RESEARCH **Open Access**

CrossMark

New rank detection methods for reduced-rank MIMO systems

Qian Wang[*] and Yindi Jing

Abstract

In practical multi-input multi-output (MIMO) systems, the channel matrices often have reduced rank. Reliable detection of the channel rank is essential in achieving the significant gain provided by MIMO configuration. Existing work on MIMO channel rank detection assume a static channel model, so the proposed methods only consider the noise distributions while the distributions of the MIMO channels are not considered. In this paper, we employ a random channel model and propose three threshold-based rank detection methods which take into account the distributions of both the channels and the noises. In our first algorithm, following existing single-threshold rank detection scheme, we rigorously derive an analytical lower bound on the correct rank detection probability and propose a systematic threshold selection method by maximizing the lower bound. Then we propose two new rank detection methods which use multiple thresholds, where each threshold corresponds to one possible rank value. The thresholds are optimized based on the derived lower bounds on the rank detection probability for different channel rank values. The convergence and complexity of the proposed algorithms are analyzed. Simulation results on the correct rank detection probability of the proposed schemes are provided to show their advantage over existing schemes. The mean-squared-error (MSE) and outage probability are also simulated to show the importance of reliable rank detection in MIMO communications.

Keywords: Rank detection; Reduced-rank MIMO; Threshold optimization

1 Introduction

In the past two decades, a configuration, called multi-input multi-output (MIMO) system, which utilizes multiple antennas at both the transmitter and the receiver, has been extensively studied [1, 2]. MIMO systems have outstanding performance in increasing the data throughput and link range over single-antenna systems without consuming additional bandwidth or transmit power. Early studies focused on MIMO systems with full-rank channels. But in many practical propagation environments such as the number of surrounding scatterers which is finite and limited, the MIMO channel matrix is likely to have reduced rank [3, 4], especially when the channel dimension is large [5]. In [3], the finite scatterer channel model was described and its capacity was analyzed. In [4], the model was applied to the large-scale MIMO system, referred to as massive MIMO. When the channel matrix

has reduced rank, the number of its entries is larger than its real dimension, and thus designs based on full-rank channels become inefficient. This motivates the research on reduced-rank technologies for MIMO systems [6–23].

In [6–14], various reduced-rank filtering technologies were proposed, where a reduced-rank transformation is first used on the observed signal vector to obtain a lower-dimension vector, then a filter is designed to estimate the desired signal vector. The reduced-rank transformation can lower the order of the filter and thus require less computation complexity and shorter training length. If the covariance matrix of the observed data vector is known, the optimal reduced-rank design is singular value decomposition (SVD)-based [13, 14]. But the covariance matrix estimation needs a very large number of training symbols. To overcome this drawback, a reduced-rank multistage Wiener filter design was proposed in [6], which relaxes the requirement on explicit estimate of the signal subspace. It was later extended to its adaptive versions in [7, 8] and applied to MIMO equalization in [9]. Another related design is the auxiliary-vector filtering algorithm [10].

*Correspondence: qw8@ualberta.ca
Department of Electrical and Computer Engineering, University of Alberta, 9107-116 Street, T6G 2V4 Edmonton, Canada

Springer

Compared with the multistage Wiener filter, it achieves better performance and does not involve matrix inversion operation. Joint transformation and filter designs can be found in [11, 12].

Another important issue for MIMO communications is the channel rank detection and channel matrix estimation, since most MIMO transceiver techniques require channel state information (CSI) at the transmitter side and/or the receiver side for smart signal processing. For example, in the MIMO multiplexing transmissions, the transmitter should align multiple data streams with the eigenspaces of the channel. To achieve this, precise estimates on the channel rank and channel matrix are needed. This paper is concerned with the channel rank detection, which is an important part of channel estimation. Thus, in what follows, we explain related literature on channel estimation and rank detection.

Some existing work on MIMO channel estimation focus on full-rank MIMO channel matrices with independent or correlated entries [24–27]. Their channel estimation schemes are entry-based, where the unknown channel matrix is parameterized by its entries. For reduced-rank channel estimation, these schemes will cause performance degradation. A more natural and efficient approach is to use SVD-based channel estimation methods [15–23]. It was shown in [16, 17] that the maximum-likelihood (ML) estimation of the reduced-rank MIMO channel with Gaussian noise is the truncated SVD method. In truncated SVD, if the channel rank is known to be r, the MIMO channel matrix is estimated from the SVD of the received signal-plus-noise matrix, by keeping the largest r singular values and their corresponding singular vectors. While the truncated SVD method traces back to the 1930s [28], it was rediscovered in [16, 17] for MIMO channel estimation and further improved to reduce the mean squared error (MSE) of the estimation [18–22]. One improved scheme is shrinkage-and-threshold SVD, where the truncated singular values are further shrunk to remove the noise effect.

In both truncated SVD and shrinkage-and-threshold SVD, the channel estimation accuracy largely depends on the correct truncation of the singular values, which is the rank detection of the MIMO channel matrix. Thus correct rank detection can improve the channel estimation quality, which is crucial for advanced MIMO techniques such as beamforming and power allocation among data streams. Furthermore, for a MIMO channel matrix, the rank is the indicator of how many data streams can be spatially multiplexed on the channel, and the data streams are represented by the singular values and the corresponding singular vectors. Similarly, for a multi-user system with multiple antennas at the base station, the rank of the MIMO channel matrix from all users to the base station determines how many users can be served by the base station within the same time-frequency bandwidth. Thus, accurate channel rank detection is an important part of channel estimation and is essential for MIMO systems.

Various rank detection methods are available in the literature [17–23]. In [17], a minimum description length (MDL)-based rank detection was used for MIMO channels. A threshold is calculated at each instance of the channel to minimize the MDL. The MDL-based detection aims at minimizing the MSE. It also requires a large number of samples to work. In [23], several rank detection methods were proposed for a real-valued channel matrix. In the proposed method, the singular values or functions of singular values are compared with a threshold for rank detection. Lower and upper bounds on the threshold selection were discussed. Josse and Sardy [22] considered both the MIMO channel rank detection and the shrinkage of the singular values for the channel estimation. The rank detection scheme is also threshold-based, where two threshold selection methods are proposed. In the first method, the threshold parameter and the shrinkage parameter are jointly optimized for each channel realization to minimize a Stein unbiased risk estimate (SURE) of the MSE of the MIMO channel estimation. In the second method, the threshold is calculated from the distribution of the largest singular value of the noise matrix. In [18–21], for asymptotic MIMO channels where both dimensions of the channel matrix approach infinity with a fixed ratio, simple closed-form thresholds were derived for threshold-based rank detection.

1.1 Summary of our results and distinction to existing work

This paper is on the rank detection of MIMO channels. We propose a threshold optimization method for the traditional single-threshold-based rank detection scheme and two new multiple-threshold-based rank detection schemes. For the threshold optimization of the traditional single-threshold scheme, we first derive a closed-form lower bound on the probability of correct rank detection based on the a priori channel rank distribution, then the optimal threshold is decided via the maximization of the lower bound. For the two new multiple-threshold schemes, different thresholds are used for different possible rank values, and each threshold is derived by maximizing the lower bound on the probability of correct rank detection when a specific rank value is assumed. Properties (e.g., well-posedness, convergence, complexity) of the two new schemes are discussed. Simulation results on the probability of correct rank detection of the proposed schemes are shown, and their advantages over existing schemes are discussed. The MSE and outage probability are also simulated to show that better rank detection improves the MSE and outage performance of MIMO systems.

Our model, problem formulation, and methods differ from existing ones in the following major aspects. First, we assume a random channel matrix where the channel entries follow Rayleigh flat-fading and the distribution of the channel coefficients is taken into account in the threshold optimization. On the contrary, in all existing work, the channels are assumed to be static and the distribution of the channel matrix is not used in the rank detection designs [14, 17–23]. Also, in our model, a general training length and unitary training matrix are considered, while most existing work (e.g., [17–23]) apply to identity training matrix only, where the training length equals the number of transmit antennas. Finally, in this work, we use the probability of correct rank detection as the performance measure and optimization objective, while existing work (except [23]) targeted at minimizing the MSE of the MIMO channel estimation [17–22].

In what follows, we clarify the major difference of our channel rank detection and channel estimation problem to the reduced-rank filter design problem in [6–14]. The goal of channel estimation is to estimate the channel matrix itself given a limited training time. In MIMO communications, channel estimation is usually required in the transceiver designs to optimize the communication performance such as outage probability and bit error rate. For the filter design, the goal is to obtain a precise estimate of the signals by filtering the observations, and usually the MSE is used as the design criterion. Regardless of the channel rank, a reduced-rank filter can be used to lower the computational complexity and required training length. Naturally, precise channel rank detection and channel estimation can be helpful in reduced-rank filter design, but it is not a necessary step in filter design. Also the optimal rank for the filter may not be the true rank of the channel matrix.

