

Proposta de um Algoritmo para Redução dos Cálculos da Inversa da Matriz de Filtragem para Detecção Linear de Sinais em Sistemas *Uplink* MIMO Massivo

Maykon Renan Pereira da Silva¹, Flávio Geraldo Coelho Rocha¹

¹Escola de Engenharia Elétrica, Mecânica e de Computação
Universidade Federal de Goiás (UFG)

maykonrenan@discente.ufg.br, flaviogcr@ufg.br

Abstract. *The massive Multiple-Input Multiple-Output (MIMO) is one of the most promising technologies for 5G. However, there are challenges in its practical implementation, such as high computational complexity in the inversion of the filtering matrix. Matrix inversion is an important issue that affects the performance of signal detection algorithms. A good detection algorithm for massive MIMO systems needs to estimate the filtering matrix with high accuracy and have low computational complexity for hardware implementation. To solve this problem, a linear signal detection method is proposed in this paper. The proposed algorithm uses an iterative scheme based on the secant method to estimate the inverse of the filtering matrix of the Minimum Mean-Squared Error (MMSE) algorithm. The iterative scheme has an advantage over other methods in the literature because it has two initial approximations, which ensure rapid convergence. Also, the stair matrix method is used to reduce the number of matrix operations and consequently the execution time. The results show that the proposed algorithm obtained near-optimal Bit Error Rate (BER) performance, which demonstrates that it has good precision in the estimation of the transmitted signal.*

Resumo. *O sistema MIMO massivo é uma das tecnologias mais promissoras para as redes 5G. Entretanto, existem problemas desafiadores na implementação prática desta tecnologia, como por exemplo, a alta complexidade computacional na inversão da matriz de filtragem. A inversão de matriz é uma questão importante que afeta o desempenho dos algoritmos de detecção de sinais. Um bom algoritmo de detecção para sistemas MIMO massivo depende de uma boa precisão na estimação da matriz de filtragem e, principalmente, ter baixa complexidade computacional para implementação em hardware. Com o objetivo de solucionar este problema, propõe-se um método de detecção linear de sinais. O algoritmo proposto utiliza um esquema iterativo baseado no método da secante para estimar a inversa da matriz de filtragem do algoritmo MMSE. O esquema iterativo tem como vantagem em relação a outros métodos da literatura, ter duas aproximações iniciais, o que garante uma rápida convergência. Além disso, o método matriz escada é utilizado com o objetivo de reduzir o número de operações matriciais e, conseqüentemente, o tempo de execução. Os resultados mostram que o algoritmo proposto obteve desempenho quase-ótimo em termos de BER, o que demonstra que o mesmo possui uma boa precisão na estimação do sinal transmitido.*

1. Introdução

A comunicação móvel se desenvolveu rapidamente, após várias gerações [Khan et al. 2020]. O aumento do tráfego de dados e a popularidade dos dispositivos móveis levam a uma alta exigência das redes de comunicações [Zhang et al. 2020]. O 4G não pode satisfazer as demandas dos usuários por capacidade, velocidade e espectro [Chataut and Akl 2020]. Neste sentido, a tecnologia MIMO massivo (*massive Multiple-Input Multiple-Output*) tem sido vista como uma das principais tecnologias habilitadoras dos sistemas de comunicação sem fio de quinta geração (5G) [Chataut and Akl 2020, Björnson et al. 2017]. O sistema MIMO massivo aumenta a capacidade da rede ao empregar centenas de antenas na Estação Base (*Base Estation*) com o objetivo de atender simultaneamente vários usuários [Zhang et al. 2020, Chataut and Akl 2020]. Essa tecnologia é capaz de trazer ganhos de desempenho em eficiência espectral, confiabilidade do *link* de comunicação e cobertura celular, quando comparada ao tradicional sistema MIMO em pequena escala utilizado nas redes 4G [Zhang et al. 2020, Chataut and Akl 2020, Björnson et al. 2017].

O uso massivo da multiplexação espacial em sistemas MIMO aumenta significativamente a dimensão das matrizes utilizadas em algoritmos de detecção de sinais. Portanto, a eficiência desses algoritmos é determinante para o sucesso do MIMO massivo em redes 5G [Albreem et al. 2019, Wu et al. 2013, Yin et al. 2013, Wu et al. 2014]. O detector ótimo baseado no estimador de Máxima Verossimilhança (ML – *Maximum Likelihood*) não é viável para este sistema em larga escala, já que sua complexidade aumenta exponencialmente com o número de antenas de transmissão [Gao et al. 2014]. Uma abordagem conhecida é o detector linear baseado no estimador do Erro Quadrático Médio-Mínimo (MMSE – *Minimum Mean-Squared Error*), que pode fornecer um desempenho de Taxa de Erro de *Bit* (BER – *Bit Error Rate*) quase-ótimo em sistemas MIMO massivo [Albreem et al. 2019]. No entanto, à medida que o número de arranjos de antenas aumenta, a complexidade computacional da inversão da matriz de filtragem do algoritmo MMSE torna-se alta para implementação em *hardware*. Na literatura, vários algoritmos têm sido propostos para detecção do sinal em sistemas *uplink* MIMO massivo, aproximando a inversa da matriz do canal por meio de métodos iterativos com o objetivo de reduzir a complexidade computacional [Wu et al. 2013, Yin et al. 2013, Wu et al. 2014, Gao et al. 2014, Yin et al. 2014, Dai et al. 2015, Wu et al. 2016, Qin et al. 2016, Tang et al. 2016, Lee 2017, Shahabuddin et al. 2017, Jiang et al. 2018, Lee 2019, Jin et al. 2019, Albreem et al. 2020, Albreem et al. 2020, Yakhelef and Saidi 2020, Naceur 2021].

Logo, o sistema MIMO massivo apresenta características que requisitam algoritmos de detecção de sinais eficazes e de baixa complexidade com capacidade de mitigar as interferências mútuas oriundas das outras antenas, a fim de ter integralmente a informação transmitida. Por este motivo, propõe-se um algoritmo de detecção linear de sinais que utiliza um esquema iterativo baseado no método da secante para sistemas MIMO massivo. O algoritmo proposto se difere dos demais algoritmos da literatura, por ter duas aproximações iniciais, o que garante uma rápida convergência com poucas iterações, reduzindo o número de operações matriciais e consequentemente o tempo de execução, ao mesmo tempo que mantém um desempenho quase-ótimo em termos de BER.

O restante do artigo está organizado da seguinte forma: A Seção 2 apresenta o mo-

delo do sistema MIMO massivo; A Seção 3 apresenta os trabalhos relacionados; A Seção 4 apresenta o algoritmo de detecção de sinais proposto com base no método da secante; A Seção 5 apresenta uma comparação da complexidade computacional do algoritmo de detecção proposto com os algoritmos presentes na literatura; A Seção 6 apresenta o desempenho do algoritmo de detecção de sinal proposto com os algoritmos presentes na literatura em termos da Função de Distribuição Cumulativa (CDF – *Cumulative Distribution Function*) dos autovalores da inversa da matriz de filtragem e BER ; Finalmente, na Seção 7 são apresentadas as conclusões obtidas.

