Fast matrix inversion based on Chebyshev acceleration for linear detection in massive MIMO systems

Salah Berra,[1,2,✉] Rui Dinis,[3] and Shahriar Shahabuddin[4]

[1] *COPELABS, Universidade Lusófona de Humanidades e Tecnologias, Lisboa, Portugal*

[2] *Department of Electronic and Telecommunications, Electrical Engineering Laboratory (LAGE), Kasdi Merbah University Ouargla, Algeria, Ouargla*

[3] *Instituto de Telecomunicações, FCT-UNL, Campus de Caparica, Caparica, Portugal*

[4] *Mobile Networks, Nokia, Oulu, Finland*

✉ Email: salahberra39@gmail.com, f6813@ulusofona.pt

To circumvent the prohibitive complexity of linear minimum mean square error detection in a massive multiple-input multiple-output system, several iterative methods have been proposed. However, they can still be too complex and/or lead to non-negligible performance degradation. In this letter, a Chebyshev acceleration technique is proposed to overcome the limitations of iterative methods, accelerating the convergence rates and enhancing the performance. The Chebyshev acceleration method employs a new vector combination, which combines the spectral radius of the iteration matrix with the receiver signal, and also the optimal parameters of Chebyshev acceleration have also been defined. A detector based on iterative algorithms requires pre-processing and initialisation, which enhance the convergence, performance, and complexity. To influence the initialisation, the stair matrix has been proposed as the first start of iterative methods. The performance results show that the proposed technique outperforms state-of-the-art methods in terms of error rate performance, while significantly reducing the computational complexity.

Introduction: Massive MIMO is a key technology for fifth-generation (5G) communication systems to achieve higher data rate, lower latency, higher reliability, and more connected users [1, 2]. Despite all of its advantages, massive MIMO requires complex transceiver signal processing. MIMO symbol detection is one of the most complex blocks of the baseband processing chain. Linear detection algorithms such as zero-forcing (ZF) and means square error (MMSE) schemes were proposed as reduced-complexity alternatives to optimum detection. However, they still involve matrix inversions, whose complexity can be very high, especially in massive MIMO scenarios. Recently, several iterative methods have been proposed to reduce the complexity of inversion matrix by one magnitude order [3], such Jacobi (JA), successive over relaxation (SOR) and, Gauss–Seidal (GS), and Richardson iteration (RI), but the performance achieved with these methods can still be visibly worse than the performance of the corresponding linear schemes, especially when the number of iterations is not high enough (which corresponds to the scenario of interest, since the complexity increases with the number of iterations).

In this letter, we propose an acceleration method that is combined with recent iterative methods to improve the convergence. Our method is able to approach the linear MMSE performance with much lower computational complexity. We use a Chebyshev acceleration [4] that rewrites the iterative method with a new vector combination using the spectral radius of iteration matrix and part of the iterative method. To influence the initialisation, the stair matrix [5] has been proposed as the first start of iterative methods. Our performance results show that the Chebyshev acceleration combined with iterative methods are able to approach the linear MMSE performance, outperforming state-of-the-art methods in terms of error rate performance and computational complexity.

System model: We consider the uplink transmission of a massive MIMO system employing N antennas at a base station that serves K users over the same time-frequency resources. The received signal can be written as size-N vector

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{1}$$

where $\mathbf{x}$ is the transmitted signal for the K users (i.e. a size-K vector), $\mathbf{H} \in \mathbb{C}^{N\times K}$ indicates the uplink channel matrix (we assume Rayleigh fading) between the user and the BS, and $\mathbf{n}$ is the additive white Gaussian noise (AWGN) with distribution $\mathcal{CN}(0, \sigma^2)$.

To estimate the transmitted signal we use a linear MMSE receiver, leading to

$$\hat{\mathbf{x}} = (\mathbf{H}^H\mathbf{H} + \sigma_n^2\mathbf{I}_N)^{-1}\mathbf{H}^H\hat{\mathbf{y}} = \mathbf{W}^{-1}\mathbf{y}^{MF}, \tag{2}$$

where $\hat{\mathbf{y}} = \mathbf{H}^H\mathbf{y}$ is corresponding the matched-filter output of $\mathbf{y}$, and the MMSE weight matrix $\mathbf{W}$ is denoted by

$$\mathbf{W} = \mathbf{G} + \sigma_n^2\mathbf{I}_N, \tag{3}$$

Calculating the MMSE inversion matrix $\mathbf{W}^{-1}$ is the main challenge, since it has cubic complexity.