The rest of this paper is organized as follows. In Section 2, the MIMO channel model, the truncated SVD-based channel estimation, and the rank detection problem are presented. In Section 3, for the traditional single-threshold-based rank detection, we derive a lower bound on the probability of correct rank detection and propose to optimize the threshold via the maximization of the lower bound. Two new rank detection algorithms based on multiple thresholds corresponding to different possible rank values are introduced in Section 4, as well as discussions on their properties. Simulation results on the probability of correct rank detection, MSE, and outage probability are shown in Section 5. Section 6 contains the conclusions.

In this paper, bold upper case letters and bold lower case letters are used to denote matrices and vectors, respectively. For a matrix $\mathbf{A}$, its Hermitian, trace, rank, and determinant are denoted by $\mathbf{A}^*$, $\mathrm{tr}(\mathbf{A})$, $\mathrm{rank}(\mathbf{A})$, and $\det(\mathbf{A})$, respectively. $\mathbf{I}_n$ is the $n \times n$ identity matrix. $\mathbb{E}(\cdot)$ denotes the average operator, and diag$\{a_1, \ldots, a_n\}$ denotes the diagonal matrix whose diagonal entries starting in the upper left corner are $a_1, \cdots, a_n$.

2 Reduced-rank MIMO channel and rank detection problem

2.1 Reduced-rank MIMO channel model

A MIMO system with M transmit antennas and N receive antennas is considered. The channels are assumed to be flat-fading and block-fading. Denote the $M \times N$ channel matrix between the transmitter and the receiver as $\mathbf{H}$ with its (i, j)-th entry being the channel from the ith transmit antenna to the jth receive antenna. Define

$$L \triangleq \max\{M, N\}, \qquad (1)$$
$$K \triangleq \min\{M, N\}. \qquad (2)$$

Denote the rank of the channel matrix as r, i.e., $\mathrm{rank}(\mathbf{H}) = r$. If $r = K$, the channel has full rank. If $r < K$, the channel has reduced rank. For a reduced-rank MIMO channel, the number of degrees of freedom in the channel matrix is less than its dimension. A rank analysis on typical propagation environments shows that MIMO channels often experience rank deficiency [3, 4, 29]. For example, in [29], the rank distribution of 8×8 MIMO channels under four scenarios, namely generalized typical urban, generalized bad urban, generalized hill terrain, and generalized rural area, is reported. For all four scenarios, the probability that the channel matrix has full rank is 0. Especially for the scenarios of generalized typical urban and generalized hill terrain, the rank of the 8×8 MIMO channel is always no higher than 4.

A typical reduced-rank channel model for MIMO system is the finite scatterers/dimensional channel [3, 4], where the number of (clusters of) scatterers is finite. The rank of the channel matrix is not only constrained by the number of transmit and receive antennas but also constrained by the number of scatterers. In the finite scatterers model in [3], under the assumptions that both the transmit and receive elements are isotropic and uncoupled and the signal bandwidth is narrow compared with the overall channel bandwidth, the MIMO channel matrix can be described as the product of a full-rank steering matrix and a propagation matrix. The propagation matrix models independent fast fading, geometric attenuation, and shadow fading. Following the model in [3] and the rank factorization in [14, 29], we assume that the MIMO channel matrix has the following decomposition:

$$\mathbf{H} \triangleq \mathbf{AB}, \qquad (3)$$

where $\mathbf{A}$ is an $M \times r$ full-rank matrix and $\mathbf{B}$ is a $r \times N$ full-rank rectangular unitary matrix. In this work, we focus on Rayleigh fading by assuming that entries of $\mathbf{A}$ are independent and identically distributed (i.i.d.) and follow circularly symmetric complex Gaussian (CSCG) distribution,

with zero mean and unit-variance, i.e., $a_{ij} \sim \mathcal{CN}(0,1)$, where a_{ij} is the (i,j)-th entry of $\mathbf{A}$. It can thus be shown straightforwardly that each entry of H has CSCG distribution and $h_{i,j} \sim \mathcal{CN}(0, \|\mathbf{b}_j\|_F^2)$, where $\mathbf{b}_j$ is the jth column of $\mathbf{B}$. It is noteworthy that the rank detection schemes proposed in this paper are not constrained to Rayleigh distribution and can be extended to more general channel fading models. The schemes are also not constrained to the unitary $\mathbf{B}$ case and can be extended to any general deterministic $\mathbf{B}$ matrix.

2.2 Training model, SVD-based channel estimation, and rank detection problem

To estimate the channel matrix, a training process is needed. Denote the length of the training period as T and average transmit power used for each training time slot as P. During the training period, the transmitter sends $\sqrt{PT/M}\mathbf{S}$, where $\mathbf{S}$ is the $T \times M$ pilot matrix. For the observability of the channel rank detection model with respect to all possible rank values, we assume that $T \geq M$, which guarantees that the number of independent equations in the training equation is no less than the number of independent unknown coefficients in the channel matrix. We further assume that $\mathbf{S}$ is unitary, i.e., $\mathbf{S}^*\mathbf{S} = I_M$, which means that the pilot vector sent from each transmit antenna is orthogonal to each other and has the same energy. Denote the $T \times N$ matrix received at the receiver as $\mathbf{Y}$. We have

$$\mathbf{Y} = \sqrt{\frac{PT}{M}}\mathbf{SH} + \mathbf{W}, \tag{4}$$

where $\mathbf{W}$ is the $T \times N$ noise matrix. Entries of the noise matrix are assumed to be i.i.d. CSCG random variables with zero mean and unit-variance. The pilot and the noise are assumed to be independent to the channel matrix, which applies to most practical systems. It is noteworthy that the channel model and training model also apply to multi-user massive MIMO systems with M single-antenna users and N base station antennas, where channel training is conducted in the uplink [5].

The channel estimation problem is to estimate $\mathbf{H}$ from the observation $\mathbf{Y}$. To do this, we first transform the training equation in (4) to obtain a more direct relationship between the channel and the (transformed) received signal. Define

$$\tilde{\mathbf{Y}} \triangleq \sqrt{\frac{M}{PT}}\mathbf{S}^*\mathbf{Y}, \tag{5}$$

which is the $M \times N$ transformed received signal matrix. By left-multiplying both sides of (4) with $\sqrt{\frac{M}{PT}}\mathbf{S}^*$, we have

$$\tilde{\mathbf{Y}} = \mathbf{H} + \tilde{\mathbf{W}}, \tag{6}$$

where

$$\tilde{\mathbf{W}} \triangleq \sqrt{\frac{M}{PT}}\mathbf{S}^*\mathbf{W}. \tag{7}$$

Since $\mathbf{S}^*$ is an $M \times T$ unitary matrix, $\tilde{\mathbf{W}}$ is an $M \times N$ random matrix. Entries of $\tilde{\mathbf{W}}$ can be shown to be i.i.d. CSCG random variables with zero mean and their variances are $M/(PT)$. From (6), $\tilde{\mathbf{Y}}$ is a noisy observation of the channel matrix $\mathbf{H}$ with white Gaussian noises.

Let $\tilde{\mathbf{Y}} = \mathbf{P}\text{diag}\{\sigma_1, \cdots, \sigma_K\}\mathbf{Q}^*$ be the SVD of $\tilde{\mathbf{Y}}$, where $\mathbf{P}$ and $\mathbf{Q}$ are $M \times K$ and $K \times N$ unitary matrices and σ_i's are in non-increasing order, i.e., $\sigma_1 \geq \cdots \geq \sigma_K \geq 0$. If the rank of $\mathbf{H}$ is known to be r, the ML estimation of $\mathbf{H}$ has been proved to be the truncated SVD of $\tilde{\mathbf{Y}}$ given as follows [28]:

$$\hat{\mathbf{H}} = \mathbf{P}\text{diag}\{\sigma_1, \cdots, \sigma_r, 0, \cdots, 0\}\mathbf{Q}^*. \tag{8}$$

In (8), an estimation with rank r is obtained by keeping the subspaces with respect to the r strongest singular values of $\tilde{\mathbf{Y}}$. Subspaces with respect to the $K - r$ smallest singular values are seen as the noise effect and are ignored. This process guarantees that the estimator has the same rank with the real channel.

On the other hand, if the channel rank is unknown or the channel has full rank, an entry-based ML estimation can be obtained as [24]

$$\hat{\mathbf{H}}_{\text{entry}} = \sqrt{\frac{M}{PT}}(\mathbf{S}^*\mathbf{S})^{-1}\mathbf{S}^*\mathbf{Y}. \tag{9}$$

The entry-based estimation leads to a full-rank matrix. When $r = K$, i.e., the channel has full rank, the SVD-based estimation and entry-based estimation are equivalent. When $r < K$, i.e., the channel has reduced rank, the entry-based estimation will contain subspaces due to the noise effect only and thus have a worse performance.

