Low-complexity sphere decoding for MIMO-SCMA systems

Zhipeng Pan1 | Jing Lei1 | Lei Wen1 | Chaojing Tang1 | Zhongfeng Wang2

1 School of Electronic Science, National University of Defense Technology, Changsha, China
2 School of Electronic Science and Engineering, Nanjing University, Nanjing, China

Correspondence
Jing Lei, School of Electronic Science, National University of Defense Technology, Changsha, China. Email: leijing@nudt.edu.cn

Funding information
National Natural Science Foundation of China, Grant/Award Number: 61702536; National Defense Technology Foundation, Grant/Award Number: 3101168

Abstract
Multiple-input multiple-output-sparse code multiple access, a non-trivial integration of sparse code multiple access and multiple-input multiple-output techniques, is able to achieve high spectrum efficiency and massive user connections. However, this integration also increases the complexity of signal detection. Here, the signal detection problem of multiple-input multiple-output sparse code multiple access is transformed into a tree search problem and use sphere decoding to detect the signal. By setting the initial radius to positive infinity, sphere decoding can achieve optimal maximum likelihood performance while the complexity is high. In order to further reduce the complexity of sphere decoding, a block-wise sorted QR decomposition algorithm is proposed. Based on block-wise sorted QR decomposition, the improved sphere decoding, namely block-wise sorted QR decomposition-sphere decoding, is able to make the tree search more efficient. Since only the detection order of each user’s signal has been changed, block-wise sorted QR decomposition-sphere decoding can maintain the optimal maximum likelihood performance. Simulation results and complexity analysis show that block-wise sorted QR decomposition-sphere decoding can achieve optimal performance and both hard-output and soft-output block-wise sorted QR decomposition-sphere decoding have much lower complexity than joint message passing algorithm. Furthermore, given the same signal-to-noise ratio, the complexity of block-wise sorted QR decomposition-sphere decoding decreases with the increase of receiving antennas, while the complexity of joint message passing algorithm increases linearly.

1 | INTRODUCTION

Future wireless communication systems demand high spectral efficiency and massive connectivity. For supporting these demands, non-orthogonal multiple access (NOMA), which improves spectral efficiency and accommodates massive connectivity by resource sharing among multiuser, has been proposed and becomes the research focus for several years [1]. Sparse code multiple access (SCMA) is one of the typical multi-carrier NOMA schemes that can be efficiently combined with the orthogonal frequency division multiple access (OFDMA) subcarriers [2]. In SCMA, the information bits of different users are directly mapped to the sparse multi-dimensional codewords. Thanks to the sparsity and the elaborately constructed codebooks, overload transmission with high reliability in SCMA can be achieved.

On the other hand, multiple-input multiple-output (MIMO) has also been viewed as the key technique that is able to improve spectral efficiency [3–7] and has been applied in several wireless communication systems for a long time [8]. To date, researchers have investigated the combination of MIMO and SCMA, namely MIMO-SCMA, to further improve the spectral efficiency and support massive connectivity [9–12].

Though MIMO-SCMA has several performance advantages, it imposes considerable difficulties on signal detection. For optimal maximum likelihood (ML) detection, the complexity grows exponentially with the number of users and thus is prohibitively complex for practical use. To reduce the complexity of MIMO-SCMA, the authors proposed the joint message passing algorithm (JMPA) [9, 13] by utilising the sparsity of SCMA codebooks. It has been shown that JMPA can achieve almost the same performance as ML while the complexity is much reduced.
Based on JMPA, many improvements have been made [14–17] to further reduce the detection complexity of MIMO-SCMA. Particularly, the complexity of Gaussian approximation MPA (GA-MPA) [15–17] is able to achieve linear complexity cost with the number of users. However, they all suffer non-ignorable performance loss.

Sphere decoding (SD), which was firstly introduced in [18] for lattice code decoding, has been widely studied in communication systems, such as MIMO detection [19], channel code decoding [20], and SCMA detection [21, 22]. If we set the initial radius to positive infinity, then based on tree search and tree pruning, SD is able to achieve optimal ML performance while has much lower complexity. However, when using SD for SCMA detection, the rank-deficient channel matrix will result in a partial exhaustive search which makes the complexity still high [23].

