



Universidade Estadual de Londrina
Centro de Tecnologia e Urbanismo
Departamento de Engenharia Elétrica

Davi Ingles Fogagnoli

DESEMPENHO EM SISTEMAS MIMO MASSIVO

Londrina
2024

Universidade Estadual de Londrina

Centro de Tecnologia e Urbanismo
Departamento de Engenharia Elétrica

Davi Ingles Fogagnoli

DESEMPENHO EM SISTEMAS MIMO MASSIVO

Trabalho de Conclusão de Curso orientado pelo Prof. Jaime Laelson Jacob intitulado “DESEMPENHO EM SISTEMAS MIMO MASSIVO” e apresentado à Universidade Estadual de Londrina, como parte dos requisitos necessários para a obtenção do Título de Bacharel em Engenharia Elétrica.

Orientador: Prof. Jaime Laelson Jacob

Londrina
2024

Ficha Catalográfica

Davi Ingles Fogagnoli

DESEMPENHO EM SISTEMAS MIMO MASSIVO - Londrina, 2024 - 83 p.,
30 cm.

Orientador: Prof. Jaime Laelson Jacob

1. MIMO Massivo. 2. ZF. 3. ML. 4. MMSE. 5. IC. 6. Detecção de sinais. 7.
Cancelamento de interferência.

I. Universidade Estadual de Londrina. Curso de Engenharia Elétrica. II.
DESEMPENHO EM SISTEMAS MIMO MASSIVO.

Davi Ingles Fogagnoli

DESEMPENHO EM SISTEMAS MIMO MASSIVO

Trabalho de Conclusão de Curso apresentado ao Curso de Engenharia Elétrica da Universidade Estadual de Londrina, como requisito parcial para a obtenção do título de Bacharel em Engenharia Elétrica.

Comissão Examinadora

Prof. Jaime Laelson Jacob
Universidade Estadual de Londrina
Orientador

Prof. Dr. Taufik Abrão
Universidade Estadual de Londrina

Prof. Dr. Rafael Augusto Pedriali
Universidade Estadual de Londrina

Londrina, 20 de março de 2025

Dedico este trabalho a todos aqueles
que acreditaram no meu potencial.

Agradecimentos

Agradeço ao meu orientador Prof. Jaime Laelson Jacob.

Agradeço também a minha família Tânia Beatriz Armanhi, Janira Ingles fogagnoli, Aristides Fogagnoli, Lidiany Ingles, Allan Ingles.

Aos meus amigos Pedro Pigatto, João Vitor, Bruno Henrique, Bruno Mazziero, Lucas Rodrigues, Lucca, Gabriel Mungo, Lucas Araujo Oliveira, Isaias Lustri e Ricardo Moreira. E, a todos aqueles que contribuíram.

Davi Ingles Fogagnoli. **DESEMPENHO EM SISTEMAS MIMO MASSIVO**. 2024. 83 p. Trabalho de Conclusão de Curso em Engenharia Elétrica - Universidade Estadual de Londrina, Londrina.

Resumo

Este trabalho aborda a análise de técnicas de detecção e pré-codificação aplicadas a sistemas MIMO massivo, considerando diferentes cenários de simulação. Inicialmente, são apresentados os conceitos fundamentais dessa tecnologia. No enlace de subida, serão investigadas as técnicas de detecção forçando a zero (ZF - *Zero-Forcing*), Erro Médio Mínimo Quadrático (MMSE - *Minimum Mean Square Error*), ZF com Cancelamento de Interferência (IC - *Interference Cancellation*) (ZF-IC), MMSE com Cancelamento de Interferência (MMSE-IC) e a Redução de Base (LLL - *Lenstra-Lenstra-Lovász*). No enlace de descida, serão analisadas as abordagens como pré-codificação ZF e pré-codificação MMSE, além também do método Busca por Descida Normativa (NDS - *Norm Descent Search*) aplicado aos precodificadores resultando nos métodos NDS-ZF e NDS-MMSE. Tanto no enlace de subida quanto no de descida foi realizada as simulações de cada método, analisando os esforços computacionais e desempenhos de cada método.

As simulações serão realizadas considerando diferentes configurações de número de antenas na estação base e de usuários, permitindo avaliar a taxa de erro de bit (BER - *Bit Error Rate*) e o impacto da interferência multiusuário no desempenho das técnicas estudadas. O objetivo é verificar como essas estratégias se comportam em diferentes regimes de relação sinal-ruído (SNR - *Signal-to-Noise Ratio*) e em cenários com diferentes antenas tanto receptoras quanto transmissoras em MIMO massivo. A comparação entre os métodos permitirá avaliar o desempenho da eficiência na detecção e transmissão dos símbolos, contribuindo para o entendimento das limitações e vantagens de cada abordagem.

Palavras-Chave: 1. MIMO Massivo. 2. ZF. 3. ML. 4. MMSE. 5. IC. 6. Detecção de sinais. 7. Cancelamento de interferência.

Abstract

This work addresses the analysis of detection and precoding techniques applied to massive MIMO systems, considering different simulation scenarios. Initially, the fundamental concepts of this technology are presented. In the uplink, detection techniques such as Zero-Forcing (ZF), Minimum Mean Square Error (MMSE), ZF with Interference Cancellation (ZF-IC), MMSE with Interference Cancellation (MMSE-IC), and Lenstra-Lenstra-Lovász Lattice Reduction (LLL) are investigated. In the downlink, approaches such as ZF precoding and MMSE precoding are analyzed, as well as the Norm Descent Search (NDS) method applied to the NDS-ZF and NDS-MMSE precoders. Both uplink and downlink simulations were performed, analyzing the computational efforts and performance of each method.

The simulations will be carried out considering different configurations of the number of antennas at the base station and users, allowing the evaluation of the Bit Error Rate (BER) and the impact of multi-user interference on the performance of the studied techniques. The objective is to verify how these strategies behave under different Signal-to-Noise Ratio (SNR) regimes and in scenarios with different numbers of receiving and transmitting antennas in massive MIMO. The comparison between methods will allow evaluating detection and transmission efficiency, contributing to the understanding of the limitations and advantages of each approach.

Key-words: 1. MIMO Massivo. 2. ZF. 3. ML. 4. MMSE. 5. IC. 6. Detecção de sinais. 7. Cancelamento de interferência.

Lista de ilustrações

Figura 1 – Constelação da modulação BPSK	29
Figura 2 – Constelação da modulação 4-QAM	30
Figura 3 – Função de densidade de Marcenko – Pastur para $H^H H$: Resultado de $f1_\beta(x)$ com diferentes valores de β	34
Figura 4 – Função de densidade de Marcenko – Pastur para $H \times H^H$: Resultado de $f2_\beta$ com diferentes valores de β	34
Figura 5 – Gráfico de enrijecimento do canal $H^H H$ para 8×8 MIMO. . .	36
Figura 6 – Gráfico de enrijecimento do canal $H^H H$ para 32×32 MIMO. . .	37
Figura 7 – Gráfico de enrijecimento do canal $H^H H$ para 64×64 MIMO. . .	38
Figura 8 – Gráfico de enrijecimento do canal $H^H H$ para 128×128 MIMO. . .	39
Figura 9 – Gráfico de enrijecimento do canal $H^H H$ para 256×256 MIMO. . .	40
Figura 10 – Imagem demonstrativa do sistema <i>uplink</i>	41
Figura 11 – Representação do funcionamento do algoritmo de ML.	44
Figura 12 – Resultado 4 antenas transmissoras e 16 antenas receptoras para comparação entre ZF, ZF-IC, MMSE, MMSE-IC no <i>uplink</i>	52
Figura 13 – Resultado 16 antenas transmissoras e 32 antenas receptoras para comparação entre ZF, ZF-IC, MMSE, MMSE-IC no <i>uplink</i>	53
Figura 14 – Resultado 32 antenas transmissoras e 64 antenas receptoras para comparação entre ZF, ZF-IC, MMSE, MMSE-IC no <i>uplink</i>	54
Figura 15 – Resultado 64 antenas transmissoras e 128 antenas receptoras para comparação entre ZF, ZF-IC, MMSE, MMSE-IC no <i>uplink</i>	55
Figura 16 – Resultado Simulação com diversos cenários de LLL no <i>uplink</i>	61
Figura 17 – Comparação entre as técnicas ZF,ZF-IC,MMSE,MMSE-IC e LLL em cenário com 40 usuários e a BS com 100 antenas para recepção	62
Figura 18 – Imagem demonstrativa do sistema <i>Downlink</i>	65
Figura 19 – Visão geral dos métodos de pré-codificação <i>Downlink</i>	67
Figura 20 – Diagrama de blocos ilustrando o funcionamento da pré-codificação	67
Figura 21 – Diagrama do sistema multiusuário MISO com a matriz de canal H	68
Figura 22 – Resultado 64 transmissoras e 4 antenas receptoras tendo a comparação entre ZF e MMSE, no <i>Downlink</i>	71
Figura 23 – Resultado 64 transmissoras e 32 antenas receptoras tendo a comparação entre ZF e MMSE, no <i>Downlink</i>	72

Figura 24 – Resultado 128 transmissoras e 32 antenas receptoras tendo a comparação entre ZF e MMSE, no <i>Downlink</i>	73
Figura 25 – Desempenho dos precodificadores NDS-ZF e NDS-MMSE para $n_t = 64$ e $n_r = 40$	76
Figura 26 – Desempenho dos precodificadores ZF,MMSE,NDS-ZF e NDS- MMSE para $n_t = 100$ e $n_r = 40$	78

Lista de tabelas

Tabela 1 – Tempos de execução para diferentes métodos	62
Tabela 2 – Tempos de execução para diferentes métodos	78

Lista de quadros

Quadro 1 – Comparação entre MIMO <i>Uplink</i> e <i>Downlink</i>	66
--	----

Lista de Acrônimos

2G	Sistemas de Comunicação de Segunda Geração
3G	Sistemas de Comunicação de Terceira Geração
4G	Sistemas de Comunicação de Quarta Geração
AWGN	<i>Additive White Gaussian Noise</i> - Ruído Branco Aditivo Gaussiano
BER, P_e	<i>Bit Error Rate</i> - Taxa de erro de bit
BPSK	<i>Binary Phase Shift Keying</i> - Modulação por Chaveamento de Fases
BW	<i>Band Width</i> - Largura de banda
CD	<i>Conventional Detector</i> - Detector Convencional
CDMA	<i>Code Division Multiple Access</i> - Acesso Múltiplo por Divisão de Código
DS	<i>Direct Sequence</i> - Espalhamento por Sequência Direta
DS/CDMA	<i>Direct Sequence CDMA</i> - CDMA por Sequência Direta
DSP	<i>Digital Signal Processing</i> - Técnica de processamento digital de sinais
ERB	Estação Rádio-base
<i>flops</i>	<i>Float Point Operations per Second</i> - Operações de Ponto Flutuante por segundo
IS-95	<i>Interim Standard</i> . Padrão CDMA de segunda geração
MAI	<i>Multiple Access Interference</i> - Interferência de Múltiplo Acesso
MuD	<i>Multi User Detection</i> - Detecção Multiusuário
NFR	<i>Near-Far Ratio</i> - Efeito “perto-longe”
OMuD	<i>Optimum Multiuser Detector</i> - Detector multiusuário ótimo
PD	<i>Power-Delay</i> - Perfil atraso-potência
PDF	<i>Probability Density Function</i> - Função Densidade de Probabilidade
PN	<i>Pseudo-Noise</i> - Trata-se de sequências pseudo-aleatórias
QAM	<i>Quadrature Amplitude Modulation</i> - Modulação por Amplitude em Quadratura
QoS	<i>Quality of Service</i> - Qualidade do serviço
QPSK	<i>Quadrature Phase Shift Keying</i> - Modulação por Chaveamento de Fases em Quadratura
RAKE	Receptor com diversos MFB para canais multipercurso
SI	<i>Self-interference</i> - Auto-interferência
SNR	<i>Signal Noise Ratio</i> - Relação Sinal Ruído
SS	<i>Spread Spectrum</i> - Espalhamento Espectral
SUB	<i>Single-User Bound</i> - Limite de BER para usuário isolado
SUD	<i>Single-User Detection</i> - Detecção Uniusuário
W-H	<i>Walsh Hadamard</i> - Família de sequências de espalhamento
MMSE	<i>Minimum Mean Square Error</i> - Erro Médio Mínimo Quadrático

ZF	<i>Zero Forcing</i> - Forçamento a Zero
ML	<i>Maximum Likelihood</i> - Máxima Verossimilhança
NDS	<i>Norm Descent Search</i> - Busca por Descida da Norma
BPSK	<i>Binary Phase Shift Keying</i> - Chaveamento de Fase Binário
PSK	<i>Phase Shift Keying</i> - Chaveamento por Deslocamento de Fase
QPSK	<i>Quadrature Phase Shift Keying</i> - Chaveamento de Fase em Quadratura
MIMO	<i>Multiple-Input Multiple-Output</i> - Múltiplas Entradas e Múltiplas Saídas
IC ou CI	<i>Interference Cancellation</i> - Cancelamento de Interferência
SE	<i>Spectral Efficiency</i> - Eficiência Espectral
LR	<i>Lattice Reduction</i> - Redução de Lattice
LLL	<i>Lenstra-Lenstra-Lovász</i> - Algoritmo de redução de lattice
VP	<i>Vector Perturbation</i> - Perturbação Vetorial
CN	<i>Complex Normal</i> - Distribuição Normal Complexa
PSK	<i>Phase Shift Keying</i> - Chaveamento por Deslocamento de Fase
QAM	<i>Quadrature Amplitude Modulation</i> - Modulação por Amplitude em Quadratura
CSI	<i>Channel State Information</i> - Informação do Estado do Canal

Lista de símbolos

\mathbf{n}_t	Número de antenas transmissoras
\mathbf{n}_r	Número de antenas receptoras
ρ	Relação sinal-ruído normalizada
M	Número de símbolos na modulação
A	Amplitude da portadora
H	Matriz do canal
H_1	Coefficiente da matriz do canal
H_2	Canal normalizado para um usuário específico
N_r	Número de antenas receptoras
N_t	Número de antenas transmissoras
\mathbf{P}	Matriz diagonal representando as potências de transmissão dos usuários
\mathbf{N}	Matriz de ruído Gaussiano
\mathbf{x}	Vetor do sinal transmitido
\mathbf{y}	Vetor do sinal recebido
\mathbf{r}_{ZF}	Sinal recebido com a matriz de pré-processamento Zero-Forcing
\mathbf{W}	Matriz de pré-processamento
$\hat{\mathbf{x}}_{MMSE}$	Sinal detectado via MMSE
\mathbf{QR}	Decomposição QR, onde \mathbf{Q} é unitária e \mathbf{R} é triangular superior
\mathbf{T}	Matriz identidade
μ	Coefficiente usado na técnica LLL
δ	Parâmetro para a condição do LLL
Θ_1	Matriz de rotação ortogonal
a_1, b_1	Coefficientes da matriz de rotação ortogonal
\mathbf{y}_p	Vetor de sinais recebidos após a aplicação do pré-processamento
\mathbf{H}_p	Matriz do canal após a transformação pelo pré-processamento
\mathbf{x}_p	Vetor de símbolos transmitidos considerando a matriz de pré-processamento
\mathbf{n}_p	Vetor de ruído aditivo gaussiano branco após a transformação do pré-processamento
\mathbf{G}	Matriz de equalização do receptor, utilizada para estimar os símbolos transmitidos
\mathbf{G}_{ZF}	Matriz de equalização baseada no critério Zero-Forcing (ZF)
\mathbf{H}_r	Matriz do canal real representando a relação entre os sinais transmitidos e recebidos
\mathbf{H}_r^T	Transposta da matriz do canal real
\mathbf{G}_{MMSE}	Matriz de equalização baseada no critério MMSE (Minimum Mean Square Error)
u	Vetor de símbolos transmitidos no domínio reduzido
$\tau \mathbf{p}$	Termo de perturbação escalado pelo fator τ aplicado ao vetor \mathbf{p}
$\tilde{\mathbf{y}}$	Vetor de sinais recebidos após a transformação QR

$\tilde{\mathbf{p}}$	Vetor perturbado utilizado na técnica de redução de redes
$\tilde{\mathbf{n}}$	Vetor de ruído transformado após a decomposição QR
$\hat{\mathbf{u}}$	Estimativa do vetor de símbolos transmitidos após a detecção
\mathbf{z}	Vetor usado no calculo do pré-processamento NDS
λ_i	Fator de ajuste
$\Delta q_{i,\text{opt}}$	Termo de otimização relacionado à escolha de perturbação
Δq_{\min}	Valor mínimo do termo $\Delta q_{i,\text{opt}}$
Índice j	Índice do elemento que minimiza $\Delta q_{i,\text{opt}}$
\mathbf{e}_j	Vetor unitário com valor 1 na posição j e zeros nos demais elementos
\mathbf{f}_j	j -ésima coluna da matriz \mathbf{F} usada na técnica de redução de redes
S_{\max}	Maior módulo dos símbolos na constelação utilizada
$\mathbf{y}_{\mathbf{c}}$	Vetor de sinais recebidos no domínio complexo
\mathbf{y}_I	Parte real do vetor de sinais recebidos
\mathbf{y}_Q	Parte imaginária do vetor de sinais recebidos
$\mathbf{u}_{\mathbf{c}}$	Vetor de símbolos transmitidos no domínio complexo
\mathbf{u}_I	Parte real do vetor de símbolos transmitidos
\mathbf{u}_Q	Parte imaginária do vetor de símbolos transmitidos
$\mathbf{x}_{\mathbf{c}}$	Vetor de sinais transmitidos no domínio complexo
\mathbf{x}_I	Parte real do vetor de sinais transmitidos
\mathbf{x}_Q	Parte imaginária do vetor de sinais transmitidos
$\mathbf{H}_{\mathbf{c}}$	Matriz do canal no domínio complexo
\mathbf{H}_I	Parte real da matriz do canal
\mathbf{H}_Q	Parte imaginária da matriz do canal
$\mathbf{n}_{\mathbf{c}}$	Vetor de ruído aditivo gaussiano branco no domínio complexo
\mathbf{n}_I	Parte real do vetor de ruído
\mathbf{n}_Q	Parte imaginária do vetor de ruído

Sumário

1	OBJETIVOS	25
2	INTRODUÇÃO AO CENÁRIO	27
2.1	Modulações de deslocamento de fase (PSK - <i>Phase Shift Keying</i>) e Modulação por Amplitude em Quadratura (QAM (Quadrature Amplitude Modulation))	28
2.1.1	Modulação PSK	28
2.1.1.1	Modulação de chaveamento de mudança de fase binária (BPSK - <i>Binary Phase Shift Keying</i>)	28
2.1.2	Modulação QAM	29
2.2	MIMO	30
2.2.1	Multi-Usuário MIMO	30
2.2.2	Propagação favorável	31
2.3	Massivo MIMO	32
2.3.1	MIMO x Massivo MIMO	32
2.4	Rigidez do Canal com Dimensões Maiores	33
3	TÉCNICAS DE RECEPÇÃO MIMO: ENLACE DE SUBIDA	41
3.1	Detecção Máxima Verossimilhança (ML - <i>Maximum Likelihood</i>)	43
3.1.1	Simulação ML	43
3.2	Detecção Forçamento a Zero (ZF - <i>Zero-Forcing</i>): <i>Uplink</i>	45
3.3	Detecção Erro Quadrático Médio Mínimo - (MMSE - <i>Minimum Mean Square Error</i>): <i>Uplink</i>	46
3.3.1	Comparação MMSE com a Detecção ML	46
3.4	Cancelamento de Interferência (CI)	47
3.4.1	Desvantagens do IC	47
3.5	Detecção ZF-IC	48
3.6	Detecção MMSE-IC	48
3.6.1	Comparação entre o MMSE-IC e MMSE	49
3.7	Simulação: Comparação entre ZF, ZF-IC, MMSE, MMSE-IC	50
3.7.1	Parâmetros da Simulação	50
3.7.1.1	Ruído	50
3.7.1.2	Simulação do IC	50
3.7.2	Resultados da Simulação: Comparação entre ZF, ZF-IC, MMSE, MMSE-IC	52

3.8	Detecção Lenstra-Lenstra-Lovász (LLL): <i>Uplink</i>	55
3.8.0.1	Passo 1: Redução do Comprimento do Segundo Vetor Coluna r_2 .	56
3.8.0.2	Passo 2: Redução do Comprimento do Terceiro Vetor Coluna r_3 .	57
3.8.0.3	Passo 3: Redução do Comprimento do Quarto Vetor Coluna r_4 . .	59
3.8.1	Simulação detecção LLL: <i>Uplink</i>	60
3.9	Análise de Desempenho em MIMO Massivo <i>Uplink</i>	61
4	TÉCNICAS DE RECEPÇÃO MIMO: ENLACE DE DE DES- CIDA	65
4.0.1	Pré-codificação	67
4.0.2	Modelo do sistema	68
4.1	Detecção ZF: <i>Downlink</i>	69
4.1.1	Vetor de Pré-codificação G	69
4.2	Detecção MMSE: <i>Downlink</i>	70
4.2.1	Simulação ZF e MMSE em <i>Downlink</i>	71
4.3	Detecção com Pré-codificação Precodificador baseado em Busca por Descida de Norma - (NDS - <i>Precoder based on Norm Des-</i> <i>cent Search</i>)	73
4.3.1	Simulação NDS	74
4.3.1.1	Resultados da Simulação NDS	76
4.4	Análise de Desempenho em MIMO Massivo <i>Downlink</i>	77
5	CONCLUSÃO	81
	REFERÊNCIAS	83

1 Objetivos

O objetivo deste trabalho é avaliar o desempenho de diferentes técnicas de detecção e pré-codificação em sistemas MIMO massivo, considerando cenários com distintas configurações de antenas. No enlace de subida, são analisados os métodos forçando a zero (ZF - *Zero-Forcing*), Erro Médio Mínimo Quadrático (MMSE - *Minimum Mean Square Error*), ZF com Cancelamento de Interferência (IC - *Interference Cancellation*) (ZF-IC), MMSE com Cancelamento de Interferência (MMSE-IC) e Redução de Lattice de Lenstra-Lenstra-Lovász (LLL - *Lenstra-Lenstra-Lovász*), verificando seu impacto na taxa de erro de bit e na qualidade da comunicação. No enlace de descida, são investigadas estratégias de pré-codificação baseadas em ZF e MMSE, além da aplicação do método de Busca por Descida Normativa (NDS - *Norm Descent Search*) aos precodificadores resultando nos métodos NDS-ZF e NDS-MMSE, buscando compreender seus efeitos na mitigação da interferência multiusuário e na eficiência espectral.