Proposed method: As already pointed out, the convergence of iterative methods like Gauss–Seidel (GS) Jacobi (JA), successive over relaxation (SOR), and Richardson iteration (RI) can be relatively slow. By employing Chebyshev acceleration we can improve the convergence rates. To do this, the matrix to be inverted, $\mathbf{W}$, is decomposed as

$$\mathbf{W} = \mathbf{D} - \mathbf{L} - \mathbf{U}, \tag{4}$$

where $\mathbf{D} = \mathrm{diag}(\mathbf{W})$, $\mathbf{L}$ and $\mathbf{U}$ are strictly lower and strictly upper triangular matrices, respectively.

The mathematical description of the different iterative methods for the ith iteration is:

- GS-based approach

$$\mathbf{x}^{i+1} = (\mathbf{D} - \mathbf{L})^{-1}\mathbf{U}\mathbf{x}^i + \mathbf{U}\mathbf{y}^{MF} = \mathbf{B}\mathbf{x}^i + \mathbf{C}. \tag{5}$$

- JA approach

$$\mathbf{x}^{i+1} = \mathbf{D}^{-1}(\mathbf{D} - \mathbf{W})\mathbf{x}^i + \mathbf{D}^{-1}\mathbf{y}^{MF} = \mathbf{B}\mathbf{x}^i + \mathbf{C}. \tag{6}$$

- SOR-based approach

$$\begin{aligned}\mathbf{x}^{i+1} &= (\mathbf{D} - \omega\mathbf{L})^{-1}((1-\omega)\mathbf{D} + \omega\mathbf{U})\mathbf{x}^i \\ &\quad + (\mathbf{D} - \omega\mathbf{L})^{-1}\mathbf{y}^{MF} = \mathbf{B}\mathbf{x}^i + \mathbf{C}\end{aligned} \tag{7}$$

- Richardson-based approach

$$\mathbf{x}^{i+1} = (\mathbf{I} - \omega\mathbf{W})\mathbf{x}^i + \omega\mathbf{y}^{MF} = \mathbf{B}\mathbf{x}^i + \mathbf{C}. \tag{8}$$

In the previous equations, ω of SOR belongs to the interval [0, 2]) and ω of Richardson belongs to the interval $[0, 2/\lambda]$, where λ is the largest eigenvalue [3]. In the initial solution $\mathbf{x}^0$, we apply a stair matrix, which is a special tridiagonal matrix, with partial pivoting [5]. The stair matrix can be expressed as

$$\mathbf{S} = \mathrm{stair}(\mathbf{W}(N, N-1); \mathbf{W}(N, N); \mathbf{W}(N, N+1). \tag{9}$$

Thus, the initial solution of $\mathbf{x}^0$ is replaced by a stair matrix, $\mathbf{S}^0$. The convergence rate of these methods can be further improved by adopting a polynomial acceleration scheme called the Chebyshev acceleration [4]. We generalise the aforementioned iterative detection schemes as (see ref. [4])

$$\mathbf{x}^{i+1} = \mathbf{B}\mathbf{x}^i + \mathbf{c}, \tag{10}$$

where $\mathbf{B} = \mathbf{M}^{-1}\mathbf{N}$ and $\mathbf{c} = \mathbf{N}^{-1}\mathbf{y}^{MF}$. Based on Equation (10), these iterative methods can generate reasonable approximations of signal vector, $\mathbf{x}$. However, their convergence rate is slow and there are rooms to accelerate the convergence of such iterative detection methods. The Chebyshev acceleration is a good method to accelerate the convergence rate, $(\mathbf{x}^i)_{i=0}^{\infty}$. Based on the signal vector sequence $(\mathbf{x}^i)_{i=0}^{\infty}$, we construct a new linear combination $(\mathbf{y}_m)_{i=1}^{\infty}$, which features faster convergence towards the accurate solution $\mathbf{x}^i$ of (10), known as secondary iteration:

