

RESEARCH

Open Access



A fast and low-complexity matrix inversion scheme based on CSM method for massive MIMO systems

Yue Xu^{1*} , Weixia Zou¹ and Liutong Du²

Abstract

Massive multiple-input-multiple-output (MIMO), also known as very-large MIMO systems, is an attracting technique in 5G and can provide higher rates and power efficiency than 4G. Linear-precoding schemes are able to achieve the near optimal performance, and thus are more attractive than non-linear precoding schemes. However, conventional linear precoding schemes in massive MIMO systems, such as regularized zero-forcing (RZF) precoding, have near-optimal performance but suffer from high computational complexity due to the required matrix inversion of large size. To solve this problem, we utilize the Cholesky-decomposition and Sherman-Morrison lemma and propose CSM (Cholesky and Sherman-Morrison strategy)-based precoding scheme to the matrix inversion by exploiting the asymptotically orthogonal channel property in massive MIMO systems. Results are evaluated numerically in terms of bit-error-rate (BER) and average sum rate. Comparing with the Neumann series approximation of inverting matrix, it is concluded that, with fewer operations, the performance of CSM-based precoding is better than conventional methods in massive MIMO configurations.

Keywords: Massive MIMO, Cholesky-decomposition, Sherman-Morrison lemma, Neumann series, RZF, CSM-based precoding

1 Introduction

Massive multiple-input multiple-output (MIMO), i.e., MIMO with large numbers of transmit and/or receive antennas (massive MIMO technology), is widely accepted as one of the key enabling technologies for next generation (e.g., 5G) wireless communication systems [1, 2]. However, with the increasing of the number of dimensions, using conventional MIMO algorithms may not be suitable any more in terms of computational efficiency and new methods must emerge.

Realizing massive MIMO systems has to solve some challenges in practice, one of which is the low-complexity and near-optimal precoding scheme. Conventional precoding methods can be divided into nonlinear-precoding and linear precoding. The optimal precoding is the nonlinear which is called dirty paper precoding (DPC) [3],

which can effectively eliminate the interference between different users and achieve optimal performance. However, the most serious drawback of the nonlinear-precoding schemes is high complexity which is unfriendly to hardware. The other nonlinear-precoding schemes, such as lattice-aided precoding [4], can achieve the close-optimal capacity with reduced complexity, but they are still unaffordable when the dimension of the MIMO system is large or the modulation order is high. Fortunately, due to the characters of massive MIMO systems, such as the columns of channel matrix are asymptotically orthogonal [5] and the channel hardening [6], the linear precoding schemes, such as RZF precoding and MMSE precoding, can also achieve the near-optimal capacity, which makes a better trade-off between the complexity and the performance. However, these schemes require computing unfavorable matrix inversion of large size. To solve this problem, based on Neumann series approximation algorithm, [7] proposed the Neumann-based precoding which can reduce the complexity by converting the matrix inversion into a series of matrix-vector multiplications. Then,

*Correspondence: yuexu24@163.com

¹KeyLab of Universal Wireless Communications, MOE, Beijing University of Posts and Telecommunications, 100876 Beijing, People's Republic of China
Full list of author information is available at the end of the article

[8–10] proposed SOR-based, LSQR-based, and TPE - based schemes, respectively, which are all based Neumann serise. But, these algorithms' the reduction in complexity is not obvious and do not consider the property of positive defined Hermitian matrix.

In this paper, we propose CSM-based precoding to reduce the complexity of matrix inversion for classical RZF or MMSE precoding. This is motivated by the fact that the matrix which needs to be inverted in RZF or MMSE precoding is a positive definite Hermitian matrix and tends to be diagonal dominant in massive MIMO systems [1], which provides the potential to utilize the Cholesky-Decomposition [11] and Sherman-Morrison lemma [12]. We also conclude that CSM-based precoding can enjoy a better performance and lower computation complexity than the Neumann-based precoding and SOR-based precoding. The impact of the algorithms in the average sum rate and BER are evaluated and compared via numerical simulations.

Notation: lower-case and upper-case boldface letters denote vectors and matrices, respectively; $(\cdot)^T, (\cdot)^H, (\cdot)^{-1}, \det(\cdot)$ and $tr(\cdot)$ denote the transpose, conjugate transpose, matrix inversion, determinant and trace, respectively; \mathbb{C} denotes the set of complex numbers, \mathbf{I}_N is the $N \times N$ identity matrix.

