Hindawi
Wireless Communications and Mobile Computing
Volume 2019, Article ID 5609740, 12 pages
https://doi.org/10.1155/2019/5609740

Research Article

Signal Detection Based on Parallel Group Detection Algorithm for High-Load Massive MIMO Systems

Thanh-Binh Nguyen,[1] **Minh-Tuan Le**,[2] **and Vu-Duc Ngo**[3]

[1]*Le Quy Don Technical University, No. 236 Hoang Quoc Viet Street, Cau Giay Dist., Hanoi, Vietnam*
[2]*MobiFone R&D Center, MobiFone Corporation, VP1 Lot, Yen Hoa Ward, Cau Giay Dist., Hanoi, Vietnam*
[3]*Hanoi University of Science and Technology, No. 01 Dai Co Viet Street, Hai Ba Trung Dist., Hanoi, Vietnam*

Correspondence should be addressed to Thanh-Binh Nguyen; nguyenthanhbinhsqtt@gmail.com

Received 20 April 2019; Revised 29 July 2019; Accepted 13 September 2019; Published 12 December 2019

Academic Editor: André L. F. de Almeida

In this paper, a parallel group detection (PGD) algorithm is proposed in order to address the degradation in the bit error rate (BER) performance of linear detectors when they are used in high-load massive MIMO systems. The algorithm is constructed by converting the equivalent extended massive MIMO system into two subsystems, which can be simultaneously detected by the classical detection procedures. Then, using the PGD and the classical ZF as well as the QR-decomposition- (QRD-) based detectors, we proposed two new detectors, called ZF-based PGD (ZF-PGD) and QRD-based PGD (QRD-PGD). The PGD is further combined with the sorted longest basis (SLB) algorithm to make the signal recovery more accurate, thereby resulting in two new detectors, namely, the ZF-PGD-SLB and the QRD-PGD-SLB. Various complexity evaluations and simulations prove that the proposed detectors can significantly improve the BER performance compared to their classical linear and QRD counterparts with the practical complexity levels. Hence, our proposed detectors can be used as efficient means of estimating the transmitted signals in high-load massive MIMO systems.

1. Introduction

In recent years, massive multiple-input-multiple-output (massive MIMO) systems have been proposed to improve the quality of signal transmission in wireless communications. In a massive MIMO system, each cell site is equipped with very large number of antennas. Therefore, massive MIMO systems can provide not only high energy efficiency but also very high spectral one [1, 2]. Currently, the system with 128 antennas deployed at the BS, which simultaneously serves 8 single-antenna users, has been built successfully in laboratory [3]. Consequently, massive MIMO is expected to be one of the most important technologies for next generation cellular networks.

In massive MIMO, all complex signal processing, including signal detection for the uplink, precoding for the downlink, and channel estimation for both, should be implemented at the BS due to large dimension of the system [1]. For uplink scenario, all active users transmit their signals to the BS using the same time-frequency resources. These transmitted signal symbols are recovered at the BS by adopting suitable detectors. The detectors used in massive MIMO systems must satisfy the following requirements: (1) they should provide good BER performance or high spectral efficiency and (2) they should have low complexities. Low complexity linear detectors, such as zero-forcing (ZF) [4–7] or minimum mean square error (MMSE), can provide near-optimal bit error rate (BER) performance when they are used in massive MIMO systems [1]. In [8], the authors proved that if the system uses the Bell Laboratory Space Time (BLAST) detector, it will obtain a huge energy efficiency compared to that of the classical MMSE.

It is worth noting that the BER performance of the system depends on the so-called load factor β, defined by the ratio of the total number of antennas equipped at the users'

side and the ones at the BS. In [9], Björnson et al. declared that there exists no specific value of system's load factor and the system can be defined unconventionally with arbitrary number of antennas and users. This means, massive MIMO can be defined as either a high-load system (i.e., $\beta \approx 1$) or a low-load one (i.e. $\beta \ll 1$). When β is high, the BER performance given by linear detectors presented in [1] becomes worse. Thus, they are not suitable detectors for signal recovery in high-load scenarios. In spite of the fact that it can provide high BER performance, the BLAST detector in [8] is also not applicable to massive MIMO systems with hundreds of antennas due to its high complexity.

In order to improve BER performance of linear detectors in high-load massive MIMO systems, Post Detection Sparse Error Recovery (PDSR) algorithm [10] can be applied. Nevertheless, this algorithm only uses binary phase shift keying (BPSK) or quadrature phase shift keying (QBSK). As a result, the overall system throughput is limited. Another way to improve BER performance of the system is to artificially reduce load factor of the system. In [11], Nguyen et al. proposed group detection (GD) algorithm and the efficient low complexity detectors called the ZF-GD and ZF-IGD. The algorithm is built by dividing a high-load massive MIMO system to two separate subsystems with smaller load factors. As a consequence, these detectors significantly outperform their classical ZF counterpart. Unfortunately, in full-load systems ($\beta = 1$), they underperform than the classical MMSE detector, thereby making them less attractive.

In this paper, we develop our idea called the parallel group detection (PGD) algorithm in [12] for high-load massive MIMO systems. The algorithm is built by dividing the equivalent extended form of high-load massive MIMO system into two smaller load subsystems, which can be detected in parallel by utilizing the classical detectors. Based on the PGD, the classical ZF, and the QRD detectors, we are able to create two new detectors, namely, ZF-PGD and QRD-PGD. In order to improve the accuracy of the signal detection process, the shortest longest basis (SLB) algorithm [13], an efficient lattice reduction technique, is further combined with the PGD to generate two other new detectors, called ZF-PGD-SLB and QRD-PGD-SLB. The empirical simulations and complexity evaluations show that the proposed detectors can significantly improve the BER performance of massive MIMO systems compared to those of the classical linear detectors and the QRD while their complexities are at the practical levels. Consequently, they are the good candidates for signal recovery in high-load massive MIMO systems.

Notations: $\mathbb{C}$ denotes set of complex numbers; $\mathcal{Q}$ is the slicing operation, which slices the received signals to the nearest values in the set of integer numbers corresponding to the QAM constellation; $\mathbf{I}_N$ is a $N \times N$ identity matrix, and 1_N is a $N \times 1$ one vector, whose elements are all 1; $(\blacksquare)^T$ and $(\blacksquare)^H$ are the transpose and Hermitian transpose operations; $\mathbb{E}[\blacksquare]$ and $\otimes$ denote the expectation operation and Kronecker product, respectively; and $\mathbf{P}^\dagger$ is the pseudoinverse of matrix $\mathbf{P}$ and $[\blacksquare]$ is the round operation.

2. Uplink Massive MIMO System Model

Let us consider an uplink scenario of a single-cell massive MIMO system, as depicted in Figure 1. In the system, the BS with N_r antennas, located at the cell's origin, communicates with K multiple-antenna users using the same frequency and time resources. Each user has N_T antennas. In order to obtain high spectral efficiency, each user is assumed to use spatial division multiplexing (SDM) scheme. Using this scheme, the serial data stream of each user is first converted into N_T substreams, and then they are transmitted simultaneously by this user's antennas to the BS. Under the above assumptions, the overall transmit signal vector from K users, $\mathbf{x} \in \mathbb{C}^{N\times 1}$, $N = KN_T$, can be expressed as

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1^T & \mathbf{x}_2^T & \cdots & \mathbf{x}_K^T \end{bmatrix}^T, \tag{1}$$

where $\mathbf{x}_i \in \mathbb{C}^{N_T\times 1}$, $i = 1, 2, \ldots, K$, is the transmit signal vector of the ith user satisfying $\mathbb{E}[\mathbf{x}_i\mathbf{x}_i^H] = E_s\mathbf{I}_{N_T}$, and E_s is the average energy of M-QAM signals.

The received vector at the BS, $\mathbf{y} \in \mathbb{C}^{N_r\times 1}$, can be modeled as follows:

$$\mathbf{y} = \sqrt{\frac{p_u}{N_T E_s}}\mathbf{H}\left(\mathbf{B} \otimes \mathbf{I}_{N_T}\right)^{1/2}\mathbf{x} + \mathbf{n}, \tag{2}$$

where p_u denotes the transmit power of each user; $\mathbf{H} \in \mathbb{C}^{N_r\times N}$ and $\mathbf{n} \in \mathbb{C}^{N_r\times 1}$ are, respectively, small-scale fading channel matrix and noise vector, whose entries are independent identically distributed (i.i.d) random variables with zero mean and unit variance; and $\mathbf{B}$ is a $K \times K$ diagonal matrix including the large-scale fading coefficients. It is reasonable to assume that the large-scale fading coefficients are the same for one user and different from those of the others because the distances between the antenna elements are very much smaller than those from the users to the BS. The large-scale fading coefficient corresponding to ith user is given by [14]

$$b_{i,i} = \frac{z_i}{\left(d_i/d_0\right)^\gamma}, \tag{3}$$

where $b_{i,i}$ is the ith diagonal element of $\mathbf{B}$; z_i is a random variable, representing the shadowing, with zero mean and variance σ^2_{Shadow}; d_i and d_0, respectively, denote the distance from the ith user to the BS and the reference distance; and γ is the path loss factor.