Here, based on SD, we propose an optimal low complexity detection approach for MIMO-SCMA systems. To the best of the authors’ knowledge, this paper is the first one to investigate SD for MIMO-SCMA systems. Owing to the multi-antenna equipped at the receiver, the rank-deficient problem of the SCMA channel matrix can be avoided. Moreover, it is well known that the complexity of SD in MIMO detection can be reduced by sorted QR decomposition (SQRD). However, there exists one major difference between the MIMO detection and MIMO-SCMA detection. While the estimated signals of the MIMO system are comprised of several one-dimensional complex symbols, the estimated signals of the MIMO-SCMA system are comprised of several multi-dimensional complex symbols. As a result, the channel matrix of the MIMO system can be permuted arbitrarily between nodes in SD, we propose a block-wise sorted QR decomposition (BSQRD) algorithm ultimately related to the nature of MIMO-SCMA. In BSQRD, the columns of the channel matrix are grouped by several blocks and only the order of the blocks can be permuted while the order of the columns in each block is not permitted to change. We call the SD that is based on BSQRD as BSQRD-SD. Since the tree search of BSQRD-SD is more efficient than SD, BSQRD-SD has the same optimal ML performance while it has lower complexity. Simulation results and complexity analysis show that BSQRD-SD can achieve optimal performance while it has lower complexity than JPA. Furthermore, it is worth noting that under the same signal-to-noise ratio (SNR), the complexity of BSQRD-SD decreases with the increase of receiving antennas while the complexity of JPA increases linearly with that.

2 | SYSTEM MODEL

As shown in Figure 1, we consider the uplink MIMO-SCMA system, where the base station (BS) is equipped with \(N_r\) receiving antennas serving \(J\) users equipped with a single antenna. At the transmitter, binary bits \(b_j = [b_{j,1}, b_{j,2}, \ldots, b_{j,\log_2(M_j)}]^T\) of each user \(j\) \((1 \leq j \leq J)\) are encoded into an \(N\)-dimensional complex SCMA codeword \(x_j = [x_{j,1}, x_{j,2}, \ldots, x_{j,N}]^T\) according to the designed codebook \(X_j\), where \(x_j \in X_j\) with cardinality \(|X_j| = M_j\), and \(M\) is the codebook size of SCMA. The codeword \(x_j\) is a sparse vector with \(d_j\) non-zero entries. Let \(\mathbf{s}_j \in S_j\) with cardinality \(|S_j| = M\) be the \(d_j\)-dimensional complex symbol that is composed of the non-zero entries of \(x_j\). Then \(x_j\) can be denoted as \(x_j = \mathbf{V}_j \mathbf{s}_j\), where \(\mathbf{V}_j \in \mathbb{B}^{N \times d_j}\) is the binary mapping matrix of user \(j\) that simply maps the \(d_j\)-dimensional complex symbol to the \(N\)-dimensional sparse complex symbol. For example, in Figure 1 where the small blank boxes denote the zero entries of the SCMA codebook, \(N = 4, d_j = 2, M_1 = 4\), and \(\mathbf{V}_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}^T\).

Let \(\mathbf{h}_{j,n} = [b_{j,1,n}, b_{j,2,n}, \ldots, b_{j,N,n}]^T\) be the channel vector between user \(j\) and the \(n\)th receiving antenna of BS. The channel between the transmitter and receiver is assumed to be i.i.d. flat Rayleigh fading. Thus, each channel gain of \(\mathbf{h}_{j,n}\) is complex Gaussian distributed according to \(\mathcal{CN}(0, 1)\). Then at the BS, the receiving signal of \(n\)th antenna \(y_{n} \in \mathbb{C}^{N \times 1}\) can be written as

\[
\begin{align*}
\mathbf{y}_n &= \sum_{j=1}^{J} \text{diag}(\mathbf{h}_{j,n}) \mathbf{V}_j \mathbf{s}_j + \mathbf{n}_n \\
&= \sum_{j=1}^{J} \mathbf{H}_{j,n} \mathbf{s}_j + \mathbf{n}_n, \\
&= \mathbf{H}_n \mathbf{s} + \mathbf{n}_n,
\end{align*}
\]

