Low-Complexity MMSE Signal Detection Based on the AOR Iterative Algorithm for Uplink Massive MIMO Systems

Zhenyu Zhang[1], Yuanyuan Dong[1], Zhongshan Zhang[1], Xiyuan Wang[1], Xiaoming Dai[1(✉)], Linglong Dai[2], and Haijun Zhang[1]

[1] University of Science and Technology Beijing, Beijing 100083, China
daixiaoming@ustb.edu.cn
[2] Tsinghua University, Beijing 100084, China

Abstract. Massive multiple-input multiple-output (MIMO) systems can substantially improve the spectral efficiency and system capacity by equipping a large number of antennas at the base station and it is envisaged to be one of the critical technologies in the next generation of wireless communication systems. However, the computational complexity of the signal detection in massive MIMO systems presents a significant challenge for practical hardware implementations. This work proposed a novel minimum mean square error (MMSE) signal detection method based on the accelerated overrelaxation (AOR) iterative algorithm. The proposed AOR-based method can reduce the overall complexity of the classical MMSE signal detection by an order of magnitude from $O(K^3)$ to $O(K^2)$, where K is the number of users. Numerical results illustrate that the proposed AOR-based algorithm can outperform the performance of the recently proposed Neumann series approximation-based algorithm and approach the conventional MMSE signal detection involving exact matrix inversion with significantly reduced complexity.

Keywords: Accelerated overrelaxation (AOR) · Iterative algorithm
Minimum mean square error (MMSE) · Convergence · Complexity

1 Introduction

Multiple-input multiple-output (MIMO) is widely acknowledged as a key technology for the fourth generation (4G) wireless communication systems [1, 2] due to the high diversity gain and system channel capacity. However, the exponential increase of mobile data traffic enabled by the wide proliferation of smartphones and tablet computers poses great challenges for the current 4G systems [3]. Massive MIMO systems which scale up the antennas at the base station (BS) by orders of magnitude contrasted to the current systems (e.g., 4 or 8 antennas in 4G system) [4] can serve multi-users on the same frequency band simultaneously [5]. Many research show that the large-scale antennas at the BS can effectively average out non-coherent interference and system noise. Massive MIMO system with large-scale antennas achieve significant enhancements in terms of spectral efficiency, multiplexing gains, and robustness compared to

K. Long et al. (Eds.): 5GWN 2017, LNICST 211, pp. 385–394, 2018.
https://doi.org/10.1007/978-3-319-72823-0_36

the conventional MIMO systems [6], and it is envisaged to be the promising critical technology in the fifth generation (5G) wireless communication systems [7].

However, the promised gains on the multiplexing capability of massive MIMO systems come at the cost of the significant increase in signal detection complexity at both sides of the wireless communication links [1]. The optimal signal detection in MIMO systems is maximum likelihood (ML) signal detection in which the complexity increases exponentially with the number of transmitting antennas, which imposes an insurmountable cost for practical implementation in the massive MIMO systems [8]. The sphere decoding (SD) [9] can be utilized to simplify the hardware implementation of the ML signal detection; however, the complexity changes along with the channel condition and it is still quite high if the modulation order and/or the number of transmitting antennas is high. The K-best algorithm [10] with fixed complexity is also a popular method to simplify the ML detection, but the linear relationship between the critical path length and the number of antennas poses serious challenges for the large-scale MIMO systems. The linear signal detection such as the minimum mean square error (MMSE) signal detection [1] is utilized in massive MIMO systems to trade off the complexity and reliability; however, it incurs a complex matrix inversion operation whose complexity is immense especially for the large dimension of antennas. The Neumann series (NS) approximation-based algorithm has been proposed recently in [11, 12] to alleviate the complex matrix inversion operation in traditional MMSE signal detection, in which algorithm transforms the matrix inversion into a series of the matrix-vector multiplications and additions. However, the bit error ratio (BER) performance is unsatisfactory with small iteration numbers when the dimension of antennas is moderately large. Furthermore, the large iteration numbers incur even higher computational complexity compared to the classical MMSE signal detection. Hence, it is highly desirable to design low-complexity high-performance signal detection schemes that deliver acceptable BER performance and scale favorably to the high-dimensional signal detection problems.