As simulações realizadas levam em conta diferentes quantidades de antenas na estação base e nos usuários, permitindo examinar como cada método se comporta diante de variações na estrutura do sistema e em diferentes níveis de relação sinal-ruído. A análise comparativa das técnicas possibilita identificar a relação entre complexidade computacional e desempenho, auxiliando na escolha da abordagem mais adequada para cada cenário. A partir dos resultados obtidos, busca-se fornecer uma visão detalhada das vantagens e limitações de cada estratégia.

2 Introdução ao Cenário

A quinta geração de tecnologia móvel, o 5G, representa uma evolução significativa no panorama das comunicações sem fio. Uma das inovações mais impactantes nesse cenário é o conceito de tecnologia de comunicação de múltiplas entradas, múltiplas saídas (MIMO - *Multiple Input, Multiple Output*) massivo, onde estações base com múltiplas antenas multiplexam espacialmente uma multidão de terminais de usuário em toda a largura de banda como abordado no livro (CHIEN; BJORNSEN, 2017).

O 5G tem como objetivo aumentar a taxa de transferência de área que é dada pela fórmula:

$$\begin{aligned} \text{Taxa de transferência de área (bit/s/km}^2\text{)} &= \\ &= \text{Largura de banda (Hz)} \times \text{Densidade de células (células/km}^2\text{)} \times \\ &\quad \text{Eficiência espectral (bit/s/Hz/células)} \end{aligned} \tag{2.1}$$

A fórmula 2.1 mostra que para maior taxa de transferência de área é necessário aumentar (1) Largura de banda designado para 5G (2) Aumentar o número de células independentes (3) Aumentar a Eficiência espectral.

- **1. Largura de Banda:** Já está sendo usada a de melhor eficiência, sendo elas entre 1 GHz e 6 GHz. Dependendo da distancia de atuação que está sendo implantada.
- **2. Densidade de Células:** Em ambientes urbanos onde há maior demanda. temos a situação que maior densificação de células certamente é possível, mas parece que estamos atingindo um ponto de saturação.
- **3. Eficiência das Transmissões de Dados:** A eficiência das transmissões de dados (por célula e para uma quantidade dada de largura de banda) pode ser melhorada e possivelmente se tornar o principal meio de alcançar alta taxa de transferência. Assim é feita implementação de MIMO Massivo que fornece os meios para melhorar a eficiência espectral.

A capacidade de transmitir dados eficientemente não teve grandes avanços em redes anteriores. Isso significa que há espaço para melhorias significativas no futuro, e a eficiência espectral pode se tornar uma maneira fundamental de alcançar altas taxas de transferência nas redes 5G. Assim irá ser explicado a lógica por trás da tecnologia MIMO massivo, que é uma forma avançada de usar múltiplas antenas nas estações base para melhorar drasticamente a eficiência da transmissão de dados. Assim, enquanto as redes anteriores se concentravam principalmente em velocidade de conexão, como o 4G, as redes 5G estão focadas em melhorar a eficiência espectral para aumentar a capacidade da rede. Tendo que a técnica MIMO massivo, que é uma tecnologia que usa múltiplas antenas

em estações base para aumentar a eficiência da transmissão de dados, é uma das principais soluções para alcançar essas melhorias significativas na eficiência espectral.

2.1 Modulações de deslocamento de fase (PSK - *Phase Shift Keying*) e Modulação por Amplitude em Quadratura (QAM (Quadrature Amplitude Modulation))

2.1.1 Modulação PSK

A modulação PSK é uma técnica de modulação digital que transmite informações variando a fase da onda portadora. Nesse esquema, a amplitude do sinal permanece constante, enquanto a fase assume valores discretos de acordo com os símbolos transmitidos. O número de fases distintas utilizadas define a ordem da modulação, sendo que o *M-PSK* representa um conjunto de M símbolos distribuídos uniformemente em um círculo no plano complexo.

A equação geral para um sinal modulado em PSK é dada por:

$$s(t) = A \cos(2\pi f_c t + \theta_k) \quad (2.2)$$

Em que A é a amplitude da portadora, f_c é a frequência da portadora e θ_k é a fase correspondente ao símbolo k . O valor de θ_k depende do número de símbolos M , sendo definido como:

$$\theta_k = \frac{2\pi k}{M}, \quad k = 0, 1, \dots, M - 1 \quad (2.3)$$

2.1.1.1 Modulação de chaveamento de mudança de fase binária (BPSK - *Binary Phase Shift Keying*)

A modulação BPSK é a forma mais simples da modulação PSK, utilizando apenas dois estados de fase, tipicamente 0° e 180° . Isso significa que cada símbolo transmitido carrega um único bit de informação ($M = 2$), sendo um dos esquemas mais robustos contra ruído e interferências. O sinal BPSK pode ser representado como:

$$s(t) = A \cos(2\pi f_c t + \pi b) \quad (2.4)$$

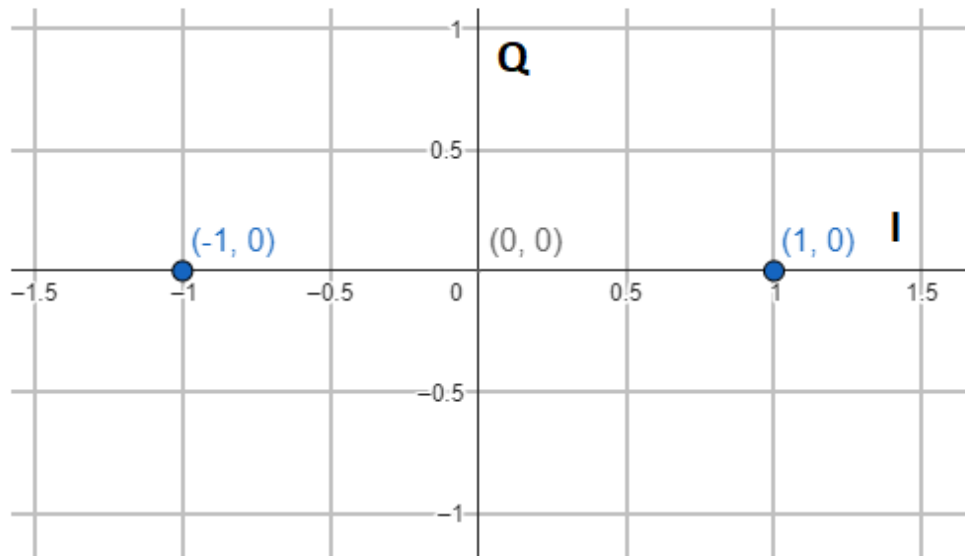
Em que b representa o bit transmitido (0 ou 1). Para $b = 0$, a fase da portadora permanece em 0° , enquanto para $b = 1$, a fase muda para 180° . Essa grande separação de fase entre os símbolos torna o BPSK altamente resistente a erros em canais ruidosos,

mas, por outro lado, limita a eficiência espectral, pois apenas um bit é transmitido por símbolo.

Por conta de sua robustez, o BPSK é amplamente utilizado em sistemas de comunicação onde a confiabilidade é a prioridade.

Tendo sua constelação representada pela figura 1:

Figura 1 – Constelação da modulação BPSK



Fonte: O próprio autor

Como visto na imagem 1, os sinais modulados são combinados em I (em fase) e Q (quadratura) para criar os sinais $(-1, 0)$ e $(1, 0)$.

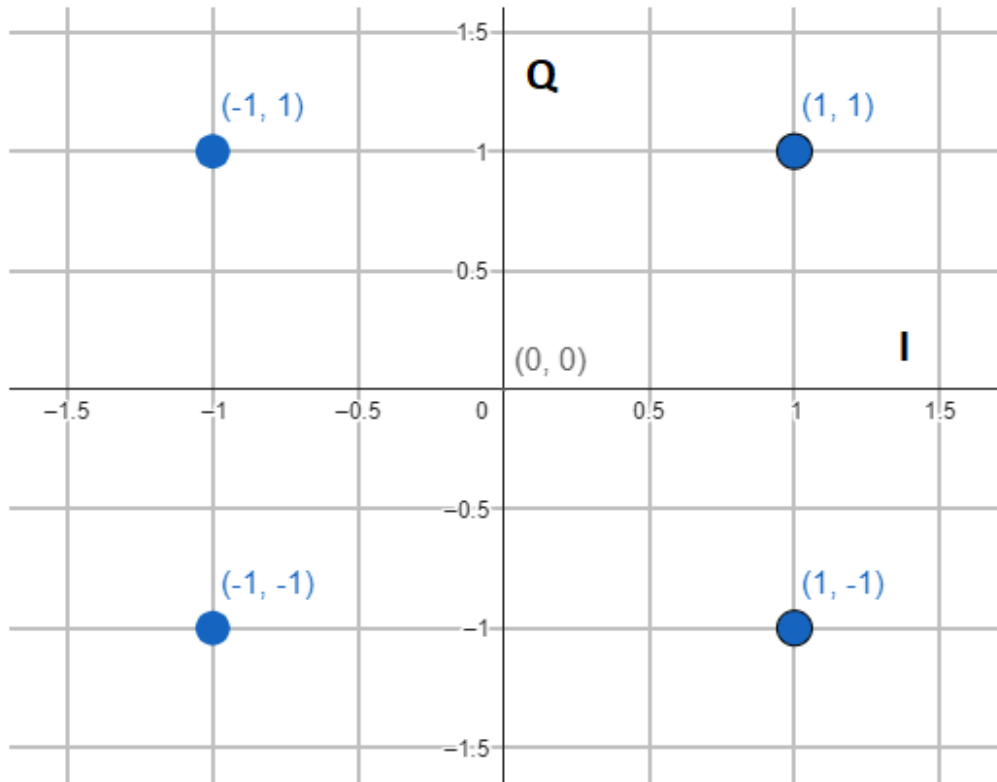
2.1.2 Modulação QAM

A modulação QAM é um método que combina variações tanto na amplitude quanto na fase da portadora para representar múltiplos bits por símbolo. Nessa modulação é necessário decompor o sinal em duas componentes ortogonais, chamadas de componente em fase (I - *In-phase*) e componente em quadratura (Q - *Quadrature*). Os valores dessas componentes definem pontos em uma constelação, permitindo que múltiplos níveis de amplitude e fase sejam utilizados para representar diferentes combinações de bits. A quantidade de bits por símbolo depende da ordem da modulação, sendo que no 16-QAM, por exemplo, cada símbolo carrega 4 bits, enquanto no 64-QAM, cada símbolo transporta 6 bits. como visto em (MARZETTA et al., 2016).

A constelação 4-QAM é representada pela figura:

Como visto na imagem 2, os sinais modulados estão combinados em I e Q para criar os sinais $(-1, 1)$, $(1, 1)$, $(-1, -1)$ e $(1, -1)$.

Figura 2 – Constelação da modulação 4-QAM



Fonte: O próprio autor

As constelações de ordem superior tornam o sistema mais suscetível a ruído e interferências, uma vez que a distância entre os símbolos na constelação diminui, aumentando a probabilidade de erros na recepção.

A vantagem da modulação QAM em relação a modulação PSK é a sua maior eficiência espectral, pois permite transmitir mais informações dentro da mesma largura de banda.

2.2 MIMO

2.2.1 Multi-Usuário MIMO

Como abordado nos livros (HAMPTON, 2014) e (CHIEN; BJORNSEN, 2017) em um sistema MIMO utilizam-se com múltiplas entradas e múltiplas saídas de forma que em cada estação base em uma rede celular atende a uma multidão de terminais de usuários. Convencionalmente, os recursos de tempo/frequência são divididos em blocos de recursos e apenas um dos terminais de usuário fica ativo por bloco. Este terminal pode então receber um único fluxo de dados com uma eficiência espectral (SE - *Spectral Efficiency*) quantificada através da Relação Sinal-Ruído (SNR - *Signal-to-Noise Ratio*) como $\log_2(1 + \text{SNR})$. A maneira eficiente de aumentar a eficiência espectral de uma rede celular é

ter múltiplas transmissões paralelas. Se houver uma quantidade U de transmissões paralelas e independentes, a SE se torna $U \log_2(1 + \text{SNR})$, onde U atua como um fator pré-log multiplicativo. Formando assim a fórmula do conceito fundamental para a Capacidade de Shannon em um canal com ruído branco gaussiano aditivo (AWGN - *Additive White Gaussian Noise*) em que B é a largura de banda.

As transmissões paralelas podem ser realizadas tendo múltiplas antenas de transmissão e múltiplas antenas de recepção. Existem dois principais casos distintos:

- **MIMO ponto a ponto:** onde uma estação base (BS - *Base Station*) com múltiplas antenas comunica-se com um único terminal de usuário tendo múltiplas antenas.
- **MIMO multiusuário:** onde uma BS com múltiplas antenas comunica-se com múltiplos terminais de usuários com cada usuário podendo ter apenas uma ou múltiplas antenas.

Assim, antes em uma rede celular apenas um terminal de usuário podia transmitir ou receber dados em um determinado bloco de recursos. Agora, para melhorar a eficiência espectral da rede, utiliza-se múltiplas transmissões paralelas. Isso significa que vários terminais de usuário podem enviar ou receber dados simultaneamente no mesmo bloco de recursos, aumentando assim a capacidade de transmissão da rede. Essa otimização é alcançada usando múltiplas antenas tanto para transmitir quanto para receber sinais, permitindo comunicações mais eficientes e rápidas entre a estação base e os terminais de usuário.

A transmissão da capacidade em sistemas MIMO multiusuário é fundamentada em processamentos de sinais. Um exemplo é a técnica Zero-Force, que será melhor explicada na seção 3.2.

Essas técnicas exigem cálculos intensivos e informação precisa do estado do canal (CSI). Sem isso, as tentativas de mitigar a interferência podem acabar sendo prejudiciais ao desempenho.

2.2.2 Propagação favorável

Suponha que $\mathbf{h}_1, \mathbf{h}_2 \in \mathbb{C}^{n_t}$ representem as respostas do canal entre uma BS e dois terminais de usuário diferentes e N_t represente a quantidade de antenas de transmissão. Se esses vetores forem não nulos e ortogonais no sentido de que $\mathbf{h}_1^H \mathbf{h}_2 = 0$, onde $(\cdot)^H$ denota o conjugado transposto, então a BS pode separar completamente os sinais $\mathbf{x}_1, \mathbf{x}_2$ transmitidos pelos usuários quando observa $\mathbf{y} = \mathbf{h}_1 \mathbf{x}_1 + \mathbf{h}_2 \mathbf{x}_2$. Ao calcular simplesmente o produto interno entre \mathbf{y} e \mathbf{h}_1 , a BS obtém

$$\mathbf{h}_1 \mathbf{y} = \mathbf{h}_1^H \mathbf{h}_1 x_1 + \mathbf{h}_1^H \mathbf{h}_2 x_2 = \|\mathbf{h}_1\|^2 x_1$$

Em que a interferência entre os usuários desapareceu devido a $\mathbf{h}_1^H \mathbf{h}_2 = 0$. A mesma coisa pode ser feita para o segundo usuário: $\mathbf{h}_2 \mathbf{y} = \|\mathbf{h}_2\|^2 \mathbf{x}_2$. Note que a BS precisa de conhecimento perfeito de \mathbf{h}_1 e \mathbf{h}_2 para calcular esses produtos internos. A ortogonalidade do canal é chamada de propagação favorável, pois os dois usuários podem se comunicar com a BS sem afetar um ao outro.

2.3 Massivo MIMO

O Massivo MIMO é uma inovação na comunicação sem fio que utiliza múltiplas antenas tanto em estações base quanto em dispositivos de usuários. Sua principal característica encontra-se na melhoria substancial do desempenho e capacidade das redes sem fio ao empregar um grande número de antenas. O Massivo MIMO busca explorar o fenômeno de diversidade espacial, combatendo interferências e aumentando a eficiência espectral. Essa tecnologia revoluciona as capacidades das redes ao permitir transmissões simultâneas para diversos usuários, o que não apenas aumenta a velocidade, mas também eleva a confiabilidade das conexões. Diferentemente das tecnologias MIMO convencionais, o Massivo MIMO utiliza um número maior de antenas, sendo igual ou superior a 64 na BS, enviando ou recebendo dados de 1 ou mais usuários, assim tendo ganhos em termos de capacidade e eficiência.

O Massivo MIMO traz consigo a promessa de taxas de transferência maiores e uma melhor capacidade para lidar com múltiplos dispositivos conectados simultaneamente. No entanto, o aumento no número de antenas acarreta complexidades consideráveis, especialmente no que diz respeito às técnicas de detecção e ao processamento dos sinais.

A complexidade dessas técnicas está diretamente relacionada à necessidade de lidar com uma quantidade massiva de informações provenientes de um grande conjunto de antenas. O processamento desses dados demanda abordagens avançadas de análise e detecção para extrair informações úteis de maneira eficiente e precisa.

2.3.1 MIMO x Massivo MIMO

O Massivo MIMO difere do MIMO convencional em dois aspectos principais:

1. **Número de Antenas na BS:** No MIMO massivo, o número de antenas na estação base é significativamente maior do que o número de receptores, que geralmente possuem apenas uma antena.
2. **Processamento de Sinal:**
 - **Uplink:** Na transmissão de enlace de subida, em resumo todos os usuários transmitem simultaneamente, e na estação base, os sinais de todos os usuários chegam juntos. Neste caso, é necessário realizar um processamento do sinal

para detectando e separando a transmissão de cada usuário, que será melhor discutido no capítulo 3.

- **Downlink:** Na transmissão de enlace de descida, a estação base gera, modula e pré-codifica os sinais dos usuários, e faz essa operação com uma soma com todos os sinais antes de transmitir. Cada usuário então recebe o sinal e realiza a demodulação para extrair sua parte. que será melhor discutido no capítulo 4.

2.4 Rigidez do Canal com Dimensões Maiores

Como abordado em (CHOCKALINGAM; RAJAN, 2014), temos que as principais vantagens são que os sistemas MIMO grandes são o aumento da taxa de dados e o ganho de diversidade e a grande dimensionalidade. Assim resultando em outras vantagens como por exemplo a medida que a matriz de canal \mathbf{H} de dimensão $N_r \times N_t$ se torna maior. Mantendo a razão N_t por N_r . por consequência temos que torna-se menos sensível à distribuição real das entradas da matriz do canal (desde que sejam independentes e distribuídas identicamente (iid)).