For the sake of simplicity, let us define $\mathbf{U} = \sqrt{p_u/(N_T E_s)}\mathbf{H}(\mathbf{B} \otimes \mathbf{I}_{N_T})^{1/2}$. Then, the system model in (2) can be rewritten as

$$\mathbf{y} = \mathbf{U}\mathbf{x} + \mathbf{n}, \tag{4}$$

or equivalently as

$$\mathbf{y}_{\text{ex}} = \mathbf{U}_{\text{ex}}\mathbf{x} + \mathbf{n}_{\text{ex}}, \tag{5}$$

where $\mathbf{y}_{\text{ex}} = \begin{bmatrix} \mathbf{y}^T & 0_N^T \end{bmatrix}^T$, $\mathbf{U}_{\text{ex}} = \begin{bmatrix} \mathbf{U}^T & (1/\sqrt{E_s})\mathbf{I}_N \end{bmatrix}^T$, and $\mathbf{n}_{\text{ex}} = \begin{bmatrix} \mathbf{n}^T & (-1/\sqrt{E_s})\mathbf{x}^T \end{bmatrix}^T$, respectively, represent the received signal vector, the channel matrix, and the noise vector of the extended massive MIMO system [15]. Normally, the

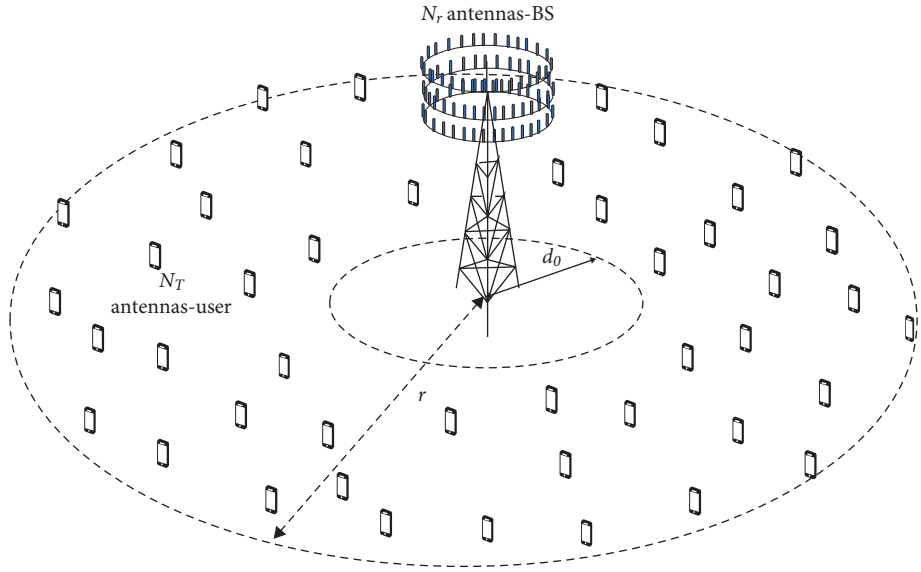

Figure 1: System model of a single-cell massive MIMO system.

receiver needs to get the channel state information (CSI), i.e., to estimate the channel matrix $\mathbf{U}$, through the adoption of some estimator, e.g., the MMSE estimator. In this paper, it is assumed that the CSI is perfectly known at the BS for the sake of simplicity.

3. Linear Detection and Its Drawbacks

Linear detectors, such as the ZF or the MMSE, have low complexity. Therefore, it is suitable for signal detection in massive MIMO systems. When the load factor of the system is sufficiently small (i.e., $\beta = N/N_r \ll 1$), the BER performances of the linear detectors become near optimal [1]. However, as the load factor approaches unit, performance of the linear detectors degrades noticeably. By applying the ZF detector to the original system in (4), all N transmitted symbols from K users can be simultaneously detected as follows:

$$\begin{aligned}\tilde{\mathbf{x}} &= \mathbf{U}^{\dagger}\mathbf{y} \\ &= \mathbf{x} + \mathbf{U}^{\dagger}\mathbf{n}.\end{aligned} \tag{6}$$

The weighted vector $\tilde{\mathbf{x}}$ is then sliced to obtain the final output of the ZF detector as

$$\hat{\mathbf{x}} = \mathcal{Q}(\tilde{\mathbf{x}}). \tag{7}$$

It is easy to observe from (6) that there is an error vector in estimation $\tilde{\mathbf{x}}$ given by $\mathbf{e} = \tilde{\mathbf{x}} - \mathbf{x} = \mathbf{U}^{\dagger}\mathbf{n}$. Therefore, the error covariance matrix is determined as follows:

$$\Phi^{\mathrm{ZF}} = \mathbb{E}\left[\mathbf{e}\mathbf{e}^{H}\right] = \left(\mathbf{U}^{H}\mathbf{U}\right)^{-1}. \tag{8}$$

The average mean square error (MSE) occurred when recovering one symbol can be determined by

$$\mathrm{MSE}^{\mathrm{ZF}} = \frac{1}{N}\mathrm{trace}\left(\Phi^{ZF}\right) = \frac{1}{N}\mathrm{trace}\left(\mathbf{U}^{H}\mathbf{U}\right)^{-1}. \tag{9}$$

Shown in Figure 2 is the empirical cumulative distribution function (ECDF) of the $\mathrm{MSE}^{\mathrm{ZF}}$ realized with 10^3 iterations when $N = 64$, $N_r = [64 : 64 : 264]$ antennas, or equivalently, the load factor $\beta = N/N_r = [1 : -0.25 : 0.25]$). The large-scale fading coefficients are determined by setting the SNR $p_u/\sigma^2 = 27$ dB (where σ^2 is the noise power), $d_0 = 100$ m, $100\,\mathrm{m} \le d_i \le 990$ m, $\sigma^2_{\mathrm{Shadow}} = 8$ dB, and $\gamma = 3.5$. The results in Figure 2 clearly show that the MSE significantly reduces as the load factor gets smaller. In the worse case, when $N = N_r = 64$ or $\beta = 1$, the MSE becomes very large. This indicates that the BER performance of the system will severely degrade. The reasons of this phenomenon can be two-fold: (1) the ZF detector suffers from the noise amplification effect and (2) the diversity order provided by the ZF detector reduces from $N_r - N + 1$ (for $N_r > N$) to 1.

It is remarkable that when the ZF detection procedures are utilized on extended system in (5), the resultant detector is called the MMSE detector. Therefore, MMSE detection and its MSE can be extended from the ZF ones by substituting channel matrix, $\mathbf{U}$, by its extended form, $\mathbf{U}_{\mathrm{ex}}$.

4. Proposed Detectors with Parallel Group Detection

4.1. Parallel Group Detection Approach. In order to address the problem of performance degradation in very high-load massive MIMO systems, we propose the so-called parallel group detection (PGD) algorithm. In the proposed approach, a massive MIMO system is divided into several parallel subsystems with lower load, which are then detected separately. As a result, the overall system performance can be noticeably improved. In general, the number of parallel branches can be from 2 to N. Nevertheless, the more the number of subsystems generated, the higher the detection complexity required. Therefore, in this paper, a system with two branches, i.e., two subsystems, is considered to keep the complexity at a reasonable level.

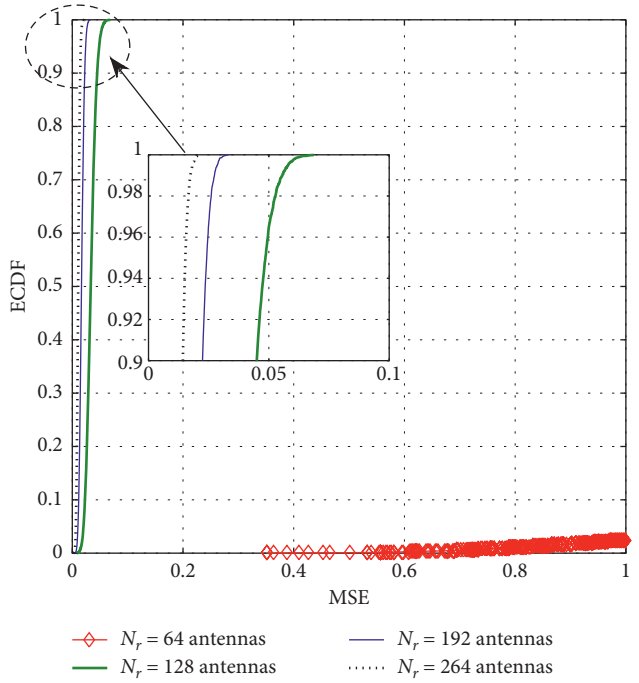

FIGURE 2: ECDF of MSE of the classical ZF detector realized in 10^3 iterations when $N_r = [64 : 64 : 264]$, $N_T = 4$, $K = 16$, $p_u/\sigma^2 = 27$ dB, $d_0 = 100$ m, $100\text{ m} \le d_i \le 990$ m, $\sigma^2_{\text{Shadow}} = 8$ dB, and $\gamma = 3.5$.

Figure 3 shows the detection scheme using the proposed PGD algorithm when number of parallel branches equals 2. The transmit signal vector $\mathbf{x}$ is estimated by the PGD algorithm through the following steps:

Step 1: convert the massive MIMO system into its equivalently extended form, as in (5)

Step 2: generate 2 subsystems in parallel

Step 3: estimate the transmitted subvectors (i.e., $\widehat{\mathbf{s}}_1$ and $\widehat{\mathbf{s}}_2$) by applying some classical detector on each subsystem

Step 4: rearrange the estimated subvectors to get the estimated signal vector $\widetilde{\mathbf{x}}$ as $\widetilde{\mathbf{x}} = \left[\widehat{\mathbf{s}}_1^T \;\; \widehat{\mathbf{s}}_2^T\right]^T$

In this section, Step 2 is described in detail. Steps 3 and 4 are presented in the following sections.