In this paper, we propose a novel MMSE signal detection method based on the accelerated overrelaxation (AOR) iterative algorithm [13] for uplink massive MIMO systems. The proposed detection scheme reconstructs the transmitted signal without the complicated matrix inversion via an iterative operation. The symmetric positive definite property of the MMSE filtering matrix is amenable to the AOR-based approach. We provide a mathematical model of the AOR-based MMSE signal detection with a convergence and complexity analysis. Numerical results show that the proposed AOR-based algorithm can approach the BER performance of the classical MMSE method in a few iterations and outperform the NS-based approach with significantly reduced complexity.

The rest of the paper is structured as follows. In Sect. 2, the system model of the uplink massive MIMO system is described. In Sect. 3, the low-complexity MMSE signal detection based on the AOR iterative algorithm for uplink massive MIMO system is proposed. The convergence and complexity analysis are presented in the same section. In Sect. 4, the numerical results of the BER performance is specified. Finally, conclusions are drawn in Sect. 5.

Notation: Throughout the paper, upper-case boldface letters **S** and lower-case boldface letters **s** denote matrices and vectors, respectively; $\mathbf{S}^{-1}$, $\mathbf{S}^{T}$ and $\mathbf{S}^{H}$ refer to the matrix inversion, matrix transpose, and matrix conjugate transpose, respectively; $\mathrm{O}(\cdot)$ and $\mathrm{E}\{\cdot\}$ stand for the order of complexity and the expectation, respectively; $\Im\{\mathbf{s}\}$ and $\Re\{\mathbf{s}\}$ represent the imaginary part and real part of the complex number, respectively; s_i denotes the *ith* element of **s**; $s_{i,j}$ denotes the *ith* row and *jth* column entry of matrix **S**; Finally, $\mathbf{I}_K$ is the $K \times K$ identity matrix.

2 System Model

Consider a representative uplink multi-user massive MIMO system consisting of N antennas at the BS to serve K single-antenna users simultaneously [1], in which we normally have $N \gg K$. The transmitted encoded and interleaved bit streams are mapped to symbols by taking values from an energy-normalized quadrature amplitude modulation (QAM) constellation.

Let $\mathbf{x}_c \in \mathbb{C}^{K\times 1}$ denotes the complex-valued transmitted signal vector. The entries of the Rayleigh flat fading channel $\mathbf{H}_c \in \mathbb{C}^{N\times K}$ are independently and identically (i.i.d.) distributed and follow the complex Gaussian distribution $CN(0,1)$ with zero mean and unit variance. $\mathbf{n}_c \in \mathbb{C}^{N\times 1}$ represents the additive white Gaussian noise (AWGN) vector whose entries are i.i.d. and follow the distribution $CN(0,\sigma^2)$. $\mathbf{y}_c \in \mathbb{C}^{N\times 1}$ denotes the received signal vector at the BS. Then the complex-valued uplink system model can be expressed as

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

For ease of representation, the complex-valued model can be converted into a corresponding real-valued one as

$$\begin{bmatrix} \Re\{\mathbf{y}_c\} \\ \Im\{\mathbf{y}_c\} \end{bmatrix} = \begin{bmatrix} \Re\{\mathbf{H}_c\} & -\Im\{\mathbf{H}_c\} \\ \Im\{\mathbf{H}_c\} & \Re\{\mathbf{H}_c\} \end{bmatrix} \begin{bmatrix} \Re\{\mathbf{x}_c\} \\ \Im\{\mathbf{x}_c\} \end{bmatrix} + \begin{bmatrix} \Re\{\mathbf{n}_c\} \\ \Im\{\mathbf{n}_c\} \end{bmatrix}. \tag{2}$$

Then the real-valued uplink system model can be described as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}. \tag{3}$$

We assume the channel state information matrix **H** is known perfectly by receiver via the assigned training sequence [14, 15]. Then the transmitted signal can be reconstructed by the MMSE signal detector as

$$\hat{\mathbf{x}} = \left(\mathbf{H}^H\mathbf{H} + \sigma^2\mathbf{I}_{2K}\right)^{-1}\mathbf{H}^H\mathbf{y} = \mathbf{W}^{-1}\hat{\mathbf{y}}, \tag{4}$$

where $\hat{\mathbf{x}}$ is the reconstructed signal vector, $\mathbf{W} = \mathbf{H}^H\mathbf{H} + \sigma^2\mathbf{I}_{2K}$ represents the MMSE filtering matrix with a size of $2K \times 2K$, and $\hat{\mathbf{y}} = \mathbf{H}^H\mathbf{y}$ denotes the matched-filter output of $\mathbf{y}$. It should be noted that the $\mathbf{W}^{-1}$ operation requires cubic computational complexity $O(K^3)$ and it is extremely high for the massive MIMO systems.