Isso é resultado da lei de Marcenko-Pastur, que afirma que se as entradas de uma matriz $nt \times nr$. A Matrix \mathbf{H} têm média zero iid com variância $1/nr$, então a distribuição empírica dos autovalores $\mathbf{H}^H \times \mathbf{H}$ converge quase certamente, conforme $nt, nr \rightarrow \infty$ com $nr/nt \rightarrow \beta$, para a função de densidade:

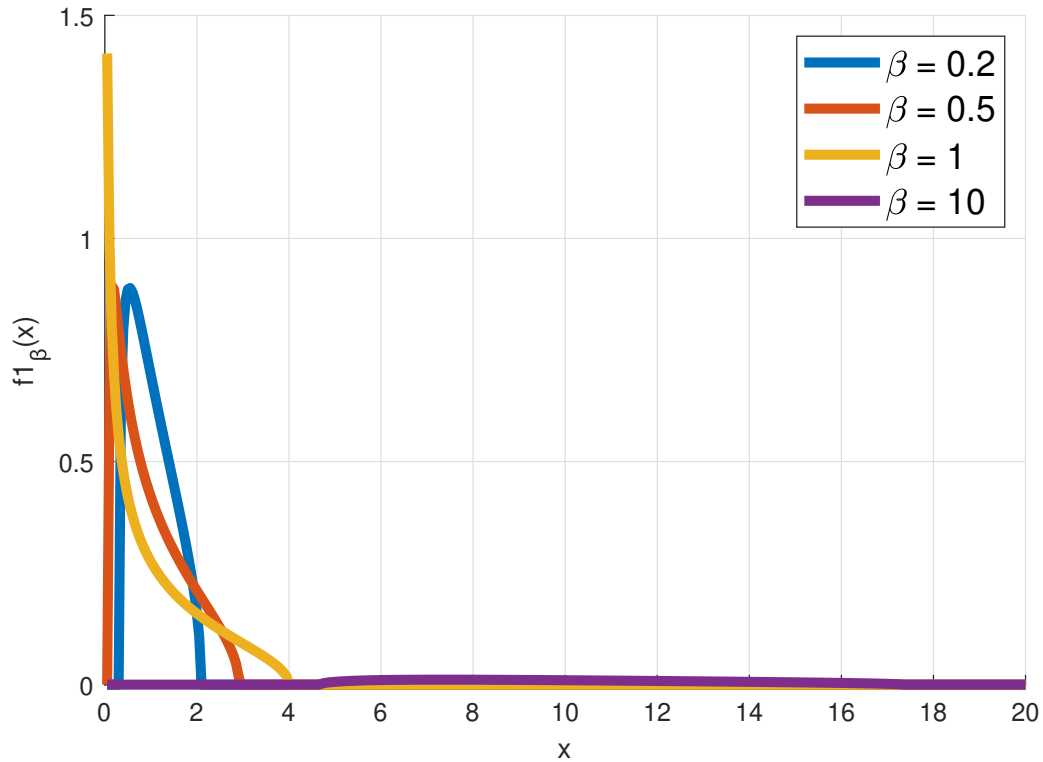
$$f1_\beta(x) = \left(1 - \frac{1}{\beta}\right)^+ \delta(x) + \frac{\sqrt{(x-a)^+(b-x)^+}}{2\pi\beta x} \quad (2.5)$$

Em que $(z)^+ = \max(z, 0)$, $a = (1 - \sqrt{\beta})^2$, e $b = (1 + \sqrt{\beta})^2$. De forma semelhante, a distribuição empírica dos autovalores de $\mathbf{H} \times \mathbf{H}^H$ converge para:

$$f2_\beta(x) = (1 - \beta)\delta(x) + \beta f1_\beta(x) \quad (2.6)$$

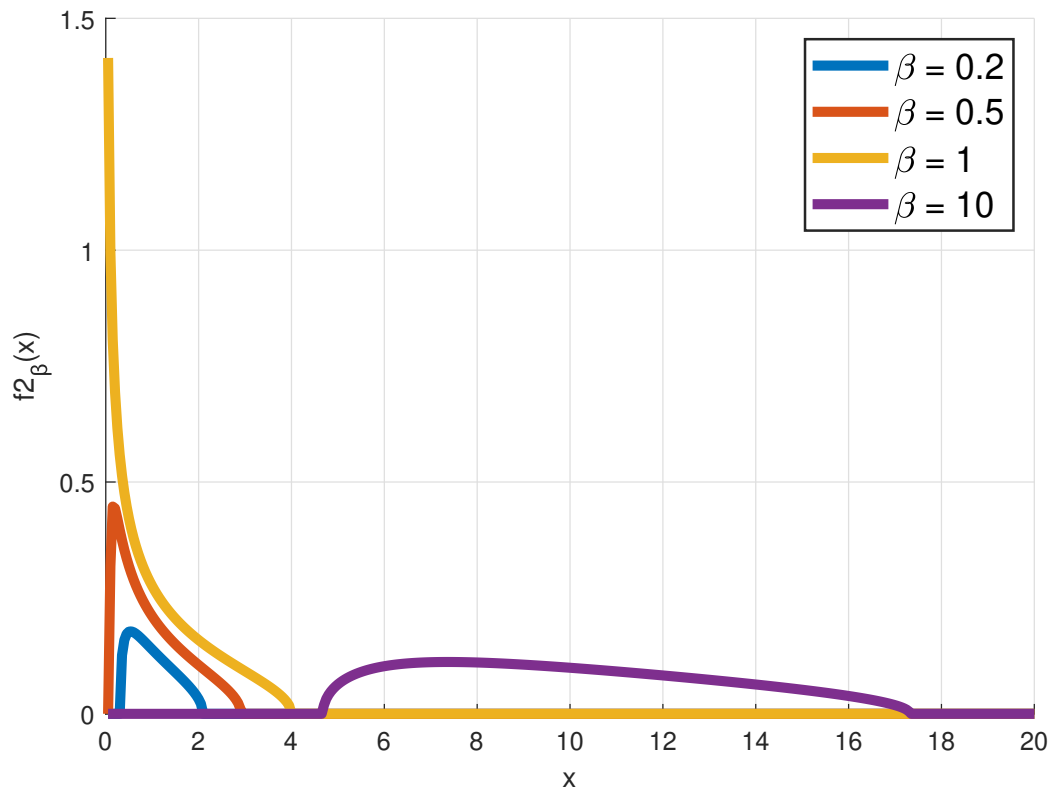
As equações 2.5 e 2.6 estão plotadas pelas figuras 3 e 4 respectivamente, para diferentes valores de β .

Figura 3 – Função de densidade de Marcenko – Pastur para $\mathbf{H}^H \mathbf{H}$:
Resultado de $f1_\beta(x)$ com diferentes valores de β .



Fonte: O próprio autor

Figura 4 – Função de densidade de Marcenko – Pastur para $\mathbf{H} \times \mathbf{H}^H$:
Resultado de $f2_\beta$ com diferentes valores de β .



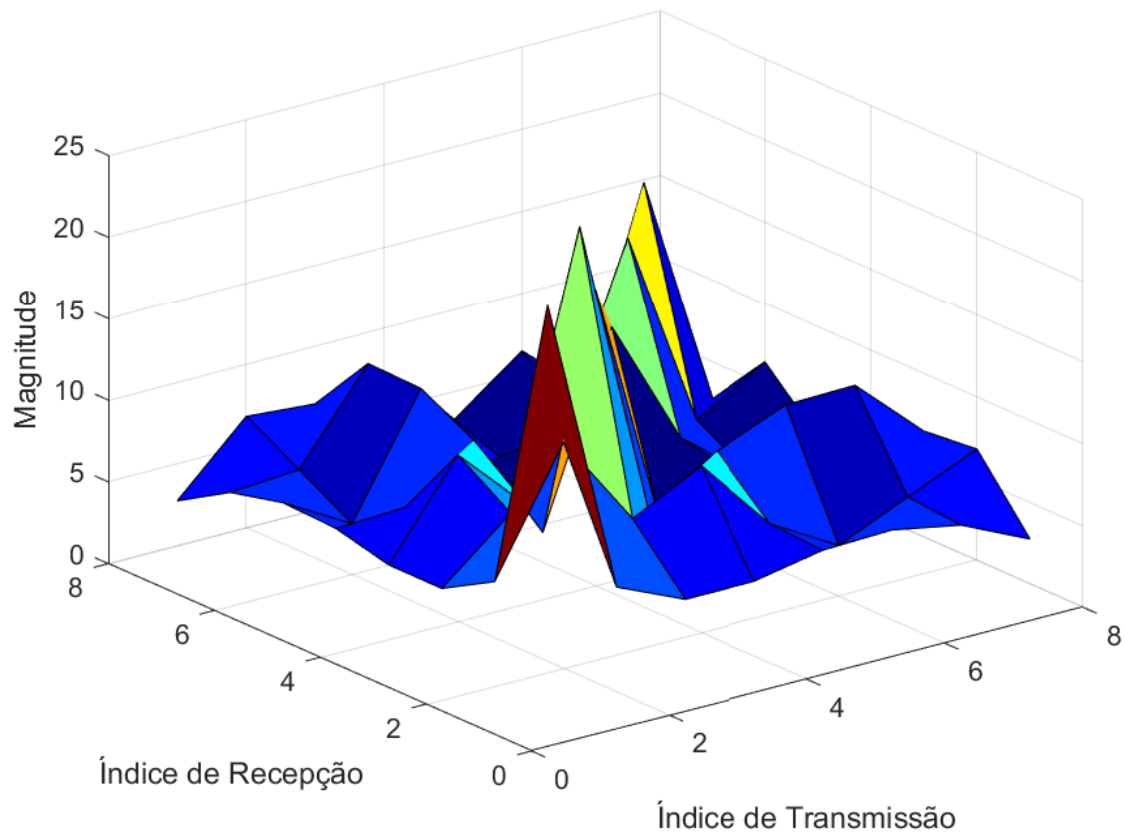
Fonte: O próprio autor

Uma das consequências decorrentes da lei de Marcenko-Pastur é que as matrizes associadas a um grande número de canais apresentam excelente condicionamento. Esse comportamento é observado nas curvas das Figuras 3 e 4 para os valores $\beta = 0.2$ e $\beta = 10$, onde se verifica a distribuição dos autovalores não nulos das matrizes $\mathbf{H}\mathbf{H}^H$ e $\mathbf{H}^H\mathbf{H}$.

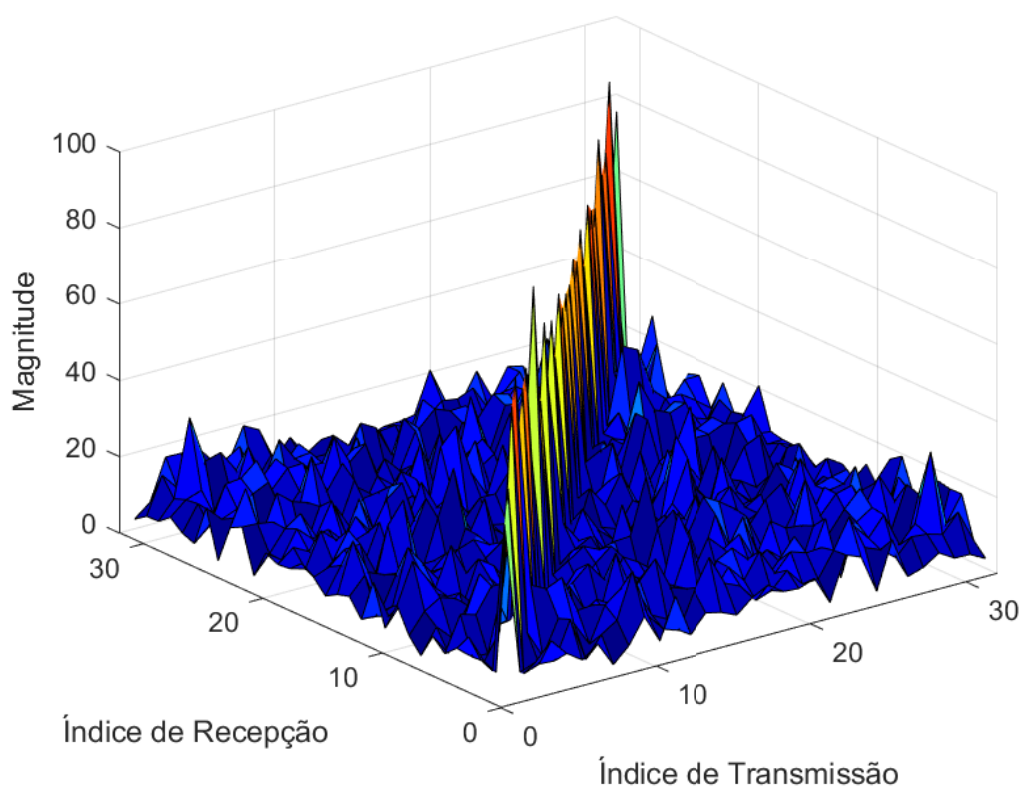
Além disso, a lei de Marcenko-Pastur estabelece que o canal sofre um processo de endurecimento, o que significa que o histograma dos autovalores de uma única realização tende a convergir para a distribuição assintótica média desses autovalores. Esse fenômeno implica que, à medida que o número de antenas cresce, a resposta do canal se torna progressivamente mais determinística.

A evidência do endurecimento do canal em sistemas de grandes dimensões pode ser analisada a partir dos gráficos de intensidade da matriz $\mathbf{H}^H\mathbf{H}$ para diferentes dimensões, correspondentes a $n_t = n_r = 8, 32, 64, 128$ e 256 , apresentados nas Figuras 5, 6, 7, 8 e 9. No gráfico o eixo x e o eixo y representa o índice do coeficiente da matriz $\mathbf{H}^H\mathbf{H}$, tendo no eixo z a magnitude do coeficiente. Assim nessas representações, considera-se que os coeficientes da matriz \mathbf{H} são amostras de um processo gaussiano independente e identicamente distribuído (iid), com média zero e variância unitária. As figuras evidenciam que, conforme o tamanho da matriz \mathbf{H} aumenta, os elementos diagonais de $\mathbf{H}^H\mathbf{H}$ passam a apresentar magnitudes significativamente superiores aos elementos fora da diagonal, reforçando a característica de endurecimento do canal.

Figura 5 – Gráfico de enrijecimento do canal $H^H H$ para 8×8 MIMO.

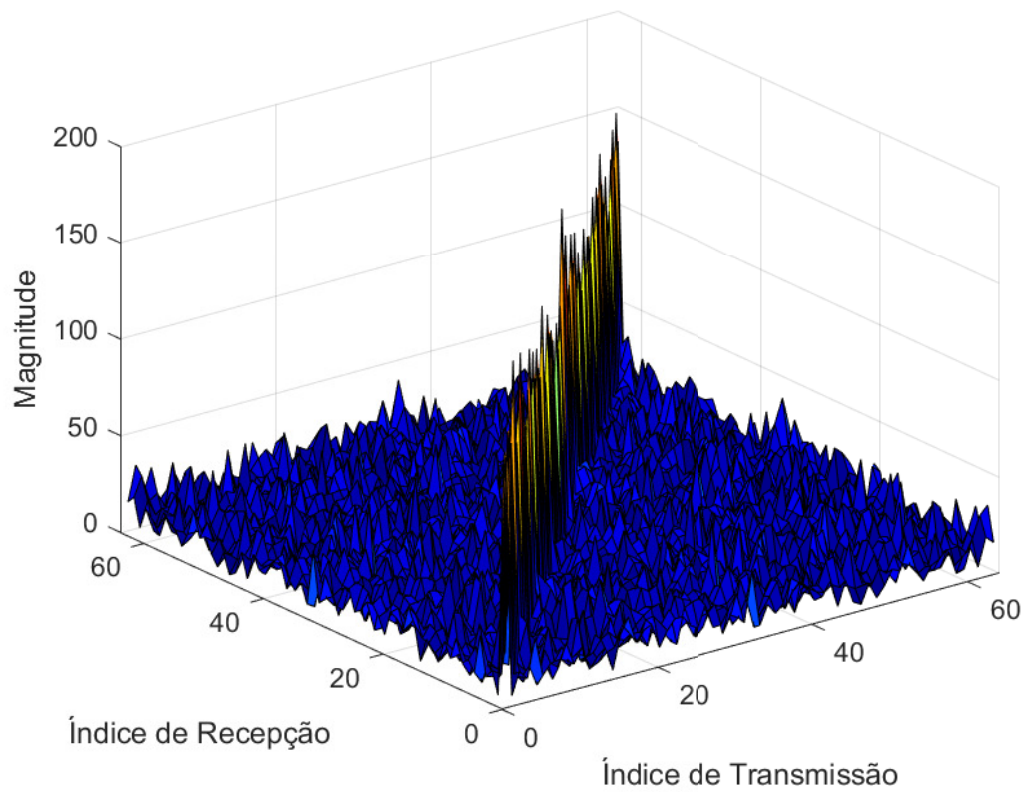


Fonte: O próprio autor

Figura 6 – Gráfico de enrijecimento do canal $H^H H$ para 32×32 MIMO.

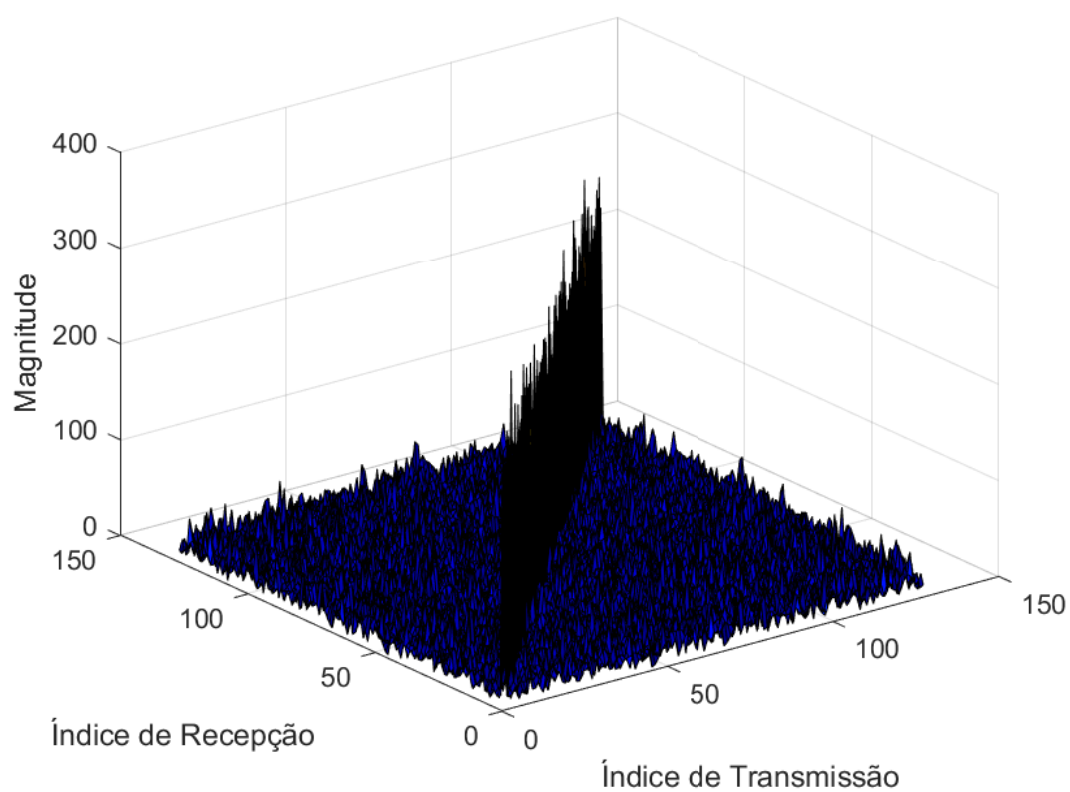
Fonte: O próprio autor

Figura 7 – Gráfico de enrijecimento do canal $\mathbf{H}^H \mathbf{H}$ para 64×64 MIMO.



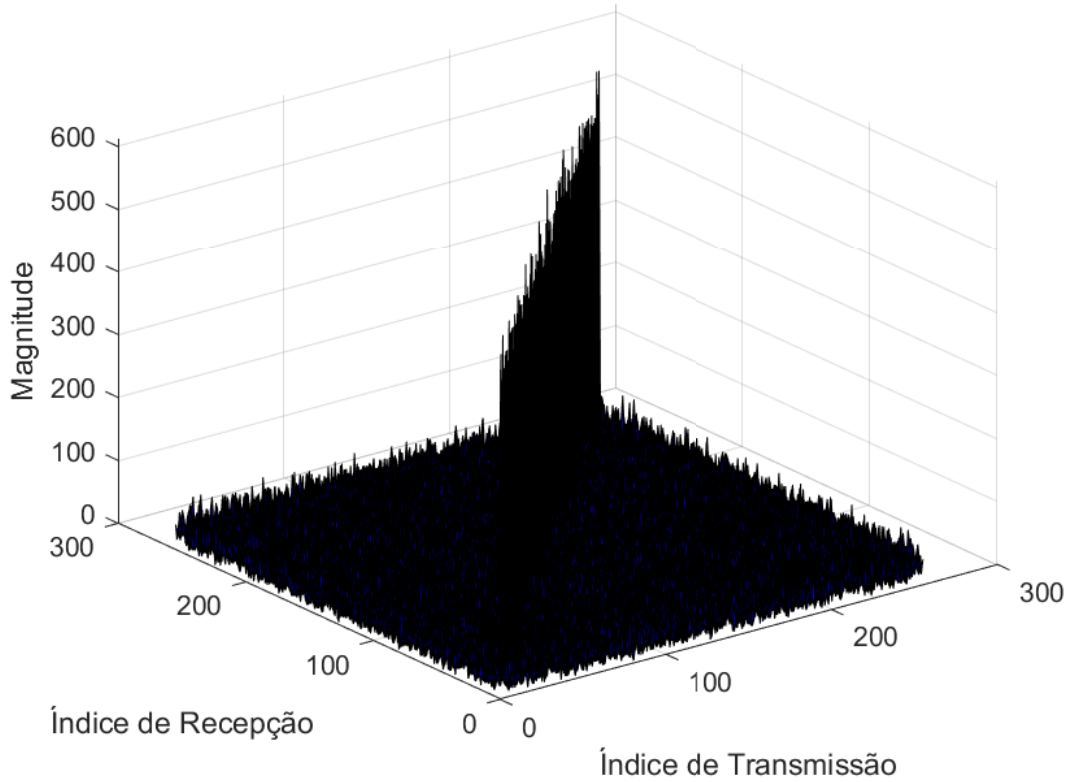
Fonte: O próprio autor

Figura 8 – Gráfico de enrijecimento do canal $\mathbf{H}^H \mathbf{H}$ para 128×128 MIMO.



Fonte: O próprio autor

Figura 9 – Gráfico de enrijecimento do canal $\mathbf{H}^H \mathbf{H}$ para 256×256 MIMO.



Fonte: O próprio autor

O enrijecimento do canal mostrado nessas figuras apresenta vantagens relevantes para o processamento de sinais em sistemas de alta dimensionalidade. Em técnicas de detecção linear a inversão de matrizes é uma etapa essencial no processamento do sinal. Para matrizes de grande dimensão, a inversão direta pode ser computacionalmente onerosa, tornando necessário o emprego de métodos eficientes, como expansões em série e aproximações determinísticas da distribuição limite dos autovalores. O endurecimento do canal possibilita a utilização dessas abordagens aproximadas com alta precisão, uma vez que a variação estatística dos coeficientes de canal se reduz significativamente conforme cresce o número de antenas na estação base.

Além da otimização do processamento matricial, o endurecimento do canal permite que algoritmos de detecção de menor complexidade alcancem desempenhos próximos aos métodos convencionais de alta precisão. Isso é ótimo em configurações de grande escala, onde a redução da carga computacional desempenha um papel fundamental na viabilização prática de sistemas massivos de múltiplas antenas.

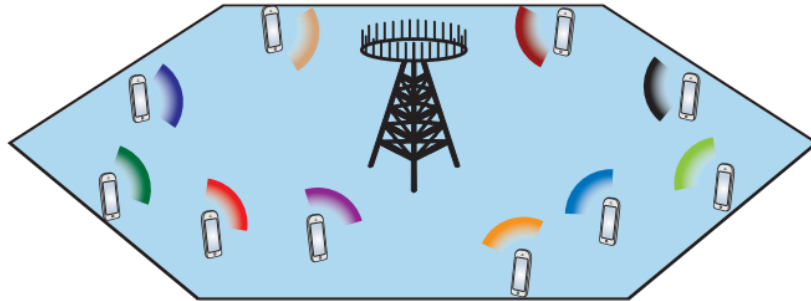
3 Técnicas de Recepção MIMO: Enlace de subida

O *uplink* é o enlace de comunicação no qual ocorre transmissão de dados de um usuário para uma estação base que a central de recepção como exemplificado pela figura 10. Esse processo é feito para realizar a troca de informações em redes sem fio, permitindo que os dispositivos enviem sinais de controle, voz, vídeo ou pacotes de dados para a infraestrutura da rede, onde serão processados.