2 System model

We consider the typical downlink transmission of a massive multi-user MIMO system, where each base station (BS) in cells is equipped with N antennas and communicates with K single-antenna users ($K \ll N$) [1]. Meanwhile, we also assume that channel state information (CSI) is known at the BS perfectly, which is a common assumption in massive MIMO systems [13, 14] and can be acquired by the training pilot [15]. More importantly, we assume that time division duplex (TDD) protocols are used so that the channel vectors are equal for both directions. In time division duplex (TDD) massive MIMO systems, the BS estimates the uplink channel by using the pilots that users send in the uplink, and then the downlink CSI can be easily acquired by using existing channel reciprocity in TDD systems.

In the downlink, the received signal at i th user in j th cell is:

$$y_{j,i} = \mathbf{h}_{j,i} \mathbf{x}_{j,i} + n_{j,i} \quad (1)$$

where, $\mathbf{x}_{j,i}$ is the transmit signal after precoding in j th cell and $\mathbf{h}_{j,i}$ represents the random channel vector from j th cell's BS to i th user. And $n_{j,i}$ is additive white Gaussian noise (AWGN) and follows the distribution $CN(0, \sigma_n^2)$. The channel vector from j th cell'BS to i th user can be specified as below form:

$$\mathbf{h}_{j,i} = \kappa_{j,i} \mathbf{R}_{j,i}^{1/2} \mathbf{z}_{j,i} \quad (2)$$

where, $\mathbf{z}_{j,i}$ is small scale fading vector that is independent and identically distributed (i.i.d) zero mean, circularly-symmetric complex Gaussian random variables $CN(0, 1)$, $\kappa_{j,i}$ is large scale fading coefficient which is accounted by path loss and shadow fading of each i th user that changes slowly and can remain constant over a coherence time interval and known prior and $\mathbf{R}_{j,i}$ is the channel covariance matrix.

Thus, the received vector in j th cell, denoted by $\mathbf{y}_j = [y_{j,1}, y_{j,2}, y_{j,3} \cdots y_{j,K}]^T \in \mathbb{C}^{K \times 1}$, at the receiver is given as:

$$\mathbf{y}_j = \mathbf{H}_j \cdot \mathbf{x}_j + \mathbf{n}_j \quad (3)$$

where, $\mathbf{H}_j = [\mathbf{h}_{j,1}, \mathbf{h}_{j,2}, \mathbf{h}_{j,3} \cdots \mathbf{h}_{j,N}] \in \mathbb{C}^{K \times N}$ is the downlink channel matrix, which contains small scale fading factor, large scale fading coefficient and channel covariance matrix. $\mathbf{x}_j \in \mathbb{C}^{N \times 1}$ is signal vector after precoding which satisfies the power limitation $E[\|\mathbf{x}_j\|_2^2] \leq K$. And $\mathbf{n}_j \in \mathbb{C}^{K \times 1}$ is additive white Gaussian noise (AWGN) and follows the distribution $CN(0, \sigma_n^2)$.

For massive MIMO systems, linear-precoding is usually considered, so we have

$$\mathbf{x}_j = \mathbf{T}_j \cdot \mathbf{s}_j, \quad (4)$$

where $\mathbf{T}_j = [\mathbf{t}_{j,1}, \mathbf{t}_{j,2}, \mathbf{t}_{j,3} \cdots \mathbf{t}_{j,K}] \in \mathbb{C}^{N \times K}$ is the precoding matrix, and $\mathbf{s}_j = [s_{j,1}, s_{j,2}, s_{j,3} \cdots s_{j,K}]^T \in \mathbb{C}^{K \times 1}$ is the transmitted signal vector for K users. We also denote the total transmit power constraints at BS in j th cell as

$$tr(\mathbf{H}_j \cdot \mathbf{H}_j^H) = P_j, \quad (5)$$

where P_j is the total transmit power in j th cell.

In the next section, we will analyze the existing classic massive MIMO precoding method (i.e., RZF precoding and MMSE precoding), as well as the proposed CSM-based scheme.

3 Low-complexity linear-precoding scheme in massive MIMO

In this section, we first give a basic background of the conventional RZF precoding and MMSE precoding. Then, we give the relative knowledge of Cholesky-decomposition [11] and Sherman-Morrison lemma [12]. After that, we prove that CSM-based precoding has a lower computational complexity and better performance than Neumann-based precoding and SOR-based precoding. Finally, we give the pseudo code of CSM-based precoding algorithm.