where \(\mathbf{H}_n = [\mathbf{H}_{1,n}, \mathbf{H}_{2,n}, \ldots, \mathbf{H}_{J,n}] \in \mathbb{C}^{N \times d_j}\), \(\mathbf{s} = [\mathbf{s}_1^T, \mathbf{s}_2^T, \ldots, \mathbf{s}_J^T] \in \mathbb{C}^{d_j \times 1}\), and \(\mathbf{n}_n \sim \mathcal{CN}(0, \sigma^2 \mathbf{I})\) denotes the complex additive white Gaussian noise (AWGN) with each element having zero mean and variance \(\sigma^2\). Consequently, the total receiving signal of the \(N_r\) receiving antennas can be given.
where \( \mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T, ..., \mathbf{y}_N^T] \in \mathbb{C}^{N_t \times 1} \), and \( \mathbf{H} = [\mathbf{H}_1^T, \mathbf{H}_2^T, ..., \mathbf{H}_N^T]^T \in \mathbb{C}^{N_t \times N_r} \). Since there are \( J \) users and the BS has \( N_r \) antennas, the MIMO-SCMA system size is represented as \( J \times N_r \). Similar to traditional MIMO systems that require receiving antennas to be more than the number of the users to make the channel matrix be column full rank, we suppose that in \( J \times N_r \) MIMO-SCMA systems \( N_r \geq \left\lceil \frac{d_j}{N_r} \right\rceil \).

### 3 | DETECTION PROBLEM FORMULATION

We assume the perfect channel information is available at the BS. Then according to (2), the optimal ML detection for MIMO-SCMA is

\[
\hat{s} = \arg \min_{s \in \mathcal{S}} \| \mathbf{y} - \mathbf{H} s \|_2^2,
\]

where \( \mathcal{S} = \{ \mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_J \} \), and \( \hat{s} = [\hat{s}_1, \hat{s}_2, ..., \hat{s}_J] \) denotes the estimated \( d_j \)-dimensional symbols of \( J \) users. Obviously, the exhaustive search for the ML detection is computationally unaffordable.

Since \( \mathbf{H} \) is column full rank, \( \mathbf{H} \) can be decomposed to \( \mathbf{H} = \mathbf{Q} \mathbf{R} \), where \( \mathbf{Q} = [\mathbf{Q}_1, \mathbf{Q}_2, ..., \mathbf{Q}_N] \in \mathbb{C}^{N_t \times N_r} \), \( \mathbf{Q}_N \in \mathbb{C}^{N_t \times (N_t - d_j)} \) is unitary, \( \mathbf{R} = \begin{bmatrix} \mathbf{R} & 0 \\ 0 & 0 \end{bmatrix} \), \( \mathbf{R} \in \mathbb{C}^{d_j \times d_j} \) and \( \mathbf{R} \) is upper triangular.

Then, left multiplying (2) by \( \mathbf{Q}^H \) we can get

\[
\hat{\mathbf{y}} = \mathbf{R} \hat{s} + \tilde{\mathbf{n}},
\]

where \( \hat{\mathbf{y}} = \mathbf{Q}^H \mathbf{y} \) and \( \tilde{\mathbf{n}} \) is the first \( d_j \) entries of \( \mathbf{Q}^H \mathbf{n} \). Since multiplying with a unitary matrix does not change the statistical properties of the noise, \( \tilde{\mathbf{n}} \) is still the complex AWGN with zero mean and \( \sigma^2 \) variance. Thus, the ML detection of (3) can be reformulated as

\[
\hat{s} = \arg \min_{s \in \mathcal{S}} \| \tilde{\mathbf{y}} - \mathbf{Rs} \|_2^2.
\]

#### 3.1 | \( M \)-ary tree model for MIMO-SCMA detection

Similar to the detection of MIMO [24], the ML detection problem of MIMO-SCMA that is formulated as (5) can be converted into a weighted tree search problem. The tree model is shown as Figure 2, and the tree has the following important properties:

1. The tree has \( Jd_j + 1 \) layers. The \( 1 \rightarrow Jd_j \) layers correspond to the \( 1 \rightarrow Jd_j \) rows of \( \mathbf{R} \) and the entries of the symbol vector of \( \mathbf{s} = \mathbf{s}_{i/J} = [\mathbf{s}_1^T, \mathbf{s}_2^T, ..., \mathbf{s}_J^T]^T \).

\[\text{FIGURE 2} \quad M \text{-ary tree model for MIMO-SCMA detection}\]

2. The nodes at layer \( Jd_j + 1 \), \( (1 \leq l \leq J) \) has \( M \) branches, while other nodes except the leaf nodes only have one branch. The \( M \) branches from the nodes at layer \( Jd_j + 1 \) to \( Jd_j \) are determined by \( \mathbf{s}_{i/J} = [\mathbf{s}_1^T, \mathbf{s}_2^T, ..., \mathbf{s}_J^T]^T \), and are assigned with the branch metric \( BM(\mathbf{s}_{i/J}) = \| \tilde{\mathbf{y}}(l) - \mathbf{R} s(l) \|_2^2 \), where \( \| \cdot \| \) is the Euclidean distance.