Therefore, the rank detection is an essential problem for the SVD-based estimation. Wrong rank detection will lead to channel estimation error. In addition, it can degrade the performance of the MIMO communications. If the detected rank is smaller than r, some singular values and the corresponding singular spaces existing in the channel matrix may be detected as the noise effect only, and the subspaces will be lost in the estimated MIMO channel. On the other hand, if the detected rank is larger than r, some singular values and the corresponding singular spaces which do not exist in the channel matrix but appear in $\tilde{\mathbf{Y}}$ because of the noise disturbance may be detected as part of the channel matrix. In MIMO communications, information and power will be allocated to such subspaces, which cause loss of information and wasting of power since the subspaces do not exist in the channel. Our problem of this paper is to detect the channel rank from the received signal $\mathbf{Y}$, or the transformed received signal $\tilde{\mathbf{Y}}$. It is noteworthy that the rank detection does not require extra training since the same observations can be used for both rank detection and channel estimation.

3 Threshold selection for single-threshold-based rank detection

We can detect the rank of $\mathbf{H}$ from the singular values of $\tilde{\mathbf{Y}}$. Threshold-based algorithm appears to be a natural and common strategy [23], where the rank of $\mathbf{H}$ is detected as the number of singular values of $\tilde{\mathbf{Y}}$ that are larger than the threshold. With this scheme, singular values of $\tilde{\mathbf{Y}}$ that are smaller than the threshold are seen as the effect of the noise only; while singular values of $\tilde{\mathbf{Y}}$ that are larger than the threshold are seen as the effect of nonzero component of the channel matrix with small noise disturbance.

Let ϵ_{th} be the threshold. Recall that σ_i's are the singular values of $\tilde{\mathbf{Y}}$ in non-increasing order. The single-threshold-based rank detection scheme, denoted as RD1, can be represented as follows:

$$\text{RD1:} \quad r = \max_{\sigma_i \geq \epsilon_{\text{th}}} \{i\}. \tag{10}$$

If no singular value is larger than ϵ_{th}, i.e., $\sigma_1 < \epsilon_{\text{th}}$, the rank detection result is set to be 1 since the rank of a MIMO channel cannot be 0. The algorithm is given in Algorithm 1.

Algorithm 1 Rank detection algorithm with single threshold

1: **for** $(i = K : 1)$ **do**
2:    **if** $\sigma_i \geq \epsilon_{\text{th}}$ **then**
3:       The rank of $\mathbf{H}$ is detected as i; break;
4:    **if** $\epsilon_{\text{th}} > \sigma_1$ **then**
5:       The rank of $\mathbf{H}$ is detected as 1.

This single-threshold-based rank detection idea is not new and was proposed and used in [17–23]. But the major challenge of using this scheme for rank detection lies in the selection of the threshold ϵ_{th}. Appropriate selection of the ϵ_{th} value is crucial to the detected result.

In this section, we first derive a lower bound on the probability of correct detection in Section 3.1, then propose a systematic method for the threshold selection based on maximizing the lower bound in Section 3.2, and finally discuss the difference of the proposed method with existing ones in Section 3.3.

3.1 Derivation of a lower bound on the conditional probability of correct rank detection

To find a systematic way of optimizing the threshold, we first derive a lower bound on the probability of correct detection conditioned on an arbitrary rank value of the MIMO channel matrix. The lower bound takes into consideration the system dimensions (e.g., T, M, and N), training power P, and the distributions of the channel coefficients and the noises. It will be used in the threshold optimization in later sections.

To help presenting our results, we first introduce the following definitions. Define the $K \times K$ matrix $\boldsymbol{F}^{(1)}(\mu)$ and the $r \times r$ matrix $\boldsymbol{F}^{(r)}(\mu)$, whose (i,j)-th entries are:

$$\left[\boldsymbol{F}^{(1)}(\mu) \right]_{i,j} \triangleq \gamma(L - K + i + j - 1, \mu), \tag{11}$$

$$\left[\boldsymbol{F}^{(r)}(\mu) \right]_{i,j} \triangleq \Gamma(M - r + i + j - 1, \mu), \tag{12}$$

where $\gamma(k, u)$ and $\Gamma(k, u)$ are the lower and upper incomplete gamma functions [30], respectively.

The following proposition on the probability of correct rank detection under the condition that the rank of the MIMO channel matrix is r (for $1 \leq r \leq K$) is derived.

Proposition 1. *If the rank of $\mathbf{H}$ is r, the probability of correct rank detection of Algorithm 1 with threshold ϵ_{th} has the following lower bound:*

$$\phi_r(\epsilon_{th}) \triangleq C_1 C_2 \cdot \det\left(\boldsymbol{F}^{(r)}\left(4\epsilon_{\text{th}}^2\right) \right) \det\left(\boldsymbol{F}^{(1)}\left(\frac{PT}{M} \epsilon_{\text{th}}^2 \right) \right), \tag{13}$$

where

$$C_1 = \prod_{i=1}^{r} [(M - i)! \, (r - i)!]^{-1}, \quad C_2 = \prod_{i=1}^{K} [(L - i)! \, (K - i)!]^{-1}. \tag{14}$$

Proof 1. *Recall that σ_i's are the singular values of $\tilde{\mathbf{Y}}$ in non-increasing order, i.e., $\sigma_1 \geq \cdots \geq \sigma_r \geq \cdots \geq \sigma_K \geq 0$. Let λ_i and γ_i be the singular values of $\mathbf{H}$ and $\tilde{\mathbf{W}}$, respectively, both in non-increasing order, i.e., $\lambda_1 \geq \cdots \geq \lambda_r \geq 0$ and $\gamma_1 \geq \cdots \geq \gamma_r \geq \cdots \geq \gamma_K \geq 0$. Since $rank(\mathbf{H}) = r$, we have $\lambda_{r+1} = \cdots = \lambda_K = 0$.*

We will first show that when $\lambda_r \geq 2\epsilon_{th} \geq 2\gamma_1$, our algorithms will detect the rank of $\mathbf{H}$ as r, which is the correct detection. According to [23], from (6), we have for all $i = 1, \cdots, K$,

$$|\sigma_i - \lambda_i| \leq \gamma_i \leq \gamma_1. \tag{15}$$

By noticing that $\lambda_{r+1} = 0$, from (15) with $i = r + 1$, we have $\sigma_{r+1} \leq \gamma_1$. Thus when $\lambda_r \geq 2\epsilon_{\text{th}} \geq 2\gamma_1$, we have $\sigma_{r+1} \leq \gamma_1 \leq \epsilon_{\text{th}}$. Also from (15) with $i = r$, $\sigma_r \geq \lambda_r - \gamma_1 \geq 2\epsilon_{\text{th}} - \gamma_1 \geq \epsilon_{\text{th}}$. By noticing that σ_i's are in non-increasing order, we can conclude that the rank detection result of Algorithm 1 is r, which is the correct detection.

Thus, a lower bound on the probability of correct detection is obtained as follows,

$$\mathbb{P}(\text{correct detection}|\text{rank}(\mathbf{H}) = r)$$
$$\geq \mathbb{P}\left(\lambda_r \geq 2\epsilon_{\text{th}} \geq 2\gamma_1 | \text{rank}(\mathbf{H}) = r \right)$$
$$= \mathbb{P}\left(\lambda_r \geq 2\epsilon_{\text{th}} \, \& \, \gamma_1 \leq \epsilon_{\text{th}} | \text{rank}(\mathbf{H}) = r \right)$$
$$= \mathbb{P}\left(\lambda_r \geq 2\epsilon_{\text{th}} | \text{rank}(\mathbf{H}) = r \right) \mathbb{P}(\gamma_1 \leq \epsilon_{\text{th}}), \tag{16}$$

where the last step is because that γ_1, the largest eigenvalue of $\tilde{\mathbf{W}}$, is independent of both λ_r and the rank of $\mathbf{H}$.

Recall that our channel is modeled as $\mathbf{H} = \mathbf{AB}$, where the $M \times r$ matrix $\mathbf{A}$ has independent entries following $\mathcal{CN}(0,1)$ and $\mathbf{B}$ is a $r \times N$ unitary matrix. Then $\mathbf{HH}^* = \mathbf{ABB}^*\mathbf{A} = \mathbf{AA}^*$, which is an $M \times M$ central Wishart matrix with degree r. The singular values of $\mathbf{H}$ are the square roots of the eigenvalues of $\mathbf{HH}^*$. The cumulative density function (CDF) of the smallest non-zero eigenvalue of $\mathbf{HH}^*$ is known as follows [30]:

$$F_{\omega_r}(\mu) = 1 - C_1 \cdot \det\left(\mathbf{F}^{(r)}(\mu)\right), \tag{17}$$

where $\mathbf{F}^{(r)}$ and C_1 are defined in (12) and (14), respectively. Thus,

$$\begin{aligned}\mathbb{P}\left(\lambda_r \geq 2\epsilon_{\text{th}}|\text{rank}(\mathbf{H}) = r\right) &= 1 - F_{\omega_r}\left(4\epsilon_{\text{th}}^2\right) \\ &= C_1 \cdot \det\left(\mathbf{F}^{(r)}\left(4\epsilon_{\text{th}}^2\right)\right).\end{aligned} \tag{18}$$

Next, we calculate $\mathbb{P}(\gamma_1 \leq \epsilon_{\text{th}})$. Recall that entries of $\tilde{\mathbf{W}}$ are i.i.d. following $\mathcal{CN}(0, M/PT)$. Thus, $(PT/M)\tilde{\mathbf{W}}\tilde{\mathbf{W}}^*$ is an $M \times M$ central Wishart matrix with degree N. The CDF of its largest eigenvalue is known to be [30]

$$F_{\omega_1}(\mu) = C_2 \cdot \det\left(\mathbf{F}^{(1)}(\mu)\right), \tag{19}$$

where $\mathbf{F}^{(1)}$ and C_2 are defined in (11) and (14), respectively. Therefore,

$$\mathbb{P}(\gamma_1 \leq \epsilon_{\text{th}}) = C_2 \det\left(\mathbf{F}^{(1)}\left(\frac{PT}{M}\epsilon_{\text{th}}^2\right)\right). \tag{20}$$

By using (18) and (20) in (16), the lower bound in (13) is obtained.

