarrow \arg \min_j (\xi_j^{(\ell)})$ ;
7    $x_{s_\ell} = \mathbf{B}_{s_\ell}^{(\ell)} \mathbf{y}^{(\ell)}$ ;
8    $\hat{q}_{s_\ell} = \mathcal{Q}_C[x_{s_\ell}]$ ;
9    $\mathbf{y}^{(\ell+1)} = \mathbf{y}^{(\ell)} - \mathbf{G}_{:,s_\ell}^{(\ell)} \hat{q}_{s_\ell}$ ;
10   $\mathbf{G}^{(\ell+1)} \leftarrow$  preencha a coluna  $s_\ell$  de  $\mathbf{G}^{(\ell)}$  com  $\mathbf{0}_{M,1}$ ;
11   $\mathbf{B}^{(\ell+1)} = \left( (\mathbf{G}^{(\ell+1)})^H \mathbf{G}^{(\ell+1)} \right)^{-1} (\mathbf{G}^{(\ell+1)})^H$ ;
12   $\ell \leftarrow \ell + 1$ 
13 end

```

Outros métodos de cancelamento de interferência foram propostos na literatura com o objetivo de reduzir a complexidade. Por exemplo, utilizando a decomposição de Gram-Schmidt (também conhecida como decomposição QR) [9, 10], que evita o cálculo da matriz pseudo-inversa do canal a cada iteração. Na decomposição de Gram-schmidt, a matriz de canal é decomposta em outras duas matrizes, $\mathbf{G} = \mathbf{U}\mathbf{R}$. A matriz \mathbf{U} é unitária, e a matriz \mathbf{R} é triangular superior. Dado que \mathbf{R} é triangular superior, ao calcularmos o sinal recebido \mathbf{y} , temos que

$$\mathbf{y} = \mathbf{G}\mathbf{q} + \boldsymbol{\omega} = \mathbf{U}\mathbf{R}\mathbf{q} + \boldsymbol{\omega} \quad (22)$$

que se multiplicado por \mathbf{U}^H , resulta no sinal transmitido com a interferência suprimida no lado direito de(22), o que leva à

$$\begin{aligned} \mathbf{z} &= \mathbf{U}^H \mathbf{y} = \mathbf{R}\mathbf{q} + \mathbf{U}^H \boldsymbol{\omega} \\ &= \begin{bmatrix} r_{1,1} & r_{1,2} & \cdots & r_{1,N} \\ 0 & r_{2,2} & \cdots & r_{2,N} \\ \vdots & \vdots & \ddots & \cdots \\ 0 & 0 & & r_{N,N} \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_N \end{bmatrix} + \mathbf{U}^H \boldsymbol{\omega}, \end{aligned} \quad (23)$$

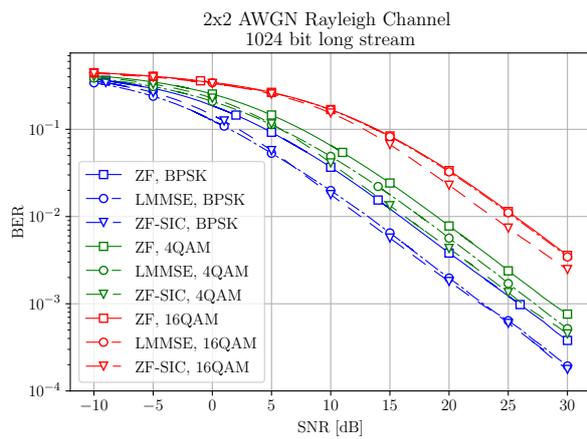
ou seja, a i -ésima coluna da matriz resultante \mathbf{z} só é afetada pela interferência dos $i - 1$ símbolos anteriores. Logo, a matriz \mathbf{R} permite fazer um cancelamento de interferências linha a linha devido a sua estrutura triangular superior. Essa transformação tem o mesmo efeito do SIC que calcula a interferência e faz o seu cancelamento a cada iteração. No entanto, se comparado ao OSIC, ainda é necessário um método de ordenação ótimo para o cancelamento das interferências. Um método de ordenação sub-ótimo é proposto em [10], calculando a matriz \mathbf{R} linha a linha de cima para baixo e a matriz \mathbf{U} da esquerda para a direita coluna a coluna, reduzindo a complexidade para $\mathcal{O}(M^2)$, mas com uma pequena degradação no desempenho em termos de BER e *Frame Error Rate* (FER).

Ainda, outros métodos do cálculo do vetor de anulação são possíveis, como o MMSE [11, 7, 12]. Conforme demonstrado em [11], com o MMSE é possível atingir BER menores para uma mesma SNR em comparação com o OSIC, se o critério de ordenação ótima for aplicado. O critério de ordenação ótima [11] é dado por

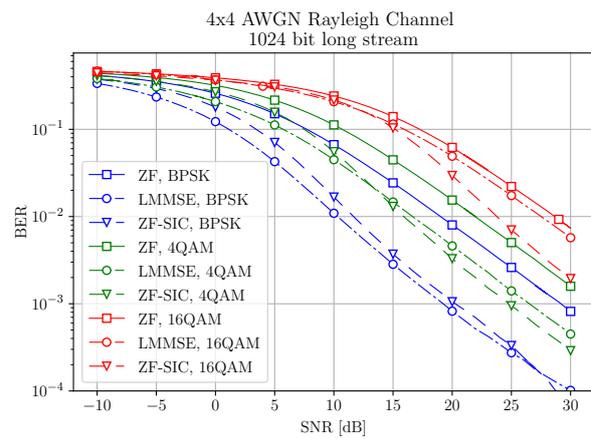
$$\rho_{s_i}^{MMSE} = \frac{\|(\mathbf{D}\mathbf{G})_{s_i}\|^2 |q_{s_i}|^2}{\sigma^2 \|\mathbf{D}\|_{s_i}^2} \quad (24)$$

onde, $\mathbf{D} = (\mathbf{G} + \sigma\mathbf{I})^\dagger$. Nesse caso, $\rho_{s_i}^{MMSE}$ denota a pós-SNR de um detector utilizando MMSE.

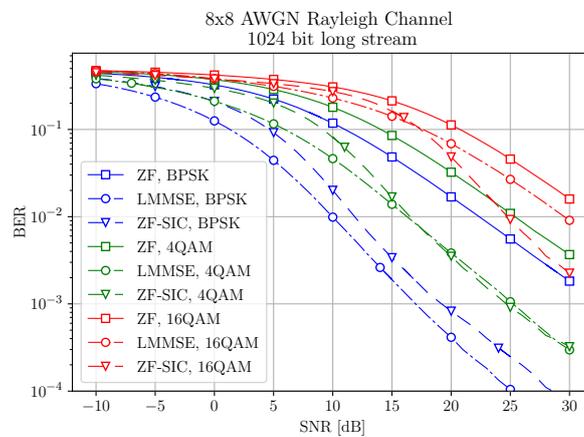
A Figura 3 apresenta a análise de desempenho das técnicas ZF, LMMSE e ZF-SIC (com ordenação ótima). É esperado que o ZF tenha o pior desempenho em geral por causa da amplificação do ruído mencionada anteriormente. Em relação ao número de antenas, podemos notar que as técnicas baseadas em ZF não são robustas ao aumento do número de antenas, devido ao aumento da interferência [2]. Devido a isto e à propagação de erros no cancelamento da interferência, o ZF-SIC não supera o LMMSE em termos de BER em cenários de SNR mais baixa quando usada a modulação *Binary Phase-Shift Keying* (BPSK). Em contrapartida, o aumento da SNR e uso das modulações *Quadrature Phase Shift Keying* (QPSK) e 16-QAM (*Quadrature Amplitude Modulation*), levam o ZF-SIC a um desempenho melhor do que o LMMSE. Nesse caso, o cancelamento de interferência do ZF-SIC é mais efetivo devido à alta SNR, enquanto o LMMSE sofre com os efeitos da interferência.



(a) Sistema MIMO com $N = M = 2$



(b) Sistema MIMO com $N = M = 4$



(c) Sistema MIMO com $N = M = 8$

Figura 3: Comparação da BER dos detectores ZF, LMMSE e ZF-SIC(V-BLAST) para as modulações BPSK (azul), 4QAM (verde) e 16QAM (vermelho), com diferentes configurações de antenas no receptor e transmissor.

3.2 Sphere Detection

O algoritmo denominado SD [13] foi proposto visando reduzir o custo computacional em relação ao algoritmo de máxima verossimilhança, mas mantendo um desempenho em termos de BER muito próximo [14, 15, 16]. Novas contribuições para reduzir ainda mais a complexidade foram apresentadas em [17]. Como descrito anteriormente, a complexidade elevada do algoritmo de máxima verossimilhança é devida ao grande número de vetores mensagem ao qual o sinal recebido deve ser comparado para encontrar a solução final [18].

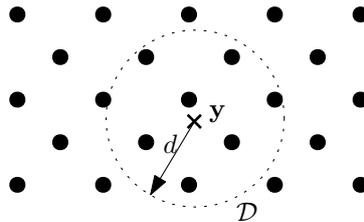


Figura 4: Ideia do algoritmo SD.

O princípio do algoritmo SD consiste em definir uma esfera multidimensional \mathcal{D} centrada no vetor de símbolos recebido \mathbf{y} [18], tal como mostrado na Figura 4. O raio da esfera d é definido de tal forma que um número limitado de símbolos estejam contidos em \mathcal{D} depois de serem transformados pelo canal \mathbf{G} . O raio de busca do SD é dado por

$$d^2 < \|\mathbf{y} - \mathbf{G}\mathbf{q}\|^2. \quad (25)$$

De modo geral, se d é grande o suficiente para englobar todos os símbolos testados na decodificação por máxima verossimilhança, ambos os algoritmos possuem a mesma complexidade computacional. Por isso, deve-se ajustar o raio da esfera d de modo que um número reduzido de símbolos seja testado. Uma maneira de encontrar d é utilizar a estimativa obtida pelo ZF [19], de modo que

$$d = \|\mathbf{y} - \mathbf{G}\hat{\mathbf{q}}_B\|, \quad (26)$$

onde

$$\hat{\mathbf{q}}_B = \lfloor \mathbf{G}^\dagger \mathbf{y} \rfloor. \quad (27)$$

Assumindo que \mathbf{G} tem colunas linearmente independentes então a matriz do canal por ser decomposta como

$$\begin{aligned} \mathbf{G} &= \mathbf{Q}\mathbf{R} \\ &= [\mathbf{Q}_1 \mathbf{Q}_2] \begin{bmatrix} \mathbf{R}' \\ \mathbf{0} \end{bmatrix} \end{aligned} \quad (28)$$

onde $\mathbf{Q} \in \mathbb{R}^{M \times M}$ é uma matriz ortonormal, $\mathbf{R} \in \mathbb{R}^{M \times N}$ é uma matriz triangular superior e $\mathbf{R}' \in \mathbb{R}^{N \times N}$ é parte de \mathbf{R} sendo também uma matriz triangular superior. Aplicando (28) em (25), tem-se [18]

$$d^2 - \|\mathbf{Q}_2^H \mathbf{y}\|^2 \geq \|\mathbf{Q}_1^H - \mathbf{R}\mathbf{q}\|^2 \quad (29)$$

Fazendo uma substituição de variáveis, empregando por $d'^2 = d^2 - \|\mathbf{Q}_2^H \mathbf{y}\|^2$ e $\mathbf{z} = \mathbf{Q}_1^H \mathbf{y}$ em (29), obtém-se

$$d'^2 \geq \sum_{i=1}^N \left(z_i - \sum_{j=i}^N r_{i,j} q_j \right)^2 \quad (30)$$

onde $r_{i,j}$ denota a entrada (i, j) de \mathbf{R} .