3. The \( d_j \) nodes at layers \( Jd_j \) to \( (J - 1)d_j + 1 \) correspond to a path which is associated with a symbol vector of \( \mathbf{s}_{i/J} = [\mathbf{s}_1^T, \mathbf{s}_2^T, ..., \mathbf{s}_J^T]^T \). The path metric is defined as \( PM(\mathbf{s}_{i/J}) = \sum_{l=1}^{Jd_j} \| \tilde{\mathbf{y}}(l) - \mathbf{R} \mathbf{s}(l) \|_2^2 \).

After initialising \( PM(\mathbf{s}_{i/J}) = 0 \), the Euclidean distances \( d(\mathbf{s}) = \| \tilde{\mathbf{y}} - \mathbf{Rs} \|_2^2 \) in (5) can be computed recursively as

\[PM(\mathbf{s}_{i/J}) = PM(\mathbf{s}_{(i+1)/J}) + BM(\mathbf{s}_{i/J}), \quad l = J, J - 1, ..., 1, \]

with \( d(\mathbf{s}) = PM(\mathbf{s}_{i/J}) \). Then according to [25], the SD based on Schnoor–Euchner (SE) enumeration strategy [26] can be applied to solve the problem of (5). The SD detection constrains the search of the nodes to lie in the sphere with \( \tilde{\mathbf{y}} \) as its center and \( C \) as its radius. SE-based SD traverses the tree depth first and in ascending order of their branch metric (which leads to efficient pruning of the tree). The radius \( C \) is adaptively updated to \( \sqrt{d(\mathbf{s})} \) once a leaf node \( \mathbf{s} \) with \( d(\mathbf{s}) < C^2 \) has been reached. The performance of SD is related to the initial radius. If we set the initial radius to \( +\infty \), then we can be assured that at least one leaf node will be found. Thus, SD can always achieve optimal ML detection.

#### 3.2 | SD based on block-wise sorted QR decomposition

A common approach to reduce the complexity of SD without sacrificing the performance for traditional MIMO detection is to reorder the columns of \( \mathbf{H} \), that is, applying QR decomposition of \( \mathbf{HP} \) where \( \mathbf{P} \) is a permutation matrix. The tree pruning
will be more efficient if the main diagonal entries of \( R \) after decomposition of \( HP \) are descending ordered.

The basic idea to reduce the complexity of SD for MIMO-SCMA detection is similar to the above approach. However, since the symbols in \( s = [s_1^T, s_2^T, \ldots, s_J^T] \) are multi-dimensional, \( H \) cannot be permuted arbitrarily. Otherwise, the entries of each block of (7) will be interleaved with each other and SD cannot be applied. In order to avoid the elements of multi-dimensional symbols being interleaved, we first divide \( H \) into \( J \) blocks by grouping every \( d_r \) columns of \( H \) as (7), and then only block-wise permutation of \( H \) is allowed.

\begin{equation}
H = [h_1, \ldots, h_{d_r}, \ldots, h_{(j-1)d_r+1}, \ldots, h_{jd_r}], \tag{7}
\end{equation}

Note that in SD, the nodes at layer \( l \) will be pruned, if the path metric

\begin{equation}
BM(s_{j:j}) > C^2, \tag{8}
\end{equation}

where \( C \) is the radius of the sphere in the current search. Thus, according to (6), the larger the dynamic range of \( BM(s_{j:j}) \), the more likely the nodes of layer \( l \) are to be pruned.

As a result, to make the whole tree be pruned efficiently, \( H \) should be block-wise reordered to make the dynamic range of \( BM(s_{j:j}), BM(s_{j-1:j}), \ldots, BM(s_{1:j}) \) in the descending order.

Reformulating the computation of \( BM(s_{j:j}) \) as

\begin{align}
BM(s_{j:j}) &= \| y_{(l-1)d_r+1:jd_r} - R_{(l-1)d_r+1:jd_r} s_{j:j} \|^2_2 \\
&= \| y_{(l-1)d_r+1:jd_r} - R_{(l-1)d_r+1:jd_r} s_{j:j} \|^2_2 - \| R_{(l-1)d_r+1:jd_r} s_{j:j} \|^2_2 \\
&= \| z_{(l-1)d_r+1:jd_r} - R_{(l-1)d_r+1:jd_r} s_{j:j} \|^2_2, \tag{9}
\end{align}

we can see that only the item \( R_{(l-1)d_r+1:jd_r} s_{j:j} \) is related to the current symbol \( s_j \). Moreover, the calculation of \( BM(s_{j:j}) \) can be further simplified as the Theorem 1.

