Efficient QAM Signal Detector for Massive MIMO Systems via PS/DPS-ADMM Approaches

Quan Zhang, Jiangtao Wang, and Yongchao Wang, *Senior Member, IEEE*

***Abstract*—In this paper, we design two efficient quadrature amplitude modulation (QAM) signal detectors for massive multiple-input multiple-output (MIMO) communication systems via the penalty-sharing alternating direction method of multipliers (PS-ADMM). The content of the paper is summarized as follows: first, we transform the maximum-likelihood detection model to a non-convex sharing optimization problem for massive MIMO-QAM systems, where a high-order QAM constellation is decomposed to a sum of multiple binary variables, integer constraints are relaxed to box constraints, and quadratic penalty functions are added to the objective function to result in a favorable integer solution; second, a customized ADMM algorithm, called PS-ADMM, is presented to solve the formulated non-convex optimization problem. In the implementation, all variables in each vector can be solved analytically and in parallel; and third, in order to solve the penalty-sharing distributively, we improve the proposed PS-ADMM algorithm to a distributed one, named DPS-ADMM. In the end, performance analyses of the proposed two algorithms, including convergence properties and computational cost, are provided. Simulation results demonstrate the effectiveness of the proposed approaches.**

***Index Terms*—Massive MIMO, maximum-likelihood detection, penalty method, sharing-ADMM, non-convex optimization.**

I. Introduction

MASSIVE multiple-input multiple-output (MIMO) technology is considered to be one of the disruptive technologies for the fifth-generation (5G) communication systems [1], [2]. The foremost benefit of massive MIMO is significant increase in the spatial degrees of freedom which can improve throughput and energy efficiency significantly in comparison with conventional MIMO systems [3]. However, numerous practical challenges arise in implementing massive MIMO technology in order to achieve such improvements. One such challenge is signal detection lying in the uplink massive systems since it is difficult to achieve an effective compromise among good detecting performance, low computational complexity, and high processing parallelism [4].

MIMO detectors can be divided into two categories: linear and nonlinear.

A. Linear Detectors

Minimum mean square error (MMSE) [5] and zero-forcing (ZF) [6], are the most classical linear detectors since both of them can achieve good detection performance when the ratio of the base station's antenna number to the user number (BS-to-user-antenna ratio) is far larger than one [7]. One main concern of the MMSE or ZF detector is that the matrix inversion operation is involved in the detection procedure. To overcome this challenge, several techniques based on the Neumann series (NS) [8], [9], Gauss-Seidel (GS) [10], [11], Richardson (RI) [12], and Conjugate gradient (CG) [13] are proposed. However, the decrease in computational complexity of these algorithms comes at the expense of the loss of detection performance; meanwhile, these methods deliver poor bit error rate (BER) performance when the BS-to-user antenna ratio is close to one.

B. Nonlinear Detectors

The most well-known nonlinear detector is the maximum-likelihood (ML) detector [14], which can obtain optimal detection performance but suffers from exponentially increasing computational complexity with the number of terminal antennas. Sphere decoding (SD) detector [15], K-best detector [16], and semidefinite relaxation (SDR) detector [17] are also well-studied. Existing research works show that they can achieve near-optimal ML detection performance with a polynomial computational complexity in high signal-to-noise (SNR) region or small-scale MIMO system. However, their computational complexity is still unaffordable in practice especially for massive MIMO systems [18]. In [19], the triangular approximate semidefinite relaxation (TASER) detector achieves near-ML performance while providing comparable hardware-efficiency, but which cannot be applied to a high-order QAM MIMO system. Besides the above referred detectors, different techniques, such as local search [20], [21], belief propagation [22], box-constrained [23], sparsity [24], and machine learning [25] are leveraged to design detectors for Massive MIMO detection. However, various defects, such as computational complexity, detection performance, or implement efficiency, limit their applications in practice.

Manuscript received 9 April 2021; revised 18 August 2021, 9 December 2021, and 7 March 2022; accepted 12 April 2022. Date of publication 5 May 2022; date of current version 11 October 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 62001349 and Grant 61771356. This article is to be presented at the 2022 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Singapore. The associate editor coordinating the review of this article and approving it for publication was M. Wang. *(Corresponding authors: Jiangtao Wang; Yongchao Wang.)*

The authors are with the State Key Laboratory of Integrated Services Network (ISN), Xidian University, Xi'an 710071, China (e-mail: qzhang_6@stu.xidian.edu.cn; jtwang@xidian.edu.cn; ychwang@mail.xidian.edu.cn).

This article has supplementary material provided by the authors and color versions of one or more figures available at https://doi.org/10.1109/TWC.2022.3170510.

Digital Object Identifier 10.1109/TWC.2022.3170510

In recent years, the alternating direction method of multipliers (ADMM) technique is widely used to solve convex and non-convex problems due to its simplicity, operator splitting capability, and guaranteed-convergence under mild conditions [26]. The ADMM strategy was first introduced in MIMO detection by Takapoui *et al.* in [27]. In [28], a detection algorithm based on ADMM with infinity norm or box-constrained equalization, named ADMIN, was proposed, which outperformed the state-of-the-art linear detectors by a large margin if the BS-to-user-antenna ratio is small. The ADMM detection algorithm was applied in various scenarios for MIMO systems [29], [30] and improved for massive MIMO systems [31]–[35]. Although the existing ADMM-based methods provide better BER performance than conventional detectors, the BER performance could be improved further since constraints set of optimization problem is over-relaxed.

In this paper, we focus on designing ADMM-based quadrature amplitude modulation (QAM) signal detectors for massive MIMO systems. By exploiting ideas of penalized bound relaxation for the ML detection formulation [36], binary transformation for high-order QAM signals [37], [38], and (distributed) sharing-ADMM techniques [26], [39], we propose two new detectors for massive MIMO system, called PS-ADMM and DPS-ADMM, respectively. Both of them have favorable BER performance and cheap computational cost. The main contributions of this paper are summarized as follows:

- *Penalty sharing formulation:* we relax the well-known MIMO detection ML model to a non-convex sharing problem. In it, high-order modulation signals are decomposed into a sum of multiple binary ones and penalty functions involving these binaries are introduced in the sharing formulation.
- *Efficient implementation:* in the proposed PS-ADMM detectors, all the variables in each subproblem can be solved analytically in each iteration step. In comparison with PS-ADMM, the proposed DPS-ADMM detector can solve the penalty-sharing problem in a distributed manner, i.e., variables in the subproblems can be solved in parallel, which is favorable from a practical viewpoint.
- *Theoretically-guaranteed convergence:* we prove that the proposed PS/DPS-ADMM algorithms are convergent to a stationary point of the non-convex optimization problem if proper parameters are chosen.

Moreover, simulation results demonstrate that both of the proposed algorithms can offer attractive tradeoffs between BER performance and computational cost, especially when BS-to-user-antenna ratio is close to one for large-scale MIMO-QAM systems.

The rest of this paper is organized as follows. In Section II, we formulate the massive MIMO detection problem to a non-convex sharing problem. In Section III, efficient PS/DPS-ADMM algorithms are customized to solve the formulated problem. Section IV presents the detailed performance analysis, including convergence, iteration complexity, and computational cost of the proposed PS/DPS-ADMM algorithms. Simulation results, which show the effectiveness of our proposed PS/DPS-ADMM algorithms, are presented in Section V and Section VI concludes this paper.

Notations: In this paper, bold lowercase, uppercase, and italic letters denote column vectors, matrices, and scalars respectively; $\mathbb{C}$ denotes the complex field; $(\cdot)^H$ and $\|\cdot\|_2$ symbolize the conjugate transpose and 2-norm operations respectively; x_{iR}, x_{iR} denote the real and imaginary parts of the ith entry of vector $\mathbf{x}$ respectively; $\prod\limits_{[a,b]}(\cdot)$ denotes the Euclidean projection operator onto interval $[a, b]$; $\nabla(\cdot)$ represents the gradient of a function; $\mathrm{Re}(\cdot)$ takes the real part of the complex variable; $\lambda_{\min}(\cdot)$ and $\lambda_{\max}(\cdot)$ denote the minimum and maximum eigenvalues of a matrix respectively; $\langle\cdot,\cdot\rangle$ denotes the dot product operator; and $\mathbf{I}$ denotes an identity matrix.

II. System Model and Problem Formulation

The considered signal detection problem lies in uplink multiuser massive MIMO systems, where BS equipped with B antennas serves U single-antenna users. Here, we assume $B \geq U$. Typically, the received signal at BS can be characterized by the following model

$$\mathbf{r} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{1}$$

where $\mathbf{x} \in \mathbb{X}^U$ is the transmitted signal from U users and $\mathbb{X}$ refers to the signal constellation set, $\mathbf{r} \in \mathbb{C}^B$ is the BS received signal vector, $\mathbf{H} \in \mathbb{C}^{B\times U}$ denotes the MIMO channel matrix, and $\mathbf{n} \in \mathbb{C}^B$ denotes additive white Gaussian noise. The entries of $\mathbf{H}$ and $\mathbf{n}$ are assumed to be independent and identically distributed (i.i.d.) complex Gaussian variables with zero mean.

The optimal MIMO ML detector for 4^Q-QAM signals, i.e., achieving minimum error probability of detecting $\mathbf{x}$ from the received signal $\mathbf{r}$, can be formulated as the following discrete least square problem [17]

$$\min_{\mathbf{x}\in\mathbb{X}^U} \|\mathbf{r} - \mathbf{H}\mathbf{x}\|_2^2, \tag{2}$$

where $\mathbb{X} = \{x = x_\mathrm{R} + jx_\mathrm{I} | x_\mathrm{R}, x_\mathrm{I} \in \{\pm1, \pm3, \cdots, \pm(2^Q - 1)\}\}$ and Q is some positive integer. Obtaining its global optimal solution is prohibitive in practice since the corresponding computational complexity grows exponentially with the users' number U, BS's antenna number B, and set $\mathbb{X}$'s size [40].

In the following, by exploiting insight structures of the model (2), we relax it to the well-known sharing problem [26]. The main challenge to solve it comes from how to handle multiple quantization levels of QAM signals, which cannot be treated by general $\{1, -1\}$ binary processing techniques directly, such as GBR-SDR [36] and HOTML/DeepHOTML [41]. To tackle this issue, we decompose high-order QAM symbols into a sum of multiple binary variables. Specifically, notice transmitted signal vector $\mathbf{x} \in \mathbb{X}^U$ can be expressed by

$$\mathbf{x} = \sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q, \tag{3}$$

where $\mathbf{x}_q \in \mathbb{X}_q^U$ and $\mathbb{X}_q = \{x_q = x_{q\mathrm{R}} + jx_{q\mathrm{I}} | x_{q\mathrm{R}}, x_{q\mathrm{I}} \in \{1, -1\}\}$. Then, model (2) can be equivalent to

$$\begin{aligned}&\min_{\mathbf{x}_q} \ \frac{1}{2}\|\mathbf{r} - \mathbf{H}\sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q\|_2^2,\\ &\text{s.t. } \mathbf{x}_q \in \mathbb{X}_q^U, \quad q = 1, \cdots, Q.\end{aligned} \tag{4}$$

Relaxing binary constraints $\{1, -1\}^U$ to box constraints $[1, -1]^U$ and then adding a sum of quadratic penalty functions $-\sum_{q=1}^{Q} \frac{\alpha_q}{2}\|\mathbf{x}_q\|_2^2$ into its objective function, we formulate (4) to the following non-convex optimization model

$$\begin{aligned}&\min_{\mathbf{x}_q} \ \frac{1}{2}\|\mathbf{r} - \mathbf{H}(\sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q)\|_2^2 - \sum_{q=1}^{Q} \frac{\alpha_q}{2}\|\mathbf{x}_q\|_2^2,\\ &\text{s.t. } \mathbf{x}_q \in \tilde{\mathbb{X}}_q^U, \quad q = 1, \cdots, Q,\end{aligned} \tag{5}$$

where penalty parameters $\alpha_q \geq 0$ and $\tilde{\mathbb{X}}_q = \{x_{q\mathrm{R}} + jx_{q\mathrm{I}} | x_{q\mathrm{R}}, x_{q\mathrm{I}} \in [1, -1]\}$.

Remarks: In the statistics learning field, a problem like (5) is called as sharing problem [26], [39], where $\frac{1}{2}\|\mathbf{r} - \mathbf{H}(\sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q)\|_2^2$ is the shared objective and $-\sum_{q=1}^{Q} \frac{\alpha_q}{2}\|\mathbf{x}_q\|_2^2$ is the local cost function for the qth subsystem. It is obvious that model (5) is a relaxation of ML detection problem (2), which will intrinsically introduce some losses of detection performance. However, we should note that the introduced penalty function $-\sum_{q=1}^{Q} \frac{\alpha_q}{2}\|\mathbf{x}_q\|_2^2$ can lead to optimal integer solution more favorable and improve the detection performance.