Como adendo, caso $\mathbf{G} \in \mathbb{C}^{M \times N}$, pode-se fazer com que

$$\hat{\mathbf{G}} = \begin{bmatrix} \Re\{\mathbf{G}\} & -\Im\{\mathbf{G}\} \\ \Im\{\mathbf{G}\} & \Re\{\mathbf{G}\} \end{bmatrix} \quad (31)$$

$$\hat{\mathbf{y}} = \begin{bmatrix} \Re\{\mathbf{y}\} \\ \Im\{\mathbf{y}\} \end{bmatrix} \quad (32)$$

$$\hat{\mathbf{q}} = \begin{bmatrix} \Re\{\mathbf{q}\} \\ \Im\{\mathbf{q}\} \end{bmatrix}, \quad (33)$$

de modo que (30) pode ser expandida como [16]

$$d'^2 \geq (z_N - r_{N,N}q_N)^2 + (z_{N-1} - r_{N-1,N-1}q_{N-1})^2 + \dots \quad (34)$$

onde o primeiro termo depende apenas de q_N , o segundo termo de q_{N-1} e assim por diante. Portanto, a condição necessária para $\mathbf{G}\mathbf{q}$ estar disposto dentro da esfera é que $d'^2 \geq (y_N - r_{N,N}q_N)^2$. Isto é equivalente a dizer que q_N pertence ao intervalo [16]:

$$\left\lceil \frac{-d' + z_N}{r_{N,N}} \right\rceil \geq q_N \geq \left\lfloor \frac{d' + z_N}{r_{N,N}} \right\rfloor \quad (35)$$

onde $\lceil \cdot \rceil$ denota a operação de arredondamento para cima.

Considerando $d'_{N-1} = d'^2 - (z_N - r_{N,N}q_N)$ e $z_{N-1|N} = z_{N-1} - r_{N-1,N}q_N$, tem-se a recursão para encontrar os próximos termos de \mathbf{q} dada por

$$\left\lceil \frac{-d'_{K-1} + z_{K-1}}{r_{K-1,K-1}} \right\rceil \geq q_K \geq \left\lfloor \frac{d'_{K-1} + z_{K-1}}{r_{K-1,K-1}} \right\rfloor \quad (36)$$

Os passos para empregar o SD estão dispostos no Algoritmo 2.

Algorithm 2: Algoritmo do *Sphere Decoding* [16]

```

1 Entrada:  $\mathbf{Q} = [\mathbf{Q}_1 \ \mathbf{Q}_2]$ ,  $\mathbf{R}$ ,  $\mathbf{y}$ ,  $\mathbf{z} = \mathbf{Q}_1^* \mathbf{y}$  e  $d$ ;
2  $i \leftarrow N$ ,  $d'_m \leftarrow d^2 - \|\mathbf{Q}_2^* \mathbf{y}\|^2$ ,  $z_{m|m+1} = z_m$ ;
3 while  $i > 1$  do
    /* Limites para  $q_i$  */
4    $UB(q_i) \leftarrow \lfloor (d'_i + z_{i|i+1})/r_{i,i} \rfloor$ ;
5    $q_i = \lceil (-d'_i + z_{i|i+1})/r_{i,i} \rceil - 1$ ;
    /* Incrementando  $q_i$  */
6    $q_i \leftarrow q_i + 1$ ;
7   if  $q_i \leq UB(q_i)$  then
    /* Decrementando  $i$  */
8     if  $i = 1$  then
9       break;
10    end
11    else
12       $i \leftarrow i - 1$ ;
13       $z_{i|i+1} \leftarrow z_i - \sum_{j=i+1}^m r_{i,j} q_j$ ;
14       $d'_i{}^2 \leftarrow d'_{i+1}{}^2 - (y_{i+1|i+2} - r_{i+1,i+1} q_{i+1})^2$ ;
15    end
16  end
17  else
    /* Incrementando  $i$  */
18     $i \leftarrow i + 1$ ;
19    if  $i = N + 1$  then
20      break;
21    end
22  end
23 end
24 Saída:  $\mathbf{q}$ 

```

3.3 Lattice Reduction Aided

Proposto por Yao *et al.* [20], o algoritmo do detector que utiliza a técnica conhecida como LRA propõe baixa complexidade para a detecção e desempenho próximo do ótimo, ou seja, próximo da solução por máxima verossimilhança.

Inicia-se o algoritmo considerando que $N > M$, ou seja, existam mais antenas receptoras do que transmissoras. Além disso, também é considerado que o receptor tenha pleno conhecimento do canal \mathbf{G} ao qual o sinal de mensagem enviado \mathbf{q} foi submetido.

Considera-se um reticulado complexo n -dimensional tal que [20]:

$$\mathcal{L} = \{s | s = \mathbf{B}\lambda\}, \quad (37)$$

onde $\mathbf{B} = [\mathbf{b}_1 \ \mathbf{b}_2 \ \dots \ \mathbf{b}_n]$ é a matriz cujas colunas são vetores da base do reticulado e $\lambda = [\lambda_1 \ \lambda_2 \ \dots \ \lambda_n]^T$ é um vetor de complexo inteiros, ou seja $\lambda_i \in \mathbb{Z} + \mathbb{Z}j$.

Como disserta o autor [20], para qualquer reticulado \mathcal{L} existem muitas possibilidades de bases. Escolhendo uma base \mathbf{B} podemos transitar entre as representações fazendo $\mathbf{B}' = \mathbf{B}\mathbf{P}$,

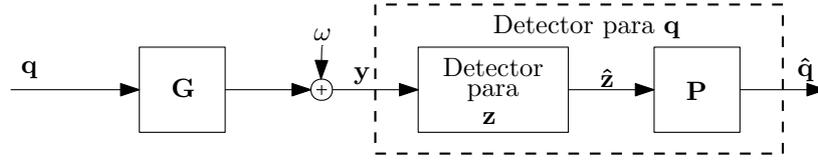


Figura 5: Fluxograma de detecção utilizando LRA.

onde \mathbf{B}' é uma nova base gerada através de uma dada matriz \mathbf{P} composta por números inteiros. Dado um símbolo \mathbf{s} representado por \mathbf{q} na base \mathbf{B} , este é representado por $\mathbf{z} = \mathbf{P}^{-1}\mathbf{q}$ na base \mathbf{B}' .

O detector de baixa complexidade baseado no algoritmo do LRA se dá na detecção de \mathbf{z} ao invés de \mathbf{q} , como indica a Figura 5.

Os autores [20] ratificam a funcionalidade do algoritmo descrito na Figura 5 realizando a mudança de base:

$$\mathbf{y} = \mathbf{G}\mathbf{q} + \omega, \quad (38)$$

$$= (\mathbf{G}\mathbf{P})(\mathbf{P}^{-1}\mathbf{q}) + \omega, \quad (39)$$

$$= \mathbf{G}'\mathbf{z} + \omega. \quad (40)$$

onde \mathbf{G}' é o canal aplicado ao transmissor caso seja enviado o símbolo transformado \mathbf{z} .

O detector para o símbolo \mathbf{z} pode ser, por exemplo, o detector por *zero-forcing*, onde $\hat{\mathbf{z}} = \mathbf{Q}(\mathbf{G}'^\dagger \mathbf{y})$, onde $\mathbf{Q}(\cdot)$ é a operação de quantização para o inteiro mais próximo [21]. Então, o problema a ser resolvido para o algoritmo do LRA é encontrar o canal efetivo \mathbf{G}' ao qual um detector linear qualquer pode ter desempenho em taxa de erros de bits semelhante ao algoritmo de detecção por máxima verossimilhança.

De fato, a ideia primordial do algoritmo do LRA é seu uso em conjunto com detectores de baixa complexidade tradicionais. Para isso, contudo, é necessário encontrar a base \mathbf{B}' que reduz o reticulado, ou seja, base formada por vetores aproximadamente ortogonais [21].

Então, o problema da detecção usando o algoritmo do LRA se reduz ao fato de encontrar tal espaço, descrito na base \mathbf{B}' . Mais precisamente, queremos encontrar a matriz \mathbf{P} ao qual $\mathbf{B}' = \mathbf{B}\mathbf{P}$. Dá-se o nome de *redução de espaço* a operação de mudança de representação através de tal matriz transformadora \mathbf{P} .

Pensando em complexidade computacional na execução da redução do espaço, existe o algoritmo proposto por Lenstra-Lenstra-Lovász (LLL) [22]. Inicia-se o algoritmo LLL tentando reduzir, iterativamente, a correlação entre os vetores que compõe a base \mathbf{B} . Então, dado que $\mathbf{B} = [\mathbf{b}_1 \ \mathbf{b}_2 \ \dots \ \mathbf{b}_M]$ e $\mathbf{B}' = [\mathbf{b}'_1 \ \mathbf{b}'_2 \ \dots \ \mathbf{b}'_M]$. Inicialmente, fazemos $\mathbf{B} = \mathbf{B}'$. A maneira intuitiva [20] de reduzir a correlação entre dois vetores da base é fazendo $\mathbf{b}'_2 = \mathbf{b}_2 - n\mathbf{b}_1$. Ou seja, retirando cópias de \mathbf{b}_1 de \mathbf{b}_2 . Dado isso, para encontrar \mathbf{b}'_2 basta encontrar o n ótimo, ou seja, n^* que minimiza a correlação entre \mathbf{b}_1 e \mathbf{b}'_2 , ou seja [20]:

$$n^* = \underset{n \in \mathbb{C}}{\operatorname{argmin}} \left| \langle \mathbf{b}_1, \mathbf{b}_2 \rangle - n \|\mathbf{b}_1\|^2 \right| = \left\lfloor \frac{\langle \mathbf{b}_1, \mathbf{b}_2 \rangle}{\|\mathbf{b}_1\|^2} \right\rfloor. \quad (41)$$

Dado o novo \mathbf{b}'_2 , podemos substituir \mathbf{b}_2 pelo vetor encontrado \mathbf{B}' . Caso \mathbf{b}'_2 seja menor que \mathbf{b}_1 , troca-se as colunas de posição (ou seja, operação de *swap*) e verifica-se as próximas subtrações possíveis até \mathbf{b}_M .

O algoritmo do LLL pode ser descrito nos seguintes passos [20]:

1. Iniciar o algoritmo com $i = 1$;
2. Checar a correlação entre \mathbf{b}_i e \mathbf{b}_{i+1} . Caso $\text{Re}\{\langle \mathbf{b}_i, \mathbf{b}_{i+1} \rangle\} \leq \frac{1}{2}\|\mathbf{b}_i\|^2$ e $\text{Im}\{\langle \mathbf{b}_i, \mathbf{b}_{i+1} \rangle\} \leq \frac{1}{2}\|\mathbf{b}_i\|^2$, parar. Caso contrário, substituir \mathbf{b}_{i+1} por $\mathbf{b}_{i+1} - \left\lfloor \frac{\langle \mathbf{b}_i, \mathbf{b}_{i+1} \rangle}{\|\mathbf{b}_i\|^2} \right\rfloor \mathbf{b}_i$ e continuar o algoritmo.
3. Checar os comprimentos de \mathbf{b}_i e \mathbf{b}_{i+1} . Se $\|\mathbf{b}_{i+1}\| > \|\mathbf{b}_i\|$, pare. Caso contrário, faça o *swap* entre \mathbf{b}_i e \mathbf{b}_{i+1} , faça $i \leftarrow i + 1$ e volte ao passo anterior.

Com isso, é possível encontrar \mathbf{B}' . Para encontrar a matriz de transformação \mathbf{P} ao qual estamos interessados, fazemos:

$$\mathbf{P} = \mathbf{B}^{-1}\mathbf{B}'. \quad (42)$$

Para reduzir ainda mais a complexidade no processo de redução de espaço, alguns algoritmos foram propostos na literatura. Iniciemos com a descrição do algoritmo conhecido como *Complex Lenstra-Lenstra-Lovász* (CLLL) [23]. A implementação de tal algoritmo se dá pela redução de espaço utilizando o LLL na matriz equivalente valor-real da matriz complexa do canal \mathbf{G} . Ou seja:

$$\mathbf{G}_R = \begin{bmatrix} \Re\{\mathbf{G}\} & -\Im\{\mathbf{G}\} \\ \Im\{\mathbf{G}\} & \Re\{\mathbf{G}\} \end{bmatrix}. \quad (43)$$

As taxas de erro de bits do LRA utilizando o CLLL e o LLL são idênticas. O ganho de tal algoritmo se dá pela redução de 50% da complexidade computacional sem perda de desempenho [23].

Outra solução para a redução de complexidade computacional foi apresentada por Zu *et al.* [24] utilizando a técnica denominada como *Regularized Block Diagonalization* (RBD). Foi proposto o algoritmo *Low Complexity Regularized Block Diagonalization Lattice Reduction* (LC-RBD-LR) que apresenta redução de 46.1% em relação ao algoritmo do CLLL tradicional. Nesta abordagem, utiliza-se a abordagem da matriz de canal \mathbf{G} por blocos para a realização da redução do espaço.