First of all, we rewrite the extended massive MIMO system in (5) as

$$\begin{aligned}\mathbf{y}_{\text{ex}} &= \begin{bmatrix}\mathbf{G}_1 & \mathbf{G}_2\end{bmatrix}\begin{bmatrix}\mathbf{s}_1\\ \mathbf{s}_2\end{bmatrix} + \mathbf{n}_{\text{ex}}\\ &= \mathbf{G}_1\mathbf{s}_1 + \mathbf{G}_2\mathbf{s}_2 + \mathbf{n}_{\text{ex}},\end{aligned} \tag{10}$$

where $\mathbf{G}_1 \in \mathbb{C}^{(N_r+N)\times L}$ and $\mathbf{G}_2 \in \mathbb{C}^{(N_r+N)\times(N-L)}$ are the subchannel matrices obtained by, respectively, taking first L and the remaining $(N - L)$ columns of the extended channel matrix $\mathbf{U}_{\text{ex}}$. Similarly, $\mathbf{s}_1 \in \mathbb{C}^{L\times 1}$ and $\mathbf{s}_2 \in \mathbb{C}^{(N-L)\times 1}$ denote the transmitted subvectors generated by taking the first L and the remaining elements of the transmitted signal vector $\mathbf{x}$. The value of L should be selected in such a way that the PGD-based detectors meet the following requirements: (1) the BER performances of users belonging to different groups are the nearly same and (2) the total complexities of proposed detectors get as low as possible. In [11], the authors pointed out that the group detection algorithm (GD) is able to achieve the best BER performance at the lowest possible computational cost when the sizes of both subsystems are identical. Following the method in [11], we can find out that $L = N/2$ is the best value of L that satisfies both of the above requirements. Therefore, throughout the paper, we select $L = N/2$ as the optimal value.

In order to generate the first subsystem, the interference term $\mathbf{G}_2\mathbf{s}_2$ in (10) needs to be eliminated. Let $\mathbf{P}_1 = (\mathbf{I}_{N_r+N} - \mathbf{G}_2\mathbf{G}_2^{\dagger})$ be the projection term. Then, multiplying both sides of (10) by $\mathbf{P}_1$, we get

$$\mathbf{P}_1\mathbf{y}_{\text{ex}} = \mathbf{P}_1\mathbf{G}_1\mathbf{s}_1 + \mathbf{P}_1\mathbf{n}_{\text{ex}}. \tag{11}$$

Now, define $\widetilde{\mathbf{y}}_1 = \mathbf{P}_1\mathbf{y}_{\text{ex}}$, $\widetilde{\mathbf{G}}_1 = \mathbf{P}_1\mathbf{G}_1$, $\widetilde{\mathbf{n}}_1 = \mathbf{P}_1\mathbf{n}_{\text{ex}}$ as received vector, channel matrix, and noise vector of the first subsystem; then, (11) is further rewritten as

$$\widetilde{\mathbf{y}}_1 = \widetilde{\mathbf{G}}_1\mathbf{s}_1 + \widetilde{\mathbf{n}}_1. \tag{12}$$

Similarly, the second subsystem is generated in parallel with the first one as follows:

$$\widetilde{\mathbf{y}}_2 = \widetilde{\mathbf{G}}_2\mathbf{s}_2 + \widetilde{\mathbf{n}}_2, \tag{13}$$

where $\widetilde{\mathbf{y}}_2 = \mathbf{P}_2\mathbf{y}_{\text{ex}}$, $\widetilde{\mathbf{G}}_2 = \mathbf{P}_2\mathbf{G}_2$, $\widetilde{\mathbf{n}}_2 = \mathbf{P}_2\mathbf{n}_{\text{ex}}$, and $\mathbf{P}_2 = (\mathbf{I}_{N_r+N} - \mathbf{G}_1\mathbf{G}_1^{\dagger})$.

It is noteworthy that the load factor given by each subsystem is $\beta_1 = \beta_2 = L/(N_r + N)$, which is a half of that of the original system. Therefore, the BER performance of the system is expected to improve as the PGD algorithm is adopted. The PGD algorithm is summarized as in Algorithm 1.

4.2. Proposed ZF-PGD and QRD-PGD Detectors. Theoretically, any classical detectors can be used as subdetectors in the PGD algorithm to estimate $\widehat{\mathbf{s}}_1$ and $\widehat{\mathbf{s}}_2$. However, low-complexity detectors should be adopted to keep the complexity at practical levels. In this section, we adopt the classical QRD and the ZF in the PGD to generate two new detectors, called the ZF-PGD and QRD-PGD, as described below.

4.2.1. ZF-PGD Detector. In the ZF-PGD algorithm, $\widehat{\mathbf{s}}_k$, $k = 1, 2$, are estimated by the ZF detection procedure to both subsystems in the PGD as follows:

$$\begin{aligned}\widehat{\mathbf{s}}_k &= \mathcal{Q}\left(\widetilde{\mathbf{G}}_k^{\dagger}\widetilde{\mathbf{y}}_k\right)\\ &= \mathcal{Q}\left(\widetilde{\mathbf{s}}_k + \widetilde{\mathbf{G}}_k^{\dagger}\widetilde{\mathbf{n}}_k\right).\end{aligned} \tag{14}$$

Similarly to the ZF detector, the error covariance matrices and MSEs for each branch in the ZF-PGD can be determined, respectively, as follows:

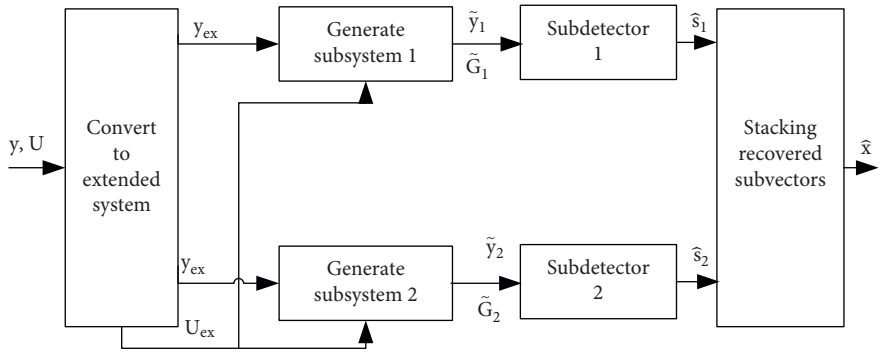

Figure 3: Block diagram of signal detection using the proposed PGD algorithm.

Input: $\mathbf{y}$, $\mathbf{U}$, K, N_T
Output: $\hat{\mathbf{x}}$
(1) Convert the system to equivalent extended form with $\mathbf{y}_{ex} = \left[\mathbf{y}^T \;\; 0_N^T\right]^T$; $\mathbf{U}_{ex} = \left[\mathbf{U}^T \;\; (1/E_s)\mathbf{I}_N\right]^T$.
(2) Set $L = (N/2)$ and define $\mathbf{G}_1 = \mathbf{U}_{ex}(:, 1:L), \mathbf{G}_2 = \mathbf{U}_{ex}(:, L+1:N)$.
(3) Compute $\mathbf{P}_1 = (\mathbf{I} - \mathbf{G}_2\mathbf{G}_2^{\dagger})$ and $\mathbf{P}_2 = (\mathbf{I} - \mathbf{G}_1\mathbf{G}_1^{\dagger})$.
(4) Compute received vectors and channel matrices of the two parallel subsystems as $\tilde{\mathbf{y}}_1 = \mathbf{P}_1\mathbf{y}_{ex}$, $\tilde{\mathbf{G}}_1 = \mathbf{P}_1\mathbf{G}_1$, $\tilde{\mathbf{y}}_2 = \mathbf{P}_2\mathbf{y}_{ex}$, $\tilde{\mathbf{G}}_2 = \mathbf{P}_2\mathbf{G}_2$.
(5) Estimate both transmitted subvectors $\hat{\mathbf{s}}_k$, $(k = 1, 2)$, by applying a suitable subdetector to the kth subsystem having $\tilde{\mathbf{y}}_k$, $\tilde{\mathbf{G}}_k$.
(6) Rearrange the recovered subvectors to get the overall estimated vector as $\hat{\mathbf{x}} = \left[\hat{\mathbf{s}}_1^T \;\; \hat{\mathbf{s}}_2^T\right]^T$.

Algorithm 1: Proposed PGD algorithm.

$$\begin{aligned}\Phi_k^{\text{ZF-PGD}} &= \mathbb{E}\left[\left(\mathbf{s}_k - \tilde{\mathbf{s}}_k\right)\left(\mathbf{s}_k - \tilde{\mathbf{s}}_k\right)^H\right] \\ &= \mathbb{E}\left[\tilde{\mathbf{G}}_k^{\dagger}\tilde{\mathbf{n}}_k\left(\tilde{\mathbf{G}}_k^{\dagger}\tilde{\mathbf{n}}_k\right)^H\right] \\ &= \left(\tilde{\mathbf{G}}_k^H\tilde{\mathbf{G}}_k\right)^{-1},\end{aligned} \tag{15}$$

$$\text{MSE}_k^{\text{ZF-PGD}} = \frac{1}{L}\text{trace}\left(\left(\tilde{\mathbf{G}}_k^H\tilde{\mathbf{G}}_k\right)^{-1}\right). \tag{16}$$

Shown in Figure 4 is the ECDF of the classical linear detectors and the ones based on PGD algorithm when the parameters are set almost the same as in Figure 2 except $N_r = 64$ antennas. The illustration shows that the MSEs of the two subsystems detected by the ZF-PGD are equal to each other and much smaller than that of the classical ZF detector. They are even comparable with the MSE of the MMSE detector. This implies that the ZF-PGD can improve the system performance significantly compared to the ZF one.