2. Modelo do Sistema *Uplink* MIMO massivo

Neste trabalho, considera-se um sistema *uplink* MIMO massivo multiusuário. Assume-se que este sistema é equipado com N antenas na estação base e K usuários, cada um deles com uma antena, onde $K \ll N$, como mostra a Figura 1. Os *bits* de informação dos K usuários são codificados e mapeados nos pontos da constelação no alfabeto finito Ω com cardinalidade $|\Omega|$ e potência de transmissão média por símbolo E_s . Os símbolos modulados são transmitidos pelo canal MIMO massivo e recebidos pela estação base. Os símbolos recebidos na estação base podem ser modelados pela seguinte equação:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \quad (1)$$

onde $\mathbf{x} = (x_1, x_2, \dots, x_K)^T \in \Omega^{K \times 1}$ é o vetor de símbolos transmitidos, $\mathbf{y} = (y_1, y_2, \dots, y_N)^T \in \mathbb{C}^{N \times 1}$ é o vetor de sinal complexo recebido, $\mathbf{H} \in \mathbb{C}^{N \times K}$ representa a matriz do canal *uplink* e $\mathbf{n} \in \mathbb{C}^{N \times 1}$ denota o Ruído Aditivo Gaussiano Branco (AWGN – *Additive White Gaussian Noise*) de média zero e variância N_0 . A Relação Sinal-Ruído (SNR – *Signal to Noise Ratio*) média pode ser definida por KE_s/N_0 . A potência de transmissão do usuário $i = 1, 2, \dots, K$ é definida como $\mathbb{E}\{|x_i|^2\} = E_s, \forall i$ [Yin et al. 2014].

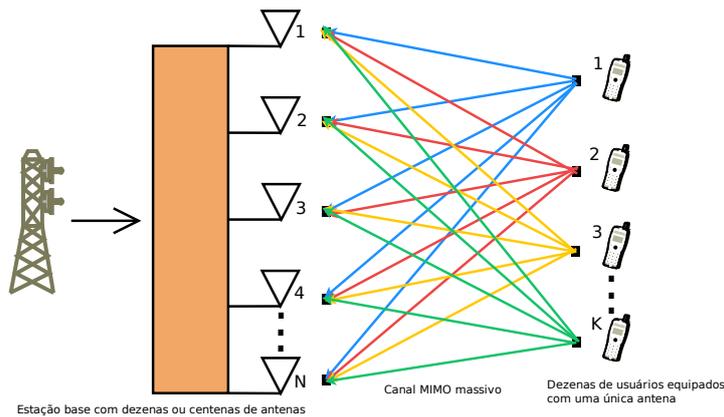


Figura 1. Exemplo de uma arquitetura do sistema MIMO massivo.

2.1. Detector Linear MMSE

Na estação base, se \mathbf{H} for conhecido, o vetor de sinal transmitido pode ser estimado a partir do sinal ruidoso recebido usando o estimador MMSE [Dai et al. 2015]:

$$\hat{\mathbf{x}} = (\mathbf{H}^H \mathbf{H} + \sigma^2 \mathbf{I}_K)^{-1} \mathbf{H}^H \mathbf{y} = \mathbf{A}^{-1} \bar{\mathbf{y}}, \quad (2)$$

onde $\hat{\mathbf{x}}$ representa uma estimativa do vetor de símbolos transmitidos, $\bar{\mathbf{y}} = \mathbf{H}^H \mathbf{y}$ representa a saída do Filtro Casado (*Matched Filter*) de \mathbf{y} e $\mathbf{A} = \mathbf{G} + \sigma^2 \mathbf{I}_K$ representa a matriz de filtragem MMSE que é definida positiva Hermitiana, onde $\mathbf{G} = \mathbf{H}^H \mathbf{H}$ é a matriz de Gram, \mathbf{I}_K é uma matriz identidade e a variância σ^2 é definida como N_0/E_s [Yin et al. 2014, Dai et al. 2015]. O algoritmo MMSE é convertido para resolver a equação linear $\hat{\mathbf{x}} = \mathbf{A}^{-1} \bar{\mathbf{y}}$. No entanto, a inversão da matriz \mathbf{A} é complexa, pois a dimensão aumenta a medida que o número de usuários K cresce. A complexidade de \mathbf{A}^{-1} é $O(K^3)$ [Albreem et al. 2019, Wu et al. 2013, Yin et al. 2013, Wu et al. 2014, Yin et al. 2014, Dai et al. 2015].

2.2. A Razão Logarítmica de Verossimilhança (LLR)

Para avaliar a estimativa da qualidade objetiva do nível de *bit*, utiliza-se o método LLR (*Log-Likelihood Ratio*). O método LLR tem se mostrado uma métrica muito eficiente para decodificação suave. Usando o método LLR, é possível reduzir consideravelmente a complexidade computacional do decodificador de canal sem prejuízos significativos em termos de BER [Studer et al. 2011].

Definindo que $\mathbf{W}^H = \mathbf{A}^{-1} \mathbf{H}^H$, e modelando o símbolo transmitido do usuário i como $\hat{x}_i = \mu_i x_i + z_i$, onde $\mu_i = \mathbf{w}_i^H \mathbf{h}_i$ representa o ganho de equalização do canal, \mathbf{w}_i^H corresponde as linhas da matriz \mathbf{W}^H , representando o vetor de filtro do algoritmo MMSE e $z_i = \sum_{j, j \neq i} \mathbf{w}_i^H \mathbf{h}_j x_j + \mathbf{w}_i^H \mathbf{n}$ representa o ruído com variância $v_i^2 = \mathbb{E}\{|z_i|^2\}$, a LLR para o *bit* b do usuário i pode ser computada por:

$$L_{i,b} = \rho_i \left(\min_{a \in \chi_b^0} \left| \frac{\hat{x}_i}{\mu_i} - a \right|^2 - \min_{a' \in \chi_b^1} \left| \frac{\hat{x}_i}{\mu_i} - a' \right|^2 \right), \quad (3)$$

onde $\rho_i = \mu_i^2 / v_i^2$ é a SINR (*Signal-to-Interference-plus-Noise Ratio*) para o i -ésimo usuário e os conjuntos χ_b^0 e χ_b^1 contém os símbolos da constelação Ω , onde o b -ésimo *bit* do símbolo é igual a 0 ou 1, respectivamente.