3.1 Conventional RZF and MMSE precoding

The conventional RZF precoding and MMSE precoding matrix can be expressed, respectively, as [6]:

$$\mathbf{T}_{RZF} = \rho_{RZF} \cdot \mathbf{H}^H \cdot (\mathbf{H} \cdot \mathbf{H}^H + \phi \cdot \mathbf{I}_K)^{-1} \quad (6)$$

$$\mathbf{T}_{MMSE} = \rho_{MMSE} \cdot \mathbf{H}^H \cdot (\mathbf{H} \cdot \mathbf{H}^H + (\sigma_n^2 \cdot n_t) \cdot \mathbf{I}_K)^{-1} \quad (7)$$

Thus, the classical RZF precoding and MMSE precoding matrix in j th cell is:

$$\mathbf{T}_{RZF,j} = \rho_{RZF,j} \cdot \mathbf{H}_j^H \cdot (\mathbf{H}_j \cdot \mathbf{H}_j^H + \phi \cdot \mathbf{I}_K)^{-1} \quad (8)$$

$$= \rho_{RZF,j} \cdot \mathbf{H}_j^H \cdot \mathbf{W}_{RZF,j}^{-1} \quad (9)$$

$$\mathbf{T}_{MMSE,j} = \rho_{MMSE,j} \cdot \mathbf{H}_j^H \cdot (\mathbf{H}_j \cdot \mathbf{H}_j^H + (\sigma_n^2 \cdot n_t) \cdot \mathbf{I}_K)^{-1} \quad (10)$$

$$= \rho_{MMSE,j} \cdot \mathbf{H}_j^H \cdot \mathbf{W}_{MMSE,j}^{-1} \quad (11)$$

where $\mathbf{W}_{RZF,j} = (\mathbf{H}_j \cdot \mathbf{H}_j^H + \phi \cdot \mathbf{I}_K)$ and $\mathbf{W}_{MMSE,j} = (\mathbf{H}_j \cdot \mathbf{H}_j^H + (\sigma_n^2 \cdot n_t) \cdot \mathbf{I}_K)$. The ϕ_j is regularized parameter which can be adaptively selected according to the different CSI, σ_n^2 is noise power and n_t is number of transmitted antenna [6, 16]. And, the $\rho_{RZF,j}$ or $\rho_{MMSE,j}$ is the power normalization factor which makes RZF precoding or MMSE precoding satisfy the power limitation. Therefore, $\rho_{RZF,j}$ and $\rho_{MMSE,j}$ can be computed by:

$$\rho_{RZF,j} = \sqrt{\frac{K}{\text{tr}(\mathbf{W}_{RZF,j} \cdot \mathbf{W}_{RZF,j}^H)}} \quad (12)$$

$$\rho_{MMSE,j} = \sqrt{\frac{K}{\text{tr}(\mathbf{W}_{MMSE,j} \cdot \mathbf{W}_{MMSE,j}^H)}} \quad (13)$$

So, the signal vector \mathbf{x} can be rewritten as:

$$\mathbf{x} = \mathbf{T}_{RZF} \cdot \mathbf{s} = \rho \cdot \mathbf{H}^H \cdot \mathbf{W}_{RZF}^{-1} \cdot \mathbf{s} \quad (14)$$

or

$$\mathbf{x} = \mathbf{T}_{MMSE} \cdot \mathbf{s} = \rho \cdot \mathbf{H}^H \cdot \mathbf{W}_{MMSE}^{-1} \cdot \mathbf{s} \quad (15)$$

We can observe from the (9) or (11) that a matrix \mathbf{W}_{RZF} or \mathbf{W}_{MMSE} inversion of $K \times K$ size is required which means that the high computational complexity is unexpected and unacceptable. If we compute the inversion directly, the resource of hardware would be wasted greatly. Therefore, combining the property of matrix \mathbf{W}_{RZF} or \mathbf{W}_{MMSE} and some mathematic knowledge such as Cholesky-Decomposition [11] and Sherman-Morrison lemma [12], we design a low-complexity precoding scheme to solve this problem.

3.2 CSM-based precoding

To reduce the computational complexity for precoding, we propose to use CSM-based scheme to avoid the complicated matrix inversion of large size in RZF or MMSE precoding. First, it is necessary to verify that the matrix \mathbf{W}_{RZF} and \mathbf{W}_{MMSE} are positive definite Hermitian

matrix. Here, assuming an arbitrary nonzero vector $\mathbf{t} \in \mathbb{C}^{N \times 1}$, then we can certificate that

$$\mathbf{t} \cdot \mathbf{W}_{RZF} \cdot \mathbf{t}^H = \mathbf{t} \cdot (\mathbf{H} \cdot \mathbf{H}^H + \phi \cdot \mathbf{I}_K) \cdot \mathbf{t}^H \quad (16)$$

$$= \mathbf{t} \cdot \mathbf{H} \cdot (\mathbf{t} \cdot \mathbf{H})^H + \mathbf{t} \cdot (\phi \cdot \mathbf{I}_K) \cdot \mathbf{t}^H > 0 \quad (17)$$

and

$$\mathbf{t} \cdot \mathbf{W}_{MMSE} \cdot \mathbf{t}^H = \mathbf{t} \cdot (\mathbf{H} \cdot \mathbf{H}^H + (\sigma_n^2 \cdot n_t) \cdot \mathbf{I}_K) \cdot \mathbf{t}^H \quad (18)$$

$$= \mathbf{t} \cdot \mathbf{H} \cdot (\mathbf{t} \cdot \mathbf{H})^H + \mathbf{t} \cdot ((\sigma_n^2 \cdot n_t) \cdot \mathbf{I}_K) \cdot \mathbf{t}^H > 0. \quad (19)$$