Por fim, é possível ainda reduzir o problema de complexidade do algoritmo LLL através da interpretação por *quaternions* [25]. Nesta interpretação, utilizamos a descrição do canal \mathbf{G} com entradas pertencentes ao conjunto \mathbb{H} tal que:

$$\mathbb{H} = \{q = q^{\{1\}} + q^{\{2\}}j \mid q^{\{1\}}, q^{\{2\}} \in \mathbb{C}\} \quad (44)$$

$$= \{q = q^1 + q^2i + q^3j + q^4k \mid q^1, q^2, q^3, q^4 \in \mathbb{R}\}. \quad (45)$$

As unidades *quaterniônicas* i , j e k mantêm relação que pode ser descrita na tabela abaixo:

Tomando a representação do canal $\mathbf{G} \in \mathbb{C}^{M \times N}$ em *quaternions* por $\tilde{\mathbf{G}} \in \mathbb{H}^{M/2 \times N/2}$. A representação de valor-real da matriz $\tilde{\mathbf{G}}$ é dada por:

$$\tilde{\mathbf{G}}_R = \begin{bmatrix} \tilde{\mathbf{G}}^{(1)} & -\tilde{\mathbf{G}}^{(2)} & -\tilde{\mathbf{G}}^{(3)} & -\tilde{\mathbf{G}}^{(4)} \\ \tilde{\mathbf{G}}^{(2)} & \tilde{\mathbf{G}}^{(1)} & -\tilde{\mathbf{G}}^{(4)} & \tilde{\mathbf{G}}^{(3)} \\ \tilde{\mathbf{G}}^{(3)} & \tilde{\mathbf{G}}^{(4)} & \tilde{\mathbf{G}}^{(1)} & -\tilde{\mathbf{G}}^{(2)} \\ \tilde{\mathbf{G}}^{(4)} & -\tilde{\mathbf{G}}^{(3)} & \tilde{\mathbf{G}}^{(2)} & \tilde{\mathbf{G}}^{(1)} \end{bmatrix}. \quad (46)$$

A partir de $\tilde{\mathbf{G}}$, pode-se estabelecer um novo reticulado através de quaisquer algoritmos de redução de espaço descrito anteriormente.

Tabela 1: Relação de multiplicação entre as unidades *quaterniônicas*.

	1	i	j	k
1	+1	i	$+j$	$+k$
i	$+i$	-1	$+k$	$-j$
j	$+j$	$-k$	-1	$+i$
k	$+k$	$+j$	$-i$	-1

3.4 *Approximate Message-passing*

Recentemente foi também proposto o algoritmo AMP, baseado na aproximação de um algoritmo muito utilizado na representação de grafos, esse por sua vez denominado de *message passing* [26]. Tipicamente o AMP pode ser empregado para reconstruir vetores esparsos a partir de observações lineares, sendo que o número de observações é menor que a dimensão desses vetores. Isso representa um problema oriundo da amostragem e compressão simultânea de sinais. No entanto, o AMP também pode ser utilizado no contexto de detecção de sinais em sistemas MIMO.

Para tanto, primeiro considere que o vetor transmitido, $\mathbf{q} \in \mathbb{C}^{K \times 1}$, seja estimado pelos seguintes passos:

$$\begin{aligned}
 1^\circ \text{ Passo: } \mathbf{z}_t &= \hat{\mathbf{q}}_t + \mathbf{G}^H (\mathbf{y} - \mathbf{G}\hat{\mathbf{q}}_t) + \mathbf{b}_t; \\
 2^\circ \text{ Passo: } \mathbf{b}_t &= \frac{\tau_t}{1 + \tau_{t-1}} [\mathbf{G}^H (\mathbf{y} - \mathbf{G}\hat{\mathbf{q}}_{t-1}) + \mathbf{b}_{t-1}]; \\
 3^\circ \text{ Passo: } \hat{\mathbf{q}}_{t+1} &= \eta_t (\mathbf{z}_t; \sigma_t);
 \end{aligned} \tag{47}$$

onde $t \in \{0, 1, \dots, T-1\}$ indica a t -ésima iteração do algoritmo AMP, $\hat{\mathbf{q}}_T$ é a estimativa do vetor transmitido, para o qual $\hat{\mathbf{q}}_0 = \mathbf{0}_K$ e $\mathbf{b}_0 = \mathbf{0}_K$, e também note que $\sigma_t = 0.5\sigma^2(1 + \tau_t)$. Além disso, perceba que

$$\tau_t = \frac{2K}{\sigma^2 M} \langle \eta'_t \rangle, \tag{48}$$

onde $\langle v \rangle = N^{-1} \sum_{i=0}^{N-1} v(i)$ e

$$\eta'_t = \sum_{s_m \in \mathbb{S}} s_m^2 P_m(\mathbf{q}_t = s_m | \mathbf{z}_t) - \left[\overbrace{\sum_{s_m \in \mathbb{S}} s_m P_m(\mathbf{q}_t = s_m | \mathbf{z}_t)}^{\eta_t} \right]^2. \tag{49}$$

O *denoiser* ótimo Gaussiano é representado pela função $f(x) = \eta_t(\cdot; \cdot)$, a qual é operada de acordo com o mapeamento $f : \mathbb{R} \rightarrow \mathbb{R}$, e onde temos

$$P_m = \frac{\exp(-2\sigma_t^{-1} \|z_t - s_m\|)}{\sum_{m=0}^{\sqrt{L}-1} \exp(-2\sigma_t^{-1} \|z_t - s_m\|)}, \tag{50}$$

representando a parcela não-linear dessa função. O intuito de $\eta_t(\cdot; \cdot)$ é aprimorar as estimativas $\hat{\mathbf{q}}_t$ a cada iteração, tendo em vista que o vetor do sinal recebido é usualmente contaminado pelo ruído AWGN. Por fim, supõe-se também que o vetor transmitido assume valores de um esquema de modulação digital qualquer como, por exemplo, constelações *L-Quadrature Amplitude Modulation* (QAM). Dessa forma, para uma constelação de energia unitária temos

$\mathbb{S} = \{\pm E_0, \pm 3E_0, \dots, \pm(\sqrt{L} - 1)E_0\}$; $E_0 = \sqrt{\frac{3}{2(L-1)}}$, representando coordenadas tanto na parte real quanto na parte imaginária do vetor transmitido.

Com isso, depreende-se de (47) que conforme $\hat{\mathbf{q}}_t \rightarrow \mathbf{q}$, também segue que $(\mathbf{y} - \mathbf{G}\hat{\mathbf{q}}_t) \rightarrow \omega$ e $\mathbf{b}_t \rightarrow 0$. Em outras palavras, a medida que as estimativas se aproximam do vetor transmitido, tem-se no 1º Passo de (47) um resultado similar àquele obtido pela equalização ZF. No entanto, percebe-se que não há cálculos de matrizes inversas no algoritmo AMP, o que diminui sua complexidade de implementação. Conseqüentemente, o AMP também se apresenta como uma boa alternativa em cenários onde o canal é mal-condicionado.

A Figura 6 apresenta análises de desempenho para sistemas MIMO 8×8 (a), 16×64 (b),

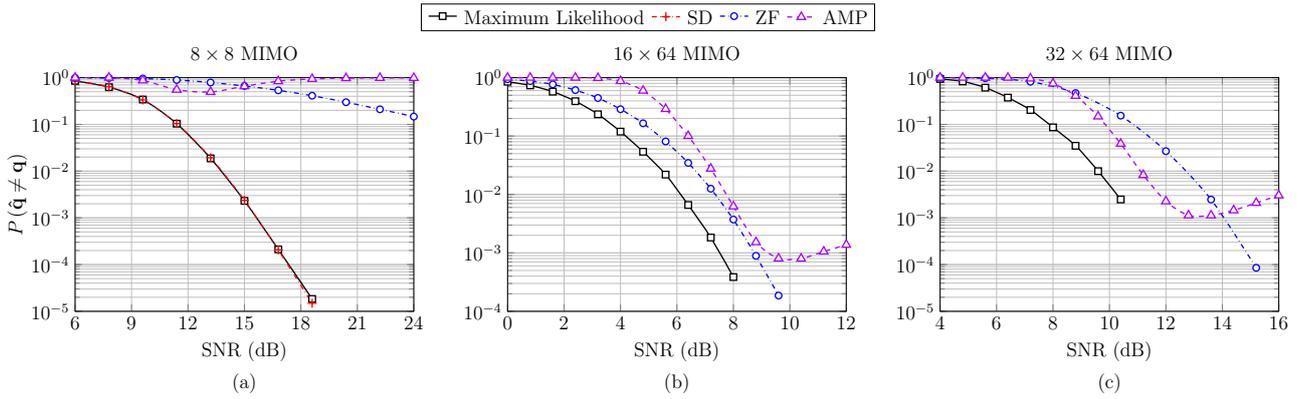


Figura 6: Desempenho dos detectores por máxima verossimilhança (5), SD, ZF (11) e AMP (47), considerando a probabilidade de erro do vetor transmitido para uma faixa de valores de SNR.

32×64 (c), considerando o número de antenas de transmissão e recepção, respectivamente. Para isso, a probabilidade de erro do vetor transmitido, $P(\hat{\mathbf{q}} \neq \mathbf{q})$, é estimada para uma faixa de valores de SNR e também para diferentes algoritmos de detecção, a saber, detector por máxima verossimilhança (5), SD, ZF (11) e o algoritmo AMP (47) estudado nesta subseção. Além disso, ressalta-se que o esquema de modulação 16-QAM e o canal Rayleigh lento-plano foram considerados, de tal forma que $G_{i,j} \sim \mathcal{CN}(0, 1/M)$, $\forall i, j$. Sendo assim, a SNR por bit pode ser calculado por

$$\text{SNR} = 10 \log \left(\frac{\mathbb{E}[\|\mathbf{G}\mathbf{q}\|_2^2]}{M\sqrt{L}\sigma^2} \right). \quad (51)$$

Verifica-se na Figura 6 (a) que os desempenhos tanto do algoritmo AMP quanto do detector ZF estão significativamente aquém daquele obtido pelo SD e, principalmente, pelo detector ótimo de máxima verossimilhança. Para valores de SNR ~ 15 dB, por exemplo, a $P(\hat{\mathbf{q}} \neq \mathbf{q})$ alcançada pelos detectores SD e ótimo é aproximadamente de 10^{-3} . Por outro lado, o ZF e o AMP mantêm seus desempenhos praticamente inalterados em $P(\hat{\mathbf{q}} \neq \mathbf{q}) \sim 10^{-1}$, para toda faixa de valores de SNR analisados. Esse desempenho proibitivo do ZF e do AMP se deve ao alto grau de IAI presente no sistema MIMO 8×8 , em que o número de antenas transmissoras é igual ao de antenas receptoras. Desta forma, observe nas Figuras 6 (b) e (c), que o desempenho dos detectores ZF e AMP se aproximam do alcançado pelo detector ótimo a medida que aumenta a razão entre o número de antenas receptoras em relação ao de antenas transmissoras. Isso acontece pois a ordem de diversidade é maior, ao passo que o número de diferentes fluxos de dados recebidos é proporcionalmente menor e, conseqüentemente, a IAI é mitigada.

Por fim, note na Figura 6 (c) que o AMP supera o detector ZF para a maior parte dos valores de SNR analisados. No entanto, para altos valores de SNR > 12 dB, o AMP atinge um *floor* de desempenho, de tal forma que os valores de BER não mais diminuem conforme o aumento da SNR. A partir disso, conjectura-se que, em grande parte, é a IAI que contamina o sinal recebido, visto que o ruído AWGN é menor para maiores valores de SNR. Portanto, o sinal recebido não mais possui distribuição predominantemente *Gaussiana*, o que prejudica o desempenho do *denoiser* aplicado no 3º Passo de (47) e, por conseguinte, o desempenho do algoritmo AMP.