3.2 Threshold optimization for Algorithm 1

Assume that the a priori probability mass function of the channel rank, $\mathbb{P}(\text{rank}(\mathbf{H}) = r)$ for $r = 1, \cdots, K$, is known. Define

$$\begin{aligned}\phi(\epsilon_{\text{th}}) \triangleq &C_2 \det\left(\mathbf{F}^{(1)}\left(\frac{PT}{M}\epsilon_{\text{th}}^2\right)\right) \cdot \sum_{r=1}^{K} \\ &C_1 \det\left(\mathbf{F}^{(r)}\left(4\epsilon_{\text{th}}^2\right)\right)\mathbb{P}(\text{rank}(\mathbf{H}) = r).\end{aligned} \tag{21}$$

Given Algorithm 1 and threshold ϵ_{th}, the overall probability of correct rank detection can be lower bounded by $\phi(\epsilon_{\text{th}})$. The derivations are as follows.

$$\begin{aligned}&\mathbb{P}(\text{correct detection}) \\ &= \sum_{r=1}^{K} \mathbb{P}(\text{correct detection}|\text{rank}(\mathbf{H}) = r)\mathbb{P}(\text{rank}(\mathbf{H}) = r) \\ &\geq \sum_{r=1}^{K} \phi_r(\epsilon_{\text{th}})\mathbb{P}(\text{rank}(\mathbf{H}) = r) \tag{22} \\ &= C_2 \det\left(\mathbf{F}^{(1)}\left(\frac{PT}{M}\epsilon_{\text{th}}^2\right)\right) \cdot \sum_{r=1}^{K} C_1 \det\left(\mathbf{F}^{(r)}\left(4\epsilon_{\text{th}}^2\right)\right) \\ &\quad \times \mathbb{P}(\text{rank}(\mathbf{H}) = r), \tag{23}\end{aligned}$$

where in (22) and (23), the results in Proposition 1 have been used.

We thus choose the threshold ϵ_{th} so that the lower bound is maximized, i.e.,

$$\epsilon_{\text{th}}^* = \arg\max_{\epsilon_{\text{th}}} \phi(\epsilon_{\text{th}}). \tag{24}$$

The optimization problem in (24) is one-dimensional and can be optimally solved via exhaustive grid search. But there is a natural tradeoff between precision and computational complexity. For low computational complexity, in solving (24), we can find a zero point of $d\ln\phi(\epsilon_{\text{th}})/d\epsilon_{\text{th}}$ via bisection method and use it as the threshold. This low-complexity method can result in sub-optimality when $d\ln\phi(\epsilon_{\text{th}})/d\epsilon_{\text{th}}$ has multiple zero points.

In this section, for the traditional single-threshold-based rank detection, we rigorously derived an analytical lower bound on the correct rank detection probability, based on which a systematic threshold optimization scheme that maximizes this lower bound is proposed. The derived optimal threshold is adaptive to the number of transmit and receive antennas of the MIMO channel, the training length and power, and the distributions of the channel coefficients and the noise. However, it is independent of the instantaneous channel values or singular values of the channel matrix. Thus the threshold optimization can be conducted off-line, which largely reduces the delay of channel rank detection in real applications.

It is noteworthy that the proposed method is not limited to Rayleigh fading channel model but can be extended to other distributions. Our lower bound calculation and threshold optimization need the distributions of the singular values of the channel matrix. For other channel fading models, even if no closed-form expressions for the distributions of the singular values are available, numerical estimations of the singular value distributions can be obtained via simulation, and our method can still be used. Especially as our methods can be conducted off-line, the computation complexity is not an issue in the real-time channel rank detection and channel estimation within a coherence interval of the channel.

Notice that in the proof of Proposition 1, we can loose our condition for the lower bound $\lambda_r \geq 2\epsilon_{\text{th}} \geq 2\gamma_1$ to $\lambda_r \geq$

$2\epsilon_{\text{th}} \geq 2\gamma_r$ without affecting the validity of the proof. With this change, another lower bounds on the conditional and overall probabilities of correct rank detection, denoted as $\tilde{\phi}_r(\epsilon_{\text{th}})$ and $\tilde{\phi}(\epsilon_{\text{th}})$, can be obtained as follows:

$$\tilde{\phi}_r(\epsilon_{\text{th}}) \triangleq \mathbb{P}\left(\lambda_r \geq 2\epsilon_{\text{th}}|\text{rank}(\mathbf{H}) = r\right)\mathbb{P}(\gamma_r \leq \epsilon_{\text{th}}),$$

$$\tilde{\phi}(\epsilon_{\text{th}}) \triangleq \sum_{r=1}^{K} \tilde{\phi}_r(\epsilon_{\text{th}})\mathbb{P}(\text{rank}(\mathbf{H}) = r).$$

Since $\gamma_r < \gamma_1$, it can be shown straightforwardly that $\tilde{\phi}_r(\epsilon_{\text{th}}) > \phi_r(\epsilon_{\text{th}})$ and $\tilde{\phi}(\epsilon_{\text{th}}) > \phi(\epsilon_{\text{th}})$, which means that the new lower bounds are tighter, and we may obtain a better threshold by maximizing $\tilde{\phi}(\epsilon_{\text{th}})$. However, the expression for $\mathbb{P}(\gamma_r \leq \epsilon_{\text{th}})$ is much more complex than $\mathbb{P}(\gamma_1 \leq \epsilon_{\text{th}})$. Notice that it is the rth largest singular value of $\tilde{\mathbf{W}}$, not the smallest singular value since $\tilde{\mathbf{W}}$ has full rank with probability 1. Thus the calculation of the derivative of $\tilde{\phi}(\epsilon_{\text{th}})$ is more involved, and the maximization of $\tilde{\phi}(\epsilon_{\text{th}})$ has higher computational complexity. Our simulations show that the use of $\tilde{\phi}(\epsilon_{\text{th}})$ improves the performance but the improvement is moderate. Thus, balancing the performance and computation complexity, we choose $\phi(\epsilon_{\text{th}})$ for our algorithms.

3.3 Difference to existing single-threshold-based rank detection schemes

In this subsection, we explain the difference of our work with existing results on threshold-based rank detection in [19, 20, 22, 23].

First, the research in [19, 20, 22, 23] are for real-valued channel matrix and real-valued noise matrix. Also, they consider the special case of $T = M$ and $\mathbf{S} = \mathbf{I}_m$. Our work applies for complex-valued channel matrix and complex-valued noise matrix, a general training length T where $T \geq M$, and unitary $T \times M$ pilot $\mathbf{S}$. In what follows, we explain existing threshold selections and calibrate the results to our model and notation.

In [23], lower and upper bounds on the threshold selection were provided. No specific threshold value or optimization method was given.

In [19] and [20], the channel rank detection and channel estimation problem were considered jointly for the asymptotic case, where the channel matrix dimensions approach to infinity but with a fixed ratio. The threshold for the rank detection was chosen to minimize the MSE of the truncated SVD channel estimation. The rank detection follows the traditional single-threshold algorithm in Algorithm 1. After calibrating their results to our model, the threshold proposed in [20] is

$$\epsilon_{\text{th},[20]} = \sqrt{\frac{MN}{PT}}\sqrt{2\left(\frac{K}{L} + 1\right) + \frac{8K}{K + L + \sqrt{K^2 + 14KL + L^2}}}$$

and the threshold proposed in [19] is

$$\epsilon_{\text{th},[19]} = \left(1 + \sqrt{\frac{K}{L}}\right)\sqrt{\frac{MN}{PT}}.$$

In [22], the single-threshold Algorithm 1 was used and the threshold selection was based on the minimization of the SURE of the MSE of the channel estimation. Two methods were proposed. The first method needs numerical threshold optimization, and the optimization problem changes with the instantaneous channel realization. Thus, the threshold optimization needs to be repeated for every coherence interval, impairing its practicality and efficiency. As in the second method, an analytical threshold was proposed. After calibrated to our model, the threshold is

$$\epsilon_{\text{th},[22]} = \mathrm{F}_{\tilde{\mathbf{W}}_1}^{-1}\left(1 - \frac{1}{\sqrt{\log L}}\right),$$

where $\mathrm{F}_{\tilde{\mathbf{W}}_1}^{-1}(x)$ represents the inverse function of the CDF of the largest singular value of the noise matrix $\tilde{\mathbf{W}}$ in (7).

