Short Communication

Low-complexity signal detection and precoding algorithms for multiuser massive MIMO systems

Abdelhak Boukharouba[1] · **Marwa Dehemchi**[1] · **Asma Bouhafer**[1]

Abstract

In this study, we present efficient detection and precoding algorithms for massive multi-user multiple-input multiple-output wireless system. To reduce the computational complexity due to large matrix inversion, the proposed algorithms are an enhanced version of zero forcing scheme based on QR matrix decomposition for both uplink and downlink systems. Through extensive numerical experiments, we demonstrate that the proposed algorithms outperform the recently published ones in terms of performance and complexity.

Keywords Multiuser MIMO wireless system · Precoding · Signal detection · QR based ZF algorithm · MMSE · 5G

1 Introduction

Massive multiple-input multiple-output systems combined with space-division multiple access (SDMA) techniques are expected to play a major role in next-generation wireless communication systems (5G) where the capability of serving a large number of users in the same time-frequency resource with higher data rates and improved spectral efficiency becomes an important requirement. Future wireless communication systems require a large number of antenna elements at the base station (BS) to increase the data rate and achieve high channel capacity [1]. Multi-user MIMO systems must adopt advanced technologies to support and serve a large number of users simultaneously whereas the high computational complexity of data processing at both sides transmitter and receiver is a major drawback and a big challenge [2–4]. In MU-MIMO systems, the multi-user interference (MUI) is the most serious problem that affects the signal detection at the receiver. Thus, the main objective of the previously published works was how to cancel the MUI using more

sophisticated algorithms for data detection in uplink and precoding in downlink.

In uplink systems, spatial demultiplexing or signal detection at the receiver is a challenging task for spatially multiplexed multiuser MIMO systems. The optimal data detection techniques in such systems are non-linear. To solve this problem, various algorithms based on Maximum likelihood (ML) criterion are proposed where optimal solutions are achieved [5]. However, their complexity increases exponentially as modulation order or the number of transmit antennas increases [6]. There are many nonlinear data detection algorithms with reduced complexity, such as the sphere decoder (SD) [7] that enables the ML detection for the MIMO system and tabu search (TS) based data detectors [8]. Unfortunately, for massive MIMO systems with large number of antennas and higher-order modulation schemes, such methods need high computation complexity [9, 10]. To make trade-off between performance and complexity, linear signal detection algorithms such as the zero-forcing and the minimum mean square error (MMSE) techniques have been proposed [5, 11]. However, both of these methods are based on a large-dimensional

✉ Abdelhak Boukharouba, boukharouba_abdelhak@hotmail.com | [1]Department of Electronics and Telecommunications, University of May 08th 1945, Guelma, Algeria.

matrix inversion, which achieve better performance at the expense of higher computational complexity. On the other hand, for the downlink transmission, a multi-antenna transmitter can communicate simultaneously with single or multiple antenna users where the pre-processing, which is called precoding, is made at the base station to facilitate the detection on the receiver side. In the literature, the proposed precoding algorithms for downlink systems can be also sub-divided into linear and nonlinear precoding types. Nonlinear algorithms achieve sum-capacity [12], whereas linear precoding approaches usually achieve best performance with much lower complexity. The famous types of linear precoding schemes are maximum ratio transmission (MRT) [13], zero-forcing [14] and minimum mean squared error [15]. The design and optimization of an optimal linear precoder is achieved using a weighted MMSE criterion where the precoding weights are computed by maximizing the ratio between the useful signal of each user and the multiuser interference plus noise [16].

Recently, several works have been published to improve the performance of both signal detection and precoding algorithms for multi-users massive MIMO mobile communication systems. In order to reduce the high complexity due to the explicit inverse of the channel matrix, Yin et al. [17] proposed an algorithm based on conjugate gradient (CG) for both data detection and precoding, whereas Zhang et al. [18] presented an efficient soft-output data detection algorithm based on Gauss–Seidel method. For large scale MIMO systems, the Gram matrix is diagonally dominant and thus the initial solution of the Gauss–Seidel method is chosen as a 2-term Neumann series expansion, which efficiently speeds up the convergence of the algorithm.