III. Solving Algorithms

In this section, we propose two efficient algorithms based on the ADMM technique, named as PS-ADMM and DPS-ADMM. The latter can be cast as a distributive version of the former.

A. PS-ADMM Algorithm

We transform problem (5) equivalently to a linearly constrained problem by introducing auxiliary variable $\mathbf{x}_0 \in \mathbb{C}^U$

$$\begin{aligned}&\min_{\mathbf{x}_0, \mathbf{x}_q} \ \frac{1}{2}\|\mathbf{r} - \mathbf{H}\mathbf{x}_0\|_2^2 - \sum_{q=1}^{Q} \frac{\alpha_q}{2}\|\mathbf{x}_q\|_2^2,\\ &\text{s.t. } \mathbf{x}_0 = \sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q,\\ &\qquad \mathbf{x}_q \in \tilde{\mathbb{X}}_q^U, \quad q = 1, \cdots, Q,\end{aligned} \tag{6}$$

whose augmented Lagrangian can be expressed as

$$\begin{aligned}&L_\rho(\{\mathbf{x}_q\}_{q=1}^{Q}, \mathbf{x}_0, \mathbf{y})\\ &= \frac{1}{2}\|\mathbf{r} - \mathbf{H}\mathbf{x}_0\|_2^2 - \sum_{q=1}^{Q} \frac{\alpha_q}{2}\|\mathbf{x}_q\|_2^2\\ &\quad + \mathrm{Re}\Big\langle \mathbf{x}_0 - \sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q, \mathbf{y}\Big\rangle + \frac{\rho}{2}\Big\|\mathbf{x}_0 - \sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q\Big\|_2^2,\end{aligned} \tag{7}$$

where $\mathbf{y} \in \mathbb{C}^U$ and $\rho > 0$ are the Lagrangian multiplier and penalty parameter respectively. Based on the above augmented Lagrangian, the proposed PS-ADMM solving algorithm framework can be described as

$$\begin{aligned}\mathbf{x}_q^{k+1} &= \arg\min_{\mathbf{x}_q \in \tilde{\mathbb{X}}_q^U} L_\rho(\mathbf{x}_1^{k+1}, \cdots, \mathbf{x}_{q-1}^{k+1}, \mathbf{x}_q, \mathbf{x}_{q+1}^k,\\ &\qquad \cdots, \mathbf{x}_Q^k, \mathbf{x}_0^k, \mathbf{y}^k), \quad q = 1, \cdots, Q,\end{aligned} \tag{8a}$$

$$\mathbf{x}_0^{k+1} = \arg\min_{\mathbf{x}_0} \ L_\rho(\{\mathbf{x}_q^{k+1}\}_{q=1}^{Q}, \mathbf{x}_0, \mathbf{y}^k), \tag{8b}$$

$$\mathbf{y}^{k+1} = \mathbf{y}^k + \rho\Big(\mathbf{x}_0^{k+1} - \sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q^{k+1}\Big), \tag{8c}$$

where k denotes iteration number.

The main challenge of implementing (8) lies in how to solve optimization subproblems (8a) and (8b) efficiently. For (8a), it can be observed that $L_\rho(\{\mathbf{x}_q\}_{q=1}^{Q}, \mathbf{x}_0^k, \mathbf{y}^k)$ is a strongly convex quadratic function [45] with respect to (w.r.t.) $\mathbf{x}_q$ when $4^{q-1}\rho - \alpha_q > 0$. It means that, if we set $4^{q-2}\rho > \alpha_q$, the optimal solution of the subproblems (8a) can be obtained through the following procedure:

Set the gradient of the corresponding augmented Lagrangian function w.r.t. $\mathbf{x}_q$ to be zero, i.e.,

$$\nabla_{\mathbf{x}_q} L_\rho(\mathbf{x}_q, \mathbf{x}_1^{k+1}, \cdots, \mathbf{x}_{q-1}^{k+1}, \mathbf{x}_{q+1}^k, \cdots, \mathbf{x}_Q^k, \mathbf{x}_0^k, \mathbf{y}^k) = 0, \tag{9}$$

which leads to the following linear equation

$$\begin{aligned}&\nabla_{\mathbf{x}_q}\Big(-\frac{\alpha_q}{2}\|\mathbf{x}_q\|_2^2 - \mathrm{Re}\langle 2^{q-1}\mathbf{x}_q, \mathbf{y}^k\rangle\\ &\quad + \frac{\rho}{2}\|\mathbf{x}_0^k - \sum_{i<q} 2^{i-1}\mathbf{x}_i^{k+1} - \sum_{i>q} 2^{i-1}\mathbf{x}_i^k - 2^{q-1}\mathbf{x}_q\|_2^2\Big) = 0.\end{aligned} \tag{10}$$

By solving (10), we can obtain

$$\begin{aligned}\mathbf{x}_q^{k+1} &= \prod_{[-1,1]}\bigg(\frac{2^{q-1}}{4^{q-1}\rho - \alpha_q}\Big(\rho\mathbf{x}_0^k - \rho\sum_{i<q} 2^{i-1}\mathbf{x}_i^{k+1}\\ &\qquad - \rho\sum_{i>q} 2^{i-1}\mathbf{x}_i^k + \mathbf{y}^k\Big)\bigg), \quad q = 1, \cdots, Q,\end{aligned} \tag{11}$$

where operator $\prod_{[-1,1]}(\cdot)$ projects every entry's real part and imaginary part of the input vector onto [−1, 1] respectively.

Moreover, $L_\rho(\{\mathbf{x}_q^{k+1}\}_{q=1}^{Q}, \mathbf{x}_0, \mathbf{y}^k)$ is a strongly convex quadratic function w.r.t. $\mathbf{x}_0$ since $\rho > 0$ and matrix $\mathbf{H}^H\mathbf{H}$ is positive semidefinite, hence the optimal solution of the subproblem (8b) can also be obtained by setting $\nabla_{\mathbf{x}_0} L_\rho(\{\mathbf{x}_q^{k+1}\}_{q=1}^{Q}, \mathbf{x}_0, \mathbf{y}^k)$ to be zero and solving the corresponding linear equation, which results in

$$\mathbf{x}_0^{k+1} = (\mathbf{H}^H\mathbf{H} + \rho\mathbf{I})^{-1}\Big(\mathbf{H}^H\mathbf{r} + \rho\sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q^{k+1} - \mathbf{y}^k\Big). \tag{12}$$

To be clear, we summarize the proposed PS-ADMM algorithm for solving model (6) in *Algorithm 1*.

B. DPS-ADMM Algorithm

The main concern of the proposed PS-ADMM algorithm is that $\mathbf{x}_q$, $q = 1, \cdots, Q$ have to be updated in serial. For a high-order QAM signal, it could be inefficient in practice.

Algorithm 1 The Proposed PS-ADMM Algorithm

Input: $\mathbf{H}$, $\mathbf{r}$, Q, ρ, $\{\alpha_q\}_{q=1}^Q$

Output: $\mathbf{x}_0^k$

1: Initialize $\{\mathbf{x}_q^1\}_{q=1}^Q, \mathbf{x}_0^1, \mathbf{y}^1$ as the all-zeros vectors.[1]

2: **For** $k = 1, 2, \cdots$

3: Step 1: Update $\{\mathbf{x}_q^{k+1}\}_{q=1}^Q$ sequentially via (11).

4: Step 2: Update $\mathbf{x}_0^{k+1}$ via (12).

5: Step 3: Update $\mathbf{y}^{k+1}$ via (8c).

6: **Until** some preset condition is satisfied.

In the following, we develop a different algorithm, named as DPS-ADMM, which can update $\mathbf{x}_q$ in parallel.

For this goal, we let $\mathbf{z}_q = 2^{q-1}\mathbf{x}_q$, $q = 1, \cdots, Q$, and transform model (5) to

$$\begin{aligned}&\min_{\mathbf{x}_0,\mathbf{x}_q}\ \frac{1}{2}\|\mathbf{r}-\mathbf{H}\sum_{q=1}^{Q}\mathbf{z}_q\|_2^2-\sum_{q=1}^{Q}\frac{\alpha_q}{2}\|\mathbf{x}_q\|_2^2,\\ &\text{s.t. } \mathbf{z}_q = 2^{q-1}\mathbf{x}_q,\\ &\qquad \mathbf{x}_q \in \tilde{\mathbb{X}}_q^U, \quad q = 1, \cdots, Q, \end{aligned} \tag{13}$$

whose augmented Lagrangian function is

$$\begin{aligned}&L_\rho(\{\mathbf{x}_q, \mathbf{z}_q, \mathbf{y}_q\}_{q=1}^Q)\\ &= \frac{1}{2}\|\mathbf{r}-\mathbf{H}\sum_{q=1}^{Q}\mathbf{z}_q\|_2^2-\sum_{q=1}^{Q}\frac{\alpha_q}{2}\|\mathbf{x}_q\|_2^2\\ &\quad + \sum_{q=1}^{Q}\mathrm{Re}\langle 2^{q-1}\mathbf{x}_q-\mathbf{z}_q, \mathbf{y}_q\rangle + \sum_{q=1}^{Q}\frac{\rho}{2}\left\|2^{q-1}\mathbf{x}_q-\mathbf{z}_q\right\|_2^2, \end{aligned} \tag{14}$$

where $\mathbf{y}_q \in \mathbb{C}^U$ and $\rho > 0$ are Lagrangian multipliers and penalty parameter respectively. Then, the corresponding ADMM algorithm framework can be written as

$$\begin{aligned}\mathbf{x}_q^{k+1} &= \arg\min_{\mathbf{x}_q\in\tilde{\mathbb{X}}_q^U}\ L_\rho(\mathbf{x}_q, \{\mathbf{z}_q^k, \mathbf{y}_q^k\}_{q=1}^Q),\\ \mathbf{z}_q^{k+1} &= \arg\min_{\mathbf{z}_q}\ L_\rho(\mathbf{z}_q, \{\mathbf{x}_q^{k+1}, \mathbf{y}_q^k\}_{q=1}^Q),\\ \mathbf{y}_q^{k+1} &= \mathbf{y}_q^k + \rho\Big(2^{q-1}\mathbf{x}_q^{k+1}-\mathbf{z}_q^{k+1}\Big). \end{aligned} \tag{15}$$

Letting $\mathbf{u}_q = \frac{\mathbf{y}_q}{\rho}$, (15) can be changed to

$$\mathbf{x}_q^{k+1} = \arg\min_{\mathbf{x}_q\in\tilde{\mathbb{X}}_q^U}\Big(-\frac{\alpha_q}{2}\|\mathbf{x}_q\|_2^2+\frac{\rho}{2}\left\|2^{q-1}\mathbf{x}_q-\mathbf{z}_q^k+\mathbf{u}_q^k\right\|_2^2\Big), \tag{16a}$$

$$\begin{aligned}\mathbf{z}_q^{k+1} &= \arg\min_{\mathbf{z}_q}\Big(\frac{1}{2}\|\mathbf{r}-\mathbf{H}\sum_{q=1}^{Q}\mathbf{z}_q\|_2^2\\ &\quad+\frac{\rho}{2}\sum_{q=1}^{Q}\left\|2^{q-1}\mathbf{x}_q^{k+1}-\mathbf{z}_q+\mathbf{u}_q^k\right\|_2^2\Big), \end{aligned} \tag{16b}$$

$$\mathbf{u}_q^{k+1} = \mathbf{u}_q^k + 2^{q-1}\mathbf{x}_q^{k+1}-\mathbf{z}_q^{k+1}. \tag{16c}$$

[1]There are no computations required for PS-ADMM to perform initialization and the initial values $\{\mathbf{x}_q^1\}_{q=1}^Q$, $\mathbf{x}_0^1$, $\mathbf{y}^1$ can be zeros, ones, minus ones, or random values.