In all three aforementioned works, the proposed threshold selection methods depend on the distribution of the noise matrix only, where [19, 20] used the asymptotic behavior of the singular values and singular vectors of the noisy observation matrix when the dimensions of the channel matrix approach infinity; and Josse and Sardy [22] used the distribution of the largest singular value of the noise matrix. It was assumed in their work that the MIMO channel matrix is deterministic and their results cannot take advantage of the distribution of the channel matrix. On the contrary, we adopt a random channel model, and our threshold selection takes into account both the distribution of the channel coefficients and the distribution of the noises.

4 Improved multiple-threshold rank detection methods

To use the rank detection scheme in Algorithm 1 with the proposed threshold in the previous section, the a priori probabilities of the channel rank need to be known. However, for some wireless communication systems, the channel rank distribution may not be precisely known due to the mobility and complexity of the signal propagation environment. In this case, a rank detection algorithm that does not rely on the channel rank distribution is required. In addition, the lower bound on the probability of correct rank detection in (21) provides the average rank detection performance over all possible rank values. It may not be sharp enough for one channel realization with a specific rank value. Thus, in this section, we propose two improved rank detection algorithms which do not need the rank distribution.

4.1 Rank detection algorithm with multiple thresholds

Instead of using only a single threshold for the rank detection as in Algorithm 1, we propose to use K thresholds $\epsilon_{\text{th},1}^*, \epsilon_{\text{th},2}^*, \cdots, \epsilon_{\text{th},K}^*$, each corresponding to one of the K possible rank values, $1, \cdots, K$. These thresholds are optimized by maximizing the lower bound on the probability of correct rank detection conditioned on the channel rank value, i.e.,

$$\epsilon_{\text{th},i}^* = \arg \max_{\epsilon} \phi_i(\epsilon), \tag{25}$$

where $\phi_i(\epsilon)$ is defined in (13).

$\epsilon_{\text{th},i}^*$ serves as the rank detection threshold when the channel rank is i, and $\epsilon_{\text{th},i+1}^*$ serves as the rank detection threshold when the channel rank is $i + 1$. Recall that $\sigma_1, \cdots, \sigma_K$ are ordered singular values of $\tilde{\mathbf{Y}}$. Thus, it is natural to detect the channel rank as i when both the following two conditions are satisfied C1) $\sigma_i \geq \epsilon_{\text{th},i}^*$ and C2) $\sigma_{i+1} < \epsilon_{\text{th},i+1}^*$. To help the presentation, we define the following set:

$$\mathcal{I} = \{i \,|\, \sigma_i \geq \epsilon_{\text{th},i}^* \,\&\, \sigma_{i+1} < \epsilon_{\text{th},i+1}^*, i = 1, \cdots, K-1\},$$

which is the set of rank detection values that satisfy the two conditions. Since it is possible that $\mathcal{I}$ has 2 or more elements, for the uniqueness of the detection result, we detect the rank as the largest index that satisfies the two conditions. In other words, the detection rule, called RD2, can be represented as follows:

$$\text{RD2:} \quad r = \max_{i \in \mathcal{I}} \{i\}. \tag{26}$$

To guarantee the existence of a rank detection result, we also need to consider the case that there is no element in $\mathcal{I}$, i.e., $\mathcal{I} = \emptyset$. The following claim is proved.

Claim 1. *For two sequences of real numbers $a_1, \cdots, a_K$ and $b_1, \cdots, b_K$. If there exists no integer i (for $1 \leq i \leq K - 1$) such that $a_i \geq b_i$ and $a_{i+1} < b_{i+1}$, one of the following two cases must be true:*

1. *there exists an integer D, $1 \leq D \leq K$, such that $a_i < b_i$ for $i < D$, and $a_i \geq b_i$ for $i \geq D$*
2. *$a_i < b_i$ for all i.*

Proof 2. *Assume that there exists no integer i (for $1 \leq i \leq K - 1$) such that $a_i \geq b_i$ and $a_{i+1} < b_{i+1}$. We consider the two cases $a_K < b_K$ and $a_K \geq b_K$ separately.*

When $a_K < b_K$, if there exists an i such that $a_i \geq b_i$, let $i_{\max}$ be the largest i satisfying $a_i \geq b_i$. We have $a_{i_{\max}} \geq b_{i_{\max}}$ and $a_{i_{\max}+1} < b_{i_{\max}+1}$, which contradicts the assumption. Thus $a_i < b_i$ for all i.

When $a_K \geq b_K$, if there exists an i such that $a_i < b_i$, let $i_{\max}$ be the largest i satisfying $a_i < b_i$. Then, we will have $a_i < b_i$ for all $i \leq i_{\max}$, based on the same reasoning as

above. In this case, $D = i_{\max} + 1$. When there does not exist an i such that $a_i < b_i$, meaning $a_i \geq b_i$ for all i, $D = 1$.

Based on Claim 1, when $\mathcal{I} = \emptyset$, we have either $\sigma_i < \epsilon_{\text{th},i}^*$ for $i < D$ and $\sigma_i \geq \epsilon_{\text{th},i}^*$ for $i \geq D$ ($1 \leq D \leq K$), in which case the rank detection result should be K; or $\sigma_i < \epsilon_{\text{th},i}^*$ for all $i = 1, \cdots, K$, in which case the rank detection result should be 1, which is the lowest possible rank for the random Rayleigh fading channel matrix $\mathbf{H}$.

Given these discussions, our second rank detection scheme with multiple thresholds is described in Algorithm 2.

Algorithm 2 Rank detection algorithm with multiple thresholds

1: **while** ($i \leq K - 1$) **do**
2:     **if** $\sigma_i \geq \epsilon_{\text{th},i}^*$ and $\sigma_{i+1} < \epsilon_{\text{th},i+1}^*$ **then**
3:         The rank of $\mathbf{H}$ is detected as i; break;
4:     $i = i + 1$;
5: **if** ($i == K$) **then**
6:     **if** $\sigma_K \geq \epsilon_{\text{th},K}^*$ **then**
7:         The rank of $\mathbf{H}$ is detected as K.
8:     **if** $\sigma_K < \epsilon_{\text{th},K}^*$ **then**
9:         The rank of $\mathbf{H}$ is detected as 1.

4.2 Iterative rank detection algorithm with multiple thresholds

In our third rank detection algorithm, we iteratively use the K thresholds $\epsilon_{\text{th},1}^*, \cdots, \epsilon_{\text{th},K}^*$ defined in (25) to refine our rank detection threshold and detection result. In each iteration, single threshold-based rank detection is performed and the threshold value is set using the rank detection result of the previous iteration. The iteration ends when the rank detection result of the current iteration is the same as the result of the previous one. More specifically, first, a rank detection result r_1 is initialized, e.g., $r_1 = 1$, then $\epsilon_{\text{th},r_1}^*$ is used for the threshold of the next iteration. The new rank detection result is found using Algorithm 1, that is, the rank is detected as the maximum index i such that $\sigma_i \geq \epsilon_{\text{th},r_1}^*$ or 1 if all singular values of $\tilde{\mathbf{Y}}$ are smaller than $\epsilon_{\text{th},r_1}^*$. This new rank detection result is denoted as r_2. If $r_2 \neq r_1$, $\epsilon_{\text{th},r_2}^*$ is used as the threshold for the next iteration, and a new rank detection result can be obtained. The scheme is described in Algorithm 3. Please note that lines 3–8 of Algorithm 3 are the same as those of Algorithm 1.

Claim 2. *If $\epsilon_{\text{th},1}^* \geq \epsilon_{\text{th},2}^* \geq \cdots \geq \epsilon_{\text{th},K}^*$, Algorithm 3 always converges.*

Proof 3. *We prove the convergence by contradiction. Assume that the algorithm does not converge. From the algorithm, r_2 is the new rank detection result when using*

threshold $\epsilon^*_{\text{th},r_1}$. For any initial value for r_1, if $r_2 = r_1$, Algorithm 3 converges and the rank detection result is r_1, which causes a contradiction. Next we consider the cases $r_2 > r_1$ and $r_2 < r_1$ separately.