An efficient detection algorithm based on the alternating direction method of multipliers (ADMM) is proposed [19]. It performs infinity-norm or box-constrained equalization, which outperforms classical linear MMSE detectors in terms of packet error-rate (PER). Another work based on optimized coordinate descent (OCD) that performs approximate MMSE is presented in [20], while a near-optimal detection algorithm with low complexity which is based on the Richardson method is proposed in [21]. These latter have lower time complexity than that of algorithms based on an exact matrix inversion and deliver near-optimal results particularly for large ratio between BS antennas and user terminals.

In this paper, we propose an enhanced version of QR based zero forcing algorithms for both detection and precoding techniques. In small-scale MIMO systems with a small number of antennas, ZF scheme results in low performance due to noise amplification especially when the channel matrix is ill-conditioned. However, the columns of channel matrix $\mathbf{H}$ are asymptotically orthogonal in massive

MIMO systems and the Gram matrix $\mathbf{HH}^H$ is Hermitian positive definite [21, 22]. Consequently, the channel matrix is well conditioned and the Gram matrix is invertible, so simple linear detection schemes can perform well using fast algorithms for either exact or approximate matrix inversion.

The rest of this paper is organized as follows. Section 2 details the massive MU-MIMO model for both uplink and downlink systems and describes the proposed algorithms using QR decomposition. Section 3 provides a complexity analysis and error-rate performance comparison with recently published works. Section 4 concludes the paper.

2 Multiuser massive MIMO systems

Consider a multiuser massive MIMO wireless system where a base station is equipped with B antennas to communicate with U single-antenna users using space-division multiple access (SDMA) technique. The received vector $\mathbf{y}$ is expressed by the following equation [10].

$$\mathbf{y} = \mathbf{Hs} + \mathbf{n} \tag{1}$$

where $\mathbf{s}$ corresponds to the transmit vector, $\mathbf{H}$ is the channel matrix, and $\mathbf{n}$ stands for independent identically distributed (i.i.d.) Gaussian noise vector with mean zero and variance N_0 per entry. The optimal detection for the MIMO system is the ML detection, in which the receiver finds a vector s such that $\|\mathbf{y} - \mathbf{Hs}\|$ is minimized.

2.1 Uplink system model and data detection

In the uplink, the symbols for all users are mapped onto constellation points in the set O. The uplink system model is expressed as follows:

$$\mathbf{y} = \mathbf{H}_u \mathbf{s} + \mathbf{n} \tag{2}$$

where $\mathbf{s} \in O^U$ is the transmit vector with modulated symbols from all users, $\mathbf{y} \in \mathbb{C}^B$ is the receive vector and $\mathbf{H}_u \in \mathbb{C}^{B \times U}$ is the uplink channel matrix. Figure 1a illustrates the uplink channel model where the Gaussian noise vector $\mathbf{n}$ is introduced as an additive noise.

When channel state information (CSI) is available, we can use the matched filtering (MF) to maximize the SNR, which is known as maximal ratio transmission (MRT) scheme in the context of multiuser MIMO systems. The output of the matched filter has the following expression:

$$\mathbf{H}_u^H \mathbf{y} = \mathbf{H}_u^H \mathbf{H}_u \mathbf{s} + \mathbf{H}_u^H \mathbf{n} \tag{3}$$

To detect the vector $\mathbf{s}$, the least squares solution can be reduced to solving a system of normal equations:
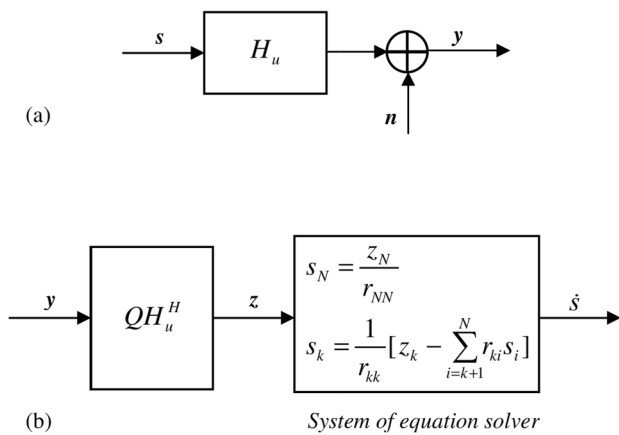
Fig. 1 Block diagram of MU-MIMO uplink system. **a** Uplink channel model, **b** data detection scheme

