Signal Detection Methods Based on Less Matrix Inversion for Massive MIMO Systems

Ding Jiarui 1,2, Ding Wei1, and Zheng Yunsheng1

1State-Operated Factory NO.760, No.760, Weiqi Road, Xinxiang, Henan, China
2State-Operated Factory NO.760, No.760, Weiqi Road, Xinxiang, Henan, China

Abstract. In this paper, we will represent several methods that can reduce the computational complexity to detect signals for Uplink Massive MIMO Systems. Then we will show the simulation performance of these methods and analyse them. Finally we will give improvement for better performance.

CCS Concepts> Information systems → Database management system engines • Computing methodologies → Massively parallel and high-performance simulations.

1. INTRODUCTION

For uplink massive multiple-input multiple-output (MIMO) systems, some traditional methods such as linear minimum mean square error (MMSE) signal detection algorithm or zero forcing (ZF) signal detection method, can achieve the near-optimal performance. However, because of their high complexity about complicated matrix inversion, they are difficult to be implemented rapidly in large-scale MIMO systems in the future. This article will represent five promising approaches that can reduce the computational complexity from \(O(K^3)\) to \(O(K^2)\) and these methods can achieve the near-optimal performance with only a small number of iterations. Meanwhile, we will show the performance of these methods using simulation results of MATLAB according to the relationship between signal to noise rate (SNR) and bit error rate (BER), and analyse the advantages and disadvantages of them. In the end of the article one of these methods will be improved for better performance.

2. SYSTEM MODEL

We consider a uplink large-scale MIMO systems which employs \(N\) antennas at the base station (BS) to simultaneously serve \(K\) single-antenna users. Commonly \(N \gg K\), and we choose \(N = 256\) and \(K = 32\) in this paper. Firstly, we take samples from a set of constellation alphabet \(Q\), using quadrature amplitude modulation (QAM). Then the transmitted bit streams of \(K\) different users are encoded by channel encoder, and we choose binary gray code to encode separate bits, for adjacent numbers in gray code only have one bit’s difference.

\[
y = Hs + n \tag{1}
\]

The system model can be expressed as (1), \(y\) is a \(N \times 1\) received signal vector at the BS, \(H\) is a complex-valued \(N \times K\) flat Rayleigh fading channel matrix, \(s\) denotes the \(K \times 1\) transmitted signal vector that contains the transmitted samples from \(K\) users, \(n\) represents a \(N \times 1\) zero-mean additive white Gaussian noise (AWGN) vector, each entry of which has the variance power \(\delta^2\).

\[
W = G + \delta^2 I \tag{2}
\]

\[
s = (H^H H + \delta^2 I)^{-1} H^H y = W^{-1} y \tag{3}
\]

In traditional methods, such as MMSE or ZF, the transmitted symbol vector \(s\) can be estimated by minimum mean square error (MMSE) detection presented as \(s\) in (2). And \(y - H^H y\) can be interpreted as the matched-filter output of \(y\) the MMSE filtering matrix is denoted by \(W\) as (3), it can be proved that \(W\) is a hermitian positive definite matrix (HPD).

Although the MMSE algorithm is near-optimal, it inevitably involves inversion of \(W\), with the complexity of \(O(K^3)\). In order to avoid the direct computation of \(W^{-1}\), we will give five methods to achieve near-optimal performance with low complexity in the following sections.

Notation: In this paper we use lower-case and upper-case boldface letters to denote vectors and matrices: \(\mathbf{v}, \mathbf{v}', \mathbf{v}^*\), and \(|\cdot|\) denote the transpose, conjugate transpose, matrix inversion, and absolute operators respectively; \(\mathbf{I}_n\) represents the \(N \times N\) identity matrix; Finally, \(\text{Re}\{\cdot\}\) and \(\text{Im}\{\cdot\}\) denote the real part and imaginary part of a complex number.

2.1 Method 1 Detection Scheme Based on Joint Steepest Descent and Jacobi Method

This method combines the advantages of the steepest descent algorithm and Jacobi iteration. The steepest descent algorithm can make the system to achieve rapid convergence at the beginning of iteration in HPD matrix, so we use it to preprocess the data. Then, Jacobi iteration
is chosen for its great performance in diagonally dominant matric. The specific steps are as follows.