**Theorem 1.** The branch metric \( BM(s_{j:j}) \) can be given by

\begin{equation}
BM(s_{j:j}) = \sum_{k = (l-1)d_r+1}^{d_r} \| z_k - R_{k,k} s_k \|^2, \tag{10}
\end{equation}

where \( z_k, R_{k,k}, s_k \) is the \( k \)th or \( (k,k) \)th element of \( z \) in (9), \( R \), and \( s \), respectively.

**Proof.** Since \( H \) can be decomposed into the form of \( H = QR \) where \( Q = [q_1, q_2, \ldots, q_{d_r}] \in \mathbb{C}^{N_r \times d_r} \) is column orthogonal and \( R \) is upper triangular, we can get

\begin{equation}
\begin{aligned}
R_{(l-1)d_r+1:jd_r} q_{j-1} &= \sum_{k = (l-1)d_r+1}^{d_r} R_{k,k} q_{j-1} \\
&= \sum_{k = (l-1)d_r+1}^{d_r} R_{k,k} q_k,
\end{aligned} \tag{11}
\end{equation}

Meanwhile, in MIMO-SCMA, the columns of \( V \) are orthogonal. Thus, the columns of each block of (7) are also orthogonal, that is, \( h_{(l-1)d_r+1}, h_{(l-1)d_r+2}, \ldots, h_{ld_r} \) in (11) are orthogonal. The orthogonality leads to \( \{ R_{i,k} = 0 \mid i \neq k, (l-1)d_r+1 \leq i \leq ld_r, (l-1)d_r+1 \leq j \leq ld_r \} \), which guarantees the equality in (10).

**Theorem 1** implies that to make the dynamic range of \( BM(s_{j:j}), BM(s_{j-1:j}), \ldots, BM(s_{1:j}) \) in the descending order is equivalent to make \( \rho_j, \rho_{j-1}, \ldots, \rho_1 \) where \( \rho = \sum_{k = (l-1)d_r+1}^{d_r} | R_{k,k} |^2 (1 \leq l \leq J) \) in the descending order. Since in QR decomposition, \( R_{k,k} (1 \leq k \leq d_r) \) are calculated in the order of \( \rho = 1, 2, \ldots, d_r \), the basic idea to make \( \rho_j, \rho_{j-1}, \ldots, \rho_1 \) in the descending order is by minimising \( \rho_2, \rho_3, \ldots, \rho_J \) when performing the QR decomposition. Now, follow the fact that after permutation of \( H \), \( \rho_j = \sum_{k = (l-1)d_r+1}^{d_r} | h_k |^2 \). Thus, the first \( d_r \) rows of \( R \) will be computed. Next, \( \rho_2 \) and the second \( d_r \) rows of \( R \) are obtained, and so on. The details of the BSQRD are shown in Algorithm 1.

We call the SD based on BSQRD as BSQRD-SD. Now, once the channel matrix \( H \) has been decomposed as \( HP = QR \) (\( P \) is the permutation matrix obtained by permuting the columns of identity matrix according to column order vector \( p \)), the detection problem is reformulated as \( \hat{z} = \arg \min_{z \in Z} \| y - Rz \|^2 \), where \( z = Ps \) is the permuted symbol vector of \( s \) and \( Z \) is the set that contains all the possible values of \( z \). Thus, in BSQRD-SD, to obtain the symbol vector that has the same order of \( \hat{s} \), we need to reorder the detected symbol vector of BSQRD-SD according to \( \hat{s} = P \hat{z} \). Note that since BSQRD-SD only changes the order of the estimated signal, it also achieves the optimal ML performance as SD if we set the initial radius to \( +\infty \).

### 3.3 Complexity analysis

In this subsection, we analyse the computational complexity of the proposed BSQRD-SD. The complexity cost is quantified by the complex floating point operations (flops) where each multiplication operation is counted as three flops and each addition is counted as one flop.

**Table 1** Complexity analysis of JMPA, SD, and BSQRD-SD

| Algorithms | Flops |
|-----------|-------|
| BSQRD-SD/SD | \( \sum_{i=0}^{J} N_r (4(V - I) + 2d_r^2 + 5d_r + 2jd_r) \) |
| JMPA | \( M^2 N_r (4d_r + 3) N_{\text{iter}} \) |
Overall, the complexity of BSQRD-SD consists of the operations performing BSQRD and tree search. Comparing BSQRD shown in Algorithm 1 and the modified Gram–Schmidt algorithm for traditional QRD, we find that only the column sorting and \( O(f^2) \) real floating point operations are required. Thus, for BSQRD, this additional computational complexity is negligible, and about \( 2(fd_v)^3 \) flops are needed \([21]\).