3.5 Deep Unfolded MIMO

Vem sendo demonstrado que algoritmos de aprendizado de máquina e arquiteturas de redes neurais também podem ser integrados a sistemas de telecomunicações, especialmente no que diz respeito a detecção de sinais [27, 28]. Mais especificamente para sistemas MIMO, recentemente foi proposta [29] uma solução que visa combinar arquiteturas de redes neurais e algoritmos que não necessariamente utilizam-se do aprendizado de máquina, como, por exemplo, o AMP. Isso é concebido de tal forma a permitir que as iterações de um dado algoritmo seja mapeada ou ‘desdobrada’ nas múltiplas camadas da rede neural. Consequentemente, o aprendizado é aprimorado pelos algoritmos que inerentemente trazem informações sobre parâmetros do sistema, enquanto que o desempenho dos mesmos são, por sua vez, aprimorados pelo processo de aprendizado.

Essa solução de arquitetura de redes neurais pode ser, portanto, definida da seguinte forma: Inicialmente considere que um algoritmo qualquer têm seus parâmetros de entrada descritos pelo vetor $\mathbf{x} \in \mathbb{R}^V$ e que a saída é dada por $\mathbf{r} \in \mathbb{R}^S$; sendo assim

$$r(s) = g(\mathbf{x}, \boldsymbol{\psi}, \Theta), \quad \forall s \in \{0, 1, \dots, S-1\}, \quad (52)$$

em que Θ é o conjunto de todos os parâmetros que descrevem o próprio algoritmo, $g(\cdot)$ representa a função mapeadora, tipicamente não linear, e $\boldsymbol{\psi}$ é atualizado iterativamente como segue

$$\psi_\ell(s) = f(\mathbf{x}, \psi_{\ell-1}(s), \Theta), \quad (53)$$

onde a ℓ -ésima iteração também utiliza-se de uma função mapeadora $f(\cdot)$ e $\boldsymbol{\psi}_0$ representa o valor inicial empregado na primeira iteração.

Desta forma, $\boldsymbol{\psi}_\ell$ descreve a relação entre entrada e saída da ℓ -ésima camada de uma arquitetura de redes neurais, ao passo que também pode ser representativo de um algoritmo iterativo qualquer. Portanto, cada camada dessa arquitetura mapeia uma iteração do algoritmo subjacente, isto é, daquele algoritmo no qual ela se baseia. Com isso, temos a arquitetura denominada *deep unfolded* ilustrada na Figura 7.

Note que os parâmetros descritos por Θ são comumente denominados de parâmetros de aprendizagem no contexto da Figura 7. Esses parâmetros são otimizados pelo processo de aprendizagem ou, mais especificamente, pelo algoritmo de treinamento implementado no contexto das redes neurais. Em princípio, o número de camadas e neurônios da arquitetura são definidos de maneira a extrair o melhor resultado no estágio de treinamento. Isso envolve um custo elevado de tempo e recursos computacionais. No entanto, na arquitetura *deep unfolding* esses parâmetros já são definidos em sua maioria pelo próprio algoritmo subjacente, simplificando assim consideravelmente o dimensionamento dessa arquitetura de redes neurais.

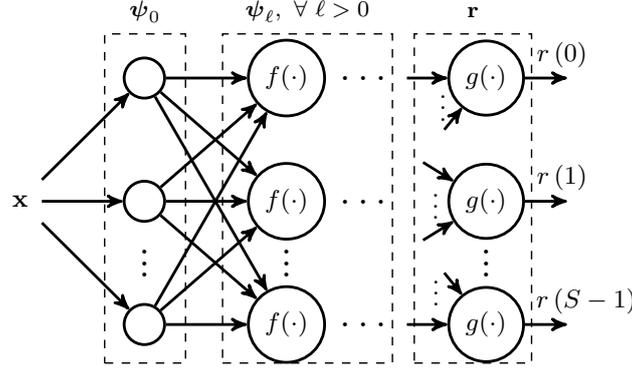


Figura 7: Arquitetura de redes neurais baseada nos princípios do *deep unfolding*.

Além disso, observe também que a representação de vetores e matrizes em valores puramente reais será utilizada daqui em diante nesta subseção. Assim, considere que o sinal recebido seja representado pela concatenação da sua parte real (\Re) e imaginária (\Im), tal que

$$\tilde{\mathbf{y}} = \tilde{\mathbf{G}}\tilde{\mathbf{q}} + \tilde{\boldsymbol{\omega}}, \quad (54)$$

onde

$$\tilde{\mathbf{y}} = [\Re(\mathbf{y})^T \ \Im(\mathbf{y})^T]^T \in \mathbb{R}^{2M}, \quad (55)$$

$$\tilde{\mathbf{G}} = \begin{bmatrix} \Re(\mathbf{G}) & -\Im(\mathbf{G}) \\ \Im(\mathbf{G}) & \Re(\mathbf{G}) \end{bmatrix} \in \mathbb{R}^{2M \times 2K}, \quad (56)$$

$$\tilde{\mathbf{q}} = [\Re(\mathbf{q})^T \ \Im(\mathbf{q})^T]^T \in \mathbb{R}^{2K}, \quad (57)$$

$$\tilde{\boldsymbol{\omega}} = [\Re(\boldsymbol{\omega})^T \ \Im(\boldsymbol{\omega})^T]^T \in \mathbb{R}^{2M}. \quad (58)$$

Sendo assim, em [29] foi proposto o DU-PDA, que combina a arquitetura *deep unfolding* com variações dos algoritmos AMP e, principalmente, com o algoritmo seminal denotado por PDA [30]. Inicialmente o sinal recebido $\tilde{\mathbf{y}}$ é pré-processado pela ℓ -ésima camada de acordo com a seguinte operação

$$\mathbf{z}_\ell = \tilde{\mathbf{q}}'_\ell + w_\ell \tilde{\mathbf{G}}^T \left(\tilde{\mathbf{y}} - \tilde{\mathbf{G}}\tilde{\mathbf{q}}'_\ell \right), \quad \forall \ell \in \{0, 1, \dots, Z-1\}, \quad (59)$$

onde $\tilde{\mathbf{q}}'_\ell \in \mathbb{R}^{2K}$ é a estimativa do vetor transmitido e o escalar $w_\ell \in \mathbb{R}$ representa o parâmetro de aprendizado. Observe que tipicamente o PDA utiliza o ZF no pré-processamento [30], mas no DU-PDA o AMP é integrado e modificado a fim de incluir a ponderação pelo parâmetro de aprendizado w_ℓ . Desta forma, $\tilde{\mathbf{q}}'_\ell$ é aprimorado pelo aprendizado e atualizado iterativamente até que seja feita uma estimativa satisfatória do vetor transmitido.

O vetor resultante de (59) é então alimentado na seguinte operação¹:

$$\begin{aligned} \psi_{\ell^*}(m) &= \text{softm} \left((\mathbf{z}_\ell - \boldsymbol{\mu}_{\ell^*} - 0.5\mathbf{e}_{\ell^*} s_m)^T \boldsymbol{\Omega}_{\ell^*}^{-1} \mathbf{e}_{\ell^*} s_m \right) \\ &\quad \forall m \in \{0, 1, \dots, \sqrt{L}-1\}, \end{aligned} \quad (60)$$

para o qual temos

$$\text{softm}(x_\ell(m)) = \frac{e^{x_\ell(m)}}{\sum_{m=0}^{Z-1} e^{x_\ell(m)}}, \quad (61)$$

¹ $\{\ell^* \in \{0, 1, \dots, 2K-1\}, n \in \{1, 2, \dots, \lceil Z/2K \rceil - 1\} \mid \ell^* = \ell - n2K; n2K \leq \ell < (n+1)2K\}$

onde também \mathbf{e}_{ℓ^*} é o vetor com 1 (um) no ℓ^* -ésimo elemento e 0 (zero) no restante e $s_m \in \mathbb{S}$ representa as coordenadas da constelação L -QAM. Adicionalmente, temos que

$$\mathbf{\Omega}_{\ell^*} = \sum_{j \neq \ell^*} \mathbf{e}_j \mathbf{e}_j^T \left((\mathbf{s}^2)^T \boldsymbol{\psi}_j - \boldsymbol{\mu}_j^2 \right) + \mathbf{e}_{\ell^*} \mathbf{e}_{\ell^*}^T \text{COV}[\mathbf{z}_\ell - \tilde{\mathbf{q}}], \quad (62)$$

onde,

$$\text{COV}[\mathbf{z}_\ell - \tilde{\mathbf{q}}] = \frac{[\epsilon_\ell]_+ \|\mathbf{I}_{2K} - w_\ell \tilde{\mathbf{G}}^T \tilde{\mathbf{G}}\|_2^2 + 0.5\sigma^2 \|w_\ell \tilde{\mathbf{G}}^T\|_2^2}{2K}, \quad (63)$$

$$\boldsymbol{\mu}_j = \sum_{i \neq j} \mathbf{e}_i (\mathbf{s}^T \mathbf{p}_i); \quad (64)$$

$$\mathbf{p}_i = P_m(\tilde{q}(i) = s_m \mid \mathbf{z}_\ell, \{\mathbf{p}_i\}_{\forall i \neq j}), \quad \forall m, \quad (65)$$

tal que $[x]_+ = \max(0, x)$, e sendo que

$$\epsilon_\ell = \frac{\|\tilde{\mathbf{y}} - \tilde{\mathbf{G}}\tilde{\mathbf{q}}'_\ell\|_2^2 - M\sigma^2}{\|\tilde{\mathbf{G}}\|_2^2}. \quad (66)$$

A partir de (60) pode-se notar que a função não linear *softmax*, dada por (61), é utilizada em cada camada da arquitetura descrita inicialmente por (52), substituindo assim o *denoiser* implementado tipicamente pelo algoritmo AMP. Ressalta-se também que $\mathbf{\Omega}_{\ell^*}$ é uma matriz diagonal, consequentemente reduzindo o custo computacional do cálculo de sua inversa em (60).

Portanto, considerando as definições apresentadas e a descrição dada em (52), podemos escrever que

$$\boldsymbol{\psi}_{\ell^*+1}(m) = \text{softm}(\mathbf{z}_\ell, \boldsymbol{\psi}_{\ell^*}(m), \{w_\ell, \boldsymbol{\mu}_{\ell^*}, \mathbf{\Omega}_{\ell^*}\}), \quad (67)$$

em que $\boldsymbol{\psi}_Z = \mathbf{r}$, ou seja, o vetor de saída da última camada. Também percebe-se que temos

$$\tilde{\mathbf{q}}'_{\ell+1} = \sum_{j \neq \ell} \mathbf{e}_j z_\ell(j) + \mathbf{e}_\ell (\mathbf{s}^T \boldsymbol{\psi}_{\ell^*}), \quad (68)$$

de tal forma que o vetor realimentado no pré-processamento da camada seguinte seja dado pela ponderação das coordenadas dos símbolos L -QAM, pela estimação das probabilidades associadas a estes símbolos. Isso visa aprimorar as propriedades de convergência do DU-PDA.

Em outras palavras, pode-se entender então que o DU-PDA essencialmente estima as probabilidades de ocorrência de um determinado símbolo da constelação, levando em conta cada antena transmissora do sistema MIMO. Essas probabilidades são aprimoradas conforme os vetores são propagados pelas sucessivas camadas que compõem a arquitetura, juntamente também com parâmetros de aprendizagem. Por consequência, note que o número total de camadas do DU-PDA, Z , é múltiplo inteiro do número de antenas de transmissão, $2K$ (considerando (54)). Como mencionado anteriormente, isso acontece pois a arquitetura do DU-PDA é baseada principalmente no mapeamento das iterações do algoritmo PDA, para o qual probabilidades *a posteriori* são calculadas por meio da forma aproximada do teorema de Bayes [8]. O Algoritmo 3 detalha o funcionamento do DU-PDA.

As amostras de referência utilizadas pelo treinamento do DU-PDA são denotadas por $\mathbf{u}_{\ell^*} = [u(0) \ u(1) \ \dots \ u(\sqrt{L}-1)]^T$, tal que $\mathcal{U}_{\mathcal{L}} = \{\mathbf{u}_{\ell^*}\}_{\forall \ell^*}$. Essas amostras indicam previamente quais coordenadas da constelação são transmitidas durante o estágio de treinamento, de

Algoritmo 3: O detector MIMO DU-PDA.

