A New Randomized Iterative Detection Algorithm For Uplink Large-scale MIMO Systems

Zheng Wang, *Member, IEEE*, Wei Xu, *Senior Member, IEEE*, Yili Xia, *Member, IEEE*, Qingjiang Shi, *Senior Member, IEEE* and Yongming Huang, *Senior Member, IEEE*

***Abstract*—In this paper, a new randomized iterative detection algorithm (NRIDA) is proposed for uplink large-scale MIMO systems, where the random iterations in it are designed to work for the detection model (denoted by $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$) directly. Different from those traditional iterations designed for the linear system (denoted by $\mathbf{A}\mathbf{x} = \mathbf{b}$ with $\mathbf{A} = \mathbf{H}^H\mathbf{H}$ and $\mathbf{b} = \mathbf{H}^H\mathbf{y}$), we show that besides the complexity reduction about the matrix inversion, in the proposed NRIDA the computational complexity of matrix multiplication for the linear detection is also greatly reduced without any performance loss, thus leading to a much lower detection complexity. Meanwhile, according to convergence analysis, we demonstrate that the proposed NRIDA enjoys a globally exponential convergence performance, enabling it well suited to the various detection cases of interest. Besides, further complexity reduction and the choices of the sampling distribution in NRIDA are studied as well in full details. Moreover, in order to achieve a better detection trade-off between performance and complexity, we introduce the concept of the conditional sampling into NRIDA, which brings significant gains in both iteration convergence and efficiency. Finally, simulations with respect to the uplink large-scale MIMO detection are presented to illustrate the remarkable gains of the proposed NRIDA in both performance and complexity.**

***Index Terms*—Massive MIMO, large-scale MIMO detection, low complexity, random sampling, iterative methods.**

I. Introduction

As a promising extension of multiple-input multiple-output (MIMO) in beyond 5G and 6G, the large-scale MIMO system is able to significantly improve the network capacity without extra bandwidth [1]–[3]. However, because of the curse of dimensionality, there is also a big challenge upon the uplink signal detection with the fast increase of the system dimension [4], [5]. In this condition, it has been proved in [6] that the optimal maximum likelihood (ML) detection can be achieved by the linear detection schemes like zero forcing (ZF) and minimum mean-square error (MMSE) if the number of received antennas at base station (BS) side (denoted by N) is sufficiently larger than the number of transmitted antennas at user side (denoted by K), i.e., $N \gg K$. Nevertheless, because of the complicated matrix inversion, the implementation of these linear detection are still challenging in practice. To this end, a number of low-complexity iterative detection schemes are proposed to bypass the matrix inversion in an iterative way [7]–[12].

In general, the low-complexity linear detection schemes based on the traditional iterative methods have a common problem solving paradigm [13], [14]. Specifically, they firstly convert the uplink detection problem (i.e., $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$) into decoding a linear system (i.e., $\mathbf{A}\mathbf{x} = \mathbf{b}$), and then employ the traditional iterative methods designed for linear systems to iteratively approximate the detection solution [15]. More precisely, low-complexity linear detection schemes based on Jacobi, Richardson, Gauss-Seidel (GS) and successive overrelaxation (SOR) iterations have different matrix splitting ways about matrix $\mathbf{A}$, which correspond to the different convergence performance and complexity cost [16]. Nevertheless, some specific requirements like $N \gg K$ should be satisfied, otherwise the convergence of these iterations can not be ensured, making them severely limited in various cases of interest [17]. For example, the convergence of Jacobi iteration is guaranteed if ZF or MMSE filtering matrix $\mathbf{A}$ is heavily diagonally dominant [9], [18]. The relaxation factor $0 < \omega < 2/\varrho(\mathbf{A})$ in Richardson iteration needs to be well chosen for the convergence [10], [19], where $\varrho(\cdot)$ indicates the spectral radius of a matrix.

Nowadays, the network of wireless communications has become much more complicated than before [20], [21], which places a higher requirement about the flexible uplink detection schemes. To well suit different cases of large-scale MIMO systems, the randomized iterative detection algorithm (RIDA) is given in [22], which adopts the random sampling into the iteration methods. It has been demonstrated that RIDA not only achieves a low computational complexity but also enjoys a global convergence performance. To further enhance the convergence performance, the modified randomized iterative detection algorithm (MRIDA) is also given in [22] by fully taking advantages of the conditional sampling. Overall, the method of random iteration greatly extends the applications of iterative detection in large-scale MIMO systems. In fact, besides the low-complexity detection, the concept of random sampling has also been introduced into large-scale MIMO systems to achieve a better detection performance [23]–[28].

However, before the iteration is carried out, both the traditional iterative detection (e.g., Jacobi, Richardson, GS, SOR) and the randomized iterative detection (e.g., RIDA, MRIDA)

This work was supported in part by the National Natural Science Foundation of China under Grant 61801216, 61771124, 62022026, 62231019, 61720106003 and 62211530108, in part by the Natural Science Foundation of Jiangsu Province under Grant BK20180420, in part by the Zhi Shan Young Scholar Program of Southeast University, in part by the Fundamental Research Funds for the Central Universities 2242022k30002.

Z. Wang, W. Xu, Y. Xia and Y. Huang are with National Mobile Communications Research Laboratory, and School of Information Science and Engineering, Southeast University, Nanjing 210096, China (e-mail: z.wang@ieee.org). Qingjiang Shi is with the School of Software Engineering, Tongji University, Shanghai 201804, China, and also with the Shenzhen Research Institute of Big Data, Shenzhen 518172, China.

schemes require to get the filtering matrix $\mathbf{A} = \mathbf{H}^H\mathbf{H}$ so that the following iterations based on matrix splitting can be performed. Essentially, this is because those iterative methods are originated from solving the problem of linear systems (i.e., $\mathbf{A}\mathbf{x} = \mathbf{b}$), so that a necessary transformation about the uplink signal detection (i.e., $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$) has to be made for such an adoption. Compared to avoiding the matrix inversion about $\mathbf{A}$ via the designed iterations, calculating the matrix $\mathbf{A}$ is also computationally expensive, but is widely ignored in the related works as a preprocessing [15], [22]. Therefore, considerable complexity cost has been consumed before the iterations, which imposes an inevitable obstacle on the implementation of these iterative detection schemes.

In this paper, a new randomized iterative detection algorithm (NRIDA) is proposed, which incorporates the random iteration directly into the detection model $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$. By doing this, we show that the computations of obtaining $\mathbf{A}$ and $\mathbf{b}$ in the iterative detection schemes can be avoided without any performance loss. Therefore, a lot of computational complexity are saved, which is greatly beneficial to the implementation of the low-complexity iterative detection. At the same time, we also demonstrate that the proposed NRIDA scheme converges exponentially and globally by means of mean squared error, which is competitive among the iterative detection schemes. Besides the convergence analysis, the choices of the sampling distribution in the proposed NRIDA are investigated while further convergence enhancement by taking full advantages of the multi-conditional sampling is also presented as well. To summarize, compared with the traditional iterative detection algorithms, the proposed NRIDA not only achieves a better signal detection trade-off between performance and complexity but also entails a global and tractable convergence, thus making it a better choice for uplink large-scale MIMO systems.

The organization of this paper is as follows. Section II describes the background of the conventional linear detection in large-scale MIMO systems and briefly introduces the state of the art of the low-complexity iterative detection schemes. Then, the proposed NRIDA is given in Section III, and the feasible complexity reduction of it is also investigated in detail. In Section IV, the convergence analysis is presented to show the globally exponential convergence of the proposed NRIDA, followed by the study of the choices of the sampling distribution. In Section V, by making use of the conditional sampling, NRIDA with multi-step conditional sampling is proposed to achieve better iteration convergence and efficiency. In Section VI, simulations about the proposed NRIDA for uplink signal detection in large-scale MIMO systems are shown, and Section VII concludes the paper finally.

Notation: Matrices and column vectors are represented by upper and lowercase boldface letters, and the conjugate transpose, inverse, pseudoinverse of a matrix $\mathbf{B}$ by $\mathbf{B}^H, \mathbf{B}^{-1}$, and $\mathbf{B}^\dagger$, respectively. We apply $\mathbf{b}_i$ for the ith column of the matrix $\mathbf{B}$, $b_{i,j}$ for the entry in the ith row and jth column of the matrix $\mathbf{B}$. Let $\langle \mathbf{X}, \mathbf{Y} \rangle_{F(\mathbf{W}^{-1})} \triangleq \text{Tr}(\mathbf{X}^H\mathbf{W}^{-1}\mathbf{Y}\mathbf{W}^{-1})$ indicate the weighting Frobenius inner product, where $\mathbf{X}, \mathbf{Y} \in \mathbb{C}^{n\times n}$ and $\mathbf{W} \in \mathbb{C}^{n\times n}$ is a symmetric positive definite matrix. Furthermore, let $\|\mathbf{X}\|^2_{F(\mathbf{W}^{-1})} \triangleq \text{Tr}(\mathbf{X}^H\mathbf{W}^{-1}\mathbf{X}\mathbf{W}^{-1}) = \|\mathbf{W}^{-\frac{1}{2}}\mathbf{X}\mathbf{W}^{-\frac{1}{2}}\|^2_F$ and $\|\cdot\|_F$ is the standard Frobenius norm with identity matrix $\mathbf{I}$, where $\text{Tr}(\cdot)$ stands for the trace of the matrix. $\lceil x \rfloor$ indicates rounding to the closest integer about x. If x is a complex value, $\lceil x \rfloor$ separately rounds the real and imaginary parts independently.

II. Preliminary

In this section, the classic linear detection in large-scale MIMO systems is introduced, followed by the state of the art about low-complexity iterative detection schemes.

A. Linear Uplink Signal Detection

Consider decoding a large-scale MIMO system, where N antennas are equipped at base station (BS) and multiple user equipments (UEs) with K antennas in all are served simultaneously, $N \geq K$. Let $\mathbf{x}$ represent the $K \times 1$ transmit signal from UEs, where the i-th element of $\mathbf{x}$ (i.e., x_i) is a symbol drawn from QAM constellation $\mathcal{X}$. Then, given the flat fading channel matrix $\mathbf{H} \in \mathbb{C}^{N\times K}$, the $N \times 1$ receive signal $\mathbf{y}$ at BS can be written by

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

where $\mathbf{n}$ is an $N \times 1$ AWGN noise vector whose elements obey $\mathcal{CN}(\mathbf{0}, \sigma^2)$. Intuitively, to restore the transmit signal $\mathbf{x}$ in (1), the optimal maximum likelihood (ML) detection aims to solve the following integer least square (ILS) problem

$$\widehat{\mathbf{x}}_{\text{ml}} = \arg\min_{\mathbf{x}\in\mathcal{X}^K} \|\mathbf{H}\mathbf{x} - \mathbf{y}\|^2, \tag{2}$$

which turns out to be NP-hard with the increase of system dimension.

To solve the problem in (2), the traditional linear detection schemes like ZF and MMSE can be applied with approximated solutions

$$\mathbf{x}_{\text{zf}} = (\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H\mathbf{y} \tag{3}$$

and

$$\mathbf{x}_{\text{mmse}} = \left(\mathbf{H}^H\mathbf{H} + \sigma^2\mathbf{I}\right)^{-1}\mathbf{H}^H\mathbf{y} \tag{4}$$

respectively. Then, the final decisions $\widehat{\mathbf{x}}_{\text{zf}}$ and $\widehat{\mathbf{x}}_{\text{mmse}}$ are determined by quantizing $\mathbf{x}_{\text{zf}}$ and $\mathbf{x}_{\text{mmse}}$ according to the modulation constellation $\mathcal{X}^K$, i.e.,

$$\widehat{\mathbf{x}}_{\text{zf}} = \lceil \mathbf{x}_{\text{zf}} \rfloor_Q \in \mathcal{X}^K \text{ and } \widehat{\mathbf{x}}_{\text{mmse}} = \lceil \mathbf{x}_{\text{mmse}} \rfloor_Q \in \mathcal{X}^K. \tag{5}$$

Theoretically, both $\mathbf{x}_{\text{zf}}$ in (3) and $\mathbf{x}_{\text{mmse}}$ in (4) are designed to solve the following least square (LS) problem

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbb{C}^K} \|\mathbf{H}\mathbf{x} - \mathbf{y}\|^2, \tag{6}$$

which is different from the ILS problem in (2). More specifically, it has been demonstrated in [6] that the optimal ML detection performance about ILS problem in (2) can be obtained by ZF or MMSE detection if $N \gg K$, making the linear detection widely applied in large-scale MIMO systems. Unfortunately, the implementation of ZF or MMSE decoding is still challenging in practice due to the complicated matrix operations in it. Typically, the computational complexities of ZE and MMSE detection mainly consist of two parts, namely, the matrix multiplication $\mathbf{H}^H\mathbf{H}$ with computational complexity $O(NK^2)$ and the matrix inversion of it with computational complexity $O(K^3)$.

B. Low-Complexity Iterative Detection

To approximate the matrix inversion in linear detection with low complexity cost, some iterative detection schemes are given based on the classic iterative methods.

In particular, the ZF decoding in (3) (or MMSE decoding in (4)) can be interpreted by decoding an equivalent linear system

$$\mathbf{A}\mathbf{x} = \mathbf{b}. \tag{7}$$

Here, $\mathbf{b} = \mathbf{H}^H\mathbf{y} \in \mathbb{C}^K$, $\mathbf{A} = \mathbf{H}^H\mathbf{H} \in \mathbb{C}^{K\times K}$ is the symmetric positive ZF filtering matrix (so is the MMSE filtering matrix $\mathbf{A} = \mathbf{H}^H\mathbf{H}+\sigma^2\mathbf{I}$). To solve this linear system, the matrix splitting about $\mathbf{A} = \mathbf{M} + \mathbf{N}$ is employed with $\mathbf{M} \in \mathbb{C}^{K\times K}$ and $\mathbf{N} \in \mathbb{C}^{K\times K}$ so that iterative methods compute the solution by repeatedly applying the following iterations [29]

$$\mathbf{x}^{k+1} = \mathbf{G}\mathbf{x}^k + \mathbf{g}, \tag{8}$$

where $\mathbf{g} = \mathbf{M}^{-1}\mathbf{b} \in \mathbb{C}^K$ and $\mathbf{G} = -\mathbf{M}^{-1}\mathbf{N} = \mathbf{I} - \mathbf{M}^{-1}\mathbf{A} \in \mathbb{C}^{K\times K}$ is known as the *iteration matrix*. Moreover, the convergence of iterative methods is guaranteed if [17]

$$\lim_{k\to\infty} \mathbf{G}^k = \mathbf{0}, \tag{9}$$

so that $\mathbf{x}^{t+1}$ will gradually approach $\mathbf{x}_{\text{zf}}$ along the iterations.