A eficiência do *uplink* é condicionada por diversos fatores técnicos, incluindo a potência limitada dos transmissores nos dispositivos terminais, as interferências decorrentes da transmissão simultânea de múltiplos usuários e a degradação do sinal devido a fenômenos como atenuação e desvanecimento.

A utilização de múltiplas antenas proporciona ganhos no desempenho do *uplink*. Assim na estação base permite a exploração da diversidade espacial para aprimorar a relação sinal-ruído e realizar a separação dos sinais transmitidos por diferentes usuários. Além disso, a estimativa precisa do canal é um essencial para o processamento do *uplink*.

Figura 10 – Imagem demonstrativa do sistema *uplink*



Fonte: (CHIEN; BJORNSEN, 2017)

Para cada símbolo transmitido no *uplink*, o sinal recebido em banda base $\mathbf{y} \in \mathbb{C}^M$ na estação base é descrito como

$$\mathbf{y}_{\text{up}} = \sum_{k=1}^K \mathbf{h}_k \sqrt{p_k} x_k + \mathbf{n},$$

Em que \mathbf{h}_k representa o vetor de canal entre o usuário k e a estação base, enquanto p_k corresponde à potência de transmissão do usuário k . O termo x_k denota o símbolo transmitido por esse usuário, assumindo normalização de potência unitária, ou seja, $\mathbb{E}[|\mathbf{x}_k|^2] = 1$. O ruído aditivo branco Gaussiano circular (AWGN - *Additive White Gaussian Noise*) na recepção é modelado por $\mathbf{n} \sim \mathcal{CN}(0, \sigma_{UP}^2 \mathbf{I}_M)$, onde σ_{UP}^2 representa a variância do ruído no enlace ascendente.

Para uma representação mais compacta, o modelo pode ser reescrito em notação matricial, levando em conta a recepção simultânea dos símbolos transmitidos pelos K usuários:

$$\mathbf{y}_{\text{up}} = \mathbf{H}\mathbf{P}^{1/2}\mathbf{x} + \mathbf{n} \quad (3.1)$$

Nesse formato, a matriz $\mathbf{H} \in \mathbb{C}^{M \times K}$ contém os vetores de canal dos K usuários, cujas colunas correspondem a \mathbf{h}_k . A matriz diagonal $\mathbf{P} \in \mathbb{C}^{K \times K}$ representa as potências de transmissão de cada usuário, sendo definida por

$$\mathbf{P} = \text{diag}(p_1, p_2, \dots, p_K).$$

O vetor $\mathbf{x} \in \mathbb{C}^K$ contém os símbolos transmitidos pelos K usuários, sendo expresso como

$$\mathbf{x} = [x_1, x_2, \dots, x_K]^T.$$

A modelagem matricial possibilita a aplicação de esquemas de detecção linear para separação dos sinais recebidos na estação base.

Para simplificar a equação 3.1 no *uplink* é considerado \mathbf{P} unitário assim simplificando para:

$$\mathbf{y}_{\text{up1}} = \mathbf{H}\mathbf{x} + \mathbf{n} \quad (3.2)$$

A fim de melhorar a modelagem da recepção de sinais MIMO no *uplink* assume nesse sistema com K usuários em que tem N_t antenas transmissoras e tem-se N_r antenas receptoras na localizados BS. O sinal recebido na estação base, representado por \mathbf{y}_{up2} , é processado pela multiplicação com uma matriz de combinação \mathbf{W} , conforme descrito na Equação 3.3:

$$\mathbf{r} = \mathbf{W}\mathbf{y} \quad (3.3)$$

Em que a relação entre os parâmetros do sistema é descrita por:

$$\mathbf{r} = \mathbf{W}(\mathbf{H}_1\mathbf{x} + \mathbf{n})$$

Em que \mathbf{r} representa o vetor de sinais recebidos pré processados na BS, \mathbf{H}_1 é a matriz do canal de comunicação com dimensão $N_r \times N_t$. Assim com esse modelo pode-se melhor analisar o impacto do canal e do ruído na detecção dos sinais transmitidos, sendo a base para as das técnicas de recepção MIMO nas próximas seções. .

3.1 Detecção Máxima Verossimilhança (ML - *Maximum Likelihood*)

A detecção ML é um método de detecção usado em sistemas de comunicação MIMO para encontrar a sequência de símbolos transmitidos que é mais provável de ter gerado o sinal recebido, com base no princípio da máxima verossimilhança.

O objetivo da detecção ML é encontrar a sequência de símbolos transmitidos $\hat{\mathbf{x}}_{\text{ML}}$ (de tamanho $N \times 1$) que maximiza a probabilidade condicional da observação \mathbf{y} dado \mathbf{x} , ou seja, maximizar $P(\mathbf{y}|\mathbf{x})$. Isso é equivalente a encontrar o símbolo transmitido $\hat{\mathbf{x}}_{\text{ML}}$ que minimiza a distância euclidiana entre \mathbf{y} e a matriz de canal \mathbf{H} :

$$\hat{\mathbf{x}}_{\text{ML}} = \arg \max_{\mathbf{x}} P(\mathbf{y}|\mathbf{x}) = \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2$$

Em que $\|\cdot\|$ denota a norma euclidiana (ou norma-2).

Esta é uma busca exaustiva que avalia todas as possíveis combinações de símbolos transmitidos. A solução ML é encontrada quando a norma euclidiana entre o sinal recebido \mathbf{y} e a matriz de canal $\mathbf{H}\mathbf{x}$ é minimizada.

A detecção ML é uma técnica ótima em termos de máxima verossimilhança, mas é computacionalmente intensiva em sistemas MIMO com alta dimensionalidade de antenas de transmissão e recepção, não sendo exequível em sistemas de massivo MIMO. Assim, em sistemas práticos, é necessário utilizar melhores técnicas de detecção.

3.1.1 Simulação ML

Para compreender e visualizar o funcionamento do critério de ML, foi simulado a transmissão de 100 símbolos modulados em 4-QAM em um canal sujeito à adição de ruído AWGN com SNR de 15.

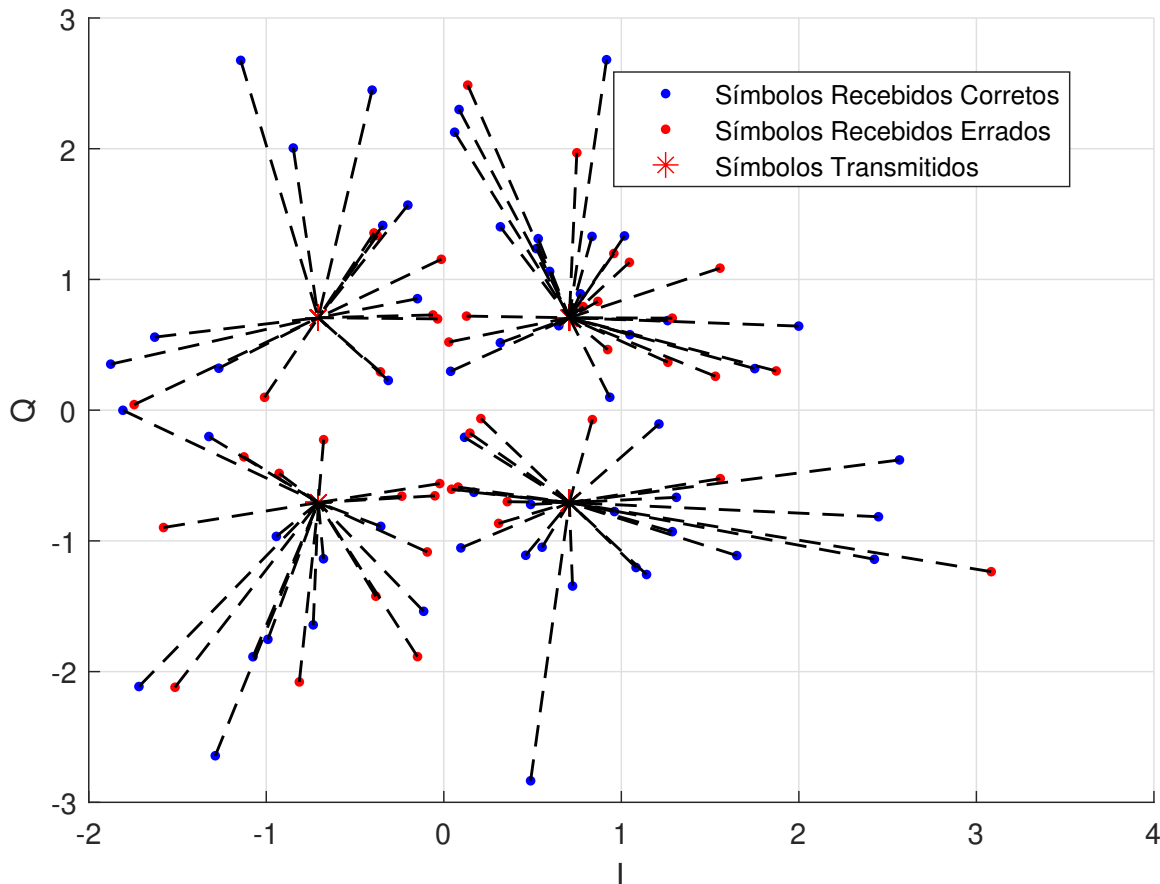
A simulação gera os símbolos 4-QAM, seguida pela introdução do efeito do canal e adição do ruído AWGN considerando uma relação sinal-ruído de 15. Os símbolos transmitidos foram então afetados pelo ruído e pelo efeito do canal, resultando em uma perturbação que desloca os pontos da constelação em relação às suas posições ideais.

A recuperação dos símbolos transmitidos foi realizada a partir da identificação dos pontos mais próximos dentro da constelação original. Para cada símbolo recebido, foi efetuada a comparação com os símbolos possíveis na constelação 4-QAM, sendo selecionado aquele cuja distância euclidiana era mínima. Assim o receptor consegue associar os símbolos distorcidos às suas posições mais prováveis na constelação.

O gráfico de constelação apresentado na Figura 11 ilustra esse processo. Os símbolos transmitidos são indicados pela estrela vermelha, enquanto os símbolos recebidos corretamente

mente, aparecem como ponto em azul. E por fim os símbolos recebidos erroneamente são apresentados com os pontos em vermelho. As linhas pontilhadas conectam cada símbolo recebido ao símbolo de transmissão mais próximo, demonstrando a decisão do receptor ao reconstruir a informação original.

Figura 11 – Representação do funcionamento do algoritmo de ML.



Fonte: O próprio autor

A Figura 11 evidencia a forma como o algoritmo de ML processa os símbolos recebidos e os associa aos símbolos mais prováveis na constelação, permitindo a reconstrução dos dados transmitidos. No entanto, conforme discutido anteriormente, sua aplicação direta em sistemas MIMO massivos não é viável do ponto de vista prático, devido à complexidade computacional exponencial associada à busca exaustiva pela melhor estimativa dos símbolos transmitidos. Apesar dessa limitação, compreender o funcionamento do ML é essencial, pois ele estabelece um limite teórico de desempenho e fundamenta o desenvolvimento de técnicas mais avançadas, capazes de oferecer um compromisso entre desempenho e viabilidade computacional.

O uso exclusivo do critério de Máxima Verossimilhança, embora eficiente para cenários de baixa dimensionalidade, torna-se inviável para sistemas com um grande número de antenas, uma vez que a complexidade da detecção cresce exponencialmente com o número

de fluxos espaciais e símbolos modulados. A busca exaustiva em grandes dimensões demanda recursos computacionais impraticáveis, tornando necessário o desenvolvimento de algoritmos alternativos que aproximem seu desempenho com menor custo computacional.

Ainda que a abordagem ML não seja implementável diretamente em MIMO massivo, sua compreensão permite avaliar o desempenho de algoritmos otimizados e justifica a necessidade de técnicas mais eficientes. A comparação com esse critério fornece um referencial teórico para mensurar a qualidade dos detectores utilizados em cenários práticos, sendo um elemento essencial na análise de desempenho e no aprimoramento das estratégias de detecção em sistemas de comunicação de grande escala.

3.2 Detecção Forçamento a Zero (ZF - Zero-Forcing): Uplink

A técnica de ZF é uma técnica de pré-processamento usada em sistemas de comunicação MIMO para mitigar as interferências e melhorar a detecção dos símbolos no receptor. Tornando o sinal recebido o mais limpo possível. A técnica é chamada de "Zero-Forcing" porque ela "força" a matriz de canal a anular a interferência.

Em um sistema MIMO com N_t antenas de transmissão e N_r antenas de recepção. O vetor de sinal transmitido é representado por \mathbf{x} , que é um vetor coluna de tamanho N_t (N transmissores). O vetor de sinal recebido é representado por \mathbf{y} , que é um vetor coluna de tamanho N_t (N_t antenas receptoras). O canal de comunicação entre as antenas de transmissão e as antenas de recepção é representado por uma matriz \mathbf{H} . Por fim o \mathbf{z} é o sinal ruído multiplicado pela matriz de combinação \mathbf{W} .

$$\mathbf{r}_{ZF} = \mathbf{W}\mathbf{H}\mathbf{x} + \mathbf{z} \quad (3.4)$$

Assim o objetivo do ZF é encontrar uma matriz de pré-processamento \mathbf{W} de tamanho $M \times N$ de modo que, quando o sinal transmitido \mathbf{x} for multiplicado por \mathbf{W} , a interferência seja eliminada, deixando apenas o sinal desejado. Matematicamente, a matriz \mathbf{W} é encontrada da seguinte forma:

$$\mathbf{W}_{ZF} = \mathbf{H}^{-1}$$

Em que $(\cdot)^{-1}$ denota a matriz inversa de \mathbf{H} . No entanto, em muitos casos práticos, a matriz \mathbf{H} pode não ser de tamanho completo ou pode não ser de posto completo. Nesse caso, para diminuir a complexidade é usado a **pseudo-inversa** (também conhecida como inversa generalizada) para calcular \mathbf{W} .

$$\mathbf{W}_{ZF} = \mathbf{H}^+ = (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H$$

O sinal recebido \mathbf{y} é pré-processado multiplicando-o pela matriz de combinação \mathbf{W}_{ZF} :

$$\hat{\mathbf{x}}_{\text{ZF}} = \mathbf{W}_{\text{ZF}}\mathbf{y}$$

Em que $\hat{\mathbf{x}}_{\text{ZF}}$ é o sinal pré-processado, que deve ser um vetor de tamanho $N \times 1$.

Assim a ideia principal do ZF é eliminar a interferência entre as antenas de transmissão, tornando o sinal recebido independente dessas interferências. No entanto, o ZF pode ser sensível ao ruído devido à inversão da matriz \mathbf{H} e pode não ser eficaz quando a relação sinal-ruído é baixa. Portanto, em cenários mais complexos, outras técnicas de processamento de sinal, como a matriz de combinação Mínima de Erros Quadráticos (MMSE) ou técnicas de cancelamento de interferência, podem ser mais adequadas.

3.3 Detecção Erro Quadrático Médio Mínimo - (MMSE - *Minimum Mean Square Error*): *Uplink*

A detecção MMSE é uma técnica usada em sistemas de comunicação MIMO que visa minimizar o erro médio quadrático e leva em consideração a qualidade do canal.

O objetivo da detecção MMSE é encontrar uma estimativa dos símbolos transmitidos, minimizando o erro médio quadrático (MMSE - *Minimum Mean Square Error*) entre os símbolos estimados e os símbolos reais.

Como empregado em (HAMPTON, 2014) A estimativa MMSE dos símbolos transmitidos $\hat{\mathbf{x}}_{\text{MMSE}}$ (de tamanho N_1) é dada por:

$$\hat{\mathbf{x}}_{\text{MMSE}} = \mathbf{W}_{\text{MMSE}}\mathbf{y}$$

Em que \mathbf{W}_{MMSE} é a matriz de de combinação MMSE definida como:

$$\mathbf{W}_{\text{MMSE}} = (\mathbf{H}^H\mathbf{H} + \frac{\sigma^2}{\rho}\mathbf{I})^{-1}\mathbf{H}^H \quad (3.5)$$

A matriz \mathbf{H} como já visto anteriormente representa o canal de comunicação, possuindo dimensões $N_r \times N_t$, em que N_r e N_t correspondem, respectivamente, ao número de antenas receptoras e transmissoras. A matriz identidade de ordem \mathbf{N} , denotada por $\mathbf{I}_{\mathbf{N}}$. A potência média do sinal transmitido, representada por ρ e por ultimo o σ^2 representa a potência do ruído AWGN.

3.3.1 Comparação MMSE com a Detecção ML

A principal diferença em relação à detecção ML é que a detecção MMSE leva em consideração a qualidade do canal, o que pode resultar em um desempenho melhor e exequível em massivo MIMO.

A detecção MMSE difere da detecção ML no sentido de que a MMSE usa uma abordagem com o objetivo de minimizar o erro médio quadrático e leva em consideração a qualidade do canal para estimar os símbolos transmitidos. A detecção ML, por outro lado, busca encontrar a sequência de símbolos que maximiza a probabilidade condicional da observação.

Enquanto a detecção ML é ótima em termos de máxima verossimilhança, ela pode ser computacionalmente intensiva e impraticável em sistemas MIMO com alta dimensionalidade. A detecção MMSE oferece um compromisso entre desempenho e complexidade, tornando-a uma escolha viável em muitos cenários de comunicação MIMO.

3.4 Cancelamento de Interferência (CI)

O Cancelamento de Interferência (IC ou CI (*Interference Cancellation*)) é uma ótima técnica que permite aprimorar a capacidade de recepção em ambientes onde a interferência é um desafio significativo. Sua aplicação pode melhorar drasticamente a qualidade do sinal e, conseqüentemente, o desempenho do sistema de comunicação.

Assim um sistema MIMO composto por N_t antenas transmissoras e N_r antenas receptoras, o processo de cancelamento de interferência ocorre de maneira iterativa, refinando a estimativa dos símbolos transmitidos a partir do sinal recebido. Inicialmente, a estimativa do primeiro símbolo transmitido é realizada com base no sinal recebido e na estrutura do canal. Essa estimativa é utilizada para mitigar a interferência que esse símbolo causa nas demais antenas receptoras, permitindo que os sinais subsequentes sejam processados com menor impacto de interferência interusuário.

A remoção da interferência é feita subtraindo-se a contribuição do símbolo previamente estimado dos sinais recebidos nas antenas restantes. Esse procedimento é repetido para cada antena transmissora, refinando progressivamente a detecção dos símbolos transmitidos. A cada iteração, um novo símbolo é estimado e a interferência causada por ele nos sinais restantes é eliminada, garantindo que o processo de detecção evolua de forma eficiente ao longo das antenas receptoras.

Ao término desse processo iterativo, obtém-se um conjunto de estimativas dos símbolos transmitidos em cada antena transmissora, onde a interferência entre os fluxos espaciais foi mitigada sucessivamente. Esse método melhora a recuperação dos dados transmitidos ao reduzir o impacto da interferência interusuário, tornando a detecção mais robusta, especialmente em cenários de alta densidade espectral.

3.4.1 Desvantagens do IC

Embora o IC melhore a qualidade da detecção, uma desvantagem significativa é a propagação de erros. Se ocorrer um erro na estimativa de símbolos para a última antena

de transmissão, esse erro pode se propagar para as estimativas das antenas anteriores durante as iterações. Isso pode resultar em uma degradação significativa no desempenho de detecção, especialmente em cenários com alta interferência ou condições adversas de canal.

3.5 Detecção ZF-IC

O detector ZF-IC é uma variação da técnica de ZF usada em sistemas de comunicação MIMO para mitigar a interferência entre as antenas de transmissão e melhorar a detecção dos símbolos no receptor. O ZF-IC leva em consideração a interferência entre as antenas de transmissão, fazendo de forma iterativa.

A técnica ZF-IC começa com uma transformação semelhante à do ZF. O objetivo é encontrar uma matriz de pré-processamento \mathbf{W}_{ZFIC} de modo que, quando o sinal transmitido \mathbf{x} for multiplicado por \mathbf{W} , a interferência seja eliminada, deixando apenas o sinal desejado. A matriz \mathbf{W}_{ZFIC} é encontrada usando a pseudo-inversa de \mathbf{H} , assim como no ZF.

O ZF-IC representa uma evolução do Zero Forcing, levando em consideração a interferência entre as antenas de transmissão e realizando sua eliminação de forma iterativa. Enquanto o ZF busca a eliminação direta da interferência, o ZF-IC adota uma melhor abordagem, aprimorando o desempenho, especialmente em cenários onde a interferência exerce um papel significativo. Ao empregar uma técnica iterativa, o ZF-IC se destaca ao lidar de maneira mais eficiente com ambientes complexos, melhorando a qualidade da detecção de símbolos no receptor e proporcionando ganhos em sistemas MIMO.

3.6 Detecção MMSE-IC

O detector MMSE-IC é uma variação da técnica MMSE utilizada em sistemas de comunicação MIMO para atenuar a interferência entre as antenas de transmissão e aprimorar a detecção dos símbolos no receptor. Assim como o ZF-IC, o MMSE-IC considera a interferência entre as antenas de transmissão e adota uma abordagem iterativa.

A técnica MMSE-IC inicia com uma abordagem similar à do MMSE. Seu objetivo é determinar uma matriz de pré-processamento \mathbf{W}_{MMSE} . Ao multiplicar o sinal transmitido \mathbf{x} por \mathbf{W} , busca-se não apenas eliminar a interferência, mas também minimizar o erro médio quadrático entre o símbolo estimado e o símbolo verdadeiro. A matriz \mathbf{W}_{MMSE} é calculada utilizando uma combinação ponderada da pseudo-inversa de \mathbf{H} e a matriz de covariância do ruído.

A principal diferença em relação ao MMSE é o cancelamento iterativo de interferência, que melhora a capacidade do detector de lidar com interferência entre antenas de transmissão.