Entrada: $T > 0$

Entrada: $\psi_{\ell^*}(m) \leftarrow \frac{1}{\sqrt{L}}, \forall m \forall \ell^*$

Entrada: $\tilde{\mathbf{q}}'_0 = \sum_{\ell^*} \mathbf{e}_{\ell^*} (\mathbf{s}^T \psi_{\ell^*})$

Entrada: $\mathcal{L}(\mathcal{I}, \boldsymbol{\psi})$ (veja (69))

```

1 Função CAMADA( $\tilde{\mathbf{y}}, \tilde{\mathbf{G}}, \boldsymbol{\psi}_{\ell^*-1}, \tilde{\mathbf{q}}'_{\ell-1}$ ):
2   Calcule (59) e (62), seguido por (60) e (68);
3   retorno  $\boldsymbol{\psi}_{\ell^*}, \tilde{\mathbf{q}}'_\ell$ 

4 Função TREINAMENTO( $\mathcal{L}(\mathcal{I}, \boldsymbol{\psi}), \boldsymbol{\psi}, \tilde{\mathbf{q}}'_0, T$ ):
5   for  $\forall$  ÉPOCAS do
6     Gere o conjunto de amostras de treinamento;;
7      $\mathcal{S}_T = \{(\tilde{\mathbf{y}}^{(1)}, \mathbf{I}^{(1)}), \dots, (\tilde{\mathbf{y}}^{(T)}, \mathbf{I}^{(T)})\}$ ;
8     for  $\ell = 1, 2, \dots, Z + 1$  do
9       Treinar: CAMADA( $\tilde{\mathbf{y}}^{(1,2,\dots,T)}, \tilde{\mathbf{G}}^{(1,2,\dots,T)}, \boldsymbol{\psi}_{\ell^*-1}, \tilde{\mathbf{q}}'_{\ell-1}$ );
10    end
11  end
12  retorno  $\boldsymbol{\psi}_{\ell^*}, \tilde{\mathbf{q}}'_\ell$ 

13 Função DETECÇÃO( $\tilde{\mathbf{y}}, \tilde{\mathbf{G}}$ ):
14   Executar forward-pass: CAMADA( $\tilde{\mathbf{y}}, \tilde{\mathbf{G}}, \boldsymbol{\psi}_{\ell^*-1}$ ),  $\forall \ell$ ;
15    $d_{\ell^*} \leftarrow \arg \max_m \{\psi_{\ell^*}(m)\}, \forall \ell^*$ ;
16  retorno  $d_{\ell^*}$ 
  
```

Saída : $\tilde{\mathbf{q}}^{(\ell^*)'} \leftarrow s_{d_{\ell^*}}, \forall \ell^*$

forma que $u_{\ell^*}(m) \in \{0, 1\} \forall m$. Desta forma, pode-se definir a função de perda empregada pelo DU-PDA como:

$$\mathcal{L}(\mathcal{U}_{\mathcal{L}}, \boldsymbol{\psi}) = \frac{-1}{Z\sqrt{L}} \sum_{\ell^*} \sum_m u_{\ell^*}(m) \log(\psi_{\ell^*}(m)). \quad (69)$$

Denominada de entropia cruzada média, essa função permite aferir qual é a distância, isto é, o erro, entre a estimação do DU-PDA e as amostras de referência. Basicamente isso indica para o algoritmo de treinamento, ou o otimizador dos parâmetros de aprendizado, quais são os melhores valores à serem definidos para maximizar o desempenho das redes neurais [28, 27].

Em seguida, será analisado o desempenho do DU-PDA em termos da BER. Considera-se que $n_b = 256$ bits de informação são codificados na transmissão usando o codificador polar [31], cuja taxa de código é dada por $R < 1$. Os elementos da matriz do canal são dados por $G_{i,j} \sim \mathcal{CN}(0, 1/M)$, $\forall i, j$, isto é, as amostras do canal são extraídas de um processo *Gaussiano* complexo, e a SNR por bit é descrita por

$$\Gamma = \left(\sqrt{LR}\right)^{-1} \frac{\mathbb{E}[\|\mathbf{G}\mathbf{q}\|_2^2]}{M\sigma^2}. \quad (70)$$

Além disso, detalhes acerca dos hiperparâmetros do DU-PDA podem ser verificados na Tabela 2.

Tabela 2: Hiperparâmetros da arquitetura DU-PDA.

Hiperparâmetros	Valores
Núm. de amostras de treinamento	10^5 amostras
Camadas	$L = 4K$
Dimensão da entrada	$\mathbb{R}^{2M}, \mathbb{R}^{2M \times 2K}, \mathbb{R}^{2K \times \sqrt{L}}, \mathbb{R}^{2K}$
Dimensão da saída	$\mathbb{R}^{2K \times \sqrt{L}}$
Núm. de parâmetros de aprendizagem	$\#\{w_\ell\}_{\forall \ell} = 4K$
Função de ativação	$\text{softmax}(\cdot), \forall \ell$
Taxa de aprendizado	10^{-3}
Otimizador	Adam

Na Figura 8, a BER do DU-PDA é ilustrada para uma faixa de valores de SNR, e comparado

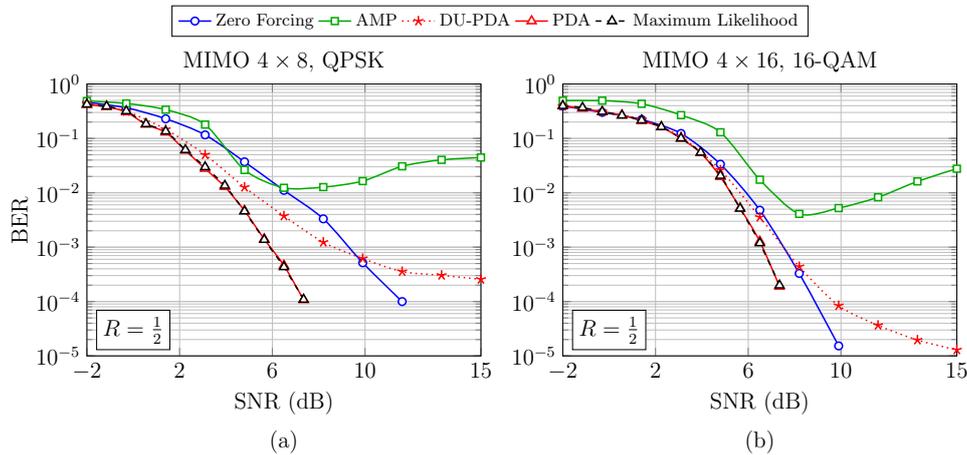


Figura 8: Desempenho dos detectores por máxima verossimilhança (5), DU-PDA, PDA, ZF (11) e AMP (47), considerando a probabilidade de erro de bit para uma faixa de valores de SNR.

com os detectores de máxima verossimilhança (5), PDA [30], ZF (11) e AMP (47). Também observe que os sistemas (a) MIMO 4×8 (QPSK) e (b) MIMO 4×16 (16-QAM) são levados em conta, ambos com a taxa de código configurado para $R = 1/2$.

Primeiramente, note nas Figuras 8 (a) e (b) que o detector ótimo de máxima verossimilhança obtém o melhor desempenho, como esperado. Contudo, o algoritmo PDA também apresenta valores de BER similares ao detector ótimo, o que demonstra sua adequação como algoritmo subjacente à arquitetura do DU-PDA. Embora o DU-PDA apresente um *floor* de desempenho para $\text{SNR} > 10$ dB, para a faixa $\text{SNR} < 10$ dB o seu desempenho é melhor do que aquele apresentado pelos detectores ZF e AMP, considerando ambas as Figuras 8 (a) e (b). Essa faixa de SNR é mais significativa para análises práticas de desempenho, pois a intensidade do ruído AWGN é ainda considerável. Além disso, ressalta-se que o DU-PDA não calcula inversas de matrizes, reduzindo significativamente sua complexidade perante o ZF, por exemplo. Por fim, observe que o *floor* de desempenho citado anteriormente, advém de um comportamento típico do detector AMP, como demonstrado na Subseção 3.4, mas é consideravelmente suprimido pelo DU-PDA.

4 Implementação de MIMO no transceptor Brasil 6G

Juliano Silveira Ferreira
silveira@inatel.br

O transceptor Brasil 6G é empregado como interface aérea da *Radio Access Network* (RAN) da rede proposta neste projeto. A Figura 9 mostra o diagrama simplificado desta RAN. A *Base Station* (BS) é o terminal responsável por prover acesso à internet e à outros servidores, além de definir os parâmetros de operação dos terminais configurados como *User Equipments* (UEs).

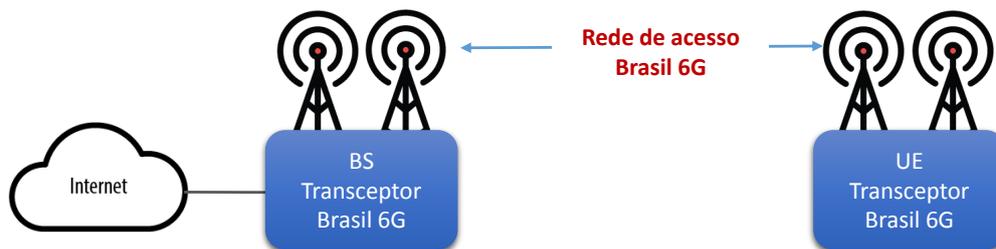


Figura 9: Diagrama simplificado da rede de acesso Brasil 6G.

Os transceptores empregam diversidade espacial de transmissão para modulação *Generalized Frequency Division Multiplexing* (GFDM) desenvolvida om base no algoritmo proposto em [32]. Além da diversidade de transmissão, os transceptores também podem operar com MIMO voltado para multiplexação espacial. Devido a complexidade de implementação, como latência e demanda de recursos computacionais, os detectores MIMO implementado foram o ZF e MMSE empregando *Software Defined Radio* (SDR) com o GNU Radio como plataforma para implementação [33]. O *hardware* previsto para a versão do transceptor Brasil 6G de alta capacidade consiste em *front-ends* de RF com 2 antenas de transmissão e 2 antenas de recepção, como mostrado em [34]. Logo, o transceptor emprega MIMO 2×2 .

A seguir estão apresentados maiores detalhes relacionados aos desafios de implementação prática de MIMO com multiplexação espacial no transceptor Brasil 6G, bem como os resultados alcançados.

4.1 Desafios da implementação

A implementação prática de sistemas MIMO no transceptor Brasil 6G demandam a superação de diferentes desafios dentre os quais pode-se destacar:

- Recursos de processamento: a implementação envolve a realização de operações sobre vetores e matrizes com grande quantidade de dados, o que impacta diretamente nos requisitos de recursos de processamento, assim como de compartilhamento e armazenamento de dados. A estrutura de quadro empregada pelo transceptor considera a transmissão simultânea de até 12.672 portadoras ativas para cada símbolo GFDM [35]. Vale ressaltar ainda que cada portadora é modulada por 4 subsímbolos de dados para a configuração em questão. Desta forma, deve-se observar a complexidade de implementação dos algoritmos envolvidos, pois, estes impactam na solução de *hardware* adotada.

- Eficiência de processamento: o desenvolvimento de soluções para operação em tempo real deve considerar a implementação de algoritmos de forma eficiente, observando o tempo demandado para a execução de cada bloco pelo processador e no impacto na vazão e latência.
- Operação em condições desfavoráveis: a implementação prática dos algoritmos MIMO devem considerar sua operação em condições reais, nas quais ocorrem condições desfavoráveis como a incidência de ruídos e interferências indesejadas, além da própria composição ou combinação dos canais de comunicação. Estes agregam complexidades e desafios à operação dos algoritmos de detecção.

Devido aos desafios acima citados, foi proposta a implementação inicial de um algoritmo MIMO para multiplexação espacial de reduzida complexidade ao transceptor Brasil 6G, com intuito de minimizar o impacto nos recursos de *hardware*, bem como na latência de operação do sistema. Desta forma, foram eleitos para a implementação os algoritmos ZF e MMSE. O condicionamento do canal é avaliado para determinar se a multiplexação espacial pode ser utilizada. Caso o posto da matriz de canal seja muito baixo, a multiplexação espacial não é empregada. As seções seguintes apresentam maiores detalhes sobre a implementação de tais algoritmos.