$$\mathbf{y}_m = \sum_{i=1}^{m} \gamma_{m,i} \mathbf{x}^i. \tag{11}$$

Note that the scalars $\gamma_{m,i}$ must satisfy $\sum_{i=0}^{m} \gamma_{m,i} = 1$. Since if $\mathbf{x^0} = \mathbf{x^1} = \cdots = \mathbf{x}$, we also have $\mathbf{y}_m = x$. Then, the error between the new combination of polynomial acceleration $\mathbf{y}_m$ and a signal vector $\mathbf{x}$ can be written as

$$\mathbf{y}_m - \mathbf{x} = \sum_{i=1}^{m} \gamma_{m,n}(\mathbf{x}_i - \mathbf{x}) = \sum_{i=1}^{m} \gamma_{m,i} \mathbf{B}^i (\mathbf{x}_0 - \mathbf{x}) = \mathbf{p}_m(\mathbf{B})(\mathbf{x}_0 - \mathbf{x}), \tag{12}$$

where $\mathbf{p}_m(\mathbf{B}) = \sum_{i=1}^{m} \gamma_{m,i} \mathbf{B}^i$ is polynomial of degree m with $p_m(1) = 1$, and $\mathbf{p}_m(\mathbf{z}) = \sum_{i=1}^{m} \gamma_{m,i} \mathbf{z}^i$.

The mth Chebyshev polynomial is defined by the three-term recurrence relation

$$\mathbf{T}_{m+1}(\mathbf{z}) = 2\mathbf{z}T_m(\mathbf{z}) - \mathbf{T}_{m-1}(\mathbf{z}), m \geq 1, \tag{13}$$

where initial parameters are $\mathbf{T_0}(\mathbf{z}) = \mathbf{1}$ and $\mathbf{T}_1(z) = z$. A polynomial parameter

$$\mathbf{p}_m(z) = \frac{\mathbf{T}_m(z/\rho)}{\mathbf{T}_m(1/\rho)}, \tag{14}$$

can provide minimum error, while ρ is spectral radius of the matrix $\mathbf{B}$ belonging to $[\lambda_{\min}, \lambda_{\max}]$. If the spectral radius of iterative methods achieves $\rho(\mathbf{B}) < 1$ or $\lim_{k\to\infty} \mathbf{B}_k = 0$, the iterative method will converge for each kth user. According to random matrix theory [6], the spectral radius of $\mathbf{B}$ is given by

$$\rho(\mathbf{B}) = \max |\lambda(\mathbf{B})|, \tag{15}$$

where $\lambda(\mathbf{B})$ donates the eigenvalue of matrix $\mathbf{B}$. The spectral radius of $\mathbf{W}$ [2] and the largest and smallest eigenvalues of $\mathbf{W}$ can be approximated as

$$\rho(\mathbf{W}) = |\lambda_{\max}(\mathbf{W})| = N\left(1 + \sqrt{\xi}\right)^2, \tag{16}$$

$$\rho(\mathbf{W}) = |\lambda_{\min}(\mathbf{W})| = N\left(1 - \sqrt{\xi}\right)^2, \tag{17}$$

where $\xi = N/K$ is the ratio between the number of antennas and the number of users. The optimal iteration polynomials $\mathbf{p}_m(\mathbf{z})$ can be obtained from the Chebyshev polynomials as

$$\mathbf{p}_m(z) = \frac{\mathbf{T}_m(1 + 2\frac{z-\epsilon}{\epsilon-\eta})}{\mathbf{T}_m(1 + 2\frac{r-\epsilon}{\epsilon-\eta})}, \epsilon = \frac{\lambda_{\max} + \lambda_{\min}}{\lambda_{\max} - \lambda_{\min}}, \eta = \frac{\lambda_{\max} + \lambda_{\min}}{2}. \tag{18}$$

