Two-Stage List Sphere Decoding for Under-Determined Multiple-Input Multiple-Output Systems

Chen Qian, Jingxian Wu, *Member, IEEE*, Yahong Rosa Zheng, *Senior Member, IEEE*, and Zhaocheng Wang, *Senior Member, IEEE*

Abstract—A two-stage list sphere decoding (LSD) algorithm is proposed for under-determined multiple-input multiple-output (UD-MIMO) systems that employ N transmit antennas and $M < N$ receive antennas. The two-stage LSD algorithm exploits the unique structure of UD-MIMO systems by dividing the N detection layers into two groups. Group 1 contains layers 1 to M that have similar structures as a symmetric MIMO system; while Group 2 contains layers $M + 1$ to N that contribute to the rank deficiency of the channel Gram matrix. Tree search algorithms are used for both groups, but with different search radii. A new method is proposed to adaptively adjust the tree search radius of Group 2 based on the statistical properties of the received signals. The employment of the adaptive tree search can significantly reduce the computation complexity. We also propose a modified channel Gram matrix to combat the rank deficiency problem, and it provides better performance than the generalized Gram matrix used in the Generalized Sphere-Decoding (GSD) algorithm. Simulation results show that the proposed two-stage LSD algorithm can reduce the complexity by one to two orders of magnitude with less than 0.1 dB degradation in the Bit-Error-Rate (BER) performance.

Index Terms—Two-stage LSD algorithm, under-determined multiple-input multiple-output (UD-MIMO), turbo detection, list sphere decoding (LSD), depth-first tree search.

I. INTRODUCTION

MULTIPLE-INPUT multiple-output (MIMO) wireless communication systems have attracted a wide range of attention due to its high spectral efficiency and good performance over fading channels. Most existing studies on MIMO systems focus on symmetric ($N = M$) or over-determined

C. Qian and Z. Wang are with Tsinghua National Laboratory for Information Science and Technology (TNList), Department of Electronic Engineering, Tsinghua University, Beijing 100084, China (e-mail: qc8802@gmail.com, zcwang@mail.tsinghua.edu.cn).

J. Wu is with the Department of Electrical Engineering, University of Arkansas, Fayetteville, AR 72701, USA (e-mail: wuj@uark.edu).

Y. R. Zheng is with the Department of Electrical and Computer Engineering, Missouri University of Science and Technology, Rolla, MO 65409, USA (e-mail: zhengyr@mst.edu).

($N < M$) systems, where N and M are the numbers of transmit antennas and receive antennas, respectively.

Under-determined MIMO (UD-MIMO) systems with more transmit antennas than receive antennas have a wide range of practical applications. The mobile terminals in a wireless system usually have fewer antennas than the base station due to the size limit, and this results in UD-MIMO systems in the downlink of cellular systems or broadcasting systems. The UD-MIMO system can also be used to model wireless sensor networks, where a large number of sensors transmit to the base station simultaneously.

Despite the importance of UD-MIMO systems, the research in this area is scarce due to the technical challenges of detecting a high dimensional signal by observing its projection onto a low dimensional space. Many existing algorithms proposed for the symmetric or the over-determined MIMO systems, for example, the sphere decoding (SD) algorithm [1], [2], the K-best algorithm [3], [4], the fixed-complexity sphere decoding (FSD) algorithm [5], [6], and the vertical Bell laboratories layered space-time (V-BLAST) [7], cannot be readily applied to UD-MIMO systems due to the extra $N - M$ dimensions in the transmitted signal. A UD-MIMO system can be detected by maximum likelihood (ML) or maximum *a posteriori* (MAP) detection, which employs exhaustive search over all the possible Q^N transmitted vectors with Q being the modulation level. The exponentially growing complexity is prohibitive for practical systems with even moderate values of Q and N.

Existing UD-MIMO detection methods can be classified into two categories: the SD-based algorithms and the interference cancelation (IC) based algorithms. For the SD-based algorithms, a generalized SD (GSD) algorithm is proposed in [8] for UD-MIMO systems, where the UD-MIMO system is partitioned into Q^{N-M} parallel symmetric subsystems such that the regular SD algorithm can be employed for each subsystem. The GSD requires an exhaustive search among the Q^{N-M} subsystems. Modified GSD algorithms are proposed in [9], [10], where the exhaustive search is replaced with low complexity alternatives at the cost of performance loss. In [11], a modified metric calculation method with diagonal loading is proposed to avoid the inversion of a rank-deficient matrix in the SD algorithm. A λ-GSD algorithm is proposed in [12] by employing a λ-loading method to convert the UD-MIMO system into a symmetric MIMO system.

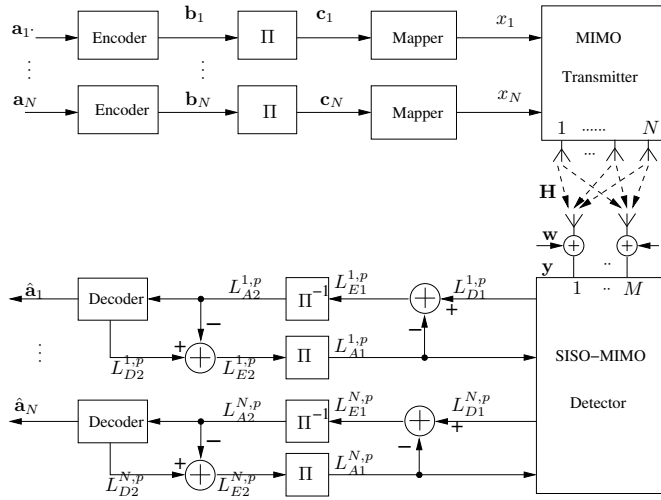For the IC-based algorithms, a generalized parallel IC

Fig. 1. Turbo-MIMO transceiver block diagram, where Π and Π^{-1} denotes the interleaver and the deinterleaver, respectively.

(GPIC) scheme is proposed in [13], and it is later applied to the iterative turbo detection of UD-MIMO systems in [14] with the aid of the block decision feedback equalization algorithm [15]. In the first iteration of the turbo detection, the GPIC algorithm partitions the UD-MIMO system into Q^{N-M} parallel subsystems as in the GSD algorithm, and parallel IC is performed with exhaustive search over the Q^{N-M}-dimensional signal space. In the second and subsequent iterations, the UD-MIMO system is detected with a generalized serial interference cancelation (GSIC) scheme, where the interference from the extra dimension is removed by using the soft decisions from the current and previous iterations. The GSIC removes the need for exhaustive search. The GPIC-GSIC scheme achieves good performance for BPSK modulated system with relatively low complexity. Its performance degrades rapidly at high modulation levels. A modified FSD algorithm is proposed in [16] to replace the GPIC in the first iteration for high-level modulation schemes, and better BER performance is achieved than the GPIC-GSIC algorithm at the cost of higher computational complexity.

In this paper, a two-stage LSD algorithm is proposed to improve the performance-complexity tradeoff in UD-MIMO detections. The N detection layers, which correspond to the N transmit antennas, are divided into two groups. Group 1 contains layers 1 to M, and group 2 contains layers $M+1$ to N corresponding to the extra $N-M$ signal dimensions at the transmitter. The layers in group 2 contribute to the rank deficiency of the channel matrix, and they have very low signal-to-noise ratio (SNR). Due to the low SNR, most existing UD-MIMO detection methods use exhaustive search over layers $M+1$ to N, and this incurs very high complexity. To address this problem, we propose to use a modified list SD (LSD) algorithm with a depth-first tree search in layers $M+1$ to N. The modified LSD dynamically adjusts the radius of the tree search based on the channel condition and the statistical properties of the metric used in the tree search, such that the probability of missing the optimum solution in the tree search is upper bounded by a very small threshold. Such a method ensures good performance with complexity much

lower than the exhaustive search employed by most existing UD-MIMO detection methods. A depth-first tree search with a conventional radius constraint is then used for the detection of layers 1 to M to further reduce the complexity. Simulation results and complexity analysis demonstrate that the proposed two-stage LSD algorithm has a much lower complexity than existing algorithms with similar performances.

Common notations used in the paper are: $\mathbf{I}_N$ is the identity matrix with size $N \times N$; $\text{Diag}(\mathbf{d})$ is a diagonal matrix with its diagonal elements being the elements of the vector $\mathbf{d}$; $\mathcal{CN}(\mu, \sigma^2)$ denotes the complex Gaussian distribution with mean μ and variance σ^2; χ_m^2 denotes the Chi-squared distribution with m degrees of freedom; $\mathbb{E}(\cdot)$ is the expectation operator; $\mathcal{C}^{M \times N}$ and $\mathcal{R}^{M \times N}$ denote the $(M \times N)$-dimensional complex and real spaces, respectively; and superscripts $(\cdot)^H$ and $(\cdot)^T$ denote Hermitian transpose and transpose, respectively.

II. THE UD-MIMO SYSTEM MODEL

Consider a UD-MIMO system with N transmit antennas and M receive antennas as depicted in Fig. 1, where $N > M$. The N independent bit streams $\{\mathbf{a}_n\}_{n=1}^N$ are encoded by channel encoders to generate the coded bit streams $\{\mathbf{b}_n\}_{n=1}^N$. The coded bit streams are interleaved by pseudo-random interleavers to obtain the interleaved bit streams $\mathbf{c}_n = \Pi(\mathbf{b}_n)$, for $n = 1 \cdots N$, where $\Pi(\cdot)$ is the interleaving operator. Then every P bits are mapped to a symbol in the modulation constellation set $\mathcal{S}$ with a cardinality $Q = 2^P$. The modulated symbol vectors, $\mathbf{x}_o = [x_1, \cdots, x_N]^T \in \mathcal{S}^{N \times 1}$, are transmitted on N antennas through a channel with flat fading and additive white gaussian noise (AWGN).

The equivalent discrete-time signals at the receiver can be represented as

$$\mathbf{y} = \mathbf{H}\mathbf{x}_o + \mathbf{w} \tag{1}$$

where $\mathbf{y} = [y_1, \cdots, y_M]^T \in \mathcal{C}^{M \times 1}$ and $\mathbf{w} = [w_1, \cdots, w_M]^T \in \mathcal{C}^{M \times 1}$ represent the received signal and the noise vector, respectively. Every element of the noise vector $\mathbf{w}$ has a complex Gaussian distribution with zero mean and variance σ^2. The matrix $\mathbf{H} \in \mathcal{C}^{M \times N}$ is the flat-fading MIMO channel matrix, with the (m, n)-th element $h_{m,n} \sim \mathcal{CN}(0, 1)$ being the complex channel coefficient between the n-th transmit antenna and the m-th receive antenna. It is assumed that the channel matrix $\mathbf{H}$ is known at the receiver.