For the tree search of BSQRD-SD, when the nodes at layer \((l-1)d_v + 1 \rightarrow ld_v\) are visited, it will compute \( \|\tilde{y}_{i(l-1)d_v+1:ld_v} - R_{(l-1)d_v+1:ld_v} s_{i/l}\|^2 \). This computation needs \((4(V-l) + 2)d_v^2 + 5d_v\) flops. Suppose that \( N_j \) nodes are visited at layer \((l-1)d_v + 1 \rightarrow ld_v\), then the overall complexity cost of BSQRD-SD can be given by

\[
J_f = \sum_{j=1}^{J} ((4(V-l) + 2)d_v^2 + 5d_v) + 2(fd_v)^3. \tag{12}
\]
The complexity analysis results of SD, BSQRD-SD, and JMPA are shown in Table 1. The flops of Log-Max JMPA are \(M^{d_f}N_rN(4d_f + 3)N_{\text{iter}}\) [27] where \(d_f\) denotes the number of users that occupy the same physical resource element in SCMA, for example, OFDMA subcarrier, and \(N_{\text{iter}}\) denotes the number of iterations used in JMPA. Note that despite the complexity results of SD and BSQRD-SD have the same expression, the number of visiting nodes at each layer, that is, \(N_l\) is different, and thus the complexity is different. Since the number of visiting nodes at each layer is a random variable related to the channel, the average complexity costs of SD and BSQRD-SD are counted by Monte Carlo simulation.

### 4 SIMULATION RESULTS

In this section, we provide some simulation results and numerical complexity results that follow from the analysis in Section 3.3 to evaluate the bit error rate (BER) performance and complexity of the proposed detection approach for MIMO-SCMA. The SCMA codebook designed in [28], where \(J = 6\), \(d_f = 2\), \(N_r = 4\), and the codebook size \(M = 4, 8, 16\), is utilised.

Figure 3a–c compares the hard-output BER performance of ML, JMPA, SD, GA-MPA, and the proposed BSQRD-SD with different codebook sizes in 6 × 3 MIMO-SCMA system. Note that since the complexity of ML detection is too large, we omit the ML simulation when codebook size \(M > 4\) and consider the optimal performance of SD as a benchmark. First, as we have analysed before that BSQRD-SD can achieve the optimal ML performance, we observe that the BER performance of BSQRD-SD is completely the same as ML and SD in these figures, which further verify this conclusion. Meanwhile, it can also be observed that JMPA achieves almost the same performance as ML and SD. However, GA-MPA suffers from significant performance loss, especially in the case of larger codebook size, for example, \(M = 8\) and \(M = 16\).

In Figure 4 we show the average flops cost comparison between JMPA, SD, and BSQRD-SD (hard output). In our simulation, the system settings of JMPA are \(d_f = 3\) and \(N_{\text{iter}} = 4\). From the figure, we can obviously see that BSQRD-SD has lower complexity than JMPA and SD. Meanwhile, it is worth noting that the complexity of SD-based methods is susceptible to SNR. Specifically, in high-SNR region, SD and BSQRD-SD require almost the same flops, which are only 9.4% of the JMPA. While in low-SNR region, for example, 0 dB, the flops of SD and BSQRD-SD are 35.1% and 25.9% of JMPA, respectively.

In order to show the reduced complexity of BSQRD-SD compared with SD, Figure 5 compares the average number of visiting nodes of SD and BSQRD-SD. From the figure, we can see that the proposed BSQRD algorithm can efficiently lower the number of visiting nodes when performing tree search in SD, which accordingly reduces the complexity efficiently.

As the number of the visiting nodes of soft-output SD and BSQRD-SD will much increase\(^1\), Figure 6 shows the average flops cost of soft-output SD, BSQRD-SD, and JMPA. From the figure, it is interesting to see that in low-SNR region (e.g., 0 dB), the complexity of SD is larger than JMPA, while the complexity of BSQRD-SD is still lower than JMPA. That is to say, without the proposed BSQRD algorithm, the complexity of SD can be larger than JMPA. Specifically, in high-SNR region, the flops of SD and BSQRD-SD are 48.9% and 24.6% of JMPA, respectively. While in low-SNR region, for example, 0 dB, the flops of SD and BSQRD-SD are 114.6% and 78.9% of JMPA, respectively.