$$\mathbf{H}_u^H \mathbf{y} = \mathbf{H}_u^H \mathbf{H}_u \mathbf{s}. \tag{4}$$

As stated above, in massive MIMO systems the Gram matrix $\mathbf{H}_u^H \mathbf{H}_u$ is Hermitian positive definite [21, 22]. Consequently, the Gram matrix is invertible and the problem has a unique solution. A straightforward way to obtain an estimate $\hat{s}$ is to compute $\hat{s} = (\mathbf{H}_u^H \mathbf{H}_u)^{-1} \mathbf{H}_u^H \mathbf{y}$, which is exactly equivalent to the ZF scheme where the interference caused by $\mathbf{H}_u$ is forced to zero.

To avoid the direct matrix inverse of $\mathbf{H}_u^H \mathbf{H}_u$, we can factorize the Gram matrix using a simple and fast QR decomposition [23] where $\mathbf{R} = [r_{kj}]$ is an upper triangular matrix and $\mathbf{Q}$ is an orthogonal matrix. Thus, the Eq. (4) can be rewritten as follows:

$$\mathbf{Q} \mathbf{H}_u^H \mathbf{y} = \mathbf{R} \mathbf{s}. \tag{5}$$

We set $\mathbf{z} = \mathbf{Q} \mathbf{H}_u^H \mathbf{y}$, thus the Eq. (5) in the form $\mathbf{R} \mathbf{s} = \mathbf{z}$ is very easy to solve by an iterative process called back substitution for upper triangular matrices. The system of equation solver works according to the following equation.

$$s_k = \frac{1}{r_{kk}} \left[z_{kk} - \sum_{j=k+1}^{N} r_{kj} s_j \right] \tag{6}$$

where $s_N = z_N / r_{NN}$ and $k = N - 1 : 1$.

Figure 1b depicts a schematic illustration of data detection at the receiver using the above system of equation solver as an alternative way to detect the transmitted data without inverting the Gram matrix.
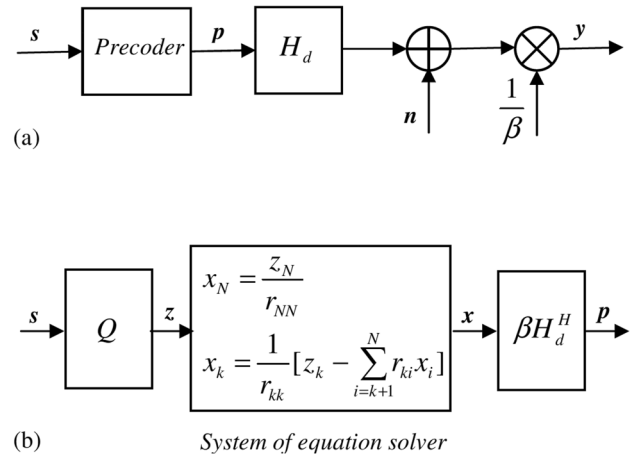
Fig. 2 MU-MIMO downlink system. **a** Block diagram of downlink model, **b** principle scheme of the precoder

2.2 Downlink system model and precoding

In downlink system, precoding is applied as pre-processing at the transmitter to facilitate detection on the receiver side without further processing. To cancel the multiuser interference, ZF scheme is introduced as pre-equalizer at the transmitter instead of the receiver. The downlink channel $\mathbf{H}_d \in \mathbb{C}^{U \times B}$ satisfies $\mathbf{H}_d = \mathbf{H}_u^H$ due to the reciprocity between uplink and downlink. Here, the BS encodes the bit streams for each user and then maps them to constellation points in O. The transmit vector $\mathbf{p} \in O^U$ containing the modulated symbols for all U users is precoded using the following linear precoder:

$$\mathbf{p} = \beta \mathbf{W} \mathbf{s} \tag{7}$$

where $\mathbf{W} = \mathbf{H}_d^H (\mathbf{H}_d \mathbf{H}_d^H)^{-1}$ is a ZF precoding matrix that allow to cancel the inter-user interference and β is a constant to meet the total transmitted power constraint after pre-coding and it is given as [5].

$$\beta = \sqrt{U / tr(\mathbf{W}\mathbf{W}^H)} \tag{8}$$

The received signal $\mathbf{H}_d \mathbf{p} + \mathbf{n}$ must be divided by β to compensate for the effect of amplification at the transmitter. Thus, the received vector becomes.

$$\mathbf{y} = \mathbf{s} + \mathbf{n}/\beta \tag{9}$$

Consequently the transmit vector s can be recovered directly from the received vector. Figure 2a shows the block diagram of MU-MIMO downlink model where the precoding is done on the transmitter side.