For the UD-MIMO system, the receiver may employ the turbo detection similar to conventional MIMO solutions, as shown in Fig. 1, where a MIMO soft-symbol detector and N soft channel decoders are connected by de-interleavers and interleavers. Soft information is iteratively exchanged between the soft-symbol detector and the soft channel decoders. Let $L_{D1}^{(n,p)}$, $L_{E1}^{(n,p)}$ and $L_{A1}^{(n,p)}$ denote the soft *a posteriori*, extrinsic, and *a priori* log-likelihood ratios (LLRs) for bit $c_{n,p}$ at the symbol detector, where the superscripts p and n denote the p-th bit from the n-th transmit antenna. Similarly, $L_{D2}^{(n,p)}$, $L_{E2}^{(n,p)}$ and $L_{A2}^{(n,p)}$ denote the *a posteriori*, extrinsic, and *a priori* LLRs of $c_{n,p}$ at the channel decoder, respectively.

The MIMO soft-symbol detector calculates $L_{D1}^{(n,p)}$ as

$$L_{D1}^{(n,p)} = \ln \frac{\sum_{\mathbf{x} \in \mathcal{S}_{n,p,0}} \exp\left(-\frac{1}{\sigma^2} ||\mathbf{y} - \mathbf{Hx}||^2\right) P(\mathbf{x})}{\sum_{\mathbf{x} \in \mathcal{S}_{n,p,1}} \exp\left(-\frac{1}{\sigma^2} ||\mathbf{y} - \mathbf{Hx}||^2\right) P(\mathbf{x})}, \quad (2)$$

where $\mathbf{x} = [x_1, \cdots, x_N]^T \in \mathcal{S}^{N \times 1}$ is the modulated symbol vector, σ^2 is the variance of the AWGN, $P(\mathbf{x})$ is the *a priori* probability of $\mathbf{x}$ that can be calculated from $L_{A1}^{n,p}$, and the set $\mathcal{S}_{n,p,b}$ contains all the vectors in $\mathcal{S}^{N \times 1}$ with $c_{n,p} = b$, for $b \in \{0, 1\}$.

Denote $\mathbf{c}_n = [c_{n,1}, \cdots, c_{n,P}]^T$ as the vector of coded bits mapped to the symbol x_n. Then the *a priori* probability is $P(\mathbf{x}) = \prod_{n=1}^{N} \prod_{p=1}^{P} P(c_{n,p})$. Given the *a priori* LLR $L_{A1}^{n,k}$ for $c_{n,k}$, we have

$$P(c_{n,k} = 1) = \frac{1}{1 + \exp(L_{A1}^{n,k})}, \quad (3a)$$

$$P(c_{n,k} = 0) = \frac{\exp(L_{A1}^{n,k})}{1 + \exp(L_{A1}^{n,k})}. \quad (3b)$$

Substituting (3) into $P(\mathbf{x})$ yields

$$P(\mathbf{x}) = \exp\left(\frac{1}{2}\mathbf{s}_\mathbf{x}^T \cdot \mathbf{L}_{A1}\right) \times \prod_{n,k} \frac{\exp\left(\frac{1}{2}L_{A1}^{n,k}\right)}{1 + \exp(L_{A1}^{n,k})} \quad (4)$$

where $\mathbf{s}_\mathbf{x} = [s_{1,1}, \cdots, s_{1,P}, \cdots, s_{N,1}, \cdots, s_{N,P}]^T$, with $s_{n,p} = 1 - 2c_{n,p}$, and $\mathbf{L}_{A1} = [L_{A1}^{1,1}, \cdots, L_{A1}^{1,P}, \cdots, L_{A1}^{N,1}, \cdots, L_{A1}^{N,P}]^T$. Combining (2) and (4), we have

$$L_{D1}^{(n,p)} = \ln \frac{\sum_{\mathbf{x} \in \mathcal{S}_{n,p,0}} \exp\left(-\frac{1}{\sigma^2} ||\mathbf{y} - \mathbf{Hx}||^2 + \frac{1}{2}\mathbf{s}_\mathbf{x}^T \cdot \mathbf{L}_{A1}\right)}{\sum_{\mathbf{x} \in \mathcal{S}_{n,p,1}} \exp\left(-\frac{1}{\sigma^2} ||\mathbf{y} - \mathbf{Hx}||^2 + \frac{1}{2}\mathbf{s}_\mathbf{x}^T \cdot \mathbf{L}_{A1}\right)}. \quad (5)$$

Equation (5) can be simplified by using the max-log-map approximation [17],

$$L_{D1}^{(n,p)} \approx \frac{1}{2} \max_{\mathbf{x} \in \mathcal{S}_{n,p,0}} \left(-\frac{2}{\sigma^2} ||\mathbf{y} - \mathbf{Hx}||^2 + \mathbf{s}_\mathbf{x}^T \cdot \mathbf{L}_{A1}\right)$$
$$- \frac{1}{2} \max_{\mathbf{x} \in \mathcal{S}_{n,p,1}} \left(-\frac{2}{\sigma^2} ||\mathbf{y} - \mathbf{Hx}||^2 + \mathbf{s}_\mathbf{x}^T \cdot \mathbf{L}_{A1}\right). \quad (6)$$

The extrinsic LLR of the MIMO soft-symbol detector is calculated as $L_{E1}^{(n,p)} = L_{D1}^{(n,p)} - L_{A1}^{(n,p)}$, and it is deinterleaved as $L_{A2}^{(n,p)} = \Pi^{-1}\left(L_{E1}^{(n,p)}\right)$, where $\Pi^{-1}(\cdot)$ is the de-interleaving operator. The channel decoders compute the *a posteriori* LLR $L_{D2}^{(n,p)}$ and obtain the extrinsic LLR $L_{E2}^{(n,p)} = L_{D2}^{(n,p)} - L_{A2}^{(n,p)}$. The extrinsic LLR of the channel decoder $L_{E2}^{(n,p)}$ is then interleaved as $L_{A1}^{(n,p)} = \Pi(L_{E2}^{(n,p)})$ and used as the soft *a priori* input to the MIMO soft-symbol detector in the next iteration. Initially, $L_{A1}^{(n,p)} = 0$ since there is no *a priori* information in the first iteration.

The computational complexity of (6) is on the order of $\mathcal{O}(Q^N)$ due to exhaustive search of all the vectors in $\mathcal{S}^N$. The complexity becomes prohibitively high when Q or N is large. Low complexity algorithms, such as LSD algorithm, are designed for symmetric-MIMO or over-determined MIMO, to simplify the calculation of (6) while achieving near-optimum performance. Instead of exhaustively searching all the possible vectors, the LSD algorithm [2] searches a hyper-sphere around the received vector as

$$\mathbf{x} \in \{\mathbf{x} | \mathbf{x} \in \mathcal{S}^{N \times 1}, ||\mathbf{y} - \mathbf{Hx}||^2 < r_1^2\}, \quad (7)$$

where r_1 is a pre-defined constraint radius. By using matrix decomposition and removing the constant terms, the inequality constraint in (7) is equivalent to

$$||\mathbf{U}(\mathbf{x} - \hat{\mathbf{x}})||^2 < r_2^2, \quad (8)$$

where the search center $\hat{\mathbf{x}} = \mathbf{H}^\dagger \mathbf{y}$ is the least squares (LS) estimate of the transmitted vector $\mathbf{x}_o$, and $\mathbf{H}^\dagger = (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H$ is the pseudo-inverse of the channel matrix $\mathbf{H}$. The upper-triangular matrix $\mathbf{U}$ can be obtained from the Cholesky decomposition of the Gram matrix $\mathbf{G} = \mathbf{H}^H \mathbf{H} = \mathbf{U}^H \mathbf{U}$. Instead of exhaustively searching all the vectors in $\mathcal{S}^N$, the LSD algorithm will only search the vectors satisfying (8).

By exploiting the upper-triangular structure of the matrix $\mathbf{U}$ in the conventional MIMO system, the LSD algorithm performs searching by following a tree structure with N layers. The nodes on the n-th layer of the tree correspond to the possible values of x_n. The N-th layer is the root. Each node on the n-th layer spawns Q child nodes in the $(n-1)$-th layer, corresponding to the Q possible values of x_{n-1}. The tree stretches from the N-th layer (the root layer) to the first layer (the leaf layer), thus a path from a leaf node to the root layer denotes a possible transmitted vector $\mathbf{x}$. A full tree has Q^N leaf nodes, thus Q^N paths, corresponding to the Q^N vectors in $\mathcal{S}^N$. The LSD algorithms will only search a subset of the tree to reduce the computation complexity.

The LSD algorithm employs a depth-first tree search. In the root layer, all the nodes that satisfy the tree search constraint survive. At the n-th layer, one of the survival nodes is chosen as the parent node, and it will spawn Q child nodes in the $(n-1)$-th layer. The child nodes in the $(n-1)$-th layer that satisfy the tree search constraint will be the survival nodes. This procedure is performed layer by layer. If layer 1 is reached and one or more leaf nodes satisfy the constraint in (8), then the paths leading to these leaf nodes are survival paths. If none of the Q child nodes in one layer satisfies the constraint, the searching operation will trace backward to choose another survival node in the previous layer as the parent node. When the constraint radius is chosen appropriately, the LSD algorithm can reduce the computational complexity with negligible performance degradation because the probability of missing the true path is very small [18].

The LSD algorithm works well for conventional MIMO systems with $N \leq M$. However, it cannot be readily applied to UD-MIMO systems for two reasons.

First, the Gram matrix $\mathbf{G} = \mathbf{H}^H \mathbf{H}$ for a UD-MIMO system is rank-deficient with $N > M$. There are always $D = N - M$ zero eigenmodes in $\mathbf{G}$. Consequently, the pseudo-inverse of $\mathbf{H}$ does not exist and the LS estimate of $\mathbf{x}_o$ cannot be obtained. In addition, the last D rows of the upper-triangular matrix $\mathbf{U}$ obtained from the Cholesky decomposition of the Gram matrix $\mathbf{G}$ or equivalent by the QR decomposition of the channel matrix $\mathbf{H}$ are all zeros, thus it cannot be used for the tree search. To overcome this problem, a generalized Gram matrix $\mathbf{G}_g = \mathbf{G} + \beta \mathbf{I}$ with β being a small positive number is used in place of $\mathbf{G}$ to estimate $\mathbf{x}_o$ for UD-MIMO systems [11]. When $\beta = \sigma^2$, the estimate of $\mathbf{x}_o$ is equivalent to a minimum mean square error (MMSE) estimation. It is reported that β will not affect the performance for constant-modulo modulation for uncoded system. However, for coded systems,

a large β will affect the LLR and thus the performance even for constant-modulo modulation.