Intuitively, the choices of $\mathbf{M}$ and $\mathbf{N}$ for the matrix splitting $\mathbf{A} = \mathbf{M}+\mathbf{N}$ are key to iterative methods. In Jacobi iteration, the matrix splitting is set with $\mathbf{M} = \mathbf{D}$ and $\mathbf{N} = \mathbf{L} + \mathbf{U}$, which leads to the following iterations [30]

$$\mathbf{D}\mathbf{x}^{k+1} = -(\mathbf{U} + \mathbf{L})\mathbf{x}^k + \mathbf{b}. \tag{10}$$

Here, $\mathbf{D} \in \mathbb{C}^{K\times K}$, $\mathbf{L} \in \mathbb{C}^{K\times K}$ and $\mathbf{U} \in \mathbb{C}^{K\times K}$ denote the diagonal elements, the lower triangular elements and the upper triangular elements of matrix $\mathbf{A}$ respectively as $\mathbf{A} = \mathbf{D}+\mathbf{L}+\mathbf{U}$ and $\mathbf{L} = \mathbf{U}^H$. As for Richardson iteration, matrices $\mathbf{M} = \frac{1}{\omega}\mathbf{I}$ and $\mathbf{N} = \mathbf{A}-\frac{1}{\omega}\mathbf{I}$ are applied, where $\omega > 0$ serves as the relaxation factor [7]. For a faster convergence performance, GS iteration with $\mathbf{M} = \mathbf{D} + \mathbf{U}$ and $\mathbf{N} = \mathbf{L}$ is introduced as [11]

$$(\mathbf{D} + \mathbf{U})\mathbf{x}^{k+1} = -\mathbf{L}\mathbf{x}^k + \mathbf{b}. \tag{11}$$

Based on GS iteration, SOR iteration is proposed for a better convergence performance, which introduces the relaxation factor $1 < \omega < 2$ into the iterations as [16]

$$(\mathbf{D} + \omega\mathbf{U})\mathbf{x}^{k+1} = [(1 - \omega)\mathbf{D} - \omega\mathbf{L}]\mathbf{x}^k + \omega\mathbf{b}. \tag{12}$$

Consequently, given the iteration results $\mathbf{x}^L$, $L \geq 1$, the final detection solution of iterative methods for the ILS problem in (2) is outputted by the direct quantization as

$$\widehat{\mathbf{x}} = \lceil \mathbf{x}^L \rfloor_Q \in \mathcal{X}^K. \tag{13}$$

Although low-complexity linear detection can be achieved by iterative detection schemes, necessary conditions (e.g., $N \gg K$) need to be fulfilled to guarantee the convergence, which severely limits their applications in practice.

C. Low-Complexity Randomized Iterative Detection

In [22], the randomized iterative detection algorithm (RIDA) is given for uplink signal detection in large-scale MIMO systems. By adopting random sampling into the iterations, it not only has a low computational complexity, but also entails a globally exponential convergence.

Specifically, in [22], a general formation of randomized iteration for large-scale MIMO detection is derived as

$$\mathbf{x}^{k+1}{=}\mathbf{x}^k{+}\mathbf{U}^{-1}\mathbf{A}^H\mathbf{S}_k(\mathbf{S}_k^H\mathbf{A}\mathbf{U}^{-1}\mathbf{A}^H\mathbf{S}_k)^{-1}\mathbf{S}_k^H(\mathbf{b}{-}\mathbf{A}\mathbf{x}^k). \tag{14}$$

Here, $\mathbf{U} \in \mathbb{C}^{K\times K}$ serves as an auxiliary matrix, which is symmetric positive definite. The matrix $\mathbf{S}_k$ is randomly sampled from a designed discrete distribution $\mathcal{D}$, namely, $\mathbf{S}_k \sim \mathcal{D}$. From it, the random iteration in RIDA follows

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \mathbf{A}^H\mathbf{S}_k(\mathbf{S}_k^H\mathbf{A}\mathbf{A}^H\mathbf{S}_k)^{-1}\mathbf{S}_k^H(\mathbf{b} - \mathbf{A}\mathbf{x}^k) \tag{15}$$

with the initial choice $\mathbf{x}^0 = \mathbf{D}^{-1}\mathbf{b}$, where $\mathbf{S}_k \in \{\mathbf{I}_{:,\mathcal{Q}_1}, \ldots, \mathbf{I}_{:,\mathcal{Q}_r}\}$ at each iteration is randomly sampled from a distribution $\mathcal{D}$. Here, $\mathbf{I}_{:,\mathcal{Q}_i}$ with $1 \leq i \leq r$ denotes a column partition of the identity matrix $\mathbf{I}_{K\times K}$ and $\mathcal{Q}_i \triangleq \{\text{index } 1, \ldots, \text{index } q_i\} \subseteq \{1, \ldots, K\}$ standards for an index set with $|\mathcal{Q}_i| = q_i$, namely,

$$\mathbf{I}_{:,\mathcal{Q}_i} = [\mathbf{I}_{:,\text{index}1}, \ldots, \mathbf{I}_{:,\text{index}q_i}], \tag{16}$$

which establishes a block operation as follows, e.g.,

$$\underbrace{\{1, 2, 5\}}_{\mathcal{Q}_1} \cup \ldots \cup \underbrace{\{4, 8, 12\}}_{\mathcal{Q}_r} = \{1, \ldots, K\} \tag{17}$$

with $\mathcal{Q}_i \cap \mathcal{Q}_j = \emptyset$, $1 \leq i \neq j \leq r$. Note that the indices in each set $\mathcal{Q}_i$ are determined initially while the following sampling with respect to $\mathbf{S}_k$ is carried out block by block (i.e., $\mathbf{I}_{:,\mathcal{Q}_i}$) rather than column by column (i.e., $\mathbf{I}_{:,\text{index}}$). For simplicity, the same block size $|\mathcal{Q}_1| = \ldots = |\mathcal{Q}_r| = K/r = q$ (i.e., $q_1 = \ldots = q_r = q$) is applied [22].

Theorem 1 ([22])**.** *As for the uplink large-scale MIMO detection, let $\mathbf{S}_k \in \{\mathbf{I}_{:,\mathcal{Q}_1}, \ldots, \mathbf{I}_{:,\mathcal{Q}_r}\}$ be sampled randomly according to the distribution $\mathcal{D}$, RIDA converges by*

$$E[\|\mathbf{x}^k - \mathbf{x}^*\|^2] \leq \rho^k\|\mathbf{x}^0 - \mathbf{x}^*\|^2 \tag{18}$$

with globally exponential convergence rate

$$\rho = 1 - \lambda_{\text{min}}(E[\mathbf{A}_{:,\mathcal{Q}_i}(\mathbf{A}_{\mathcal{Q}_i,:}\mathbf{A}_{:,\mathcal{Q}_i})^{-1}\mathbf{A}_{\mathcal{Q}_i,:}]) < 1. \tag{19}$$

Here, $\lambda_{\text{min}}(\cdot)$ represents the minimum eigenvalue of a matrix, $\mathbf{x}^* = \mathbf{A}^{-1}\mathbf{b}$ indicates the detection solution of (6) and (7), $E[\cdot]$ represents the expectation of a random variable, $\mathbf{A}_{\mathcal{Q}_i,:} = \mathbf{I}^H_{:,\mathcal{Q}_i}\mathbf{A}$ and $\mathbf{A}_{:,\mathcal{Q}_i} = \mathbf{A}^H\mathbf{I}_{:,\mathcal{Q}_i}$ stand for the row and column partitions respectively with index set $\mathcal{Q}_i$. Clearly, in [22], the exponential convergence of RIDA by means of the mean squared error (MSE) is demonstrated in a statistic way, where the global convergence can be verified because the related convergence rate ρ is always less than 1.

Based on RIDA, further optimization and enhancement by well exploiting the conditional sampling are given in [22], where the modified randomized iterative detection algorithm (MRIDA) is given for better iteration convergence and efficiency. Typically, the iteration in MRIDA follows

$$\mathbf{x}^{k+1}=\mathbf{x}^k+\mathbf{S}_k(\mathbf{S}_k^H\mathbf{A}\mathbf{S}_k)^{-1}(\mathbf{S}_k^H\mathbf{b}-\mathbf{S}_k^H\mathbf{A}\mathbf{x}^k) \tag{20}$$

with $\mathbf{S}_k=\mathbf{D}^{-\frac{1}{2}}\mathbf{I}_{:,\mathcal{Q}_i}\in\mathbb{C}^{K\times q}$ and $\mathbf{S}_k\notin\{\mathbf{S}_{k-1},\ldots,\mathbf{S}_{k-r+1}\}$. Considering the block size $1\leq q\leq\sqrt{K}$, the overall computational complexity of MRIDA by updating all the elements of $\mathbf{x}$ in a full iteration is $O(K^2)$, which is comparable to the conventional iterative detection schemes. Note that obtaining the filtering matrix $\mathbf{A}=\mathbf{H}^H\mathbf{H}$ and $\mathbf{b}=\mathbf{H}^H\mathbf{y}\in\mathbb{C}^K$ is deemed as a preprocessing for these iterative detection schemes, where the related complexity is not considered here. Finally, the iterations results of RIDA and MRIDA are rounded according to (13) to yield the detection solution.

III. New randomized iterative detection algorithm

In order to solve the ILS problem in (2), the existing low-complexity iterative detection schemes try to firstly convert the original system model in (1) into the classic linear system in (7), where iterations (like Jacobi and GS) or random iterations (like RIDA and MRIDA) are carried out thereafter. However, during such an equivalent transformation from $\mathbf{y}=\mathbf{H}\mathbf{x}+\mathbf{n}$ to $\mathbf{A}\mathbf{x}=\mathbf{b}$, the symmetric positive matrix $\mathbf{A}=\mathbf{H}^H\mathbf{H}$ (or $\mathbf{A}=\mathbf{H}^H\mathbf{H}+\sigma^2\mathbf{I}$) and vector $\mathbf{b}=\mathbf{H}^H\mathbf{y}$ need to be calculated as a preprocessing stage, so that the matrix splitting in iterative methods or the column partition in random iteration can be performed. In fact, due to $N\geq K$, matrix multiplication about $\mathbf{A}=\mathbf{H}^H\mathbf{H}$ with computational complexity (i.e., $O(NK^2)$) is also computational expensive compared to the matrix inversion. This is accordance with the computational complexity of the traditional linear detection schemes like ZF in (3) and MMSE in (4), which consists of matrix multiplication with complexity $O(NK^2)$ and matrix inversion with complexity $O(K^3)$.

However, most of current works about low-complexity iterative detection mainly focus on how to effectively reduce the complexity cost of matrix inversion $\mathbf{A}^{-1}$. For this reason, a new randomized iterative detection algorithm (NRIDA) is proposed, which is directly designed for the overdetermined detection system in (1) rather than the linear system in (7). By doing this, we show the preprocessing stage (i.e., the matrix multiplication about $\mathbf{A}$ and $\mathbf{b}$ can be avoided without any performance loss. Intuitively, this corresponds to the complexity reduction with respect to the matrix multiplication $\mathbf{H}^H\mathbf{H}$ in linear detection, where considerable complexity cost is saved accordingly.

A. Algorithm Description

Specifically, based on the randomized iteration in (14), by letting the auxiliary symmetric positive definite matrix $\mathbf{U}=\mathbf{A}$ and $\mathbf{S}_k\in\{\mathbf{I}_{:,\mathcal{Q}_1},\ldots,\mathbf{I}_{:,\mathcal{Q}_r}\}$, the random iteration in the proposed NRIDA is designed as follows

$$\mathbf{x}^{k+1}=\mathbf{x}^k+\mathbf{W}_k(\mathbf{G}_k^H\mathbf{H}\mathbf{W}_k)^{-1}\mathbf{G}_k^H(\mathbf{y}-\mathbf{H}\mathbf{x}^k) \tag{21}$$

with

$$\mathbf{W}_k\triangleq(\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H\mathbf{G}_k, \tag{22}$$

where $\mathbf{G}_k=\mathbf{H}\mathbf{S}_k=\mathbf{H}\mathbf{I}_{:,\mathcal{Q}_i}=\mathbf{H}_{:,\mathcal{Q}_i}\in\mathbb{C}^{N\times q}$ is randomly sampled from the distribution $\mathcal{D}$, i.e.,

$$p_i=\mathcal{D}(\mathbf{G}_k=\mathbf{H}_{:,\mathcal{Q}_i}). \tag{23}$$

Here, the same block size $|\mathcal{Q}_1|=\ldots=|\mathcal{Q}_r|=K/r=q$ is applied. Similar to RIDA, the indices in set $\mathcal{Q}_i$, $1\leq i\leq r$ are determined initially while a simple way is to group these indices in a sequential order, i.e., $\mathbf{G}_k=\mathbf{H}_{:,\mathcal{Q}_i}=[\mathbf{h}_{(i-1)q+1},\ldots,\mathbf{h}_{iq}]$, $\mathcal{Q}_i=\{(i-1)q+1,\ldots,iq\}$. Based on $\mathcal{Q}_i$, the sampling about $\mathbf{G}_k$ is performed block by block.