Figure 7a–c compares the BER performance of SD, JMPA, GA-MPA, and the proposed BSQRD-SD with different codebook sizes in the LDPC coded MIMO-SCMA system.

\(^1\) Soft output is achieved by single tree search [25].
LDPC code is of length 648 and rate 5/6 that used in the IEEE 802.11 standard. It can be observed that BSQRD-SD always has the same BER performance as SD and JMPA no matter what codebook size is. However, due to the error caused by the Gaussian approximation, GA-MPA suffers non-ignorable performance loss, in which the performance gap increases with the increase of the codebook size.

To show the relationship between the complexity of different detection approaches and the number of receiving antennas, we further present the flops cost of JMPA, SD, and BSQRD-SD with a different number of receiving antennas (N_r = 3, 4, 5, 6) in Figure 8. The SNR is fixed at 8 dB. It can be seen from the figure that with the increase of N_r, the complexity of JMPA increases linearly, while the computation overheads of SD and BSQRD-SD slightly decrease with the increase of the number of receiving antennas. This is because the complexity of SD is related to its performance. The better BER performance the fewer nodes have to be visited during tree search (i.e. lower complexity). Increasing the number of receive antennas can lead to a larger diversity gain. Thus, SD-based methods have lower complexity when increasing the number of receiving antennas.

Since massive MIMO has become a trend, more antennas have been equipped at the BS and thus the proposed MMSEBSQRD-SD can be a very promising detection approach for MIMO-SCMA systems.

To further analyse the BER performance of the proposed algorithm in the scenario that the BS is equipped with large-scale antennas, performance evaluations are performed in 6 × 32 and 6 × 64 MIMO-SCMA systems. As shown in Figure 9, the performance of BSQRD-SD is similar to that of JMPA. This performance result along with the complexity advantage illustrated before indicates that BSQRD-SD has more potential to be used than JMPA in this scenario. Moreover, Figure 9 also illustrates the impact of channel estimation errors on the proposed algorithm. In our simulation, we consider that the noise channel estimation is modelled as \( H_{est} = H + \Delta \) where \( \Delta \) is the channel estimation error that independent with \( H \) and \( \Delta \) is complex Gaussian distributed according to \( \mathcal{CN}(0, 1) \). Three cases, that is \( \epsilon = 5\%, 10\%, 20\% \) are considered and compared with the perfect channel, that is, \( \epsilon = 0\% \). As shown in the figure, when \( \epsilon \) is less than 10\%, the performance of BSQRD-SD under imperfect channel is almost the same as that under perfect channel. While \( \epsilon = 20\% \), the proposed algorithm only suffers a slight performance loss. In summary, the performance of the proposed algorithm under imperfect channel is consistent with JMPA and is robust to the channel estimation error.
5 | CONCLUSIONS

In this paper, SD has been investigated to detect signals in the multiuser MIMO-SCMA system. Since the detection problem of SD is equivalent to the ML detection problem and the tree search of SD is more efficient than ML, SD is able to achieve optimal performance while having lower complexity. Moreover, we proposed the BSQRD algorithm for MIMO-SCMA channel matrix QR decomposition. Based on BSQRD, we proposed a low complexity SD-based detection approach, namely BSQRD-SD, for MIMO-SCMA system. Simulation and analysis results show that both hard-output and soft-output BSQRD-SD have lower complexity than SD and traditional JMPA while can achieve optimal ML performance. Meanwhile, simulation results also show that the complexity of BSQRD-SD decreases with the increase of the number of receiving antennas, which makes it a very suitable detection approach for the future massive MIMO-SCMA system.

ACKNOWLEDGMENTS
This work was supported by the National Natural Science Foundation of China (No. 61702536), and the National Defense Technology Foundation under Grant 3101168.