3. Trabalhos Relacionados

As Séries de Neumann (NS – *Neumann Series*) [Albreem et al. 2019, Wu et al. 2013, Yin et al. 2013] são um método para aproximar a inversa da matriz de filtragem com o objetivo de reduzir a complexidade do detector linear MMSE. A matriz \mathbf{A} pode ser decomposta em $\mathbf{A} = \mathbf{D} + \mathbf{E}$, onde \mathbf{D} é a matriz diagonal principal e \mathbf{E} contém os elementos fora da diagona principal. A expansão NS de \mathbf{A} é dada por:

$$\mathbf{A}^{-1} = \sum_{\theta=0}^{\Theta-1} (-\mathbf{D}^{-1} \mathbf{E})^\theta \mathbf{D}^{-1}, \text{ for } \theta = 0, 1, 2, \dots, \Theta - 1, \quad (4)$$

onde θ representa os termos em (4). À medida que k aumenta, a precisão na estimação da inversa da matriz de filtragem é alcançada, entretanto, à custa de uma maior complexidade [Wu et al. 2013, Wu et al. 2014]. Em [Yin et al. 2013], o algoritmo de detecção baseado em NS foi testado em um cenário MIMO massivo com a tecnologia OFDM (*Orthogonal Frequency Division Multiplexing*) composto por $K = 4, 8$ usuários e $N = 16, 32, 64, 128$ antenas na estação base, utilizando um esquema de modulação 16-QAM. A complexidade computacional do algoritmo de detecção baseado em NS é $O(K^3)$ para $\Theta > 2$.

O método do Gradiente Conjugado (CG – *Conjugate Gradient*) [Albreem et al. 2019, Yin et al. 2014] é outro método eficaz para resolver as equações lineares por meio de k iterações. O sinal estimado ($\hat{\mathbf{x}}$) pode ser obtido da seguinte maneira:

$$\hat{\mathbf{x}}^{(k+1)} = \hat{\mathbf{x}}^{(k)} + \alpha^{(k)} \mathbf{p}^{(k)}, \quad (5)$$

onde $\alpha^{(k)}$ é um parâmetro escalar e $\mathbf{p}^{(k)}$ representa a direção conjugada em relação à matriz \mathbf{A} , ou seja

$$\left(\mathbf{p}^{(k)}\right)^H \mathbf{A} \mathbf{p}^{(l)} = 0, \quad k \neq l. \quad (6)$$

Em [Yin et al. 2014], o algoritmo de detecção baseado no método baseado no CG foi testado em um cenário MIMO massivo composto por $K = 8, 16$ usuários e $N = 32, 128$ antenas na estação base. O método CG supera o esquema de detecção baseado em NS em termos de desempenho e complexidade. A complexidade computacional do algoritmo de detecção baseado no GC é $O((k+1)(4K^2 + 8K))$.

O método de Gauss-Seidel (GS) [Albreem et al. 2019, Dai et al. 2015] pode ser usado para resolver o sistema linear mostrado em (2). A matriz \mathbf{A} pode ser decomposta em $\mathbf{A} = \mathbf{D} + \mathbf{L} + \mathbf{U}$, onde \mathbf{D} , \mathbf{L} e \mathbf{U} representam a matriz diagonal, a matriz estritamente triangular inferior e a matriz estritamente triangular superior, respectivamente. O método GS pode ser usado para estimar o vetor do sinal transmitido ($\hat{\mathbf{x}}$), sendo dado por:

$$\hat{\mathbf{x}}^{(k+1)} = (\mathbf{D} + \mathbf{L})^{-1} (\bar{\mathbf{y}} - \mathbf{U} \hat{\mathbf{x}}^{(k)}), \quad (7)$$

para $k = 0$, se não houver informação a priori sobre a solução inicial $\hat{\mathbf{x}}^{(0)}$, essa pode ser considerada zero. Em [Dai et al. 2015], o método GS foi testado em um cenário MIMO massivo composto por $K = 16$ usuários e $N = 128$ antenas na estação base, utilizando um esquema de modulação 64-QAM. De acordo com [Albreem et al. 2019, Dai et al. 2015] o método de iteração GS supera o método NS em termos de desempenho e complexidade. A complexidade computacional do algoritmo de detecção baseado em GS é $O((k+1)4K^2 + 4K)$.

Recentemente, Albreem *et al.* propuseram um algoritmo de detecção linear de sinais [Albreem et al. 2020], que explora os benefícios dos métodos de Jacobi (JA) e GS. Neste método é utilizada uma matriz escada (\mathbf{S}) com o objetivo de garantir uma taxa de convergência rápida. O método proposto em [Albreem et al. 2020] é construído da seguinte maneira:

$$\hat{\mathbf{x}}_{(0)} = \mathbf{S}^{-1} \bar{\mathbf{y}}, \quad (8)$$

onde $\hat{\mathbf{x}}_{(0)}$ representa a solução inicial calculada com base na matriz escada. Posteriormente, a primeira iteração é conduzida utilizando o método de Jacobi da seguinte maneira:

$$\hat{\mathbf{x}}_{(1)} = \mathbf{D}^{-1} (\bar{\mathbf{y}} + (\mathbf{D} - \mathbf{A}) \hat{\mathbf{x}}_{(0)}). \quad (9)$$

Na próxima etapa, ocorre um processo de refinamento com base no método de Gauss-Seidel [Albreem et al. 2020]. O método proposto por Albreem *et al.* foi testado em [Albreem et al. 2020] em um cenário MIMO massivo composto por $K = 30, 40, 50, 60$ usuários e $N = 160$ antenas na estação base, utilizando um esquema de modulação 64-QAM. A complexidade computacional do algoritmo de detecção é $O(K^2(5 + 4k) + K - 3)$.

4. Proposta de Algoritmo de Detecção de Sinais Baseado no Método da Secante

Na abordagem proposta, primeiramente, aproxima-se a inversa da matriz \mathbf{A} por meio do algoritmo iterativo proposto recentemente por [Artidiello et al. 2020], que baseia-se no método da secante. O método da secante é dado por:

$$x_{k+1} = x_k - \frac{f(x_k)}{\alpha_k}, \quad (10)$$

onde α_k satisfaz $f(x_k) - f(x_{k-1}) = \alpha_k(x_k - x_{k-1})$, $k \geq 0$. Para obter o processo iterativo, substitui-se (10) na função $f(x) = x^{-1} - \mathbf{A}$. Assim, obtém-se o algoritmo iterativo do método da secante para aproximação da matriz inversa, dado por:

$$\mathbf{X}_{k+1} = \mathbf{X}_{k-1} + \mathbf{X}_k - \mathbf{X}_{k-1} \mathbf{A} \mathbf{X}_k, \quad (11)$$

para $k = 0$, \mathbf{X}_{-1} e \mathbf{X}_0 são aproximações iniciais. Quando $k \rightarrow \infty$, $\lim_{k \rightarrow \infty} \mathbf{X}_k = \mathbf{A}^{-1}$. A convergência do método é garantida por $\|\mathbf{X}_{k+1} - \mathbf{A}^{-1}\|_2 \leq \sqrt{\bar{n}m_k} \|\mathbf{X}_k - \mathbf{A}^{-1}\|_2$, onde $m_k = \max_{1 \leq j \leq n} \{c_k^j\}$, $c_k^j \geq 0$, $\forall k$ [Artidiello et al. 2020].