Assim o MMSE-IC representa um avanço em relação ao MMSE convencional, lidando de forma mais sofisticada com a interferência entre antenas. Enquanto o MMSE se concentra na minimização do erro médio quadrático global, o MMSE-IC, por meio de sua abordagem iterativa, aprimora a detecção em cenários com interferência significativa, contribuindo para um desempenho melhorado nos sistemas MIMO ao mitigar os efeitos adversos dos canais e da interferência.

3.6.1 Comparação entre o MMSE-IC e MMSE

A principal diferença entre o MMSE-IC e o MMSE é o cancelamento iterativo de interferência. Enquanto o MMSE produz uma única estimativa linear dos símbolos transmitidos, o MMSE-IC leva em consideração a interferência iterativamente, o que pode melhorar o desempenho em cenários com interferência significativa entre as antenas de transmissão.

No entanto, o MMSE-IC é mais computacionalmente complexo devido à iteração. Portanto, a escolha entre MMSE e MMSE-IC depende das necessidades específicas do sistema e das restrições de complexidade.

3.7 Simulação: Comparação entre ZF, ZF-IC, MMSE, MMSE-IC

3.7.1 Parâmetros da Simulação

A taxa de erro por bit (BER - *Bit Error Rate*) é analisada em função da SNR, variando no intervalo de 0 a 20 dB, permitindo avaliar o desempenho da detecção sob diferentes condições de canal. A modulação utilizada na simulação corresponde à modulação BPSK, caracterizada por uma constelação de ordem unitária. O número de símbolos transmitidos é definido em conformidade com o número de antenas transmissoras, garantindo que o sistema opere de forma consistente com a estrutura de um sistema MIMO. A BER é calculada realizando a comparação entre os bits transmitidos e os bits recebidos.

3.7.1.1 Ruído

A modelagem do ruído na antena de recepção segue a abordagem descrita em (VISWANATHAN, 2020), considerando um componente complexo cujas partes real e imaginária são amostras de uma distribuição normal de média zero e variância unitária. O sinal é normalizado ao ser dividido por um fator $\sqrt{2SNR}$, garantindo a consistência na escala de potência. Para levar em conta a influência do número de antenas transmissoras no desempenho do sistema, o termo resultante é ainda ponderado pelo fator $\sqrt{N_t}$, ajustando assim o impacto do ruído em função da dimensionalidade do sistema. Como a potência do sinal é normalizado, para calcular o ruído basta fazer o inverso do SNR.

3.7.1.2 Simulação do IC

O cancelamento de interferência é um processo iterativo em que os símbolos transmitidos são estimados e progressivamente refinados para mitigar os efeitos da interferência interusuário. Inicialmente, é realizada a estimativa do símbolo transmitido na primeira antena de transmissão, baseando-se no sinal recebido e comparando-o com os possíveis símbolos que poderiam ter sido enviados. A estimativa obtida é então utilizada para ajustar os sinais recebidos nas demais antenas, removendo a interferência causada pelo primeiro símbolo detectado.

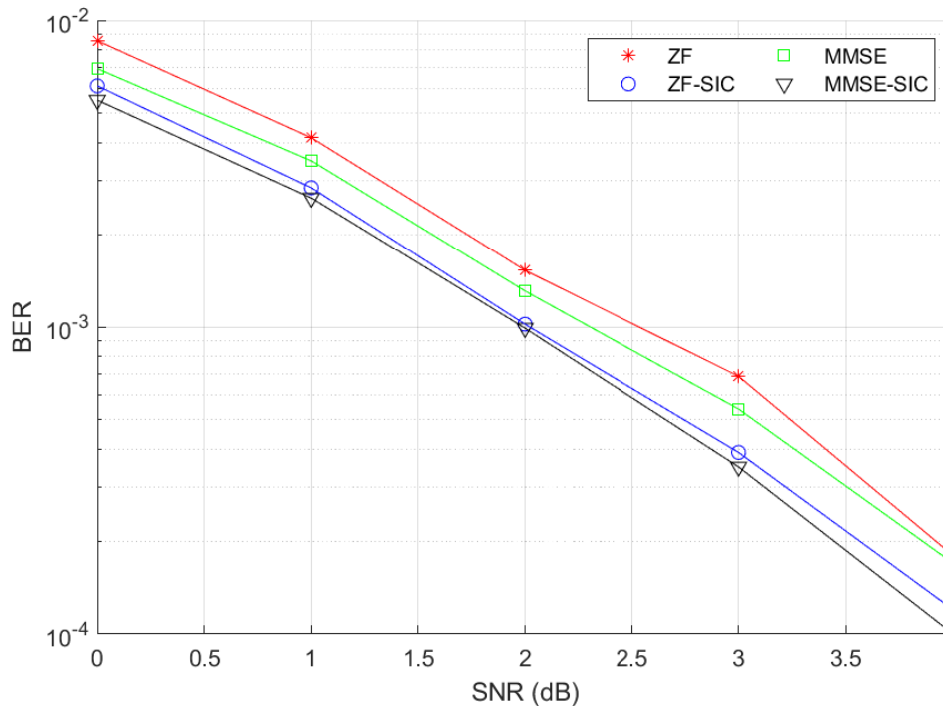
Com a interferência parcialmente suprimida, o algoritmo prossegue estimando os símbolos transmitidos pelas antenas subsequentes, aplicando a mesma abordagem iterativa até que todos os símbolos tenham sido processados. Esse refinamento contínuo permite que os sinais sejam reconstruídos com maior precisão, garantindo um melhor desempenho do sistema na separação dos fluxos de transmissão. O resultado final consiste em um conjunto de estimativas dos símbolos transmitidos, em que a interferência foi cancelada.

lada progressivamente ao longo do processamento, aprimorando a qualidade da detecção e reduzindo a degradação causada pela superposição de sinais em canais compartilhados.

3.7.2 Resultados da Simulação: Comparação entre ZF, ZF-IC, MMSE, MMSE-IC

As simulações apresentadas a seguir comparam o desempenho dos detectores ZF, ZF-IC, MMSE e MMSE-IC no *uplink* de um sistema MIMO Massivo. Os resultados ilustram a influência da relação entre o número de antenas transmissoras e receptoras na BER em função da SNR. Diferentes configurações de antenas são analisadas, evidenciando o impacto do aumento da diversidade espacial e do uso de técnicas de cancelamento de interferência. As Figuras 12, 13, 14, e 15 apresentam os resultados obtidos.

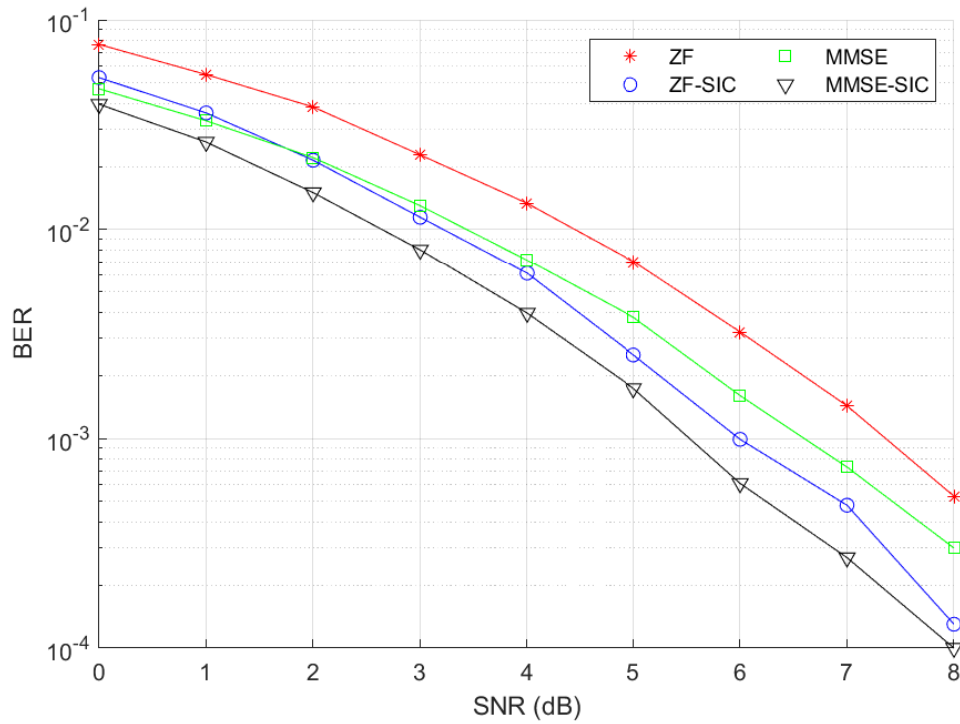
Figura 12 – Resultado 4 antenas transmissoras e 16 antenas receptoras para comparação entre ZF, ZF-IC, MMSE, MMSE-IC no *uplink*



Fonte: O próprio autor

A Figura 12 exibe o impacto do aumento do número de antenas receptoras para 16. Esse crescimento melhora a diversidade espacial, resultando em uma detecção mais eficiente e, conseqüentemente, em uma redução na BER para um mesmo intervalo de SNR. A maior quantidade de elementos de recepção proporciona ganhos significativos para os algoritmos IC, uma vez que eles exploram a diversidade para mitigar interferências. Nota-se uma maior separação entre as curvas dos métodos tradicionais e aqueles baseados em SIC, evidenciando a importância da diversidade espacial na melhoria da detecção.

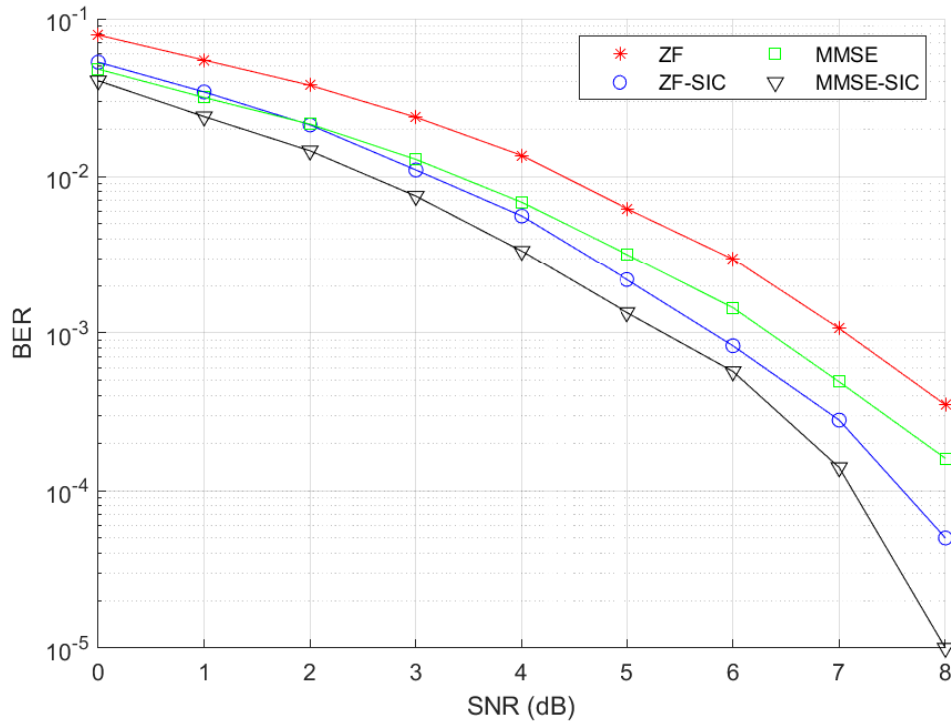
Figura 13 – Resultado 16 antenas transmissoras e 32 antenas receptoras para comparação entre ZF, ZF-IC, MMSE, MMSE-IC no *uplink*



Fonte: O próprio autor

A Figura 13 apresenta um novo incremento no número de antenas receptoras, agora para 32. Esse aumento proporciona uma melhora significativa na separação dos sinais transmitidos, permitindo que a BER reduza para um mesmo intervalo de SNR. A diversidade espacial ampliada favorece principalmente os métodos MMSE-IC e ZF-IC, que aproveitam melhor a estrutura do canal para minimizar interferências. O desempenho geral do sistema melhora, destacando o impacto positivo da relação entre elementos de recepção e transmissão.

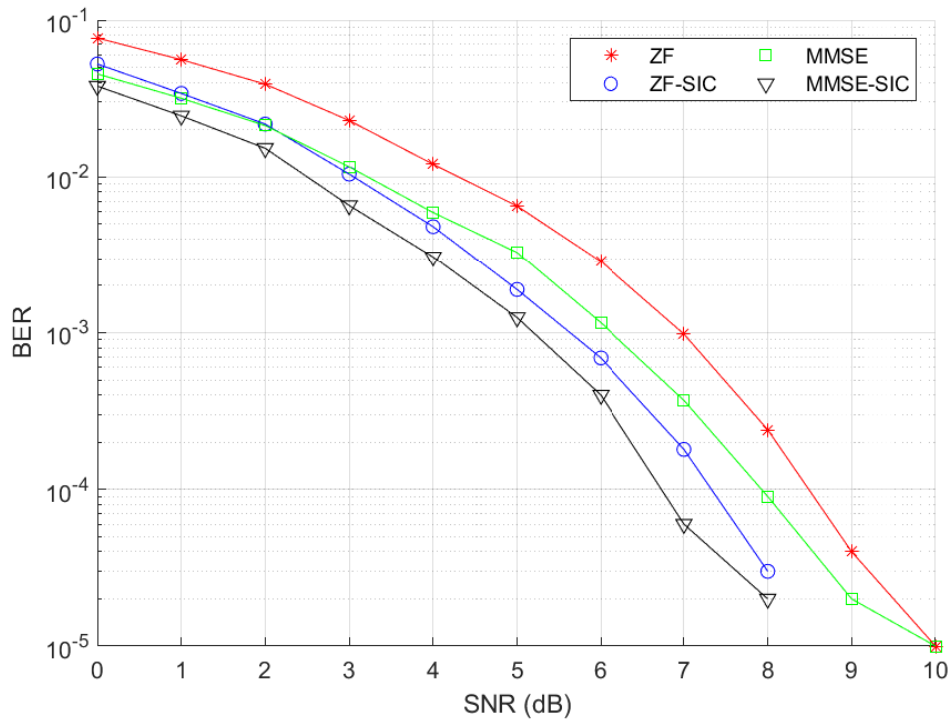
Figura 14 – Resultado 32 antenas transmissoras e 64 antenas receptoras para comparação entre ZF, ZF-IC, MMSE, MMSE-IC no *uplink*



Fonte: O próprio autor

A Figura 14 representa a transição para o regime de MIMO Massivo, onde há um número significativamente maior de antenas receptoras. O sistema exibe um comportamento mais próximo do ideal, com uma redução mais acentuada da BER à medida que a SNR cresce. A separação espacial dos fluxos de dados melhora consideravelmente, favorecendo os esquemas baseados em SIC. Nota-se que, enquanto ZF e MMSE tradicionais ainda apresentam limitações em baixos valores de SNR, suas versões aprimoradas demonstram uma convergência mais rápida para baixos níveis de BER.

Figura 15 – Resultado 64 antenas transmissoras e 128 antenas receptoras para comparação entre ZF, ZF-IC, MMSE, MMSE-IC no *uplink*



Fonte: O próprio autor

Por fim, a Figura 15 apresenta um cenário com 64 antenas transmissoras e 128 antenas receptoras, consolidando-se no regime de MIMO Massivo. O aumento expressivo no número de transmissores impõe desafios na recuperação do sinal devido à interferência interusuário intensificada, exigindo algoritmos mais sofisticados para mitigação eficiente. Por outro lado, a presença de 128 antenas receptoras compensa esse efeito ao proporcionar um nível elevado de diversidade espacial. Os métodos baseados em SIC continuam a apresentar vantagens significativas, especialmente para baixos valores de SNR, enquanto para valores elevados de SNR, todas as técnicas convergem para uma BER reduzida.

3.8 Detecção Lenstra-Lenstra-Lovász (LLL): *Uplink*

Como abordado no livro (CHO et al., 2010) e nos artigos (AGRELL T. ERIKSSON; ZEGGER, 2002), (WINDPASSINGER; FISCHER, 2003) e (LARSSON, 2009) a detecção auxiliada por redução de redes em sistemas MIMO tem como principal objetivo minimizar o aumento do ruído inerente aos métodos lineares de detecção, como Zero-Forcing e MMSE. Esse efeito de amplificação do ruído se torna crítico especialmente quando a matriz do canal possui um número de condição elevado. A técnica de redução de redes (LR - *Lattice Reduction*), baseada no algoritmo de LLL, permite transformar a matriz do canal em uma base melhor condicionada, reduzindo significativamente os impactos negativos no desempenho da detecção.

Seja o modelo do sistema MIMO descrito por:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \quad (3.6)$$

Em que \mathbf{y} é o vetor recebido, \mathbf{H} representa a matriz do canal, \mathbf{x} é o vetor de símbolos transmitidos e \mathbf{z} é o ruído aditivo gaussiano branco. Para melhorar a condição da matriz do canal, emprega-se a decomposição QR:

$$\mathbf{H} = \mathbf{Q}\mathbf{R} \quad (3.7)$$

em que \mathbf{Q} é uma matriz unitária e \mathbf{R} é uma matriz triangular superior. Multiplicando ambos os lados da equação do sistema por \mathbf{Q}^H , obtém-se:

$$\tilde{\mathbf{y}} = \mathbf{Q}^H\mathbf{y} = \mathbf{R}\mathbf{x} + \tilde{\mathbf{z}} \quad (3.8)$$

Em que $\tilde{\mathbf{z}} = \mathbf{Q}^H\mathbf{z}$ preserva as características estatísticas do ruído original. A condição da matriz \mathbf{H} é equivalente à de \mathbf{R} , ou seja, a decomposição QR por si só não reduz o número de condição do sistema. No entanto, a técnica de redução de redes aplicada a \mathbf{R} resulta em uma base melhor condicionada, minimizando a amplificação do ruído durante a filtragem linear.

A redução de base LLL é um método que permite melhorar a ortogonalidade das colunas da matriz do canal, reduzindo o número de condicionamento da matriz triangular superior obtida pela decomposição QR. Esse procedimento é iterativo e baseia-se na modificação e troca de colunas de forma a tornar a matriz mais bem condicionada para algoritmos de detecção. O processo consiste em três etapas principais, onde sucessivamente são reduzidos os comprimentos das colunas, aplicadas trocas e rotações de Givens para preservar a estrutura triangular superior da matriz.

A seguir, detalham-se os passos desse processo. tendo como exemplo R e T inicial como

$$R = [r_1, r_2, r_3, r_4] \quad \text{e} \quad T = [t_1, t_2, t_3, t_4] \quad (3.9)$$

em que $\{\mathbf{r}_i\}_{i=1}^4$ e i é a i -ésima coluna de R , e $\{\mathbf{t}_i\}_{i=1}^4$ e i é a i -ésima coluna de T que é a matriz identidade,

3.8.0.1 Passo 1: Redução do Comprimento do Segundo Vetor Coluna r_2

tendo $\mu_{1,2}$ como:

$$\mu_{1,2} = \left\lceil \frac{r_{1,2}}{r_{1,1}} \right\rceil \quad (3.10)$$

Utilizando $\mu_{1,2}$ na Equação (3.10), o segundo vetor coluna das matrizes \mathbf{R} e \mathbf{T} é modificado como:

$$r_2 \leftarrow r_2 - \mu_{1,2}r_1 \quad (3.11)$$

$$t_2 \leftarrow t_2 - \mu_{1,2}t_1 \quad (3.12)$$

Se $\mu_{1,2} = 0$ nada é alterado. Caso contrário, se a matriz modificada \mathbf{R} não satisfizer a seguinte condição:

$$\delta r_{1,1}^2 > r_{1,2}^2 + r_{2,2}^2, \quad \frac{1}{4} < \delta \leq 1 \quad (3.13)$$

Assim os vetores coluna correspondentes de \mathbf{R} e \mathbf{T} são trocados:

$$\mathbf{R} = [r_2, r_1, r_3, r_4] \quad \text{e} \quad \mathbf{T} = [t_2, t_1, t_3, t_4] \quad (3.14)$$

A matriz \mathbf{R} agora não é triangular superior, exigindo a seguinte rotação de Givens para manter essa propriedade:

$$\mathbf{R} \leftarrow \mathbf{\Theta}_1 \mathbf{R} \quad (3.15)$$

A modificação é refletida na matriz unitária \mathbf{Q} :

$$\mathbf{Q} \leftarrow \mathbf{Q} \mathbf{\Theta}_1^T \quad (3.16)$$

Em que $\mathbf{\Theta}_1$ é uma matriz de rotação ortogonal definida como:

$$\mathbf{\Theta}_1 = \begin{bmatrix} a_1 & b_1 & 0 & 0 \\ -b_1 & a_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.17)$$

com a_1 e b_1 definidos como:

$$a_1 = \frac{r_{1,1}}{\sqrt{r_{1,1}^2 + r_{2,1}^2}}, \quad b_1 = \frac{r_{2,1}}{\sqrt{r_{1,1}^2 + r_{2,1}^2}} \quad (3.18)$$

O processo se repete até que a condição na Equação (3.13) seja satisfeita. Após isso, as operações das Equações (3.15) e (3.16) são ignoradas, e passa-se para o próximo passo.