In the following, we consider to how to implement the above ADMM algorithm efficiently. First, focus on (16b). Let

$$\bar{\mathbf{z}} = \frac{1}{Q}\sum_{q=1}^{Q}\mathbf{z}_q, \quad \mathbf{a}_q^{k+1} = \mathbf{u}_q^k + 2^{q-1}\mathbf{x}_q^{k+1}.$$

Then, we obtain the following optimization problem from (16b)

$$\min_{\bar{\mathbf{z}},\mathbf{z}_q}\ \frac{1}{2}\|\mathbf{r}-Q\mathbf{H}\bar{\mathbf{z}}\|_2^2+\frac{\rho}{2}\sum_{q=1}^{Q}\left\|\mathbf{z}_q-\mathbf{a}_q^{k+1}\right\|_2^2, \tag{17a}$$

$$\text{s. t. } \bar{\mathbf{z}} = \frac{1}{Q}\sum_{q=1}^{Q}\mathbf{z}_q. \tag{17b}$$

Its Lagrangian function can be written as

$$\begin{aligned}L(\bar{\mathbf{z}}, \mathbf{z}_q, \boldsymbol{\gamma}) &= \frac{1}{2}\|\mathbf{r}-Q\mathbf{H}\bar{\mathbf{z}}\|_2^2+\frac{\rho}{2}\sum_{q=1}^{Q}\left\|\mathbf{z}_q-\mathbf{a}_q^{k+1}\right\|_2^2\\ &\quad+\mathrm{Re}\Big(\boldsymbol{\gamma}^H(\bar{\mathbf{z}}-\frac{1}{Q}\sum_{q=1}^{Q}\mathbf{z}_q)\Big), \end{aligned} \tag{18}$$

where $\boldsymbol{\gamma}$ is Lagrangian multiplier. According to Karush-Kuhn-Tucker (KKT) conditions [45], optimal solutions $\bar{\mathbf{z}}^{k+1}$, $\mathbf{z}_q^{k+1}$, and $\boldsymbol{\gamma}^*$ should satisfy $\nabla_{\mathbf{z}_q}L(\bar{\mathbf{z}}^{k+1}, \mathbf{z}_q^{k+1}, \boldsymbol{\gamma}^*) = 0$, which results in

$$\mathbf{z}_q^{k+1} = \mathbf{a}_q^{k+1}+\frac{\boldsymbol{\gamma}^*}{\rho Q}. \tag{19}$$

Plugging (19) into (17b), we obtain

$$\boldsymbol{\gamma}^* = \rho Q(\bar{\mathbf{z}}^{k+1}-\frac{1}{Q}\sum_{q=1}^{Q}\mathbf{a}_q^{k+1}).$$

Taking $\boldsymbol{\gamma}^*$ back into (19)'s RHS, we can derive $\mathbf{z}_q^{k+1}$ as

$$\mathbf{z}_q^{k+1} = \mathbf{a}_q^{k+1}+\bar{\mathbf{z}}^{k+1}-\bar{\mathbf{a}}^{k+1}, \tag{20}$$

where $\bar{\mathbf{a}}^{k+1} = \frac{1}{Q}\sum_{q=1}^{Q}\mathbf{a}_q^{k+1}$. Plugging (20) into (17a), the problem can be simplified as the following unconstrained problem

$$\bar{\mathbf{z}}^{k+1} = \arg\min_{\bar{\mathbf{z}}}\ \Big(\frac{1}{2}\|\mathbf{r}-Q\mathbf{H}\bar{\mathbf{z}}\|_2^2+\frac{\rho Q}{2}\left\|\bar{\mathbf{z}}-\bar{\mathbf{a}}^{k+1}\right\|_2^2\Big). \tag{21}$$

Second, we consider (16c). Define

$$\bar{\mathbf{u}} = \frac{1}{Q}\sum_{q=1}^{Q}\mathbf{u}_q, \quad \bar{\mathbf{x}} = \frac{1}{Q}\sum_{q=1}^{Q}2^{q-1}\mathbf{x}_q.$$

We have $\bar{\mathbf{a}}^k = \bar{\mathbf{u}}_q^k + \bar{\mathbf{x}}^{k+1}$. Plugging it and $\mathbf{a}_q^{k+1} = \mathbf{u}_q^k + 2^{q-1}\mathbf{x}_q^{k+1}$ into (20), we obtain

$$\mathbf{z}_q^{k+1} = \mathbf{u}_q^k + 2^{q-1}\mathbf{x}_q^{k+1}+\bar{\mathbf{z}}^{k+1}-\bar{\mathbf{u}}^k-\bar{\mathbf{x}}^{k+1}. \tag{22}$$

Taking (22) into (16c), we get

$$\mathbf{u}_q^{k+1} = \bar{\mathbf{u}}^k+\bar{\mathbf{x}}^{k+1}-\bar{\mathbf{z}}^{k+1}. \tag{23}$$

It is interesting to see that all the dual variables $\mathbf{u}_q^k$, $q = 1, \cdots, Q$, are equal. In the following, we will use $\mathbf{u}^k$ to take

Algorithm 2 The Proposed DPS-ADMM Algorithm

Input: $\mathbf{H}$, $\mathbf{r}$, Q, ρ, $\{\alpha_q\}_{q=1}^{Q}$
Output: $Q\bar{\mathbf{z}}^k$
1: Initialize $\{\mathbf{x}_q^1\}_{q=1}^{Q}, \bar{\mathbf{x}}^1, \bar{\mathbf{z}}^1, \mathbf{u}^1$ as the all-zeros vectors.
2: **For** $k = 1, 2, \cdots$
3: Step 1: Update $\{\mathbf{x}_q^{k+1}\}_{q=1}^{Q}$ in parallel via (25a).
4: Step 2: Compute $\bar{\mathbf{x}}^{k+1}$ and update $\bar{\mathbf{z}}^{k+1}$ via (25b).
5: Step 3: Update $\mathbf{u}^{k+1}$ via (24c).
6: **Until** some preset conditions are satisfied.

the place of $\mathbf{u}_q^k$. Third, plugging (22) into (16a) and combining (21) and (23), ADMM algorithm (16) can be equivalent to

$$\mathbf{x}_q^{k+1} = \arg\min_{\mathbf{x}_q \in \tilde{\mathbb{X}}_q^U} \Big(-\frac{\alpha_q}{2}\|\mathbf{x}_q\|_2^2 + \frac{\rho}{2}\big\|2^{q-1}\mathbf{x}_q - 2^{q-1}\mathbf{x}_q^k - \bar{\mathbf{z}}^k + \bar{\mathbf{x}}^k + \mathbf{u}^k\big\|_2^2\Big), \quad (24a)$$

$$\bar{\mathbf{z}}^{k+1} = \arg\min_{\bar{\mathbf{z}}} \Big(\frac{1}{2}\|\mathbf{r} - Q\mathbf{H}\bar{\mathbf{z}}\|_2^2 + \frac{\rho Q}{2}\big\|\bar{\mathbf{z}} - \bar{\mathbf{x}}^{k+1} - \mathbf{u}^k\big\|_2^2\Big), \quad (24b)$$

$$\mathbf{u}^{k+1} = \mathbf{u}^k + \bar{\mathbf{x}}^{k+1} - \bar{\mathbf{z}}^{k+1}, \quad (24c)$$

in the sense that their solutions are one-to-one correspondent. Moreover, we should note that the Q optimization problems in (24a) can be implemented in parallel, which is much more efficient than its original version (16) and the PS-ADMM solving algorithm in the previous subsection.

Solving (24a) and (24b) is quite similar to solving (8a) and (8b). Their optimal solutions can be determined analytically as follows

$$\mathbf{x}_q^{k+1} = \prod_{[-1,1]}\bigg(\frac{2^{q-1}\rho}{4^{q-1}\rho - \alpha_q}\Big(2^{q-1}\mathbf{x}_q^k + \bar{\mathbf{z}}^k - \bar{\mathbf{x}}^k - \mathbf{u}^k\Big)\bigg), \quad (25a)$$

$$\bar{\mathbf{z}}^{k+1} = (Q\mathbf{H}^H\mathbf{H} + \rho\mathbf{I})^{-1}\Big(\mathbf{H}^H\mathbf{r} + \rho\big(\bar{\mathbf{x}}^{k+1} + \mathbf{u}^k\big)\Big). \quad (25b)$$

The proposed DPS-ADMM algorithm for solving model (13) is summarized in *Algorithm 2*.

IV. Performance Analysis

In this section, a detailed analysis of the proposed PS-ADMM algorithm on convergence, iteration complexity, and computational cost is provided. Since the performance analysis procedure and result of the DPS-ADMM detector related to convergence and iteration complexity are quite similar to the PS-ADMM detector, we post the corresponding analysis in the arXiv website as a supplementary material of this paper.

A. Convergence Property

We have the following theorem to characterize convergence property of the proposed PS-ADMM algorithm.

Theorem 1: Assume $4^{q-1}\rho > \alpha_q, q = 1, \cdots, Q$ and $\rho > \sqrt{2}\lambda_{\max}(\mathbf{H}^H\mathbf{H})$ are satisfied. Then, sequence $\{\{\mathbf{x}_q^k\}_{q=1}^Q, \mathbf{x}_0^k, \mathbf{y}^k\}$ generated by *Algorithm 1* is convergent, i.e.,

$$\lim_{k\to+\infty}\mathbf{x}_q^k = \mathbf{x}_q^*, \quad \lim_{k\to+\infty}\mathbf{x}_0^k = \mathbf{x}_0^*, \quad \lim_{k\to+\infty}\mathbf{y}^k = \mathbf{y}^*, \quad \forall\ \mathbf{x}_q \in \tilde{\mathbb{X}}_q^U, \quad q = 1, \cdots, Q. \quad (26)$$

Moreover, $\{\mathbf{x}_q^*\}_{q=1}^Q$ is a stationary point of problem (5), i.e., $\forall\ \mathbf{x}_q \in \tilde{\mathbb{X}}_q^U$, $q = 1, \cdots, Q$, which satisfies

$$\mathrm{Re}\Big\langle \nabla_{\mathbf{x}_q}\Big(\frac{1}{2}\|\mathbf{r} - \mathbf{H}(\sum_{q=1}^{Q}2^{q-1}\mathbf{x}_q^*)\|_2^2 - \sum_{q=1}^{Q}\frac{\alpha_q}{2}\|\mathbf{x}_q^*\|_2^2\Big), \mathbf{x}_q - \mathbf{x}_q^*\Big\rangle \geq 0. \quad (27)$$

Proof: See Appendix A. ■

Theorem 1 indicates that the proposed PS-ADMM algorithm is theoretically-guaranteed converged to some stationary point of the model (5) under some mild conditions. Here, we should note that these conditions are easily satisfied since the values of penalty parameters ρ and $\{\alpha_q\}_{q=1}^Q$ can be set accordingly when the channel matrix $\mathbf{H}$ is known. The key idea of proving Theorem 1 is to find out that the potential function $L_\rho(\{\mathbf{x}_q\}_{q=1}^Q, \mathbf{x}_0, \mathbf{y})$ *decreases sufficiently* in every PS-ADMM iteration and is lower-bounded.

B. Iteration Complexity

We use the following residual

$$\sum_{q=1}^{Q}\|\mathbf{x}_q^{k+1} - \mathbf{x}_q^k\|_2^2 + \|\mathbf{x}_0^{k+1} - \mathbf{x}_0^k\|_2^2 \quad (28)$$

to measure convergence progress of the proposed PS-ADMM algorithm. Then, similar to [39], we have Theorem 2 to characterize its convergence progress.

Theorem 2: Let t be the minimum iteration index such that the residual in (28) is less than ϵ, where ϵ is the desired precise parameter for the solution. Then, we have the following iteration complexity result

$$t \leq \frac{1}{C\epsilon}\bigg(L_\rho(\{\mathbf{x}_q^1\}_{q=1}^Q, \mathbf{x}_0^1, \mathbf{y}^1) - \Big(\ell\left(\mathbf{x}_0^*\right) - \sum_{q=1}^{Q}\frac{\alpha_q}{2}\|\mathbf{x}_q^*\|_2^2\Big)\bigg),$$

where constant

$$C = \min\bigg\{\{\frac{\gamma_q(\rho)}{2}\}_{q=1}^Q, \Big(\frac{\gamma_0(\rho)}{2} - \frac{\lambda_{\max}^2(\mathbf{H}^H\mathbf{H})}{\rho}\Big)\bigg\},$$

$\gamma_q(\rho) = 4^{q-1}\rho - \alpha_q$, $\gamma_0(\rho) = \rho + \lambda_{\min}(\mathbf{H}^H\mathbf{H})$, and $\ell\left(\mathbf{x}\right) = \frac{1}{2}\|\mathbf{r} - \mathbf{H}\mathbf{x}\|_2^2$.

Proof: See Appendix B. ■

Theorem 2 indicates that the "minimum iteration" has a maximum value, which gives a low bound for the computational complexity of the proposed PS-ADMM algorithm.

C. Computational Cost

The overall computational complexity[2] of the PS-ADMM detection algorithm consists of two parts: the first part, which

[2]The computational complexity is measured by the number of complex-valued multiplications for T iterations. It should be noted that a complex-valued multiplication can be implemented via four real-valued multiplications.