Meanwhile,

$$\mathbf{W}_{RZF}^H = (\mathbf{H} \cdot \mathbf{H}^H + \phi \cdot \mathbf{I}_K)^H = \mathbf{W}_{RZF} \quad (20)$$

and

$$\mathbf{W}_{MMSE}^H = (\mathbf{H} \cdot \mathbf{H}^H + (\sigma_n^2 \cdot n_t) \cdot \mathbf{I}_K)^H = \mathbf{W}_{MMSE}. \quad (21)$$

So, the important conclusion that \mathbf{W}_{RZF} and \mathbf{W}_{MMSE} are positive definite Hermitian matrix is clear. Due to RZF precoding is as same as MMSE precoding in decomposition and inversion process. Therefore, we just utilize the RZF precoding as an example to demonstrate the CSM-based scheme. Now, we can utilize the Cholesky-decomposition [11] to decompose the matrix \mathbf{W}_{RZF} as

$$\mathbf{W}_{RZF} = \mathbf{L} \cdot \mathbf{L}^H, \quad (22)$$

where matrix \mathbf{L} is the lower triangular matrix. Then, we have knowledge that

$$\mathbf{W}_{RZF}^{-1} = (\mathbf{L}^H)^{-1} \cdot \mathbf{L}^{-1}. \quad (23)$$

Thus, computing the inversion of matrix \mathbf{W}_{RZF} can be transformed into computing the inversion of matrix \mathbf{L} . Continue to utilize the Sherman-Morrison lemma [12] to iterate the process of computing the \mathbf{L} 's matrix inversion. Here, we should introduce the Sherman-Morrison lemma [12].

Suppose \mathbf{A} is an invertible square matrix and \mathbf{x}, \mathbf{y} are column vectors. Suppose furthermore that $1 + \mathbf{y}^H \cdot \mathbf{A}^{-1} \cdot \mathbf{x} \neq 0$ and $(\mathbf{A} + \mathbf{x} \cdot \mathbf{y}^H)$ is invertible. Then, the Sherman-Morrison formula states that

$$(\mathbf{A} + \mathbf{x} \cdot \mathbf{y}^H)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1} \cdot \mathbf{x} \cdot \mathbf{y}^H \cdot \mathbf{A}^{-1}}{1 + \mathbf{y}^H \cdot \mathbf{A}^{-1} \cdot \mathbf{x}}. \quad (24)$$

This lemma inspires us that we can utilize iteratively method to calculate several times' simple matrix inversion instead of computing directly complex matrix inversion and eventually simplify the high computational complexity to lower computational complexity.

In order to keep our computational complexity lower, we have to continue to decompose matrix \mathbf{L} :

$$\mathbf{L} = \mathbf{D} + \mathbf{L}', \quad (25)$$

where $\mathbf{D} = \text{diag}(\mathbf{l}_{1,1}, \mathbf{l}_{2,2}, \dots, \mathbf{l}_{K,K})$ is the diagonal matrix which is diagonal component of \mathbf{L} and $\mathbf{L}' = (\mathbf{l}'_1, \mathbf{l}'_2, \dots, \mathbf{l}'_{K-1}, 0)$ is a matrix based on \mathbf{L} which the diagonal elements of \mathbf{L} are replaced by zero.

Based on above the idea, we can calculate the inversion of matrix \mathbf{L} . First, the matrix should be rewritten as the following form:

$$\mathbf{L} = \mathbf{D} + \mathbf{L}' \tag{26}$$

$$= \mathbf{D} + \mathbf{l}'_1 \cdot \mathbf{e}_1 + \mathbf{l}'_2 \cdot \mathbf{e}_2 + \dots + \mathbf{l}'_{K-1} \cdot \mathbf{e}_{K-1} + \mathbf{0} \cdot \mathbf{e}_K, \tag{27}$$

where \mathbf{e}_i is the i th row of identity matrix \mathbf{I}_K , \mathbf{l}'_i is the i th column of matrix \mathbf{L}' . Thus, we can compute the inversion by:

$$\mathbf{L}^{-1} = (\mathbf{F}_{K-2} + \mathbf{l}'_{K-1} \cdot \mathbf{e}_{K-1})^{-1} \tag{28}$$