Specifically, given the detection model in (1), To be more specific, given $\mathbf{G}_k=\mathbf{H}_{:,\mathcal{Q}_i}$, it follows that

$$\mathbf{W}_k=\mathbf{I}_{:,\mathcal{Q}_i}\in\mathbb{C}^{K\times q} \tag{24}$$

while the random iteration in (21) can be rewritten as

$$\mathbf{x}^{k+1}=\mathbf{x}^k+\mathbf{I}_{:,\mathcal{Q}_i}(\mathbf{H}_{:,\mathcal{Q}_i}^H\mathbf{H}_{:,\mathcal{Q}_i})^{-1}\mathbf{H}_{:,\mathcal{Q}_i}^H(\mathbf{y}-\mathbf{H}\mathbf{x}^k) \tag{25}$$

with random $\mathcal{Q}_i$. Consequently, by iterating $\mathbf{x}$ according to (25), the desired detection solution $\mathbf{x}^*=\mathbf{H}^\dagger\mathbf{y}=\mathbf{x}_{\text{zf}}$ in (6) can be approximated asymptotically, where the globally exponential convergence will be demonstrated in the following. As for the choice of the initial setup $\mathbf{x}^0$, it can be an arbitrary point from $\mathcal{X}^K$ and we set $\mathbf{x}^0=\mathbf{0}$ for simplicity. Finally, after $k=L$ times iteration, the iteration results $\mathbf{x}^L$ is rounded based on the constellation $\mathcal{X}^K$ in (13) to output the detection solution.

Another point should be noticed is the choice of the block size q, where a small $q\ll N$ is preferable due to a small size of matrix inversion $(\mathbf{H}_{:,\mathcal{Q}_i}^H\mathbf{H}_{:,\mathcal{Q}_i})^{-1}$ in (25). Nevertheless, a larger size q also implies more elements of $\mathbf{x}$ are updated at each iteration, which naturally results in a better convergence performance. Clearly, there is a latent trade-off in the proposed NRIDA between convergence and complexity about the choice of q, and here the choice $1<q\leq\sqrt{K}$ is recommended. Further investigation about the choice of q will be one of our work in future.

Now we go through the computational complexity of the proposed NRIDA. In particular, the computational complexity of computing $(\mathbf{H}_{:,\mathcal{Q}_i}^H\mathbf{H}_{:,\mathcal{Q}_i})^{-1}$ is q^3+q^2N; the multiplication among $\mathbf{I}_{:,\mathcal{Q}_i}$ and $(\mathbf{H}_{:,\mathcal{Q}_i}^H\mathbf{H}_{:,\mathcal{Q}_i})^{-1}$ and $\mathbf{H}_{:,\mathcal{Q}_i}^H$ costs q^2K+qKN while the computational complexity of multiplying it with $(\mathbf{y}-\mathbf{H}\mathbf{x}^k)$ requires $2KN$. Hence, the total computational complexity of NRIDA at each iteration (i.e., compute $\mathbf{x}^{k+1}$ in (25)) is $q^3+q^2N+q^2K+qKN+2KN$. Given the choice $1<q\leq\sqrt{K}$, the complexity of NRIDA at each iteration is no more than $O(K^{1.5}N)$. Note that the complicated matrix multiplication like $\mathbf{H}^H\mathbf{H}$ is not involved in the proposed NRIDA so that significant computational complexity is reduced.

B. Complexity Reduction

The computational complexity of NRIDA can be further reduced by making use of the inherent structure of $\mathbf{I}_{:,\mathcal{Q}_i}$. As shown in (16), the matrix $\mathbf{I}_{:,\mathcal{Q}_i}^H$ can be depicted in the following

Algorithm 1 The proposed NRIDA for Large-scale MIMO Systems

Require: $\mathbf{H}, \mathbf{y}$, $\mathbf{x}^0 = \mathbf{0}$, L
Ensure: near linear detection solution $\widehat{\mathbf{x}}$
1: **for** $k = 1, \ldots, L$ **do**
2:     **for** $t = 1, \ldots, K/q$ **do**
3:         randomly sample index set $\mathcal{Q}_i$ according to (23)
4:         update $\mathbf{x}$ based on (25)
5:     **end for**
6: **end for**
7: output $\widehat{\mathbf{x}} = \lceil \mathbf{x}^L \rfloor_Q \in \mathcal{X}^K$

way:

$$\mathbf{I}^H_{:,\mathcal{Q}_i} = \begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \cdots & \vdots & 0 & \ddots & \ddots & \vdots & \vdots & \cdots & \vdots \\ \vdots & \cdots & \vdots & \vdots & \ddots & \ddots & 0 & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{bmatrix}. \tag{26}$$

Intuitively, the operations of $\mathbf{I}_{:,\mathcal{Q}_i}$ and $\mathbf{I}^H_{:,\mathcal{Q}_i}$ are essentially determined by the $q \times q$ nonzero matrix within it. Meanwhile, every row or column of this submatrix only has a nonzero element, i.e., 1, where all other elements in it are 0 as well. From it, further complexity reduction with respect to the proposed NIRDA can be achieved.

In particular, given the structure of $\mathbf{I}_{:,\mathcal{Q}_i}$ in (26), the multiplication of $\mathbf{I}_{:,\mathcal{Q}_i}$ and $(\mathbf{H}^H_{:,\mathcal{Q}_i}\mathbf{H}_{:,\mathcal{Q}_i})^{-1}$ can be finished instantaneous by the simple shift and matrix augmentation with 0, e.g.,

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ g & h & i \\ 0 & 0 & 0 \\ a & b & c \\ d & e & f \end{bmatrix}. \tag{27}$$

Since the complicated matrix operations can be avoided, the complexity of multiplying $\mathbf{I}_{:,\mathcal{Q}_i}$ with $(\mathbf{H}^H_{:,\mathcal{Q}_i}\mathbf{H}_{:,\mathcal{Q}_i})^{-1}$ turns out to be negligible. On the other hand, due to the fact that the $K \times q$ matrix $\mathbf{I}_{:,\mathcal{Q}_i}(\mathbf{H}^H_{:,\mathcal{Q}_i}\mathbf{H}_{:,\mathcal{Q}_i})^{-1}$ is essentially operated by the $q \times q$ nonzero matrix in it (i.e., all the other columns in it are 0), the computational complexity of multiplying it with $\mathbf{H}^H_{:,\mathcal{Q}_i}$ can be reduced to q^2N. Subsequently, as the $K \times N$ matrix $\mathbf{I}_{:,\mathcal{Q}_i}(\mathbf{H}^H_{:,\mathcal{Q}_i}\mathbf{H}_{:,\mathcal{Q}_i})^{-1}\mathbf{H}^H_{:,\mathcal{Q}_i}$ contains $K - q$ zero rows, the computational complexity of multiplying it with $(\mathbf{y} - \mathbf{H}\mathbf{x}^k)$ can be reduced as $NK + qN$. Overall, the total computational complexity of the proposed NRIDA can be reduced to $q^3 + qN + 2q^2N + KN$, which is $O(KN)$ given the choice of $q = \sqrt{K}$.

Here, we point out that at each of iteration in NRIDA only q elements of $\mathbf{x}$ are updated according to (25). For this reason, to achieve a fair comparison with other low-complexity iterative detection schemes, a full iteration which contains K/q iterations is employed by NRIDA. Consequently, with $q = \sqrt{K}$, this leads to the complexity $O(K^{1.5}N)$ of a full iteration, which is competitive compared to other iterative detection schemes in both theoretic and simulation results. Note that both K/q and q are integers in practice.

Overall, the operations of the proposed NRIDA for uplink large-scale MIMO systems is summarized in Algorithm 1. Typically, the loop between step 1 and 6 denotes a full iteration of NRIDA, which contains K/q times random iteration of (25). For a better understanding, the related complexity comparison of various low-complexity iterative detection schemes is presented in Table I, where k denotes the number of full iteration. Clearly, besides the global convergence, considerable computational complexity is also saved by the proposed NRIDA without incurring the preprocessing about the matrix multiplication of $\mathbf{H}^H\mathbf{H}$.

IV. Convergence Analysis

In this section, the convergence of the proposed NRIDA is investigated in terms of the statistic mean squared error (MSE), where the globally exponential convergence is demonstrated in detail. Meanwhile, the choice of the sampling distribution $\mathcal{D}$ for the random iteration is also studied.

A. Globally Exponential Convergence

To start with, considering the full column rank channel matrix $\mathbf{H}$[1], the multiplication $\mathbf{H}\mathbf{w}$ with vector $\mathbf{w} \in \mathbb{C}^K$ is $\mathbf{0}$ if and only if $\mathbf{w} = \mathbf{0}$. Therefore, it is clear to see that the matrix $\mathbf{H}^H\mathbf{H}$ is positive definite because of

$$\mathbf{w}^H\mathbf{H}^H\mathbf{H}\mathbf{w} = (\mathbf{H}\mathbf{w})^H\mathbf{H}\mathbf{w} > 0 \text{ for } \mathbf{w} \neq \mathbf{0}. \tag{28}$$

Clearly, since $\mathbf{H}^H\mathbf{H}$ is also symmetric, it is essentially a positive definite *Hermitain* matrix.

Then, given the symmetric positive definite matrix $\mathbf{H}^H\mathbf{H}$, we define matrix $\mathbf{V} \in \mathbb{C}^{K \times K}$ as

$$\mathbf{V}\mathbf{V} = \mathbf{H}^H\mathbf{H}, \tag{29}$$

where $\mathbf{V}$ can be easily calculated based on eigenvalue matrix and eigenvector matrix [31][2]. Besides, since $\mathbf{V}$ is symmetric as well, the following relationship holds

$$\mathbf{V}\mathbf{V} = \mathbf{V}\mathbf{V}^H = \mathbf{V}^H\mathbf{V}. \tag{30}$$

Next, based on the matrix $\mathbf{V}$, we have the following result, where the symmetric matrix $\mathbf{Z} \in \mathbb{C}^{K \times K}$ is defined by

$$\mathbf{Z} \triangleq \mathbf{I}_{:,\mathcal{Q}_i}(\mathbf{H}^H_{:,\mathcal{Q}_i}\mathbf{H}_{:,\mathcal{Q}_i})^{-1}\mathbf{I}^H_{:,\mathcal{Q}_i}. \tag{31}$$

Theorem 2. *As for the uplink large-scale MIMO detection, let $\mathbf{S}_k = \mathbf{H}_{:,\mathcal{Q}_i}$ be sampled randomly according to the distribution $\mathcal{D}$, the proposed NRIDA based on (25) converges by*

$$E[\|\mathbf{V}(\mathbf{x}^k - \mathbf{x}^*)\|^2] \leq \rho^k \|\mathbf{V}(\mathbf{x}^0 - \mathbf{x}^*)\|^2 \tag{32}$$

with globally exponential convergence rate

$$\rho = 1 - \lambda_{\min}(\mathbf{H}^H\mathbf{H}E[\mathbf{Z}]) < 1. \tag{33}$$

Proof: According to (31), the random iteration in (25) can be rewritten as

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \mathbf{Z}\mathbf{H}^H\mathbf{H}(\mathbf{x}^* - \mathbf{x}^k), \tag{34}$$

[1]Note that a full column rank channel matrix $\mathbf{H}$ is a general configuration for MIMO detection, otherwise the solution for $\mathbf{x}$ will not be unique. For example, any two UEs can not have the same channel responses in scale, e.g., $\mathbf{h}_i \neq \alpha\mathbf{h}_j$, α is a constant.

[2]$\mathbf{V}$ is applied to depict the convergence process of NRIDA but its computation is not needed in the implementation of NRIDA.

which is further expressed as

$$\mathbf{x}^{k+1} - \mathbf{x}^* = (\mathbf{I} - \mathbf{Z}\mathbf{H}^H\mathbf{H})(\mathbf{x}^k - \mathbf{x}^*). \tag{35}$$

Next, for a better presentation, by denoting $\mathbf{r}_k = \mathbf{x}^k - \mathbf{x}^*$, we have

$$\|\mathbf{V}(\mathbf{x}^k - \mathbf{x}^*)\|^2 = \|\mathbf{V}\mathbf{r}_k\|^2 \tag{36}$$

$$\begin{aligned}
&= \mathbf{r}_k^H \mathbf{V}^H \mathbf{V} \mathbf{r}_k \\
&\overset{(a)}{=} \mathbf{r}_{k-1}^H (\mathbf{I} - \mathbf{Z}\mathbf{H}^H\mathbf{H})^H \mathbf{V}^H \mathbf{V} (\mathbf{I} - \mathbf{Z}\mathbf{H}^H\mathbf{H}) \mathbf{r}_{k-1} \\
&= \mathbf{r}_{k-1}^H (\mathbf{V} - \mathbf{V}\mathbf{Z}\mathbf{H}^H\mathbf{H})^H (\mathbf{V} - \mathbf{V}\mathbf{Z}\mathbf{H}^H\mathbf{H}) \mathbf{r}_{k-1} \\
&= \mathbf{r}_{k-1}^H (\mathbf{V}^H\mathbf{V} - \mathbf{V}^H\mathbf{V}\mathbf{Z}\mathbf{H}^H\mathbf{H} - \mathbf{H}^H\mathbf{H}\mathbf{Z}^H\mathbf{V}^H\mathbf{V} \\
&\qquad + \mathbf{H}^H\mathbf{H}\mathbf{Z}^H\mathbf{V}^H\mathbf{V}\mathbf{Z}\mathbf{H}^H\mathbf{H}) \mathbf{r}_{k-1} \\
&\overset{(b)}{=} \mathbf{r}_{k-1}^H (\mathbf{V}^H\mathbf{V} - \mathbf{V}^H\mathbf{V}\mathbf{Z}\mathbf{H}^H\mathbf{H}) \mathbf{r}_{k-1}.
\end{aligned} \tag{37}$$

Here, the equality (a) holds due to (35), the equality (b) comes from the fact

$$\mathbf{Z}^H\mathbf{V}^H\mathbf{V}\mathbf{Z} = \mathbf{Z} \tag{38}$$

and $\mathbf{V}^H\mathbf{V} = \mathbf{H}^H\mathbf{H}$.