To avoid matrix inversion, we adopt the same technique proposed above in the Sect. 2.1. The matrix $\mathbf{H}_d \mathbf{H}_d^H$ must be factorized into QR factorization, $\mathbf{H}_d \mathbf{H}_d^H = \mathbf{Q} \mathbf{R}$. We can set $\mathbf{x} = (\mathbf{H}_d \mathbf{H}_d^H)^{-1} \mathbf{s}$ and the precoded vector

becomes $\mathbf{p} = \beta\mathbf{H}_d^H\mathbf{x}$ where $\mathbf{x}$ is the solution of the following equation.

$$\mathbf{Qs} = \mathbf{Rx} \qquad (10)$$

To compute $\mathbf{x}$, we can set $\mathbf{z} = \mathbf{Qs}$ and use the same system of equation solver described in the uplink system. Thus, instead of transmitting the original symbol s, we must transmit the precoded vector $\mathbf{p}$ which is the output of the precoder as explained in Fig. 2b.

3 Simulation results

To evaluate the performance of the proposed algorithm, we implemented several simulation experiments for single carrier based systems using high-order modulation scheme (64-QAM). We assume perfect knowledge of the channel state information at the base station in order to ensure multiuser detection in the uplink and precoding in the downlink. In our study, the estimate vector $\hat{s}$ and the precoded vector p are the exact solution of a system of normal equations using QR decomposition. Luckily, the QR decomposition is done during the preprocessing step, and thereby detection and precoding steps are extremely simplified. For QR decomposition, we have used a new simple and fast algorithm recently proposed in [23] that reduces the time complexity from $O(n^3)$ to $O(n^{2.529})$ using the fastest known matrix multiplication. The first part of the simulation is focused basically on signal detection in uplink channel. To analyze the computational complexity of the proposed algorithm, the evaluation is based on the number of complex-valued multiplications needed to detect one single transmitted symbol vector.

Table 1 compares the computational complexity of the proposed algorithm to other, recently published massive MIMO data detectors, namely conjugate gradient (CG)-based detector [17], Improved GS [18], ADMIN detector [19], OCDBOX [20], and Richardson detector [21].

In this table, the second column represents the computational complexity of the pre-processing step whereas the third column represents the number of complex-valued

multiplications needed for k iterations in the detection step where U is the number of users and B the number of BS antennas. Note that for ADMIN, the LDL decomposition is used with a time complexity of $O(u^{2.529})$ instead $O(u^3)$ in this comparison where, to the best of our knowledge, it is the first time that this algorithm is used in such applications.

In terms of number of operations, our algorithm provides an exact solution using a simple back substitution algorithm that costs only $U(U-1)/2$ complex-valued multiplications and consequently it is the fastest one compared to the other algorithms even for one single iteration (k = 1).

Table 2 shows the computational complexity in pre-processing and detection steps for two iterations (k = 2) and four iterations (k = 4) with B = 128 and U = 16, which are used in the next part to evaluate the bit error rate (BER) performance.

In terms of bit error rate (BER) performance analysis, the BER performance of the proposed algorithm is first compared with that of the well-known MMSE signal detection algorithm, which is near-optimal for multi-user MIMO systems, but requires the statistical information of noise and high complexity matrix inversion. In Fig. 3, we compare the BER of the MIMO systems based on the MMSE algorithm and the proposed algorithm. The modulation constellation is 64-QAM, B = 128 and U = 16, 8.

We observe that both algorithms attain practically the same BER as shown in Fig. 3a. However, if we zoom in on any part of the BER curves to make it appear much larger and more clearly, we can notice that the proposed algorithm presents a slight improvement in terms of BER performance as illustrated in Fig. 3b for the two cases.