$$= (\mathbf{F}_{K-2})^{-1} - \frac{(\mathbf{F}_{K-2})^{-1} \cdot \mathbf{l}'_{K-1} \cdot \mathbf{e}_{K-1} \cdot (\mathbf{F}_{K-2})^{-1}}{1 + \mathbf{e}_{K-1} \cdot (\mathbf{F}_{K-2})^{-1} \cdot \mathbf{l}'_{K-1}}, \tag{29}$$

where $\mathbf{F}_{K-2} = \mathbf{D} + \sum_{i=1}^{K-2} \mathbf{l}'_i \cdot \mathbf{e}_i$. Therefore, we should continue to the inversion of \mathbf{F}_{K-2} :

$$\mathbf{F}_{K-2}^{-1} = (\mathbf{F}_{K-3} + \mathbf{l}'_{K-2} \cdot \mathbf{e}_{K-2})^{-1} \tag{30}$$

$$= (\mathbf{F}_{K-3})^{-1} - \frac{(\mathbf{F}_{K-3})^{-1} \cdot \mathbf{l}'_{K-2} \cdot \mathbf{e}_{K-2} \cdot (\mathbf{F}_{K-3})^{-1}}{1 + \mathbf{e}_{K-2} \cdot (\mathbf{F}_{K-3})^{-1} \cdot \mathbf{l}'_{K-2}}, \tag{31}$$

where $\mathbf{F}_{K-3} = \mathbf{D} + \sum_{i=1}^{K-3} \mathbf{l}'_i \cdot \mathbf{e}_i$. And so on, we can compute that

$$\mathbf{F}_1^{-1} = (\mathbf{F}_0 + \mathbf{l}'_1 \cdot \mathbf{e}_1)^{-1} \tag{32}$$

$$= (\mathbf{F}_0)^{-1} - \frac{(\mathbf{F}_0)^{-1} \cdot \mathbf{l}'_1 \cdot \mathbf{e}_1 \cdot (\mathbf{F}_0)^{-1}}{1 + \mathbf{e}_1 \cdot (\mathbf{F}_0)^{-1} \cdot \mathbf{l}'_1}, \tag{33}$$

where $\mathbf{F}_0 = \mathbf{D}$. Due to the \mathbf{D} matrix is diagonal matrix, its inversion process is simple and efficient. Therefore, the whole process of inversion can be calculate by basic mathematical process and simple process of iteration.

Finally, we can compute the inversion of \mathbf{L} by iterating $K - 1$ times.

3.3 Complexity

We evaluate the computational complexity in terms of required number of complex multiplications which is more dominant and complex than other operations for the total computational complexity.

According to [17], we have knowledge that the Cholesky-decomposition can be decomposed quickly and accurately by hardware such as FPGA. So, the computational complexity of the decomposition can be ignored. Observing the numerator $(\mathbf{F}_{i-1})^{-1} \cdot \mathbf{l}'_i \cdot \mathbf{e}_i \cdot (\mathbf{F}_{i-1})^{-1}$ in (29) or in (31), the column vector of \mathbf{l}'_i and the row vector of \mathbf{e}_i have many zero elements (much more than half). Utilizing computing features of sparse matrix and vector from [18, 19], we have knowledge that the computational complexity of $\frac{(\mathbf{F}_{i-1})^{-1} \cdot \mathbf{l}'_i \cdot \mathbf{e}_i \cdot (\mathbf{F}_{i-1})^{-1}}{1 + \mathbf{e}_i \cdot (\mathbf{F}_{i-1})^{-1} \cdot \mathbf{l}'_i}$ is $O(4K + 1)$. Thus, after $K - 1$ times of the iterative process, the whole computational complexity of \mathbf{L}^{-1} is $O(4K^2 - 3K - 1)$. Thus, we achieve that the computational complexity of CSM-based RZF precoding is $O(4K^2)$.

In [7], the Neumann-based precoding is

$$\mathbf{W}^{-1} \approx \sum_{n=0}^N (\mathbf{I}_K - \mathbf{D} \cdot \mathbf{Z})^n \cdot \mathbf{D}, \tag{34}$$

where $\mathbf{D} = \text{diag}(\frac{1}{w_{1,1}}, \frac{1}{w_{2,2}}, \dots, \frac{1}{w_{K,K}})$. Thus, when the $N \geq 3$, the complexity of the Neumann-based precoding is $O(K^3)$, which means the reduction in the complexity of RZF precoding is not obvious. On the other hand, the computational complexity of proposed CSM-based precoding is $O(K^2)$. When $L = 2$, complexity of Neumann-based precoding reduces to $O(K^2)$, but its performance would be greatly reduced.