3.8.0.2 Passo 2: Redução do Comprimento do Terceiro Vetor Coluna r_3

Defina $\mu_{2,3}$ como:

$$\mu_{2,3} = \left\lfloor \frac{r_{2,3}}{r_{2,2}} \right\rfloor \quad (3.19)$$

Utilizando $\mu_{2,3}$, os terceiros vetores coluna de \mathbf{R} e \mathbf{T} são modificados como:

$$r_3 \leftarrow r_3 - \mu_{2,3}r_2 \quad (3.20)$$

$$t_3 \leftarrow t_3 - \mu_{2,3}t_2 \quad (3.21)$$

Defina $\mu_{1,3}$ como:

$$\mu_{1,3} = \left\lfloor \frac{r_{1,3}}{r_{1,1}} \right\rfloor \quad (3.22)$$

Utilizando $\mu_{1,3}$, os terceiros vetores coluna de \mathbf{R} e \mathbf{T} são novamente modificados:

$$r_3 \leftarrow r_3 - \mu_{1,3}r_1 \quad (3.23)$$

$$t_3 \leftarrow t_3 - \mu_{1,3}t_1 \quad (3.24)$$

A seguinte condição é verificada:

$$\delta r_{2,2}^2 > r_{2,3}^2 + r_{3,3}^2 \quad (3.25)$$

Se a matriz \mathbf{R} modificada pela Equação (3.23) não satisfizer essa condição, os vetores coluna segundo e terceiro de \mathbf{R} e \mathbf{T} são trocados:

$$\mathbf{R} = [r_1, r_3, r_2, r_4] \quad \text{e} \quad \mathbf{T} = [t_1, t_3, t_2, t_4] \quad (3.26)$$

A matriz \mathbf{R} é ajustada via rotação de Givens novamente:

$$\mathbf{R} \leftarrow \Theta_2 \mathbf{R} \quad (3.27)$$

$$\Theta_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & a_2 & b_2 & 0 \\ 0 & -b_2 & a_2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.28)$$

Com os coeficientes:

$$a_2 = \frac{r_{2,2}}{\sqrt{r_{2,2}^2 + r_{3,2}^2}}, \quad b_2 = \frac{r_{3,2}}{\sqrt{r_{2,2}^2 + r_{3,2}^2}} \quad (3.29)$$

A redução do comprimento do segundo vetor coluna já foi efetuada no primeiro passo. Entretanto, caso a condição definida na Equação (3.25) não seja satisfeita, a troca entre as colunas segunda e terceira se torna necessária. Como consequência dessa permutação, o novo vetor segunda coluna precisa ser ajustado novamente, exigindo a repetição do processo de redução. Dessa forma, o algoritmo retorna ao Passo 1 e reexecuta a modificação até que a matriz resultante atenda à condição imposta. Assim que todas as exigências sejam cumpridas, o procedimento avança para o Passo 3.

3.8.0.3 Passo 3: Redução do Comprimento do Quarto Vetor Coluna r_4

Defina $\mu_{3,4}$:

$$\mu_{3,4} = \left\lfloor \frac{r_{3,4}}{r_{3,3}} \right\rfloor \quad (3.30)$$

Utilizando $\mu_{3,4}$, os quarta coluna de vetores de \mathbf{R} e \mathbf{T} são modificados:

$$r_4 \leftarrow r_4 - \mu_{3,4} r_3 \quad (3.31)$$

$$t_4 \leftarrow t_4 - \mu_{3,4} t_3 \quad (3.32)$$

Finalmente, define-se $\mu_{1,4}$:

$$\mu_{1,4} = \left\lfloor \frac{r_{1,4}}{r_{1,1}} \right\rfloor \quad (3.33)$$

Se a condição

$$\delta r_{3,3}^2 > r_{3,4}^2 + r_{4,4}^2 \quad (3.34)$$

for satisfeita, o procedimento de redução de rede LLL está completo. Caso contrário, os vetores coluna terceiro e quarto de \mathbf{R} e \mathbf{T} são trocados:

$$\mathbf{R} = [r_1, r_2, r_4, r_3] \quad \text{e} \quad \mathbf{T} = [t_1, t_2, t_4, t_3] \quad (3.35)$$

A matriz \mathbf{R} é convertida novamente em triangular superior via rotação de Givens:

$$\mathbf{R} \leftarrow \Theta_3 \mathbf{R} \quad (3.36)$$

$$\Theta_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \alpha_3 & \beta_3 \\ 0 & 0 & -\beta_3 & \alpha_3 \end{bmatrix} \quad (3.37)$$

com os coeficientes:

$$\alpha_3 = \frac{r_{3,3}}{\sqrt{r_{3,3}^2 + r_{4,3}^2}}, \quad \beta_3 = \frac{r_{4,3}}{\sqrt{r_{3,3}^2 + r_{4,3}^2}} \quad (3.38)$$

Assim o algoritmo de LLL opera iterativamente sobre a matriz \mathbf{R} para encontrar uma matriz \mathbf{R}_{LLL} associada a uma transformação unimodular $\mathbf{T}_{LLL} = I$, tal que:

$$\bar{\mathbf{y}} = \bar{\mathbf{H}}\bar{\mathbf{x}} + \bar{\mathbf{z}} \quad (3.39)$$

Resulta em:

$$= \mathbf{Q}_{LLL} \mathbf{R}_{LLL} \mathbf{T}_{LLL}^{-1} \bar{\mathbf{x}} + \bar{\mathbf{z}} \quad (3.40)$$

A nova equação do sistema, após a aplicação da redução de redes, assim pode-se é reescrever a equação 3.8 como:

$$\tilde{\mathbf{y}} = \mathbf{R}_{LLL} \tilde{\mathbf{x}}_{LLL} + \tilde{\mathbf{z}} \quad (3.41)$$

Em que $\tilde{\mathbf{x}}_{LLL} = \mathbf{T}_{LLL}^{-1} \mathbf{x}$. Como a matriz \mathbf{T}_{LLL} é unimodular, suas entradas são inteiras e sua inversa também contém apenas valores inteiros. Isso garante que, mesmo após a transformação, os valores dos símbolos permanecem pertencentes ao mesmo conjunto de constelação, possibilitando um processo de detecção mais robusto.

A estimativa do vetor transmitido utilizando a detecção MMSE sobre o sistema bem condicionado é dada por:

$$\hat{\mathbf{x}}_{MMSE} = (\mathbf{R}_{LLL}^H \mathbf{R}_{LLL} + \sigma_z^2 \mathbf{I})^{-1} \mathbf{R}_{LLL}^H \tilde{\mathbf{y}} \quad (3.42)$$

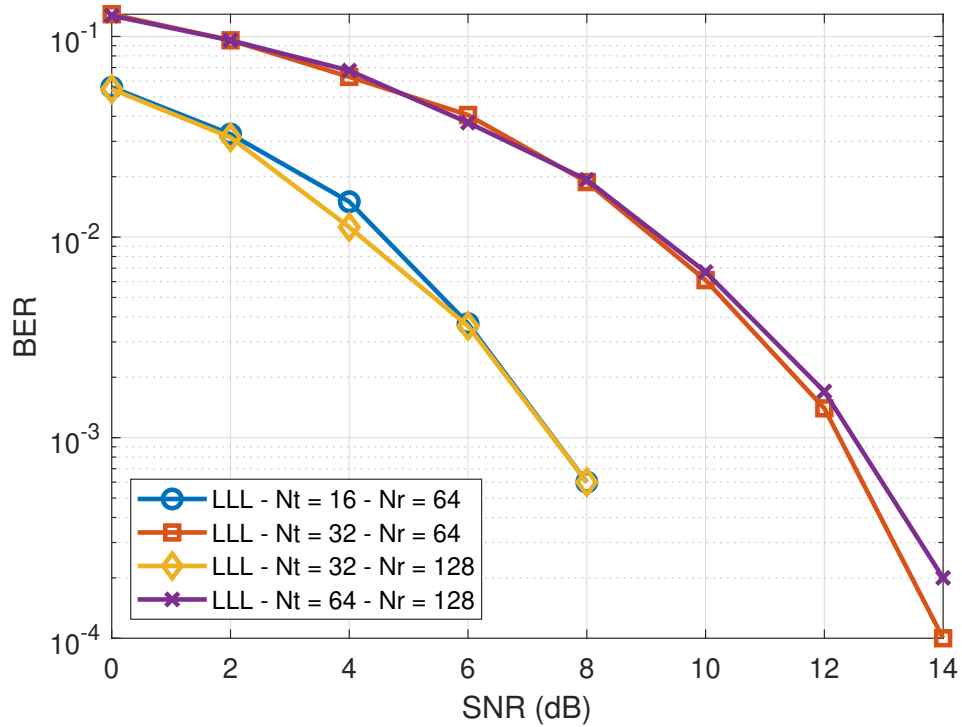
Como $\hat{\mathbf{x}}_{MMSE}$ representa uma versão suavizada do vetor $\tilde{\mathbf{x}}$, suas componentes podem ser aproximadas para os valores mais próximos da constelação. Definindo $\tilde{\mathbf{x}}_{sliced}$ como os valores quantizados, a estimativa final do vetor transmitido no domínio original é obtida por:

$$\hat{\mathbf{x}} = \mathbf{T}_{LLL}^{-1} \tilde{\mathbf{x}}_{sliced} \quad (3.43)$$

Assim pode-se aplicar em uma simulação conforme a seção 3.8.1 a seguir

3.8.1 Simulação detecção LLL: *Uplink*

A Figura 16 apresenta a BER em função da relação SNR para diferentes configurações do sistema MIMO Massivo no *uplink*, considerando a aplicação do algoritmo de detecção LLL. O objetivo é analisar o impacto do número de antenas transmissoras (N_t) e receptoras (N_r) no desempenho do sistema.

Figura 16 – Resultado Simulação com diversos cenários de LLL no *uplink*

Fonte: O próprio autor

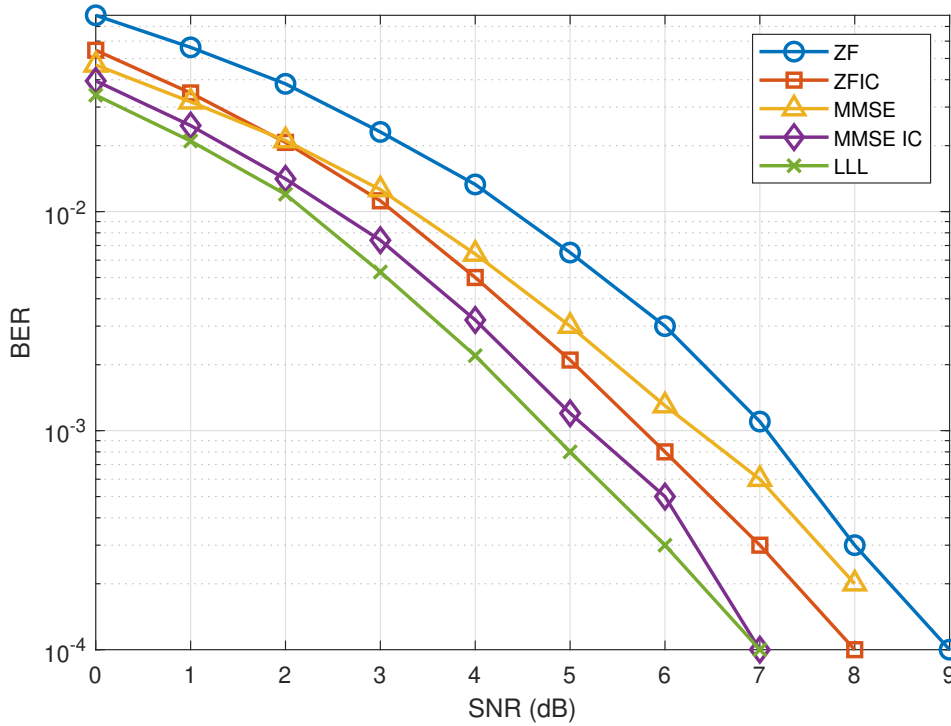
Os resultados indicam que a BER reduz conforme o número de antenas transmissoras aumenta para um mesmo número de antenas receptoras. Esse comportamento pode ser observado ao comparar os cenários com $N_t = 16, N_r = 64$ e $N_t = 32, N_r = 64$, onde o aumento de N_t resulta em uma queda mais acentuada na BER para valores crescentes de SNR. Esse efeito é atribuído ao ganho de diversidade proporcionado por um maior número de antenas transmissoras, que melhora a capacidade do receptor de distinguir os sinais transmitidos.

Também há o fato que quando o número de antenas receptoras é ampliado, há um impacto positivo na detecção. Comparando os casos $N_t = 32, N_r = 128$ e $N_t = 64, N_r = 128$, nota-se que o acréscimo de antenas no receptor melhora a rejeição de interferências, proporcionando um desempenho superior em termos de BER. Esse resultado reforça a importância do dimensionamento adequado da estrutura MIMO Massivo para a otimização do sistema.

3.9 Análise de Desempenho em MIMO Massivo *Uplink*

A seguir mostra um gráfico com o desempenho de 100 antenas de recepção na BS com 40 usuários.

Figura 17 – Comparação entre as técnicas ZF,ZF-IC,MMSE,MMSE-IC e LLL em cenário com 40 usuários e a BS com 100 antenas para recepção



Fonte: O próprio autor

Na tabela a seguir (1) mostra o desempenho relacionando o tempo computacional para executar cada método:

Método	ZF	MMSE	ZF-IC	MMSE-IC	LLL
Tempo	0.000492	0.000321	0.002353	0.00943	0.0151

Tabela 1 – Tempos de execução para diferentes métodos

Com os dados acima é possível analisar o desempenho de diferentes técnicas de detecção em termos de tempo de execução. O método ZF apresenta o menor tempo, com um valor de 0.000492 segundos, indicando uma abordagem computacionalmente eficiente. No entanto, apesar da sua rapidez, essa técnica tende a apresentar um desempenho inferior em termos de taxa de erro de bit (BER) quando comparada a métodos mais sofisticados.

O ZF-IC, por outro lado, tem um tempo de execução significativamente maior, chegando a 0.002353 segundos. Esse aumento no tempo pode ser justificado pelo processamento adicional envolvido na sua implementação, o que geralmente melhora o desempenho na detecção de símbolos, especialmente em cenários com interferências interusuário mais acentuadas.

O MMSE, com um tempo de 0.000321 segundos, é ainda mais rápido que o ZF, o que sugere que sua implementação não adiciona um custo computacional significativo em relação ao método mais simples. Esse método é conhecido por oferecer uma melhor supressão

de ruído quando comparado ao ZF, tornando-se uma opção atrativa para ambientes com uma relação sinal-ruído moderada.

Já o MMSE-IC, com um tempo de 0.00943 segundos, apresenta um tempo de execução um maior que o MMSE tradicional, refletindo o custo adicional do cancelamento iterativo de interferência. Essa abordagem tende a apresentar melhor desempenho quando a interferência entre os usuários é significativa, justificando o leve acréscimo na complexidade computacional.

Por fim, o método LLL, com um tempo de 0.0151 segundos, é o que apresenta a maior complexidade computacional entre os métodos analisados, embora ainda seja inferior ao ZF-IC. Essa técnica se destaca por proporcionar um melhor compromisso entre desempenho e complexidade, ajustando a base do canal para minimizar os erros na detecção de símbolos. Em cenários com alta interferência, essa técnica pode ser mais vantajosa, ainda que seu tempo de execução seja superior ao dos métodos mais simples.

Assim, escolha das melhores técnicas depende do critério adotado para a análise, seja ele o tempo de execução ou o desempenho em termos de taxa de erro de bit (BER). No entanto, ao considerar um equilíbrio entre ambas as métricas, algumas técnicas se destacam.

O método MMSE surge como uma das melhores opções, pois apresenta um tempo de execução extremamente baixo (0.000321 segundos) e um desempenho significativamente melhor que o ZF em relação ao BER. Ele consegue mitigar o impacto do ruído de forma mais eficiente, sendo uma escolha vantajosa para cenários com relação sinal-ruído (SNR) moderada a alta.

Já o MMSE-IC, apesar de ter um tempo de execução superior (0.00943 segundos), apresenta um desempenho mais robusto, especialmente em ambientes onde há baixo SNR. O ganho na redução do BER compensa o pequeno aumento na complexidade computacional, tornando-o uma opção superior para sistemas mais exigentes.

O LLL, embora apresente um tempo de execução mais elevado (0.0151 segundos), se destaca em cenários onde o BER precisa ser minimizado ao máximo. Sua abordagem baseada na redução da base do canal oferece uma melhora expressiva na detecção de símbolos, sendo ideal para aplicações que demandam uma maior robustez, mesmo com um custo computacional um pouco maior.

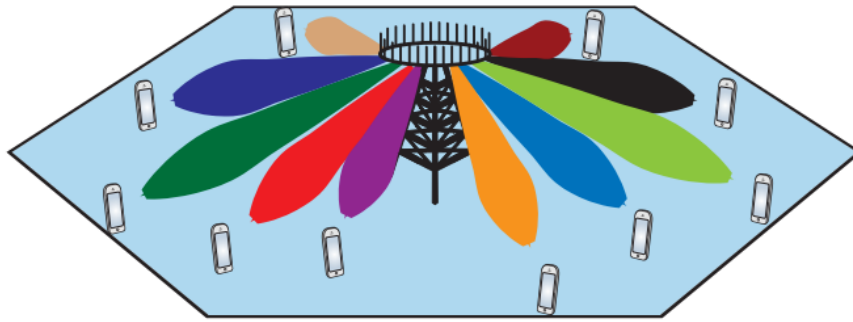
Por outro lado, o ZF e o ZF-IC, apesar de serem métodos tradicionais, acabam não sendo as melhores opções. O ZF, apesar do baixo tempo de execução, não apresenta um bom desempenho em termos de BER, tornando-se pouco eficiente para cenários com baixo SNR. O ZF-IC, por sua vez, tem um tempo de execução muito alto (0.002353 segundos) sem necessariamente apresentar um ganho proporcional na redução do BER, o que o torna menos atrativo frente a alternativas como o MMSE-IC e o LLL.

Dessa forma, as melhores opções considerando o equilíbrio entre desempenho e complexidade computacional são MMSE e MMSE-IC, pois combinam tempos de execução

baixos com uma significativa redução da taxa de erro de bit. O LLL também se destaca como uma excelente escolha para cenários onde o BER é um fator crítico, mesmo com um tempo de execução mais elevado.

4 Técnicas de Recepção MIMO: Enlace de descida

Figura 18 – Imagem demonstrativa do sistema *Downlink*



Autor: (CHIEN; BJORNSEN, 2017)

No enlace de descida, a comunicação ocorre da BS para os usuários terminais, sendo esta responsável por todo o esforço computacional associado à transmissão. Diferentemente do enlace de subida, onde os sinais transmitidos pelos usuários precisam ser detectados na BS, no *downlink* a BS possui informações completas sobre os dados que serão enviados. Para garantir que os usuários recebam os sinais corretamente e sem interferências significativas, é necessário aplicar técnicas de pré-codificação, que ajustam a transmissão de maneira a minimizar distorções causadas pelo canal de propagação.

A BS, equipada com múltiplas antenas, transmite simultaneamente para diversos usuários, explorando os graus de liberdade espaciais disponíveis no canal. Esse procedimento envolve a aplicação de uma matriz de pré-codificação que otimiza a transmissão dos símbolos, ajustando-os de forma a minimizar a interferência interusuário e garantir que cada terminal receba apenas as informações destinadas a ele. Como resultado, os usuários finais necessitam apenas de um processamento simples para demodular os sinais recebidos.

A eficiência dessa transmissão depende diretamente da matriz de canal e do método de pré-codificação adotado. Técnicas como ZF e MMSE são amplamente utilizadas para esse fim, permitindo que a BS realize a transmissão com interferência reduzida e melhor aproveitamento da capacidade espectral.

A Figura 18 ilustra a configuração do sistema no enlace de descida, destacando a distribuição dos sinais da BS para os usuários e o papel essencial da pré-codificação na mitigação de interferências. Esse modelo reforça a importância do processamento na BS, que busca maximizar a eficiência espectral e garantir comunicações confiáveis, mesmo em cenários com um grande número de usuários simultâneos.

Embora o processamento no enlace de descida dependa fortemente da pré-codificação realizada na estação base, é importante destacar as diferenças fundamentais entre as

operações de *uplink* e *downlink* em sistemas MIMO massivos. No *uplink*, os sinais transmitidos pelos usuários chegam simultaneamente à estação base, que precisa separá-los e remover interferências. Já no *downlink*, a estação base realiza a combinação dos sinais destinados aos diferentes usuários e transmite um único sinal otimizado para minimizar interferências na recepção.

A quantidade de antenas na estação base e nos terminais influencia diretamente a eficiência dessas operações. Em configurações com um grande número de antenas, a estação base assume um papel central na manipulação dos sinais, seja no processamento da recepção no *uplink* ou na pré-codificação no *downlink*. Além disso, no *uplink*, cada usuário transmite apenas seu próprio sinal, enquanto no *downlink* os usuários recebem um sinal pré-processado pela estação base, necessitando apenas de um procedimento de demodulação.

O quadro 1 apresenta uma comparação entre as características do *uplink* e do *downlink*, destacando aspectos como o sentido da comunicação, a necessidade de pré-codificação e o papel desempenhado pela estação base e pelos usuários em cada caso.