Second, the simulation results of the BER performance with respect to the signal-to-noise ratio (SNR) are analyzed to compare the proposed signal detection algorithm with the recently published ones whose computational complexity is analyzed above. Figure 4 illustrates the BER performance of the proposed algorithm and the recently published ones for two iterations (k = 2) with B = 128, U = 16 and 64-QAM modulation scheme.

Table 1 Computational complexity comparison

Algorithm	Preprocessing	Detection
CG [17]	$2BU$	$k(U^2 + 6U)$
IGS [18]	$2BU$	$(k+2)U^2$
ADMIN [19]	$2BU + U^{2.529} + U(U-1)/2$	$2k(U^2 + U)$
OCD [20]	$BU + U$	$K(2BU + U)$
Richardson [21]	$2BU$	$k(U^2 + U)$
Proposed	$2BU + U^2 + U^{2.529}$	$U(U-1)/2$

Table 2 Computational complexity for $k = 2$ and $k = 4$

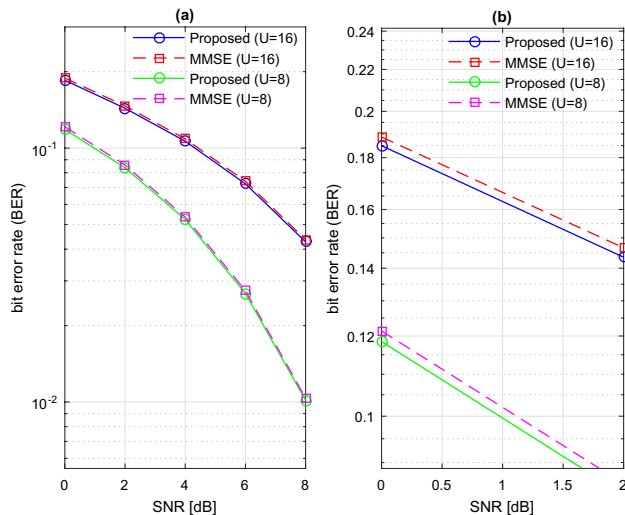| Algorithm	Preprocessing	Detection (k = 2)	Detection (k = 4)
CG [17]	4096	704	1408
IGS [18]	4096	1024	1536
ADMIN [19]	5325	1088	2176
OCDBOX [20]	2064	8224	16,448
Richardson [21]	4096	544	1088
Proposed	5461	120	120

Fig. 3 BER comparison between the MMSE and the proposed algorithm

Fig. 5 BER performance comparison for k = 4, B = 128 and U = 16
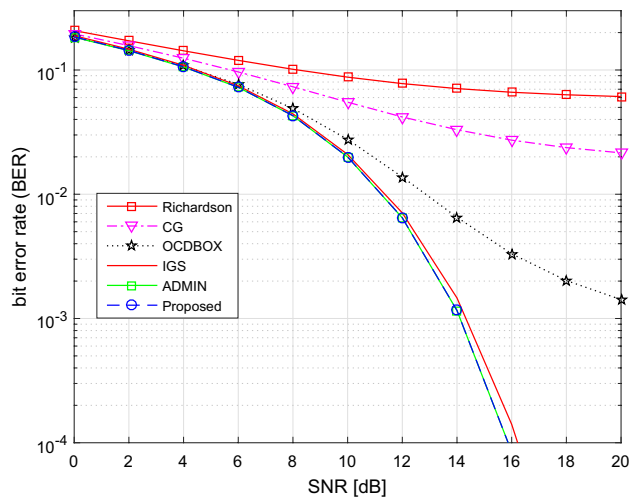

Fig. 4 BER performance comparison for k = 2, B = 128 ansd U = 16

From Fig. 4, one can observe that the proposed detector has a better error performance compared to all the others over the range of SNR for two iterations (k = 2), except ADMIN detector, which performs similar to our algorithm in this scenario, but at the cost of 1088 complex-valued multiplications in the detection step whereas the proposed algorithm requires only 120 operations while the pre-processing step needs practically the same number of operations for both detectors as illustrated in Table 2.