3 Proposed Low-Complexity MMSE Signal Detection Method

3.1 MMSE Signal Detection Based on the AOR Iterative Algorithm

Unlike the conventional MIMO systems, things that were random before and now start to look deterministic in large-scale MIMO systems [1]. Owing to the fact that the column vectors of $\mathbf{H}$ are asymptotically orthogonal [1], it is obvious that the MMSE filtering matrix $\mathbf{W}$ is symmetric positive definite in uplink massive MIMO systems. This property inspires us to employ the AOR iterative algorithm to solve (4). The AOR iterative algorithm [13] is a classical iterative scheme for the numerical solution of the linear system $\hat{\mathbf{x}} = \mathbf{W}^{-1}\hat{\mathbf{y}}$. Splitting $\mathbf{W}$ as $\mathbf{W} = \mathbf{D} - \mathbf{U} - \mathbf{L}$, where $\mathbf{D}$ denotes the diagonal element of $\mathbf{W}$, $\mathbf{U}$ and $\mathbf{L}$ represent the negative of the strictly upper and lower triangular element of $\mathbf{W}$, respectively. Then the transmitted signal reconstructs by the AOR scheme can be denoted as

$$\hat{\mathbf{x}}^{(n)} = (\mathbf{D} - r\mathbf{L})^{-1}\left\{[(1-\omega)\mathbf{D} + (\omega - r)\mathbf{L} + \omega\mathbf{U}]\hat{\mathbf{x}}^{(n-1)} + \omega\hat{\mathbf{y}}\right\}, \quad (5)$$

where the coefficient ω and r represent the relaxation parameter and the acceleration parameter, respectively; and the superscript n denotes the iteration number. The initial iteration $\hat{\mathbf{x}}^{(0)}$ is set as a zero vector and the notation $\mathbf{L}_{r,\omega} = (\mathbf{D} - r\mathbf{L})^{-1}[(1-\omega)\mathbf{D} + (\omega - r)\mathbf{L} + \omega\mathbf{U}]$ denotes the iterative matrix.

3.2 Convergence Analysis

For uplink massive MIMO systems, the necessary and sufficient condition for the AOR-based MMSE signal detection algorithm being convergent is $\rho(\mathbf{L}_{r,\omega}) < 1$ [16], where the notation $\rho(\mathbf{L}_{r,\omega})$ denotes the spectral radius of the iterative matrix $\mathbf{L}_{r,\omega}$. The spectral radius is defined as $\rho(\mathbf{L}_{r,\omega}) = \max|\lambda_{r,\omega}|$, where $\lambda_{r,\omega}$ is the eigenvalue of $\mathbf{L}_{r,\omega}$. Then the convergence condition for the AOR iterative algorithm can be given as follows [17, 18]

$$\begin{aligned} &(1)\, 0 < \omega \leq r < 2 \\ &(2)\, 0 < r < \omega < 2 \left(\min\lambda_{0,1} \geq 0\right) \\ &(3)\, \max\left(0, \omega + \frac{2-\omega}{\min\lambda_{0,1}}\right) < r < \omega < 2 \left(\min\lambda_{0,1} < 0\right), \end{aligned} \quad (6)$$

where $\min\lambda_{0,1}$ denotes the minimum eigenvalue of the Jacobi iteration matrix as $\mathbf{L}_{0,1} = \mathbf{D}^{-1}(\mathbf{L} + \mathbf{U})$.