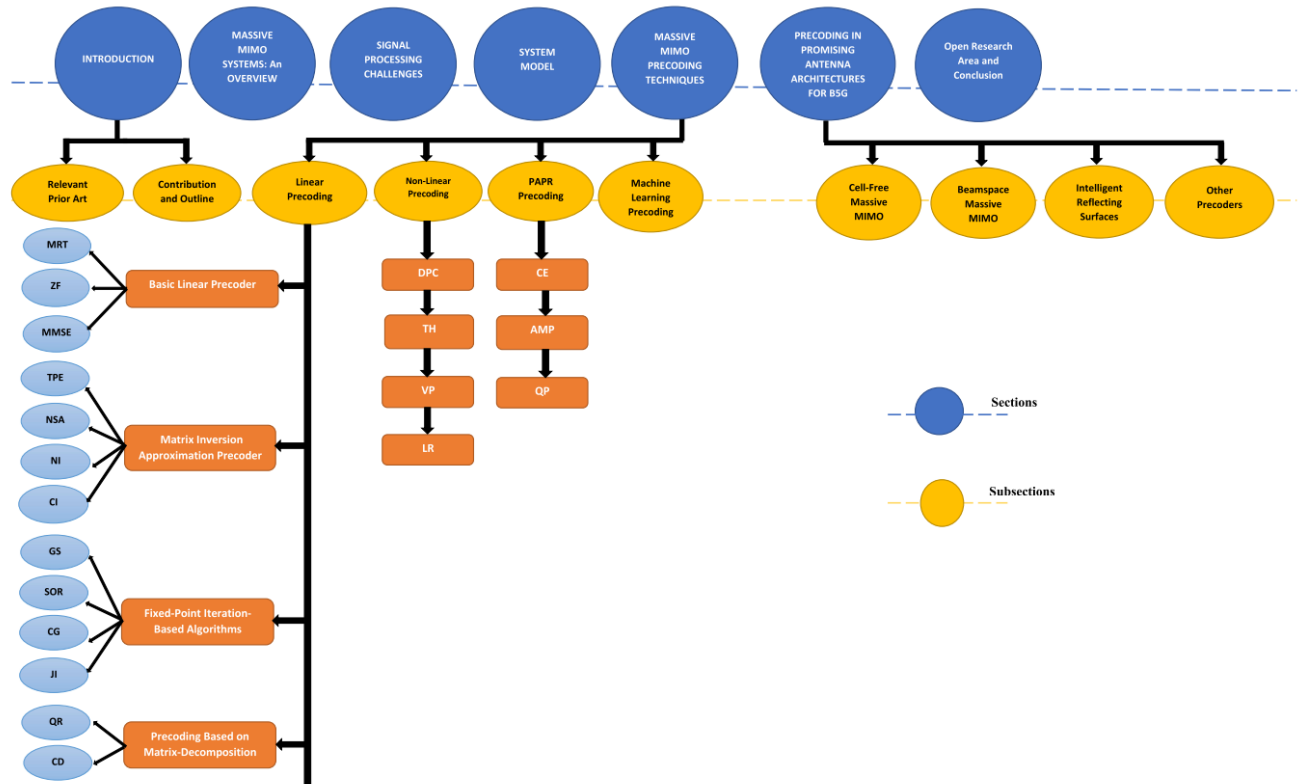
Quadro 1 – Comparação entre MIMO *Uplink* e *Downlink*

Aspecto	<i>Uplink</i>	<i>Downlink</i>
Sentido da comunicação	Usuários \rightarrow BS	BS \rightarrow Usuários
Número mínimo de antenas	≥ 64 na recepção (BS)	≥ 64 na transmissão (BS)
Pré-codificação	Não há	Sim, necessária
Função da BS	Separar sinais e remover interferência	Combinar sinais e aplicar pré-codificação
Sinais na BS	Recebidos juntos e são separados	Combinados e são transmitidos juntos
Função do usuário	Transmite apenas seu sinal	Recebe todos os sinais, mas separa apenas seu sinal

No livro (SAMSUZZAMAN MD. SHAI FUR RAHMAN, 2021) mostra diversas formas de precoding através da figura 19 irá ser feito a pré-codificação linear com ZF e MMSE. e pré-codificação não linear através da perturbação do vetor (VP - *Vetor perturbation*) que é utilizado no método NDS.

para melhor entendimento, o estudo apresentado em (SAMSUZZAMAN MD. SHAI FUR RAHMAN, 2021) explora diferentes técnicas de pré-codificação, destacando suas características e aplicações. A partir da análise da Figura 19, será realizada a implementação da pré-codificação linear utilizando os métodos ZF e MMSE. Além disso, será aplicada a pré-codificação não linear baseada no método VP, uma abordagem que busca minimizar o erro ao introduzir uma perturbação controlada nos símbolos transmitidos. Essa técnica será integrada ao esquema NDS, permitindo uma avaliação comparativa entre as diferentes estratégias de pré-codificação no contexto de sistemas Massive MIMO.

Figura 19 – Visão geral dos métodos de précodificação *Downlink*

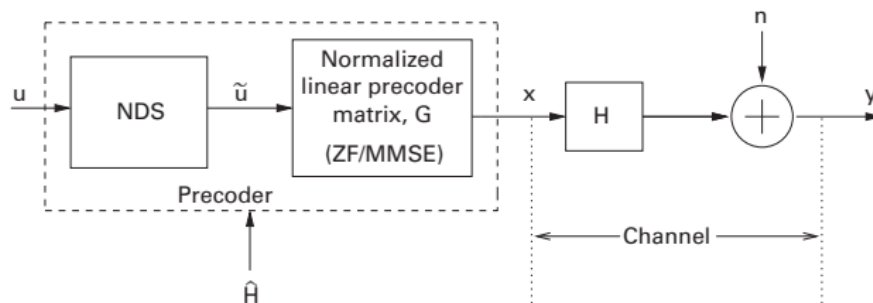


Autor: (SAMSUZZAMAN MD. SHAFUR RAHMAN, 2021)

4.0.1 Pré-codificação

A pré-codificação ou *precoding* é uma técnica utilizada em sistemas MIMO multiusuário para melhorar a transmissão de sinais na *Downlink* em que a estação base transmite para todos os usuários de uma vez. A técnica consiste em aplicar uma transformação linear ou não linear entre os símbolos de dados, usando uma matriz de pré-codificação. O objetivo dessa técnica é transmitir os dados minimizando as interferências e ruídos, otimizando o sistema MIMO, assim melhorando o SNR. Como mostrado na figura 21

Figura 20 – Diagrama de blocos ilustrando o funcionamento da pré-codificação



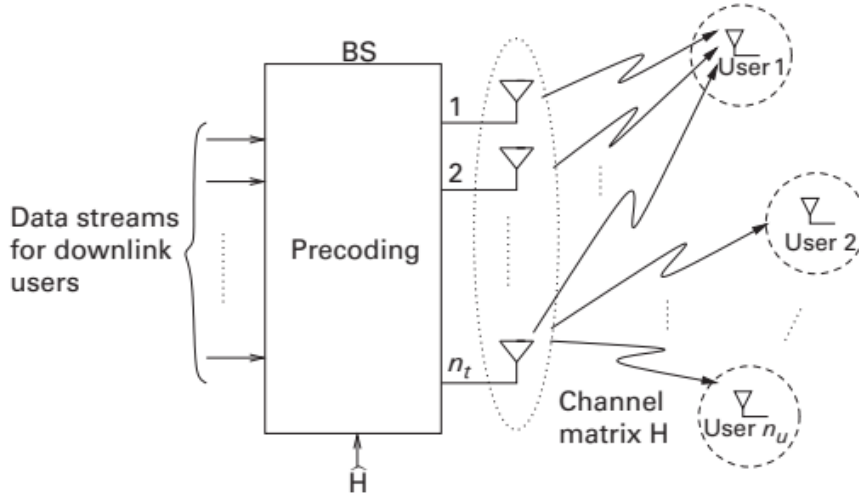
Autor: (CHOCKALINGAM; RAJAN, 2014).

4.0.2 Modelo do sistema

Considerando um sistema multiusuário MISO no *downlink*, onde a estação rádio-base (BS) se comunica com n_u usuários, cada um equipado com n_r antenas, sendo $n_r = 1$.

Usando o índice i , para representar o identificador de cada usuário no sistema, variando de 1 até n_u , ou seja, $i = 1, \dots, n_u$. Dessa forma $\mathbf{x} \in \mathbb{C}^{n_i}$, a matriz de dados complexos que contém os símbolos de informação destinados a cada usuário antes da aplicação do *precoding*. Neste sistema, o *precoding* é aplicado sobre os dados \mathbf{x} para gerar a matriz de dados pré-codificados $\mathbf{u} \in \mathbb{C}^{n_t}$, que será transmitida através das n_t antenas da estação rádio-base. E também é usado o \mathbf{H} para a matriz do canal, assim temos o sistema mostrado pela figura 21,

Figura 21 – Diagrama do sistema multiusuário MISO com a matriz de canal \mathbf{H}



Autor: (CHOCKALINGAM; RAJAN, 2014).

O modelo pode ser representado, inicialmente, no domínio dos números complexos por meio da seguinte equação:

$$\mathbf{y}_c = \mathbf{H}_c \mathbf{u}_c + \mathbf{n}_c \quad (4.1)$$

Neste modelo, $\mathbf{y}_c \in \mathbb{C}^{n_u}$ é o vetor de sinais complexos recebidos pelos N_u usuários. Cada componente de \mathbf{y}_c corresponde ao sinal recebido por um usuário específico. As demais variáveis são definidas como:

- $\mathbf{u}_c \in \mathbb{C}^{n_u}$ representa o vetor de símbolos complexos de informação, onde o i -ésimo elemento de \mathbf{u}_c é destinado ao i -ésimo usuário;
- $\mathbf{H}_c \in \mathbb{C}^{n_u \times n_t}$ é a matriz de canal, cujos elementos $\mathbf{h}_{i,j}$ modelam o ganho de canal complexo entre a j -ésima antena de transmissão da BS e o i -ésimo usuário. Assume-

se que os elementos de \mathbf{H}_c são variáveis aleatórias independentes e identicamente distribuídas (iid) segundo $\mathcal{CN}(0, 1)$;

- $\mathbf{n}_c \in \mathbb{C}^{n_u}$ é o vetor de ruído aditivo branco Gaussiano, cujas componentes são iid $\mathcal{CN}(0, \sigma^2)$.

O modelo acima pode ser convertido para uma representação equivalente em termos de variáveis reais. Tendo que:

$$\mathbf{y}_c = \mathbf{y}_I + j\mathbf{y}_Q, \quad \mathbf{u}_c = \mathbf{u}_I + j\mathbf{u}_Q, \quad \mathbf{x}_c = \mathbf{x}_I + j\mathbf{x}_Q, \quad \mathbf{H}_c = \mathbf{H}_I + j\mathbf{H}_Q, \quad \mathbf{n}_c = \mathbf{n}_I + j\mathbf{n}_Q \quad (4.2)$$

Utilizando essa composição, a equação do sistema é reescrita do domínio complexo para o domínio real como:

$$\mathbf{y}_p = \mathbf{H}_p \mathbf{x}_p + \mathbf{n}_p \quad (4.3)$$

onde:

$$\mathbf{y}_r = [\mathbf{y}_I^T \quad \mathbf{y}_Q^T]^T \in \mathbb{R}^{2n_u}, \quad \mathbf{x}_r = [\mathbf{x}_I^T \quad \mathbf{x}_Q^T]^T \in \mathbb{R}^{2n_t}, \quad \mathbf{n}_r = [\mathbf{n}_I^T \quad \mathbf{n}_Q^T]^T \in \mathbb{R}^{2n_u} \quad (4.4)$$

A matriz de canal \mathbf{H} no modelo real é definida como:

$$\mathbf{H}_r = \begin{pmatrix} \mathbf{H}_I & -\mathbf{H}_Q \\ \mathbf{H}_Q & \mathbf{H}_I \end{pmatrix} \in \mathbb{R}^{2n_u \times 2n_t} \quad (4.5)$$

Com essa transformação, o modelo real permite que tanto os componentes reais quanto os componentes imaginários das variáveis complexas sejam manipulados separadamente. Dependendo da modulação empregada, como M-PAM ou M-QAM, certos componentes do vetor de símbolos podem ser nulos ou representar apenas valores reais, facilitando o processamento no domínio real.

4.1 Detecção ZF: Downlink

4.1.1 Vetor de Pré-codificação G

No contexto de sistemas MIMO de múltiplos usuários como abordado em (CHOCKALINGAM; RAJAN, 2014), a técnica de pré-codificação desempenha um papel essencial na mitigação de interferências entre os usuários e na maximização da eficiência espectral. O vetor de sinais transmitidos \mathbf{x} pode ser expresso como:

$$\mathbf{x}_r = \frac{\mathbf{G}\mathbf{u}}{\|\mathbf{G}\mathbf{u}\|} \quad (4.6)$$

onde \mathbf{G} é a matriz de pré-codificação com dimensões $\mathbf{G} \in \mathbb{R}^{2n_t \times 2n_u}$ e \mathbf{u} representa o vetor de símbolos de informação já modulados a serem transmitidos. A normalização $\|\mathbf{Gu}\|$ é necessária para garantir que a potência total transmitida permaneça dentro dos limites especificados.

Adotando a técnica de pré-codificação ZF, que busca eliminar completamente a interferência entre os fluxos de dados dos diferentes usuários, a matriz de pré-codificação é definida como:

$$\mathbf{G} = \mathbf{G}_{ZF} = \mathbf{H}_r^T (\mathbf{H}_r \mathbf{H}_r^T)^{-1} \quad (4.7)$$

Então, \mathbf{H}_p é a matriz de canal, que contém os ganhos de canal complexos entre as antenas de transmissão e recepção. Essa definição de \mathbf{G} permite que o vetor de sinais recebidos \mathbf{y}_r seja dado por:

$$\mathbf{y}_r = \frac{\mathbf{u}}{\|\mathbf{Gu}\|} + \mathbf{n}_r \quad (4.8)$$

4.2 Detecção MMSE: *Downlink*

A técnica de pré-codificação baseada no critério MMSE busca otimizar a transmissão no *downlink* ao minimizar simultaneamente a interferência interusuário e os efeitos da amplificação do ruído. Diferente de abordagens como ZF, que anulam completamente a interferência sem levar em consideração o impacto do ruído, o MMSE equilibra essas duas influências, proporcionando maior robustez em canais mal condicionados. Esse método leva em conta a presença do ruído ao incluir um termo regularizador na inversão da matriz do canal, evitando a amplificação excessiva que ocorre quando os valores singulares da matriz de canal são muito pequenos.

A pré-codificação MMSE segue a mesma estrutura da abordagem ZF, diferenciando-se apenas pela definição da matriz \mathbf{G} em que a matriz de pré-codificação para MMSE é definida como:

$$\mathbf{G} = \mathbf{G}_{MMSE} = \mathbf{H}_r^T (\mathbf{H}_r \mathbf{H}_r^T + \sigma^2 \mathbf{I})^{-1} \quad (4.9)$$

onde σ^2 representa a variância do ruído aditivo branco Gaussiano. Com essa substituição, a estrutura matemática do modelo de transmissão permanece inalterada, mantendo a normalização da potência transmitida e a relação entre os vetores de entrada e saída do sistema. O vetor de sinais transmitidos mantém a mesma formulação:

$$\mathbf{x}_r = \frac{\mathbf{Gu}}{\|\mathbf{Gu}\|} \quad (4.10)$$

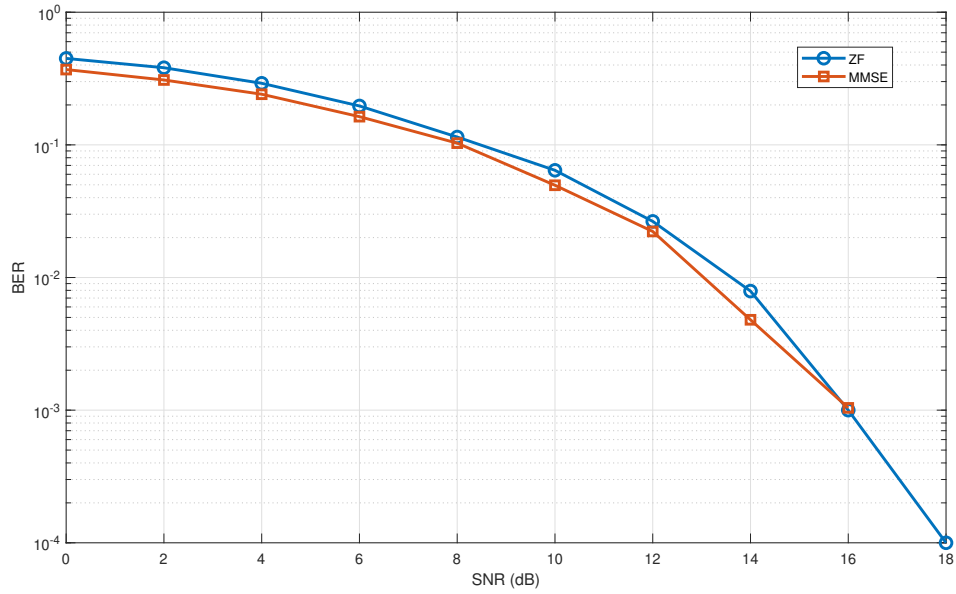
garantindo que a potência total transmitida respeite os limites estabelecidos. Da mesma forma, o sinal recebido continua sendo descrito por:

$$\mathbf{y}_r = \frac{\mathbf{u}}{\|\mathbf{G}\mathbf{u}\|} + \mathbf{n}_r \quad (4.11)$$

preservando a modelagem do ruído e da detecção no receptor. A única diferença entre as abordagens está na definição da matriz de pré-codificação, que no caso do MMSE inclui um termo adicional para mitigar a amplificação do ruído. Dessa forma, ambas as técnicas compartilham a mesma estrutura matemática, diferenciando-se apenas pela forma como tratam a interferência e o impacto do ruído na transmissão.

4.2.1 Simulação ZF e MMSE em *Downlink*

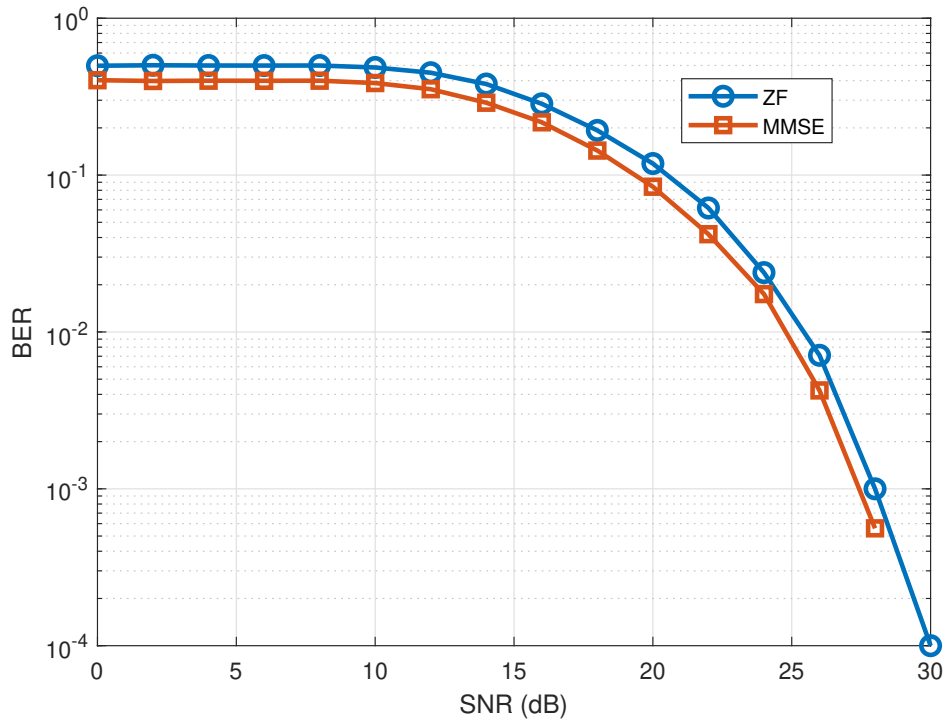
Figura 22 – Resultado 64 transmissoras e 4 antenas receptoras tendo a comparação entre ZF e MMSE, no *Downlink*



Fonte: O próprio autor

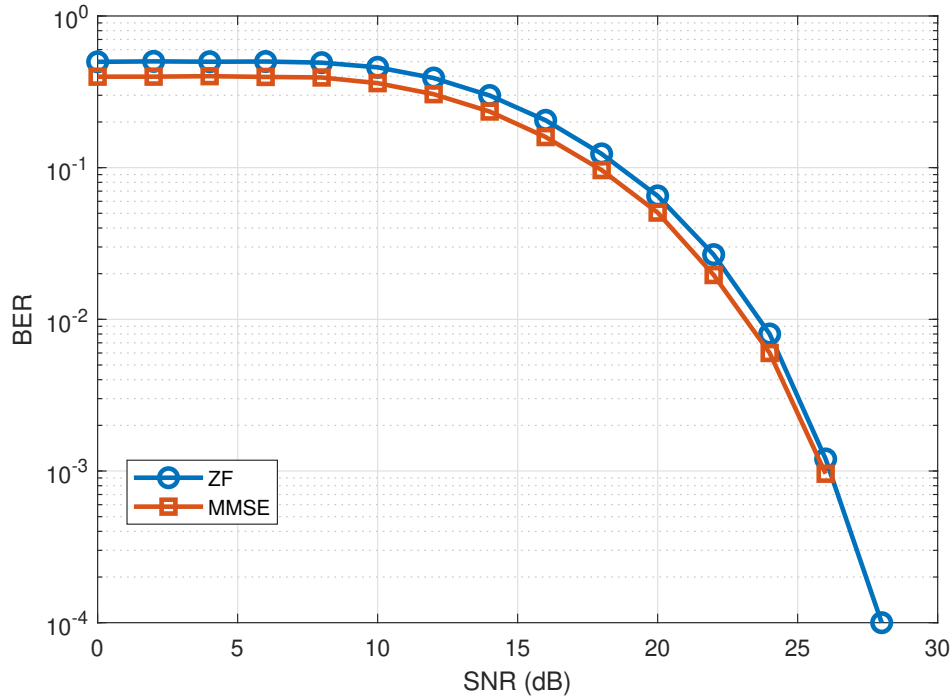
A Figura 22 apresenta a BER em função da SNR para uma configuração 64 transmissoras e 4 antenas receptoras. O número reduzido de antenas transmissoras implica menor interferência entre os sinais dos usuários, facilitando a separação espacial e a recuperação dos dados. Por outro lado, a presença de 64 antenas receptoras proporciona uma elevada diversidade espacial, melhorando significativamente a detecção do sinal. Esse fator contribui para que o desempenho do sistema alcance um patamar aceitável para SNRs mais baixas, atingindo até 18 dB antes que a BER se torne desprezível. O comportamento dos detectores ZF e MMSE segue a tendência esperada, com o MMSE apresentando uma leve vantagem devido à sua melhor capacidade de mitigar o ruído.