The ϵ and η donate the optimal parameters of polynomials and they are defined by the smallest and the largest eigenvalues of the iteration matrix $\mathbf{B}$, respectively. Here, r is an scalar satisfying the coefficients of polynomials. Let us assume that $\mu_m = 1/\mathbf{T}_m(1/\rho)$, so $\mathbf{p}_m(\mathbf{B}) = \mu_m/T_m(\mathbf{B})/\rho)$, and $\gamma_{0,0} = 1$, $\gamma_{1,0} = 0$, $\gamma_{1,1} = 1$. Then, the properties of μ_m becomes

$$\frac{1}{\mu_{m+1}} = 2\frac{1}{\rho\mu_m} - \frac{1}{\mu_{m-1}}, \tag{19}$$

To employ the Chebyshev polynomial method in our purposed technique, we rewrite the equation of $\mathbf{y}_m$, which also satisfies a three-term-recurrence. After some mathematical manipulations, the secondary iteration can be defined as

$$\mathbf{y}_{m+1} = 2\frac{\mu_m}{\rho\mu_{m+1}}\mathbf{B}\mathbf{y}_m - \frac{\mu_{m-1}}{\mu_{m+1}} y_{m-1} + 2\frac{\mu_m}{\rho\mu_{m+1}}\mathbf{c}, \tag{20}$$

where $y_0 = x_0$ and $y_1 = x_1$. Therefore Chebyshev polynomials can be employed to design iterative methods.

Table 1. Computational complexity comparison

Method	Number of multiplications
GS [3]	$i4K^2 + K - 3$
JA [3]	$i(4K^2 - 2K) + K - 3$
RI [3]	$i(4K^2 + 2K) + K - 3$
SOR	$i(16K + 8K^2) + K - 3$
Chebyshev-RI	$i(4K^2 + 4K + 1) + K - 3$
Chebyshev-SOR	$i(16K + 8K^2) + K - 3$
Chebyshev-GS	$i\left(\frac{5}{2}K^2 + \frac{5}{2}K\right) + K - 3$
Chebyshev-JA	$i(K^2 + 3K) + K - 3$
MMSE	$K^3 + K^2$

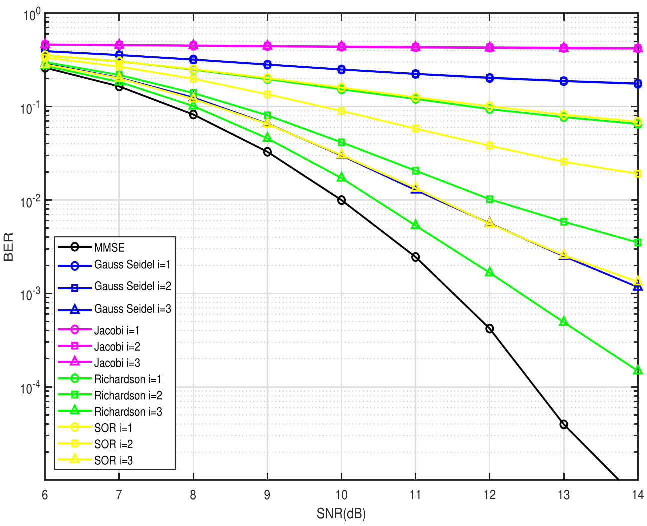

Fig. 1 *BER performance for 64-QAM in a* $N \times K = 128 \times 16$ *scheme and* i *iterations*

Computational complexity: In this section, we outline the computational complexity of the computation of $\mathbf{W^{-1}}$ in terms of the number of multiplications required. The computational complexities are roughly divided into two parts: 1) The iterative approach; 2) The iterative method combined with Chebyshev acceleration method. In general, the iterative methods have complexity of $\mathcal{O}(K^2)$, which is lower than the conventional MMSE procedure, which has $\mathcal{O}(K^3)$ complexity. Although the current iterative methods can also reduce the complexity to $\mathcal{O}(K^2)$, but they still require a large number of iterations, which compromises the overall complexity. As a result, the number of iterations necessary to attain a certain level of estimation accuracy reduces. The first step of stair matrix require a $K - 3$ multiplications. Table 1 shows the overall complexities of the proposed methods.