4.2 Implementação de detectores MIMO com multiplexação espacial

Conforme mencionado anteriormente, a implementação de técnica MIMO com multiplexação espacial no transceptor Brasil 6G considera, inicialmente, o emprego dos detectores ZF e MMSE. Os algoritmos foram desenvolvidos utilizando-se a linguagem de programação C++ visando otimizar o uso de recursos de *hardware* e reduzir a latência de execução. Os blocos foram implementados empregando os recursos de processamento paralelo, baseados em *Single Instruction Multiple Data* (SIMD). Vale destacar que os detectores implementados empregam a resposta do canal estimado e interpolado para a largura de banda alocada pela BS para o terminal UE em teste.

4.3 Análise do condicionamento da matriz de canal

Para que os detectores ZF e MMSE funcionem adequadamente, a matriz de canal deve apresentar linhas e colunas com baixa correlação, ou seja, o posto da matriz deve ser elevado. Um canal de comunicação cuja matriz é mal condicionada resulta em detectores lineares com alto nível de amplificação de ruído e sensível aos erros de aproximações nas estimativas do canal. Uma matriz de canal bem condicionada, por outro lado, viabiliza a correta detecção dos múltiplos sinais.

O condicionamento da matriz de \mathbf{G} é dada por [36]

$$\Phi = \frac{v_{\max}}{v_{\min}} \geq 1 \quad (71)$$

onde v_{\max} e v_{\min} são o maior e menor valor singular de \mathbf{G} , respectivamente. Quanto mais próximos de 1 for o valor do condicionamento do canal, mais propício será o canal para a transmissão com multiplexação espacial. Caso exista uma grande diferença entre os valores singulares, maior o valor do condicionamento do canal menos propício será o canal para tal transmissão.

O condicionamento do canal de comunicação afeta a operação de certas técnicas MIMO e que impacta na capacidade do canal de comunicação. Em [36] é considerado, por exemplo, um sistema com eficiência espectral de 10 bps/Hz e que requer SNR de 15 dB para operação com canal de comunicação com condicionamento igual a 1. Para o mesmo sistema operando em um canal com condicionamento igual a 40B, seria necessária uma SNR de 20 dB para se obter o mesmo desempenho. Ainda em [36] é destacado que, para operação prática, valores de condicionamento de canal de até 10 são considerados como adequados para a exploração de técnicas MIMO com multiplexação espacial. Canais com condicionamento acima de 100, entretanto, não são recomendados para o uso de tais técnicas.

Com base nestas informações citadas, foi implementado no transceptor Brasil 6G o algoritmo para cálculo em tempo real do condicionamento do canal de comunicação. A rede de acesso Brasil 6G permite que seja definido pela BS qual será a técnica MIMO a ser empregada em cada enlace de comunicação: diversidade ou multiplexação. A métrica do condicionamento do canal foi calculada para servir de base para tal definição. O algoritmo desenvolvido utiliza as condições do canal estimado pelas pilotos. No caso da numerologia adotada para os testes iniciais do transceptor, na numerologia 0 [35] são consideradas 24 portadoras pilotos a cada 180 KHz de largura de banda. O valor final do condicionamento do canal é calculado a partir da média dos valores de condicionamento para toda a banda alocada para o terminal UE em questão. O valor calculado é inserido nas informações de controle da UE, sendo estas transmitidas para a BS.

4.4 Ajustes na definição dos parâmetros do enlace de comunicação

Uma vez recebidas as informações de controle pela BS, estas são utilizadas nas tomadas de decisões sobre o enlace de comunicação com as UE. A definição do MCS de operação, tanto de *uplink* quanto de *downlink*, é realizada de maneira dinâmica pela BS em função das últimas condições do canal de comunicação. O MCS a ser utilizado em um quadro é enviado para as UEs usando a camada de controle da rede. A Figura 10 ilustra o envio das informações de condição de recepção da UE para a BS e o envio da sinalização pela BS do *Modulation Coding Scheme* (MCS) a ser empregado pela UE.

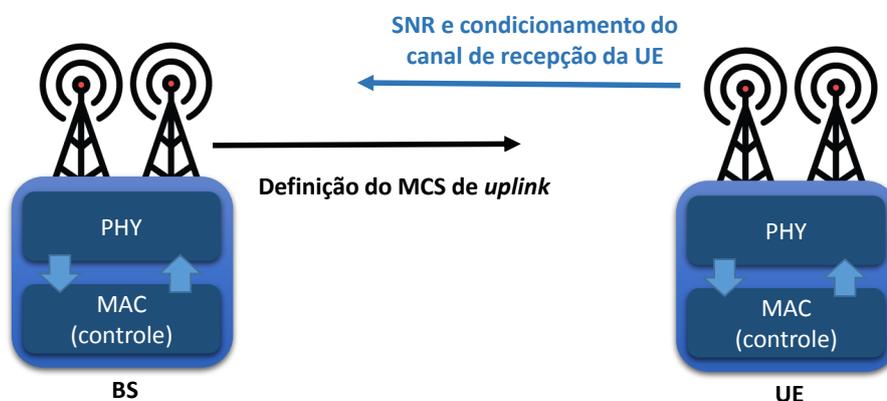


Figura 10: Ilustração da troca de informações do canal entre BS e UE.

A partir das informações do condicionamento do canal, citadas na Seção 4.3 deste documento, foram considerados ajustes na definição de parâmetros do enlace de comunicação. Mais

especificamente, foram realizados ajustes na definição do MCS empregado nos enlaces de *down-link* e *uplink* estabelecidos com a UE que opera com multiplexação espacial. O ajuste, no caso, se refere a adoção de um fator dependente do condicionamento do canal que impacta na definição do MCS do enlace de comunicação. Como resultado do uso deste fator, tem-se a definição de um MCS mais robusto quando comparado com o MCS definido para operação com diversidade espacial dada a mesma SNR. O fator de ajuste adotado, definido com base em [36], deve ser melhor avaliado e ajustado em função dos resultados de testes práticos. O MCS é definido em função da SNR do enlace. O ajuste implementado consiste em reduzir o valor de SNR a ser empregado na definição do MCS. Inicialmente foi adotada um fator de redução de 3 e 6 dB no valor de SNR para os casos onde o condicionamento do canal medido for inferior a 10 e 100, respectivamente. Para os casos de condicionamento de canal superior a 100, considera-se inviável o uso de multiplexação espacial.

4.5 Resultados de testes práticos

Testes práticos para avaliação de desempenho dos detectores foram realizados e a Figura 11 mostra a foto dos equipamentos utilizados para a geração e recepção dos sinais, onde é possível observar os front-end de RF utilizados pela BS e UE.

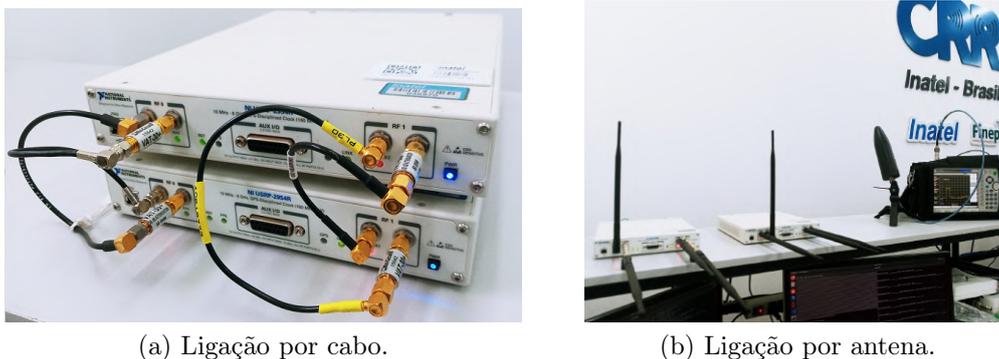


Figura 11: *Set-ups* de testes.

A Tabela 3 mostra o resultado do tempo de processamento demandado para a execução dos algoritmos de decodificação ou detecção MIMO. Para efeitos de comparação, foi realizada a medida também para a técnica MIMO com diversidade espacial. Pode-se verificar que os detectores para multiplexação espacial demandam mais tempo de processamento do que o detector para diversidade espacial. O detector MMSE, por sua vez, requereu o maior tempo de processamento, uma vez que o algoritmo calcula o determinante da matriz do canal para cada subportadora. Este último deve passar por otimizações visando a redução do tempo envolvido. Vale destacar que nenhum dos algoritmos em teste impactou na latência total do sistema considerando a realização do teste de através do comando *ping* entre BS e UE.

As Figuras 12 e 13 mostram os resultados da vazão de dados a partir do emprego das técnicas de MIMO com diversidade espacial e multiplexação espacial, respectivamente. Nas figuras aparecem o terminal do transceptor Brasil 6G configurado para operar como UE. Na parte superior das figuras aparecem status de operação e das condições do canal, como SNR e condicionamento do canal. Na parte inferior da figura aparece o resultado do teste de vazão realizado empregando-se a ferramenta de redes denominada de *iperf*. Em destaque estão a

Tabela 3: Tempo de processamento dos algoritmos.

Técnica MIMO	Tempo de processamento*
Diversidade espacial	$\sim 2 \mu s$
Multiplexação espacial Detector ZF	$\sim 6 \mu s$
Multiplexação espacial Detector MMSE	$\sim 46 \mu s$

*Pacote *ping* com comprimento de 1.000 bytes

configuração de MIMO empregada e a vazão de dados resultantes do teste com a ferramenta citada.

Os dois testes mostrados nas Figuras 12 e 13 consideram a vazão de dados úteis no enlace de *uplink*, empregando 6 MHz de largura de banda para a UE em teste. Conforme pode-se verificar na Figura 12, com a operação com diversidade espacial foi alcançada a taxa útil da ordem de 22 Mbps, enquanto para configuração com multiplexação espacial, mostrada na Figura 13, a taxa alcançada foi de 44 Mbps. Desta forma, a taxa do enlace foi 2 vezes maior com o emprego da técnica MIMO com multiplexação espacial, sem demandar o aumento da largura de banda ocupada para o estabelecimento do enlace de comunicação. Vale ressaltar que este resultado foi alcançado no *set-up* via cabo, que permitiu alcançar um canal com condicionamento unitário, sendo esta uma condição muito favorável para o emprego de MIMO. Pode-se notar a partir do status de operação do transceptor, mostrado na parte superior das duas figuras, que o emprego da técnica de multiplexação espacial impactou na redução da SNR de recepção estimada pela UE em aproximadamente de 3 dB, em relação ao valor estimado com o emprego da técnica de diversidade espacial, uma vez que a energia é dividida entre dois símbolos distintos no sistema MIMO com multiplexação espacial.

Para as condições de teste considerando a transmissão via antenas, é possível observar o aumento da vazão de dados com o emprego da multiplexação espacial, entretanto, a vazão alcançada e até mesmo o estabelecimento de um enlace de comunicação é altamente dependente do posicionamento das antenas, que altera drasticamente o condicionamento do canal. Conforme comentado na Seção 4.4 deste documento, foi considerado um ajuste para a definição do MCS a ser empregado no enlace de comunicação com multiplexação espacial em função do condicionamento do canal. Este ajuste impacta na redução da vazão de dados alcançada, porém, busca priorizar a robustez do enlace.

Pode-se destacar assim, que uma versão funcional dos detectores MIMO com multiplexação espacial foi desenvolvida e integrada ao transceptor Brasil 6G. Estes devem passar por otimizações e validações complementares a fim de se alcançar maior redução do tempo de processamento envolvido, bem como de ajustes para maior robustez de operação na condição de comunicação via antena. Vale ressaltar que os algoritmos selecionados em função de sua baixa complexidade de implementação apresentam desempenho reduzido quanto comparado com técnicas mais complexas, conforme demonstra os resultados apresentados na Seção 3.1. Como possibilidades de implementação futura, cita-se o desenvolvimento de algoritmos de técnicas mais avançadas e de maior desempenho.