Second, it is difficult to obtain an appropriate value of the constraint radius r_2 in (8) for UD-MIMO systems, even if the initial estimate of $\mathbf{x}_o$ is made available with the generalized Gram matrix. If r_2 is selected based on the rules for conventional MIMO systems [2], it would only be suitable for the layers corresponding to the M non-zero eigenmodes, but too large for the D layers corresponding to the zero eigenmodes. A unnecessarily large value of r_2 means exhaustive search for the D layers, which leads to high computation complexity. This fact means that, for UD-MIMO systems, the layers corresponding to the zero eigenmodes have different properties than the layers corresponding to the non-zero eigenmodes, and they should be treated differently.

III. THE PROPOSED TWO-STAGE LSD ALGORITHM

This section proposes a two-stage LSD algorithm to balance the complexity-performance tradeoff for UD-MIMO systems. The two-stage LSD algorithm explores the channel structure of the UD-MIMO systems and uses different constraint radii synergistically to suit the structure of UD-MIMO systems. A new channel ordering method is also proposed so that the layers corresponding to stronger channel modes are moved to the bottom D layers of the tree. A new modified Gram matrix is then utilized to compute the initial estimate of $\mathbf{x}_o$, which is used as the search center for the proposed algorithm. With the initial estimate, the bottom D layers of the tree are searched with a modified constraint radius selection method while the top M layers are searched with the conventional radius selection method.

The following subsections will detail the channel ordering, initial symbol estimation, and tree search for the proposed two-stage LSD algorithm.

A. Channel Ordering

The main objective of channel ordering for UD-MIMO systems is to classify the N layers of the channel Gram matrix into two groups: the top M layers that correspond to weaker channel modes, and the bottom D layers that correspond to the stronger channel modes. The ordering of the D layers inside the bottom group plays a very small role on the performance of the tree search. A three-step heuristic approach is proposed here for channel ordering of the UD-MIMO system:

1) Calculate the channel Gram matrix: $\mathbf{G} = \mathbf{H}^H\mathbf{H} = [\mathbf{g}_1, \cdots, \mathbf{g}_N]$, where $\mathbf{g}_i$ is the i-th column of $\mathbf{G}$.
2) Calculate the norms, $\|\mathbf{g}_i\|$, and sort them in an ascending order, such that $\|\mathbf{g}_{p_1}\| \leq \cdots \leq \|\mathbf{g}_{p_M}\| \leq \cdots \leq \|\mathbf{g}_{p_N}\|$. Note that the sorting operation can stop after finding the smallest M norms of $\mathbf{g}_i$ because the ordering of the bottom D layers has negligible effects;
3) Permute the channel matrix into $\mathbf{H}_p = [\mathbf{h}_{p_1}, \cdots, \mathbf{h}_{p_N}]$, where $[p_1, \cdots, p_N]$ is the permutation of $[1, \cdots, N]$ according to the order obtained from Step 2) and $\mathbf{h}_{p_i}$ is the p_i-th column vector of the original channel matrix $\mathbf{H}$.

Using the norms of the columns of the Gram matrix $\mathbf{G}$ as the measurement of the channel reliability avoids the inversion operation of a rank-deficient matrix. Our simulation results indicate that the proposed ordering scheme has lower complexity and better performance than the channel ordering scheme in [6].

B. Initial symbol estimation

The Gram matrix of the permuted channel matrix $\mathbf{G}_p = \mathbf{H}_p^H\mathbf{H}_p$ is rank deficient with D zero eigenvalues. The eigenvalue decomposition of $\mathbf{G}_p$ takes the form

$$\mathbf{G}_p = \mathbf{V}_p^H \boldsymbol{\Gamma} \mathbf{V}_p, \tag{9}$$

where the diagonal matrix $\boldsymbol{\Gamma} = \text{Diag}\{\gamma_1, \cdots, \gamma_M, 0, \cdots, 0\}$, with γ_i being the M non-zero eigenvalues of $\mathbf{G}_p$ and the columns of $\mathbf{V}_p \in \mathcal{C}^{N \times N}$ are the corresponding eigenvectors of $\mathbf{G}_p$.

Instead of using the generalized Gram matrix $\mathbf{G}_g = \mathbf{G}_p + \beta\mathbf{I}_N$ as in [11], we propose to use a modified Gram matrix $\mathbf{G}_m = \mathbf{G}_p + \mathbf{B}$ to solve the problem of rank-deficiency, where

$$\mathbf{B} = \mathbf{V}_p^H \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \beta\mathbf{I}_D \end{pmatrix} \mathbf{V}_p, \tag{10}$$

with β being a small positive number. The modified Gram matrix only adds a small offset to the layers with zero eigenvalues without affecting the layers with non-zero eigenvalues. In addition, this modification can facilitate the derivation of the constraint radius used in the tree search as shown in the next subsection.

With $\mathbf{G}_m$, the initial estimate $\hat{\mathbf{x}}$ of the transmitted vector is calculated as

$$\hat{\mathbf{x}} = \mathbf{G}_m^{-1}\mathbf{H}_p^H\mathbf{y}. \tag{11}$$

It is noted that (11) is a biased estimate of $\mathbf{x}_o$ due to the introduction of the bias matrix $\mathbf{B}$. That is $\mathbb{E}[\hat{\mathbf{x}}] = \mathbf{G}_m^{-1}\mathbf{H}_p^H\mathbf{H}_p\mathbf{x}_o$. However, it serves as a good starting point for the subsequent tree search.

C. Tree search

The proposed tree search is a combination of two depth-first searches with different constraint radii. The metric for the overall tree search is

$$\Delta' = ||\mathbf{y} - \mathbf{H}_p\mathbf{x}||^2. \tag{12}$$

Without loss of generality, we assume the elements in $\mathbf{x}$ have been reordered according to the ordering of the columns of $\mathbf{H}_p$.

With the modified Gram matrix, define an alternative metric as

$$\Delta = ||\mathbf{U}_m(\mathbf{x} - \hat{\mathbf{x}})||^2 = \sum_{i=1}^{N} u_{ii}^2 |x_i - z_i|^2, \tag{13}$$

where the matrix $\mathbf{U}_m = \{u_{ij}\} \in \mathcal{C}^{N \times N}$ is an upper-triangular matrix obtained from the Cholesky decomposition of the modified Gram matrix $\mathbf{G}_m = \mathbf{U}_m^H\mathbf{U}_m$, and z_i can be calculated as

$$z_i = \hat{x}_i - \sum_{j=i+1}^{N} \frac{u_{ij}}{u_{ii}}(x_j - \hat{x}_j). \tag{14}$$

$$z_N = \hat{x}_N. \tag{15}$$

With the definition of $\mathbf{U}_m$ and $\mathbf{G}_m$, it is straightforward to show that

$$\Delta' = \Delta + C - \mathbf{x}^H \mathbf{B} \mathbf{x} \qquad (16)$$

where C is a constant independent of $\mathbf{x}$. There is a very small difference between Δ' and Δ contributed by the last term in (16). However, since the value of β is usually very small ($\beta = 10^{-6}$ is used in this paper), the difference between the two metrics is negligible for practical systems.

With the Cholesky decomposition $\mathbf{G}_m = \mathbf{U}_m^H \mathbf{U}_m$, the bottom $D = N - M$ layers correspond to the $M + 1$ to N column dimensions. Consequently, the values of u_{ii}, for $i = M + 1, \cdots, N$, are very small, which results in very low SNR at these layers. Ordering the channel matrix such that the columns of $\mathbf{H}$ with larger norms are in the bottom D layers can slightly alleviate this problem, but the SNR of the bottom D layers is still significantly lower than that of the first M layers. As a result, if the same constraint radius is applied to all the layers of the LSD algorithm, the detector will have unnecessarily high computational complexity if a large radius is used, or unsatisfied performance if a small radius is used.

For this reason, we propose to use different search radii for the two groups. For the first group corresponding to layers N to $M + 1$, we define a partial metric as

$$\Delta_D = \sum_{i=M+1}^{N} u_{ii}^2 |x_i - z_i|^2. \qquad (17)$$

The depth-first tree search is performed for these layers with respect to a new constraint, $\Delta_D < r_D^2$. The radius r_D is a key design parameter for the proposed two-stage LSD algorithm. A larger r_D means higher complexity, yet a smaller r_D might miss the optimum paths, especially in UD-MIMO systems with a very low SNR in the bottom D layers. The choice of r_D depends on the structure of the receiver and the statistics of the signals in the bottom D layers. The design of r_D for UD-MIMO systems will be addressed in section III-D.

The tree search of the first stage starts from layer N and stops at layer $M + 1$. In the N-th layer, the values of x_N that satisfy the following condition are denoted as the survival nodes,

$$u_{NN}^2 |x_N - z_N|^2 < r_D^2. \qquad (18)$$

In the $(n+1)$-th layer for $M+1 \le n < N$, one of the survival nodes is chosen as the parent node, which will spawn Q child nodes in the n-th layer. The Q child nodes in the n-th layer will be checked against the following condition

$$\sum_{i=n}^{N} u_{ii}^2 |x_i - z_i|^2 < r_D^2 \qquad (19)$$

The child nodes that satisfy the above condition will be denoted as the survival nodes in the n-th layer. If none of the child nodes in the n-th layer satisfies (19), then the search will back up to the previous layer and choose another survival node as the parent node. The above procedure is performed from layer N to layer $M + 1$ along the path of the tree. The path from each survival node at layer $M+1$ traced back to layer N is denoted as a survival path, and it will be added to the set of survival paths. Once layer $M + 1$ is reached, another survival

nodes in the root layer will be selected to start the search again. The depth-first search terminates when all the survival nodes at layer N have been used as a parent node. After the depth-first search, N_{LSD} paths with the smallest metrics are chosen as the partial candidate paths. The number N_{LSD} is usually much smaller than Q^D. We choose $N_{LSD} = \frac{1}{8} Q^D$ in this paper. If the number of the survival paths is less than N_{LSD}, the tree search will continue and moves to the next stage.

Layers M to 1 have higher SNR than the bottom D layers. Therefore, the corresponding tree search adopts the conventional radius calculation method as in [2]. The search process is nearly the same with the tree search of the first stage. The N_{LSD} survival nodes in layer $M + 1$ will be used as the parent nodes for layer M. For a given parent node in the $(j + 1)$-th layer, the following constraint condition will be checked

$$\sum_{i=j}^{N} u_{ii}^2 |x_i - z_i|^2 < R^2, \qquad (20)$$

where R is the constraint radius for the second stage and according to the method in [2], R is derived as

$$R^2 = 2\sigma^2 KM - \mathbf{y}^H (\mathbf{I} - \mathbf{H}(\mathbf{G}_m)^{-1} \mathbf{H}^H) \mathbf{y}^H + E(\mathbf{x}^H \mathbf{B} \mathbf{x}), \qquad (21)$$

where $K \ge 1$ is chosen to ensure there are enough survival paths for the LLR calculation, and the last term in (21) is caused by the modification of the Gram matrix $\mathbf{G}$. Since β is small, the last term is negligible compared to the other terms.