TABLE I
COMPUTATIONAL COMPLEXITY FOR DIFFERENT ALGORITHMS

Algorithm	Computational complexity
MMSE [5]	$\mathcal{O}(U^3+BU^2+BU)$
NS [8]	$\mathcal{O}(\frac{1}{2}(T-1)(U^3+U^2+U))$
GS [10]	$\mathcal{O}(BU^2+BU+\frac{3}{2}TU^2)$
OCD-BOX [23]	$\mathcal{O}(T(2BU+U))$
TASER [19][3]	$\mathcal{O}(T(\frac{2}{3}U^3+\frac{5}{2}U^2+\frac{37}{12}U))$
SDR [17]	$\mathcal{O}((1+U\log_2 4^Q)^{3.5})$
K-best [16]	$\mathcal{O}(U^3)\sim\mathcal{O}(4^{QU})$
SD [42]	$\mathcal{O}(4^{\beta QU}),\quad \beta\in(0,1]$
ADMM [27]	$\mathcal{O}(U^3+BU^2+BU+T(U^2+\frac{1}{2}U))$
ADMIN [28]	$\mathcal{O}(U^3+BU^2+BU+T(U^2+U))$
PS-ADMM	$\mathcal{O}(U^3+BU^2+BU+T(U^2+\frac{1}{2}Q(Q+1)U))$
DPS-ADMM	$\mathcal{O}(U^3+BU^2+BU+T(U^2+QU))$

is independent of the number of iterations, is required to compute the $\mathbf{x}_0$ update of PS-ADMM in (12); then, it needs to be calculated only once when detecting each transmitted symbol vector. The first part of the calculations is performed in three steps: first, the multiplication of the $U\times B$ matrix $\mathbf{H}^H$ by the $B\times U$ matrix $\mathbf{H}$; second, the computation of the inversion of the regularized Gramian matrix $\mathbf{H}^H\mathbf{H}+\rho\mathbf{I}$; and third, the computation of the $U\times B$ matrix $\mathbf{H}^H$ by the $B\times 1$ vector $\mathbf{y}$ to obtain the matched-filter vector $\mathbf{H}^H\mathbf{y}$. These steps require BU^2, U^3, and BU complex multiplications, respectively. The second part, which is iteration dependent, needs to be repeated every iteration in two steps: first, the Q scalar multiplications by the U vectors in (11) for Q $\mathbf{x}_q$-updates are implemented sequentially; second, a multiplication of the $U\times U$ matrix by the $U\times 1$ vector and the Q scalar multiplications by the $U\times 1$ vector in (12). These steps require $\frac{1}{2}Q^2U$, $U^2+\frac{1}{2}QU$ complex multiplications, respectively. Combining this result with Theorem 2, we conclude that the total computational cost to attain an ϵ-optimal solution is roughly $U^3+BU^2+BU+T(U^2+\frac{1}{2}Q^2U+\frac{1}{2}QU)$, where the number of iterations $T=\frac{1}{C\epsilon}\Big(L_\rho(\{\mathbf{x}_q^1\}_{q=1}^Q,\mathbf{x}_0^1,\mathbf{y}^1)-\big(\frac{1}{2}\|\mathbf{r}-\mathbf{H}\mathbf{x}_0^*\|_2^2-\sum_{q=1}^Q\frac{\alpha_q}{2}\|\mathbf{x}_q^*\|_2^2\big)\Big)$. The main difference between the complexity of the DPS-ADMM algorithm and the PS-ADMM algorithm is that DPS-ADMM implement parallelly one scalar multiplications by the $U\times 1$ vector operations in Q $\mathbf{x}_q$-updates, which require QU complex multiplications. Since $T\ll B$ for massive MIMO detection, the computational complexity of the PS-ADMM or DPS-ADMM mainly lies in matrix multiplication and inversion computations, which is comparable to that of the linear detector.

In Table I, we compare the total computational complexity of our proposed PS-ADMM and DPS-ADMM algorithms with other state-of-the-art massive MIMO algorithms mentioned in the introduction. Moreover, we also present the total complexity of the classical linear MMSE and nonlinear SD algorithms as comparison. We see that the complexity of MMSE and NS scales with TU^3, SD has exponential complexity, which is much higher than MMSE, GS is slightly less and TASER is slightly more complex than MMSE, and K-best and SDR are between MMSE and SD. Evidently, OCD-BOX scales with TBU, which has the lowest complexity of all detectors. However, OCD-BOX could not be applied to high-speed MIMO system since its variables cannot be implemented in parallel. Finally, we note that DPS-ADMM algorithm can be implemented distributively while other algorithms do not have this nice merit. In summary, one can conclude that the computational complexity of the proposed PS/DPS-ADMM algorithms is higher than the linear algorithms, but less than the nonlinear algorithms. In the next section, we will show that proposed detectors can offer better detection performance than linear ones, such as NS, GS, and MMSE and some nonlinear ones, such as OCD-BOX, TASER, and ADMIN etc.

[3]The systems employing 4-QAM constellations, TASER algorithm cannot be applied to high-order modulation scenarios because the computational complexity is prohibitive in practice.

V. SIMULATION RESULTS

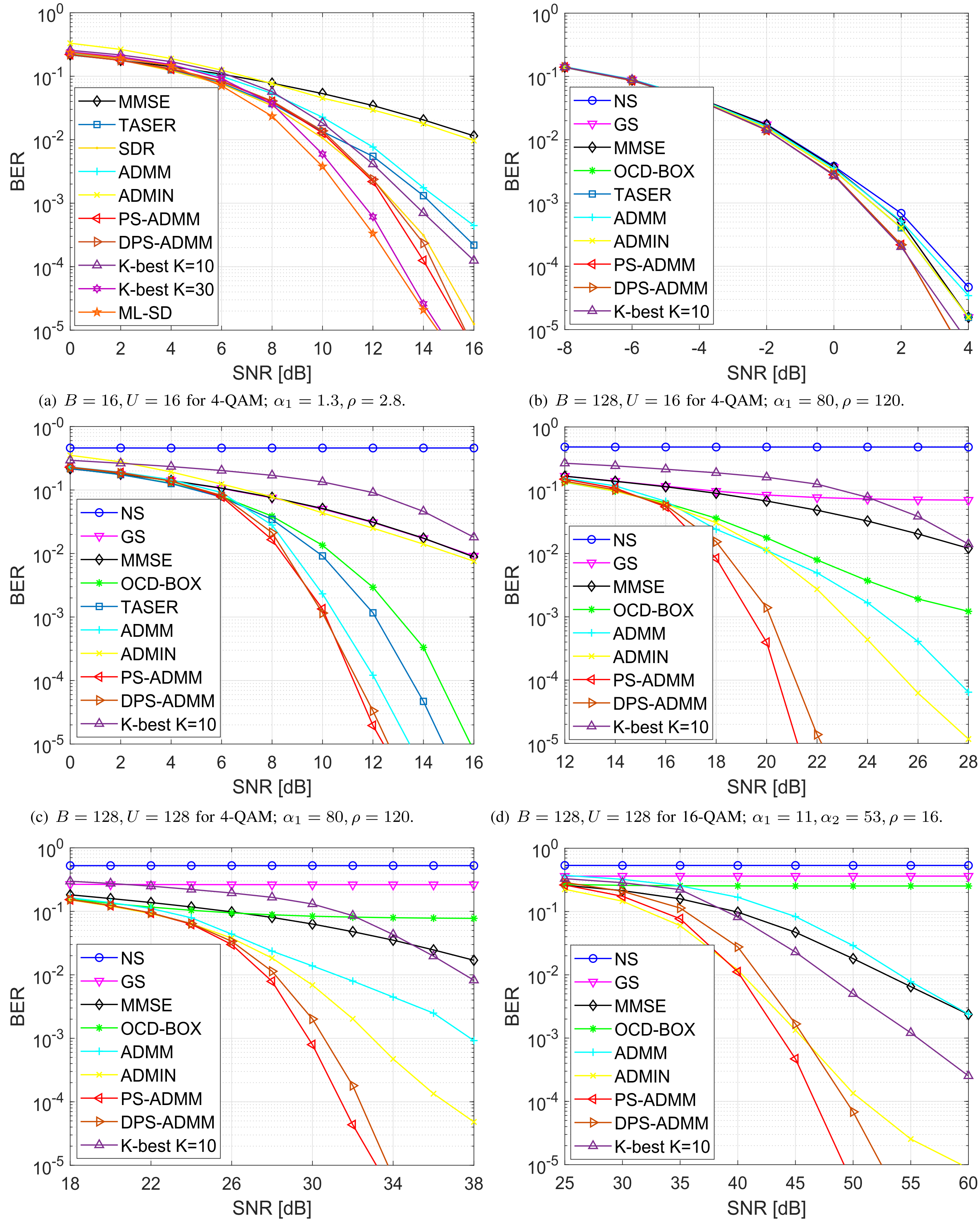
In this section, numerical results are presented to show the effectiveness of the proposed PS/DPS-ADMM detectors. Specifically, in Subsection V-A, we focus on comparing BER performance and computational complexity of the proposed PS/DPS-ADMM detectors with the detectors listed in the previous subsection. In Subsection V-B, we focus on analyzing the impact of parameters ρ, $\{\alpha_q\}_{q=1}^Q$, and T on the performance of the PS-ADMM detector.

Throughout this section, we show simulation results for uncoded and hard decision-based signal detection with the i.i.d. Rayleigh fading channel in different $B\times U$ $(B\geq U)$ multiuser massive MIMO systems. The modulation schemes of 4-QAM, 16-QAM, 64-QAM, and 256-QAM are considered. We assume that perfect knowledge of the channel state information is known at the receiver side. For a fair comparison, all the algorithms are implemented using Matlab 2019a/Windows 7 environment on a computer with 3.7GHz Intel i3-6100 $\times$ 2 CPU and 16GB RAM.

A. BER Performance and Computational Complexity

In this subsection, the BER performance and computational complexity of the proposed PS/DPS-ADMM detectors are evaluated and compared with conventional and state-of-the-art MIMO detectors by numerical simulations, which are the classical MMSE detector, NS detector [8], GS detector [11], SD detector [42], K-best detector [16], SDR detector [17], TASER detector [19], OCD-BOX detector [23], and two ADMM-based detectors ADMM and ADMIN in [27] and [28] respectively. Because the SDR and TASER detectors are difficult to support high-order modulations, we only show them in the 4-QAM systems. The termination criterium is that iteration number T reaches 30 or the residual in (28) is less than 10^{-5}. The data points plotted in all curves are averaged over 1000 Monte-Carlo trials.

Fig. 1 compares BER performance of several detectors for 4-QAM, 16-QAM, 64-QAM, and 256-QAM modulation with different numbers of antennas in the massive MIMO systems. Fig. 1(a) shows BER performance of a small-scale MIMO system, where BS equipped with 16 antennas serves 16 users and modulation is 4-QAM. Specifically, since implementing exhaustive search approach is prohibitive in practice,



(a) $B=16, U=16$ for 4-QAM; $\alpha_1=1.3, \rho=2.8$.

(b) $B=128, U=16$ for 4-QAM; $\alpha_1=80, \rho=120$.

(c) $B=128, U=128$ for 4-QAM; $\alpha_1=80, \rho=120$.

(d) $B=128, U=128$ for 16-QAM; $\alpha_1=11, \alpha_2=53, \rho=16$.

(e) $B=128, U=128$ for 64-QAM; $\alpha_1=2.17$, $\alpha_2=2.8$, $\alpha_3=10$, $\rho=3.9$.

(f) $B=128, U=128$ for 256-QAM; $\alpha_1=0.01, \alpha_2=0.01, \alpha_3=0.01, \alpha_4=0.01, \rho=0.03$.