ORCID
Zhipeng Pan https://orcid.org/0000-0003-4577-7705

REFERENCES
1. Dai, L., et al.: A survey of non-orthogonal multiple access for 5G. IEEE Communications Surveys Tutorials 20(3), 2294–2323 (2018)
2. Dabiri, M., Saeedi, H.: Dynamic SCMA codebook assignment methods: A comparative study. IEEE Communications Letters 22(2), 364–367 (2018)
3. Challal, N.R., Bagadi, K.: Design of near-optimal local likelihood search-based detection algorithm for coded large-scale MU-MIMO system. International Journal of Communication Systems 33(12), e4436 (2020)
4. Challal, N.R., Bagadi, K.: Lattice reduction assisted likelihood ascent search algorithm for multiuser detection in massive MIMO system. In: 2018 15th IEEE India Council International Conference (INDICON), pp. 1–5 (2018)
5. Menon, U.V., et al.: Lenstra lenstra lovász (lll) assisted likelihood ascent search (LAS) algorithm for signal detection in massive MIMO. In: 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (VITECoN), pp. 1–4 (2019)
6. Challal, N.R., Bagadi, K.: Design of massive multiuser MIMO system to mitigate inter antenna interference and multiuser interference in 5G wireless networks. Journal of Communications 15(9), 693–701 (2020)
7. Challal, N.R., Bagadi, K.: Likelihood ascent search detection for coded massive MU-MIMO systems to mitigate IAI and MUI. Radioelectronics and Communications Systems 63(5), 223–234 (2020)
8. Li, Q., et al.: MIMO techniques in WiMAX and LTE: A feature overview. IEEE Communications Magazine 48(5), 86–92 (2010)
9. Tang, S., et al.: Low complexity joint MPA detection for downlink MIMO-SCMA. In: 2016 IEEE Global Communications Conference (GLOBECOM), pp. 1–4 (2016)
10. Eilamati, S., et al.: Performance analysis of sparse code multiple access MIMO systems. In: 2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), pp. 1–6 (2019)
11. Pan, Z., et al.: Multi-dimensional space-time block coding aided downlink MIMO-SCMA. IEEE Transactions on Vehicular Technology 68(7), 6657–6669 (2019)
12. Abdessamael, W., et al.: On the performance evaluation of MIMO-SCMA systems. In: 2016 8th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), pp. 135–140 (2016)
13. Du, Y., et al.: Joint sparse graph-detector design for downlink MIMO-SCMA systems. IEEE Wireless Communications Letters 6(1), 14–17 (2017)
14. Dai, J., et al.: Partially active message passing receiver for MIMO-SCMA systems. IEEE Wireless Communications Letters 7(2), 222–225 (2018)
15. Dai, J., et al.: Iterative Gaussian-approximated message passing receiver for MIMO-SCMA system. IEEE Journal of Selected Topics in Signal Processing 13(3), 753–765 (2019)
16. Yuan, W., et al.: Iterative receivers for downlink MIMO-SCMA: Message passing and distributed cooperative detection. IEEE Transactions on Wireless Communications 17(5), 3444–3458 (2018)
17. Wang, P., et al.: Near-optimal MIMO-SCMA uplink detection with low-complexity expectation propagation. IEEE Transactions on Wireless Communications 19(2), 1025–1037 (2019)
18. Viterbo, E., Boutros, J.: A universal lattice code decoder for fanning channels. IEEE Transactions on Information Theory 45(5), 1639–1642 (1999)
19. Damen, O., et al.: Lattice code decoder for space-time codes. IEEE Communications Letters 4(5), 161–163 (2000)
20. Niu, K., et al.: Low-complexity sphere decoding of polar codes based on optimum path metric. IEEE Communications Letters 18(2), 332–335 (2014)
21. Chen, G., et al.: Optimal receiver design for SCMA system. In: 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), pp. 1–6 (2017)
22. Vameghestahbanati, M., et al.: Enabling sphere decoding for SCMA. IEEE Communications Letters 21(12), 2750–2753 (2017)
23. Vameghestahbanati, M., et al.: A novel SD-based detection for generalized SCMA constellations. IEEE Transactions on Vehicular Technology 68(10), 10278–10282 (2019)
24. Larsson, E.G.: MIMO detection methods: How they work. IEEE Signal Process. Mag. 26(3), 91–95 (2009)
25. Studer, C., et al.: Soft-output sphere decoding: Performance and implementation aspects. In: 2006 Fortieth Asilomar Conference on Signals, Systems and Computers, pp. 2071–2076 (2006)
26. Schnorr, C.P., Euchner, M.: Lattice basis reduction: Improved practical algorithms and solving subset sum problems. Mathematical Programming 66(1), 181–199 (1994)
27. Pan, Z., et al.: Uplink spatial modulation SCMA system. IEEE Communications Letters 23(1), 184–187 (2019)
28. Taherzadeh, M., et al.: SCMA codebook design. In: 2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall), pp. 1–5 (2014)

How to cite this article: Pan Z, Lei J, Wen L, Tang C, Wang Z. Low-complexity sphere decoding for MIMO-SCMA systems. IET Commun. 2021;15:537–545. https://doi.org/10.1049/cmu2.12085