After the second stage of the tree search, the N_{cand} paths with the smallest metrics are chosen to form the set of candidates $\mathcal{L}$ for LLR calculation. Usually N_{cand} is chosen such that it is much smaller than Q^N, thus the complexity of LLR calculation is significantly reduced. Similar with N_{LSD}, if the number of the survival paths is less than N_{cand}, the tree search will terminate.

The set of survival paths obtained through the two-stage tree search is used to calculate the soft LLR. The calculation of the soft LLR is the same as (6) except that $\mathbf{x} \in \mathcal{L}_{n,p,b}$ replaces $\mathbf{x} \in \mathcal{S}_{n,p,b}$, where the set $\mathcal{L}_{n,p,b}$ contains the vectors in the set $\mathcal{L}$ with $c_{n,p} = b$, for $b \in \{0, 1\}$.

D. Radius for Tree Search of the first stage

For UD-MIMO systems, the constraint radius r_D used for the first stage tree search is a key parameter for achieving the near-optimal performance while keeping the complexity low. A smaller radius will reduce the number of nodes to be visited, thus reducing the complexity, but at the cost of a higher probability of missing the optimum path. The radius selection method used for the conventional LSD algorithm is no longer applicable to the first stage tree search of the UD-MIMO systems. We propose a new method for calculating the appropriate constraint radius r_D based on the statistical properties of the signals at layers $M + 1$ to N in UD-MIMO systems.

The radius r_D will be selected to ensure that the probability of missing the optimum path is less than a specified threshold value η. From (13), the metric for layers $M + 1$ to N given

in (17) can be alternatively represented by

$$\Delta_D = \mathbf{e}_D \mathbf{U}_D^H \mathbf{U}_D \mathbf{e}_D, \tag{22}$$

where $\mathbf{e}_D = [x_{M+1} - \hat{x}_{M+1}, \cdots, x_N - \hat{x}_N]^T \in \mathcal{C}^{D \times 1}$ is the error vector in layers $M+1$ to N, and the matrix $\mathbf{U}_D \in \mathcal{C}^{D \times D}$ is a sub-matrix of $\mathbf{U}_m$ that contains the $D \times D$ elements on the bottom right corner of $\mathbf{U}_m$. The matrix $\mathbf{U}_D$ is also an upper triangular matrix.

We now study the statistical properties of Δ_D, and the results will help us identify the radius r_D given the threshold probability η. From (11), the error vector $\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}}$ can be written as

$$\mathbf{e} = \mathbf{G}_m^{-1}\mathbf{B}\mathbf{x} - \mathbf{G}_m^{-1}\mathbf{H}_p^H\mathbf{w}. \tag{23}$$

Consequently, the partial error vector $\mathbf{e}_D$ containing the last D elements of $\mathbf{e}$ can be expressed as

$$\mathbf{e}_D = \bar{\mathbf{G}}_D\mathbf{B}\mathbf{x} - \bar{\mathbf{G}}_D\mathbf{H}_p^H\mathbf{w}, \tag{24}$$

where $\bar{\mathbf{G}}_D \in \mathcal{C}^{D \times N}$ is the sub-matrix of $\mathbf{G}_m^{-1}$ containing the bottom D rows of $\mathbf{G}_m^{-1}$.

Substituting (24) into (22) yields

$$\Delta_D = \mathbf{x}^H \mathbf{A}_1 \mathbf{x} - \mathbf{x}^H \mathbf{A}_2 \mathbf{w} - \mathbf{w}^H \mathbf{A}_2^H \mathbf{x} + \mathbf{w}^H \mathbf{A}_3 \mathbf{w}, \tag{25}$$

where $\mathbf{A}_1 = \mathbf{B}^H \bar{\mathbf{G}}_D^H \mathbf{U}_D^H \mathbf{U}_D \bar{\mathbf{G}}_D \mathbf{B}$, $\mathbf{A}_2 = \mathbf{B}^H \bar{\mathbf{G}}_D^H \mathbf{U}_D^H \mathbf{U}_D \bar{\mathbf{G}}_D \mathbf{H}_p^H$, and $\mathbf{A}_3 = \mathbf{H}_p \bar{\mathbf{G}}_D^H \mathbf{U}_D^H \mathbf{U}_D \bar{\mathbf{G}}_D \mathbf{H}_p^H$. Denote $\Delta_1 = \mathbf{x}^H \mathbf{A}_1 \mathbf{x}$, $\Delta_2 = -\mathbf{x}^H \mathbf{A}_2 \mathbf{w}$, and $\Delta_3 = \mathbf{w}^H \mathbf{A}_3 \mathbf{w}$. Then $\Delta_D = \Delta_1 + 2\Re(\Delta_2) + \Delta_3$, where $\Re$ is the real part operator.

It is shown in the Appendix that if β is small, for example, $\beta = 10^{-6}$, then $\Delta_2 \approx 0$, and Δ_1 is upper bounded as $\Delta_1 \leq \bar{\Delta}_1$, where $\bar{\Delta}_1$ is a constant defined as

$$\bar{\Delta}_1 = \beta \cdot \max_{\mathbf{x} \in \mathcal{S}^N} \|x\|^2 \tag{26}$$

Therefore, the metric (25) can be approximated by

$$\Delta_D \leq \bar{\Delta}_1 + \Delta_3. \tag{27}$$

Since β is small, the statistical behavior of Δ_D is dominated by Δ_3.

Let the eigenvalue decomposition of the matrix $\mathbf{A}_3$ be $\mathbf{A}_3 = \mathbf{Q}^H \mathbf{\Lambda} \mathbf{Q}$, where $\mathbf{\Lambda} = \text{Diag}(\lambda_1, \cdots, \lambda_M)$ is the diagonal matrix consisting of all the eigenvalues of $\mathbf{A}_3$, and $\mathbf{Q}$ is the orthogonal matrix containing the corresponding eigenvectors. Then we have

$$\Delta_3 = \sigma^2 \mathbf{z}^H \mathbf{\Lambda} \mathbf{z} = \sigma^2 \sum_{i=1}^M \lambda_i |Z_i|^2, \tag{28}$$

where $\mathbf{z} = \mathbf{Q}\mathbf{w}/\sigma$ follows a complex Gaussian distribution with zero mean and covariance matrix $\mathbf{I}_N$, and Z_i is the i-th element of $\mathbf{z}$. The metric Δ_3 is in the quadratic form of the complex Gaussian vector. Since the elements of $\mathbf{z}$ are mutually independent, the moment generating function (MGF) of Δ_3 conditioned on the channel matrix $\mathbf{H}_p$ is

$$M_{\Delta_3}(t|\mathbf{H}_p) = \mathbb{E}\left[e^{t\Delta_3}|\mathbf{H}_p\right] = \prod_{i=1}^M \frac{1}{1 - \sigma^2 \lambda_i t}. \tag{29}$$

The eigenvalues λ_i depend on the permuted channel matrix $\mathbf{H}_p$. For practical systems, the values of λ_i are usually

different from each other. Under the assumption of unique eigenvalues, the MGF in (29) can be expanded by means of partial fraction expansion as

$$M_{\Delta_3}(t|\mathbf{H}_p) = \sum_{i=1}^M \frac{\zeta_i}{1 - \sigma^2 \lambda_i t}, \tag{30}$$

where $\zeta_i = \prod_{l=1,l\neq i}^M \frac{\lambda_i}{\lambda_i - \lambda_l}$ is the coefficients obtained from partial fraction expansion.

The conditional probability density function (pdf) of Δ_3 can be obtained from (30) as

$$p_{\Delta_3}(x|\mathbf{H}_p) = \sum_{i=1}^M \frac{\zeta_i}{\sigma^2 \lambda_i} \exp\left(-\frac{x}{\sigma^2 \lambda_i}\right), \tag{31}$$

Then the conditional probability that Δ_D is less than the squared radius r_D^2 is

$$P(\Delta_D \leq r_D^2|\mathbf{H}_p) \leq P(\Delta_3 \leq r_D^2 - \bar{\Delta}_1|\mathbf{H}_p) = \\ \sum_{i=1}^M \zeta_i \left[1 - \exp\left(-\frac{r_D^2 - \bar{\Delta}_1}{\sigma^2 \lambda_i}\right)\right]. \tag{32}$$

The radius r_D can then be obtained by numerically solving the equation

$$\sum_{i=1}^M \zeta_i \left[1 - \exp\left(-\frac{r_D^2 - \bar{\Delta}_1}{\sigma^2 \lambda_i}\right)\right] = \eta, \tag{33}$$

where $\bar{\Delta}_1$ is a constant defined in (26). In this paper we set $\eta = 0.99$ to ensure that the probability of missing the optimum solution during the depth-first tree search is less than 1%.

To further improve the probability of finding the optimum path, we can choose a search radius qr_D^2 with $q \geq 1$. The parameter q allows us to adjust the complexity-performance tradeoff.

It should be noted that, even if there are identical eigenvalues, we can still obtain the MGF with partial fraction expansion, the form of which would be more complicated than (30) due to the identical eigenvalues. However, due to the randomness of $\mathbf{H}_p$, the probability of identical eigenvalues is negligible for practical applications. In case the identical eigenvalues do occur, the metric Δ_3 can be analyzed in the similar manner. The derivation is omitted here for brevity.

The radius obtained by solving (33) is a function of $\mathbf{H}$ through λ_i, therefore, different radii shall be selected for different channel conditions. In most practical systems with low or moderate Doppler spread, the channel usually changes slowly, thus the receiver can adaptively adjust the search radius only when the channel conditions change significantly.

IV. NUMERICAL RESULTS

A. Capacity analysis

In this subsection, the capacity and the mutual information (MI) of UD-MIMO systems are analyzed and compared to symmetric MIMO systems. If we assume that every antenna at the transmitter has equal power, the capacity of the MIMO channel due to Gaussian inputs is [7]

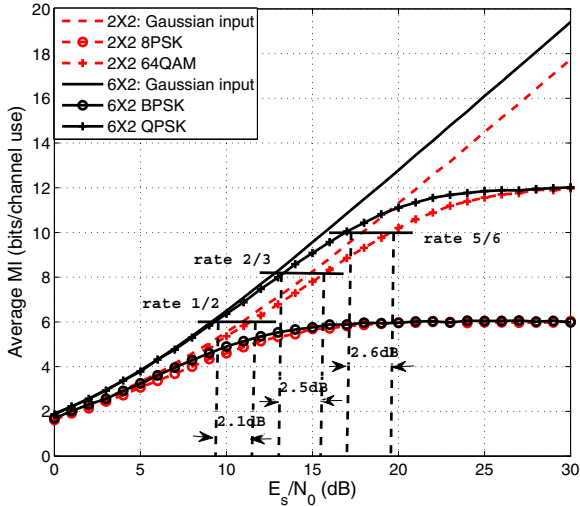$$C = \mathbb{E}\left[\log_2(\mathbf{I}_M + \frac{\rho}{N}\mathbf{H}\mathbf{H}^H)\right], \tag{34}$$

Fig. 2.   Capacity comparison between 6×2 MIMO and 2×2 MIMO.