Um bom valor inicial pode tornar a taxa de convergência mais rápida, o que diminui o número de iterações, reduzindo o número de operações matriciais. Logo, para a abordagem proposta, \mathbf{X}_0 é dado por:

$$\mathbf{X}_0 = \alpha \mathbf{I}_K + \phi \mathbf{A}, \quad (12)$$

onde

$$\alpha = \frac{2\eta(1+\eta)}{N(1+\eta^2)}, \quad \phi = -\frac{\eta^2}{N^2(1+\eta^2)}, \quad (13)$$

para $\eta = \frac{N}{K}$. Neste caso, α e ϕ dependem apenas do número de usuários e de antenas na estação base [Minango and de Almeida 2018]. A matriz \mathbf{X}_1 é dada por:

$$\mathbf{X}_{-1} = \Delta_1 \mathbf{I}_K - \Delta_2 \mathbf{A}, \quad (14)$$

onde $\Delta_1 = \frac{4\delta}{2\delta^2 - \varphi^2}$, $\Delta_2 = \frac{2}{2\delta^2 - \varphi^2}$, $\delta = \frac{\max(\lambda) + \min(\lambda)}{2}$, $\varphi = \max(\lambda) - \delta$ e λ representa os autovalores de \mathbf{A} .

Substituindo (12) e (14) em (11), \mathbf{A}^{-1} pode ser aproximado por \mathbf{X}_{k+1} onde $k \leq 1$ (Proposta 1). A matriz de filtragem (\mathbf{A}) é diagonalmente dominante para sistemas *uplink* MIMO massivo [Dai et al. 2015], conforme ilustrado na Figura 2. Além disso, percebe-se que os termos fora da diagonal da matriz de filtragem são menos influentes quando comparados aos termos da diagonal principal. Consequentemente, as características da matriz do canal são mantidas na diagonal principal. Neste caso, usar uma matriz escada (mais detalhes podem ser obtidos no Apêndice A) como aproximação inicial pode ser uma opção interessante para reduzir o número de operações matriciais, ou seja, $\mathbf{X}_0 = \text{stair}(\alpha \mathbf{I}_K + \phi \mathbf{A})$ (Proposta 2), mantendo bom desempenho na detecção de sinal. Substituindo (11) em (2) o algoritmo proposto é obtido. Finalmente, substituindo \mathbf{A}^{-1} por \mathbf{X}_{k+1} , $\mathbf{W} = \mathbf{X}^{k+1} \mathbf{H}^H$, a LLR pode ser calculada usando (3). A Proposta 2 é representada pelo **Algoritmo 1**.

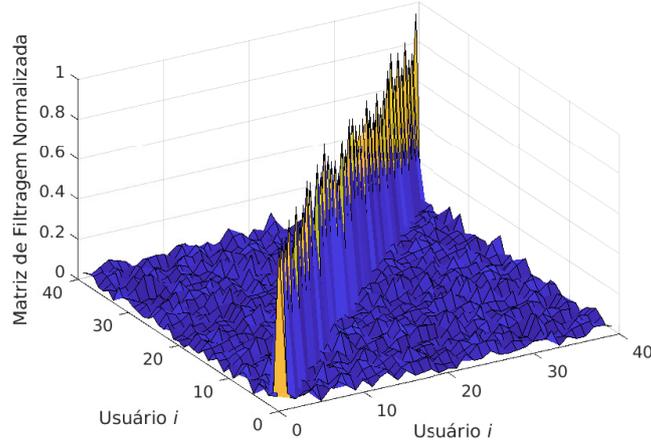


Figura 2. Matriz de filtragem (A) normalizada com 40 usuários.

Algoritmo 1 Proposta 2

- 1: **Entrada:** \mathbf{H} , σ^2 , \mathbf{y} , K , N .
 - 2: **Saída:** $\hat{\mathbf{x}}$, $L_{i,b}$.
 - 3: **Inicializar:** $\mathbf{G} = \mathbf{H}^H \mathbf{H}$, $\mathbf{A} = \mathbf{G} + \sigma^2 \mathbf{I}_K$, $\bar{\mathbf{y}} = \mathbf{H}^H \mathbf{y}$, $\eta = \frac{N}{K}$, $\alpha = \frac{2\eta(1+\eta)}{N(1+\eta^2)}$, $\phi = -\frac{\eta^2}{N^2(1+\eta^2)}$, $\Delta_1 = \frac{4\delta}{2\delta^2 - \phi^2}$, $\Delta_2 = \frac{2}{2\delta^2 - \phi^2}$;
 - 4: $\mathbf{X}_{-1} = \Delta_1 \mathbf{I}_K - \Delta_2 \mathbf{A}$;
 - 5: $\mathbf{X}_0 = \text{stair}(\alpha \mathbf{I}_K + \phi \mathbf{A})$; ▷ Utilizar Matriz Escada (S)
 - 6: **for** $k = 0, 1$ **do**
 - 7: $\mathbf{X}_{k+1} = \mathbf{X}_{k-1} + \mathbf{X}_k - \mathbf{X}_{k-1} \mathbf{A} \mathbf{X}_k$;
 - 8: **end**
 - 9: $\hat{\mathbf{x}} = \mathbf{X}_{k+1} \bar{\mathbf{y}}$; ▷ Detecção do Sinal
 - 10: **Computar LLR:** $\mathbf{W}^H = \mathbf{X}_{k+1} \mathbf{H}^H$, ρ_i , μ_i , χ_b^0 , χ_b^1 ;
 - 11: $L_{i,b} = \rho_i \left(\min_{a \in \chi_b^0} \left| \frac{\hat{x}_i}{\mu_i} - a \right|^2 - \min_{a' \in \chi_b^1} \left| \frac{\hat{x}_i}{\mu_i} - a' \right|^2 \right)$.
-

O **Algoritmo 1** estima o sinal transmitido (\mathbf{x}) em apenas duas iterações (ou seja, $k = 1$), e tem como principal vantagem, possuir duas aproximações iniciais. Como já foi dito, isso acelera a convergência do método, reduzindo o número de operações matriciais e tornando o método proposto mais simples de ser implementado em *hardware*. A solução inicial \mathbf{X}_0 , por depender apenas do número de usuários (K) e antenas na estação base (N) mostra que o **Algoritmo 1** pode adaptar-se a qualquer cenário de simulação MIMO massivo que envolva uma variação do número de usuários e antenas na estação base.