4.2.2. QRD-PGD Detector. In order to further improve the BER performance of the PGD algorithm, the conventional QRD detector is applied to the subsystems to generate the QRD-PGD detector as follows.

First, using the QR decomposition to decompose subchannel matrices, $\tilde{\mathbf{G}}_k$, $k = 1, 2$, we get

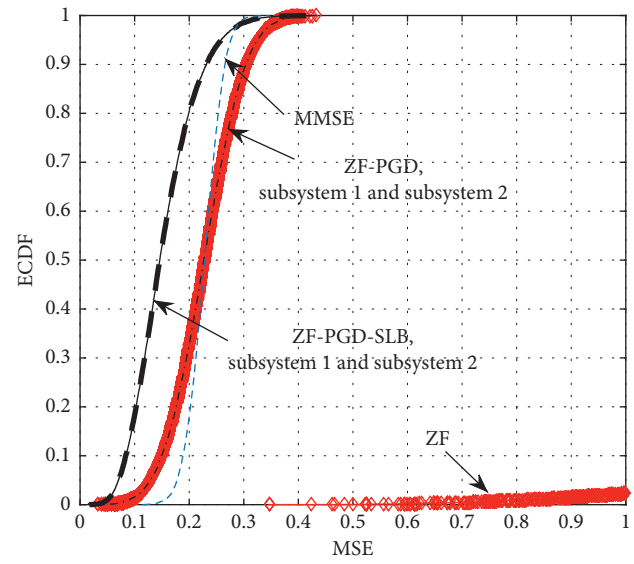

Figure 4: ECDF curves of the MSEs of the classical linear detectors: the ZF-PGD and the ZF-PGD-SLB detectors realized in 10^3 iterations when $N_r = 64$, $N_T = 4$, $K = 16$, $p_u/\sigma^2 = 27$ dB, $d_0 = 100$ m, $100\text{ m} \le d_i \le 990$ m, $\sigma^2_{\text{Shadow}} = 8$ dB, and $\gamma = 3.5$.

$$\tilde{\mathbf{G}}_k = \mathbf{Q}_k\mathbf{R}_k, \tag{17}$$

where $\mathbf{Q}_k$ is $(N_r + N) \times L$ unitary matrices satisfying that $\mathbf{Q}_k^H\mathbf{Q}_k = \mathbf{I}_L$ and $\mathbf{R}_k$ be $L \times L$ upper triangle matrices, which can be presented in the matrix form asfollows:

$$\mathbf{R}_k = \begin{bmatrix} r_{k_{1,1}} & r_{k_{1,2}} & \cdots & r_{k_{1,L}} \\ 0 & r_{k_{2,2}} & \cdots & r_{k_{2,L}} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & r_{k_{L,L}} \end{bmatrix}, \tag{18}$$

where $r_{k_{i,j}}$ denotes the ith row and jth column entry of $\mathbf{R}_k$. Next, multiplying both sides of (17) by $\mathbf{Q}_k^H$, we obtain

$$\mathbf{v}_k = \mathbf{Q}_k^H \widetilde{\mathbf{y}}_k = \mathbf{R}_k \mathbf{s}_k + \mathbf{Q}_k^H \widetilde{\mathbf{n}}_k. \tag{19}$$

Finally, ignoring the noise term $\mathbf{Q}_k^H \widetilde{\mathbf{n}}_k$ in (19), all entries of $\widehat{\mathbf{s}}_k$ (denoted by $\widetilde{\mathbf{s}}_{k_i}, i = L, L-1, \ldots, 1$) are determined layer by layer using the following rule:

$$\widehat{\mathbf{s}}_{k_i} = \mathcal{Q}\left(\widetilde{\mathbf{s}}_{k_i}\right) = \mathcal{Q}\left(\begin{cases} \dfrac{v_{k_i}}{r_{k_{i,i}}}, & i = L, \\ \dfrac{v_{k_i} - \sum_{j=i+1}^{L} \left(r_{k_{i,j}} \widehat{\mathbf{s}}_{k_j} \right)}{r_{k_{i,i}}}, & i \neq L, \end{cases} \right), \tag{20}$$

where v_{k_i} is ith entry of $\mathbf{v}_k$ and $\widehat{\mathbf{s}}_{k_j}$ is the sliced value of $\widetilde{\mathbf{s}}_{k_j}$.

Note that the BER performance given by the QRD-PGD is usually better than that of its ZF-PGD counterpart due to the advantage of the successive interference cancelation (SIC) technique. However, it suffers from the larger delay of the SIC technique. The higher the dimensions of the system are, the higher delay the QRD-PGD suffers. Therefore, the QRD-PGD detector is suitable to be used in small- or medium-size systems.

5. Proposed ZF-PGD-SLB and QRD-PGD-SLB Detectors

5.1. Lattice Reduction-Aided PGD Detection Approach. In this section, a new detection approach, called PGD-LR, is considered by combining the lattice reduction (LR) techniques with the PGD to further enhance the BER performance of massive MIMO systems.

First of all, the subsystems, $\widetilde{\mathbf{y}}_k$, $k = 1, 2$, are converted into the ones in LR domain as follows:

$$\begin{aligned} \widetilde{\mathbf{y}}_k &= \widetilde{\mathbf{G}}_k \mathbf{s}_k + \widetilde{\mathbf{n}}_k \\ &= \widetilde{\mathbf{G}}_k \mathbf{T}_k \mathbf{T}_k^{-1} \mathbf{s}_k + \widetilde{\mathbf{n}}_k \\ &= \widetilde{\mathbf{G}}_k^{(LR)} \mathbf{c}_k + \widetilde{\mathbf{n}}_k, \end{aligned} \tag{21}$$

where $\mathbf{c}_k = \mathbf{T}_k^{-1} \mathbf{s}_k$, $\widetilde{\mathbf{G}}_k^{(\mathrm{LR})} = \widetilde{\mathbf{G}}_k \mathbf{T}_k$ is the channel matrix in the LR domain and $\mathbf{T}_k$ is a unimodular matrix (i.e., $\det(\mathbf{T}_k) = \pm 1$), whose entries are integer numbers. $\widetilde{\mathbf{G}}_k^{(\mathrm{LR})}$ and $\mathbf{T}_k$ are determined by applying LR techniques on the channel matrix $\widetilde{\mathbf{G}}_k$.

Next, we apply some classical detectors to (21) to estimate the transmitted vector in LR domain, $\widetilde{c}_k$. It is worth noting that the entries of $\mathbf{c}_k$ must be selected from a set of consecutive integer numbers. Therefore, if the signal is modulated by M-QAM method, it needs to be converted to the set of consecutive integer number by the shift-and-scaling operation. Let us define $m = \log(M), \overline{\alpha} = 1/2, \overline{\beta} = (m-1)(1+j)/2$, where $j^2 = -1$. Then, the shift-and-scaling transmitted signal vector is $\overline{\mathbf{s}}_k = \overline{\alpha}\mathbf{s}_k + \overline{\beta}$, thereby leading to the shift-and-scaling signal vector in the LR as follows:

$$\overline{\mathbf{c}}_k = \mathbf{T}_k^{-1} \overline{\mathbf{s}}_k = \overline{\alpha} \mathbf{c}_k + \overline{\beta} \mathbf{T}_k^{-1} 1_L. \tag{22}$$

Therefore, the hard decision of $\overline{\mathbf{c}}_k$ is given by

$$\widehat{\mathbf{c}}_k = \frac{1}{\overline{\alpha}} \left(\lfloor \overline{\alpha} \widetilde{\mathbf{c}}_k + \overline{\beta} \mathbf{T}_k^{-1} 1_L \rfloor - \overline{\beta} \mathbf{T}_k^{-1} 1_L \right). \tag{23}$$

Once $\widehat{\mathbf{c}}_k$ is determined, the estimated signal subvector, $\widetilde{\mathbf{s}}_k$, is obtained as follows:

$$\widetilde{\mathbf{s}}_k = \mathbf{T}_k \widehat{\mathbf{c}}_k. \tag{24}$$

Finally, $\widetilde{\mathbf{s}}_k$ is sliced to get the final output of kth subsystem as $\widehat{\mathbf{s}}_k = \mathcal{Q}(\widetilde{\mathbf{s}}_k)$.

5.2. Proposed ZF-PGD-SLB and QRD-PGD-SLB Detectors. In practice, $\widetilde{\mathbf{G}}_k^{(\mathrm{LR})}$ and $\mathbf{T}_k$ can be carried out by utilizing LR techniques such as Lenstra–Lenstra–Lovasz (LLL) algorithm [16], Seysen algorithm (SA) [17], or element-based lattice reduction (ELR) [13] on the subchannel matrix, $\widetilde{\mathbf{G}}_k$. Among these LR techniques, ELR [13] is the most suitable to be used in massive MIMO systems because of its high performance and low complexity. The ELR finds out $\widetilde{\mathbf{G}}_k^{(\mathrm{LR})}$ and $\mathbf{T}_k$ matrices from $\widetilde{\mathbf{G}}_k$ by minimizing the diagonal entries of error covariance matrix $\Phi_k = (\widetilde{\mathbf{G}}_k^H \widetilde{\mathbf{G}}_k)^{-1}$ in (15). This technique includes two versions called shortest longest vector (SLV) algorithm and shortest longest basis (SLB). While the SLB minimizes all diagonal entries of Φ_k, the SLV only investigates the maximum one. As a sequence, BER performance given by the SLB is better than that of the SLV, yet at the cost of higher complexity. In this work, we use the SLB as the LR technique and combine it with the proposed ZF-PGD and QRD-PGD to create the ZF-PGD-SLB and QRD-PGD-SLB ones. Due to space limitation, we just summarize the SLB adopted on channel matrix $\widetilde{\mathbf{G}}_k$ in Algorithm 2. For more details of this technique, the readers can find them in [13].