On the other hand, based on *law of total probability* (i.e., $E[E[A|B]] = E[A]$), it follows that

$$E[\|\mathbf{V}\mathbf{r}_k\|^2] = E[E[\|\mathbf{V}\mathbf{r}_k\|^2|\mathbf{r}_{k-1}]]. \tag{39}$$

Then, based on (37), we have

$$\begin{aligned}
E[\|\mathbf{V}\mathbf{r}_k\|^2|\mathbf{r}_{k-1}] &= E[\mathbf{r}_{k-1}^H(\mathbf{V}^H\mathbf{V} - \mathbf{V}^H\mathbf{V}\mathbf{Z}\mathbf{H}^H\mathbf{H})\mathbf{r}_{k-1}] \\
&= E[\mathbf{r}_{k-1}^H(\mathbf{V}^H\mathbf{V} - \mathbf{V}^H\mathbf{V}\mathbf{Z}\mathbf{V}^H\mathbf{V})\mathbf{r}_{k-1}] \\
&= E[\mathbf{r}_{k-1}^H\mathbf{V}^H(\mathbf{I} - \mathbf{V}\mathbf{Z}\mathbf{V}^H)\mathbf{V}\mathbf{r}_{k-1}] \\
&= E[\langle(\mathbf{I} - \mathbf{V}\mathbf{Z}\mathbf{V}^H)\mathbf{V}\mathbf{r}_{k-1}, \mathbf{V}\mathbf{r}_{k-1}\rangle] \\
&= \langle(\mathbf{I} - \mathbf{V}E[\mathbf{Z}]\mathbf{V}^H)\mathbf{V}\mathbf{r}_{k-1}, \mathbf{V}\mathbf{r}_{k-1}\rangle \\
&\le \|\mathbf{I} - \mathbf{V}E[\mathbf{Z}]\mathbf{V}^H\| \cdot \|\mathbf{V}\mathbf{r}_{k-1}\|^2 \\
&\overset{(c)}{=} \lambda_{\max}(\mathbf{I} - \mathbf{V}E[\mathbf{Z}]\mathbf{V}^H)\|\mathbf{V}\mathbf{r}_{k-1}\|^2 \\
&= (1 - \lambda_{\min}(\mathbf{V}E[\mathbf{Z}]\mathbf{V}^H))\|\mathbf{V}\mathbf{r}_{k-1}\|^2 \\
&= (1 - \lambda_{\min}(\mathbf{V}\mathbf{V}^H E[\mathbf{Z}]))\|\mathbf{V}\mathbf{r}_{k-1}\|^2 \\
&= (1 - \lambda_{\min}(\mathbf{H}^H\mathbf{H}E[\mathbf{Z}]))\|\mathbf{V}\mathbf{r}_{k-1}\|^2 \\
&= \rho\|\mathbf{V}\mathbf{r}_{k-1}\|^2,
\end{aligned} \tag{40}$$

where the change in (c) from operator norm into spectral radius holds because of the symmetry of $\mathbf{I} - \mathbf{V}E[\mathbf{Z}]\mathbf{V}^H$.

Subsequently, by simply substituting (40) into (39), we can arrive at

$$\begin{aligned}
E[\|\mathbf{V}(\mathbf{x}^k - \mathbf{x}^*)\|^2] &= E[E[\|\mathbf{V}(\mathbf{x}^k - \mathbf{x}^*)\|^2|\mathbf{r}_{k-1}]] \\
&\le \rho E[\|\mathbf{V}(\mathbf{x}^{k-1} - \mathbf{x}^*)\|^2] \\
&\le \cdots \\
&\le \rho^k E[\|\mathbf{V}(\mathbf{x}^0 - \mathbf{x}^*)\|^2] \\
&= \rho^k \|\mathbf{V}(\mathbf{x}^0 - \mathbf{x}^*)\|^2
\end{aligned} \tag{41}$$

where $\mathbf{x}^0$ is set as an initial setup.

Meanwhile, given the full rank matrix $\mathbf{H}$ as well as the full row rank matrix $\mathbf{H}_{:,\mathcal{Q}_i}^H$, the expectation of symmetric matrix $\mathbf{Z}$ can be demonstrated to be positive definite as

$$\begin{aligned}
E[\mathbf{Z}] &= \sum_{i=1}^r p_i \mathbf{I}_{:,\mathcal{Q}_i}(\mathbf{H}_{:,\mathcal{Q}_i}^H\mathbf{H}_{:,\mathcal{Q}_i})^{-1}\mathbf{I}_{:,\mathcal{Q}_i}^H \\
&= \left(\sum_{i=1}^r p_i^{\frac{1}{2}} \mathbf{I}_{:,\mathcal{Q}_i}(\mathbf{H}_{:,\mathcal{Q}_i}^H\mathbf{H}_{:,\mathcal{Q}_i})^{-\frac{1}{2}}(\mathbf{H}_{:,\mathcal{Q}_i}^H\mathbf{H}_{:,\mathcal{Q}_i})^{-\frac{1}{2}}\mathbf{I}_{:,\mathcal{Q}_i}^H p_i^{\frac{1}{2}}\right) \\
&= (\mathbf{I}\mathbf{J})(\mathbf{J}\mathbf{I}^H) \\
&= \mathbf{J}^2
\end{aligned} \tag{42}$$

with the invertible matrix $\mathbf{J} \in \mathbb{C}^{K\times K}$, i.e.,

$$\mathbf{J} = \mathrm{diag}(p_1^{\frac{1}{2}}(\mathbf{H}_{:,\mathcal{Q}_1}^H\mathbf{H}_{:,\mathcal{Q}_1})^{-\frac{1}{2}}, \ldots, p_r^{\frac{1}{2}}(\mathbf{H}_{:,\mathcal{Q}_r}^H\mathbf{H}_{:,\mathcal{Q}_r})^{-\frac{1}{2}}), \tag{43}$$

which is block diagonal.

Therefore, because of the symmetric positive definite matrix $E[\mathbf{Z}]$, it follows that

$$\lambda_{\min}(E[\mathbf{Z}]) > 0. \tag{44}$$

Since $\mathbf{H}^H\mathbf{H}$ is also symmetric positive definite, we can get $\lambda_{\min}(\mathbf{H}^H\mathbf{H}E[\mathbf{Z}]) > 0$ so as to

$$\rho = 1 - \lambda_{\min}(\mathbf{H}^H\mathbf{H}E[\mathbf{Z}]) < 1, \tag{45}$$

thus completing the proof. ■

From Theorem 2, the proposed NRIDA converges to the desired detection solution $\mathbf{x}^*$ of (6) in an exponential way. Note that the convergence of NRIDA is always ensured (i.e., $\rho < 1$), making it well fitted to the various scenarios of large-scale MIMO systems. Typically, to make sure the approximation error of the proposed random iteration smaller than a certain value

$$E[\|\mathbf{V}(\mathbf{x}^k - \mathbf{x}^*)\|^2] \le \epsilon\|\mathbf{V}(\mathbf{x}^0 - \mathbf{x}^*)\|^2 \tag{46}$$

with $0 < \epsilon < 1$, the iteration number k should satisfy

$$k \ge \frac{1}{1-\rho}\log\left(\frac{1}{\epsilon}\right), \tag{47}$$

where the inequality $\ln(1-\delta) < -\delta$ for $0 < \delta < 1$ is employed. Hence, a tractable random iteration can be achieved by flexibly adjusting k. Note that a better choice of $\mathbf{x}^0$ also enables a positive impact upon the convergence in (32) so that a closer choice of $\mathbf{x}^0$ to the detection solution $\mathbf{x}^*$ is preferable for the proposed NRIDA.

B. The Choice of Sampling Distribution $\mathcal{D}$

As for the choice of the sampling distribution $\mathcal{D}$ in (23), a natural solution is to apply the uniform distribution, where the following result can be obtained.

Corollary 1. *With $\mathbf{G}_k = \mathbf{H}_{:,\mathcal{Q}_i}$ following the uniform sampling probability*

$$p_i = \frac{1}{r}, \tag{48}$$

the proposed NRIDA converges by

$$E[\|\mathbf{V}(\mathbf{x}^k - \mathbf{x}^*)\|^2] \le \rho_{\text{uniform}}^k\|\mathbf{V}(\mathbf{x}^0 - \mathbf{x}^*)\|^2 \tag{49}$$

with

$$\rho_{\text{uniform}} = 1 - \frac{1}{r}\cdot\min_i\left\{\frac{\lambda_{\min}(\mathbf{H}^H\mathbf{H})}{\lambda_{\max}(\mathbf{H}_{:,\mathcal{Q}_i}^H\mathbf{H}_{:,\mathcal{Q}_i})}\right\}. \tag{50}$$

TABLE I
COMPARISON OF VARIOUS LOW-COMPLEXITY ITERATIVE DETECTION SCHEMES.

	Complexity of Preprocessing	Complexity of Iterative Detection	Global Convergence
Jacobi [9]	$O(NK^2)$	$O(K^2 \cdot k)$	
Richardson [19]	$O(NK^2)$	$O(K^2 \cdot k)$	
Gauss Seidel [11]	$O(NK^2)$	$O(K^2 \cdot k)$	✓
SOR [32]	$O(NK^2)$	$O(K^2 \cdot k)$	
RIDA [22]	$O(NK^2)$	$O(K^{2.5} \cdot k)$	✓
MRIDA [22]	$O(NK^2)$	$O(K^2 \cdot k)$	✓
NRIDA	No need	$O(K^{1.5}N \cdot k)$	✓

Proof: To start with, from (42), we can obtain that

$$\lambda_{\text{min}}(\mathbf{H}^H\mathbf{H}E[\mathbf{Z}]) = \lambda_{\text{min}}(\mathbf{H}^H\mathbf{H}\mathbf{J}^2) \overset{(d)}{\geq} \lambda_{\text{min}}(\mathbf{H}^H\mathbf{H})\lambda_{\text{min}}(\mathbf{J}^2) \tag{51}$$

where the inequality (d) holds due to $\lambda_{\text{min}}(\mathbf{EF}) \geq \lambda_{\text{min}}(\mathbf{E})\lambda_{\text{min}}(\mathbf{F})$ for positive definite matrices $\mathbf{E}$, $\mathbf{F}$. Meanwhile, according to (43), it follows that

$$\mathbf{J}^2 = \text{diag}(p_1(\mathbf{H}^H_{:,\mathcal{Q}_1}\mathbf{H}_{:,\mathcal{Q}_1})^{-1}, \ldots, p_r(\mathbf{H}^H_{:,\mathcal{Q}_r}\mathbf{H}_{:,\mathcal{Q}_r})^{-1}). \tag{52}$$

Then, given the uniform sampling probability $p_i = 1/r$, we have

$$\lambda_{\text{min}}(\mathbf{J}^2) = \frac{1}{r} \cdot \frac{1}{\lambda_{\text{max}}(\mathbf{H}^H_{:,\mathcal{Q}_i}\mathbf{H}_{:,\mathcal{Q}_i})}, \tag{53}$$

Therefore, based on (51) and (53), the convergence rate ρ in (33) is upper bounded by

$$\begin{aligned}\rho &= 1 - \lambda_{\text{min}}(\mathbf{H}^H\mathbf{H}E[\mathbf{Z}]) \\ &\leq 1 - \lambda_{\text{min}}(\mathbf{H}^H\mathbf{H})\lambda_{\text{min}}(\mathbf{J}^2) \\ &\leq 1 - \frac{1}{r} \cdot \min_i \left\{ \frac{\lambda_{\text{min}}(\mathbf{H}^H\mathbf{H})}{\lambda_{\text{max}}(\mathbf{H}^H_{:,\mathcal{Q}_i}\mathbf{H}_{:,\mathcal{Q}_i})} \right\} \\ &= \rho_{\text{uniform}},\end{aligned} \tag{54}$$

which completes the proof. ■

According to Corollary 1, the convergence rate depends on the block size q (i.e., $q = K/r = |\mathcal{Q}_i|$) as well as the partition of $\mathbf{H}$ denoted by $\mathbf{H}_{:,\mathcal{Q}_i}$. To further investigate this point, an alternative sampling choice is presented in the following.

Corollary 2. *With $\mathbf{G}_k = \mathbf{H}_{:,\mathcal{Q}_i}$ following the sampling probability*

$$p_i = \frac{\|\mathbf{H}_{:,\mathcal{Q}_i}\|_F^2}{\|\mathbf{H}\|_F^2}, \tag{55}$$

the proposed NRIDA converges by

$$E[\|\mathbf{V}(\mathbf{x}^k - \mathbf{x}^*)\|^2] \leq \rho^k_{\text{norm}}\|\mathbf{V}(\mathbf{x}^0 - \mathbf{x}^*)\|^2 \tag{56}$$

with

$$\rho_{\text{norm}} = 1 - \alpha \cdot \frac{\lambda_{\text{min}}(\mathbf{H}^H\mathbf{H})}{\text{Tr}(\mathbf{H}^H\mathbf{H})}, \tag{57}$$

where

$$\alpha \triangleq \min_i \left\{ \frac{\text{Tr}(\mathbf{H}^H_{:,\mathcal{Q}_i}\mathbf{H}_{:,\mathcal{Q}_i})}{\lambda_{\text{max}}(\mathbf{H}^H_{:,\mathcal{Q}_i}\mathbf{H}_{:,\mathcal{Q}_i})} \right\} \geq 1. \tag{58}$$

Proof: Considering the sampling probability p_i in (55), because of $\|\mathbf{B}\|_F^2 = \sum_i^m \sum_j^n |b_{i,j}|^2 = \text{Tr}(\mathbf{B}^H\mathbf{B})$, we have

$$\begin{aligned}\lambda_{\text{min}}(\mathbf{J}^2) &= \frac{1}{\text{Tr}(\mathbf{H}^H\mathbf{H})} \min_i \left\{ \frac{\text{Tr}(\mathbf{H}^H_{:,\mathcal{Q}_i}\mathbf{H}_{:,\mathcal{Q}_i})}{\lambda_{\text{max}}(\mathbf{H}^H_{:,\mathcal{Q}_i}\mathbf{H}_{:,\mathcal{Q}_i})} \right\} \\ &= \frac{\alpha}{\text{Tr}(\mathbf{H}^H\mathbf{H})},\end{aligned} \tag{59}$$

with $\alpha \geq 1$ due to

$$\text{Tr}(\mathbf{A}) = \sum_i \lambda_i(\mathbf{A}) \geq \lambda_{\text{max}}(\mathbf{A}) \geq \lambda_{\text{min}}(\mathbf{A}). \tag{60}$$

Then, according to (51) and (59), the convergence rate ρ in (33) is upper bounded by

$$\rho \leq 1 - \alpha \cdot \frac{\lambda_{\text{min}}(\mathbf{H}^H\mathbf{H})}{\text{Tr}(\mathbf{H}^H\mathbf{H})} = \rho_{\text{norm}}, \tag{61}$$

completing the proof. ■

From (58), the coefficient $\alpha \geq 1$ is chiefly determined by the block size q (i.e., $|\mathcal{Q}_i|$). More specifically, it is clear to see that $\alpha = 1$ if $q = 1$ while a larger size α can be achieved with the increment of q. Therefore, the convergence gain by the block operation can be confirmed explicitly. On the other hand, because the matrix $\mathbf{H}^H\mathbf{H}$ is symmetric positive definite, all its eigenvalues are larger than 0, which means the global convergence $\rho < 1$ still holds under the sampling probabilities in (48) and (55). Meanwhile, from Corollary 2, it is clear to see that the convergence rate ρ_{norm} heavily relies on the condition number of the matrix $\mathbf{H}^H\mathbf{H}$ (i.e., $\kappa = \lambda_{\text{max}}(\mathbf{H}^H\mathbf{H})/\lambda_{\text{min}}(\mathbf{H}^H\mathbf{H})$), where a smaller κ is highly desired for a better convergence performance.