5. Complexidade Computacional

Sabe-se, que uma baixa complexidade computacional é fundamental para que os algoritmos de detecção de sinais sejam implementados de forma prática (em *hardware*) em cenários que envolvam sistemas MIMO massivo multiusuário. A complexidade computacional das Propostas 1 e 2 foram obtidas através do cálculo do número de operações

matriciais de multiplicação e adição. Assim, a Proposta 1 obteve uma complexidade computacional $O(10K^2)$. Ao decompor a matriz de solução inicial \mathbf{X}_0 utilizando o método da matriz escada, foi possível reduzir o número de operações matriciais para $O(5K^2)$ (Proposta 2, representada pelo **Algoritmo 1**). Como as Propostas 1 e 2 utilizam apenas duas iterações para estimar a inversa da matriz de filtragem, calculou-se a complexidade computacional para duas e três iterações dos seguintes algoritmos da literatura: NS, CG, GS e Albreem *et al.* com o objetivo de comparar com as Propostas 1 e 2 (Tabela 1). Analisando a Tabela 1, percebe-se que todos os métodos estão com complexidade na ordem $O(K^2)$, entretanto, a Proposta 2, dentre todos os algoritmos, é o método que utiliza o menor número de operações matriciais para o cálculo da inversa da matriz \mathbf{A} . A Tabela 2 mostra o tempo de execução (em segundos) dos algoritmos presentes na literatura e das Propostas 1 e 2, variando o número de usuários (K). O tempo de execução foi calculado a partir da média de 300 simulações. Percebe-se que os algoritmos que utilizam um número menor de operações para estimar a inversa da matriz de filtragem e o sinal transmitido (Tabela 1) tiveram um tempo de execução menor, conforme o esperado. Pode-se observar que a Proposta 2, devido ao menor número de operações matriciais, obteve um tempo de execução ligeiramente menor que a Proposta 1 e os demais algoritmos da literatura com complexidade quadrática. Além disso, vale ressaltar o alto tempo de execução dos métodos MMSE e Neumann (para $\Theta = 3$), ambos com complexidade cúbica, principalmente para $K = 2^{11}$ usuários, tornando-os inviáveis para implementação em *hardware* para um grande número de usuários.

Tabela 1. Complexidade computacional dos algoritmos de detecção linear com duas e três iterações.

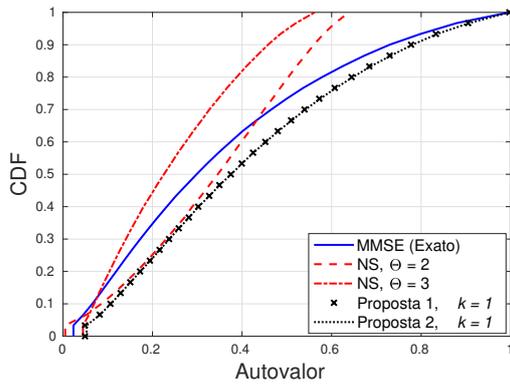
Métodos	$k = 1$	$k = 2$
NS [Yin et al. 2013]	$O(6K^2 - 6K)$, for $\Theta = 2$	$O(K^3)$, for $\Theta = 3$
CG [Yin et al. 2014]	$O(8K^2 + 8K)$	$O(12K^2 + 8K)$
GS [Dai et al. 2015]	$O(8K^2 + 4K)$	$O(12K^2 + 4K)$
Albreem <i>et al.</i> [Albreem et al. 2020]	$O(9K^2 + K - 3)$	$O(13K^2 + K - 3)$
Proposta 1	$O(10K^2)$	–
Proposta 2	$O(5K^2)$	–

6. Resultados

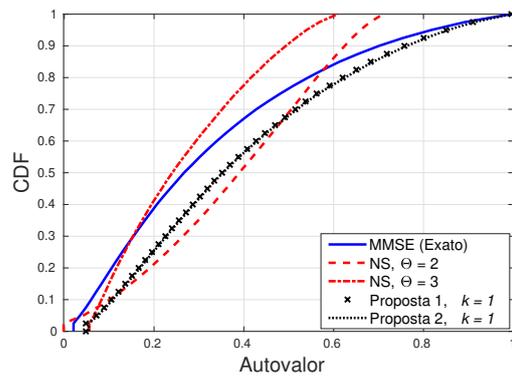
Nesta seção, a precisão da estimação do sinal transmitido é avaliada em termos da função de distribuição cumulativa dos autovalores da inversa da matriz de filtragem e BER para as Propostas 1 e 2. Além disso, seu desempenho é comparado aos seguintes algoritmos: MMSE (Exato), NS [Yin et al. 2013], GC [Yin et al. 2014], GS [Dai et al. 2015] e método de Albreem *et al.* [Albreem et al. 2020]. Os algoritmos da literatura foram simulados com duas e três iterações ($k = 1$ e $k = 2$, respectivamente). Neste trabalho, o cenário de simulação *uplink* MIMO massivo é composto por: $K = 30, 40, 50$ e 60 usuários; $N = 256$ antenas na estação base; o esquema de modulação 64-QAM foi adotado; e utilizando 300 simulações de Monte-Carlo. Os resultados foram obtidos usando o *software* Matlab R2020a e as seguintes configurações de *hardware*: Processador Intel Core i5-9300H 2,40 GHz, 16 GB de memória RAM e placa gráfica GeForce 1650 com 4 GB de memória dedicada.

Tabela 2. Tempo de execução em segundos (s) dos algoritmos de detecção linear com duas ($k = 1$) e ($k = 2$) três iterações.