5.2.1. ZF-PGD-SLB Detector. In the ZF-PGD-SLB detection algorithm, the transmitted symbols in LR domain $\widetilde{\mathbf{c}}_k$, $k = 1, 2$, are determined by utilizing the ZF detection procedure as follows:

$$\widetilde{\mathbf{c}}_k = \left(\widetilde{\mathbf{G}}_k^{(\mathrm{LR})H} \widetilde{\mathbf{G}}_k^{(\mathrm{LR})} \right)^{-1} \widetilde{\mathbf{G}}_k^{(\mathrm{LR})H} \widetilde{\mathbf{y}}_k. \tag{25}$$

Once $\widetilde{\mathbf{c}}_k$ is obtained, the estimated subvector, $\widehat{\mathbf{s}}_k$, can be easily determined using (23) and (24).

Similar to the ZF detector, the error covariance matrix and the MSE for the kth subsystems in the LR domain are, respectively, given by

$$\begin{aligned}\boldsymbol{\Phi}_k^{\mathrm{ZF-PGD-SLB}} &= \mathrm{E}\left[\left(\mathbf{c}_k - \widetilde{\mathbf{c}}_k\right)\left(\mathbf{c}_k - \widetilde{\mathbf{c}}_k\right)^H\right] \\ &= \left(\widetilde{\mathbf{G}}_k^{(\mathrm{LR})H}\widetilde{\mathbf{G}}_k^{(\mathrm{LR})}\right)^{-1},\end{aligned} \tag{26}$$

$$\mathrm{MSE}_k^{\mathrm{ZF-PGD-SLB}} = \frac{1}{L}\mathrm{trace}\left(\widetilde{\mathbf{G}}_k^{(\mathrm{LR})H}\widetilde{\mathbf{G}}_k^{(\mathrm{LR})}\right)^{-1}. \tag{27}$$

Note that the error in estimating $\widehat{\mathbf{s}}_k$ depends on the accuracy of $\widetilde{\mathbf{c}}_k$ in (25). Hence, (27) can be used as the MSE for estimating $\widehat{\mathbf{s}}_k$. It can be seen from Figure 4 that the MSEs of the subsystems detected by the ZF-PGD-SLB are almost identical and much smaller than those of the ZF and the ZF-PGD counterparts. This indicates that the ZF-PGD-SLB detector is able to noticeably outperform the ZF and the ZF-PGD ones.

5.2.2. QRD-PGD-SLB Detector. In the QRD-PGD-SLB detection algorithm, the transmitted symbols in LR domain $\widetilde{\mathbf{c}}_k$, $k = 1, 2$, is determined by adopting the SIC detection technique as follows.

First, use QR decomposition to decompose $\widetilde{\mathbf{G}}_k^{(\mathrm{LR})}$ as

$$\widetilde{\mathbf{G}}_k^{(\mathrm{LR})} = \mathbf{Q}_k^{(\mathrm{LR})}\mathbf{R}_k^{(\mathrm{LR})}, \tag{28}$$

where $\mathbf{Q}_k^{(\mathrm{LR})}$ and $\mathbf{R}_k^{(\mathrm{LR})}$ are, respectively, the unitary matrix and upper triangle one.

Next, multiplying both sides of (28) by $\mathbf{Q}_k^{(\mathrm{LR})H}$, we get

$$\begin{aligned}\mathbf{v}_k^{(\mathrm{LR})} &= \mathbf{Q}_k^{(\mathrm{LR})H}\widetilde{\mathbf{y}} \\ &= \mathbf{R}_k^{(\mathrm{LR})}\mathbf{c}_k + \mathbf{Q}_k^{(\mathrm{LR})H}\widetilde{\mathbf{n}}_k.\end{aligned} \tag{29}$$

Let $r_{k_{i,j}}^{(\mathrm{LR})}$ be the entry at the ith row, jth column of $\mathbf{R}_k^{(\mathrm{LR})}$ and c_{k_i} and $v_{k_i}^{(\mathrm{LR})}$, are respectively, the ith entries of $\mathbf{c}_k$ and $\mathbf{v}_k^{(\mathrm{LR})}$. Applying SIC detection technique, the last entry of $\widehat{\mathbf{c}}_k$ (i.e., $\widehat{\mathbf{c}}_{k_L}$) is determined as

$$\widehat{\mathbf{c}}_{k_L} = \frac{1}{\overline{\alpha}}\left(\lfloor\overline{\alpha}\widetilde{c}_{k_L} + \overline{\beta}\mathbf{t}_{k_L}1_L\rfloor - \overline{\beta}\mathbf{t}_{k_L}1_L\right), \tag{30}$$

where $\widetilde{\mathbf{c}}_{k_L} = v_{k_L}^{(\mathrm{LR})}/r_{k_{L,L}}^{(\mathrm{LR})}$ and $\mathbf{t}_{k_L}$ is Lth row of $\mathbf{T}_k^{-1}$. Once $\widehat{\mathbf{c}}_{k_L}$ is determined, the $(L-1)$th entry of $\widehat{\mathbf{c}}_k$ will be estimated by canceling the interference of $\widehat{\mathbf{c}}_{k_L}$ out of $v_{k_{L-1}}^{(\mathrm{LR})}$. In general, The ith entry of $\widetilde{\mathbf{c}}_k$ ($\widetilde{\mathbf{c}}_{k_i}$, $i = 1, 2, \ldots, L-1$) is determined as

$$\widetilde{\mathbf{c}}_{k_i} = \left(v_{k_i}^{(\mathrm{LR})} - \sum_{j=i+1}^{L} r_{k_{i,j}}^{(\mathrm{LR})}\widehat{\mathbf{c}}_{k_j}\right)\left(r_{k_{i,i}}^{(\mathrm{LR})}\right)^{-1}, \tag{31}$$

and its hard-decision value is given by

$$\widehat{\mathbf{c}}_{k_i} = \frac{1}{\overline{\alpha}}\left(\lfloor\overline{\alpha}\widetilde{\mathbf{c}}_{k_i} + \overline{\beta}\mathbf{t}_{k_i}1_L\rfloor - \overline{\beta}\mathbf{t}_{k_i}1_L\right). \tag{32}$$

This process is repeated until all entries of $\widehat{\mathbf{c}}_k$ are obtained.

6. Complexity Analysis

In this section, we evaluate the complexity of the classical linear detectors, the QRD, the BLAST, and all the proposed ones. In order to do so, we follow the method in [11, 18, 19] to count the number of necessary floating point operations (flops) to estimate a transmitted signal vector. It is assumed that each real arithmetic operation such as a real addition, a real subtraction, a real multiplication, a real division, or a square root of a real number is counted as a flop. It is worth noting that a multiplication of two matrices with dimensions of $c \times d$ and $d \times e$ needs a total of $ce\,d$ multiplications and $ce\,(d-1)$ additions. In addition, the inversion of $c \times c$ matrix requires c^3 multiplications and c^3 additions [20]. Furthermore, we would like to note that the complexity of a detector can be computed using either the complex or the equivalent real system. However, as shown in [21], the real system requires more flops than the complex one.

Under the above assumptions, the complexity of the classical ZF, the MMSE, the QRD, and the BLAST detectors in terms of flops is computed directly on the complex system as follows:

$$C_{\mathrm{ZF}} = 8N^3 + 16N^2N_r - 2N^2 + 6NN_r - 2N, \tag{33}$$

$$C_{\mathrm{MMSE}} = 8N^3 + 16N^2N_r - 2N^2 + 6NN_r, \tag{34}$$

$$C_{\mathrm{QRD}} = 6N^2N_r + 3N^2 + 12NN_r + 4N, \tag{35}$$

$$C_{\mathrm{BLAST}} = \frac{15}{4}N^4 + 2N^3N_r + N^2N_r^2 + N\left(16N_r - 2\right). \tag{36}$$

Because the proposed detectors are built based on the extended form of the massive MIMO system with the channel matrix $\mathbf{U}_{\mathrm{ex}}$ including both real and complex values, to evaluate the complexity more exactly, we represent $\mathbf{U}_{\mathrm{ex}}$ as follows:

$$\mathbf{U}_{\mathrm{ex}} = \begin{bmatrix}\mathbf{G}_1 & \mathbf{G}_2\end{bmatrix} = \begin{bmatrix}\mathbf{F}_1 & \mathbf{F}_2 \\ \mathbf{D}_1 & \mathbf{0}_2 \\ \mathbf{0}_1 & \mathbf{D}_2\end{bmatrix}, \tag{37}$$

where both $\mathbf{F}_1$ and $\mathbf{F}_2$ are $N_r \times L$ complex matrices; $\mathbf{D}_1$ and $\mathbf{D}_2$ are $L \times L$ diagonal real matrices; and $\mathbf{0}_1$ and $\mathbf{0}_2$ denote $L \times L$ zeros matrices, $L = N/2$.