Performance results: To evaluate the proposed methods based on the Chebyshev acceleration, we provide the BER performance as a function of the signal-to-noise ratio (SNR) simulation, which are compared with conventional iterative methods. We consider a 64-QAM modulation scheme and a massive MIMO configuration of $N = 128$ and $K = 16$. We adopt Rayleigh fading channel model and a perfect synchronisation and channel estimation at the receiver.

Figure 1 shows the BER performance comparison between the iterative methods based on the GS, JA, RI, SOR, where i represents the number of iterations. It is clear that the iterative methods require more iterations to approach the linear MMSE performance, which leads to increased computation complexity. We can also observe that the performance of RI method since it has a factor of a overexertion ω [3].

Figure 2 shows the BER performance comparison between the iterative methods combined with the Chebyshev acceleration method. Figure 2 also demonstrates that we get a faster convergence with our acceleration methods compared to conventional iterative methods. For

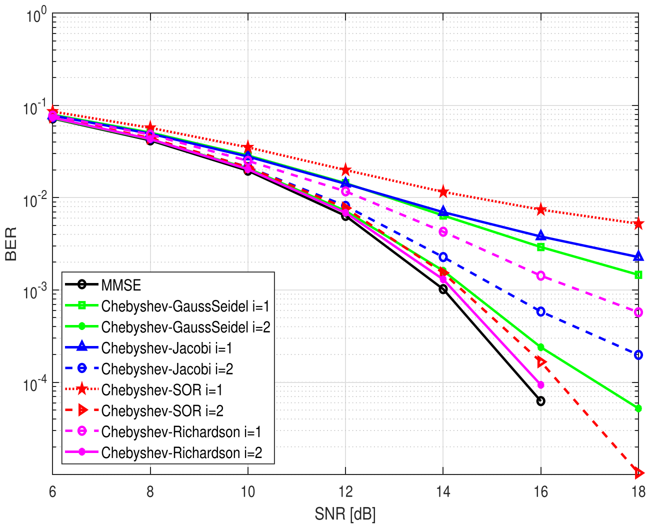

Fig. 2 *BER performance for 64-QAM in a* $N \times K = 128 \times 16$ *scheme and* i *iterations*

instance, the RI method has almost the MMSE performance with two iterations (the same applies to the other methods). It is clear from this result that Chebyshev acceleration method can improve the performance and achieve a fast convergence rate, requiring only two iterations to approach the linear MMSE performance.

Conclusion: In this letter we considered low complexity detection for massive MIMO schemes and proposed an acceleration technique based on the Chebyshev method to speed up the existing iterative methods. It was shown that our technique can be employed in different methods, outperforming state-of-the-art methods in terms of error rate performance and computational complexity.

Acknowledgement: This work was supported by Fundação para a Ciência e Tecnologia and Instituto de Telecomunicações under projects UIDB/50008/2020 and MASSIVE5G (SAICT-45-2017-02), and Copelabs under project UIDB/04111/2020.

Conflict of interest: The authors declare no conflict of interest.

Data availability statement: Author elects to not share data

Received: *12 January 2022* Accepted: *17 March 2022*
doi: 10.1049/ell2.12486

References

1 Björnson, E., Sanguinetti, L., Wymeersch, H., Hoydis, J., Marzetta, T.L.: Massive MIMO is a reality-what is next? five promising research directions for antenna arrays. *Digital Signal Process.* **94**, 3–20 (2019)
2 Berra, S., Dinis, R., Rabie, K., Shahabuddin, S.: Efficient iterative massive mimo detection using Chebyshev acceleration. *Phys. Commun.* **52**, 101651 (2022)
3 Albreem, M.A., Juntti, M., Shahabuddin, S.: Massive MIMO detection techniques: a survey. *IEEE Commun. Surv. Tutorials* **21**(4), 3109–3132 (2019)
4 Young, D.M.: *Iterative Solution of Large Linear Systems*, Elsevier, New York (2014)
5 Jiang, F., Li, C., Gong, Z., Su, R.: Stair matrix and its applications to massive mimo uplink data detection. *IEEE Trans. Commun.* **66**(6), 2437–2455 (2018)
6 Tulino, A.M., Verdú, S., et al.: Random matrix theory and wireless communications. *Found. Trends Commun. Inf. Theory* **1**(1), 1–182 (2004)