In [8], the SOR-based precoding which is an improvement of Neumann is utilizing (22) to calculate the result of $\mathbf{W}^{-1} \cdot \mathbf{s}$ by iterating

$$\mathbf{t}^{(i+1)} = (\mathbf{D} + \omega \mathbf{L})^{-1} \left(\omega \mathbf{s} + ((1 - \omega) \mathbf{D} - \omega \mathbf{L}^H) \mathbf{t}^{(i)} \right), \tag{35}$$

where the \mathbf{D} and \mathbf{L} are diagonal component and lower triangular component respectively and satisfy $\mathbf{W} = \mathbf{D} + \mathbf{L} + \mathbf{L}^H$. $\mathbf{t}^{(i+1)}$ denote the result after $(i + 1)$ th iterator. ω is the relaxation parameter and can be computed by

$$\omega = a \cdot e^{-b \cdot (M/K)} + c. \tag{36}$$

Here, $a = 0.404$, $b = 0.323$ and $c = 1.035$. According the analyze the SOR scheme's computational complexity and performance, we have knowledge that when the times of iterating is $i = 3$, its performance of sum rate is near optimal but the performance of BER worse than CSM-based precoding, even its complexity is $O(4K^2)$ which is as same as CSM-based scheme. Therefore, CSM-based precoding has lower computational complexity than Neumann-based precoding and SOR-based precoding and better performance.

3.4 Pseudo code of CSM-based precoding algorithm

In this part, we only give the pseudo code of core part of the CSM-based precoding which is the process of calculating the inversion of matrix \mathbf{W} .

Algorithm 1 CSM-based precoding algorithm

- 1: INPUT: matrix W
- 2: initialize I_k and decompose matrix W to matrix L
- 3: decompose the matrix L to the matrix D and matrix L'
- 4: decompose the matrix L' to the form as $l'_1 \cdot e_1 + l'_2 \cdot e_2 + \dots + l'_{K-1} \cdot e_{K-1} + 0 \cdot e_K$
- 5: **for** $m = 0$ to $K - 1$ **do**
- 6: **if** $m = 0$ **then**
- 7: calculating the inversion of F_0
- 8: calculating the formula that $(F_0)^{-1} = \frac{(F_0)^{-1} \cdot l'_1 \cdot e_1 \cdot (F_0)^{-1}}{1 + e_1 \cdot (F_0)^{-1} \cdot l'_1}$
- 9: earning the inversion of F_1
- 10: **else**
- 11: utilizing the results F_{m-1}^{-1} which is obtained last iteration
- 12: calculating the formula that $(F_m)^{-1} = \frac{(F_m)^{-1} \cdot l'_{m+1} \cdot e_{m+1} \cdot (F_m)^{-1}}{1 + e_{m+1} \cdot (F_m)^{-1} \cdot l'_{m+1}}$
- 13: earning the inversion of F_{m+1}
- 14: **end if**
- 15: inversion of F_K is the matrix L^{-1}
- 16: **end for**
- 17: calculating the product of matrix L and matrix L^{-1}
- 18: OUTPUT: matrix W^{-1}

4 Simulation result

We provide the simulation results of average sum rate and BER of the proposed CSM-based precoding in a 256×16 massive MIMO system and a 256×32 massive MIMO system. For convenience, we set the regularized parameter $\phi = K/SNR$. The RZF precoding with exact matrix inversion is also set as benchmark. We compare the performance among RZF precoding, CSM-based scheme, Neumann approximate series and SOR-based precoding in one cell.

Figure 1 compares the average sum rate among RZF precoding, CSM-based scheme, Neumann-based precoding and SOR-based precoding. From the figure, we can observe that performance of CSM-based scheme as well as RZF and better than the SOR-based precoding and Neumann-based method. In addition, although during the i increasing, e.g. $i = 3$ which means the computational complexity of the SOR-based precoding and Neumann-based method precoding are $O(K^3)$ and $O(4K^2)$, respectively, the performance of SOR-based and Neumann-based have improvement in different level, the performance of Neumann and SOR are still close to but less than CSM-based scheme. So, the performance of CSM-based is the best among them.