6.1. Complexity Evaluation of the ZF-PGD and QRD-PGD Detectors. It can be seen that the computational costs of the ZF-PGD and the QRD-PGD detectors comprised of two parts: (1) generating the two subsystems and (2) utilizing detection procedures to detect $\widehat{\mathbf{s}}_1$ and $\widehat{\mathbf{s}}_2$. Because the subsystems in the PGD algorithm are created exactly the same way and their dimensions are equal to each other, the complexity of the ZF-PGD and QRD-PGD detectors can be computed by using the following equation:

$$C_{\mathrm{subD-PGD}} = 2\left(C_{\mathrm{Ge}} + C_{\mathrm{subD}}\right), \tag{38}$$

where $C_{\mathrm{subD-PGD}}$ is the total complexity of the ZF-PGD/QRD-PGD, C_{Ge} denotes the computational cost of generating a subsystem, and C_{subD} is number of flops given by applying the classical D detector (D is either the ZF or the QRD) to each subsystem. From the PGD algorithm, C_{Ge} is evaluated as

$$\begin{aligned} C_{\mathrm{Ge}} &= C_{\mathbf{P}_1} + C_{\widetilde{G}_1} + C_{\widetilde{y}_1} \\ &= N^3 + 4N^2N_r + 4a^2N + N - NN_r + 4aNN_r \\ &\quad - 2aN_r + 8aN - a\,(\text{flops}), \end{aligned} \tag{39}$$

where $C_{\mathbf{P}_1}$, $C_{\widetilde{G}_1}$, and $C_{\widetilde{y}_1}$ are, respectively, the numbers of flops to compute the projection term $\mathbf{P}_1$, the sub-channel matrix $\widetilde{\mathbf{G}}_1$, and the received vector $\widetilde{\mathbf{y}}_1$; $a = N_r + N$.

Because the dimensions of the subsystems are $a \times L$, C_{subD} can be obtained based on (33) and (35) by, respectively, replacing N_r and N with $a = N_r + N$ and $L = N/2$ as follows:

$$C_{\mathrm{subZF}} = N^3 + 4aN^2 - \frac{1}{2}N^2 + 3aN - N\,(\text{flops}), \tag{40}$$

$$C_{\mathrm{subQRD}} = \frac{3}{2}aN^2 + \frac{3}{4}N^2 + 6aN + 2N\,(\text{flops}). \tag{41}$$

6.2. Complexity Evaluation of the ZF-PGD-SLB and the QRD-PGD-SLB Detectors. From the signal detection procedures of the ZF-PGD-SLB and the QRD-PGD-SLB, we can easily observe that the complexity of these detectors are different from those of their ZF-PGD and the QRD-PGD counterparts by three aspects: (1) converting the subsystems to the ones in the LR domain; (2) utilizing hard decision to estimate transmitted symbols as in (23); and (3) converting the estimated vector $\widehat{\mathbf{c}}_k$ in the LR domain to $\widehat{\mathbf{s}}_k$ in the transmit constellation. Besides, the dimensions of subsystems in the PGD and the ones in the LR domain are exactly the same. Therefore, the computational complexity of the ZF-PGD-SLB and the QRD-PGD-SLB detectors can be determined based on the following equation:

$$C_{\mathrm{subD-PGD-SLB}} = C_{\mathrm{subD-PGD}} + 2C_{\mathrm{SLB}} + 2C_{\mathrm{Sli}} + 2C_{\mathrm{conv}}, \tag{42}$$

where $C_{\mathrm{subD-PGD-SLB}}$ denotes the complexity of either the ZF-PGD-SLB or the QRD-PGD-SLB detector, C_{SLB} is the number of flops to covert one subsystem to the one in the LR domain (e.g., convert $\widetilde{\mathbf{G}}_1$ to $\widetilde{\mathbf{G}}_1^{(\mathrm{LR})}$), C_{Sli} and C_{conv} are, respectively, the total complexity of hard decision operation in (23) and of converting $\widehat{\mathbf{c}}_k$ to $\widehat{\mathbf{s}}_k$ in (24).

Following Algorithm 2 step by step, C_{SLB} can be evaluated to be

$$\begin{aligned} C_{\mathrm{SLB}} &= 16L^3 + 16aL^2 - 2L^2 - 2aL + 4C_\lambda + 10C_\Delta \\ &\quad + 24LC_{\mathrm{update}} \\ &= 2N^3 + 4aN^2 - \frac{1}{2}N^2 - aN + 4C_\lambda + 10C_\Delta \\ &\quad + 12NC_{\mathrm{update}}\,(\text{flops}), \end{aligned} \tag{43}$$

where C_λ, C_Δ, and C_{update} are, respectively, the average number of times needed to compute $\lambda_{i,j}$ and $\Delta_{i,j}$ and to update Φ and $\mathbf{T}'$ so that $\widetilde{\mathbf{G}}_1$ and $\widetilde{\mathbf{T}}_1$ are generated successfully.

From (23) and (24), the remaining two terms in (42) can be evaluated to be

$$C_{\mathrm{Sli}} = 2L^2 + 4L = \frac{N^2}{2} + 2N\,(\text{flops}), \tag{44}$$

$$C_{\mathrm{conv}} = 3L^2 - L = \frac{3}{4}N^2 - \frac{N}{2}\,(\text{flops}). \tag{45}$$

Note that the complexity of $\mathbf{T}_k^{-1}$ in (23) is already included in the SLB algorithm.

The computational complexities of all the detectors are summarized in Table 1. The results in Table 1 show that all proposed detectors have the same complexity order as those of linear detectors (i.e., $\mathbb{O}(N^3)$).

The computational costs of all the detectors with different antenna's configurations are further illustrated in Figures 5 and 6.

It can be observed from Figures 5 and 6 that the complexities of the detectors increase proportionally with the number of antennas. Although the computational costs of the proposed detectors are higher than those of the classical MMSE and the ZF-SLB, they are very much smaller than those of the BLAST one in high-load systems. It is worth emphasizing that the higher complexities of our proposed detectors than those of the MMSE/ZF-SLB are rewarded by huge SNR gains as illustrated in the following section.

It can also be seen from Figure 6 that when the load factor reduces from $\beta = 0.8$ to $\beta = 0.27$ (or equivalently, the number of data streams reduces from 96 to 32), the gaps between the complexities of the detectors under consideration get much smaller. This implies that the computational costs of proposed detectors and the BLAST one become comparable in very low load systems. Therefore, our proposed detectors are only suitable to use in the systems with sufficiently high load factors.

7. Simulation Results

In this section, the BER performances of all the proposed detectors as well as the classical linear detectors, the QRD, and the MMSE-BLAST (in the paper, the detector is referred to as the BLAST for short) ones are investigated. The proposed detectors are also compared with the ZF-SLB one proposed in [13] regarding the BER performance. In the simulation, the channel is assumed to be block fading, which is unchanged within the signal block and changed from one block to another. The main simulation parameters and assumptions are set up as follows: the cell radius, r, and reference distance, d_0, are 1000 meters and 100 meters, respectively; all active users are uniformly distributed in the cell such that the distance from the ith user to the BS satisfying $100\,\mathrm{m} \le d_i \le 990\,\mathrm{m}$; the large-scale fading coefficients are generated according to (3), where $\sigma^2_{\mathrm{Shadow}} = 8$ dB and $\gamma = 3.5$. All BER curves are drawn versus SNR defined by p_u/σ^2 in dB.

Figures 7 and 8 illustrate the BER curves of the aforementioned detectors for 2 different system configurations: (1) $N_r = 64$, $K = 16$, and $N_T = 4$ and (2) $N_r = 128$, $K = 32$, and $N_T = 4$ antennas. The results in both figures show that

Input: $\widetilde{\mathbf{G}}_k$
Output: $\widetilde{\mathbf{G}}_k^{(\mathrm{LR})}$, $\mathbf{T}_k$
(1) Set $\widetilde{\mathbf{U}} = \widetilde{\mathbf{G}}_k$; $\mathbf{T}' = \mathbf{I}_N$; and compute $\Phi = (\widetilde{\mathbf{U}}^H \widetilde{\mathbf{U}})^{-1}$.
(2) **Do**
(3) Find j index such that $\Phi_{j,j}$ is maximum reducible value of Φ.
(4) $\lambda_{i,k} = -\Phi_{i,j}/\Phi_{i,i}$, $\forall\, i \neq j$
(5) Compute $\Delta_{i,j} = -|\lambda_{i,j}|^2 \Phi_{i,i} - \lambda_{i,j}^* \Phi_{i,j} - \lambda_{i,j} \Phi_{i,j}^*$ and chose $i = \mathrm{argmax}_{i=1:L; i\neq j} \Delta_{i,j}$
(6) If $\Delta_{i,j} = 0$, $\forall i, j = 1 : L$ go to (10)
(7) $\mathbf{t}'_j = \mathbf{t}'_j + \lambda_{i,j} \mathbf{t}'_i$; % $\mathbf{t}'_j$ is jth column of $\mathbf{T}'$
(8) $\phi_j = \phi_j + \lambda_{i,j} \phi_i$; % ϕ_j is jth column of Φ
(9) $\phi^j = \phi^j + \lambda_{i,j} \phi^i$; % ϕ^j is jth row of Φ
(10) **While** (true)
(11) $\mathbf{T}_k = (\mathbf{T}'^{-1})^H$; $\widetilde{\mathbf{G}}_1^{(\mathrm{LR})} = \widetilde{\mathbf{U}} \mathbf{T}_k$

Algorithm 2: SLB algorithm.

Table 1: Computational complexity comparison.