Fig. 1. Comparisons of BER performance using various massive MIMO detectors.

we use SD detector to produce ML performance. From the figure, we see that the proposed PS/DPS-ADMM detectors achieve competitive detection performances, which are lower than ML-SD and K-best ($K=30$), but outperform MMSE, ADMIN, ADMM, TASER, SDR, and K-best ($K=10$). In Fig. 1(b), one can find that all of the detectors have comparable BER performance when the BS-to-user-antenna ratio is more than two. In Fig. 1(c), 1(d), 1(e), and 1(f), for

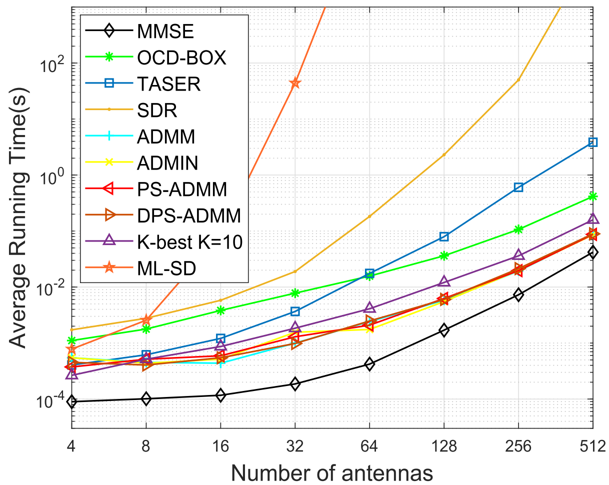
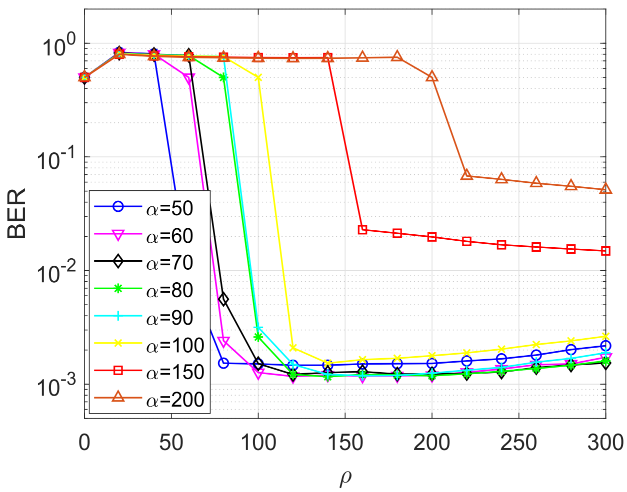
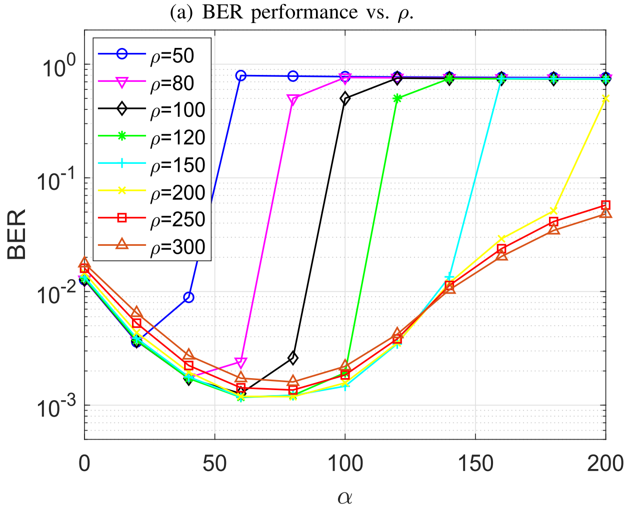


Fig. 2. Complexity comparison of massive MIMO detectors with $U = B$ for 4-QAM, SNR $= 12$dB.

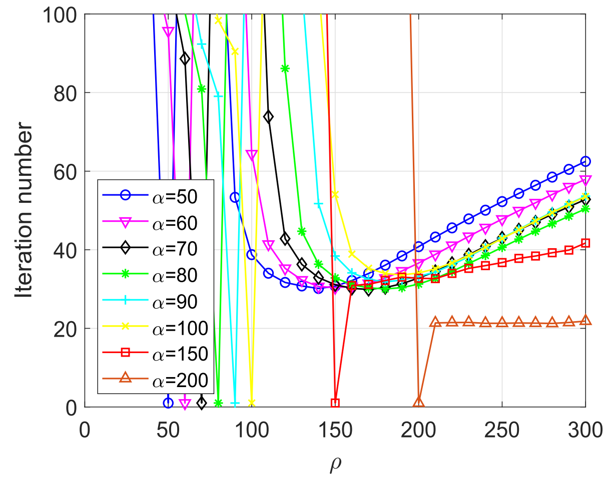


(a) BER performance vs. ρ.

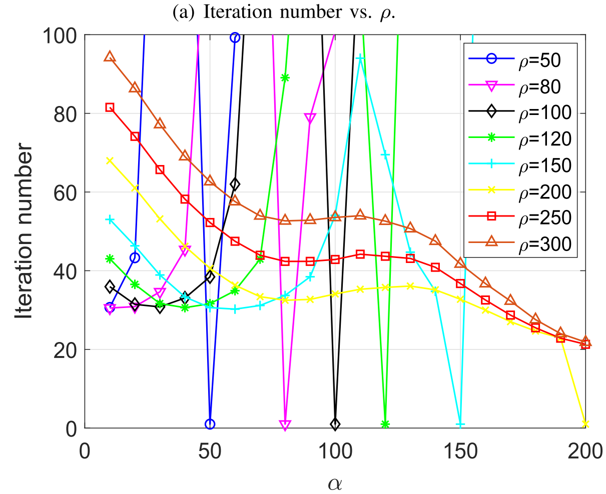


(b) BER performance vs. α.

Fig. 3. The impact of ρ and α on BER performance of the PS-ADMM detector.



(a) Iteration number vs. ρ.



(b) Iteration number vs. α.

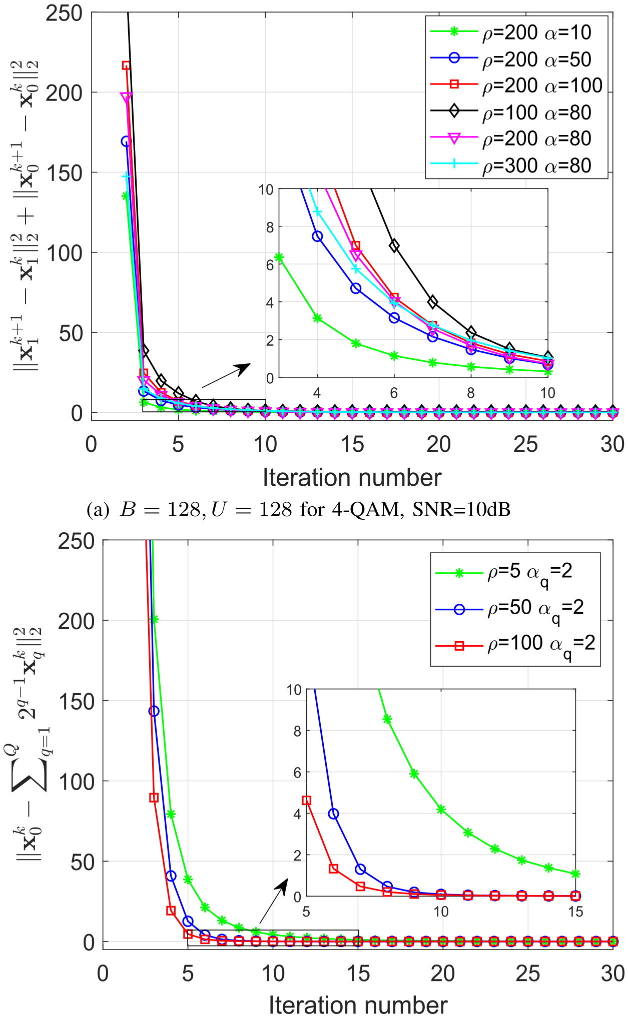
Fig. 4. The impact of ρ and α on convergence performance of the PS-ADMM detector.

the challenging 128×128 massive MIMO system, one can see that the proposed PS/DPS-ADMM detectors outperform other detectors and their BER curves are almost coincident. However, it should be noted that PS-ADMM is slightly better than DPS-ADMM in terms of the BER performance, but DPS-ADMM has the advantage of distributed processing. Moreover, one can also see that for large-scale MIMO systems with high-order QAM signals, BER curves of the proposed PS/DPS-ADMM detectors still keep waterfall style, but other detectors do not have this nice property.

Fig. 2 compares computational complexity of the massive MIMO detectors. The averaged running time refers to detection time required for a symbol vector sent by the multiple transmit antennas. From the figure, one can see that the running time of the proposed PS/DPS-ADMM detectors is higher than MMSE, similar to ADMM, ADMIN, and K-best (K $=$ 10), and lower than other ones. However, we should note that PS/DPS-ADMM have much better detection performance than MMSE and K-best (K $=$ 10). Therefore, one can conclude that the proposed PS/DPS-ADMM detectors can deliver an attractive tradeoff between BER performance and computational complexity.

B. Choice of Parameters

In this subsection, we show that the proper parameters ρ and α can achieve lower BER performance and speed



(a) $B = 128, U = 128$ for 4-QAM, SNR=10dB

(b) $B = 128, U = 128$ for 64-QAM, SNR=32dB

Fig. 5. The impact of the maximum iteration number T on convergence performance of the PS-ADMM detector.

up convergence of the proposed PS-ADMM detector. The considered modulation scheme is 4-QAM and the simulation parameters are $B = 128$, $U = 128$, and SNR $= 10$dB. Since the impact of parameters on the performance of the DPS-ADMM detector is very similar to that of the PS-ADMM detector, we do not provide the relevant simulation results for the DPS-ADMM detector.

In Fig. 3, it shows the effects on BER performance when the different values of the penalty parameters ρ and α are chosen. From the figure, one can have the following observations: first, both ρ and α can affect BER performance of the proposed PS-ADMM decoder; second, too large or too small values of α and ρ can worsen BER performance of the detector. For the case of the presented simulation, one can see the proper $\rho \in [120\ 200]$ and $\alpha \in [60\ 80]$ respectively. Moreover, we note that when α approaches zero, the PS-ADMM detector degenerates to the conventional ADMM detector with a box constraint.

In Fig. 4, we study the effects of the penalty parameters ρ and α on the convergence characteristic of the proposed PS-ADMM decoder. From the figures, one can observe that the proposed PS-ADMM algorithm can always converge with different settings of the parameters ρ and α when $\rho > \alpha$ is satisfied, and these parameters can affect its convergence rate. From the figure, it shows that the larger the penalty parameter ρ and the value of α is close to ρ, the faster the algorithm converges, but these too large parameters ρ and α are not a good choice for BER performance of the PS-ADMM detector. There is no need to sacrifice a lot of BER performance just to reduce a dozen iterations. We can observe that the optimal value of ρ and α in terms of convergence rate agrees with the optimal value of ρ and α in terms of BER performance when iteration number reaches 30.

In Fig. 5, by checking the residuals $\|\mathbf{x}_1^{k+1} - \mathbf{x}_1^k\|_2^2 + \|\mathbf{x}_0^{k+1} - \mathbf{x}_0^k\|_2^2$ and $\|\mathbf{x}_0^k - \sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q^k\|_2^2$, respectively, one can observe the impact of ρ and α on convergence performance. Moreover, it can also be seen that the proposed PS-ADMM algorithm can converge within a few tens of iterations to converge.

Specifically, for higher order modulations, proper BER and convergence performance can be obtained only when $4^{q-1}\rho > \alpha_q, q = 1, \cdots, Q$ is satisfied. Moreover, we find that the simulation results are sensitive to T, ρ, and α_q, especially for a high-order QAM scheme. Here, we provide some tips on choosing their values:

- A smaller α_q is a good choice.
- Set ρ meets $4^{q-1}\rho > \alpha_q$.
- Set $T \in [20, 30]$.

VI. Conclusion

In this paper, we proposed two QAM detectors named PS-ADMM and DPS-ADMM for the uplink massive MIMO system. We show that the proposed PS/DPS-ADMM detectors have competitive BER performance and cheap computational complexity, especially when the BS-to-user-antenna ratio is close to one. The correspondent theoretical analysis on convergence and computational complexity are provided. Since channel coding and soft decision-based MIMO detection is essentially used in practical systems, it would be meaningful to further study how the proposed PS/DPS-ADMM approaches can be applied to such scenarios. In addition, how to choose the optimal penalty parameter is also an interesting research topic. In the end, we note that even though this work focuses on a massive MIMO detection problem, the proposed PS/DPS-ADMM methods could also be applied to solving some large-scale optimization problems in practice [46], [48]–[50].

Appendix A
Proof of Theorem 1

Before proving convergence of the proposed PS-ADMM algorithm, we give several lemmas and their proofs as follows.