Another experience is performed for four iterations (k = 4) where the achieved results of the aforementioned detectors are compared and shown in Fig. 5

From this figure, it is shown that the BER performance of the proposed algorithm achieves higher throughputs
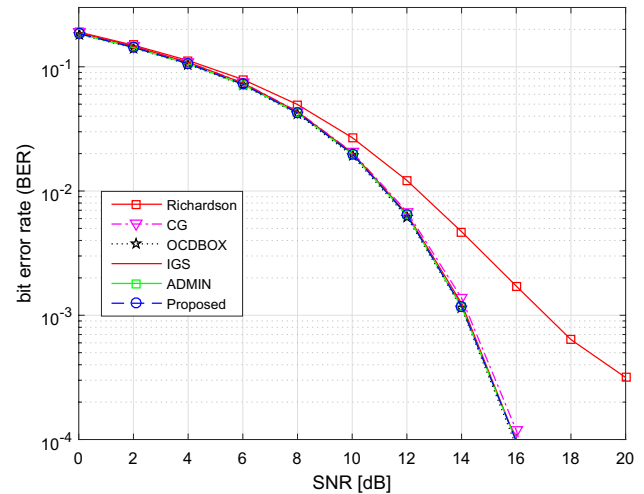
than those that approximate MMSE detector as demonstrated above in Fig. 3. However, the algorithms based on constraint optimization problems such as (ADMIN) [19] and (OCD) detector [20] perform slightly better in terms of error-rate performance but at an enormous cost where the computational complexity is 2176 and 16,448 complex-valued multiplications respectively in the detection phase (Table 2).

Finally, the performance of the proposed QR decomposition based ZF precoder and the MMSE precoding is analyzed. To avoid an explicit matrix inversion in the downlink system, most of linear precoders try to approximate the MMSE precoding matrix using low-complexity iterative algorithms by solving systems of linear equations. Consequently, we compare the proposed precoder with the exact matrix inversion based MMSE precoder. As the total number of arithmetic operations required is the same as in the detection algorithm, the comparison is only made in terms of bit error rate performance. Similar to the configuration above, the modulation constellation is 64-QAM, B = 128 and U = 16, 8 as illustrated in Fig. 6. We can notice that the BER performance of the proposed precoder is able to outperform that of the MMSE precoder at a lower cost in terms of time complexity.

4 Conclusion

We have proposed an efficient low complexity near-optimal detection and precoding algorithms for multi-user large-scale MIMO systems. By exploiting a very interesting property that the Gram matrix in massive MIMO systems is symmetric positive definite, we adopted an enhanced version of zero forcing scheme based on QR decomposition
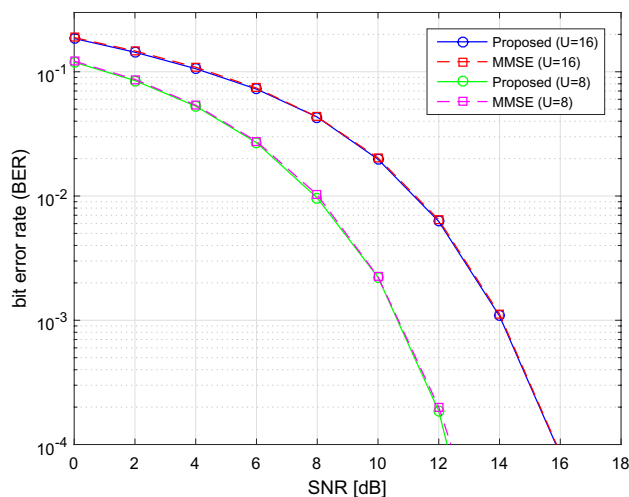
Fig. 6 BER performance of the MMSE and the proposed precoders

using a new fast algorithm that reduces considerably the computational complexity from $O(n^3)$ to $O(n^{2.529})$. The proposed scheme outperforms linear MMSE algorithm which is near-optimal for multi-user MIMO systems, but requires the statistical information of noise. Simulation results also show that the proposed algorithms outperform recently published ones in terms of computational complexity and BER performance. Thus, this approach can allow the implementation of realistic data detection and decoding in massive MU-MIMO systems, especially if we can develop a faster algorithm for QR matrix factorisation. Although the computational complexity is significantly reduced, the proposed algorithm is near optimal. The optimal solution which is obtained by solving a Maximum likelihood criterion may be achieved using the sphere decoder combined with the proposed scheme in the phase detection.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Chen X, Ng DWK, Yu W, Larsson EG, Al-Dhahir N, Schober R (2020) Massive access for 5G and beyond. arXiv preprint arXiv:2002.03491