Detectors	Number of necessary flops to recover a transmitted signal vector
ZF	$8N^3 + 16N^2N_r - 2N^2 + 6NN_r - 2N$
MMSE	$8N^3 + 16N^2N_r - 2N^2 + 6NN_r$
QRD	$6N^2N_r + 3N^2 + 12NN_r + 4N$
BLAST	$(15/4)N^4 + 2N^3N_r + N^2N_r^2 + N(16N_r - 2)$
ZF-PGD	$4N^3 + 8N^2N_r + 8aN^2 - N^2 + 8a^2N - 2NN_r + 8aNN_r - 4aN_r + 22aN - 2a$
QRD-PGD	$2N^3 + 8N^2N_r + 8a^2N + 3aN^2 + (3/2)N^2 + 6N - 2NN_r + 8aNN_r - 4aN_r + 28aN - 2a$
ZF-SLB	$24N^3 + 32N^2N_r + 20N^2 + 4NN_r + 1 + 4C_{\lambda_1} + 10C_{\Delta_1} + 24NC_{\mathrm{update1}}$
ZF-PGD-SLB	$8N^3 + 8N^2N_r + 16aN^2 + (N^2/4) + 8a^2N - 2NN_r + 8aNN_r - 4aN_r + 20aN - 2a + (7/2)N + 8C_\lambda + 20C_\Delta + 24NC_{\mathrm{update}}$
QRD-PGD-SLB	$6N^3 + 8N^2N_r + 11aN^2 + (9/4)N^2 + 8a^2N + (19/2)N - 2NN_r + 8aNN_r - 4aN_r + 26aN - 2a + 8C_\lambda + 20C_\Delta + 24NC_{\mathrm{update}}$

Note. $a = N_r + N, N = KN_T$; $C_\lambda, C_\Delta, C_{\mathrm{update}}, C_{\lambda_1}, C_{\Delta_1}$, and C_{update1} are given by simulations.

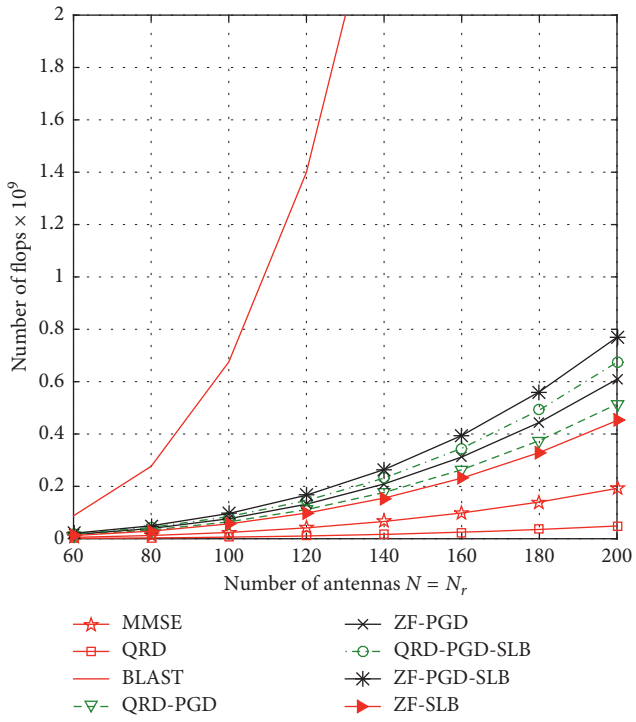

Figure 5: Complexity curves of the proposed detectors as well as the classical ZF, the MMSE, the QRD, and the BLAST detectors when $N_r = N = [60 : 20 : 200]$.

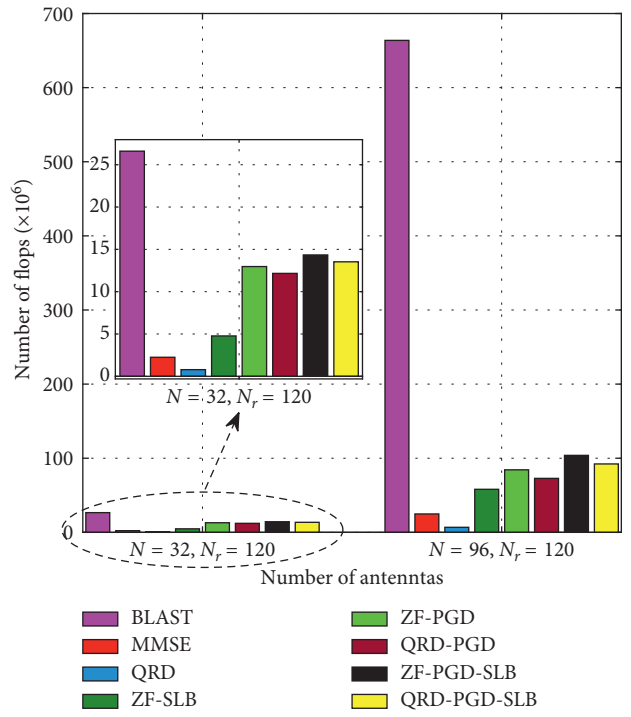

Figure 6: Complexities of the proposed detectors as well as of the classical ZF, the MMSE, the QRD, and the BLAST detectors when $N_r = 120$; $K = 8, 24$; $N_T = 4$; and 64-QAM.

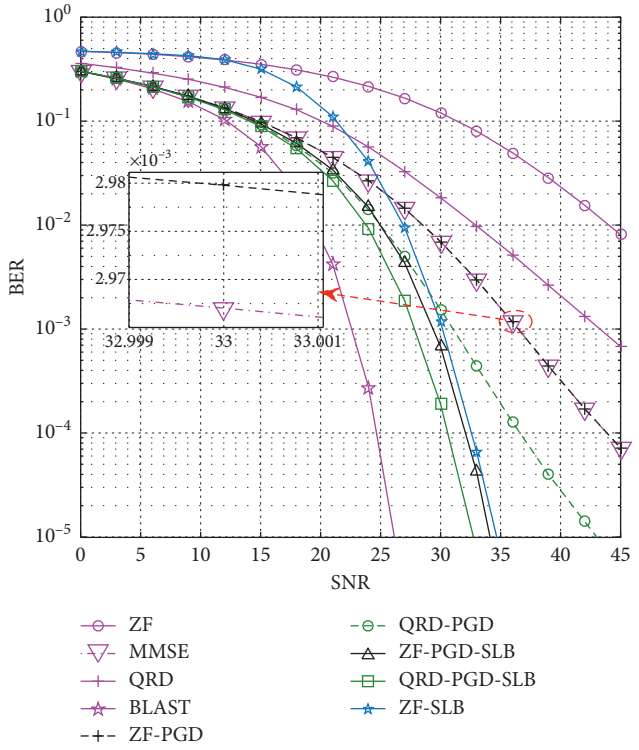

Figure 7: BER curves of the proposed detectors as well as of the ZF, the MMSE, the QRD, the BLAST, and the ZF-SLB ones when $N_r = 64, K = 16, N_T = 4$, and 4-QAM.

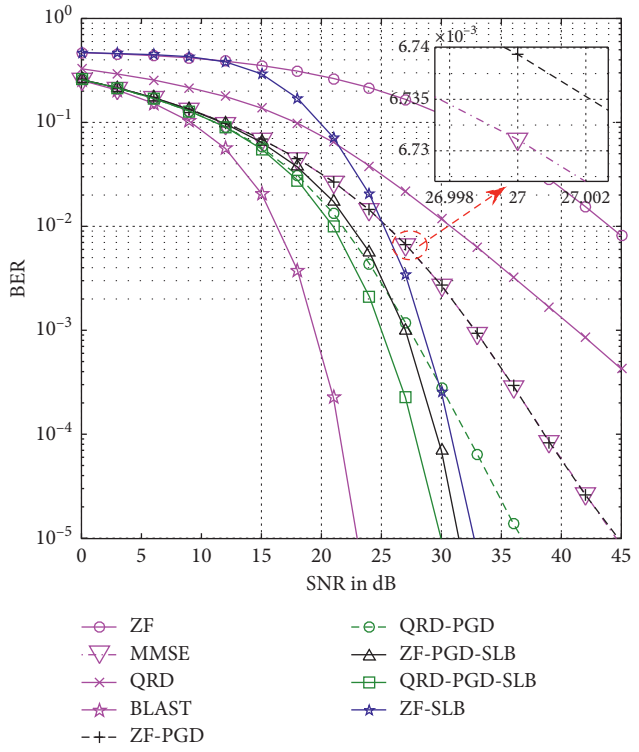

Figure 8: BER curves of the proposed detectors as well as of the ZF, the MMSE, the QRD, the BLAST, and the ZF-SLB ones when $N_r = 128, K = 32, N_T = 4$, and 4-QAM.