For a better understanding, in Fig. 1, the convergence rates ρ_{uniform} and ρ_{norm} in the proposed NRIDA with both sampling probabilities in (48) and (55) are illustrated in a Monte Carlo way, where 64×128 large-scale MIMO systems are applied. Here, for the sake of efficient computation, in both uniform and norm sampling distributions the partition is carried out in

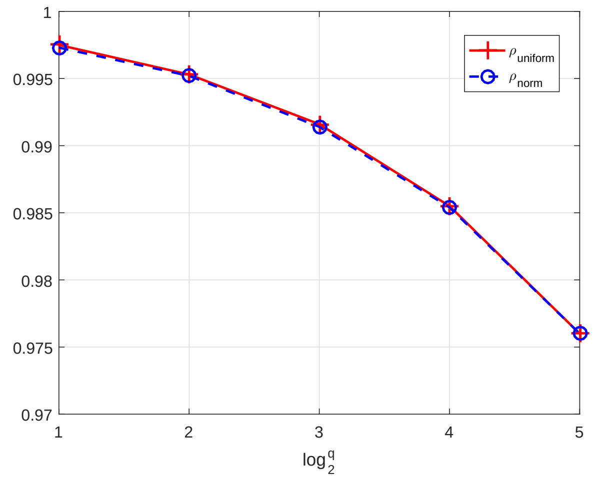

Fig. 1. Convergence comparison of NRIDA in 64×128 large-scale MIMO systems.

a forward sequential order, namely,

$$\mathbf{H}_{:,\mathcal{Q}_1}=[\mathbf{h}_1,\ldots,\mathbf{h}_q],\ldots,\mathbf{H}_{:,\mathcal{Q}_r}=[\mathbf{h}_{(r-1)q+1},\ldots,\mathbf{h}_K]. \quad (62)$$

In this way, the impact of partitioning $\mathbf{H}$ is removed from the comparison. Nevertheless, further optimization still can be made by seeking for the optimal partition in the expense of extra computational complexity. On one hand, both the convergence of NRIDA are nearly the same, making the uniform sampling a promising choice due to simplicity. On the other hand, as expected, with the increment of the block size $q = 2, 4, 8, 16, 32$, both the convergence become better gradually.

V. Convergence Enhancement

In this section, to improve the convergence performance, the concept of conditional sampling is introduced to the random iteration in the proposed NRIDA. By well exploiting it, a derandomized-like iteration can be achieved for both better convergence and efficiency.

A. Enhancement by Conditional Sampling

In particular, define the conditional sampling probability in the proposed NRIDA as

$$\begin{aligned}\overline{p}_i &\triangleq \mathcal{D}(\mathbf{G}_k=\mathbf{H}_{:,\mathcal{Q}_i}|\mathbf{G}_{k-1}=\mathbf{H}_{:,\mathcal{Q}_j}),\ i\neq j\\ &=\frac{p_i}{1-p_j},\ i\neq j,\end{aligned} \quad (63)$$

where the last sampling choice $\mathbf{G}_{k-1}$ is taken into account at the current sampling of $\mathbf{G}_k$. Clearly, by doing this, the sampling choice $\mathbf{G}_k=\mathbf{H}_{:,\mathcal{Q}_j}$ given $\mathbf{G}_{k-1}=\mathbf{H}_{:,\mathcal{Q}_j}$ will be avoided, which effectively improves the sampling diversity. In what follows, we show that more convergence gain can be obtained from the application of conditional sampling.

Typically, according to the condition sampling probability $\overline{p}_i$ in (63), the conditional expectation of $\mathbf{Z}$ based on the last sampling choice $\mathbf{G}_{k-1}=\mathbf{H}_{:,\mathcal{Q}_j}$ turns out to be

$$\begin{aligned}E[\mathbf{Z}|\mathbf{G}_{k-1}] &= \sum_{i=1,i\neq j}^{r}\overline{p}_i\mathbf{I}_{:,\mathcal{Q}_i}(\mathbf{H}^H_{:,\mathcal{Q}_i}\mathbf{H}_{:,\mathcal{Q}_i})^{-1}\mathbf{I}^H_{:,\mathcal{Q}_i}\\ &=\left(\sum_{i=1,i\neq j}^{r}\overline{p}_i^{\frac{1}{2}}\mathbf{I}_{:,\mathcal{Q}_i}(\mathbf{H}^H_{:,\mathcal{Q}_i}\mathbf{H}_{:,\mathcal{Q}_i})^{-\frac{1}{2}}(\mathbf{H}^H_{:,\mathcal{Q}_i}\mathbf{H}_{:,\mathcal{Q}_i})^{-\frac{1}{2}}\mathbf{I}^H_{:,\mathcal{Q}_i}\overline{p}_i^{\frac{1}{2}}\right)\\ &=(\overline{\mathbf{I}}\overline{\mathbf{J}})(\overline{\mathbf{J}}\overline{\mathbf{I}}^H),\end{aligned} \quad (64)$$

which is still symmetric positive definite. Here, matrix $\overline{\mathbf{I}} \in \mathbb{C}^{K\times(K-q)}$ is a partition of the identity matrix $\mathbf{I}$ by removing the related columns in the index set $\mathcal{Q}_j$, matrix $\overline{\mathbf{J}} \in \mathbb{C}^{(K-q)\times(K-q)}$ takes the form $\overline{\mathbf{J}} = \text{diag}(\overline{p}_1^{\frac{1}{2}}(\mathbf{H}^H_{:,\mathcal{Q}_1}\mathbf{H}_{:,\mathcal{Q}_1})^{-\frac{1}{2}},\ldots,\overline{p}_{j-1}^{\frac{1}{2}}(\mathbf{H}^H_{:,\mathcal{Q}_{j-1}}\mathbf{H}_{:,\mathcal{Q}_{j-1}})^{-\frac{1}{2}},\overline{p}_{j+1}^{\frac{1}{2}}(\mathbf{H}^H_{:,\mathcal{Q}_{j+1}}\mathbf{H}_{:,\mathcal{Q}_{j+1}})^{-\frac{1}{2}},\ldots,\overline{p}_r^{\frac{1}{2}}(\mathbf{H}^H_{:,\mathcal{Q}_r}\mathbf{H}_{:,\mathcal{Q}_r})^{-\frac{1}{2}})$.

Based on $E[\mathbf{Z}|\mathbf{G}_{k-1}]$ in (64), the globally exponential convergence of conditional random iteration in the proposed NRIDA can be easily demonstrated. Note that the related proof is omitted here, and more details can be found in Theorem 2.

Theorem 3. *As for the uplink large-scale MIMO detection, let $\mathbf{G}_k = \mathbf{H}_{:,\mathcal{Q}_i}$ be sampled randomly according to the conditional sampling probability $\overline{p}_i$ in (63), the proposed NRIDA converges by*

$$E[\|\mathbf{V}(\mathbf{x}^k-\mathbf{x}^*)\|^2] \leq \overline{\rho}\|\mathbf{V}(\mathbf{x}^{k-1}-\mathbf{x}^*)\|^2 \quad (65)$$

with globally exponential convergence rate

$$\overline{\rho} = 1-\lambda_{\text{min}}(\mathbf{H}^H\mathbf{H}E[\mathbf{Z}|\mathbf{G}_{k-1}]) < 1. \quad (66)$$

We then show that the proposed NRIDA with conditional sampling probability $\overline{p}_i$ has a better convergence performance than that with sampling probability p_i.

Corollary 3. *With the conditional sampling probability $\overline{p}_i$ in (63), the proposed NRIDA has a better convergence performance than that with the sampling probability p_i in (23) because of a smaller upper bound of the convergence rate.*

Proof: On one hand, the convergence rate under the conditional sampling probability $\overline{p}_i$ is upper bounded as

$$\begin{aligned}\overline{\rho} &= 1-\lambda_{\text{min}}(\mathbf{H}^H\mathbf{H}E[\mathbf{Z}|\mathbf{G}_{k-1}])\\ &\leq 1-\lambda_{\text{min}}(\mathbf{H}^H\mathbf{H})\lambda_{\text{min}}(\overline{\mathbf{J}}^2)\\ &\leq 1-\min_i\left\{\overline{p}_i\cdot\frac{\lambda_{\text{min}}(\mathbf{H}^H\mathbf{H})}{\lambda_{\text{max}}(\mathbf{H}^H_{:,\mathcal{Q}_i}\mathbf{H}_{:,\mathcal{Q}_i}|i\neq j)}\right\}.\end{aligned} \quad (67)$$

On the other hand, the convergence rate with sampling probability p_i is upper bounded as

$$\begin{aligned}\rho &= 1-\lambda_{\text{min}}(\mathbf{H}^H\mathbf{H}E[\mathbf{Z}])\\ &< 1-\min_i\left\{p_i\cdot\frac{\lambda_{\text{min}}(\mathbf{H}^H\mathbf{H})}{\lambda_{\text{max}}(\mathbf{H}^H_{:,\mathcal{Q}_i}\mathbf{H}_{:,\mathcal{Q}_i})}\right\}.\end{aligned} \quad (68)$$

Therefore, it is clear to see that $\overline{\rho}$ has a smaller upper bound

TABLE II
ILLUSTRATION OF THE CONVERGENCE RATE UPPER BOUND.

	$f=0$	$f=1$	$f=2$	$f=3$	$\cdots$	$f=r-1$
$q=2$	0.9975	0.9974	0.9973	0.9972	$\cdots$	0.9062
$q=4$	0.9953	0.9950	0.9946	0.9942	$\cdots$	0.9160
$q=8$	0.9916	0.9903	0.9887	0.9864	$\cdots$	0.9273
$q=16$	0.9855	0.9805	0.9703	0.9390	$\cdots$	0.9390

than ρ by

$$\min_i \left\{ \frac{\overline{p}_i}{\lambda_{\max}(\mathbf{H}^H_{:,\mathcal{Q}_i}\mathbf{H}_{:,\mathcal{Q}_i}|i \neq j)} \right\} > \min_i \left\{ \frac{p_i}{\lambda_{\max}(\mathbf{H}^H_{:,\mathcal{Q}_i}\mathbf{H}_{:,\mathcal{Q}_i})} \right\}, \tag{69}$$

where $\lambda_{\max}(\mathbf{H}^H_{:,\mathcal{Q}_i}\mathbf{H}_{:,\mathcal{Q}_i}|i \neq j) \leq \lambda_{\max}(\mathbf{H}^H_{:,\mathcal{Q}_i}\mathbf{H}_{:,\mathcal{Q}_i})$ and $\overline{p}_i > p_i$ can be verified in a straightforward way. ■

B. Extension by Multi-Step Conditional Sampling

To further exploit the convergence gain brought by conditional sampling, more previous sampling results can be taken into account so that the multi-step conditional sampling probability in the proposed NRIDA is defined as

$$\overline{p}_i^f \triangleq \mathcal{D}(\mathbf{G}_k = \mathbf{H}_{:,\mathcal{Q}_i}|\mathbf{G}_{k-1}, \ldots, \mathbf{G}_{k-f}) \tag{70}$$

with $\mathbf{H}_{:,\mathcal{Q}_i} \notin \{\mathbf{G}_{k-1}, \ldots, \mathbf{G}_{k-f}\}$, where $1 \leq f \leq r-1$ indicates the length of the multi-step conditional sampling. Clearly, the conditional sampling probability $\overline{p}_i$ in (63) can be viewed as a special case of $\overline{p}_i^f$ with $f = 1$. According to the multi-step conditional sampling probability $\overline{p}_i^f$, we arrive at the following result, and the proof is omitted for the sake of simplicity.

Theorem 4. *As for the uplink large-scale MIMO detection, let $\mathbf{G}_k = \mathbf{H}_{:,\mathcal{Q}_i}$ be randomly sampled following the conditional sampling probability $\overline{p}_i^f$ in (70), the proposed NRIDA converges by*

$$E[\|\mathbf{V}(\mathbf{x}^k - \mathbf{x}^*)\|^2] \leq \overline{\rho}^f \|\mathbf{V}(\mathbf{x}^{k-1} - \mathbf{x}^*)\|^2 \tag{71}$$

with globally exponential convergence rate

$$\overline{\rho}^f = 1 - \lambda_{\min}(\mathbf{H}^H\mathbf{H}E[\mathbf{Z}|\mathbf{G}_{k-1}, \ldots, \mathbf{G}_{k-f}]) < 1. \tag{72}$$

Based on Theorem 4, the convergence gain of multi-step conditional sampling can be verified as follows. Here, the proof is omitted where the related details can be found in Corollary 3.