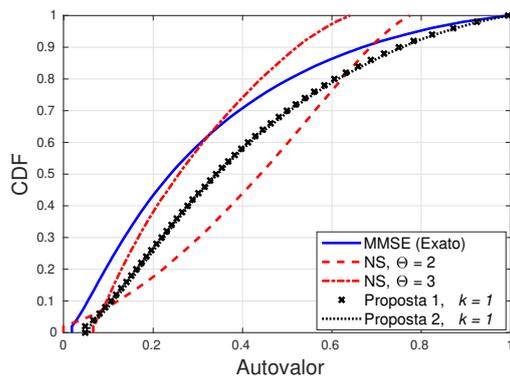
Métodos	$K = 2^5$	$K = 2^7$	$K = 2^9$	$K = 2^{11}$
MMSE (Exato)	$3,0 \times 10^{-4}$	$1,9 \times 10^{-3}$	$2,1 \times 10^{-2}$	1,7879
NS, $\Theta = 3$	$4,0 \times 10^{-4}$	$4,5 \times 10^{-3}$	$6,6 \times 10^{-2}$	2,2701
NS, $\Theta = 2$	$1,6 \times 10^{-6}$	$9,8 \times 10^{-6}$	$2,1 \times 10^{-4}$	$6,2 \times 10^{-3}$
CG, $k = 1$	$1,7 \times 10^{-6}$	$1,7 \times 10^{-5}$	$3,0 \times 10^{-4}$	$7,8 \times 10^{-3}$
CG, $k = 2$	$4,1 \times 10^{-6}$	$5,0 \times 10^{-5}$	$4,7 \times 10^{-4}$	$4,2 \times 10^{-2}$
GS, $k = 1$	$1,5 \times 10^{-6}$	$1,7 \times 10^{-5}$	$2,8 \times 10^{-4}$	$7,8 \times 10^{-3}$
GS, $k = 2$	$3,6 \times 10^{-6}$	$4,9 \times 10^{-5}$	$4,8 \times 10^{-4}$	$3,9 \times 10^{-2}$
Albreem <i>et al.</i> , $k = 1$	$1,9 \times 10^{-6}$	$2,3 \times 10^{-5}$	$3,5 \times 10^{-4}$	$8,7 \times 10^{-3}$
Albreem <i>et al.</i> , $k = 2$	$5,6 \times 10^{-6}$	$6,2 \times 10^{-5}$	$5,7 \times 10^{-4}$	$5,3 \times 10^{-2}$
Proposta 1	$2,6 \times 10^{-6}$	$3,0 \times 10^{-5}$	$4,2 \times 10^{-4}$	$9,7 \times 10^{-3}$
Proposta 2	$9,7 \times 10^{-7}$	$7,8 \times 10^{-6}$	$9,9 \times 10^{-5}$	$4,9 \times 10^{-3}$



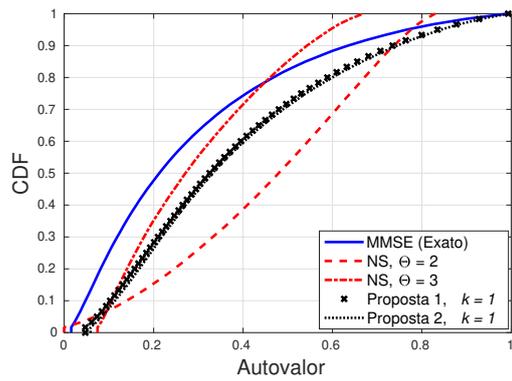
(a) 30 usuários



(b) 40 usuários



(c) 50 usuários



(d) 60 usuários

Figura 3. Distribuição normalizada dos autovalores da matriz A^{-1} .

Os autovalores da matriz do canal MIMO podem ser interpretados como os ganhos de potência dos canais paralelos. Em comunicações digitais, a distribuição de autovalores determina a capacidade do sistema. Neste trabalho, foi avaliada a distribuição dos

autovalores da inversa da matriz filtragem, uma vez que os autovalores da matriz inversa são os inversos dos autovalores da matriz original. Portanto, avaliando a distribuição dos autovalores da matriz inversa, pode-se obter o desempenho de estimação da inversa da matriz filtragem através das Propostas 1 e 2, e o impacto na BER. A Figura 3 mostra o CDF dos autovalores do método MMSE, Neumann e as Propostas 1 e 2. Analisando a Figura 3, pode-se observar que o método de Neumann possui uma distribuição de autovalores que difere do método MMSE, especialmente quando o número de usuários aumenta. As Propostas 1 e 2, por outro lado, apresentaram essencialmente a mesma distribuição, e mantiveram quase o mesmo comportamento que a CDF dos autovalores do método MMSE, mesmo com o aumento do número de usuários.

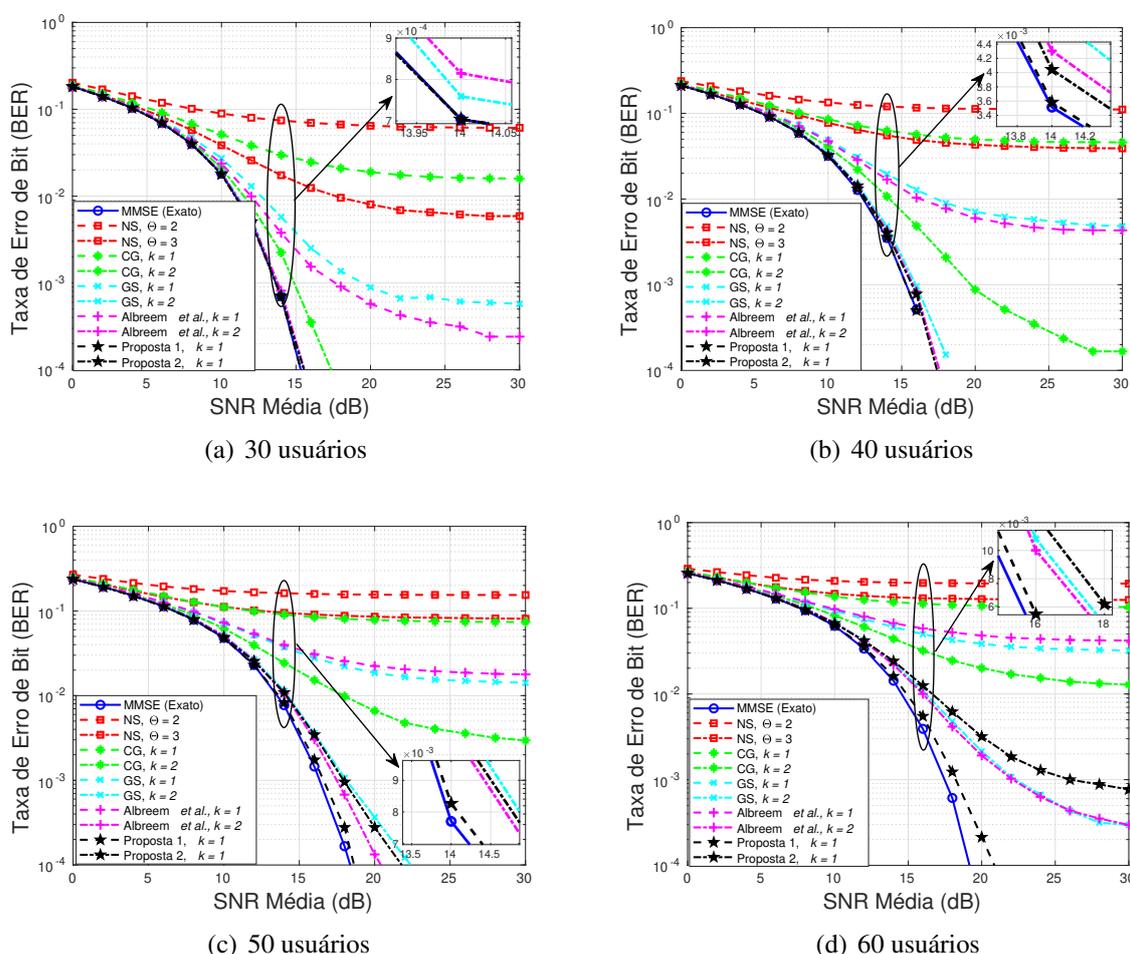


Figura 4. BER versus SNR média.