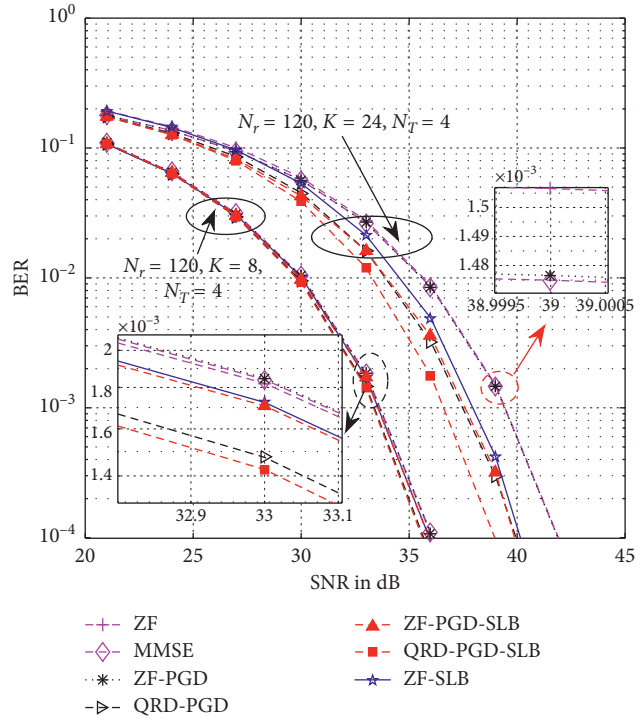

Figure 9: BER curves of the proposed detectors as well as the ZF, the MMSE, and the ZF-SLB ones when $N_r = 120$; $K = 8, 24$; $N_T = 4$; and 64-QAM.

all the proposed detectors significantly outperform their classical ZF and the QRD counterparts. In addition, the ZF-PGD-SLB and the QRD-PGD-SLB detectors provide much better BER performance than the ZF-SLB one. More specifically, BER performance of the ZF-PGD is much better than that of the classical ZF and even comparable to the MMSE detector's performance (it is well understood that the MMSE detector can be implemented by applying the ZF detector to the extended system defined in equation (5); therefore, the ZF-PGD is able to approach the performance of the MMSE detector, and due to the effects of the projection terms $\mathbf{P}_1$ and $\mathbf{P}_2$, when generating the two subsystems, the ZF-PGD slightly underperforms than its MMSE counterpart). This result confirms our analysis in Section 3. The illustrations in both these figures show that there are big gaps between the BER curve of the MMSE and those of the remaining proposed detectors. For example, for the first antenna configuration, i.e., $N_r = 64$, $K = 16$, $N_T = 4$, the gaps between the MMSE curve and the ones of the QRD-PGD, the ZF-PGD-SLB, and the QRD-PGD-SLB are about 7.4 dB, 11.9 dB, and 13.4 dB at BER= 10^{-4}, respectively. These gaps reduce to 6.52 dB, 9.02 dB, and 10.82 dB for $N_r = 128$, $K = 32$, and $N_T = 4$. These results imply that in full-load scenarios, as the system size increases, the BER performance gaps between our proposed detectors and those of the MMSE get smaller. This is probably due to the fact that the columns of the channel matrix becomes more orthogonal in pairs when the system size increases, thereby reducing the effect of cochannel interference at the BS.

It is worth noting that the BER performance of the ZF detector is worse than that of the MMSE even if the SNR is very high. The reason is that there always exists a significant gap between the output SNRs of these two detectors when the SNR reaches to infinity [6]. The results in Figures 7 and 8 also show that the BLAST detector gives the best BER performance among all the detectors under consideration. However, its excessively high complexity prevents it from being practically adopted in massive MIMO systems.

Shown in Figure 9 are the BER curves of the detectors under consideration when the load factors of the system are $\beta_1 = 0.27$ and $\beta_2 = 0.8$. The system parameters are set as follows: $N_r = 120$ antennas, $K = 8, 24$, and $N_T = 4$ antennas, leading to the equivalent dimensions as 120×32 and 120×96. It can be seen from Figure 9 that when the load factor of the system reduces, the BER performances of all considered detectors improve significantly. Besides, the SNR gains between our proposed detectors and the classical ones get much smaller when the load factor reduces. Specifically, when $\beta_2 = 0.8$, the ZF-PGD, the ZF, and MMSE detectors have nearly the same BER performance, but they underperform than the ZF-SLB, the proposed ZF-PGD-SLB, and QRD-PGD-SLB detectors by about 2 dB, 2.2 dB, and 3.2 dB at $\text{BER} = 10^{-4}$, respectively. When the load factor reduces to $\beta_1 = 0.27$, performances of all considered detectors are just slightly different from each others. This implies that when the system's load factor is sufficiently small, the proposed detectors are no longer more advantageous than the classical linear detectors with respect to the BER performance. If the detection complexity is taken into account (as illustrated in Figure 6), then the classical ZF and MMSE detectors are the most efficient for signal recovery in massive MIMO systems with low load factors.

8. Conclusion

In this paper, the parallel group detection algorithm has been proposed to enhance BER performance of high-load massive MIMO systems. By combining the PGD, the classical ZF, and the QRD detectors, we are able to construct the ZF-PGD and QRD-PGD detectors. Moreover, the PGD is further combined with shortest longest basis (SLB) algorithm to generate two other new detectors, called ZF-PGD-SLB and QRD-PGD-SLB. Various simulation and analytical results show that the proposed detectors remarkably outperform their classical linear detectors and the QRD counterparts, yet at the cost of higher detection complexities. Fortunately, the complexities of the proposed detectors are very much smaller than those of the BLAST in high-load scenario. Consequently, our proposed detectors are suitable for signal detection in high-load massive MIMO systems.

Data Availability

All data included in this paper are the results of our research and are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] H. Q. Ngo, *Massive MIMO: Fundamentals and System Designs*, Vol. 1642, Linköping University Electronic Press, Linköping, Sweden, 2015.

[2] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*, Cambridge University Press, Cambridge, UK, 2016.

[3] C. Shepard, J. Ding, R. E. Guerra, and L. Zhong, "Understanding real many-antenna mu-MIMO channels," in *Proceedings of the 2016 50th Asilomar Conference on Signals, Systems and Computers*, pp. 461–467, Pacific Grove, CA, USA, November 2016.

[4] H. Q. Ngo, M. Matthaiou, T. Q. Duong, and E. G. Larsson, "Uplink performance analysis of multicell mu-simo systems with zf receivers," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 9, pp. 4471–4483, 2013.

[5] Y. Li, Y.-H. Nam, and B. L. Ng, "Zero-forcing receiver in uplink massive mimo," in *Proceedings of the 2013 IEEE Globecom Workshops (GC Wkshps)*, pp. 140–144, IEEE, Atlanta, GA, USA, December 2013.

[6] Y. Jiang, M. K. Varanasi, and J. Li, "Performance analysis of ZF and mmse equalizers for MIMO systems: an in-depth study of the high SNR regime," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 2008–2026, 2011.

[7] T. Li and M. Torlak, "Performance of ZF linear equalizers for single carrier massive MIMO uplink systems," *IEEE Access*, vol. 6, pp. 32156–32172, 2018.

[8] T. Liu, J. Tong, Q. Guo, J. Xi, Y. Yu, and Z. Xiao, "Energy efficiency of uplink massive mimo systems with successive interference cancellation," *IEEE Communications Letters*, vol. 21, no. 3, pp. 668–671, 2017.

[9] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: ten myths and one critical question," *IEEE Communications Magazine*, vol. 54, no. 2, pp. 114–123, 2016.

[10] J. W. Choi and B. Shim, "New approach for massive MIMO detection using sparse error recovery," in *Proceedings of the 2014 IEEE Global Communications Conference (GLOBECOM)*, pp. 3754–3759, IEEE, Austin, TX, USA, December 2014.

[11] T.-B. Nguyen, T.-D. Nguyen, M.-T. Le, and V.-D. Ngo, "Efficiency zero-forcing detectors based on group detection for massive mimo systems," in *Procedings of the 2017 International Conference on Advanced Technologies for Communications (ATC)*, pp. 48–53, Quy Nhon, Vietnam, October 2017.

[12] T.-B. Nguyen, M.-T. Le, V.-D. Ngo, and V.-G. Nguyen, "Parallel group detection approach for massive MIMO systems," in *Proceedings of the 2018 International Conference on Advanced Technologies for Communications (ATC)*, pp. 94–99, IEEE, Ho Chi Minh City, Vietnam, October 2018.

[13] Q. Zhou and X. Ma, "Element-based lattice reduction algorithms for large MIMO detection," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 274–286, 2013.

[14] T. Li, S. Patole, and M. Torlak, "A multistage linear receiver approach for MMSE detection in massive MIMO," in *Proceedings of the 2014 48th Asilomar Conference on Signals, Systems and Computers*, pp. 2067–2072, IEEE, Pacific Grove, CA, USA, November 2014.

[15] L. Bai and J. Choi, *Low Complexity MIMO Detection*, Springer Science & Business Media, Berlin, Germany, 2012.

[16] A. K. Lenstra, H. W. Lenstra, and L. Lovász, "Factoring polynomials with rational coefficients," *Mathematische Annalen*, vol. 261, no. 4, pp. 515–534, 1982.

[17] M. Seysen, "Simultaneous reduction of a lattice basis and its reciprocal basis," *Combinatorica*, vol. 13, no. 3, pp. 363–376, 1993.

[18] T.-D. Nguyen, X.-N. Tran, T.-M. Do, V.-D. Ngo, and M.-T. Le, "Low-complexity detectors for high-rate spatial modulation," in *Proceddings of the 2014 International Conference on Advanced Technologies for Communications (ATC 2014)*, pp. 652–656, IEEE, Hanoi, Vietnam, October 2014.

[19] X. N. Tran, H. C. Ho, T. Fujino, and Y. Karasawa, "Performance comparison of detection methods for combined stbc and sm systems," *IEICE Transactions on Communications*, vol. 91, no. 6, pp. 1734–1742, 2008.

[20] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Vol. 3, JHU Press, Baltimore, MD, USA, 2012.

[21] R. T. Kobayashi, F. Ciriaco, and T. Abrão, "Efficient near-optimum detectors for large MIMO systems under correlated channels," *Wireless Personal Communications*, vol. 83, no. 2, pp. 1287–1311, 2015.