Corollary 4. *With the increment of $1 \leq f \leq r-1$, the convergence of the proposed NRIDA with multi-step conditional sampling probability $\overline{p}_i^f$ in (70) improves gradually due to a smaller upper bound of convergence rate.*

According to Corollary 4, we can see that the smallest upper bounded of the convergence rate is obtained when $f = r-1$ is applied. More specifically, given $\mathbf{G}_{k-1}, \ldots, \mathbf{G}_{k-r+1}$, $k > r-1$, we have

$$\overline{\rho}^{r-1} = 1 - \lambda_{\min}(\mathbf{H}^H\mathbf{H}E[\mathbf{Z}|\mathbf{G}_{k-1}, \ldots, \mathbf{G}_{k-r+1}])$$

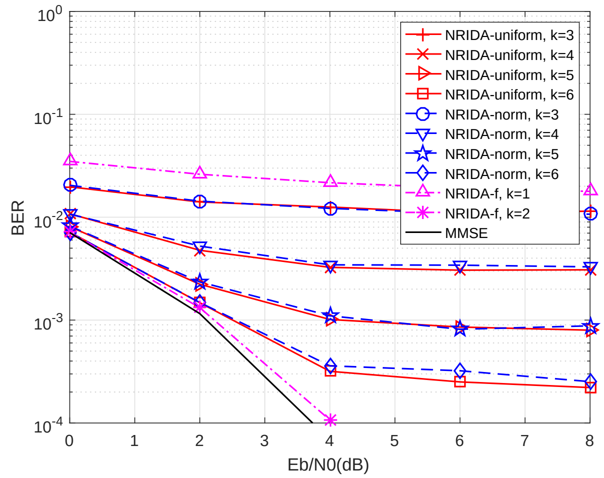

Fig. 2. BER performance versus average SNR per bit for 16×128 large-scale MIMO with 16-QAM.

$$\leq 1 - \frac{\lambda_{\min}(\mathbf{H}^H\mathbf{H})}{\lambda_{\max}(\mathbf{H}^H_{:,\mathcal{Q}_i}\mathbf{H}_{:,\mathcal{Q}_i})} \tag{73}$$

with $\mathbf{H}_{:,\mathcal{Q}_i} \notin \{\mathbf{G}_{k-1}, \ldots, \mathbf{G}_{k-r+1}\}$. Interestingly, under the $f = r-1$ steps conditional sampling, only one sampling choice is left for $\mathbf{G}_k$ when $k > r-1$, which makes the sampling about $\mathbf{G}_k$ become deterministic. Undoubtedly, this is greatly beneficial to the implementation in practice since the process of random sampling can be effectively avoided without any performance loss, thus leading to a more efficient iteration. Hence, for the consideration of better convergence and efficiency, multi-step conditional sampling with $f = r-1$ is highly recommended in the proposed NRIDA, which simply replaces p_i (23) in step 3 of Algorithm 1 with $\overline{p}_i^f$ in (70).

For a better understanding, as shown in Table II, the convergence rate upper bounds of NRIDA with different q and f in 64×128 large-scale MIMO systems are presented in a Monte Carlo way. Clearly, with the increase of the multiple steps of the conditional sampling (i.e., f), a better convergence performance can be obtained by NRIDA with uniform sampling probability, which is accordance with the results of Corollary 3 and 4.

VI. SIMULATIONS

In this section, the detection performance and complexity of the proposed NRIDA for uplink large-scale MIMO systems are investigated by simulations.

In Fig. 2, the bit error rate (BER) performance of the proposed NRIDA is illustrated in 16×128 large-scale MIMO systems with 16-QAM. In particular, NRIDA-uniform and NRIDA-norm denote the proposed NRIDA with sampling probabilities in (48) and (55) respectively, and NRIDA-f represents NRIDA with $f = r-1$ multi-step conditional uniform sampling. Meanwhile, the initial vector $\mathbf{x}^0 = \mathbf{0}$ is applied, the block size is set as $q = 4$. Also, for a fair comparison, k is the number of full iteration for NRIDA in this simulation part. Clearly, as can be seen from Fig. 2, the BER performance of all the NIRDA schemes improve gradually with the increment

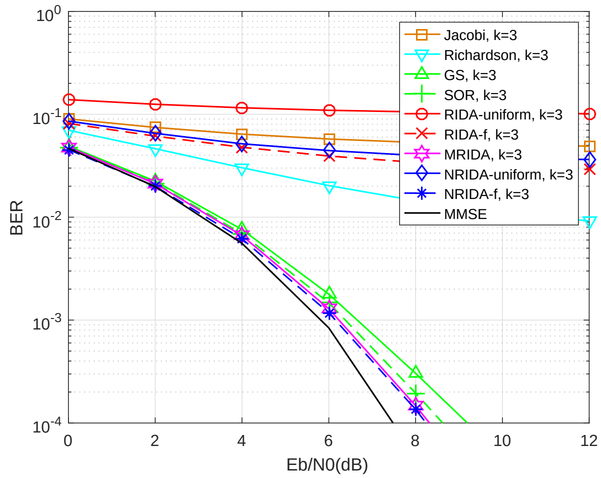

Fig. 3. BER performance versus average SNR per bit for 32×128 large-scale MIMO with 16-QAM.

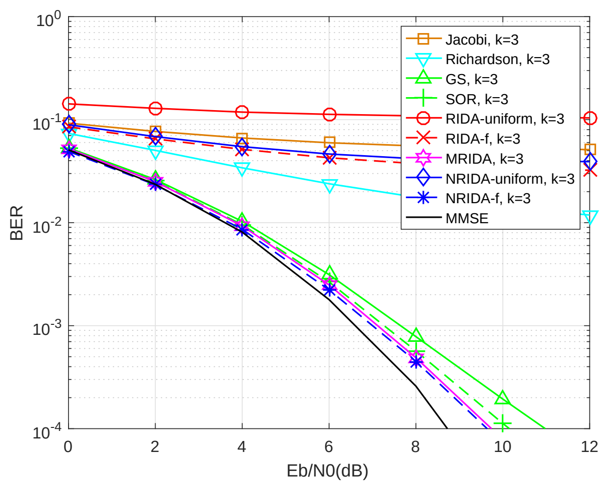

Fig. 4. BER performance versus average SNR per bit for 32×128 large-scale MIMO with the imperfect CSI using 16-QAM.

of full iterations, which is accordance with the convergence results given in Theorem 2 and 4. More specifically, the proposed NRIDA with sampling probabilities in (48) and (55) have nearly the same BER performance, which is in line with the convergence rates provided in Fig. 1. Since the uniform sampling is more straightforward to implement, it is more preferable for NRIDA compared to the sampling choice in (55). On the other hand, as expected, the proposed NRIDA with $f=r-1$ multi-step conditional sampling significantly outperforms those with random sampling probabilities. To be more specific, the detection performance of near-MMSE can be obtained by NRIDA-f with $k=2$. This implies a faster convergence, which has been explained in Corollary 4. Note that the complexity cost of each iteration in NRIDA-f is also less than NRIDA-uniform or NRIDA-norm as the sampling becomes derandomized gradually.

In Fig. 3, the BER performance comparison between the proposed NRIDA and other iterative detection schemes are presented with respect to 32×128 large-scale MIMO systems with 16-QAM. Apart from MMSE detection scheme, the traditional iterative detection schemes such as Jacobi iteration in [9], Richardson iteration in [19] with the relaxation factor $\omega=1/(N+K)$, Gauss Seidel iteration in [11], successive over-relaxation (SOR) iteration in [32] with the relaxation factor $\omega=\frac{2}{1+\sqrt{1-[\varrho(\mathbf{I}-\mathbf{D}^{-1}\mathbf{A})]^2}}$ are also employed. Meanwhile, the randomized iterative detection schemes like RIDA and MRIDA in [22] are applied as well for a better comparison. To make a fair comparison, the block sizes in RIDA, MRIDA and NRIDA are set as $q=8$ with the initial setup $\mathbf{x}^0=\mathbf{0}$. In particular, with the same full iteration number $k=3$, the proposed NRIDA with uniform sampling achieves a better BER performance than RIDA with uniform sampling. Meanwhile, NRIDA with $f=r-1$ multiple step conditional sampling has a comparable BER performance with MRIDA, and both of them outperform other iterative detection schemes like Jacobi, GS and SOR. Note that NRIDA has a much lower computational complexity than MRIDA by removing the preprocessing stage, making it a better choice in uplink signal detection.

Fig. 4 is presented to illustrate the BER performance of the proposed NRIDA without the perfect channel state information (CSI) in 32×128 large-scale MIMO systems using 16-QAM. In particular, $\widehat{\mathbf{H}}=\mathbf{H}+\Delta\mathbf{H}$ denotes imperfect CSI at the receiver, where $\Delta\mathbf{H}\sim\mathcal{CN}(\mathbf{0},\sigma_e^2\mathbf{I}_N)$ represents the channel estimation errors with $\sigma_e^2=\frac{K}{n_p\cdot E_p}$ [33]. Here, n_p and E_p respectively stand for the number and power of pilot symbols, and we apply $\sigma_e^2=0.1$ in the simulation. Compared with the detection performance under perfect CSI in Fig. 3, the BER performance of detection schemes with imperfect CSI decline accordingly in Fig. 4. Nevertheless, the performance gain of NRIDA with $f=r-1$ multi-step conditional sampling still can be verified.

Furthermore, in Fig. 5, the BER performance comparison based on the common pilot channel estimation is shown with respect to 32×128 large-scale MIMO systems using 16-QAM. In particular, we consider a pilot matrix $\mathbf{X}$ composed of K columns of the $\tau\times\tau$ discrete Fourier transform (DFT) operator and we set $\tau=32$. Based on it, the model of channel estimation is formulated as $\mathbf{Y}_p=\mathbf{H}\mathbf{X}+\mathbf{W}_p$ where $\mathbf{W}_p$ and $\mathbf{Y}_p$ denote the noise matrix and the receive signal matrix during the training period respectively. Then, the classic least-squares (LS) channel estimate $\widehat{\mathbf{H}}=\mathbf{Y}_p\mathbf{X}^H(\mathbf{X}\mathbf{X}^H)^{-1}$ is applied as the imperfect channel state information (CSI) with the estimation errors $\Delta\mathbf{H}=\widehat{\mathbf{H}}-\mathbf{H}=\mathbf{W}_p\mathbf{X}^H(\mathbf{X}\mathbf{X}^H)^{-1}$ [34]–[36]. Similar to the imperfect CSI case in Fig. 4, the BER performance of all detection schemes in Fig. 5 are degraded compared to the perfect case in Fig. 3. This is easy to understand as the errors $\Delta\mathbf{H}$ in channel estimation are also taken into account. Nevertheless, the BER performance of the proposed NRIDA is easy to confirm, which is due to a faster convergence performance during the iteration process.

Apart from the independent, identically distributed (i.i.d.) Rayleigh channels, the correlated channels of large-scale MIMO systems are also studied to illustrate the convergence performance of NRIDA. Typically, following the configurations in

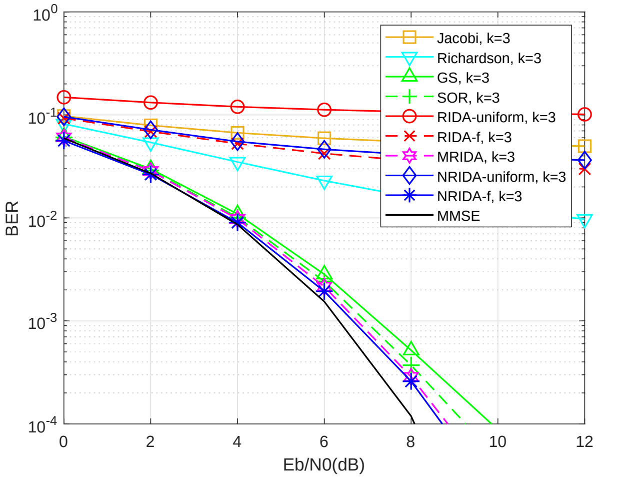

Fig. 5. BER performance versus average SNR per bit for 32×128 large-scale MIMO using 16-QAM with common pilot based channel estimation.

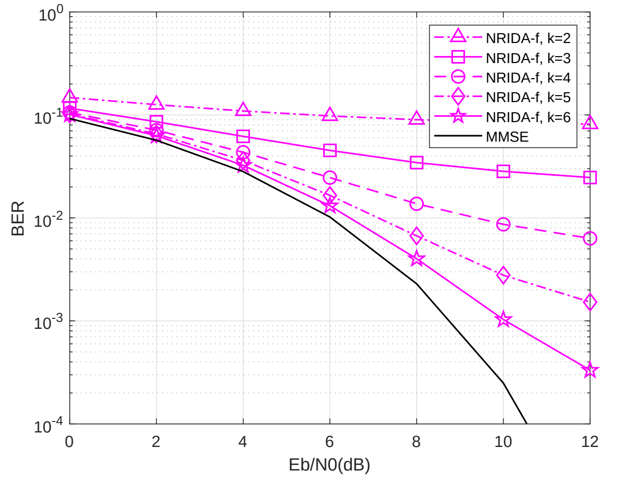

Fig. 7. BER performance versus average SNR per bit for 32×128 large-scale MIMO using 16-QAM with $\psi = 0.1$.

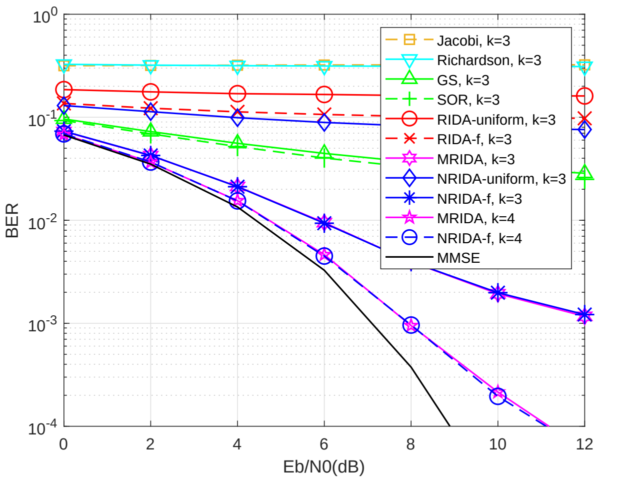

Fig. 6. BER performance versus average SNR per bit for 32×128 large-scale MIMO using 16-QAM with $\psi = 0.05$.