Step 1: we use an initial estimation (4) instead of a zero vector to get desired performance within limited numbers of iteration.

\[ x^{(0)} = D^{-1}b = D_{nn}^{-1}b \]

where \( k \)

D denote the diagonal component and \( k \) denote the inversion of D. It is obvious that the computation of D’s inversion is quite cheap.

\[ p^{(0)} = Wp^{(0)} \quad \rho^{(0)} = b - Wx^{(0)} \]

Step 2: use hybrid iteration as the first order of iteration

\[ u = \frac{\rho^{(0)}r}{\rho^{(0)} - \rho^{(0)}} \quad x^{(1)} = x^{(0)} + ur^{(0)} + D_{nn}^{-1}(r^{(0)} - up^{(0)}) \]

Step 3: use Jacobi iteration to K-1 order of iteration

\[ x^{(k)} = D_{nn}^{-1}[(D - A)x^{(k-1)} + b] \]

2.2 Method 2: symmetric successive over-relaxation (SSOR)

Please leave 3.81 cm (1.5") of blank text box at the bottom of the left column of the first page for the copyright notice.

Step 1: decompose W into three parts to calculate.

\[ W = D + L + L^H \]

D, L, and \( L^H \) denote the diagonal component, the strictly lower triangular component, and the strictly upper triangular component.

Step 2: separate the SOR iteration into two half iterations, which is symmetric successive over-relaxation (SSOR) iteration.

\[
(D + wL)_{s} = (1 - w)D_{s} - wL_{s} + wy
\]

\[
(D + wL^H)_{s} = (1 - w)D_{s} - wL_{s} + wy
\]

And we can find the optimal w as:

\[
w = \frac{2}{1 + \sqrt{2(1 - a)}} \cdot a = \left(1 + \frac{K}{\sqrt{N}}\right)^2 - 1
\]

In the system that involves relaxation parameter w, w is of great importance to a system’s performance. How to set w appropriately has been a problem in many systems. Compared with other methods that need to use w, the advantage of this method is that we need not to set the figure of w ourselves. w is a fixed number which is related to \( K/N \). Attention that every system has its special w, and it is not suitable to use the way we set w in other systems.

2.3 Method 3: Richardson

As the vectors in model (1) is complex value, we rewrite (1) as

\[
y_{c} = H_{c}s_{c} + n_{c}
\]

\[
y = [Re[y_{c}] \quad Im[y_{c}]]
\]

\[
s = [Re[s_{c}] \quad Im[s_{c}]]
\]

where \( n = [Re[n_{c}] \quad Im[n_{c}]] \)

In this method, we can get the value range of relaxation parameter w, \( 0 < w < \lambda \), \( \lambda \) is the largest
The advantage of this method is that it is too sensitive to the figure of \( w \). As we only have the value range of \( w \), how to set \( w \) appropriately is an important step in this method. If \( w \) is bigger that its range, the result will increase sharply. The result of computation shows that the range of \( w \) is \( 0<w<2 \). In order to get the best performance, we test the value of \( w \) manually and finally we choose \( w=1.1 \). But it is not wise to test \( w \) by people, we will improve this method in next part.

### 2.5 Method 5 successive over-relaxation (SOR)

Like method 3, we rewrite (1) as, \( y=Hs+n \) then we can get \( y=Hs+n \) where

\[
\begin{align*}
  y &= \begin{bmatrix} \text{Re}(y_1) \\ \text{Im}(y_1) \end{bmatrix} \\
  s &= \begin{bmatrix} \text{Re}(s_1) \\ \text{Im}(s_1) \end{bmatrix} \\
  n &= \begin{bmatrix} \text{Re}(n_1) \\ \text{Im}(n_1) \end{bmatrix}
\end{align*}
\]

Separate \( W \) into 3 parts as \( W = D + L + L' \). We set initial solution \( s^{(0)} \) as a \( 2K \times 1 \) zero vector, the iteration:

\[
s^{(i+1)} = (L + D)^{-1}(y - L'n^{(i)})
\]

In this system, the value range of \( w \) is \( 0<w<2 \). In order to get the best performance, we test the value of \( w \) manually and finally we choose \( w=1.1 \). But it is not wise to test \( w \) by people, we will improve this method in next part.