Considerando as curvas BER-SNR representadas na Figura 4, pode-se observar que as Propostas 1 e 2 (com apenas duas iterações) tiveram um melhor desempenho em termos de BER em relação aos métodos NS, CG (com três iterações e maior número de operações, como mostrado na Tabela 1), e aos métodos GS e Albreem *et al.* (com duas iterações). Analisando os métodos GS e Albreem *et al.* (com três iterações), houve avanços em termos de redução de BER (mas a um custo de um maior número de operações matriciais), obtendo desempenhos semelhantes aos das Propostas 1 e 2, apesar da vantagem da Proposta 1, principalmente com 60 usuários. Ainda que a Proposta 2 tenha duas

vezes menos operações de matriz em comparação aos outros métodos, em geral, ela manteve uma boa precisão na estimativa do sinal transmitido, embora em raros casos ficando atrás dos métodos GS e Albreem *et al.*.

Sabe-se, que as redes 5G requerem algoritmos de baixa complexidade a fim de reduzir o atraso de processamento do *hardware* e, conseqüentemente, reduzir o consumo da bateria dos terminais móveis. Portanto, as Propostas 1 e 2 atendem a esse requisito, utilizando menos interações e operações matriciais, porém, sem perder o bom desempenho na estimação do sinal.

7. Conclusão

Neste artigo foi proposto um algoritmo de detecção de sinais para sistemas *uplink* MIMO massivo multiusuário. Aplicou-se um método iterativo para o cálculo da inversa da matriz de filtragem do algoritmo MMSE. O método da secante tem como vantagem possuir duas aproximações iniciais, o que garante uma rápida convergência e reduz o número de operações matriciais. Utilizando a abordagem proposta foi possível reduzir o número de operações matriciais e conseqüentemente o tempo de execução, porém, mantendo um desempenho “quase-ótimo” na estimação do sinal transmitido (\mathbf{x}), apresentando, de maneira geral, um desempenho superior aos principais algoritmos de detecção linear da literatura. Logo, as Propostas 1 e 2 tornam-se soluções promissoras para aplicações práticas envolvendo a tecnologia MIMO massivo.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior -Brasil (CAPES) - Código de Financiamento 001.

Apêndice

A. Matriz Escada e suas Propriedades

A Matriz Escada (*Stair Matrix*) é uma matriz tridiagonal especial. A matriz tridiagonal é uma matriz que possui elementos diferentes de zero na diagonal principal, na primeira diagonal abaixo da diagonal principal e na primeira diagonal acima da mesma. A matriz de escada é utilizada para garantir uma taxa de convergência rápida. Na literatura, a maioria dos métodos iterativos utilizam a matriz diagonal, entretanto, não há garantia de se atingir uma taxa de convergência satisfatória, o que pode comprometer o desempenho no cálculo da matriz inversa [Lu 1999, Jiang et al. 2018].

Definição A.1. *Se uma matriz tridiagonal satisfaz uma das seguintes condições:*

- (I) $\mathbf{S}_{(u,u-1)} = 0$, $\mathbf{S}_{(u,u+1)} = 0$, onde $u = 2v - 1$, para $v = 1, 2, \dots, \lfloor (K+1)/2 \rfloor$. Alternadamente, os elementos não diagonais nas linhas ímpares da matriz tridiagonal são zeros;
- (II) $\mathbf{S}_{(u,u-1)} = 0$, $\mathbf{S}_{(u,u+1)} = 0$, onde $u = 2v$, para $v = 1, 2, \dots, \lfloor K/2 \rfloor$. Em outras palavras, os elementos não diagonais nas linhas pares da matriz tridiagonal são zeros;

então pode-se considerá-la como uma matriz escada, denotada por $\mathbf{S} = \text{stair}(\mathbf{S}_{(u,u-1)}, \mathbf{S}_{(u,u)}, \mathbf{S}_{(u,u+1)})$. Uma matriz escada é do tipo I se a condição (I) for

satisfeita e é do tipo II se a condição (II) for satisfeita. Por exemplo, uma matriz escada 6×6 pode ser construída da seguinte forma:

$$\mathbf{S} = \begin{bmatrix} \times & \times & 0 & 0 & 0 & 0 \\ 0 & \times & 0 & 0 & 0 & 0 \\ 0 & \times & \times & \times & 0 & 0 \\ 0 & 0 & 0 & \times & 0 & 0 \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & 0 & \times \end{bmatrix} \quad \text{ou} \quad \mathbf{S} = \begin{bmatrix} \times & 0 & 0 & 0 & 0 & 0 \\ \times & \times & \times & 0 & 0 & 0 \\ 0 & 0 & \times & 0 & 0 & 0 \\ 0 & 0 & \times & \times & \times & 0 \\ 0 & 0 & 0 & 0 & 0 & \times \\ 0 & 0 & 0 & 0 & \times & \times \end{bmatrix}. \quad (15)$$

A matriz escada do lado esquerdo é do tipo I e do lado direito é do tipo II. A seguir, é apresentada as seguintes propriedades da matriz escada nos Corolários A.1 e A.2.

Corolário A.1. *Seja \mathbf{S} uma matriz escada. Então, \mathbf{S}^H também é uma matriz escada. Além disso, se \mathbf{S} é do tipo I, então \mathbf{S}^H é do tipo II e vice-versa.*

Demonstração. Usando a **Definição A.1**, obtém-se o **Corolário A.1**. □

Corolário A.2. *Seja \mathbf{S} uma matriz escada. \mathbf{S} é uma matriz não singular se e somente se $\mathbf{S}_{(u,u)}$, para $u = 1, 2, \dots, K$, for diferente de zero. Além disso, se \mathbf{S} for não singular, então $\mathbf{S}^{-1} = \mathbf{D}^{-1} (2\mathbf{D} - \mathbf{S}) \mathbf{D}^{-1}$.*

Demonstração. Como $\det(\mathbf{S}) = \prod_{u=1}^K \mathbf{S}_{(u,u)}$, percebe-se que \mathbf{S} é não singular se e somente se $\mathbf{S}_{(u,m)}$, para $u = 1, 2, \dots, K$, for diferente de zero. □

A partir do **Corolário A.2**, tem-se o **Algoritmo 2** para obter \mathbf{S}^{-1} . A complexidade para obter a inversa da matriz escada é $O(K)$ utilizando o **Algoritmo 2**, sendo da mesma ordem de complexidade da inversa da matriz diagonal \mathbf{D}^{-1} .