3.3 Computational Complexity Analysis

In this subsection, we evaluate the complexity of the proposed AOR-based MMSE signal detection algorithm. The hardware complexity is mainly dominated by the multipliers, so we define the complexity as the number of multiplications. Rewrite (5) as

$$\hat{\mathbf{x}}^{(n)} = (1-\omega)\hat{\mathbf{x}}^{(n-1)} + \omega\mathbf{D}^{-1}\left\{\hat{\mathbf{y}} + \frac{r}{\omega}\mathbf{L}\hat{\mathbf{x}}^{(n)} + \left[\left(1-\frac{r}{\omega}\right)\mathbf{L}+\mathbf{U}\right]\hat{\mathbf{x}}^{(n-1)}\right\}. \quad (7)$$

Considering the definition of the matrix **D**, **L** and **U**, Eq. (7) can be expressed as

$$\hat{x}_i^{(n)} = (1-\omega)\hat{x}_i^{(n-1)} + \frac{\omega}{w_{i,i}}\left(\hat{y}_i + \frac{r}{\omega}\sum_{j=1}^{i-1} w_{i,j}\hat{x}_j^{(n)} + \left(1-\frac{r}{\omega}\right)\sum_{j=1}^{i-1} w_{i,j}\hat{x}_j^{(n-1)} + \sum_{j=i+1}^{2K} w_{i,j}\hat{x}_j^{(n-1)}\right). \quad (8)$$

where the parameter $\frac{r}{\omega}$ can be calculated separately. The number of multiplications required for computation of $(1-\omega)\hat{x}_i^{(n-1)}$, $\frac{r}{\omega}\sum_{j=1}^{i-1} w_{i,j}\hat{x}_j^{(n)}$, $\sum_{j=i+1}^{2K} w_{i,j}\hat{x}_j^{(n-1)}$ and $\left(1-\frac{r}{\omega}\right)\sum_{j=1}^{i-1} w_{i,j}\hat{x}_j^{(n-1)}$ are 1, i, $2K-i$ and i, respectively, so the required number of multiplications for $\hat{x}_i^{(n)}$ is $2K+i+3$. Thus, the overall complexity for one iteration is $\sum_{i=1}^{2K} 2K+i+3 = 6K^2+7K$.

Table 1 compares the complexity of the proposed AOR-based algorithm with that of NS-based algorithm. The classical MMSE signal detection with exact matrix inversion operation has cubic computational complexity $O(K^3)$. Table 1 illustrates the complexity of the NS-based algorithm scales with $O(K^2)$ only for $n=2$. However, the proposed AOR-based algorithm can decrease the complexity by an order of magnitude from $O(K^3)$ to $O(K^2)$.

Table 1. Computational complexity

Iteration number	NS-based algorithm	AOR-based algorithm
$n=2$	$12K^2-2K$	$12K^2+14K$
$n=3$	$8K^3+4K^2$	$18K^2+21K$
$n=4$	$16K^3-4K^2+2K$	$24K^2+28K$

4 Simulation Results

To verify the validity of the proposed AOR-based MMSE signal detection, we evaluated the BER performance against the signal-to-noise ratio (SNR) and compared it with the recently proposed NS-based algorithm [11, 12]. The BER performance of the classical MMSE signal detection with exact matrix inversion was also given as the benchmark for comparison. A representative massive MIMO system scenario with $N \times K = 128 \times 16$ and the 64-QAM modulation scheme was adopted.

Figure 1 shows the BER performance of the proposed AOR-based MMSE signal detection versus the relaxation parameter ω and the acceleration parameter r in a three-dimensional way. The SNR is 15 dB and the iteration number is $n = 3$. The optimal parameters can achieve a faster convergence rate and a preferable performance. As shown in Fig. 1, the BER performance difference is negligible when the parameters are close to $\omega = 1.10$ and $r = 1.05$, indicating the robustness of the proposed method. So we chose the optimal relaxation and acceleration parameter as $\omega_{opt} = 1.10$ and $r_{opt} = 1.05$ in the following simulations.