Fig. 3.   Performance comparison between a 6×2 UD-MIMO and a 2×2 symmetric MIMO.

where $\rho = E_s/N_0$ with E_s being the average symbol energy and N_0 the one-sided power spectral density of the AWGN. For commonly used digital modulations, for example, MPSK or MQAM, the average MI is calculated as [2]

$$I(\mathbf{x}; \mathbf{y}) =$$
$$-\mathbb{E} \log \left[\frac{1}{Q^N} \frac{1}{(2\pi\sigma^2)^M} \sum_{\mathbf{x} \in \mathcal{S}^{N \times 1}} \exp \left(-\frac{1}{2\sigma^2} ||\mathbf{y} - \mathbf{Hx}||^2 \right) \right]$$
$$-[-M \log_2(2\pi e \sigma^2)], \qquad (35)$$

where $\sigma^2 = N_0/2$. We use Monte-Carlo simulation to obtain the value of the first term in (35).

Figure 2 shows the comparison of average MI between a 6×2 UD-MIMO and a 2×2 symmetric MIMO. For systems with Gaussian input, the capacity of the 6×2 UD-MIMO system is better than that of the 2×2 system due to the extra multiplexing gain provided by the additional transmit antennas. At high E_s/N_0, the two capacity curves with Gaussian inputs have the same slope because they share the same number of receive antennas. The modulation schemes of the finite alphabet cases are chosen such that a pair of UD-MIMO and symmetric MIMO systems have the same maximum MI when the code rate approaches 1. For example, the maximum MI for the 6×2 UD-MIMO system with BPSK modulation is 6 bps per channel use, which is the same as that of the 2×2 MIMO system with 8PSK modulation. Similarly, the maximum MI of the 6×2 system with QPSK or the 2×2 system with 64QAM is 12 bps per channel use. For the finite alphabet inputs cases, the average MI saturates at high E_s/N_0. The UD-MIMO systems with larger numbers of antennas and low order constellations can achieve the same average MI as symmetric MIMO systems with smaller numbers of antennas and high order constellations. For the commonly used code rate such as rate $5/6$ or rate $2/3$, the UD-MIMO with QPSK modulation outperforms its symmetric MIMO counterpart with 64QAM modulation by 2.6 dB or 2.5 dB, respectively. This result indicates that practical UD-MIMO systems outperform their symmetric MIMO system counterparts by leveraging the multiplexing gain.
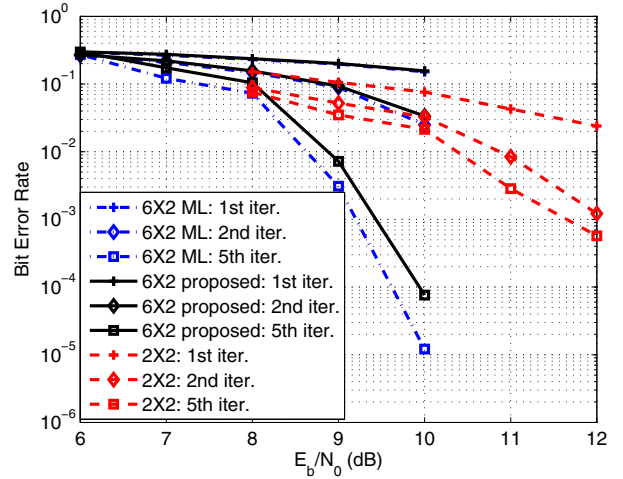
B. Performance analysis

This subsection presents the simulation results to demonstrate the performance of the proposed two-stage LSD algorithm. An UD-MIMO system with $N = 6$ transmit antennas and $M = 2$ receive antennas was compared to a 2×2 symmetric MIMO system with the same information rates. The information bits were divided into blocks with a block length of 1024 bits. A rate $1/2$ convolutional code with the generator polynomial $G = [7, 5]_8$ was employed to encode the information bits. We assumed that the receiver had perfect knowledge of the channel matrix $\mathbf{H} = \{h_{m,n}\}$. The MIMO channel coefficients $h_{m,n} \sim \mathcal{CN}(0, 1)$ were identically and independently distributed complex Gaussian random variables. Turbo equalization was applied at the receiver and the maximum number of iteration was five. For the proposed two-stage LSD algorithm, the number of survival paths for the first search stage was $N_{LSD} = \frac{1}{8}Q^{(N-M)}$. Thus $N_{LSD} = 32$ for QPSK constellation, $N_{LSD} = 512$ for 8PSK constellation and $N_{LSD} = 8192$ for 16QAM constellation. The value of β was 10^{-6}. The radius of the first search stage was chosen as $r_D'^2 = qr_D$ with r_D solved from (33) and $q = 2$. The number of survival paths for the LLR calculation was $N_{cand} = 64$ and was much smaller than Q^N.

Figure 3 shows the performance comparison between the 6×2 UD-MIMO with QPSK constellation and a 2×2 symmetric MIMO with 64QAM constellation. The information rates of the two systems were the same at 6 bits per channel use with a rate $1/2$ channel code. For the symmetric MIMO system, the conventional LSD algorithm proposed in [2] was applied. For the UD-MIMO system, the proposed algorithm, as well as the optimum ML algorithm were applied. At BER = 10^{-3}, the 6×2 UD-MIMO with QPSK modulation and the proposed two-stage LSD algorithm outperforms the 2×2 symmetric MIMO with 64QAM modulation and the conventional LSD algorithm by more than 2 dB at the fifth iteration. The gap between the proposed algorithm and the optimum algorithm is small. At BER = 10^{-4}, the performance loss of the proposed algorithm is less than 0.3 dB.
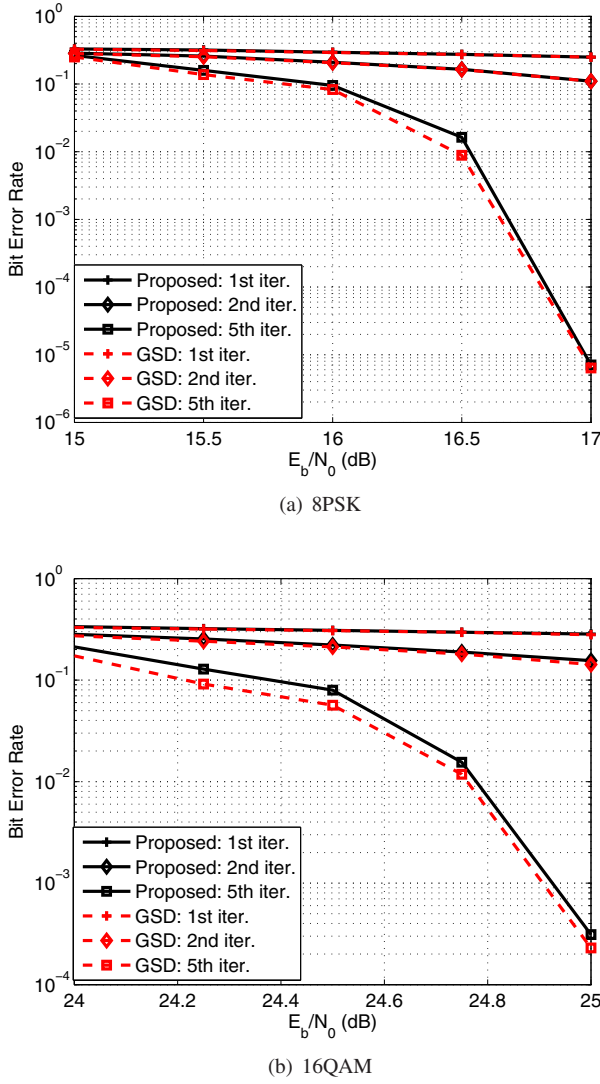
The proposed algorithm was compared to the generalized

Fig. 4. Simulated BER results for 6×2 UD-MIMO using the proposed algorithm in comparison to the GSD algorithm.

sphere decoding (GSD) algorithm proposed in [11] for 8PSK and 16QAM constellation, as shown in Fig. 4. The gap between the two algorithms is smaller than 0.1 dB at the fifth iteration at BER = 10^{-4} for 8PSK modulation. Similarly, for 16QAM modulation, at BER = 10^{-3}, the gap between the proposed algorithm and the GSD algorithm is also less than 0.1 dB at the fifth iteration. It can be concluded that the performance of the proposed two-stage LSD algorithm is nearly the same as that of the conventional GSD algorithm. However, we will show in Section V that the proposed two-stage LSD algorithm exhibits lower computational complexity than the GSD algorithm.

V. COMPLEXITY ANALYSIS

In this section, the complexity of the proposed algorithm is analyzed in comparison to the GSD algorithm proposed in [11]. The detection algorithm for UD-MIMO systems consists of four parts: channel ordering, initialization, tree search, and LLR calculation. The complexity of the tree search is the highest among the four parts and the complexity of the other

three parts is negligible. We will focus on complexity analysis of the tree search part here. The number of the arithmetic operations, including real multiplications and real additions, is used to measure the complexity. One complex multiplication is counted as four real multiplications and two real additions, and one complex addition is counted as two real additions.

A. Approximation of the Complexity

For the SD-based detection algorithm, the average number of arithmetic operations can be calculated as

$$N_{oper} = \sum_{n=1}^{N} S_n N_n, \tag{36}$$

where S_n is the number of nodes survived in layer $n+1$ and N_n is the number of arithmetic operations performed over one node in the n-th layer. Layer N is the root layer and $S_N = 1$. The value of N_n is computed as follows. According to (19), for the n-th layer, the metric is written as

$$\sum_{i=n+1}^{N} u_{ii}^2 |x_i - z_i|^2 + u_{nn}^2 |x_n - z_n|^2$$
$$= d_{n+1} + u_{nn}^2 |x_n - z_n|^2, \tag{37}$$

where $d_{n+1} = \sum_{i=n+1}^{N} u_{ii}^2 |x_i - z_i|^2$ is the metric calculated by the pervious layers. Substituting (14) into (37), we get an expression that is easier to analyze,

$$d_{n+1} + \left| u_{nn}x_n - \left[u_{nn}\hat{x}_n - \sum_{j=n+1}^{N} u_{ij}(x_j - \hat{x}_j) \right] \right|^2. \tag{38}$$

For the root layer, $d_{N+1} = 0$ and eqn. (38) is simplified as

$$u_{NN}^2 |x_N - \hat{x}_N|^2, \tag{39}$$

which requires one complex subtraction, four real multiplications and one real addition for one possible value of x_N. Thus for the root layer, $N_N = 7M$.