Figures 2 and 3 show that the BER performance comparison in Rayleigh fading channels. From the two figures,

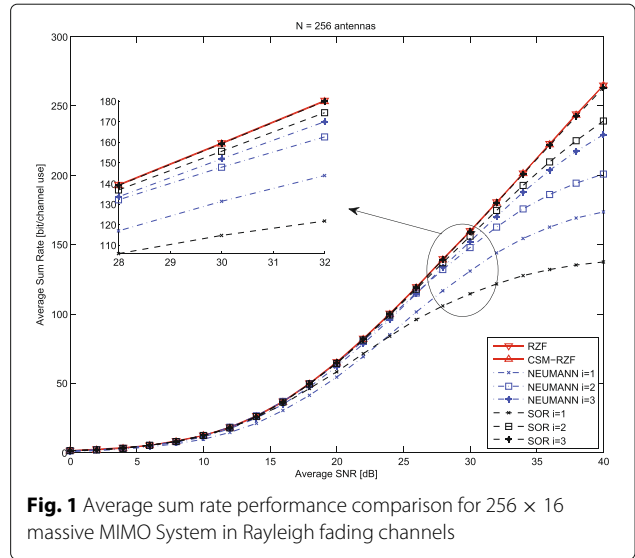


Fig. 1 Average sum rate performance comparison for 256×16 massive MIMO System in Rayleigh fading channels

we can obviously find that when the number of users goes to large in massive MIMO systems, the BER performance of all precoding schemes suffers from severe performance loss due to the limited number of BS antennas in practical systems, but the CSM-based scheme loss less than Neumann-based and SOR-based precoding which means its Robust is the best among them. Then, it is clear that when the i increasing, BER performance of Neumann-based and SOR-based precoding have improvement in some extend. But the BER performance of CSM-based precoding is still better than that of Neumann-based precoding and SOR-based precoding even $i = 3$ and close to the RZF. In addition, as SNR increasing, the performance of the proposed CSM-based precoding improves faster.

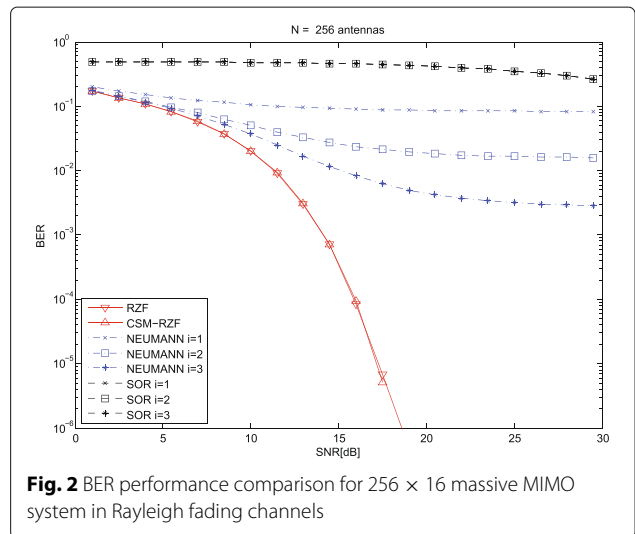
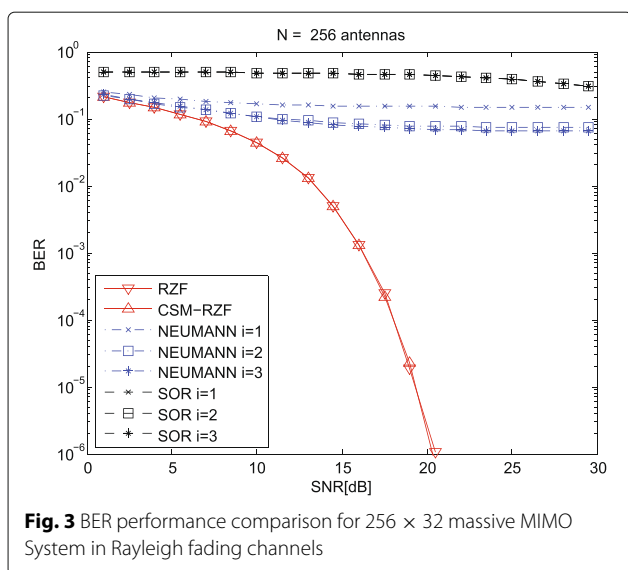


Fig. 2 BER performance comparison for 256×16 massive MIMO system in Rayleigh fading channels



5 Conclusions

In this paper, we exploit the special channel property of massive MIMO systems and some mathematic lemmas to propose the CSM-based scheme to reduce the computational complexity from $O(K^3)$ to $O(K^2)$. Meanwhile, CSM-based precoding scheme is able to achieve the near-optimal performance by decomposition and iteratively approach the exact matrix inversion of large size in RZF precoding or MMSE precoding. Simulation results utilize the RZF-based precoding as an example to illustrate that when increasing the SNR and keeping the N/K fixed, CSM-based precoding performances of BER and average sum rate are better than Neumann series and other schemes proposed based on Neumann series such as SOR-based precoding. Moreover, CSM-based scheme approaches the near-optimal performance of RZF precoding in Rayleigh fading channels.