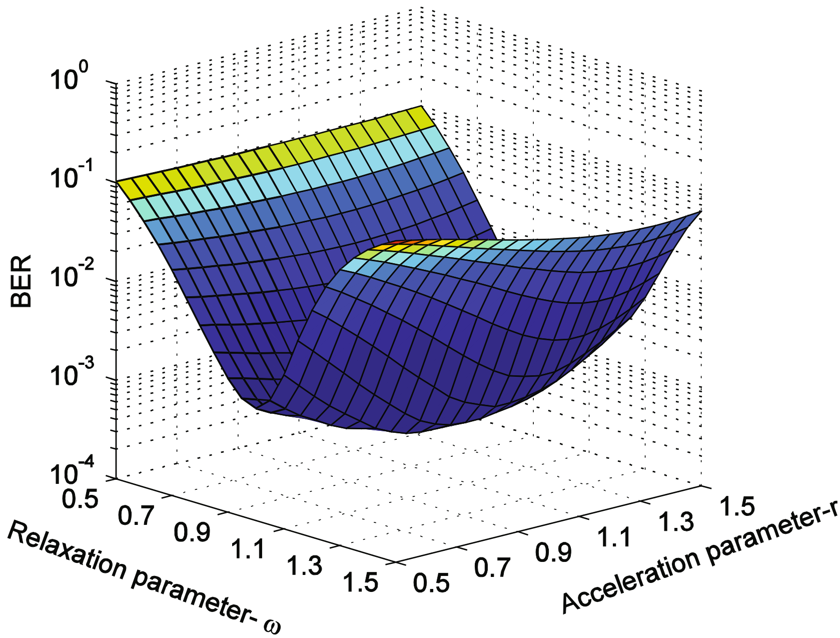

Fig. 1. BER performance of the AOR-based algorithm versus ω and r. The SNR is 15 dB and n is 3.

Next we evaluate the influence of the relaxation and acceleration parameters on the convergence rate. Figure 2 illustrates the BER performance of the proposed AOR-based MMSE signal detection versus the relaxation parameter ω with the acceleration parameter r being fixed. Figure 3 illustrates the BER performance of the proposed AOR-based MMSE signal detection versus the acceleration parameter r with the relaxation parameter ω being fixed. The BER curve behaves like a quadratic function curve. The BER performance initially improves with the value of ω up to approximately 1.10 and then starts to deteriorate for higher values. It is worth noting that the BER performance are extremely sensitive to the variation of the relaxation parameter with the fixed acceleration parameter. So we conclude that the relaxation parameter ω plays a dominant role in the convergence of the algorithm while the

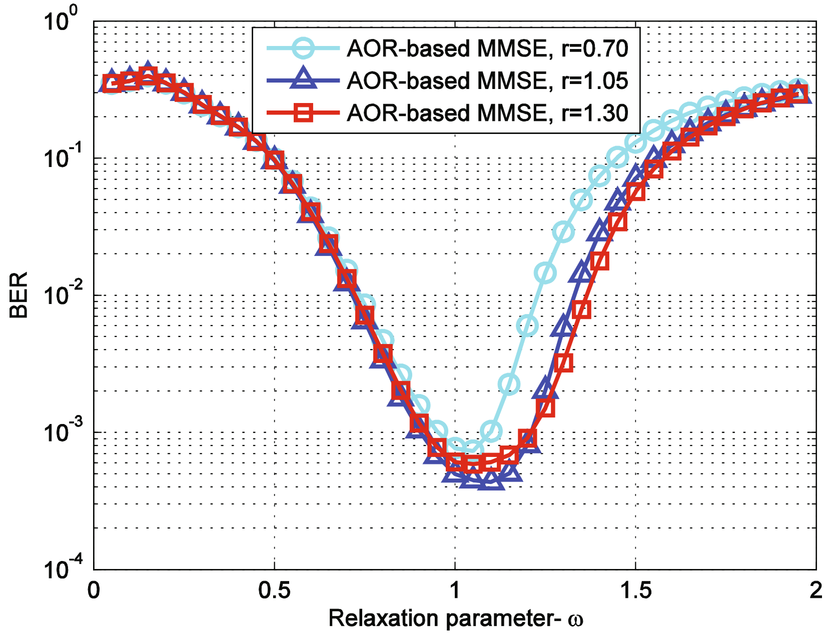

Fig. 2. BER performance of the AOR-based algorithm versus ω. The SNR is 15 dB and n is 3.