Algoritmo 2 Computar a Inversa da Matriz Escada

- 1: **Entrada:** $\mathbf{S} = \text{stair}(\mathbf{S}_{(u,u-1)}, \mathbf{S}_{(u,u)}, \mathbf{S}_{(u,u+1)})$.
 - 2: **Saída:** $\mathbf{S}^{-1} = \tilde{\mathbf{S}} = \text{stair}(\tilde{\mathbf{S}}_{(u,u-1)}, \tilde{\mathbf{S}}_{(u,u)}, \tilde{\mathbf{S}}_{(u,u+1)})$.
 - 3: **for** $u = 1 : 1 : K$ **do**
 - 4: $\tilde{\mathbf{S}}_{(u,u)} = 1/\mathbf{S}_{(u,u)}$;
 - 5: **end**
 - 6: **for** $u = 2 : 2 : 2 \lfloor K/2 \rfloor$ **do**
 - 7: $\tilde{\mathbf{S}}_{(u,u-1)} = -\mathbf{S}_{(u,u-1)} \tilde{\mathbf{S}}_{(u,u)} \tilde{\mathbf{S}}_{(u-1,u-1)}$;
 - 8: $\tilde{\mathbf{S}}_{(u,u+1)} = -\mathbf{S}_{(u,u+1)} \tilde{\mathbf{S}}_{(u,u)} \tilde{\mathbf{S}}_{(u+1,u+1)}$;
 - 9: **end**
-

Referências

- Albreem, M., Alsharif, M., and Kim, S. (2020). A robust hybrid iterative linear detector for massive mimo uplink systems. *Symmetry*, 12(306):1–12.
- Albreem, M., Juntti, M., and Shahabuddin, S. (2020). Efficient initialisation of iterative linear massive mimo detectors using a stair matrix. *Electronics Letters*, 56(1):50–52.

- Albreem, M. A., Juntti, M., and Shahabuddin, S. (2019). Massive mimo detection techniques: A survey. *IEEE Communications Surveys Tutorials*, 21(4):3109–3132.
- Artidiello, S., Cordero, A., Torregrosa, J., and Vassileva, M. (2020). Generalized inverses estimations by means of iterative methods with memory. *Mathematics*, 8(2):1–13.
- Björnson, E., Hoydis, J., and Sanguinetti, L. (2017). Massive mimo networks: Spectral, energy, and hardware efficiency.
- Chataut, R. and Akl, R. (2020). Systems for 5g and beyond networks—overview, recent trends, challenges, and future research direction. *Sensors*, 20(10):2753.
- Dai, L., Gao, X., Su, X., Han, S., I, C., and Wang, Z. (2015). Low-complexity soft-output signal detection based on gauss–seidel method for uplink multiuser large-scale mimo systems. *IEEE Transactions on Vehicular Technology*, 64(10):4839–4845.
- Gao, X., Dai, L., Ma, Y., and Wang, Z. (2014). Low-complexity near-optimal signal detection for uplink large-scale mimo systems. *Electronics Letters*, 50(18):1326–1328.
- Jiang, F., Li, C., Gong, Z., and Su, R. (2018). Stair matrix and its applications to massive mimo uplink data detection. *IEEE Transactions on Communications*, 66(6):2437–2455.
- Jin, F., Liu, Q., Liu, H., and Wu, P. (2019). A low complexity signal detection scheme based on improved newton iteration for massive mimo systems. *IEEE Communications Letters*, 23(4):748–751.
- Khan, R., Kumar, P., Jayakody, D. N. K., and Liyanage, M. (2020). A survey on security and privacy of 5g technologies: Potential solutions, recent advancements, and future directions. *IEEE Communications Surveys Tutorials*, 22(1):196–248.
- Lee, Y. (2017). Decision-aided jacobi iteration for signal detection in massive mimo systems. *Electronics Letters*, 53(23):1552–1554.
- Lee, Y. (2019). Hybrid kacmarz and coordinate-descent iterations for signal detection in massive mimo systems. *Electronics Letters*, 55(11):665–667.
- Lu, H. (1999). Stair matrices and their generalizations with applications to iterative methods i: A generalization of the successive overrelaxation method. *SIAM Journal on Numerical Analysis*, 37(1):1–17.
- Minango, J. and de Almeida, C. (2018). Low complexity zero forcing detector based on newton-schultz iterative algorithm for massive mimo systems. *IEEE Transactions on Vehicular Technology*, 67(12):11759–11766.
- Naceur, A. (2021). Damped jacobi methods based on two different matrices for signal detection in massive mimo uplink. *Journal of Microwaves, Optoelectronics and Electromagnetic Applications*, 20(1):92–104.
- Qin, X., Yan, Z., and He, G. (2016). A near-optimal detection scheme based on joint steepest descent and jacobi method for uplink massive mimo systems. *IEEE Communications Letters*, 20(2):276–279.
- Shahabuddin, S., Juntti, M., and Studer, C. (2017). Admm-based infinity norm detection for large mu-mimo: Algorithm and vlsi architecture. *2017 IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 1–4.

- Studer, C., Fateh, S., and Seethaler, D. (2011). Asic implementation of soft-input soft-output mimo detection using mmse parallel interference cancellation. *IEEE Journal of Solid-State Circuits*, 46(7):1754–1765.
- Tang, C., Liu, C., Yuan, L., and Xing, Z. (2016). High precision low complexity matrix inversion based on newton iteration for data detection in the massive mimo. *IEEE Communications Letters*, 20(3):490–493.
- Wu, M., Dick, C., Cavallaro, J. R., and Studer, C. (2016). High-throughput data detection for massive mu-mimo-ofdm using coordinate descent. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 63(12):2357–2367.
- Wu, M., Yin, B., Vosoughi, A., Studer, C., Cavallaro, J. R., and Dick, C. (2013). Approximate matrix inversion for high-throughput data detection in the large-scale mimo uplink. *2013 IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 2155–2158.
- Wu, M., Yin, B., Wang, G., Dick, C., Cavallaro, J. R., and Studer, C. (2014). Large-scale mimo detection for 3gpp lte: Algorithms and fpga implementations. *IEEE Journal of Selected Topics in Signal Processing*, 8(5):916–929.
- Yakhelef, M. and Saidi, L. (2020). Low-complexity iterative detection based on parametric aor for uplink massive mimo systems. *Telecommunications and Radio Engineering*, 79(18):1609–1623.
- Yin, B., Wu, M., Cavallaro, J. R., and Studer, C. (2014). Conjugate gradient-based soft-output detection and precoding in massive mimo systems. *2014 IEEE Global Communications Conference*, pages 3696–3701.
- Yin, B., Wu, M., Studer, C., Cavallaro, J. R., and Dick, C. (2013). Implementation trade-offs for linear detection in large-scale mimo systems. *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 2679–2683.
- Zhang, J., Björnson, E., Matthaiou, M., Ng, D. W. K., Yang, H., and Love, D. J. (2020). Prospective multiple antenna technologies for beyond 5g. *IEEE Journal on Selected Areas in Communications*, 38(8):1637–1660.