For the other layers, the summation in (38) requires $N - n$ complex subtractions, $N - n$ complex multiplications, which yield $8(N - n)$ real operations. The last two terms in (38) incur $8(N - n) + 4$ operations. For each possible x_n, 8 arithmetic operations are required. Thus the number of operations required for one node in layer n is

$$N_n = 8M + 8(N - n) + 4, 1 \leq n < N. \tag{40}$$

We let $k_n = N + 1 - n$ denote the dimension of the vectors in layer n. The number of survival paths in layer n equals to the number of vectors $\mathbf{x}^{(n)} \in \mathcal{S}^{k_n \times 1}$ satisfying

$$||\mathbf{x}_0^{(n)} - \mathbf{U}_n\mathbf{x}^{(n)}||^2 \leq r^2, n = 1, \cdots, N \tag{41}$$

where the vector $\mathbf{x}_0^{(n)} \in \mathcal{S}^{k_n \times 1}$ is formed by the last k_n components of $\mathbf{x}_0 = \mathbf{U}\hat{\mathbf{x}}$, and r is the constraint radius. The matrix $\mathbf{U}_n \in \mathcal{C}^{k_n \times k_n}$ is the right-lower sub-matrix of the matrix $\mathbf{U}_m$.

The number of vectors satisfying (41) can be identified with the method of lattice packing, which is suitable for lattice-based constellations such as QAM modulations. According to

[19], for a infinite lattice, the number of vectors lying in a hyper-sphere with radius r can be approximated by

$$J_r^{(n)} \simeq \frac{V_r^{(n)}}{V_{basis}^{(n)}}, \qquad (42)$$

where $V_r^{(n)}$ is the volume of a n-dimensional hyper-sphere with radius r, and $V_{basis}^{(n)}$ is the volume of the fundamental region of the lattice under consideration. For layer n, the lattice is a set of k_n-dimensional complex vectors and $V_r^{(k_n)}$ is computed as follows.

$$V_r^{(k_n)} = \frac{\pi^{k_n} r^{2k_n}}{k_n!}. \qquad (43)$$

In order to obtain the volume of the fundamental region V_{basis}, we consider layers N to $M+1$ and layers M to 1 separately. For layers N to $M+1$, according to [20], for a real lattice with bases defined by the matrix $\mathbf{\Lambda}$, the volume of the fundamental region is computed as

$$V_{basis} = \sqrt{\det(\mathbf{\Lambda}^T \mathbf{\Lambda})}, \qquad (44)$$

For the equivalent system described in (41), the matrix $\mathbf{U}_n$ can be regarded as the base matrix. Since the lattice of interests is of complex value, we can convert it to a real one by using the following equivalent real-valued representation of the complex multiplication $c = a \cdot b$

$$\begin{bmatrix} \mathcal{R}(c) \\ \mathcal{I}(c) \end{bmatrix} = \begin{bmatrix} \mathcal{R}(a) & -\mathcal{I}(a) \\ \mathcal{I}(a) & \mathcal{R}(a) \end{bmatrix} \begin{bmatrix} \mathcal{R}(b) \\ \mathcal{I}(b) \end{bmatrix}, \qquad (45)$$

where a, b and c are complex numbers. The real-valued equivalent representation of the complex matrix $\mathbf{U}_m$ can be obtained by replacing each element in $\mathbf{U}_m$ with a 2×2 matrix,

$$\begin{bmatrix} \mathcal{R}(u_{i,j}) & -\mathcal{I}(u_{i,j}) \\ \mathcal{I}(u_{i,j}) & \mathcal{R}(u_{i,j}) \end{bmatrix}, \qquad (46)$$

which yields a real matrix $\tilde{\mathbf{U}}_m \in \mathcal{R}^{2N \times 2N}$. For layer n, the base matrix $\tilde{\mathbf{U}}_n \in \mathcal{R}^{2k_n \times 2k_n}$ is the sub-matrix on the right lower corner of $\tilde{\mathbf{U}}_m$.

Since the transmitted vector is normalized to ensure unit transmit energy, a normalization factor α_n is used to multiply to the matrix $\tilde{\mathbf{U}}_n$. The volume of the fundamental region for layer n can then be written as

$$V_{basis}^{(k_n)} = \sqrt{\det(\alpha_n^2 \tilde{\mathbf{U}}_n^T \tilde{\mathbf{U}}_n)}. \qquad (47)$$

The normalization factor is determined by the constellation and lattice dimension. For example, $\alpha_n = \frac{1}{\sqrt{k_n E_{sym}}}$, where E_{sym} is the average energy of symbols before power normalization. The normalization factor α_n is different from each layers since the power of the vectors in each layer is different.

From (43) and (47), the approximated number of survival nodes in layer n for the first stage of the proposed algorithm can be calculated as

$$S_{n-1} = \frac{V_{r_D}^{(k_n)}}{V_{basis}^{(k_n)}} \simeq \frac{1}{\delta_n} \frac{\pi^{k_n} r_D^{2k_n}}{k_n! \sqrt{\det(\alpha_n^2 \tilde{\mathbf{U}}_n^T \tilde{\mathbf{U}}_n)}},$$
$$\text{for } n = M+1, \cdots, N, \qquad (48)$$

where δ_n is an adjustment factor used to account for the fact that the set of possible vectors is finite and it is different for each layers.

During the second stage, which corresponds to the tree search of the second part of the UD-MIMO system, if we still calculate the fundamental region $V_{basis}^{(k_n)}$ by using (47), then the value is usually very small, since the $\det(\mathbf{U}_m^H \mathbf{U}_m)$ is proportional to β^N. For layer N to $M+1$, the problem caused by small β is not serious because the contribution to the overall metric is also related to β. However, for layer M to 1, since these layers correspond to the full rank rows of the Gram matrix, the value of $\det(\mathbf{U}_m^H \mathbf{U}_m)$ is too small to describe the fundamental region.

In order to mitigate the effect of β, we rewrite the approximation of $V_{basis}^{(k_n)}$ as follows

$$V_{basis}^{(k_n)} \simeq \frac{1}{\beta^{k_n/2}} \sqrt{\det(\alpha_n^2 \tilde{\mathbf{U}}_n^T \tilde{\mathbf{U}}_n)}. \qquad (49)$$

Then the approximated number of the survival nodes in layer n for the second stage is written as

$$S_{n-1} = \frac{V_R^{(k_n)}}{V_{basis}^{(k_n)}} \simeq \frac{\beta^{k_n/2}}{\delta_n} \frac{\pi^{k_n} R^{2k_n}}{k_n! \sqrt{\det(\alpha_n^2 \tilde{\mathbf{U}}_n^T \tilde{\mathbf{U}}_n)}},$$
$$\text{for } n = 1, \cdots, M. \qquad (50)$$

For the conventional GSD algorithm proposed in [11], eqn. (48) and (51) still can be applied as approximations of the number of survival nodes in each layer. Since the radius used by the GSD algorithm is a constant R, we can just replace r_D in (48) with R.

Substituting S_n and N_n into (36) yields an approximation of the number of arithmetic operations. It should be noted that for the proposed algorithm, considering after the first stage of the tree search, only N_{LSD} paths with smallest metrics are chosen as the candidate paths, we let $S_M = N_{LSD}$ when calculating the number of operations.

B. Numerical Analysis

In this subsection, numerical results are presented to demonstrate the efficiency of the proposed algorithm. For the proposed two-stage LSD algorithm, part of the complexity reduction is achieved by using the proposed adaptive search radius for layers N to $M+1$, while the GSD algorithm searches nearly all the possible paths in these layers.

Table I shows the number of the survival paths of each layer for a 6×2 UD-MIMO system with 16QAM modulation at $E_b/N_0 = 24.5$ dB. The average number of the survival paths is obtained through Monte-Carlo simulation as well as by the approximation presented in Section V-A. For the approximation, the parameter δ_n is chosen as $\delta_n = 2^{4(k_n-1)}$ for $3 \leq n \leq 6$, $\delta_n = 64$ for $n = 2$ and $\delta_n = 2$ for $n = 1$. For the GSD algorithm proposed in [11], nearly all the possible paths of the first D layers are visited, which means that an exhaustive search is performed over the first D layers. On the other hand, for the proposed algorithm, some paths are pruned from layer 5. Less than 25% paths survived at layer 3, which indicates that the exhaustive search is avoided. From Table I we observe that the number of survival paths for layer

TABLE I
THE NUMBER OF SURVIVAL PATHS OF EACH LAYER FOR 6×2 UD-MIMO SYSTEM WITH 16QAM MODULATION AT $E_b/N_0 = 24.5\ dB$

Layer	GSD, simulated	GSD, approximated	Proposed, simulated	Proposed, approximated
6	16	16	16	15.81
5	256	256	233.41	239.75
4	4.10×10^3	4.10×10^3	2.60×10^3	2.37×10^3
3	6.55×10^4	6.55×10^4	1.49×10^4	1.99×10^4
2	3.81×10^3	4.43×10^3	3.81×10^3	4.57×10^3
1	147.29	136.77	147.29	140.61

TABLE II
THE NUMBER OF SURVIVAL PATHS OF THE FIRST D LAYERS

Modulation	E_b/N_0 (dB)	GSD proposed in [11]	Proposed algorithm	%
QPSK	10	256	59.42	23.21
8PSK	17	4093.33	516.85	12.63
16QAM	25	6.53×10^4	1.47×10^4	22.67

M to 1 is nearly the same for the GSD algorithm and the proposed algorithm. This is due to the fact that we count all the paths that satisfy the metric constraint in (20). However, for the proposed algorithm, at most N_{LSD} paths survive after the first stage. Thus the actual number of survival paths at layers M to 1 is less than that in Table I.

Table II shows the number of survival paths of the first D layers for a 6×2 UD-MIMO system with different constellations. The E_b/N_0 was chosen to make sure that the UD-MIMO system achieved a low BER on the order of 10^{-3}. Results in Table II indicate that, the proposed algorithm avoids the exhaustive search over the first D layers for different constellations, thus significant complexity reduction is achieved. Besides, since the average number of survival paths is close to $N_{LSD} = \frac{1}{8}Q^D$, the complexity of the sorting operation is also smaller than that of the GSD algorithm.

Table I and II indicate that, the complexity reduction of the proposed algorithm mainly comes from the avoidance of the exhaustive search over the first D layers. The probability of missing the transmitted paths is also slightly higher than that of the GSD algorithm, which causes a slight degradation of the performance.