```
[Brasil 6G][BR6G Transceptor][PHY RX STATUS] =====
[UE - User Equipment] 1
Numerology : 0
MIMO : DIVERSITY
Mod : 6
Code rate : 0,877976 (0,877976)
[PHY RX] MSE : -41,902
RX Data Throughput (Mbps): 0.970369
RX SNR : 41.902 [dB]
RX Channel Condition Number: 0 [dB]

[ 4] 20.00-21.00 sec 2.64 MBytes 22.1 Mbits/sec 3 2.16 MBytes
[ 4] 21.00-22.00 sec 2.62 MBytes 22.0 Mbits/sec 3 2.16 MBytes
[ 4] 22.00-23.00 sec 2.54 MBytes 21.3 Mbits/sec 5 2.16 MBytes
[ 4] 23.00-24.00 sec 2.61 MBytes 21.9 Mbits/sec 3 2.16 MBytes
[ 4] 24.00-25.00 sec 2.67 MBytes 22.4 Mbits/sec 0 2.16 MBytes
[ 4] 25.00-26.00 sec 2.56 MBytes 21.5 Mbits/sec 8 2.16 MBytes
[ 4] 26.00-27.00 sec 2.61 MBytes 21.9 Mbits/sec 0 2.16 MBytes
[ 4] 27.00-28.00 sec 2.61 MBytes 21.9 Mbits/sec 0 2.17 MBytes
[ 4] 28.00-29.00 sec 2.62 MBytes 22.0 Mbits/sec 0 2.17 MBytes
[ 4] 29.00-30.00 sec 2.65 MBytes 22.2 Mbits/sec 0 2.19 MBytes

[ ID] Interval      Transfer      Bandwidth      Retr
[ 4] 0.00-30.00 sec 80.8 MBytes 22.6 Mbits/sec 61      sender
[ 4] 0.00-30.00 sec 79.9 MBytes 22.3 Mbits/sec      receiver

iperf Done.
root@modem:~# iperf3 -c 10.0.0.10 -t 30
modem:~#
```

Figura 12: MIMO com diversidade espacial: taxa de dados alcançada pela UE.

```
[Brasil 6G][BR6G Transceptor][PHY RX STATUS] =====
[UE - User Equipment] 1
Numerology : 0
MIMO : MULTIPLEXING
Precode : 0
Mod : 6
Code rate : 0,420238 (0,420238)
[PHY RX] MSE : -39,2518
RX Data Throughput (Mbps): 0.700812
RX SNR : 39.2518 [dB]
RX Channel Condition Number: 0 [dB]

[ 4] 20.00-21.00 sec 5.54 MBytes 46.5 Mbits/sec 0 2.47 MBytes
[ 4] 21.00-22.00 sec 5.48 MBytes 45.9 Mbits/sec 0 2.49 MBytes
[ 4] 22.00-23.00 sec 5.45 MBytes 45.7 Mbits/sec 0 2.51 MBytes
[ 4] 23.00-24.00 sec 5.53 MBytes 46.4 Mbits/sec 0 2.55 MBytes
[ 4] 24.00-25.00 sec 5.51 MBytes 46.2 Mbits/sec 0 2.60 MBytes
[ 4] 25.00-26.00 sec 5.53 MBytes 46.3 Mbits/sec 20 2.64 MBytes
[ 4] 26.00-27.00 sec 5.37 MBytes 45.2 Mbits/sec 0 2.65 MBytes
[ 4] 27.00-28.00 sec 5.56 MBytes 46.7 Mbits/sec 2 2.65 MBytes
[ 4] 28.00-29.00 sec 5.46 MBytes 45.7 Mbits/sec 0 2.65 MBytes
[ 4] 29.00-30.00 sec 5.49 MBytes 46.0 Mbits/sec 0 2.65 MBytes

[ ID] Interval      Transfer      Bandwidth      Retr
[ 4] 0.00-30.00 sec 160 MBytes 44.8 Mbits/sec 53      sender
[ 4] 0.00-30.00 sec 159 MBytes 44.6 Mbits/sec      receiver

iperf Done.
root@modem:~# iperf3 -c 10.0.0.10 -t 30
modem:~#
```

Figura 13: MIMO com multiplexação espacial: taxa de dados alcançada pela UE.

5 Avaliações complementares envolvendo simulações computacionais com condições reais do canal de comunicação

Davi da Silva Brilhante, Juliano Silveira Ferreira
 dbrilhante@land.ufrj.br, silveira@inatel.br

Uma vez que as simulações dos detectores MIMO foram desenvolvidas, foi possível criar um *set-up* para a avaliação dos mesmos, porém, considerando condições mais realistas. Para tal, foi desenvolvida uma funcionalidade de captura e armazenamento de sinais na camada física do transceptor Brasil 6G. Esta funcionalidade viabiliza que informações obtidas a partir da operação real do transceptor possam ser utilizadas em pesquisas e avaliações de novos algoritmos ou técnicas que possam ser futuramente integradas ao transceptor e à rede de acesso Brasil 6G. Este tipo de integração permite agregar às simulações computacionais informações inerentes à implementação prática como, por exemplo, arredondamentos e erros do processo de conversão analógico-digital e digital-analógico, estimação e equalização de canal e, para o caso do enlace de RF ser estabelecido via antena, permite agregar interferências, ruídos, influência da polarização das antenas, dentre outros.

A funcionalidade desenvolvida foi utilizada para a captura da resposta do canal de comunicação e esta foi utilizada para a realização de avaliações iniciais de algumas técnicas MIMO. Em uma primeira etapa foram consideradas os detectores MIMO de menor complexidade, visando validar o processo de integração proposto. A seguir estão apresentados detalhes do *set-up* empregado para a captura das condições do canal pelo transceptor, bem como os resultados de simulações dos detectores de menor complexidade considerando tais canais.

5.1 *Set-up* de captura das condições reais do canal

A realização da captura dos canais de comunicação pelo transceptor Brasil 6G considerou dois *set-ups* distintos, de acordo com o enlace estabelecido entre BS e UE: via cabo de RF e via antena. As configurações principais consideradas em tais *set-ups* são:

- Captura realizada pelo terminal UE (enlace de *downlink*);
- Frequência empregada no enlace: 728 MHz;
- Largura de banda do enlace: 6 MHz;
- Total de portadoras na banda: 3168;
- Arranjo MIMO: 2×2 .

As condições principais do *set-up* para a captura da resposta do canal considerando a comunicação entre BS e UE estabelecida via cabo de RF estão apresentadas na Tabela 4. O *set-up* empregado, neste caso, é o mesmo mostrado na parte (a) da Figura 11.

As condições principais do *set-up* via antena, por sua vez, estão apresentadas na Tabela 5. A Figura 14 mostra o *set-up* em questão. O modelo da antena empregado nos testes é o ANT-TRI, do fabricante Nuande.

Tabela 4: Detalhes do *set-up* de captura considerando comunicação via cabo de RF.

Potência de transmissão	-10 dBm
Comprimento do cabo	22 cm
Atenuador na recepção	30 dB

Tabela 5: Detalhes do *set-up* de captura considerando comunicação via antenas.

Potência de transmissão	-30 dBm
Ganho das antenas	5 dBi
Polarização das antenas	Cruzadas
Distância mínima entre antenas	0,86 m

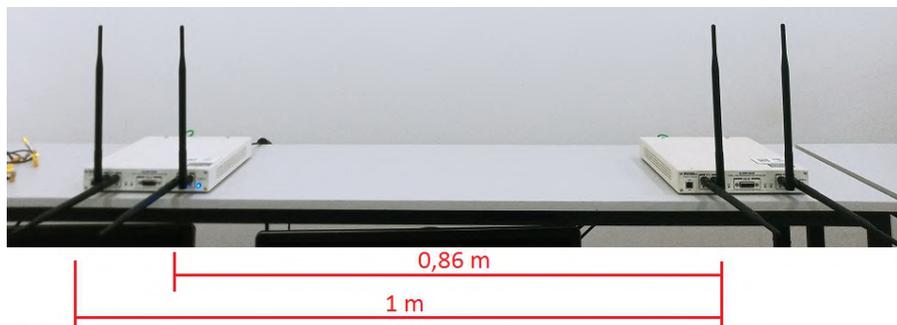


Figura 14: *Set-up* empregado para a captura da resposta do canal com enlace via antena.

5.2 Resultado inicial das simulações considerando os canais reais

Com objetivo de complementar a avaliação de desempenho de algumas das técnicas MIMO considerando condições mais realistas, foram executadas simulações utilizando a resposta do canal capturada pelo transceptor Brasil 6G. Apesar de não retratar com total exatidão a execução dos algoritmos em teste no transceptor, esse método de experimentação propicia a avaliação das técnicas de detecção com respeito à BER sem perda de generalização.

Para a execução da simulação, os canais capturados pelo transceptor foram salvos em arquivos e estes foram lidos por um código *Python*. Após esta leitura, as amostras de canal são multiplicadas pelos símbolos gerados por uma sequência aleatória de 1024 *bits* e empregou-se uma modulação de amplitude em quadratura, 4, 16, ou 64 QAM. Em seguida, uma das técnicas de detecção MIMO é aplicada e a demodulação do símbolo detectado é então realizada. Ao fim desta última operação, comparam-se os *bits* resultantes com a sequência original para a estimativa da BER.

As técnicas de detecção simuladas nesta fase foram: MF, ML, ZF-SIC, e *Linear Minimum Mean Square Error* (LMMSE)-SIC. O gráfico da Figura 15 mostra o resultado da simulação considerando a resposta do canal capturado na condição de enlace via cabo de RF entre a BS e UE, conforme *set-up* descrito anteriormente. Neste caso nota-se que o desempenho das técnicas avaliadas estão muito próximas do detector ótimo, o ML, para a modulação 4QAM. Contudo, ao aumentar a ordem da modulação, o detector MF piora o seu desempenho, devido a simplicidade do algoritmo de detecção, como abordado na Seção 2. Por outro lado, o MF é o detector que oferece a menor complexidade [8]. Para os canais testados, ZF-SIC e LMMSE-SIC, oferecem desempenho praticamente igual ao do detector ML, mantendo a robustez ao aumento da ordem de modulação. É importante salientar que o canal é estático.

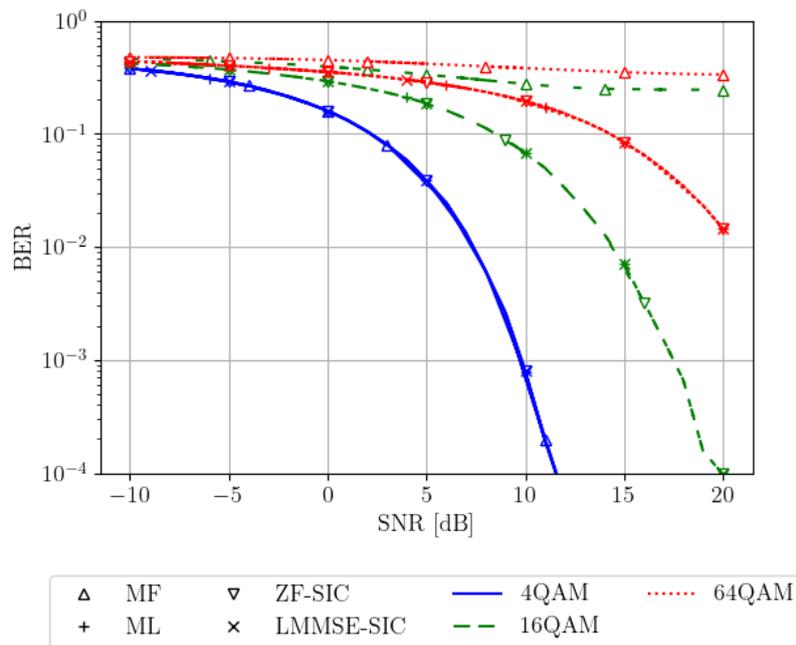


Figura 15: Comparação da BER dos detectores MF, ML, LMMSE-SC e ZF-SIC(V-BLAST) para as modulações 4QAM (azul), 16QAM (verde) e 64QAM (vermelho), com resposta do canal real capturada a partir do transceptor com enlace via cabo de RF.

A Figura 15 apresenta, por sua vez, o resultado da simulação considerando a resposta do canal capturado na condição de enlace via antena entre a BS e UE, conforme *set-up* descrito anteriormente. Neste caso nota-se que o desempenho das técnicas avaliadas sofreram uma piora em relação à simulação anterior. Isto se deve ao fato, principalmente, das condições do canal de comunicação proporcionado pelas condições do *set-up*, bem como pela correlação entre as respostas dos canais existente nesta condição.

Considera-se que em trabalhos futuros, novas campanhas de captura das condições do canal possam ser realizadas variando parâmetros do cenário de captura como polarização das antenas, distância entre transmissor e receptor e até mesmo condições de operação em campo. Além disso, podem ser melhor avaliadas as técnicas de detecção MIMO mais complexas, antes de sua implementação na camada física do transceptor Brasil 6G.