### 2.4 Method 4 : Gauss-Seidel Iteration

Like SSOR, we first decompose \( W \) into three parts as

\[
\begin{align*}
  s^{(i)} &= (D + L)^{-1}(y - L'n^{(i)}) \\
  s^{(0)} &= D^{-1}y
\end{align*}
\]

This method can achieve great performance when \( i=2 \), but with the increase of iteration number, we can not get more distinct anti-noise ability. So this method is suitable for the situation that need low computation complexity. What’s more, we set initial solution in this system, which help to accelerate the convergence of the system, and we can find that because we choose initial solution this system has better performance than those chose use zero vector.
2.6 Improvement Method

We compare the performance of these method when i=4. From the picture we can find method 5 has the optical anti-noise ability. From the table we can see method 4 need the lowest complexity, and other methods do not have big difference. However, method 4 has the worst anti-noise ability. So finally we choose this method 6 to do some improvement, because it has near-best performance in these method and it still have obvious defect.

| Table 1. computational statistics of these methods |
|-----------------------------------------------|
| Method  | Initial Solution | time of iteration |
|---------|------------------|-------------------|
| Method 1 | 2K(N^2+4KN+4K/2) | (4K-2-2K)^2+6K |
| Method 2 | 0                | (4K-2-4K)^2+6K   |
| Method 3 | 4K^2-2K          | (4K-2-2K)^2+6K   |
| Method 4 | 2K^2-2K          | (2K-2-2K)^2+6K   |
| Method 5 | 0                | (4K-2-4K)^2+6K   |

First, we combine zone-based initial solution with SOR, and from the picture we can see that initial solution obviously helps the system to converge in low orders.

Second, as we say before, we only have the value range 0<w<2, in order to explore the optical value of w, we design a program to text in what situation we can get the smallest spectral radius of W, then the corresponding w is the optical value. In workspace, we can see the value of w is also 1.10, which is equal to what we test manually. So this method can reduce the workload of people.

3. CONCLUSION

From analysis and experiment data above, we can draw a conclusion: we can take many measures that comes from approximately mathematical formulas to reduce the computational complexity to detect signals for Uplink Massive MIMO Systems, because the direct computation of matrix inversion is so complex. As different methods have different features, we can adopt these methods in different situations according to their characteristics. We can also combine these methods to remove defects and get better performance. We need to attention that when we try to find the joint method, we need careful analysis and accurate computational data, instead of simply applying one method to another.

ACKNOWLEDGMENTS

Thanks to the help of Mircoelectronics and Nanoengineering Division of Tsinghua National Laboratory for Information Science and Technology. And thanks to the help of PHD Peng Guiqiang of Tsinghua University.

REFERENCES

1. B. Yin, M. Wu, C. Studer, J. R. Cavallaro, and C. Dick, “Large-scale MIMO detection for 3GPP LTE: Algorithms and FPGA implementations,” IEEE J.Sel.Topics Signal Process., vol. 8, no.5, pp. -916-929, Mar.2014.
2. F. Rusek, D. Persson, B. K.Lau, E.G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, “Scaling up MIMO:Opportunities and challenges with very large arrays,” IEEE Signal Process. Mag., vol.30, no.1, pp.40-60.Jan.2013.
3. G. H. Golub and C. F. Van Loan, Matrix computations. JHU Press,2012.
4. H. D. Tiwari, N. B.Huynh, and B. C. Yong, "Flexible LDPC decoder using stream data processing for 802.11n and 802.16e," IEEE
5. J. Goldberger and A. Leshem, "MIMO detection for high-order QAM based on a Gaussian tree approximation," IEEE Trans. Inf. Theory, vol. 57, no. 8, pp. 4973-4982, Aug. 2011.

6. L. Dai, X. Gao, X. Su, S. Han, C. I, and Z. Wang, “Low-complexity soft-output signal detection based on gauss-seidel method for uplink multiuser large-scale MIMO systems,” IEEE Trans. Veh. Technol., vol. 64, no. 10, pp. 4839-4845, Oct. 2015.

7. L. Dai, Z. Wang, and Z. Yang, “Compressive sensing based time domain synchronous OFDM transmission for vehicular communication,” IEEE J. Sel. Areas Commun., vol. 31, no. 9, pp. 460-469, Sep. 2013.

8. L. G. Barbero and J. S. Thompson, “Fixing the complexity of the sphere decoder for MIMO detection,” IEEE Trans. Wireless Commun., vol. 7, no. 6, pp. 2131-2142, Jun. 2008.

9. R. W. Health, A. Lozano, T. L. Marzetta, and P. Popovski, “Five disruptive technology directions for 5G,” IEEE Commun. Mag., vol. 52, no. 2, pp. 74-80, Feb. 2014.

10. T. L. Marzetta, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” IEEE Trans. Wireless Commun., vol. 9, no. 11, pp. 3590-3600, Nov. 2010.