Figure 5 shows the ratio of the average number of arithmetical operations of the proposed algorithm over that of the GSD algorithm when $M = 2$. From Fig. 5, it is observed that the complexity of the proposed algorithm is lower than that of the GSD algorithm. The larger the values of D and Q are, the more the benefit the proposed algorithm can provide. For example, when $D = 4$, less than $1/2$ of the GSD operations are required for QPSK modulation while less than $1/5$ of the GSD operations are required for 8PSK. For 16QAM, this ratio is reduced to 0.15. These results indicate that the order of the complexity for the proposed algorithm is less than that of the GSD algorithm.

The average numbers of arithmetic operations, including multiplications and additions required in the tree search, of different algorithms with 8PSK and 16QAM modulations are shown in Fig. 6. The results in the figures were obtained by counting and accumulating the number of arithmetic operations used during simulations. In Fig. 6(b), the approximated numbers of arithmetic operations calculated by using the an-
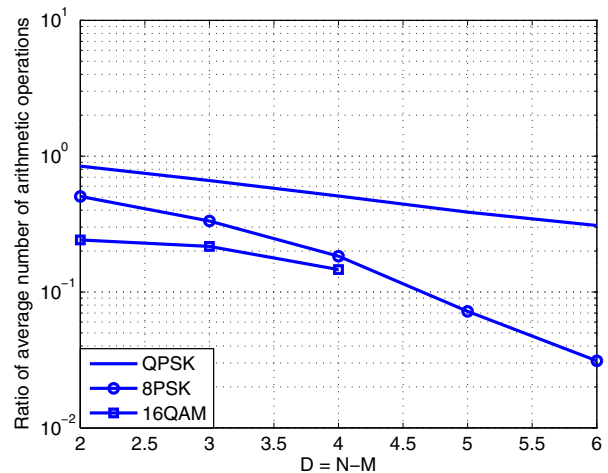

Fig. 5. The ratio of average number of mathematical operations of the proposed algorithm over that of the GSD algorithm when fixed $M = 2$.

alytical expressions in the previous subsection is also plotted. As expected, the proposed algorithm has lower complexity than the GSD algorithm [11]. For 8PSK modulation, the complexity of the proposed algorithm is about $1/5$ of the GSD algorithm while the performance loss is smaller than 0.1 dB. For 16QAM modulation, the complexity of the proposed algorithm is only $1/10$ of the GSD algorithm with a negligible performance loss.

VI. CONCLUSIONS

In this paper, a two-stage LSD algorithm was proposed for UD-MIMO systems with N transmit antennas and $M < N$ receive antennas. The proposed algorithm utilizes the unique structure of the UD-MIMO channel matrix. The N detection layers were divided into two groups. Layers 1 to M have a similar structure as a symmetric MIMO system, and layers $M + 1$ to N correspond to the extra signal dimensions in an UD-MIMO system. A modified depth-first tree search was applied to layers $M + 1$ to N to replace the exhaustive search utilized by most other detection techniques for UD-MIMO systems. A new method was proposed to adaptively
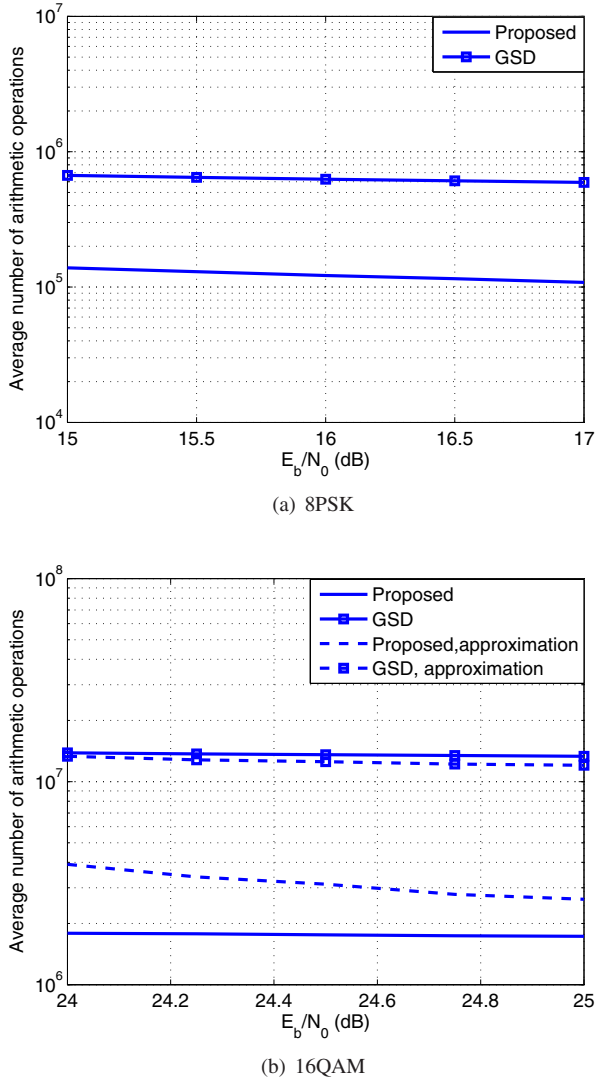
(a) 8PSK

(b) 16QAM

Fig. 6. The average number of arithmetic operations for a 6×2 UD-MIMO detector.

adjust the search radius of the depth-first tree search, such that the overall complexity was significantly reduced while maintaining a good performance. The depth-first tree search with the conventional constraint radius was employed for layers 1 to M to further reduce the complexity. Simulation results and complexity analysis showed that the proposed algorithm can achieve a similar performance as the GSD algorithm but with a much lower complexity.

APPENDIX

Evaluation of Δ_1 and Δ_2

The contributions of the metric Δ_1 and Δ_2 to the overall metric Δ_D defined in (25) are evaluated here.

Decompose $\mathbf{G}_m^{-1}$ and $\mathbf{U}_m$ as

$$\mathbf{G}_m^{-1} = \begin{pmatrix} \bar{\mathbf{G}}_1 & \bar{\mathbf{G}}_2^H \\ \bar{\mathbf{G}}_2 & \bar{\mathbf{G}}_3 \end{pmatrix}, \quad \mathbf{U}_m = \begin{pmatrix} \mathbf{U}_1 & \mathbf{U}_2 \\ \mathbf{0}_{D \times M} & \mathbf{U}_D \end{pmatrix}. \quad (51)$$

where $\bar{\mathbf{G}}_1 \in \mathcal{C}^{M \times M}$, $\bar{\mathbf{G}}_2 \in \mathcal{C}^{D \times M}$, $\bar{\mathbf{G}}_3 \in \mathcal{C}^{D \times D}$, $\mathbf{U}_1 \in \mathcal{C}^{M \times M}$, $\mathbf{U}_2 \in \mathcal{C}^{M \times D}$, and $\mathbf{U}_D \in \mathcal{C}^{D \times D}$.

Based on the fact that $\mathbf{G}_m^{-1} \mathbf{U}_m^H \mathbf{U}_m = \mathbf{I}_N$, we have

$$\begin{pmatrix} \bar{\mathbf{G}}_1 & \bar{\mathbf{G}}_2^H \\ \bar{\mathbf{G}}_2 & \bar{\mathbf{G}}_3 \end{pmatrix} \begin{pmatrix} \mathbf{U}_1^H \mathbf{U}_1 & \mathbf{U}_1^H \mathbf{U}_2 \\ \mathbf{U}_2^H \mathbf{U}_1 & \mathbf{U}_2^H \mathbf{U}_2 + \mathbf{U}_D^H \mathbf{U}_D \end{pmatrix} = \mathbf{I}_N \quad (52)$$

The following equations are obtained from (52)

$$\bar{\mathbf{G}}_2 \mathbf{U}_1^H \mathbf{U}_1 + \bar{\mathbf{G}}_3 \mathbf{U}_2^H \mathbf{U}_1 = \mathbf{0} \quad (53)$$
$$\bar{\mathbf{G}}_2 \mathbf{U}_1^H \mathbf{U}_2 + \bar{\mathbf{G}}_3 \mathbf{U}_2^H \mathbf{U}_2 + \bar{\mathbf{G}}_3 \mathbf{U}_D^H \mathbf{U}_D = \mathbf{I}_D. \quad (54)$$

Since $\mathbf{U}_1$ is an upper-triangular matrix, $\bar{\mathbf{G}}_2 \mathbf{U}_1^H = -\bar{\mathbf{G}}_3 \mathbf{U}_2^H$ from (53). Substituting the result into (54) yields $\bar{\mathbf{G}}_3 \mathbf{U}_D^H \mathbf{U}_D = \mathbf{I}$.

Define $\mathbf{E} = \bar{\mathbf{G}}_D^H \mathbf{U}_D^H \mathbf{U}_D \bar{\mathbf{G}}_D$. Based on the fact that $\bar{\mathbf{G}}_D = [\bar{\mathbf{G}}_2, \bar{\mathbf{G}}_3]$ and the results from (53) and (54), $\mathbf{E}$ can be decomposed as

$$\mathbf{E} = \begin{pmatrix} \bar{\mathbf{G}}_2^H \mathbf{U}_D^H \mathbf{U}_D \bar{\mathbf{G}}_2 & \bar{\mathbf{G}}_2^H \\ \bar{\mathbf{G}}_2 & \bar{\mathbf{G}}_3 \end{pmatrix}. \quad (55)$$

We are going to show next that, when β is small, $\bar{\mathbf{G}}_2^H \mathbf{U}_D^H \mathbf{U}_D \bar{\mathbf{G}}_2 \approx \bar{\mathbf{G}}_1$, or equivalently, $\mathbf{E} \approx \mathbf{G}_m^{-1}$.

From the decomposition in (51), the inverse of the upper triangular matrix $\mathbf{U}_m$ can be written as

$$\mathbf{U}_m^{-1} = \begin{pmatrix} \mathbf{U}_1^{-1} & -\mathbf{U}_1^{-1} \mathbf{U}_2 \mathbf{U}_D^{-1} \\ \mathbf{0} & \mathbf{U}_D^{-1} \end{pmatrix}. \quad (56)$$

Since $\mathbf{G}_m^{-1} = (\mathbf{U}_m^H \mathbf{U}_m)^{-1}$, we have

$$\mathbf{G}_m^{-1} = \mathbf{U}_m^{-1}(\mathbf{U}_m^{-1})^H = \\ \begin{pmatrix} \mathbf{U}_1^{-1}(\mathbf{U}_1^{-1})^H + \mathbf{T}\mathbf{T}^H & \mathbf{T}(\mathbf{U}_D^{-1})^H \\ \mathbf{U}_D^{-1}\mathbf{T}^H & \mathbf{U}_D^{-1}(\mathbf{U}_D^{-1})^H \end{pmatrix}, \quad (57)$$

where $\mathbf{T} = -\mathbf{U}_1^{-1} \mathbf{U}_2 \mathbf{U}_D^{-1}$.