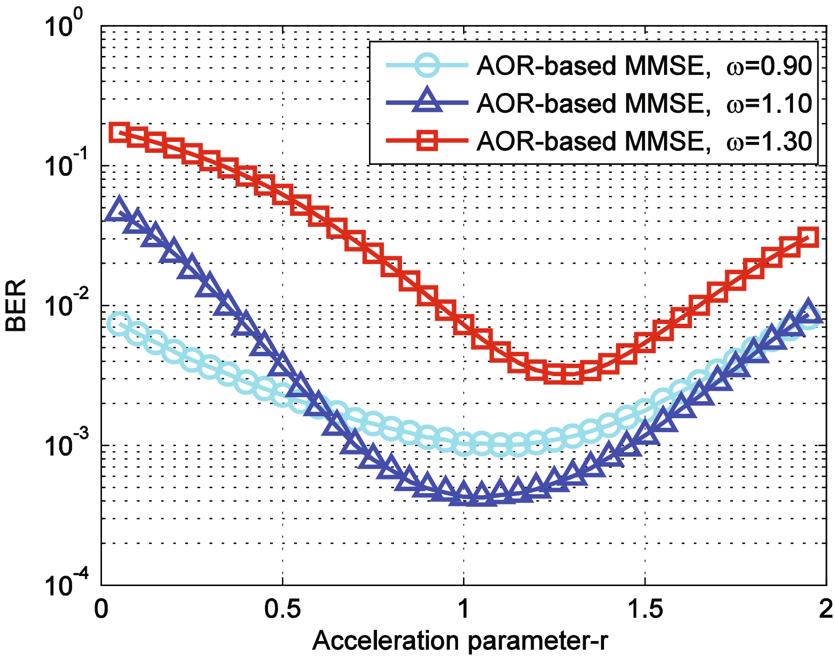

Fig. 3. BER performance of the AOR-based algorithm versus r. The SNR is 15 dB and n is 3.

acceleration parameter r can expedite the convergence rate to some extent in the proposed AOR-based MMSE signal detection. Based on this finding, a more robust system can be realized by selecting a proper acceleration parameter despite a small variation of the relaxation parameter. Numerical results show that the optimal relaxation parameter ω_{opt} can be set around 1.10 while the corresponding optimal acceleration parameter r_{opt} is close to ω_{opt}.

Figure 4 provides the BER performance comparison for different detection algorithms. It is apparent that the BER performance improves significantly with the increasing iteration number of the proposed algorithm. The BER performance of the AOR-based algorithm outperforms the NS-based algorithm with the same iteration number due to the faster convergence rate of the AOR iterative algorithm. For example, the proposed AOR-based method with n = 2 can even achieve similar performance levels as the NS-based method with n = 4. The BER performance of the AOR-based algorithm approaches the traditional MMSE method within 0.1 dB when the iteration number is n = 3 for the $N \times K = 128 \times 16$ massive MIMO systems.

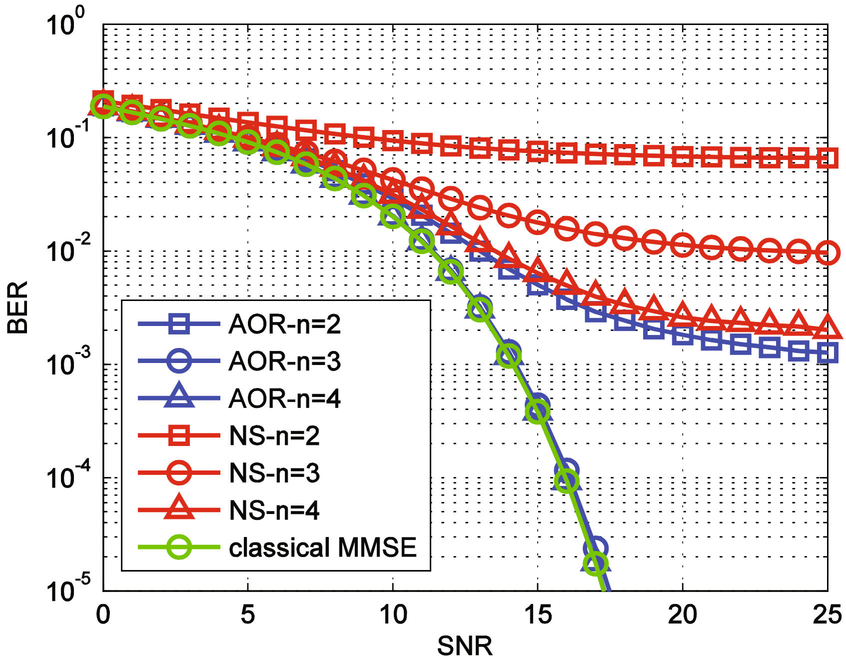

Fig. 4. BER performance comparison versus SNR.