Lemma 1: For Algorithm 1, the following inequality holds

$$\|\mathbf{y}^{k+1} - \mathbf{y}^k\|_2^2 \leq \lambda_{\max}^2(\mathbf{H}^H\mathbf{H})\|\mathbf{x}_0^{k+1} - \mathbf{x}_0^k\|_2^2. \qquad (29)$$

Proof: Since $\mathbf{x}_0^{k+1}$ is a minimizer of problem (8b), it should satisfy the following optimality condition [43]

$$\nabla_{\mathbf{x}_0}\ell\left(\mathbf{x}_0^{k+1}\right) + \mathbf{y}^k + \rho(\mathbf{x}_0^{k+1} - \sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q^{k+1}) = 0. \qquad (30)$$

Plugging $\mathbf{y}^{k+1}$ in (8c) into the above equation, we obtain

$$\mathbf{y}^{k+1} = -\nabla_{\mathbf{x}_0}\ell\left(\mathbf{x}_0^{k+1}\right). \tag{31}$$

According to Lagrange's mean value theorem, since $\ell(\mathbf{x}_0)$ is continuous and differentiable, there exists some point $\bar{\mathbf{x}}_0$ between $\mathbf{x}_0^k$ and $\mathbf{x}_0^{k+1}$ which satisfies

$$\frac{\nabla_{\mathbf{x}_0}\ell\left(\mathbf{x}_0^{k+1}\right) - \nabla_{\mathbf{x}_0}\ell\left(\mathbf{x}_0^k\right)}{\mathbf{x}_0^{k+1} - \mathbf{x}_0^k} = \nabla_{\bar{\mathbf{x}}_0}^2\ell\left(\mathbf{x}_0\right). \tag{32}$$

Moreover, since $\nabla_{\bar{\mathbf{x}}_0}^2\ell(\mathbf{x}_0) = \mathbf{H}^H\mathbf{H} \preceq \lambda_{\max}(\mathbf{H}^H\mathbf{H})\mathbf{I}$, we have

$$\|\nabla_{\mathbf{x}_0}\ell\left(\mathbf{x}_0^{k+1}\right) - \nabla_{\mathbf{x}_0}\ell\left(\mathbf{x}_0^k\right)\|_2^2 \leq \lambda_{\max}^2(\mathbf{H}^H\mathbf{H})\|\mathbf{x}_0^{k+1} - \mathbf{x}_0^k\|_2^2. \tag{33}$$

From (33), we can see that $\nabla_{\mathbf{x}_0}\ell(\mathbf{x}_0)$ is Lipschitz continuous with constant $\lambda_{\max}(\mathbf{H}^H\mathbf{H})$. Plugging (31) into LHS of equation (33), we can obtain

$$\begin{aligned}\|\mathbf{y}^{k+1} - \mathbf{y}^k\|_2^2 &= \|\nabla_{\mathbf{x}_0}\ell\left(\mathbf{x}_0^{k+1}\right) - \nabla_{\mathbf{x}_0}\ell\left(\mathbf{x}_0^k\right)\|_2^2 \\ &\leq \lambda_{\max}^2(\mathbf{H}^H\mathbf{H})\|\mathbf{x}_0^{k+1} - \mathbf{x}_0^k\|_2^2.\end{aligned}$$

This completes the proof. ∎

Lemma 2: Let α_q and ρ satisfy $4^{q-1}\rho > \alpha_q$, $\forall q \in \{1, \cdots, Q\}$. Then, for Algorithm 1, we have the following inequality

$$\begin{aligned}&L_\rho\left(\{\mathbf{x}_q^{k+1}\}_{q=1}^Q, \mathbf{x}_0^{k+1}, \mathbf{y}^{k+1}\right) - L_\rho\left(\{\mathbf{x}_q^k\}_{q=1}^Q, \mathbf{x}_0^k, \mathbf{y}^k\right) \\ &\leq -\sum_{q=1}^Q \frac{\gamma_q(\rho)}{2}\|\mathbf{x}_q^{k+1} - \mathbf{x}_q^k\|_2^2 \\ &\quad - \left(\frac{\gamma_0(\rho)}{2} - \frac{\lambda_{\max}^2(\mathbf{H}^H\mathbf{H})}{\rho}\right)\|\mathbf{x}_0^{k+1} - \mathbf{x}_0^k\|_2^2,\end{aligned} \tag{34}$$

where $\gamma_q(\rho) = 4^{q-1}\rho - \alpha_q$ and $\gamma_0(\rho) = \rho + \lambda_{\min}(\mathbf{H}^H\mathbf{H})$.

Proof: We split LHS of the inequality (34) into two terms

$$\begin{aligned}&L_\rho(\{\mathbf{x}_q^{k+1}\}_{q=1}^Q, \mathbf{x}_0^{k+1}, \mathbf{y}^{k+1}) - L_\rho(\{\mathbf{x}_q^k\}_{q=1}^Q, \mathbf{x}_0^k, \mathbf{y}^k) \\ &= \underbrace{\left(L_\rho(\{\mathbf{x}_q^{k+1}\}_{q=1}^Q, \mathbf{x}_0^{k+1}, \mathbf{y}^{k+1}) - L_\rho(\{\mathbf{x}_q^{k+1}\}_{q=1}^Q, \mathbf{x}_0^{k+1}, \mathbf{y}^k)\right)}_{\text{term 1}} \\ &\quad + \underbrace{\left(L_\rho(\{\mathbf{x}_q^{k+1}\}_{q=1}^Q, \mathbf{x}_0^{k+1}, \mathbf{y}^k) - L_\rho(\{\mathbf{x}_q^k\}_{q=1}^Q, \mathbf{x}_0^k, \mathbf{y}^k)\right)}_{\text{term 2}}.\end{aligned}$$

For the first term, we have the following derivations

$$\begin{aligned}&L_\rho(\{\mathbf{x}_q^{k+1}\}_{q=1}^Q, \mathbf{x}_0^{k+1}, \mathbf{y}^{k+1}) - L_\rho(\{\mathbf{x}_q^{k+1}\}_{q=1}^Q, \mathbf{x}_0^{k+1}, \mathbf{y}^k) \\ &= \mathrm{Re}\left\langle \mathbf{x}_0^{k+1} - \sum_{q=1}^Q 2^{q-1}\mathbf{x}_q^{k+1}, \mathbf{y}^{k+1}\right\rangle \\ &\quad - \mathrm{Re}\left\langle \mathbf{x}_0^{k+1} - \sum_{q=1}^Q 2^{q-1}\mathbf{x}_q^{k+1}, \mathbf{y}^k\right\rangle \\ &= \mathrm{Re}\left\langle \mathbf{x}_0^{k+1} - \sum_{q=1}^Q 2^{q-1}\mathbf{x}_q^{k+1},\ \mathbf{y}^{k+1} - \mathbf{y}^k\right\rangle \\ &\overset{\text{(a)}}{=} \frac{1}{\rho}\|\mathbf{y}^{k+1} - \mathbf{y}^k\|_2^2 \overset{\text{(b)}}{\leq} \frac{\lambda_{\max}^2(\mathbf{H}^H\mathbf{H})}{\rho}\|\mathbf{x}_0^{k+1} - \mathbf{x}_0^k\|_2^2,\end{aligned} \tag{35}$$

where "$\overset{\text{(a)}}{=}$" and "$\overset{\text{(b)}}{\leq}$" comes from (8c) and (29) respectively. For the second term, we have the following derivations

$$\begin{aligned}&L_\rho(\{\mathbf{x}_q^{k+1}\}_{q=1}^Q, \mathbf{x}_0^{k+1}, \mathbf{y}^k) - L_\rho(\{\mathbf{x}_q^k\}_{q=1}^Q, \mathbf{x}_0^k, \mathbf{y}^k) \\ &= L_\rho(\{\mathbf{x}_q^{k+1}\}_{q=1}^Q, \mathbf{x}_0^k, \mathbf{y}^k) - L_\rho(\{\mathbf{x}_q^k\}_{q=1}^Q, \mathbf{x}_0^k, \mathbf{y}^k) \\ &\quad + L_\rho(\{\mathbf{x}_q^{k+1}\}_{q=1}^Q, \mathbf{x}_0^{k+1}, \mathbf{y}^k) - L_\rho(\{\mathbf{x}_q^{k+1}\}_{q=1}^Q, \mathbf{x}_0^k, \mathbf{y}^k) \\ &\leq \sum_{q=1}^Q \Big(\mathrm{Re}\left\langle \nabla_{\mathbf{x}_q} L_\rho(\mathbf{x}_1^{k+1}, \cdots, \mathbf{x}_{q-1}^{k+1}, \mathbf{x}_q^{k+1}, \mathbf{x}_{q+1}^k,\right. \\ &\quad \left.\cdots, \mathbf{x}_Q^k, \mathbf{x}_0^k, \mathbf{y}^k), \mathbf{x}_q^{k+1} - \mathbf{x}_q^k\right\rangle \\ &\quad - \frac{4^{q-1}\rho - \alpha_q}{2}\|\mathbf{x}_q^{k+1} - \mathbf{x}_q^k\|_2^2\Big) \\ &\quad + \mathrm{Re}\left\langle \nabla_{\mathbf{x}_0} L_\rho(\{\mathbf{x}_q^{k+1}\}_{q=1}^Q, \mathbf{x}_0^{k+1}, \mathbf{y}^k), \mathbf{x}_0^{k+1} - \mathbf{x}_0^k\right\rangle \\ &\quad - \frac{\rho + \lambda_{\min}(\mathbf{H}^H\mathbf{H})}{2}\|\mathbf{x}_0^{k+1} - \mathbf{x}_0^k\|_2^2 \\ &\leq -\sum_{q=1}^Q \frac{4^{q-1}\rho - \alpha_q}{2}\|\mathbf{x}_q^{k+1} \\ &\quad - \mathbf{x}_q^k\|_2^2 - \frac{\rho + \lambda_{\min}(\mathbf{H}^H\mathbf{H})}{2}\|\mathbf{x}_0^{k+1} - \mathbf{x}_0^k\|_2^2,\end{aligned} \tag{36}$$

where the first inequality holds since the corresponding augmented Lagrangian functions are strongly convex [44] and the second inequality holds since $\mathbf{x}_q^{k+1}$ and $\mathbf{x}_0^{k+1}$ are minimizers of the problems (8a) and (8b), i.e.,

$$\begin{aligned}&\left\langle \nabla_{\mathbf{x}_q} L_\rho(\mathbf{x}_q^{k+1}, \mathbf{x}_1^{k+1}, \cdots, \mathbf{x}_{q-1}^{k+1}, \mathbf{x}_{q+1}^k, \cdots, \mathbf{x}_Q^k, \mathbf{x}_0^k, \mathbf{y}^k),\right. \\ &\quad \left.\mathbf{x}_q^k - \mathbf{x}_q^{k+1}\right\rangle \geq 0 \\ &\quad \nabla_{\mathbf{x}_0} L_\rho(\{\mathbf{x}_q^{k+1}\}_{q=1}^Q, \mathbf{x}_0^{k+1}, \mathbf{y}^k) = 0.\end{aligned}$$

Adding both sides of inequalities (35) and (36) and letting $\gamma_q(\rho) = 4^{q-1}\rho - \alpha_q$ and $\gamma_0(\rho) = \rho + \lambda_{\min}(\mathbf{H}^H\mathbf{H})$, we can obtain

$$\begin{aligned}&L_\rho(\{\mathbf{x}_q^{k+1}\}_{q=1}^Q, \mathbf{x}_0^{k+1}, \mathbf{y}^{k+1}) - L_\rho(\{\mathbf{x}_q^k\}_{q=1}^Q, \mathbf{x}_0^k, \mathbf{y}^k) \\ &\leq -\sum_{q=1}^Q \frac{\gamma_q(\rho)}{2}\|\mathbf{x}_q^{k+1} - \mathbf{x}_q^k\|_2^2 \\ &\quad - \left(\frac{\gamma_0(\rho)}{2} - \frac{\lambda_{\max}^2(\mathbf{H}^H\mathbf{H})}{\rho}\right)\|\mathbf{x}_0^{k+1} - \mathbf{x}_0^k\|_2^2,\end{aligned}$$

which completes the proof. ∎

Lemma 3: Let α_q, $\forall q \in \{1, \cdots, Q\}$, and ρ satisfy $4^{q-1}\rho > \alpha_q$ and $\rho > \sqrt{2}\lambda_{\max}(\mathbf{H}^H\mathbf{H})$. Assume tuples $\{\{\mathbf{x}_q^k\}_{q=1}^Q, \mathbf{x}_0^k, \mathbf{y}^k\}$ is generated by Algorithm 1, then $L_\rho(\{\mathbf{x}_q^k\}_{q=1}^Q, \mathbf{x}_0^k, \mathbf{y}^k)$ is lower bounded as follows

$$L_\rho(\{\mathbf{x}_q^k\}_{q=1}^Q, \mathbf{x}_0^k, \mathbf{y}^k) \geq \ell\Big(\sum_{q=1}^Q 2^{q-1}\mathbf{x}_q^k\Big) - \sum_{q=1}^Q \frac{\alpha_q}{2}\|\mathbf{x}_q^k\|_2^2. \tag{37}$$