[37], [38], the correlated channel matrix $\mathbf{R}_{\text{cor}}^{\frac{1}{2}}\mathbf{H}\mathbf{T}_{\text{cor}}^{\frac{1}{2}}$ is applied with the normalized correlation coefficient $1 \geq \psi \geq 0$, where $\mathbf{R}_{\text{cor}} \in \mathbb{C}^{N\times N}$ and $\mathbf{T}_{\text{cor}} \in \mathbb{C}^{K\times K}$ represent the receive and transmit correlation matrices respectively, i.e.,

$$\mathbf{R}_{\text{cor}} = \begin{bmatrix} 1 & \psi & \psi^4 & \cdots & \psi^{(N-1)^2} \\ \psi & 1 & \psi & \cdots & \vdots \\ \psi^4 & \psi & 1 & \cdots & \psi^4 \\ \vdots & \vdots & \vdots & \ddots & \psi \\ \psi^{(N-1)^2} & \cdots & \psi^4 & \psi & 1 \end{bmatrix},$$

$$\mathbf{T}_{\text{cor}} = \begin{bmatrix} 1 & \psi & \psi^4 & \cdots & \psi^{(K-1)^2} \\ \psi & 1 & \psi & \cdots & \vdots \\ \psi^4 & \psi & 1 & \cdots & \psi^4 \\ \vdots & \vdots & \vdots & \ddots & \psi \\ \psi^{(K-1)^2} & \cdots & \psi^4 & \psi & 1 \end{bmatrix}.$$

Specifically, an uncorrelated scenario entails $\psi = 0$ and a completely correlated case corresponds to $\psi = 1$. Intuitively, with $\psi = 0.05$ in Fig. 6, the BER performance of all the detection degrade accordingly compared to the results given in Fig. 3. On one hand, it is clear to see that iterative detection schemes like Jacobi and Richardson fail to converge as they are not globally convergent. In contrast to them, detection schemes based on random iteration enjoy the global convergence so that their convergence are still guaranteed. More specifically, under the same full iteration, the proposed NRIDA with $f = r - 1$ multi-step conditional sampling achieves the comparable BER performance with MRIDA. As expected, with the increase of iterations, both of their performance improve accordingly. Meanwhile, the similar observations can also be found in Fig. 7 with $\psi = 0.1$, where the channel matrix is getting more correlated. In this case, the BER performance of MMSE detection becomes worse, and the proposed NRIDA with $f = r - 1$ multi-step conditional sampling still works due to its global convergence. However, because of the correlated channel matrix, the convergence performance of NRIDA-f is slower than before. Nevertheless, by increasing the iteration number k, the detection performance of near-MMSE still can be achieved by the proposed NRIDA.

To further illustrate the proposed NRIDA in correlated channel models, Fig. 8 is presented to show the BER performance comparison in a 32×128 uncoded MIMO system with 16-QAM under local scattering spatial correlation model in [39]. In particular, in the local scattering spatial correlation model, the uniformly distributed deviations are applied with $\theta \sim U[-\sqrt{3}\sigma_\varphi, \sqrt{3}\sigma_\varphi]$, and we set the angular standard deviation (ASD) $\sigma_\varphi = 10°$, where the correlation matrices are normalized with matrix trace equaling to $N = 128$. As can be seen clearly in Fig. 8, under local scattering spatial correlation model, the proposed NRIDA still outperforms the traditional iterative detection schemes by a better BER performance. Meanwhile, with the increment of the full iteration number k, the BER performance of NRIDA improves gradually, which is in line with the results derived in Theorem 2 and 4. On the other hand, compared to the i.i.d. Gaussian channel model

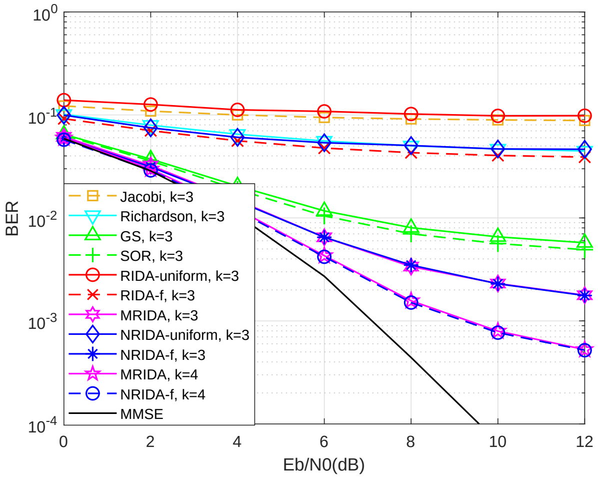

Fig. 8. BER performance versus average SNR per bit for 32×128 large-scale MIMO using 16-QAM under local scattering spatial correlation model with $\sigma_\varphi = 10°$.

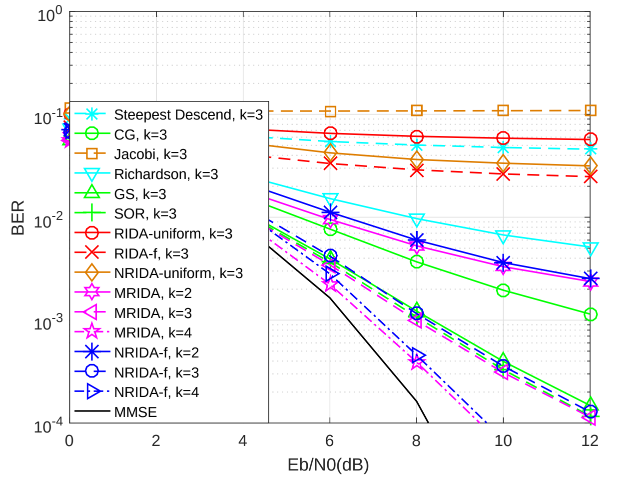

Fig. 10. BER performance versus average SNR per bit for 64×128 large-scale MIMO with 4-QAM.

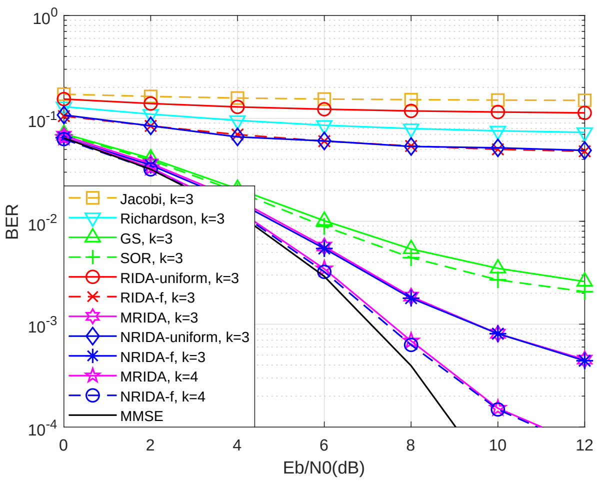

Fig. 9. BER performance versus average SNR per bit for 32×128 large-scale MIMO using 16-QAM under local scattering spatial correlation model with $\sigma_\varphi = 20°$.

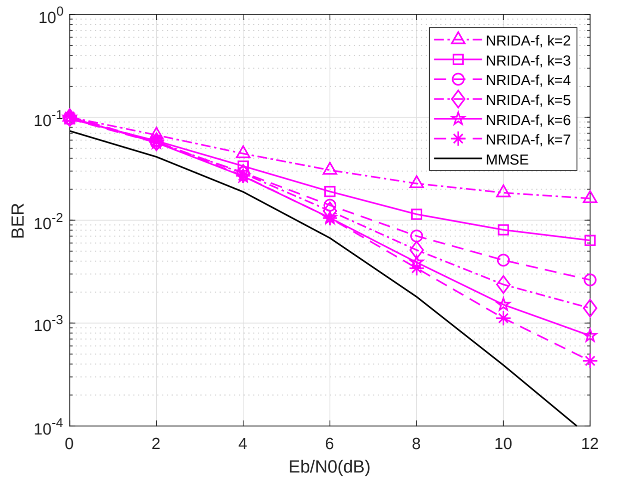

Fig. 11. BER performance versus average SNR per bit for 64×128 large-scale MIMO using 4-QAM under local scattering spatial correlation model with $\sigma_\varphi = 15°$.

shown in Fig. 3, we can observe that the BER performance of all the detection schemes including MMSE, NRIDA and so on degrade accordingly. This is due to the fact that more channel correlation are introduced by the local scattering spatial correlation model. To make it clear, Fig. 9 is added as a complement, where the the angular standard deviation is set as $\sigma_\varphi = 20°$. Clearly, with the increase of σ_φ, the spatial channel correlation of the local scattering spatial correlation model is reduced gradually, so that the BER performance of all the detection schemes in Fig. 9 are better than those in Fig. 8. Meanwhile, as expected, the performance gain of the proposed NRIDA still can be found while near-MMSE performance can be achieved by simply increasing the iteration number.

In Fig. 10, the BER performance comparison about the proposed NRIDA is extended to 64×128 large-scale MIMO systems with 4-QAM. Clearly, the antenna ratio N/K decreases, which imposes a higher requirement upon the detection schemes. Nevertheless, detection schemes like RIDA, MRIDA and NRIDA still work, and their BER performance gradually improve with the increase of iterations. Here, the block size of them are set as $q = 8$ with initial setup $\mathbf{x}^0 = \mathbf{0}$. Meanwhile, the detection schemes based on the steepest descent method in [40] and the conjugate gradient (CG) iterations in [41] are also applied for a better comparison. Compared to the detection cases 16×128 in Fig. 2 and 32×128 in Fig. 3, more iterations are needed to achieve the near-MMSE detection performance. This is straightforward to interpret since less receive diversity can be exploited with the increment of K. As a complement of Fig. 10, the BER performance comparison in a 64×128 large-scale MIMO using 4-QAM is illustrated in Fig. 11 under local scattering spatial correlation model with angular standard deviation $\sigma_\varphi = 15°$. Specifically, the uniformly distributed deviations are applied with $\theta \sim U[-\sqrt{3}\sigma_\varphi, \sqrt{3}\sigma_\varphi]$, and the correlation matrices are normalized with matrix trace equaling to $N = 128$. Clearly, because of the introduced channel cor-

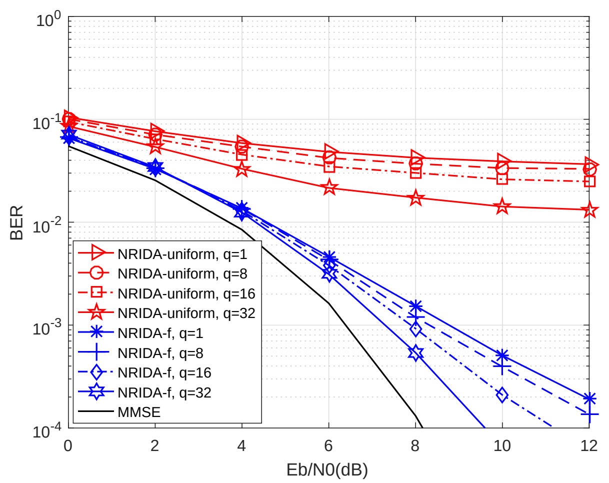

Fig. 12. BER performance versus average SNR per bit for 64×128 large-scale MIMO with 4-QAM.

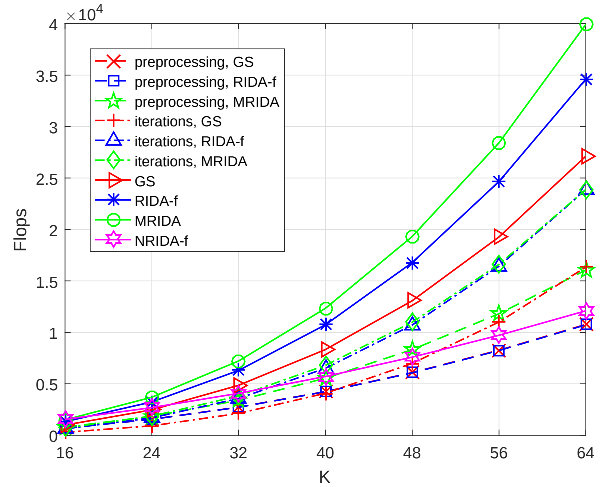

Fig. 13. Complexity comparison in flops for $K \times 128$ large-scale MIMO using 16-QAM at SNR per bit = 7dB.

relation by local scattering spatial correlation model, both the BER performance of NRIDA and MMSE degrade accordingly compared to their counterparts in Fig. 10. Nevertheless, due to the global convergence, the proposed NRIDA still works as usual, and its BER performance improves gradually with the increment of the full iteration number k.

In order to study the choice of the block size q in NRIDA, Fig. 12 is given to show the BER performance of different choices of q in 64×128 large-scale MIMO systems with 4-QAM. Typically, the choices $q = 1, 8, 16, 32$ are applied while the number of iterations for NRIDA is set as $k = 3$. As can be seen, with the improvement of size q, both the BER performance of NRIDA-uniform and NRIDA-f improve gradually. This is easy to understand since a larger q means more components of $\mathbf{x}$ are updated each time, where the correlations between components of $\mathbf{x}$ can be further exploited. As explained before, this accounts for a better convergence performance, which is accordance with the results shown in Table II. However, according to the random iteration in (25), a larger size q also accounts for more computational cost. Therefore, there is a latent trade-off between the convergence and complexity with respect to the choice of q, and how to well balance such a trade-off will be one of our work in future.

To illustrate the computational cost of the proposed NRIDA, Fig. 13 is presented to show the complexity comparison in flops among different iterative detection schemes. Here, the method of flops evaluation that we used comes from [42]. Specifically, the large-scale MIMO system with $N = 128$ received antennas is applied using 16-QAM at SNR per bit $= 7$dB, and the simulation is performed on the platform of MATLAB R2021a on a computer, with a processor at 2.8GHz and 16GB RAM. Meanwhile, the block size of RIDA, MRIDA and NRIDA are set as $q = 8$ with initial setup $\mathbf{x}^0 = \mathbf{0}$. As can be seen clearly, with the improvement of K, the complexities measured by flops of all the detection schemes grow gradually. This is easy to understand as a higher system dimension improves the difficulty of the signal detection problem in large-scale MIMO systems.