Case 1: $r_1 < r_2$. Recall that r_2 is the new rank detection result when using threshold $\epsilon^*_{\text{th},r_1}$. Thus, we have either

- *Case A: $r_2 = K$ and $\sigma_K \geq \epsilon^*_{\text{th},r_1}$; or*

- *Case B: $\sigma_{r_2} \geq \epsilon^*_{\text{th},r_1}$ and $\sigma_{r_2+1} < \epsilon^*_{\text{th},r_1}$.*

Since $\epsilon^*_{\text{th},i}$'s are in non-increasing order, for Case A, we have $\sigma_K \geq \epsilon^*_{\text{th},r_1} \geq \epsilon^*_{\text{th},K}$. Thus the new rank detection result is K, i.e., $r_3 = K$. Then $r_3 = r_2$ and Algorithm 3 terminates, which contradicts the assumption. For Case B, since $r_1 < r_2$, we have $\sigma_{r_2} \geq \epsilon^*_{\text{th},r_1} \geq \epsilon^*_{\text{th},r_2}$. Thus the next rank detection result cannot be smaller than r_2, i.e., $r_2 \leq r_3$. If $r_2 = r_3$, Algorithm 3 terminates, which contradicts the assumption. Thus $r_2 < r_3$. The same situation happens for the next iterations. So, if Algorithm 3 does not converge, we will find an infinite strictly increasing integer sequence $r_1 < r_2 < r_3 < \cdots$. This contradicts the fact that r_m's are in the range of $[1, K]$.

Case 2: $r_1 > r_2$. Similarly, we have either Case A ($r_2 = 1$ and $\sigma_1 < \epsilon^*_{\text{th},r_1}$) or Case B listed above. For Case A, we have $\sigma_1 < \epsilon^*_{\text{th},r_1} \leq \epsilon^*_{\text{th},1}$. Thus the new rank detection result is 1, i.e., $r_3 = 1$ and Algorithm 3 terminates, which contradicts the assumption. For Case B, since $r_1 > r_2$, we have $\sigma_{r_2+1} < \epsilon^*_{\text{th},r_1} \leq \epsilon^*_{\text{th},r_2}$. Thus the next rank detection result cannot be larger than r_2, i.e., $r_3 \leq r_2$. If $r_2 = r_3$, Algorithm 3 terminates, which contradicts the assumption. Thus $r_2 < r_3$. The same situation happens for the next iterations. So, if Algorithm 3 does not converge, we will find an infinite strictly decreasing integer sequence $r_1 > r_2 > r_3 > \cdots$. This contradicts the fact that r_m's are in the range of $[1, K]$.

Claim 2 shows that when the thresholds corresponding to the K rank values are in non-increasing order,

---

Algorithm 3 Iterative rank detection algorithm with multiple thresholds

---

1: $r_0 = 0$, $i = 1$. Choose an integer value for r_1 between $[1, K]$, e.g., $r_1 = 1$.
2: **while** $r_i! = r_{i-1}$ **do**
3:    **for** $m = K : 1$ **do**
4:       **if** $\sigma_m \geq \epsilon^*_{\text{th},r_i}$ **then**
5:          $r_{i+1} = m$; break;
6:       **if** $\epsilon^*_{\text{th},r_i} > \sigma_1$ **then**
7:          $r_{i+1} = 1$.
8:    $i = i + 1$;
9: The rank of **H** is detected as r_i;

Algorithm 3 is guaranteed to converge. Also, from the proof of Claim 2, we can see that the convergence is guaranteed within K iterations. Our limited simulation results indicate that the algorithm converges very fast (within 2–3 iterations). Intuitively, as $\epsilon^*_{\text{th},i}$ is the threshold when the channel rank is i and the singular values are in a non-increasing order, it is natural to have $\epsilon^*_{\text{th},1}, \epsilon^*_{\text{th},2}, \cdots, \epsilon^*_{\text{th},K}$ in a non-increasing order. However, we cannot prove this analytically. When violation of the ordering happens on one threshold, we can simply reset the threshold to be the average of the one before and the one after to fix the ordering problem and have guaranteed convergence.[1]

Although the convergence of Algorithm 3 is guaranteed, we cannot not guarantee the uniqueness of the rank detection solution with respect to different initial values for r_1. In other words, for different initial rank values, the algorithm may converge to different solutions. An example is as follows. Assume that $\sigma_1 > \epsilon^*_{\text{th},1} > \sigma_2 > \epsilon^*_{\text{th},2} > \cdots > \sigma_K > \epsilon^*_{\text{th},K}$. Then the final rank detection result of Algorithm 3 will equal to r_1 for any initial r_1 value.

4.3 Discussion on complexity

The computational complexity of Algorithms 2 and 3 is composed of two parts: the calculations of $\epsilon^*_{\text{th},1}, \cdots, \epsilon^*_{\text{th},K}$ and the rank detection part.

The optimization of $\epsilon^*_{\text{th},i}$ only depends on the dimensions of the channel matrix M and N, the training time T, and the training power P. It is independent of the channel realization of each coherence interval. Thus the optimization can be conducted off-line. Further, the following lemma is proved which can be used to reduce the computational complexity of the optimization.

Lemma 1. *The function $\phi_r(\epsilon_{\text{th}})$ in (13) is a log-concave function of ϵ_{th}.*

Proof 4. *For the Hermitian matrices $\boldsymbol{F}^{(1)}(\mu)$ and $\boldsymbol{F}^{(r)}(\mu)$, we can show that all their leading principal minors are positive when $\mu > 0$ from the definitions in (11), (12), and the CDFs in (17), (19). Thus the two matrices are positive definite and $\det\left(\boldsymbol{F}^{(1)}\left(\frac{PT}{M}\epsilon_{\text{th}}^2\right)\right)$ and $\det\left(\boldsymbol{F}^{(r)}\left(4\epsilon_{\text{th}}^2\right)\right)$ are log-concave functions since the determinant of a positive definite matrix is log-concave [31]. Based on [31], the product of log-concave functions is also log-concave. This ends the proof.*

Notice that $\epsilon^*_{\text{th},r}$ is the maximum point of $\ln \phi_r(\epsilon_{\text{th}})$. With the log-concavity of $\phi_r(\epsilon_{\text{th}})$, we can find $\epsilon^*_{\text{th},r}$ by finding the unique zero point of $d \ln \phi_r(\epsilon_{\text{th}})/d\epsilon_{\text{th}}$, using bisection method. The calculations are as follows. From (13) and the definitions in (11) and (12), we have

$$\ln \phi_r(\epsilon_{\text{th}}) = \ln C_1 + \ln C_2 + \ln \det\left(\boldsymbol{F}^{(r)}\left(4\epsilon_{\text{th}}^2\right)\right)$$
$$+ \ln \det\left(\boldsymbol{F}^{(1)}\left(\frac{PT}{M}\epsilon_{\text{th}}^2\right)\right),$$

and

$$\frac{d(\ln \phi_r(\epsilon_{\text{th}}))}{d\epsilon_{\text{th}}} = \text{tr}\left[\left(\boldsymbol{F}^{(r)}\left(4\epsilon_{\text{th}}^2\right)\right)^{-1}\boldsymbol{D}_r\right] \qquad (27)$$
$$+ \text{tr}\left[\left(\boldsymbol{F}^{(1)}\left(\frac{PT}{M}\epsilon_{\text{th}}^2\right)\right)^{-1}\boldsymbol{D}_1\right],$$

where $\boldsymbol{D}_r \triangleq d\left(\boldsymbol{F}^{(r)}\left(4\epsilon_{\text{th}}^2\right)\right)/d\epsilon_{\text{th}}$ and $\boldsymbol{D}_1 \triangleq d\left(\boldsymbol{F}^{(1)}\left(\frac{PT}{M}\epsilon_{\text{th}}^2\right)\right)/d\epsilon_{\text{th}}$. The (i,j)-th entries of $\boldsymbol{D}_r$ and $\boldsymbol{D}_1$ are respectively

$$[\boldsymbol{D}_r]_{i,j} = -8\epsilon_{\text{th}}e^{-4\epsilon_{\text{th}}^2}\left(4\epsilon_{\text{th}}^2\right)^{M-r+i+j-2}, \qquad (28)$$
$$[\boldsymbol{D}_1]_{i,j} = 2\frac{PT}{M}\epsilon_{\text{th}}e^{-\frac{PT}{M}\epsilon_{\text{th}}^2}\left(\frac{PT}{M}\epsilon_{\text{th}}^2\right)^{i+j-2}.$$

Next, we analyze the complexity of the rank detection part. For Algorithm 2, the total number of comparisons in the worst scenario is $2K$, thus the complexity is $\mathcal{O}(K)$. For Algorithm 3, in the worst case, the number of iterations is K; and for each iteration, at most $K + 1$ comparisons are needed. The overall number of comparisons is $K(K + 1)$. Thus the complexity is $\mathcal{O}(K^2)$. Notice that $K = \min\{M, N\}$. Even for massive MIMO systems with N base station antennas and M single-antenna users, where N is large (e.g., hundreds), the complexities of the two proposed rank detection algorithms are linear and quadratic in the number of users, respectively.