Proof: Plugging (31) into (7), we obtain

$$\begin{aligned}&L_\rho(\{\mathbf{x}_q^k\}_{q=1}^Q, \mathbf{x}_0^k, \mathbf{y}^k) \\ &= \ell\left(\mathbf{x}_0^k\right) - \sum_{q=1}^Q \frac{\alpha_q}{2}\|\mathbf{x}_q^k\|_2^2\end{aligned}$$

$$+\mathrm{Re}\Big\langle \sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q^k - \mathbf{x}_0^k, \nabla_{\mathbf{x}_0}\ell\left(\mathbf{x}_0^k\right)\Big\rangle + \frac{\rho}{2}\Big\|\mathbf{x}_0^k - \sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q^k\Big\|_2^2. \quad (38)$$

Since we show that gradient $\|\nabla_{\mathbf{x}_0}\ell\left(\mathbf{x}_0\right)\|_2$ is Lipschitz continuous in Lemma 1 and $\|\nabla^2_{\mathbf{x}_0}\ell\left(\mathbf{x}_0\right)\|_2 \le \lambda_{\max}(\mathbf{H}^H\mathbf{H})$, according to the Decent Lemma [45], we can obtain

$$\begin{aligned}&\ell\Big(\sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q^k\Big)\\ &\le \ell\left(\mathbf{x}_0^k\right) + \mathrm{Re}\Big\langle \nabla_{\mathbf{x}_0}\ell\left(\mathbf{x}_0^k\right), \sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q^k - \mathbf{x}_0^k\Big\rangle\\ &\quad + \frac{\lambda_{\max}(\mathbf{H}^H\mathbf{H})}{2}\Big\|\sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q^k - \mathbf{x}_0^k\Big\|_2^2,\end{aligned}$$

which can be further derived to the following inequality

$$\begin{aligned}&\ell\left(\mathbf{x}_0^k\right) + \mathrm{Re}\Big\langle \sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q^k - \mathbf{x}_0^k, \nabla_{\mathbf{x}_0}\ell\left(\mathbf{x}_0^k\right)\Big\rangle\\ &\ge \ell\Big(\sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q^k\Big) - \frac{\lambda_{\max}(\mathbf{H}^H\mathbf{H})}{2}\Big\|\sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q^k - \mathbf{x}_0^k\Big\|_2^2.\end{aligned} \quad (39)$$

Plugging (39) into (38), we can get

$$\begin{aligned}&L_\rho(\{\mathbf{x}_q^k\}_{q=1}^Q, \mathbf{x}_0^k, \mathbf{y}^k)\\ &\ge \ell\Big(\sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q^k\Big) - \sum_{q=1}^{Q}\frac{\alpha_q}{2}\|\mathbf{x}_q^k\|_2^2\\ &\quad + \frac{\rho - \lambda_{\max}(\mathbf{H}^H\mathbf{H})}{2}\Big\|\mathbf{x}_0^k - \sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q^k\Big\|_2^2.\end{aligned} \quad (40)$$

Since $\ell\big(\sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q^k\big) - \sum_{q=1}^{Q}\frac{\alpha_q}{2}\|\mathbf{x}_q^k\|_2^2$ is bounded over $\mathbf{x}_{qR}, \mathbf{x}_{qI} \in [-1\ \ 1]^U$, as well as the fact that $\rho - \lambda_{\max}(\mathbf{H}^H\mathbf{H}) > 0$ comes from $\rho > \sqrt{2}\lambda_{\max}(\mathbf{H}^H\mathbf{H})$. Using these two cases leads to the desired result that $L_\rho(\{\mathbf{x}_q^k\}_{q=1}^Q, \mathbf{x}_0^k, \mathbf{y}^k)$ is lower bounded and Lemma 3 has been proved. ■

Based on the above lemmas, we are ready to prove Theorem 1. First, we consider to prove (26) in Theorem 1. According to Lemma 2, summing both sides of the inequality (34) when $k = 1, 2, \cdots, +\infty$, we can obtain

$$\begin{aligned}&L_\rho\left(\{\mathbf{x}_q^1\}_{q=1}^Q, \mathbf{x}_0^1, \mathbf{y}^1\right) - \lim_{k\to+\infty} L_\rho\left(\{\mathbf{x}_q^k\}_{q=1}^Q, \mathbf{x}_0^k, \mathbf{y}^k\right)\\ &\ge \sum_{k=1}^{+\infty}\sum_{q=1}^{Q}\frac{\gamma_q(\rho)}{2}\|\mathbf{x}_q^{k+1} - \mathbf{x}_q^k\|_2^2\\ &\quad + \sum_{k=1}^{+\infty}\Big(\frac{\gamma_0(\rho)}{2} - \frac{\lambda^2_{\max}(\mathbf{H}^H\mathbf{H})}{\rho}\Big)\|\mathbf{x}_0^{k+1} - \mathbf{x}_0^k\|_2^2.\end{aligned} \quad (41)$$

From Lemma 3, we have $\lim_{k\to+\infty} L_\rho\left(\{\mathbf{x}_q^k\}_{q=1}^Q, \mathbf{x}_0^k, \mathbf{y}^k\right) > -\infty$. Moreover, since $\gamma_0(\rho) > 0$ and $\frac{\gamma_0(\rho)}{2} - \frac{\lambda^2_{\max}(\mathbf{H}^H\mathbf{H})}{\rho} > 0$, we can obtain

$$\lim_{k\to+\infty}\|\mathbf{x}_0^{k+1} - \mathbf{x}_0^k\|_2 = 0, \quad (42a)$$

$$\lim_{k\to+\infty}\|\mathbf{x}_q^{k+1} - \mathbf{x}_q^k\|_2 = 0, \quad q \in \{1, \cdots, Q\}. \quad (42b)$$

Plugging (42a) into RHS of equation (29), we get

$$\lim_{k\to+\infty}\|\mathbf{y}^{k+1} - \mathbf{y}^k\|_2 = 0. \quad (43)$$

Plugging (43) into (8c), we get

$$\lim_{k\to+\infty}\|\mathbf{x}_0^{k+1} - \sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q^{k+1}\|_2 = 0. \quad (44)$$

Since $\mathbf{x}_{q\mathrm{R}}, \mathbf{x}_{q\mathrm{I}} \in [-1\ \ 1]^U$, we can obtain the following convergence results from (42b).

$$\lim_{k\to+\infty}\mathbf{x}_q^k = \mathbf{x}_q^*, \quad \forall\ q = 1, 2, \cdots, Q. \quad (45)$$

Plugging (45) into (44), we can see

$$\lim_{k\to+\infty}\mathbf{x}_0^k = \mathbf{x}_0^* = \sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q^*. \quad (46)$$

From (31), we can derive

$$\lim_{k\to+\infty}\mathbf{y}^k = \lim_{k\to+\infty} -\nabla_{\mathbf{x}_0}\ell\left(\mathbf{x}_0^k\right). \quad (47)$$

Since $\|\nabla_{\mathbf{x}_0}\ell\left(\mathbf{x}_0^{k+1}\right) - \nabla_{\mathbf{x}_0}\ell\left(\mathbf{x}_0^k\right)\|_2^2 \le \lambda^2_{\max}(\mathbf{H}^H\mathbf{H})\|\mathbf{x}_0^{k+1} - \mathbf{x}_0^k\|_2^2$ and $\mathbf{x}_0^k$ is bounded, we can conclude that all the elements in $\nabla_{\mathbf{x}_0}\ell\left(\mathbf{x}_0\right)$ are also bounded. Therefore, equation (43) indicates

$$\lim_{k\to+\infty}\mathbf{y}^k = \mathbf{y}^*. \quad (48)$$

Second, we consider to prove (27). Since $\{\mathbf{x}_q^{k+1}\}_{q=1}^Q = \arg\min_{\mathbf{x}_q\in\tilde{\mathbb{X}}_q^U} L_\rho\left(\{\mathbf{x}_q\}_{q=1}^Q, \mathbf{x}_0^k, \mathbf{y}^k\right)$ (see (8a)) and $L_\rho\left(\{\mathbf{x}_q\}_{q=1}^Q, \mathbf{x}_0^k, \mathbf{y}^k\right)$ is strongly convex w.r.t. $\mathbf{x}_q$, we have the following optimality conditions

$$\begin{aligned}&\mathrm{Re}\Big\langle \nabla_{\mathbf{x}_q}\Big(\ell\left(\mathbf{x}_0^k\right) - \sum_{q=1}^{Q}\frac{\alpha_q}{2}\|\mathbf{x}_q^{k+1}\|_2^2\\ &\quad + \big\langle \mathbf{x}_0^k - \sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q^{k+1}, \mathbf{y}^k\big\rangle + \frac{\rho}{2}\big\|\mathbf{x}_0^k - \sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q^{k+1}\big\|_2^2\Big),\\ &\quad \mathbf{x}_q - \mathbf{x}_q^{k+1}\Big\rangle \ge 0, \quad \forall\ \mathbf{x}_q \in \tilde{\mathbb{X}}_q^U,\ q = 1, 2, \cdots, Q.\end{aligned} \quad (49)$$

When $k \to +\infty$, plugging convergence results (45) and (46) into (49), it can be simplified as

$$\mathrm{Re}\Big\langle \nabla_{\mathbf{x}_q}\Big(\ell\Big(\sum_{q=1}^{Q} 2^{q-1}\mathbf{x}_q^*\Big) - \sum_{q=1}^{Q}\frac{\alpha_q}{2}\|\mathbf{x}_q^*\|_2^2\Big), \mathbf{x}_q - \mathbf{x}_q^*\Big\rangle \ge 0,$$

where $\forall\ \mathbf{x}_q \in \tilde{\mathbb{X}}_q^U,\ q = 1, 2, \cdots, Q$. It concludes the proof.

APPENDIX B
PROOF OF THEOREM 2

To be clear, here we rewrite (34) as

$$\begin{aligned} &L_\rho\left(\{\mathbf{x}_q^k\}_{q=1}^Q, \mathbf{x}_0^k, \mathbf{y}^k\right) - L_\rho\left(\{\mathbf{x}_q^{k+1}\}_{q=1}^Q, \mathbf{x}_0^{k+1}, \mathbf{y}^{k+1}\right) \\ &\geq \sum_{q=1}^Q \frac{\gamma_q(\rho)}{2}\|\mathbf{x}_q^{k+1} - \mathbf{x}_q^k\|_2^2 \\ &\quad + \Big(\frac{\gamma(\rho)}{2} - \frac{\lambda_{\max}^2(\mathbf{H}^H\mathbf{H})}{\rho}\Big)\|\mathbf{x}_0^{k+1} - \mathbf{x}_0^k\|_2^2. \end{aligned}$$

According to Lemma 2, there exists a constant $C = \min\left\{\{\frac{\gamma_q(\rho)}{2}\}_{q=1}^Q, \left(\frac{\gamma(\rho)}{2} - \frac{\lambda_{\max}^2(\mathbf{H}^H\mathbf{H})}{\rho}\right)\right\}$ such that

$$\begin{aligned} &L_\rho\left(\{\mathbf{x}_q^k\}_{q=1}^Q, \mathbf{x}_0^k, \mathbf{y}^k\right) - L_\rho\left(\{\mathbf{x}_q^{k+1}\}_{q=1}^Q, \mathbf{x}_0^{k+1}, \mathbf{y}^{k+1}\right) \\ &\qquad\qquad \geq C\Big(\sum_{q=1}^Q \|\mathbf{x}_q^{k+1} - \mathbf{x}_q^k\|_2^2 + \|\mathbf{x}_0^{k+1} - \mathbf{x}_0^k\|_2^2\Big). \end{aligned}$$

Summing both sides of the above inequality from $k = 1, \cdots, T$, we have

$$\begin{aligned} &L_\rho\left(\{\mathbf{x}_q^1\}_{q=1}^Q, \mathbf{x}_0^1, \mathbf{y}^1\right) - L_\rho\left(\{\mathbf{x}_q^{T+1}\}_{q=1}^Q, \mathbf{x}_0^{T+1}, \mathbf{y}^{T+1}\right) \\ &\geq \sum_{k=1}^T \left(C\Big(\sum_{q=1}^Q \|\mathbf{x}_q^{k+1} - \mathbf{x}_q^k\|_2^2 + \|\mathbf{x}_0^{k+1} - \mathbf{x}_0^k\|_2^2\Big)\right). \end{aligned} \tag{50}$$

Since $t = \min_k\{k | \sum_{q=1}^Q \|\mathbf{x}_q^{k+1} - \mathbf{x}_q^k\|_2^2 + \|\mathbf{x}_0^{k+1} - \mathbf{x}_0^k\|_2^2 \leq \epsilon\}$, we can change (50) to

$$\begin{aligned} &L_\rho\left(\{\mathbf{x}_q^1\}_{q=1}^Q, \mathbf{x}_0^1, \mathbf{y}^1\right) - L_\rho\left(\{\mathbf{x}_q^{T+1}\}_{q=1}^Q, \mathbf{x}_0^{T+1}, \mathbf{y}^{T+1}\right) \\ &\qquad\qquad\qquad \geq tC\epsilon. \end{aligned} \tag{51}$$

Since we have $L_\rho\left(\{\mathbf{x}_q^{T+1}\}_{q=1}^Q, \mathbf{x}_0^{T+1}, \mathbf{y}^{T+1}\right) \geq L_\rho\left(\{\mathbf{x}_q^*\}_{q=1}^Q, \mathbf{x}_0^*, \mathbf{y}^*\right)$, (51) can be reduced to

$$t \leq \frac{1}{C\epsilon}\left(L_\rho\left(\{\mathbf{x}_q^1\}_{q=1}^Q, \mathbf{x}_0^1, \mathbf{y}^1\right) - L_\rho\left(\{\mathbf{x}_q^*\}_{q=1}^Q, \mathbf{x}_0^*, \mathbf{y}^*\right)\right),$$

where $L_\rho\left(\{\mathbf{x}_q^*\}_{q=1}^Q, \mathbf{x}_0^*, \mathbf{y}^*\right) = \ell\left(\mathbf{x}_0^*\right) - \sum_{q=1}^Q \frac{\alpha_q}{2}\|\mathbf{x}_q^*\|_2^2$, which concludes the proof.