From (57), we have $\bar{\mathbf{G}}_1 = \mathbf{U}_1^{-1}(\mathbf{U}_1^{-1})^H + \mathbf{T}\mathbf{T}^H$, and $\bar{\mathbf{G}}_2 = \mathbf{U}_D^{-1}\mathbf{T}^H$. Therefore, the sub-matrix $\bar{\mathbf{G}}_2^H \mathbf{U}_D^H \mathbf{U}_D \bar{\mathbf{G}}_2$ in (55) can be expressed as

$$\bar{\mathbf{G}}_2^H \mathbf{U}_D^H \mathbf{U}_D \bar{\mathbf{G}}_2 = \mathbf{T}(\mathbf{U}_D^{-1})^H \mathbf{U}_D^H \mathbf{U}_D \mathbf{U}_D^{-1} \mathbf{T}^H = \mathbf{T}\mathbf{T}^H. \quad (58)$$

If β is small, for example, $\beta = 10^{-6}$, then $\mathbf{U}_1^{-1}(\mathbf{U}_1^{-1})^H$ is negligible compared to $\mathbf{T}\mathbf{T}^H$ because $\mathbf{U}_1^{-1}$ is not affected by β yet $\mathbf{U}_D^{-1}$ in $\mathbf{T}$ is scaled by $\beta^{-1/2}$. Therefore $\mathbf{G}_1 \simeq \mathbf{T}\mathbf{T}^H$. Based on (51), (55) and (58), we have $\mathbf{E} \approx \mathbf{G}_m^{-1}$.

Substituting $\mathbf{E} \approx \mathbf{G}_m^{-1}$ into the definitions of Δ_1 yields $\Delta_1 \approx \beta \sum_{i=M+1}^{N} |x_i'|^2 \leq \beta \max_{\mathbf{x} \in \mathcal{S}^N} \|\mathbf{x}\|^2$, where x_i' is the i-th element of $\mathbf{V}^H \mathbf{x}$, with $\mathbf{V}$ being a unitary matrix containing the eigenvectors of $\mathbf{G}_m$.

The metric Δ_2 can be approximated by $\Delta_2 \approx \Delta_2' = \mathbf{x}^H \mathbf{B}^H \mathbf{G}_m^{-1} \mathbf{H}_p^H \mathbf{w}$. The mean of Δ_2' is 0. The variance of Δ_2' is

$$\mathbb{E}[|\Delta_2'|^2] = \sigma^2 \mathbb{E}\left(\mathbf{x}^H \mathbf{B}^H \mathbf{G}_m^{-1} \mathbf{G}_p \mathbf{G}_m^{-1} \mathbf{B}\mathbf{x}\right) \quad (59)$$

Since $\mathbf{B}$, $\mathbf{G}_m$, and $\mathbf{G}_p$ share the same eigenvectors, it is straightforward to show that $\mathbf{B}^H \mathbf{G}_m^{-1} \mathbf{G}_p = \mathbf{0}_{N \times N}$. Thus $\mathbb{E}[|\Delta_2'|^2] = 0$ and $\Delta_2' = 0$. As a result, $\Delta_2 \approx 0$.

REFERENCES

[1] U. Fincke and M. Pohst, "Improved methods for calculating vectors of short length in a lattice, including a complexity analysis," *Math. Comput.*, vol. 44, pp. 463–471, Apr. 1985.

[2] B. M. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 389–399, Mar. 2003.

[3] K. Wong, C. Tsui, R. S. Cheng, and W. Mow, "A VLSI architecture of a K-best lattice decoding algorithm for MIMO channels," in *IEEE Int. Symp. Circuits Syst.*, 2002, pp. 273–276.

[4] L. Wang, L. Xu, S. Chen, and L. Hanzo, "Generic iterative search-centre-shifting k-best sphere detection for rank-deficient SDM-OFDM systems," *Electron. Lett.*, vol. 44, no. 8, pp. 552–553, Apr. 2008.

[5] L. G. Barbero and J. S. Thompson, "Extending a fixed-complexity sphere decoder to obtain likelihood information for turbo-MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 57, no. 5, pp. 2804–2814, Sept. 2008.

[6] ——, "Fixing the complexity of the sphere decoder for MIMO detection," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2132–2142, June 2008.

[7] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Lab. Tech. J.*, vol. 1, pp. 41–59, 1996.

[8] M. Damen, K. Abed-Meraim, and J. C. Belfiore, "Generalized sphere decoding for asymmetrical space-time communication architecture," *Electron. Lett.*, vol. 36, pp. 166–167, Jan. 2000.

[9] P. Dayal and M. K. Varanasi, "A fast generalized sphere decoder for optimum decoding of under-determined MIMO systems," in *Proc. 2003 Allerton Conf. Communication, Control, and Computing*.

[10] Z. Yang, C. Liu, and J. He, "A new approach for fast generalized sphere decoding in MIMO systems," *IEEE Signal Process. Lett.*, vol. 12, no. 1, pp. 41–44, Jan. 2005.

[11] T. Cui and C. Tellambura, "An efficient generalized sphere decoder for rank-deficient MIMO systems," *IEEE Commun. Lett.*, vol. 9, no. 5, pp. 423–425, May 2005.

[12] P. Wang and T. Le-Ngoc, "A low-complexity generalized sphere decoding approach for underdetermined linear communication systems: performance and complexity evaluation," *IEEE Trans. Commun.*, vol. 57, no. 11, pp. 3376–3388, Nov. 2009.

[13] Z. Luo, S. Liu, M. Zhao, and Y. Liu, "Generalized parallel interference cancellation algorithm for V-BLAST systems," in *Proc. 2006 IEEE Int. Conf. Commun.*, pp. 3207–3212.

[14] M. Walker, J. Tao, J. Wu, and Y. Zheng, "Low complexity turbo detection of coded under-determined MIMO systems," in *Proc. 2011 IEEE Int. Conf. Commun.*, pp. 1–5.

[15] J. Wu and Y. R. Zheng, "Low complexity soft-input soft-output block decision feedback equalization," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 2, pp. 281–289, Feb. 2008.

[16] C. Qian, J. Wu, Y. R. Zheng, and Z. Wang, "A modified fixed sphere decoding algorithm for under-determined MIMO systems," in *Proc. 2012 IEEE GLOBECOM*, pp. 4482–4487.

[17] P. Robertson, E. Villebrun, and P. Hoeher, "A comparison of optimal and sub-optimal map decoding algorithms operating in the log domain," in *Proc. 1995 IEEE Int. Conf. Commun.*, vol. 2, pp. 1009–1013.

[18] J. Jalden, L. Barbero, B. Ottersten, and J. Thompson, "Full diversity detection in MIMO systems with a fixed-complexity sphere decoder," in *Proc. 2007 IEEE Int. Conf. Acoustics, Speech and Signal Processing,*, vol. 3, pp. 49–52.

[19] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups: Charpter 1.*   New York: Springer, 1999.

[20] D. Wübben, D. Seethaler, J. Jaldén, and G. Matz, "Lattice reduction," *IEEE Signal Process. Mag.*, vol. 28, no. 3, pp. 70–91, May 2011.

Chen Qian received his B.S. degree in electronic engineering from Tsinghua University, Beijing, China in July 2010. From September 2010, he has been a Ph. D. candidates at the Department of Electronic Engineering of Tsinghua University. His research interests include MIMO detection algorithm and channel coding and modulation.

Jingxian Wu (S'02-M'06) received the B.S. (EE) degree from the Beijing University of Aeronautics and Astronautics, Beijing, China, in 1998, the M.S. (EE) degree from Tsinghua University, Beijing, China, in 2001, and the Ph.D. (EE) degree from the University of Missouri at Columbia, MO, USA, in 2005. He is currently an Assistant Professor with the Department of Electrical Engineering, University of Arkansas, Fayetteville. His research interests mainly focus on wireless communications and wireless networks, including ultra-low power communications, energy efficient communications, high mobility communications, and cross-layer optimization, etc. He is an Associate Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, an Associate Editor of the IEEE ACCESS, and served as an Associate Editor of the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY from 2007 to 2011. He served as a co-chair for the 2012 Wireless Communication Symposium of the IEEE International Conference on Communication, and a co-chair for the 2009 Wireless Communication Symposium of the IEEE Global Telecommunications Conference. Since 2006, he has served as a Technical Program Committee Member for a number of international conferences, including the IEEE Global Telecommunications Conference, the IEEE Wireless Communications and Networking Conference, the IEEE Vehicular Technology Conference, and the IEEE International Conference on Communications.
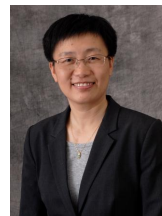
Yahong Rosa Zheng (SM'07) received the B.S. degree from the University of Electronic Science and Technology of China, Chengdu, China, in 1987, and the M.S. degree from Tsinghua University, Beijing, China, in 1989, both in electrical engineering. She received the Ph.D. degree from the Department of Systems and Computer Engineering, Carleton University, Ottawa, Canada, in 2002. She was an NSERC Postdoctoral Fellow from Jan. 2003 to April, 2005 at University of Missouri-Columbia. In fall 2005, she joined the Department of Electrical and Computer Engineering at the Missouri University of Science and Technology where, currently, she is an Associate Professor. Her research interests include array signal processing, wireless communications, and wireless sensor networks. She has served as a Technical Program Committee (TPC) member for many IEEE international conferences, including IEEE Vehicular Technology Conference, IEEE GlobeCom, and IEEE ICC, and IEEE Wireless Communications and Networking Conference, etc. She served as Wireless Communications Symposium co-chair for ICC 2014 and Globecom 2013. She also served as an Associate Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS for 2006-2008. She is currently Associate Editor for IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. She has been a Senior Member of the IEEE since 2007. She is the recipient of an NSF CAREER award in 2009.

Zhaocheng Wang (SM'10) received his B.S., M.S. and Ph.D. degrees from Tsinghua University in 1991, 1993 and 1996, respectively. From 1996 to 1997, he was with Nanyang Technological University (NTU) in Singapore as a Post Doctoral Fellow. From 1997 to 1999, he was with OKI Techno Centre (Singapore) Pte. Ltd., firstly as a research engineer and then as a senior engineer. From 1999 to 2009, he worked at SONY Deutschland GmbH, firstly as a senior engineer and then as a principal engineer. He is currently a Professor at the Department of Electronic Engineering, Tsinghua University. His research areas include wireless communications, digital broadcasting and millimeter wave communications. He holds 31 granted US/EU patents and has published over 80 technical papers. He has served as technical program committee co-chair/member of many international conferences. He is a Fellow of the Institution of Engineering and Technology.