5 Simulation results

5.1 Simulation on the probability of correct rank detection

In this section, simulation results are shown for Algorithm 1 with our proposed threshold optimization in Section 3, and the two new rank detection algorithms with multiple thresholds, Algorithm 2 and Algorithm 3, proposed in Section 4. We simulate the probability of correct rank detection for different parameters, such as the average training power P, the training length T, and the numbers of transmitter and receiver antennas M and N. In our simulations, channel coefficients are generated as Rayleigh fading following the model in (3). $\boldsymbol{S}$ is generated as a random $T \times M$ unitary matrix following the isotropic distribution. For comparison, we also show the rank detection accuracy of Algorithm 1 with the threshold values proposed in [19, 20, 22]. While, the results in [19, 20, 22] are for $T = M$ and $\boldsymbol{S} = \boldsymbol{I}_M$, we extend them for a general $T \geq M$ and unitary $\boldsymbol{S}$ as explained in Section 3.3.

We first consider an 8×8 MIMO system, where $M = N = 8$. This configuration is typical in 4G standards (e.g., LTE-advance, WiMAX Release 2) [32]. Figure 1 shows the probabilities of correct rank detection for different

average training powers, where $T = M = 8$. The channel rank is randomly generated with uniform distribution, i.e., $\mathbb{P}(\text{rank}(\boldsymbol{H}) = i) = 1/8$ for $i = 1, \cdots, 8$. We can observe from this figure that for the single-threshold algorithm, Algorithm 1, the proposed threshold has about the same performance as the threshold in [22] for the whole training power range and are much better than the ones in [19, 20]. The proposed multiple-threshold algorithms, Algorithm 2 and Algorithm 3, achieve considerably higher detection rate than Algorithm 1 when the training power is higher than 6.5 dB. Algorithm 2 is slightly better than Algorithm 3 at high training power, but is slightly worse at low training power.

To better understand the performance of the schemes for different channel ranks, for the same network, in Figs. 2 and 3, we show the probabilities of correct rank detection for low-rank channel matrix where the rank can be 1, 2, or 3 with equal probability and for high-rank channel matrix where the rank can be 4, 5, or 6 with equal probability.[2] For the low-rank case, Algorithm 1 with the proposed threshold outperforms the proposed Algorithms 2 and 3. Compared with existing work, the proposed threshold has better performance than [22] at high training power but worse performance at low training power. It is significantly better than results in [19, 20] for most training power values. For the high-rank case, among the three proposed schemes, Algorithm 2 has the best performance, followed by Algorithm 3. They achieve significantly higher detection rate than schemes in [19, 20], especially for low training power. Compared with [22], the proposed schemes are largely better for small and medium P and slightly worse for large P.

In Fig. 4, for the same 8×8 MIMO system with uniform rank value from 1 to 8, we simulate the probability of correct rank detection for different training lengths. The figure shows that the proposed Algorithms 2 and 3 are significantly better than Algorithm 1. For Algorithm 1, the proposed threshold is slightly worse than the one in [22], but better than those in [19, 20]. The figure also shows that the proposed Algorithm 2 achieves the highest correct detection probability.

In Figs. 5 and 6, we show the rank detection performance for different numbers of transmit antennas while the number of receive antennas is fixed as 8 and different numbers of receive antennas while the number of transmit antennas is fixed as 8. The channel rank is random and uniformly distributed from 1 to K. The training length is $T = 12$. For Fig. 5, the average transmit power of each transmit antenna is set to be 1 dB, thus, $P/M = 1$ dB, or equivalently, $P = 10^{0.1}M = (1 + 10\log_{10} M)$ dB. This is for fair comparison among systems with different M values. For Fig. 6, we set $P = 10$ dB. From the figures, we can see that our proposed Algorithms 2 and 3 always achieve significantly higher detection rate than existing
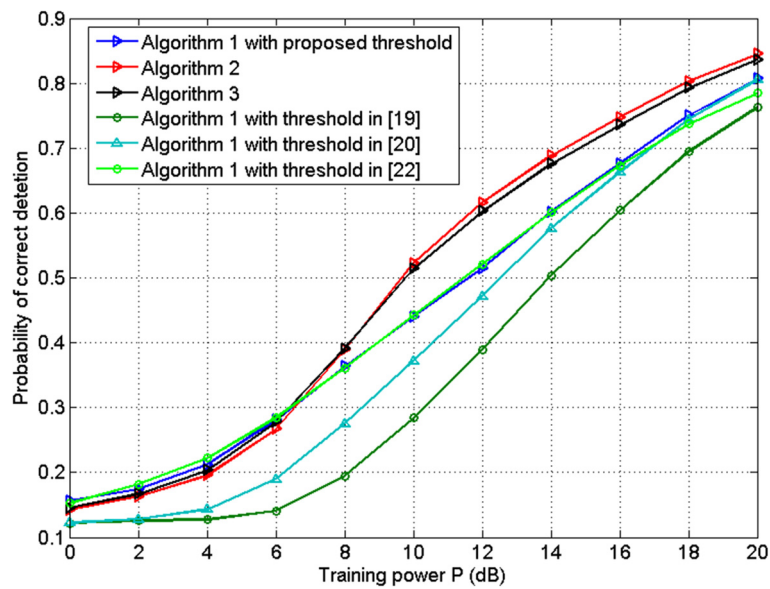
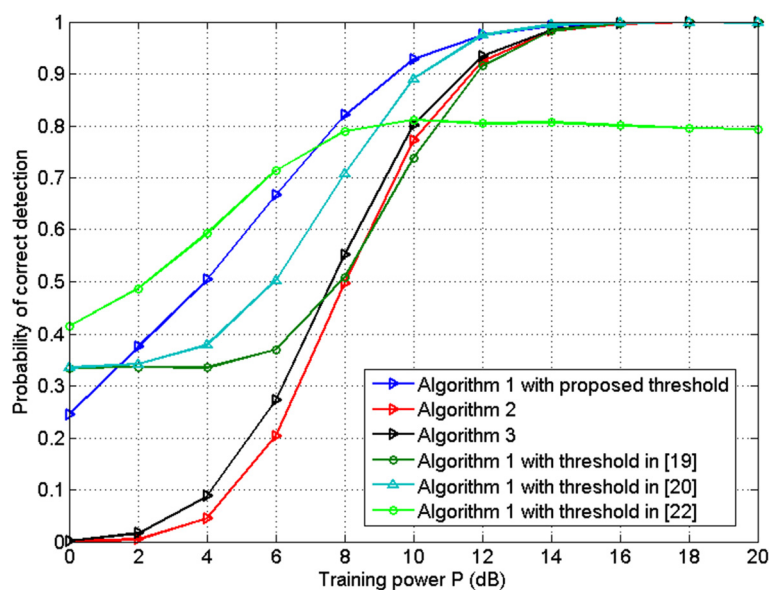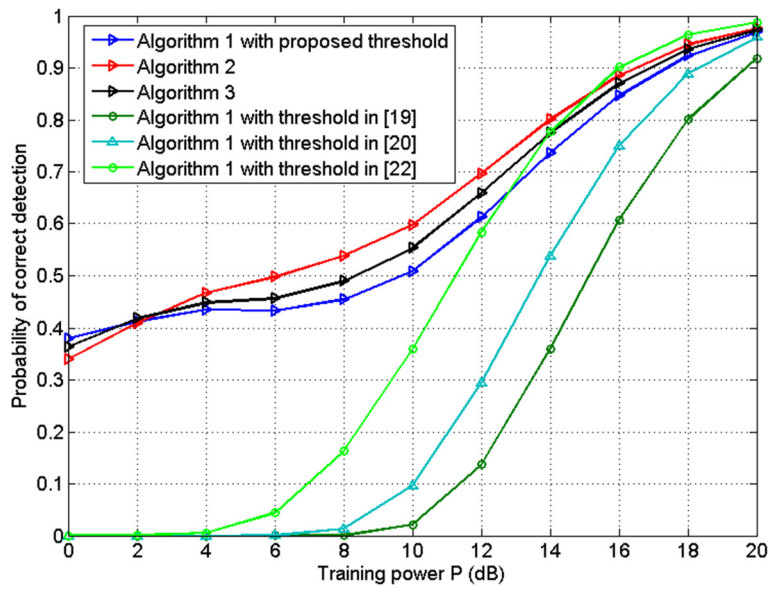**Fig. 1** Probability of correct rank detection of 8 × 8 MIMO systems for different average training power. The rank of the channel matrix is uniformly distributed between 1 and 8

ones. Algorithm 1 with the proposed threshold is also superior to existing ones in general except for large M values in Fig. 5, where it is slightly worse than [22].

In Fig. 7, we consider a 10×100 uplink massive MIMO system. The receiver is a base station where 100 antennas are deployed. The transmitters are 10 single-antenna users or one user with 10 antennas. The training length is $T = 20$. The channel rank is assumed to be uniformly distributed over $\{5,6,7,8,9\}^3$. The probabilities of correct rank detection for different average training powers are shown. We can see that the proposed schemes outperform all existing ones for the whole range of the training power.