As demonstrated, we can see that the proposed NRIDA with $f = r - 1$ multi-step conditional sampling achieves the smallest computational cost among them. This is because the random iteration in the proposed NRIDA is designed to work for the detection model in (1) directly without incurring the matrix multiplication about $\mathbf{H}^H\mathbf{H}$. Different from NRIDA, the iterations in GS, RIDA and MRIDA are designed for the linear system in (7) so that considerable computational complexity have to be paid in the preprocessing stage for matrix multiplication. To make it more specific, both the complexity costs of preprocessing and iteration in GS, RIDA and MRIDA are depicted in Fig. 13. Clearly, the preprocessing stage takes a significant part in running GS, RIDA and MRIDA. Note that the preprocessing of MRIDA has a higher complexity cost than those of GS and RIDA, which is because an extra approximated diagonal matrix needs to be computed in MRIDA. Interestingly, the proposed NRIDA still achieves a lower complexity cost than RIDA and MRIDA even the preprocessing of them are not taken into account, making the proposed NRIDA perspective in the uplink large-scale MIMO signal detection due to a better detection trade-off.

VII. Conclusion

In this paper, a new randomized iterative detection algorithm (NRIDA) is proposed for large-scale MIMO systems. By adopting the random iteration directly into the detection model, considerable complexity cost is saved by the proposed NRIDA without the preprocessing stage for the problem transformation. Meanwhile, we also demonstrate that the proposed NRIDA achieves a globally exponential convergence, making it a promising choice in the uplink signal detection. Besides, the possible complexity reduction of NRIDA is presented while the choices of the sampling distributions in NRIDA are also investigated as well. In addition, to further enhance the convergence performance, the concept of conditional sampling is introduced into the proposed NRIDA, where significant gains in both convergence and complexity can be obtained.

REFERENCES

[1] I. Tomkos, D. Klonidis, E. Pikasis, and S. Theodoridis, "Toward the 6G network era: Opportunities and challenges," *IT Professional*, vol. 22, no. 1, pp. 34–38, 2020.

[2] F. Tariq, M. R. A. Khandaker, K.-K. Wong, M. A. Imran, M. Bennis, and M. Debbah, "A speculative study on 6G," *IEEE Wireless Communications*, vol. 27, no. 4, pp. 118–125, 2020.

[3] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Network*, vol. 34, no. 3, pp. 134–142, 2020.

[4] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, February 2014.

[5] D. Tse and P. Viswanath, Fundamentals of Wireless Communication. Cambridge University Press, 2005.

[6] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, November 2010.

[7] F. Jin, Q. Liu, H. Liu, and P. Wu, "A low complexity signal detection scheme based on improved Newton iteration for massive MIMO systems," *IEEE Communications Letters*, vol. 23, no. 4, pp. 748–751, 2019.

[8] C. Tang, C. Liu, L. Yuan, and Z. Xing, "High precision low complexity matrix inversion based on Newton iteration for data detection in the massive MIMO," *IEEE Communications Letters*, vol. 20, no. 3, pp. 490–493, 2016.

[9] X. Qin, Z. Yan, and G. He, "A near-optimal detection scheme based on joint steepest descent and Jacobi method for uplink massive MIMO systems," *IEEE Communications Letters*, vol. 20, no. 2, pp. 276–279, 2016.

[10] J. Tu, M. Lou, J. Jiang, D. Shu, and G. He, "An efficient massive MIMO detector based on second-order Richardson iteration: From algorithm to flexible architecture," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 67, no. 11, pp. 4015–4028, 2020.

[11] L. Dai, X. Gao, X. Su, S. Han, C.-L. I, and Z. Wang, "Low-complexity soft-output signal detection based on Gauss-Seidel method for uplink multiuser large-scale MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 10, pp. 4839–4845, 2015.

[12] D. Zhu, B. Li, and P. Liang, "On the matrix inversion approximation based on Neumann series in massive MIMO systems," in *2015 IEEE International Conference on Communications (ICC)*, 2015, pp. 1763–1769.

[13] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, Jan 2013.

[14] 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 Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 916–929, 2014.

[15] A. Björck, Numerical Methods for Least Squares Problems. Philadelphia, PA, USA: SIAM, 1996.

[16] A. Greenbaum, Iterative Methods for Solving Linear Systems. Philadelphia, PA, USA: SIAM, 1997.

[17] Y. Saad, Iterative Methods for Sparse Linear Systems. Philadelphia, PA, USA: SIAM, 1997.

[18] Y. Zhang, A. Yu, X. Tan, Z. Zhang, X. You, and C. Zhang, "Adaptive damped Jacobi detector and architecture for massive MIMO uplink," in *2018 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS)*, 2018, pp. 203–206.

[19] X. Gao, L. Dai, C. Yuen, and Y. Zhang, "Low-complexity MMSE signal detection based on Richardson method for large-scale MIMO systems," in *2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall)*, 2014, pp. 1–5.

[20] 3GPP, "User equipment (UE) conformance specification," in *TS 38.521-4 v.15.0.0, Tech. Spec.*, May 2019, pp. 1–5.

[21] ——, "Study on scenarios and requirements for next generation access technologies," in *TR 38.913 v.14.3.0, Tech. Rep.*, June 2017, pp. 1–41.

[22] Z. Wang, R. M. Gower, Y. Xia, L. He, and Y. Huang, "Randomized iterative methods for low-complexity large-scale MIMO detection," *IEEE Transactions on Signal Processing*, vol. 70, pp. 2934–2949, 2022.

[23] S. Liu, C. Ling, and D. Stehlé, "Decoding by sampling: A randomized lattice algorithm for bounded distance decoding," *IEEE Trans. Inform. Theory*, vol. 57, pp. 5933–5945, Sep. 2011.

[24] P. Klein, "Finding the closest lattice vector when it is unusually close," in *ACM-SIAM Symp. Discr. Algorithms*, 2000, pp. 937–941.

[25] Z. Wang and C. Ling, "On the geometric ergodicity of Metropolis-Hastings algorithms for lattice Gaussian sampling," *IEEE Transactions on Information Theory*, vol. 64, no. 2, pp. 738–751, Feb. 2018.

[26] ——, "Lattice Gaussian sampling by Markov chain Monte Carlo: Bounded distance decoding and trapdoor sampling," *IEEE Transactions on Information Theory*, vol. 65, no. 6, pp. 3630–3645, June 2019.

[27] Z. Wang, S. Lyu, Y. Xia, and Q. Wu, "Expectation propagation-based sampling decoding: Enhancement and optimization," *IEEE Transactions on Signal Processing*, vol. 69, pp. 195–209, 2021.

[28] Z. Wang, L. Liu, and C. Ling, "Sliced lattice Gaussian sampling: Convergence improvement and decoding optimization," *IEEE Transactions on Communications*, vol. 69, no. 4, pp. 2599–2612, 2021.

[29] A. Björck, *Numerical Methods in Matrix Computations*. Springer, 2015.

[30] J. Minango, C. de Almeida, and C. Daniel Altamirano, "Low-complexity MMSE detector for massive MIMO systems based on damped Jacobi method," in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2017, pp. 1–5.

[31] G. H. Golub and C. F. V. Loan, *Matrix Computations*. United States: Johns Hopkins University Press, 4th ed., 2013.

[32] P. Zhang, L. Liu, G. Peng, and S. Wei, "Large-scale MIMO detection design and FPGA implementations using SOR method," in *2016 8th IEEE International Conference on Communication Software and Networks (ICCSN)*, 2016, pp. 206–210.

[33] Q. Chen, S. Zhang, S. Xu, and S. Cao, "Efficient MIMO detection with imperfect channel knowledge - A deep learning approach," in *2019 IEEE Wireless Communications and Networking Conference (WCNC)*, 2019, pp. 1–6.

[34] M. Biguesh and A. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 884–893, 2006.

[35] M. Ke, Z. Gao, Y. Wu, X. Gao, and R. Schober, "Compressive sensing-based adaptive active user detection and channel estimation: Massive access meets massive MIMO," *IEEE Transactions on Signal Processing*, vol. 68, pp. 764–779, 2020.

[36] Z. Gao, M. Wu, C. Hu, F. Gao, G. Wen, D. Zheng, and J. Zhang, "Data-driven deep learning based hybrid beamforming for aerial massive MIMO-OFDM systems with implicit CSI," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 10, pp. 2894–2913, 2022.

[37] B. Costa, A. Mussi, and T. Abrao, "MIMO detectors under correlated channels," *Semina: Ciencias Exatas e Tecnologicas*, vol. 37, no. 1, pp. 3–12, 2016.

[38] R. De Lamare and R. Sampaio-Neto, "Detection and estimation algorithms in massive MIMO systems," 08 2014.

[39] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Foundations and Trends® in Signal Processing*, vol. 11, no. 3-4, pp. 154–655, 2017. [Online]. Available: http://dx.doi.org/10.1561/2000000093

[40] Y. Xue, C. Zhang, S. Zhang, Z. Wu, and X. You, "Steepest descent method based soft-output detection for massive MIMO uplink," in *2016 IEEE International Workshop on Signal Processing Systems (SiPS)*, 2016, pp. 273–278.

[41] Y. Hu, Z. Wang, X. Gaol, and J. Ning, "Low-complexity signal detection using CG method for uplink large-scale MIMO systems," in *2014 IEEE International Conference on Communication Systems*, 2014, pp. 477–481.

[42] H. Qian, "Counting the floating point operations (FLOPS)," *MATLAB Central File Exchange*, no. 50608, 2022.

Zheng Wang (Member, IEEE) received the B.S. degree in electronic and information engineering from Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2009, and the M.S. degree in communications from University of Manchester, Manchester, U.K., in 2010. He received the Ph.D degree in communication engineering from Imperial College London, UK, in 2015.

Since 2021, he has been an Associate Professor in the School of Information and Engineering, Southeast University, Nanjing, China. From 2015 to 2016 he served as a Research Associate at Imperial College London, UK. From 2016 to 2017 he was an senior engineer with Radio Access Network R&D division, Huawei Technologies Co.. From 2017 to 2020 he was an Associate Professor at the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics (NUAA), Nanjing, China. His current research interests include massive MIMO systems, machine learning and data analytics over wireless networks, and lattice theory for wireless communications.

Wei Xu (Senior Member, IEEE) received the B.Sc. degree in electrical engineering and the M.S. and Ph.D. degrees in communication and information engineering from Southeast University, Nanjing, China, in 2003, 2006, and 2009, respectively. From 2009 to 2010, he was a Post-Doctoral Research Fellow with the Department of Electrical and Computer Engineering, University of Victoria, Canada. He was an Adjunct Professor of the University of Victoria from 2017 to 2020 and a Distinguished Visiting Fellow of the Royal Academy of Engineering, U.K., in 2019. He is currently a Professor at the National Mobile Communications Research Laboratory, Southeast University. He has coauthored over 100 refereed journal articles in addition to 36 domestic patents and four U.S. patents granted. His research interests include information theory, signal processing, and machine learning for wireless communications. He received the Best Paper Awards from a number of prestigious IEEE conferences including IEEE GLOBECOM/ICCC. He received the Science and Technology Award for Young Scholars of the Chinese Institute of Electronics in 2018. He was an Editor of IEEE Communications Letters from 2012 to 2017. He is currently an Editor of IEEE Transactions on Communications and a Senior Editor of IEEE Communications Letters.

Yili Xia (Member, IEEE) received the B.Eng. degree in Information Engineering from Southeast University, Nanjing, China, in 2006, and the M.Sc. (Distinction) degree in Communications & Signal Processing from Department of Electrical and Electronic Engineering, Imperial College London, London, U.K., in 2007. He obtained his Ph.D. degree in adaptive signal processing from Imperial College London in 2011.

Since 2013, he has been an Associate Professor in Signal Processing with the School of Information Science and Engineering, Southeast University, Nanjing, China, where he is currently the Deputy Head of the Department of Information and Signal Processing Engineering. His research interests include complex and hypercomplex statistical analysis, detection and estimation, linear and nonlinear adaptive filters, as well as their applications on communications and power systems. Dr. Xia was a recipient of the Best Student Paper Award at the International Symposium on Neural Networks (ISNN) in 2010 (coauthor), and the Education Innovation Award at the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) in 2019. He is currently an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING.

Qingjiang Shi (Senior Member, IEEE) received the Ph.D. degree in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 2011. From September 2009 to September 2010, he visited Prof. Z.-Q. (Tom) Luo's Research Group with the University of Minnesota, Twin Cities, MN, USA. In 2011, he was a Research Scientist with Bell Labs China. From 2012, he was with the School of Information and Science Technology, Zhejiang Sci-Tech University, Hangzhou, China. From February 2016 to March 2017, he was a Research Fellow with Iowa State University, Ames, IA, USA. From March 2018, he is currently a Full Professor with the School of Software Engineering, Tongji University, Shanghai, China. He is also with the Shenzhen Research Institute of Big Data. He has authored or coauthored more than 80 IEEE journals and filed about 40 national patents. His interests include algorithm design and analysis with applications in machine learning, signal processing and wireless networks. Dr. Shi was an Associate Editor for the IEEE Transactions on Signal Processing. He was the recipient of IEEE Signal Processing Society Best Paper Award in 2022, the Huawei Technical Cooperation Achievement Transformation Award (2nd Prize) in 2022, the Huawei Outstanding Technical Achievement Award in 2021, the Golden Medal at the 46th International Exhibition of Inventions of Geneva in 2018, the First Prize of Science and Technology Award from China Institute of Communications in 2017, National Excellent Doctoral Dissertation Nomination Award in 2013, Shanghai Excellent Doctorial Dissertation Award in 2012, and the Best Paper Award from the IEEE PIMRC'09 conference.

Yongming Huang (Senior Member, IEEE) received the B.S. and M.S. degrees from Nanjing University, Nanjing, China, in 2000 and 2003, respectively, and the Ph.D. degree in electrical engineering from Southeast University, Nanjing, China, in 2007.

Since March 2007, he has been a Faculty with the School of Information Science and Engineering, Southeast University, where he is currently a Full Professor. During 2008-2009, he visited the Signal Processing Lab, Royal Institute of Technology, Stockholm, Sweden. He has authored or coauthored more than 200 peer-reviewed papers, and holds more than 80 invention patents. His research interests include intelligent 5G/6G mobile communications and millimeter wave wireless communications. He submitted around 20 technical contributions to IEEE standards, and was awarded a certificate of appreciation for outstanding contribution to the development of IEEE standard 802.11aj. He was an Associate Editor for the IEEE Transactions on Signal Processing and a Guest Editor of the IEEE Journal Selected Areas in Communications. He is currently an Editor-at-Large of the IEEE Open Journal of the Communications Society and an Associate Editor for the IEEE Wireless Communications Letters.