REFERENCES

[1] 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.

[2] F. Boccardi, R. W. Heath, Jr., 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.

[3] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Feb. 2013.

[4] M. A. Albreem, M. Juntti, and S. Shahabuddin, "Massive MIMO detection techniques: A survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3109–3132, 4th Quart., 2019.

[5] D. A. Shnidman, "A generalized Nyquist criterion and an optimum linear receiver for a pulse modulation system," *Bell Syst. Tech. J.*, vol. 46, no. 9, pp. 2163–2177, Nov. 1967.

[6] W. van Etten, "An optimum linear receiver for multiple channel digital transmission systems," *IEEE Trans. Commun.*, vol. COM-23, no. 8, pp. 828–834, Aug. 1975.

[7] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[8] M. Wu, B. Yin, G. Wang, C. Dick, J. R. Cavallaro, and C. Studer, "Large-scale MIMO detection for 3GPP LTE: Algorithms and FPGA implementations," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 916–929, Oct. 2014.

[9] M. Čirkič and E. G. Larsson, "On the complexity of very large multi-user MIMO detection," in *Proc. IEEE 15th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2014, pp. 55–59.

[10] L. Dai *et. al.*, "Low-complexity soft-output signal detection based on Gauss–Seidel method for uplink multiuser large-scale MIMO systems," *IEEE Trans. Veh. Tech.*, vol. 64, no. 10, pp. 4839–4845, Oct. 2015.

[11] Z. Wu, C. Zhang, Y. Xue, S. Xu, and X. You, "Efficient architecture for soft-output massive MIMO detection with Gauss–Seidel method," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2016, pp. 1886–1889.

[12] X. Gao, L. Dai, Y. Ma, and Z. Wang, "Low-complexity near-optimal signal detection for uplink large-scale MIMO systems," *Electron. Lett.*, vol. 50, no. 18, pp. 1326–1328, Aug. 2014.

[13] B. Yin, M. Wu, J. R. Cavallaro, and C. Studer, "Conjugate gradient-based soft-output detection and precoding in massive MIMO systems," in *Proc. IEEE Global Commun. Conf.*, Austin, TX, USA, Dec. 2014, pp. 3696–3701.

[14] S. Verdu, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Trans. Inf. Theory*, vol. IT-32, no. 1, pp. 85–96, Jan. 1986.

[15] M. O. Damen, H. El Gamal, and G. Caire, "On maximum-likelihood detection and the search for the closest lattice point," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2389–2402, Oct. 2003.

[16] K.-W. Wong, C.-Y. Tsui, R. S.-K. Cheng, and W.-H. Mow, "A VLSI architecture of a K-best lattice decoding algorithm for MIMO channels," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, vol. 3, May 2002, p. 3.

[17] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.

[18] S. Yang and L. Hanzo, "Fifty years of MIMO detection: The road to large-scale MIMOs," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 1941–1988, Sep. 2015.

[19] O. Castañeda, T. Goldstein, and C. Studer, "Data detection in large multi-antenna wireless systems via approximate semidefinite relaxation," *IEEE Trans. Circuits Syst.*, vol. 63, no. 12, pp. 2334–2346, Dec. 2016.

[20] K. V. Vardhan, S. K. Mohammed, A. Chockalingam, and B. S. Rajan, "A low-complexity detector for large MIMO systems and multicarrier CDMA systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 3, pp. 473–485, Apr. 2008.

[21] N. Srinidhi, S. K. Mohammed, A. Chockalingam, and B. S. Rajan, "Low-complexity near-ML decoding of large non-orthogonal STBCs using reactive Tabu search," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2009, pp. 1993–1997.

[22] M. Suneel, P. Som, A. Chockalingam, and B. S. Rajan, "Belief propagation based decoding of large non-orthogonal STBCs," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2009, pp. 2003–2007.

[23] M. Wu, C. Dick, J. R. Cavallaro, and C. Studer, "High-throughput data detection for massive MU-MIMO-OFDM using coordinate descent," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 63, no. 12, pp. 2357–2367, Dec. 2016.

[24] A. K. Sah and A. K. Chaturvedi, "An MMP-based approach for detection in large MIMO systems using sphere decoding," *IEEE Wireless Commun. Lett.*, vol. 6, no. 2, pp. 158–161, Apr. 2017.

[25] N. Samuel, T. Diskin, and A. Wiesel, "Learning to detect," *IEEE Trans. Signal Process.*, vol. 67, no. 10, pp. 2554–2564, May 2019.

[26] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.

[27] R. Takapoui, N. Moehle, S. Boyd, and A. Bemporad, "A simple effective heuristic for embedded mixed-integer quadratic programming," in *Proc. Amer. Control Conf.*, Boston, MA, USA, Jul. 2016, pp. 5619–5625.

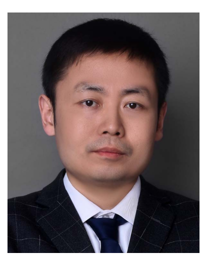
[28] S. Shahabuddin, I. Hautala, M. Juntti, and C. Studer, "ADMM-based infinity-norm detection for massive MIMO: Algorithm and VLSI architecture," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 29, no. 4, pp. 747–759, Apr. 2021.

[29] N. Souto and R. Dinis, "MIMO detection and equalization for single-carrier systems using the alternating direction method of multipliers," *IEEE Signal Process. Lett.*, vol. 23, no. 12, pp. 1751–1755, Dec. 2016.

[30] N. Souto and R. Dinis, "Efficient MIMO detection for high-order QAM constellations in time dispersive channels," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2018, pp. 1–6.
[31] H. Lopes and N. Souto, "Iterative signal detection for large-scale GSM-MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7734–7738, Aug. 2018.
[32] A. Elgabli, A. Elghariani, V. Aggarwal, and M. R. Bell, "A low-complexity detection algorithm for uplink massive MIMO systems based on alternating minimization," *IEEE Wireless Commun. Lett.*, vol. 8, no. 3, pp. 917–920, Jun. 2019.
[33] A. Elgabli, A. Elghariani, V. Aggarwal, M. Bennis, and M. R. Bell, "A proximal Jacobian ADMM approach for fast massive MIMO signal detection in low-latency communications," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.
[34] H. Feng, X. Zhao, Z. Li, and S. Xing, "A novel iterative discrete estimation algorithm for low-complexity signal detection in uplink massive MIMO systems," *Electronics*, vol. 8, no. 9, p. 980, Sep. 2019.
[35] M.-W. Un, M. Shao, W.-K. Ma, and P. C. Ching, "Deep MIMO detection using ADMM unfolding," in *Proc. IEEE Data Sci. Workshop (DSW)*, Jun. 2019, pp. 333–337.
[36] W.-K. Ma, T. N. Davidson, K. M. Wong, Z.-Q. Luo, and P.-C. Ching, "Quasi-maximum-likelihood multiuser detection using semi-definite relaxation with application to synchronous CDMA," *IEEE Trans. Signal Process.*, vol. 50, no. 4, pp. 912–922, Apr. 2002.
[37] Z. Mao, X. Wang, and X. Wang, "Semidefinite programming relaxation approach for multiuser detection of QAM signals," *IEEE Trans. Wireless Commun.*, vol. 6, no. 12, pp. 4275–4279, Dec. 2007.
[38] W.-K. Ma, C.-C. Su, J. Jalden, T.-H. Chang, and C.-Y. Chi, "The equivalence of semidefinite relaxation MIMO detectors for higher-order QAM," *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 6, pp. 1038–1052, Dec. 2009.
[39] M. Hong, Z.-Q. Luo, and M. Razaviyayn, "Convergence analysis of alternating direction method of multipliers for a family of nonconvex problems," *SIAM J. Optim.*, vol. 26, no. 1, pp. 337–364, 2016.
[40] J.-C. Chen, "A low complexity data detection algorithm for uplink multiuser massive MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 8, pp. 1701–1714, Aug. 2017.
[41] M. Shao and W.-K. Ma, "Binary MIMO detection via homotopy optimization and its deep adaptation," *IEEE Trans. Signal Process.*, vol. 69, pp. 781–796, 2021.
[42] A. Burg, M. Borgmann, M. Wenk, M. Zellweger, W. Fichtner, and H. Bolcskei, "VLSI implementation of MIMO detection using the sphere decoding algorithm," *IEEE J. Solid-State Circuits*, vol. 40, no. 7, pp. 1566–1577, Jul. 2005.
[43] D. P. Bertsekas, *Convex Optimization Theory*. Belmont, MA, USA: Athena Scientific, 2009.
[44] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[45] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA, USA: Athena Scientific, 1999.
[46] K. Xu, Z. Shen, Y. Wang, X. Xia, and D. Zhang, "Hybrid time-switching and power splitting swipt for full-duplex massive MIMO systems: A beam-domain approach," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7257–7274, Aug. 2018.
[47] Y. Wang, Y. Wang, and Q. Shi, "Optimized signal distortion for PAPR reduction of OFDM signals with IFFT/FFT complexity via ADMM approaches," *IEEE Trans. Signal Process.*, vol. 67, no. 2, pp. 399–414, Jan. 2019.
[48] Z. Shen, K. Xu, and X. Xia, "Beam-domain anti-jamming transmission for downlink massive MIMO systems: A Stackelberg game perspective," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 2727–2742, 2021.
[49] Z. Shen, K. Xu, X. Xia, W. Xie, and D. Zhang, "Spatial sparsity based secure transmission strategy for massive MIMO systems against simultaneous jamming and eavesdropping," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 3760–3774, 2020.
[50] K. Xu, Z. Shen, Y. Wang, and X. Xia, "Location-aided mMIMO channel tracking and hybrid beamforming for high-speed railway communications: An angle-domain approach," *IEEE Syst. J.*, vol. 14, no. 1, pp. 93–104, Mar. 2020.

Quan Zhang received the B.E. degree in satellite communication from Air Force Engineering University, Xi'an, China, in 2001, and the M.E. degree in information and communication engineering from Xi'an Jiaotong University, Xi'an, in 2008. He is currently pursuing the Ph.D. degree in communication and information systems with Xidian University, Xi'an. His research interests include communication signal processing, optimization algorithms, and their applications.

Jiangtao Wang received the B.E. degree in electronic information engineering from Shandong Normal University, Jinan, China, in 2011, and the Ph.D. degree in communication and information systems from Xidian University, Xi'an, in 2020. He is currently a Post-Doctoral Researcher with the State Key Laboratory of Integrated Services Network (ISN), Xidian University. His research interests include convex optimization and efficient algorithms with applications in sequence set design, MIMO radar waveform design, and communication signal processing.



Yongchao Wang (Senior Member, IEEE) received the B.E. degree in communication engineering and the M.E. and Ph.D. degrees in information and communication engineering from Xidian University, Xi'an, China, in 1998, 2004, and 2006, respectively. From September 2008 to January 2010, he was a one-year Visiting Scholar and then a Post-Doctoral Fellow with the Departmen t of Electronics and Communication Engineering, University of Minnesota, USA. From March 2016 to March 2017, he was also a Visiting Scholar with Purdue University. Since 2012, he has been a Professor with the State Key Laboratory of Integrated System Network (ISN), Xidian University. Moreover, he was a recipient of several awards, such as prize in progress of science and technology from the Ministry of Education of the People's Republic of China, Shaanxi government, and Xidian University. His research works have been applied to several real telecommunication systems.