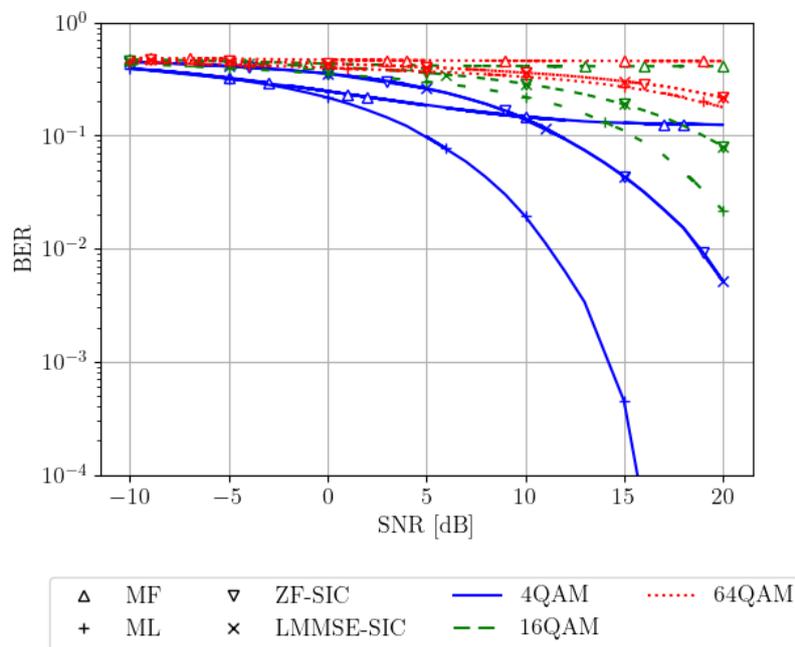


Figura 16: Comparação da BER dos detectores MF, ML, LMMSE-SC e ZF-SIC(V-BLAST) para as modulações 4QAM (azul), 16QAM (verde) e 64QAM (vermelho), com resposta do canal real capturada a partir do transceptor com enlace via antenas.

6 Conclusão

O projeto Brasil 6G tem avançado a execução de suas atividades visando conceber e implementar uma plataforma de testes e avaliação de diferentes tecnologias candidatas e potenciais para integrar a rede 6G. Dentre as tecnologias candidatas pode-se destacar aquelas associadas a MIMO. Estas técnicas podem ser consideradas para o adequado suporte à futuras aplicações que requeiram tanto a alta confiabilidade na comunicação, quanto alta vazão de dados. O presente relatório abordou técnicas que visam proporcionar o aumento da vazão de dados a partir de técnicas de multiplexação espacial.

As tecnologias MIMO têm sido alvo de pesquisas e diversos avanços têm sido alcançados a partir do Sphere Detection, LRA, *Approximate Message-Passing* e *Deep Unfolded MIMO*. Destaca-se nesta última, o emprego de redes neurais para a detecção MIMO.

Uma versão funcional de equalizadores lineares para multiplexação espacial foi desenvolvida e integrada ao transceptor Brasil 6G. Testes iniciais foram realizados, conforme apresentado no presente relatório, onde pode-se observar o impacto na vazão de dados em função do número de antenas utilizado. Destaca-se, entretanto, que o desempenho destes detectores é afetado pelas condições de canal e demandam condições específicas e favoráveis para sua operação.

Além disso, foi desenvolvida uma metodologia para a avaliação dos detectores MIMO considerando condições mais realistas de operação. A metodologia definida considera o uso das informações do canal de comunicação obtidas a partir da operação real do transceptor Brasil 6G e sua utilização na execução das simulações computacionais. Os detectores MIMO de menor complexidade foram avaliados, demonstrando que a metodologia criada pode ser bastante versátil e útil. Além disso, a metodologia pode ser adotada para o estudo e análises de diferentes algoritmos e técnicas, incluindo detectores MIMO de maior complexidade, que possam vir a compor a rede de acesso Brasil 6G.

Sendo assim, o presente relatório apresentou uma introdução à diferentes técnicas e algoritmos MIMO para multiplexação espacial e apresentou uma visão dos avanços principais relacionados à implementação de MIMO no transceptor Brasil 6G. O embasamento e análise de desempenho das técnicas foi devidamente apresentado no relatório, assim como uma metodologia para a avaliação dos detectores considerando condições de operação mais realistas. Desta forma, espera-se que o presente relatório possa servir como um material de referência inicial para futuras pesquisas e implementações considerando técnicas MIMO com multiplexação espacial.

Referências

- [1] J. Mietzner, R. Schober, L. Lampe, W. H. Gerstacker, e P. A. Hoeher, “Multiple-antenna techniques for wireless communications - a comprehensive literature survey,” *Commun. Surveys Tuts.*, v. 11, n. 2, p. 87–105, Apr. 2009. [Online]. Disponível em: <http://dx.doi.org/10.1109/SURV.2009.090207>
- [2] J. R. Hampton, *Introduction to MIMO Communications*. New York, NY, USA: Cambridge University Press, 2014.
- [3] J. Jalden, “Maximum likelihood detection for the linear mimo channel,” Ph.D. dissertation, Royal Institute of Technology (KTH), 2004.
- [4] M. A. Albreem, M. Juntti, e S. Shahabuddin, “Massive mimo detection techniques: A survey,” *IEEE Communications Surveys & Tutorials*, v. 21, n. 4, p. 3109–3132, 2019.
- [5] C. D. Altamirano, J. Minango, H. C. Mora, e C. De Almeida, “Ber evaluation of linear detectors in massive mimo systems under imperfect channel estimation effects,” *IEEE access*, v. 7, p. 174 482–174 494, 2019.
- [6] R. T. Kobayashi, F. Ciriaco, e T. Abrão, “Performance and complexity analysis of sub-optimum MIMO detectors under correlated channel,” in *2014 International Telecommunications Symposium (ITS)*. IEEE, 2014, p. 1–5.
- [7] D. Wübben, R. Böhnke, V. Kühn, e K.-D. Kammeyer, “MMSE extension of V-BLAST based on sorted QR decomposition,” in *Proceedings of the IEEE 58th Vehicular Technology Conference, 2003 (VTC’03-Fall)*, v. 1, 2003, p. 6.
- [8] S. Yang, T. Lv, R. G. Maunder, e L. Hanzo, “From nominal to true a posteriori probabilities: An exact bayesian theorem based probabilistic data association approach for iterative MIMO detection and decoding,” *IEEE Transactions on Communications*, v. 61, n. 7, p. 2782–2793, 2013.
- [9] B. Hassibi, “An efficient square-root algorithm for BLAST,” in *2000 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No. 00CH37100)*, v. 2. IEEE, 2000, p. II737–II740.
- [10] D. Wübben, J. Rinas, R. Böhnke, V. Kühn, e K. Kammeyer, “Efficient algorithm for detecting layered space-time codes,” *order*, v. 1, p. 1, 2002.
- [11] A. Benjebbour, H. Murata, e S. Yoshida, “Comparison of ordered successive receivers for space-time transmission,” in *IEEE 54th Vehicular Technology Conference. VTC Fall 2001. Proceedings (Cat. No. 01CH37211)*, v. 4. IEEE, 2001, p. 2053–2057.
- [12] Zanella, Alberto and Chiani, Marco and Win, Moe Z, “MMSE reception and successive interference cancellation for MIMO systems with high spectral efficiency,” *IEEE Transactions on wireless Communications*, v. 4, n. 3, p. 1244–1253, 2005.
- [13] U. Fincke e M. Pohst, “Improved methods for calculating vectors of short length in a lattice, including a complexity analysis,” *Mathematics of computation*, v. 44, n. 170, p. 463–471, 1985.

- [14] O. Damen, A. Chkeif, e J.-C. Belfiore, “Lattice code decoder for space-time codes,” *IEEE Communications letters*, v. 4, n. 5, p. 161–163, 2000.
- [15] M. O. Damen, H. El Gamal, e G. Caire, “On maximum-likelihood detection and the search for the closest lattice point,” *IEEE Transactions on information theory*, v. 49, n. 10, p. 2389–2402, 2003.
- [16] B. Hassibi e H. Vikalo, “On the sphere-decoding algorithm i. expected complexity,” *IEEE transactions on signal processing*, v. 53, n. 8, p. 2806–2818, 2005.
- [17] J. Jaldén e B. Ottersten, “On the complexity of sphere decoding in digital communications,” *IEEE transactions on signal processing*, v. 53, n. 4, p. 1474–1484, 2005.
- [18] C.-Y. Hung e T.-H. Sang, “A sphere decoding algorithm for mimo channels,” in *2006 IEEE International Symposium on Signal Processing and Information Technology*. IEEE, 2006, p. 502–506.
- [19] M. Grötschel, L. Lovász, e A. Schrijver, *Geometric algorithms and combinatorial optimization*. Springer Science & Business Media, 2012, v. 2.
- [20] H. Yao e G. W. Wornell, “Lattice-reduction-aided detectors for mimo communication systems,” in *Global Telecommunications Conference, 2002. GLOBECOM’02. IEEE*, v. 1. IEEE, 2002, p. 424–428.
- [21] C. Ling, “On the proximity factors of lattice reduction-aided decoding,” *IEEE Transactions on Signal Processing*, v. 59, n. 6, p. 2795–2808, 2011.
- [22] L. Lovász, *An algorithmic theory of numbers, graphs and convexity*. SIAM, 1986.
- [23] Y. H. Gan e W. H. Mow, “Complex lattice reduction algorithms for low-complexity mimo detection,” in *GLOBECOM ’05. IEEE Global Telecommunications Conference, 2005.*, v. 5, 2005, p. 5 pp.–2957.
- [24] K. Zu e R. C. De Lamare, “Low-complexity lattice reduction-aided regularized block diagonalization for mu-mimo systems,” *IEEE Communications Letters*, v. 16, n. 6, p. 925–928, 2012.
- [25] S. Stern, C. Ling, e R. F. Fischer, “Algorithms and bounds for complex and quaternionic lattices with application to mimo transmission,” *IEEE Transactions on Information Theory*, v. 68, n. 7, p. 4491–4517, 2022.
- [26] M. Bayati e A. Montanari, “The dynamics of message passing on dense graphs, with applications to compressed sensing,” *IEEE Transactions on Information Theory*, v. 57, n. 2, p. 764–785, 2011.
- [27] P. H. C. De Souza, L. L. Mendes, e M. Chaffi, “Compressive learning in communication systems: A neural network receiver for detecting compressed signals in OFDM systems,” *IEEE Access*, v. 9, p. 122 397–122 411, 2021.
- [28] A. Zappone, M. Di Renzo, e M. Debbah, “Wireless networks design in the era of deep learning: Model-based, AI-based, or both?” *IEEE Transactions on Communications*, v. 67, n. 10, p. 7331–7376, 2019.

- [29] P. H. C. d. Souza e L. L. Mendes, “Low-complexity deep unfolded neural network receiver for MIMO systems based on the probability data association detector,” *J Wireless Com Network*, v. 69, 2022.
- [30] D. Pham, K. Pattipati, P. Willett, e J. Luo, “A generalized probabilistic data association detector for multiple antenna systems,” *IEEE Communications Letters*, v. 8, n. 4, p. 205–207, 2004.
- [31] I. Tal e A. Vardy, “List decoding of polar codes,” *IEEE Transactions on Information Theory*, v. 61, n. 5, p. 2213–2226, 2015.
- [32] S. M. Alamouti, “A simple transmit diversity technique for wireless communications,” *IEEE Journal on select areas in Communications*, v. 16, n. 8, October 1998.
- [33] Centro de Referência em Radiocomunicações (CRR) Inatel, “Atividade 2.1 - Definição das arquiteturas física e lógica da rede protótipo,” *Projeto Brasil 6G*, Jul. 2022.
- [34] —, “Atividade 2.2 - efnição dos componentes de hardware da plataforma,” *Projeto Brasil 6G*, Jul. 2022.
- [35] Projeto Brasil 6G, “ Atividade 3.6 - Formatação de Quadro e Integração,” *Inatel*, agosto 2023.
- [36] Rohde & Schwarz, “ Assessing a MIMO Channel,” https://scdn.rohde-schwarz.com/ur/pws/dl_downloads/dl_application/application_notes/1sp18/1SP18_10e.pdf, 2012, [*Online*; acessado em 10 de julho de 2023].