Acknowledgements

This work was supported by the 863 Program of China under Grant (No.2015AA01A703), the Fundamental Research Funds for the Central Universities under grant (No.2014ZD03-02)NSFC (No. 61571055) and fund of SKL of MMW (No.K201501). Here, it is very grateful to the instructor and above funding for helping my subject.

Competing interests

The authors declare that they have no competing interests.

Author details

¹KeyLab of Universal Wireless Communications, MOE, Beijing University of Posts and Telecommunications, 100876 Beijing, People's Republic of China. ²State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, 100876 Beijing, People's Republic of China.

References

- EG Larsson, O Edfors, F Tufvesson, TL Marzetta, Massive MIMO for next generation wireless systems. *IEEE Commun. Mag.* **52**(2), 186–195 (2014)
- L Lu, GY Li, AL Swindlehurst, A Ashikhmin, R Zhang, An overview of massive MIMO: benefits and challenges. *IEEE J. Sel. Topics Signal Process.* **8**(5), 742–758 (2014)
- MH Costa, Writing on dirty paper (corresp). *IEEE Trans. Inf. Theory.* **29**(3), 439–441 (1983)
- JH Lee, Lattice precoding and pre-distorted constellation in degraded broadcast channel with finite input alphabets. *IEEE Trans. Commun.* **58**(5), 1315–1320 (2010)
- F Rusek, D Persson, BK Lau, EG Larsson, TL Marzetta, O Edfors, F Tufvesson, Scaling up MIMO: opportunities and challenges with very large arrays. *IEEE Signal Process. Mag.* **30**(1), 40–60 (2013)
- A Chockalingam, B Sundar Rajan, *Large MIMO Systems*, India Institute of Science. (Foundations: CAMBRIDGE UNIVERSITY PRESS, Inc., Bangalore, 2004), pp. 17–19
- H Prabhu, J Rodrigues, O Edfors, F Rusek, in *Wireless Communications and Networking Conference (WCNC), 2013*. Approximativematrix inverse computations for very-large MIMO and applications to linear pre-coding systems (IEEE, Shanghai, 2013), pp. 2710–2715
- T Xie, Q Han, H Xu, Z Qi, W Shen, in *81st Vehicular Technology Conference (VTC Spring)*. A Low-Complexity Linear Precoding Scheme Based on SOR Method for Massive MIMO Systems, (IEEE, Glasgow, 2015), pp. 1–5
- T Xie, Z Lu, Q Han, J Quan, B Wang, in *82nd Vehicular Technology Conference (VTC Fall)*. Low-Complexity LSQR-Based Linear Precoding for Massive MIMO Systems (IEEE, Boston, 2015), pp. 1–5
- A Muller, A Kammoun, E Bjornson, M Debbah, in *8th Sensor and Multichannel Signal Processing Workshop (SAM)*. Efficient Linear Precoding for Massive MIMO Systems using Truncated Polynomial Expansion (IEEE, A Coruña, 2014), pp. 273–276
- A Rontogiannis, V Kekatos, K Berberidis, A square-root adaptive V-BLAST algorithm for fast time-varying MIMO channels. *IEEE Signal Process. Lett.* **13**(5), 265–268 (2006)
- J Sherman, WJ Morrison, Adjustment of an inverse matrix corresponding to change in the elements of a given column or a given row of the original matrix (abstract). *Ann. Math. Statist.* **20**, 621 (1949)
- TL Marzetta, Noncooperative cellular wireless with unlimited numbers of base station antennas. *IEEE Trans. Wirel. Commun.* **9**(11), 3590–3600 (2010)
- L Dai, J Wang, Z Wang, P Tsiaflakis, M Moonen, Spectrum and energy-efficient OFDM based on simultaneous multi-channel reconstruction. *IEEE Trans. Signal Process.* **61**(23), 6047–6059 (2013)
- L Dai, Z Wang, C Pan, Z Yang, Wireless positioning using TD-SOFDM signals in single-frequency networks. *IEEE Trans. Broadcast.* **58**(2), 236–246 (2012)
- BL Ng, JS Evans, SV Hanly, D Aktas, Distributed downlink beamforming with cooperative base station. *IEEE Trans. Inf. Theory.* **54**(12), 5491–5499 (2008)
- C-J Wei, C-S Zhang, J Liu, A new method of solving inverse matrix based on Cholesky matrix, *Electron. Des. Eng.* **22**(1), 159–164 (2014)
- N Goharian, D Grossman, T El-Ghazawi, in *Proceedings: International Conference on Information Technology: Coding and Computing (ITCC), 2001*. Enterprise text processing: A sparse matrix approach (IEEE, Las Vegas, 2001)
- N Goharian, A Jain, Q Sun, Comparative analysis of sparse matrix algorithms for information retrieval. *J. Syst. Cybern. Inform.* **1**(1) (2010)