2. Marzetta TL (2010) Noncooperative cellular wireless with unlimited numbers of base station antennas. IEEE Trans Wirel Commun 9(11):3590–3600

3. Larsson EG, Edfors O, Tufvesson F, Marzetta TL (2014) Massive MIMO for next generation wireless systems. IEEE Commun Mag 52(2):186–195

4. Hoydis J, Ten Brink S, Debbah M (2013) Massive MIMO in the UL/DL of cellular networks: how many antennas do we need? IEEE J Sel Areas Commun 31(2):160–171

5. Cho YS, Kim J, Yang WY, Kang CG (2010) MIMO–OFDM wireless communications with MATLAB. Wiley, London

6. Kim J, Kim Y, Kim K (2007) Computationally efficient signal detection method for next generation mobile communications using multiple antennas. SK Telecommun Rev 17(1C):183–191

7. Barbero LG, Thompson JS (2008) Fixing the complexity of the sphere decoder for MIMO detection. IEEE Trans Wirel Commun 7(6):2131–2142

8. Srinidhi N, Datta T, Chockalingam A, Rajan BS (2011) Layered Tabu search algorithm for large-MIMO detection and a lower bound on ML performance. IEEE Trans Commun 59(11):2955–2963

9. Jaldén J, Ottersten B (2005) On the complexity of sphere decoding in digital communications. IEEE Trans Signal Process 53(4):1474–1484

10. Verdú S (1989) Computational complexity of optimum multiuser detection. Algorithmica 4(1–4):303–312

11. Mahanta S, Rajauria A (2013) Analysis of MIMO system through ZF & MMSE detection scheme. Int J Electron Commun Technol (IJECT) 4

12. Weingarten H, Steinberg Y, Shamai SS (2006) The capacity region of the Gaussian multiple-input multiple-output broadcast channel. IEEE Trans Inf Theory 52(9):3936–3964

13. Lo TK (1999) Maximum ratio transmission. In: 1999 IEEE international conference on communications (Cat. No. 99CH36311), vol 2. IEEE, pp 1310–1314

14. Hashima S, Muta O (2020) Fast matrix inversion methods based on Chebyshev and Newton iterations for zero forcing precoding in massive MIMO systems. EURASIP J Wirel Commun Netw 1:1–12

15. Joham M, Utschick W, Nossek JA (2005) Linear transmit processing in MIMO communications systems. IEEE Trans Signal Process 53(8):2700–2712

16. Sheikh TA, Bora J, Hussain MA (2019) Capacity maximizing in massive MIMO with linear precoding for SSF and LSF channel with perfect CSI. Digit Commun Netw. https://doi.org/10.1016/j.dcan.2019.08.002

17. Yin B, Wu M, Cavallaro JR, Studer C (2014) Conjugate gradient-based soft-output detection and precoding in massive MIMO systems. In: IEEE global communications conference. IEEE 2014, pp 3696–3701

18. Zhang C, Wu Z, Studer C, Zhang Z, You X (2018) Efficient soft-output Gauss-Seidel data detector for massive MIMO systems. IEEE Trans Circuits Syst I Regul Pap. https://doi.org/10.1109/TCSI.2018.2875741

19. Shahabuddin S, Juntti M, Studer C (2017) ADMM-based infinity norm detection for large MU-MIMO: algorithm and VLSI architecture. In: IEEE international symposium on circuits and systems (ISCAS). IEEE 2017, pp 1–4

20. Wu M, Dick C, Cavallaro JR, Studer C (2016) High-throughput data detection for massive MU–MIMO–OFDM using coordinate descent. IEEE Trans Circuits Syst I Regul Pap 63(12):2357–2367

21. Gao X, Dai L, Ma Y, Wang Z (2014) Low-complexity near-optimal signal detection for uplink large-scale MIMO systems. Electron Lett 50(18):1326–1328

22. Rusek F, Persson D, Lau BK, Larsson EG, Marzetta TL, Edfors O, Tufvesson F (2012) Scaling up MIMO: opportunities and challenges with very large arrays. IEEE Signal Process Mag 30(1):40–60

23. Camarero C (2018) Simple, fast and practicable algorithms for Cholesky, LU and QR decomposition using fast rectangular matrix multiplication. arXiv preprint arXiv:1812.02056

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.