5 Conclusions

In this paper, we proposed a low-complexity MMSE signal detection based on the AOR iterative algorithm for uplink massive MIMO systems. The proposed AOR-based algorithm reconstructs the transmitting signal via a relaxation and acceleration operation to refine the solution estimate in an iterative manner. Numerical results illustrate that the proposed AOR-based algorithm can obtain better BER performance than the recently proposed NS-based algorithm and approach the performance of the classical MMSE signal detection with a significantly reduced complexity. Owing to the superiority of the low-complexity AOR iteration algorithm, it can be utilized to solve other problems involving complicated matrix inversion operations such as the pre-coding technique in massive MIMO systems.

Acknowledgments. This work was supported by Huawei Innovation Research Program, the key project of the National Natural Science Foundation of China (No. 61431001), the 5G research program of China Mobile Research Institute (Grant No. [2015] 0615), the open research fund of National Mobile Communications Research Laboratory Southeast University (No. 2017D02), Key Laboratory of Cognitive Radio and Information Processing, Ministry of Education (Guilin University of Electronic Technology), the Foundation of Beijing Engineering and Technology Center for Convergence Networks and Ubiquitous Services, and Keysight.

References

1. Rusek, F., et al.: Scaling up MIMO: opportunities and challenges with very large arrays. IEEE Sig. Process. Mag. **30**(1), 40–60 (2013)
2. Zheng, L., Tse, D.N.C.: Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels. IEEE Trans. Inf. Theory **49**(5), 1073–1096 (2003)
3. Dai, X., et al.: Successive interference cancelation amenable multiple access (SAMA) for future wireless communications. In: 2014 IEEE International Conference on Communication Systems, Macau, pp. 222–226 (2014)
4. Marzetta, T.L.: Noncooperative cellular wireless with unlimited numbers of base station antennas. IEEE Trans. Wirel. Commun. **9**(11), 3590–3600 (2010)
5. Lu, L., Li, G.Y., Swindlehurst, A.L., Ashikhmin, A., Zhang, R.: An overview of massive MIMO: benefits and challenges. IEEE J. Sel. Top. Sig. Process. **8**(5), 742–758 (2014)
6. Xie, H., Gao, F., Zhang, S., Jin, S.: A unified transmission strategy for TDD/FDD massive MIMO systems with spatial basis expansion model. IEEE Trans. Veh. Technol. **66**(4), 3170–3184 (2017)
7. Boccardi, F., Heath, R.W., Lozano, A., Marzetta, T.L., Popovski, P.: Five disruptive technology directions for 5G. IEEE Commun. Mag. **52**(2), 74–80 (2014)
8. Artes, H., Seethaler, D., Hlawatsch, F.: Efficient detection algorithms for MIMO channels: a geometrical approach to approximate ML detection. IEEE Trans. Signal Process. **51**(11), 2808–2820 (2003)
9. Jalden, J., Ottersten, B.: On the complexity of sphere decoding in digital communications. IEEE Trans. Sig. Process. **53**(4), 1474–1484 (2005)
10. Bello, I.A., Halak, B., El-Hajjar, M., Zwolinski, M.: VLSI implementation of a scalable K-best MIMO detector. In: 2015 15th International Symposium on Communications and Information Technologies (ISCIT), Nara, pp. 281–286 (2015)
11. Wu, M., Yin, B., Wang, G., Dick, C., Cavallaro, J.R., Studer, C.: Large-scale MIMO detection for 3GPP LTE: algorithms and FPGA implementations. IEEE J. Sel. Top. Sig. Process. **8**(5), 916–929 (2014)
12. Zhu, D., Li, B., Liang, P.: On the matrix inversion approximation based on neumann series in massive MIMO systems. In: 2015 IEEE International Conference on Communications (ICC), London, pp. 1763–1769 (2015)

13. Hadjidimos, A.: Accelerated overrelaxation method. Math. Comput. **32**(141), 149–157 (1978)
14. Bölcskei, H.: Space-Time Wireless Systems: From Array Processing to MIMO Communications. Cambridge University Press, New York (2006)
15. Xie, H., Gao, F., Jin, S.: An overview of low-rank channel estimation for massive MIMO systems. IEEE Access **4**, 7313–7321 (2016)
16. Björck, Å.: Numerical Methods in Matrix Computations. TAM, vol. 59. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-05089-8
17. Zeng, W.: On convergence of AOR method. Journal of Huqiao University (1985)
18. Chen, P.X.: Convergence of AOR Method. Mathematica Numerica Sinica (1983)