Figura 23 – Resultado 64 transmissoras e 32 antenas receptoras tendo a comparação entre ZF e MMSE, no *Downlink*



Fonte: O próprio autor

Na Figura 23, observa-se o número de antenas transmissoras sendo 64, e o aumento de antenas receptoras para 32. Esse crescimento no número de usuários ativos intensifica a interferência interusuário, tornando a recuperação do sinal mais desafiadora. O efeito desse aumento de interferência é perceptível na BER, que se mantém relativamente elevada para baixos valores de SNR e exige uma relação SNR maior para atingir o mesmo desempenho observado no caso 4x64. O MMSE continua apresentando desempenho superior ao ZF, especialmente na faixa intermediária de SNR, onde o impacto da interferência residual é mais relevante. No entanto, para SNRs elevadas, ambos os esquemas convergem para um mesmo nível de erro.

Figura 24 – Resultado 128 transmissoras e 32 antenas receptoras tendo a comparação entre ZF e MMSE, no *Downlink*

Fonte: O próprio autor

A Figura 24 ilustra o impacto do aumento do número de antenas transmissoras para 128, mantendo 32 antenas receptoras. Comparado ao caso anterior, o sistema exibe um desempenho aprimorado, evidenciado por uma redução mais acentuada na BER para uma mesma faixa de SNR. O aumento da diversidade espacial oferecido pelo maior número de antenas transmissoras possibilita uma diversidade melhor ao transmitir, reduzindo os efeitos da interferência e do ruído. Assim, a necessidade de uma relação SNR elevada para garantir baixos níveis de BER é atenuada, tornando esta configuração mais robusta em relação ao cenário de 64 antenas transmissoras. O MMSE segue mostrando uma vantagem sobre o ZF, especialmente na transição entre regimes intermediários e elevados de SNR, mas para valores muito altos de SNR, ambos os detectores convergem para o mesmo limite inferior de erro.

4.3 Detecção com Pré-codificação Precodificador baseado em Busca por Descida de Norma - (NDS - Precoder based on Norm Descent Search)

A técnica de pré-codificação baseada em NDS é utilizada para minimizar a interferência em sistemas MIMO de múltiplos usuários. O algoritmo busca uma configuração ótima para o vetor de símbolos transmitidos, iterando sobre ajustes baseados na redução de

norma. O NDS utiliza-se da técnica de perturbação vetorial que é aplicada para melhorar o desempenho, ajustando o vetor \mathbf{u} de acordo com a seguinte expressão:

$$\tilde{\mathbf{u}} = \mathbf{u} + \tau \mathbf{p} \quad (4.12)$$

onde τ é um escalar positivo e \mathbf{p} representa um vetor de perturbação pertencente a um conjunto discreto de pontos que minimiza a norma da transformação do vetor perturbado pelo canal. O vetor de perturbação ótimo \mathbf{p}_{opt} é definido como:

$$\mathbf{p}_{\text{opt}} = \arg \min_{\mathbf{p} \in \mathbb{Z}^{2n_u}} \|\mathbf{G}(\mathbf{u} + \tau \mathbf{p})\|^2. \quad (4.13)$$

A busca exata pelo vetor \mathbf{p}_{opt} exige complexidade exponencial em relação ao número de usuários n_u , tornando sua aplicação inviável para grandes sistemas MIMO. Alternativamente, métodos aproximados com complexidade polinomial são adotados para resolver esse problema de forma eficiente, permitindo uma solução subótima que mantém um bom compromisso entre desempenho e viabilidade computacional.

No processo de detecção no receptor, considera-se uma solução aproximada $\tilde{\mathbf{p}}$ para \mathbf{p}_{opt} . O vetor de sinal recebido, após a compensação pelo fator de ganho $\|\mathbf{G}(\mathbf{u} + \tau \tilde{\mathbf{p}})\|$, é expresso como:

$$\tilde{\mathbf{y}} = (\mathbf{u} + \tau \tilde{\mathbf{p}}) + \tilde{\mathbf{n}}, \quad (4.14)$$

onde o ruído efetivo $\tilde{\mathbf{n}}$ é dado por:

$$\tilde{\mathbf{n}} = \|\mathbf{G}(\mathbf{u} + \tau \tilde{\mathbf{p}})\| \mathbf{n}. \quad (4.15)$$

Com base nesse modelo, o vetor de símbolos detectados no receptor é obtido aplicando a seguinte expressão:

$$\hat{\mathbf{u}} = \tilde{\mathbf{y}} - \tau \left[\frac{\tilde{\mathbf{y}} + \frac{\tau}{2}}{\tau} \right]. \quad (4.16)$$

Esse método permite que o receptor estime os símbolos transmitidos de forma eficiente, aproveitando a pré-codificação realizada na estação base para reduzir os efeitos da interferência e melhorar a confiabilidade da detecção.

4.3.1 Simulação NDS

O primeiro passo para simular consiste na definição dos parâmetros fundamentais do sistema: o número de antenas transmissoras, o número de usuários e o número de antenas receptoras. Além disso, define-se a modulação a ser utilizada, que neste caso é o QAM-4. Para essa modulação, o valor máximo da constelação é dado por $s_{\text{max}} = \sqrt{2}$.

A inicialização do algoritmo requer três elementos principais: o vetor de símbolos ajustado inicialmente como $\tilde{\mathbf{u}}^{(0)} = \mathbf{u}$, a matriz de pré-codificação $\mathbf{F} = \mathbf{G}^\top \mathbf{G}$ e o índice de iteração $k = 0$.

Em seguida, o vetor $\mathbf{z}^{(0)}$ é calculado a partir da matriz \mathbf{F} e do vetor inicial $\tilde{\mathbf{u}}^{(0)}$:

$$\mathbf{z}^{(0)} = \mathbf{F}\tilde{\mathbf{u}}^{(0)}.$$

Paralelamente, define-se o parâmetro τ para o ajuste de símbolos perturbados. Este parâmetro é calculado como:

$$\tau = 2|c_{\max}| + d.$$

Considerando a modulação QAM-4, temos $|c_{\max}| = 1$ e $d = 2$, resultando em $\tau = 4$.

A quantidade de símbolos modulados, n_{symb} , depende do tipo de modulação utilizada. Para QAM, define-se $n_{\text{symb}} = 2n_u$, enquanto para PAM, $n_{\text{symb}} = n_u$.

O algoritmo então entra em um laço principal iterativo. Para cada índice $i = 1, 2, \dots, n_{\text{symb}}$, realiza-se o cálculo de $\lambda_i^{(k)}$ como:

$$\lambda_i^{(k)} = -\text{sgn}(z_i^{(k)}) \left\lfloor \frac{|z_i^{(k)}|}{\tau F_{i,i}} \right\rfloor.$$

Caso o valor absoluto de $\lambda_i^{(k)}$ ultrapasse s_{\max} , ele é redefinido como zero. Posteriormente, calcula-se $\Delta q_{i,\text{opt}}^{(k+1)}$:

$$\Delta q_{i,\text{opt}}^{(k+1)} = (\lambda_i^{(k)})^2 F_{i,i} + 2(\lambda_i^{(k)}) z_i^{(k)}.$$

Ao final do laço, determina-se o menor valor Δq_{\min} :

$$\Delta q_{\min} = \min_i \Delta q_{i,\text{opt}}^{(k+1)}.$$

Se $\Delta q_{\min} \geq 0$, o algoritmo avança para o término. Caso contrário, identifica-se o índice j que minimiza $\Delta q_{i,\text{opt}}$:

$$j = \arg \min_i \Delta q_{i,\text{opt}}^{(k+1)}.$$

A partir desse índice, atualizam-se os vetores $\tilde{\mathbf{u}}$ e \mathbf{z} :

$$\tilde{\mathbf{u}}^{(k+1)} = \tilde{\mathbf{u}}^{(k)} + \tau \lambda_j \mathbf{e}_j, \quad \mathbf{z}^{(k+1)} = \mathbf{z}^{(k)} + \tau \lambda_j \mathbf{f}_j.$$

O índice de iteração é incrementado:

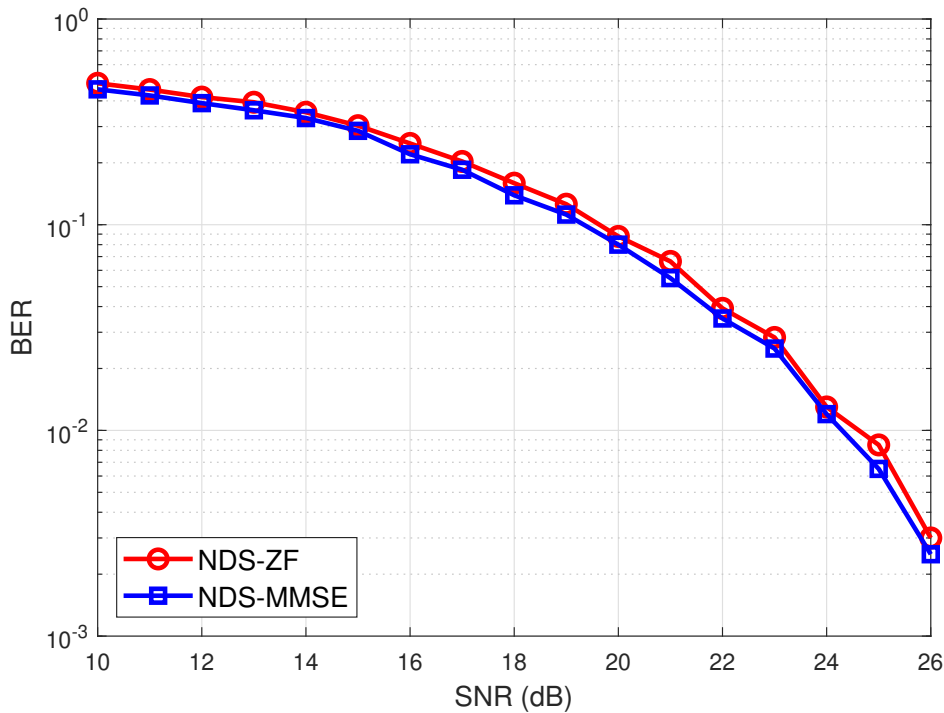
$$k = k + 1.$$

O processo retorna ao cálculo inicial de \mathbf{z} até que a condição de término seja satisfeita, completando assim o algoritmo de pré-codificação NDS.

4.3.1.1 Resultados da Simulação NDS

Nesta simulação, são analisadas a BER dos precodificadores NDS-ZF e NDS-MMSE em um sistema MIMO de grande escala. A simulação considera um cenário com $n_t = 64$ antenas transmissoras e $n_r = 40$ antenas receptoras, empregando modulação 4-QAM. O objetivo é avaliar o desempenho da precodificação no cenário com interferência multiusuário e na preservação da integridade do sinal em diferentes SNR.

Figura 25 – Desempenho dos precodificadores NDS-ZF e NDS-MMSE para $n_t = 64$ e $n_r = 40$



Fonte: O próprio autor

A análise do desempenho dos precodificadores NDS-ZF e NDS-MMSE demonstra a superioridade do critério MMSE em relação ao ZF no que diz respeito a (BER). Conforme ilustrado na Figura 25, a curva de BER referente ao NDS-MMSE exibe melhor performance em comparação com o NDS-ZF, evidenciando que a minimização conjunta da interferência e do ruído resulta em um desempenho mais eficiente. O método ZF, apesar de mitigar a interferência inter-usuário, não leva em consideração a amplificação do ruído, o que impacta negativamente seu desempenho.

Além disso, observa-se que a diferença de desempenho entre os dois métodos reduz conforme o SNR aumenta. Esse motivo se realiza porque em condições de SNR elevado, a influência do ruído no sistema se torna menos significativa, permitindo que ambos os precodificadores apresentem um desempenho próximo ao ideal. Entretanto, em faixas intermediárias de SNR, o NDS-MMSE apresenta uma vantagem mais evidente, propor-

cionando uma melhor relação entre a supressão da interferência e a preservação do sinal desejado, o que justifica sua maior eficiência frente ao NDS-ZF.

Os resultados obtidos confirmam que a utilização do NDS-MMSE em sistemas MIMO massivos não apenas se adapta bem a arquiteturas de grande escala, mas também proporciona uma melhoria no desempenho da comunicação, especialmente em cenários em que há interferência multiusuário.

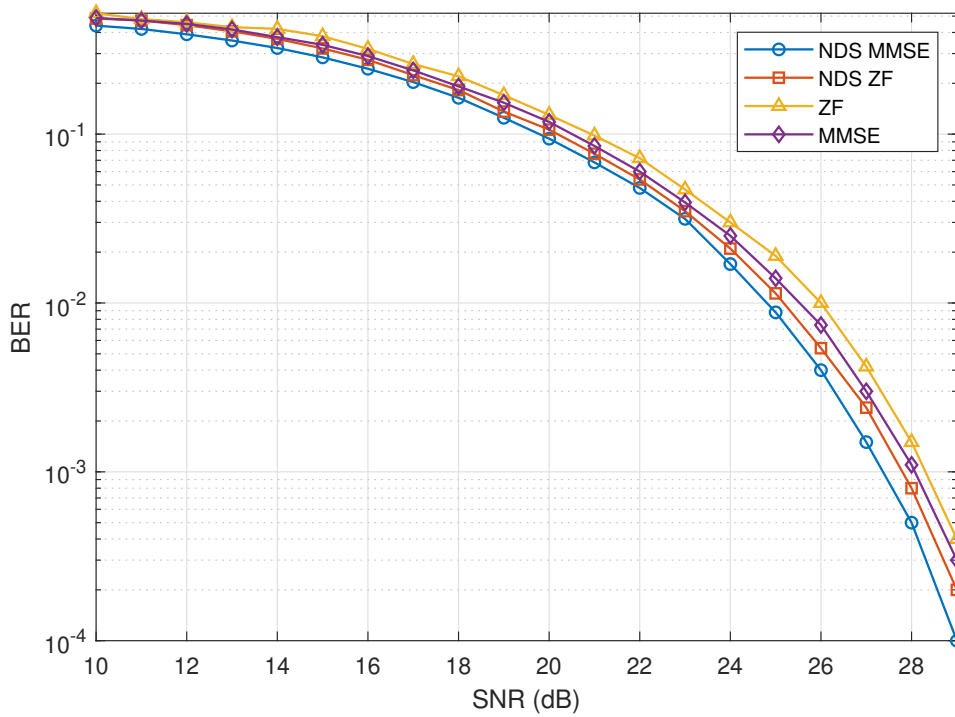
4.4 Análise de Desempenho em MIMO Massivo *Downlink*

A eficiência dos precodificadores é um fator crítico para garantir um bom desempenho em sistemas MIMO Massivo no *downlink*. Neste contexto, são analisadas as técnicas ZF, MMSE, NDS-ZF e NDS-MMSE no cenário com $N_t = 100$ antenas transmissoras e $N_r = 40$ antenas receptoras. A Figura 26 apresenta a BER dessas técnicas em função da SNR, enquanto a Tabela 2 compara os tempos de execução de cada método.

Os métodos ZF e MMSE são amplamente utilizados para mitigar a interferência entre usuários em sistemas MIMO. No entanto, o compromisso entre complexidade computacional e desempenho ainda é um desafio nesses sistemas. A introdução da técnica NDS, que pode ser aplicada ao ZF e ao MMSE, busca otimizar esse compromisso ao aprimorar a eficiência sem aumentar significativamente a complexidade computacional.

A seguir, através da imagem 26 e a tabela 2 é analisado os resultados obtidos, destacando as vantagens e desvantagens de cada método e avaliando o impacto da técnica NDS na melhoria do desempenho do sistema.

Figura 26 – Desempenho dos precodificadores ZF,MMSE,NDS-ZF e NDS-MMSE para $n_t = 100$ e $n_r = 40$



Fonte: O próprio autor

Método	MMSE	ZF	NDS
Tempo	0.000387	0.000287	0.000124

Tabela 2 – Tempos de execução para diferentes métodos

Assim, fazendo a análise dos tempos de execução apresentados na Tabela 2 em conjunto com os resultados da Figura 26 revela aspectos importantes sobre o desempenho dos diferentes métodos. O método MMSE, que apresenta um tempo de execução de 0.000387 segundos, é amplamente reconhecido por sua capacidade de minimizar o erro quadrático médio, proporcionando um desempenho robusto na redução da BER. No entanto, essa melhoria vem acompanhada de um custo computacional relativamente maior em comparação com outras técnicas.

O método ZF, com um tempo de 0.000287 segundos, se destaca por sua eficiência computacional, sendo mais rápido que o MMSE. Apesar disso, sua abordagem de cancelamento de interferência pode amplificar o ruído em cenários de alta interferência, impactando a performance em determinadas condições. Seu menor tempo de processamento pode ser vantajoso em aplicações onde a latência é um fator crítico.

A técnica NDS, que pode ser aplicada tanto ao MMSE quanto ao ZF, mostra-se promissora ao combinar eficiência computacional com melhorias no desempenho do sistema. Seu tempo de execução, de apenas 0.000124 segundos, indica que pode ser integrada às outras técnicas sem comprometer significativamente a latência. Na Figura 26, observa-se

que a aplicação do NDS ao ZF e ao MMSE melhora o BER, tornando a técnica uma alternativa viável para otimização dos sistemas de comunicação sem fio.

Dessa forma, a escolha do método ideal depende do compromisso entre tempo de processamento e desempenho na mitigação de erros. O MMSE continua sendo uma opção sólida para cenários mais exigentes, enquanto o ZF pode ser útil em condições mais favoráveis. A introdução do NDS a esses métodos representa uma evolução, permitindo que ambos melhorem seu desempenho sem penalizar significativamente o tempo de execução.

5 Conclusão

Este estudo analisou o desempenho de diferentes técnicas de detecção e pré-codificação aplicadas a sistemas Massive MIMO, considerando tanto o enlace de subida quanto o enlace de descida. A avaliação dos métodos revelou que a escolha da técnica influencia diretamente a eficiência do sistema, impactando a taxa de erro de bit, a complexidade computacional e a robustez frente a interferências.

No enlace de subida, os resultados demonstraram que os métodos analisados apresentam compromissos distintos entre tempo de execução e qualidade da detecção. Enquanto o ZF oferece uma solução computacionalmente eficiente, seu desempenho em cenários de baixa relação sinal-ruído é inferior. Por outro lado, o MMSE mostrou-se uma alternativa mais robusta ao reduzir os efeitos do ruído sem comprometer significativamente a eficiência computacional. A aplicação de cancelamento iterativo de interferência, como no MMSE-IC, demonstrou ganhos expressivos no BER, mas com um custo computacional superior. O método LLL, apesar de apresentar maior complexidade, provou-se eficaz na mitigação de interferências interusuário, sendo uma alternativa viável para ambientes com alta densidade de usuários.

No enlace de descida, a implementação de pré-codificadores foi essencial para minimizar a interferência multiusuário e garantir a transmissão eficiente dos sinais. A introdução do método NDS aos esquemas convencionais demonstrou vantagens significativas, proporcionando melhorias no BER sem impactar de maneira expressiva o tempo de execução. Dentre as abordagens testadas, o NDS-MMSE se destacou ao manter um desempenho consistente em diferentes faixas de SNR, garantindo um melhor compromisso entre qualidade de detecção e eficiência computacional.

Os resultados das simulações indicaram que a utilização de um grande número de antenas na estação base melhora a capacidade do sistema, reforçando a importância de técnicas avançadas de processamento de sinal para lidar com os desafios de cenários de alta dimensionalidade. A análise confirmou que a seleção adequada da técnica de detecção e pré-codificação é crucial para maximizar a eficiência espectral e minimizar os efeitos adversos da interferência interusuário.

Dessa forma, conclui-se que a escolha entre as técnicas deve considerar tanto os requisitos computacionais quanto o nível de interferência presente no sistema. Métodos como LMMSE e LLMMSE-IC se mostraram eficientes para cenários onde a mitigação de ruído é prioritária, enquanto o uso de técnicas baseadas em NDS revelou-se promissor para sistemas de transmissão que necessitam equilibrar desempenho e complexidade computacional. A combinação dessas estratégias representa um caminho promissor para a otimização de sistemas Massive MIMO, garantindo maior confiabilidade e eficiência na comunicação sem fio.

Referências

AGRELL T. ERIKSSON, A. V. E.; ZEGGER, K. *Closest point search in lattices*.

Desconhecido: IEEE Transactions on Information Theory, 2002. ISBN N/A. 55

CHIEN, T. V.; BJORNSSON, E. *Massive MIMO Communications*. Linköping University, Sweden: Springer, 2017. ISBN 978-3-319-34206-1. 27, 30, 41, 65

CHO, Y. S. et al. *MIMO-OFDM Wireless Communications with MATLAB*. United Kingdom: John Wiley Sons, 2010. ISBN 978-0-470-82556-0. 55

CHOCKALINGAM, A.; RAJAN, B. S. *Large MIMO Systems*. United States of America: Cambridge University Press, 2014. ISBN 978-1-107-02665-0. 33, 67, 68, 69

HAMPTON, J. R. *Introduction to MIMO Communications*. University Printing House, Cambridge CB2 8BS, United Kingdom: Cambridge University Press, 2014. ISBN 978-1-107-04283-4. 30, 46

LARSSON, E. G. *MIMO Detection Methods: How They Work*. Linköping University, Sweden: IEEE Signal Processing Magazine, 2009. ISBN N/A. 55

MARZETTA, T. L. et al. *Fundamentals of Massive MIMO*. University Printing House, Cambridge CB2 8BS, United Kingdom: Cambridge University Press, 2016. ISBN 978-1-107-17557-0. 29

SAMSUZZAMAN MD. SHAIKUR RAHMAN, M. A. M. M. *Overview of Precoding Techniques for Massive MIMO*. USA: IEEE Communications Surveys Tutorials, 2021. ISBN N/A. 66, 67

VISWANATHAN, M. *Wireless Communication Systems in MATLAB, Second Edition*. [S.l.]: Mathuranathan Viswanathan, 2020. 50

WINDPASSINGER, C.; FISCHER, R. *Low-complexity near-maximum-likelihood detection and precoding*. Desconhecido: IEEE Information Theory Workshop, 2003. ISBN N/A. 55