Fig. 2 Probability of correct rank detection of 8 × 8 MIMO systems for different average training power. The rank of the channel matrix is uniform between 1 and 3

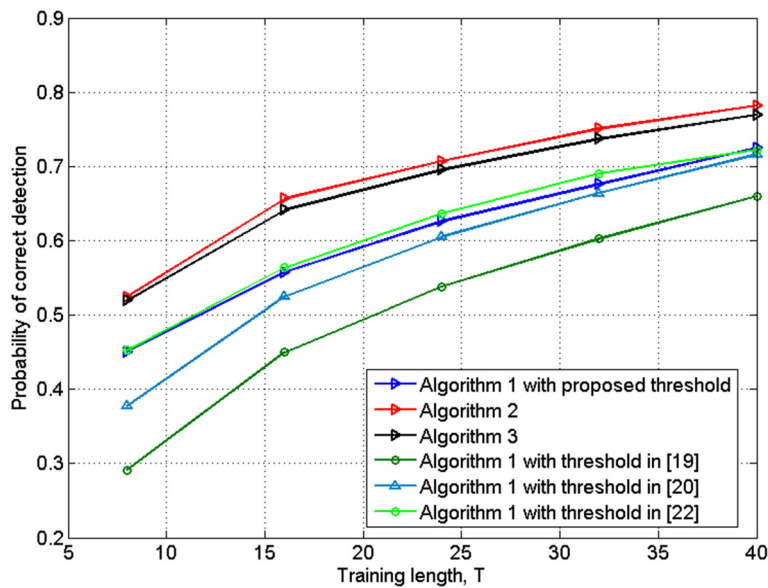**Fig. 3** Probability of correct rank detection of 8 × 8 MIMO systems for different average training power. The rank of the channel matrix is uniform between 4 and 6

The figure also shows that Algorithm 3 is slightly better than Algorithm 2 at low and medium training power, but slightly worse at high training power.

5.2 Simulation on MSE and outage probability

In this section, to show the effect of rank detection on the channel estimation and the MIMO communication performance via simulations on the MSE of the truncated SVD channel estimation and the outage probability of the MIMO multiplexing communications. The truncated SVD channel estimation given the rank detection result has been shown in (8) in Section 2.2. The MSE is defined as

$$\mathrm{MSE}(\hat{\mathbf{H}}) \triangleq \frac{1}{MN}\mathbb{E}\{\|\hat{\mathbf{H}} - \mathbf{H}\|_F^2\}, \tag{29}$$

Fig. 4 Probability of correct rank detection of 8 × 8 MIMO systems for different training lengths *T*. The rank of the channel matrix is uniformly distributed between 1 and 8
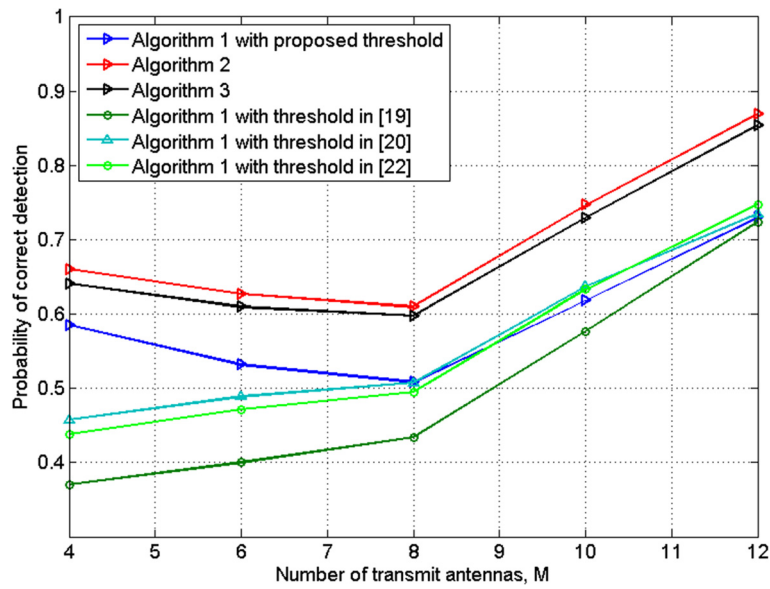
Fig. 5 Probability of correct rank detection of 8×8 MIMO systems for different M and $N = 8$. $P/M = 1$ dB. The channel rank is uniform from 1 to K

which is the MSE per channel coefficient. Traditionally, the average in (29) is over the noises. But here it is over both the noises and the channels, since the channel coefficients are also random variables. In the conducted Monte-Carlo simulation, a distinct channel realization is used for each iteration.

In Fig. 8, we show the MSE for an 8×8 MIMO system with different rank detection methods. The rank of the channel matrix is uniformly distributed between 1 and 8. The training length is $T = 8$. The simulated training power range is from 8 to 18 dB, which is the medium and high power range. Previous simulations show that our algorithms achieve higher rank detection probability in this training power range. From the figure, we can see that our algorithms achieve lower MSE than the existing algorithms. For comparison, we also simulated the MSE
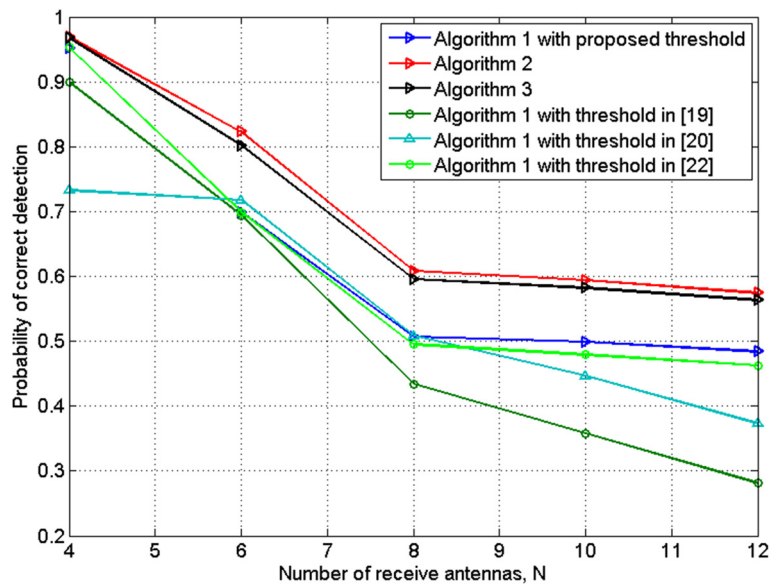

Fig. 6 Probability of correct rank detection of 8×8 MIMO systems for different N and $M = 8$. $P = 10$ dB. The channel rank is uniform from 1 to K
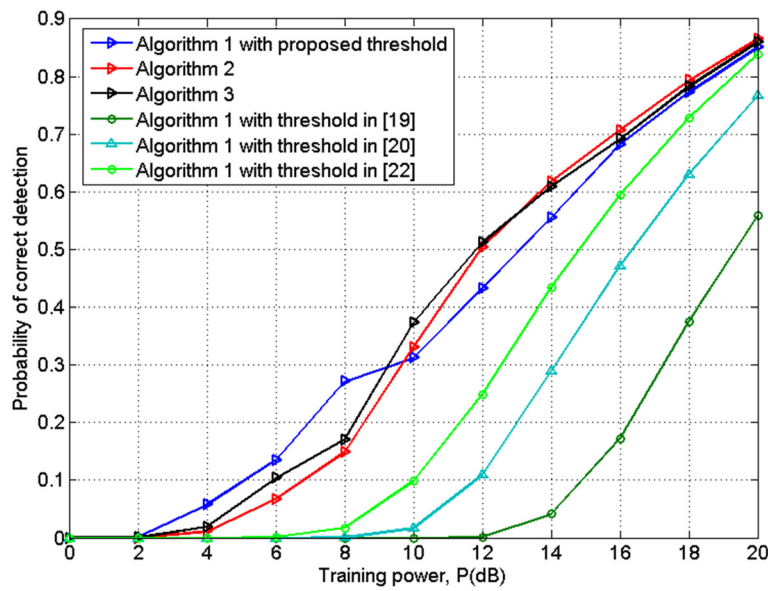
Fig. 7 Probability of correct rank detection of 10 × 100 massive MIMO system for different average training power, with high channel rank values

for SVD-based estimation with perfect rank detection. We can see that it has the lowest MSE, which helps verify that better rank detection will improve the SVD-based channel estimation. The figure also shows that the SVD-based estimation achieves lower MSE than entry-based estimation.

For the outage probability, MIMO multiplexing transmission is adopted, where multiple data streams are sent

through different channel eigenspaces. First we rewrite the channel estimate in (8) as follows

$$\hat{\mathbf{H}} = \hat{\mathbf{U}}_{\hat{r}}\mathrm{diag}\{\sigma_1,\ldots,\sigma_{\hat{r}}\}\hat{\mathbf{V}}_{\hat{r}}^*,$$

where $\hat{r}$ is the rank detection result, $\hat{\mathbf{U}}_{\hat{r}}$ is the $M \times \hat{r}$ matrix composed of the first $\hat{r}$ columns of $\mathbf{P}$, and $\hat{\mathbf{V}}_{\hat{r}}$ is the $M \times \hat{r}$ matrix composed of the first $\hat{r}$ columns of $\mathbf{Q}$. Since the rank detection result is $\hat{r}$, only $\hat{r}$ users
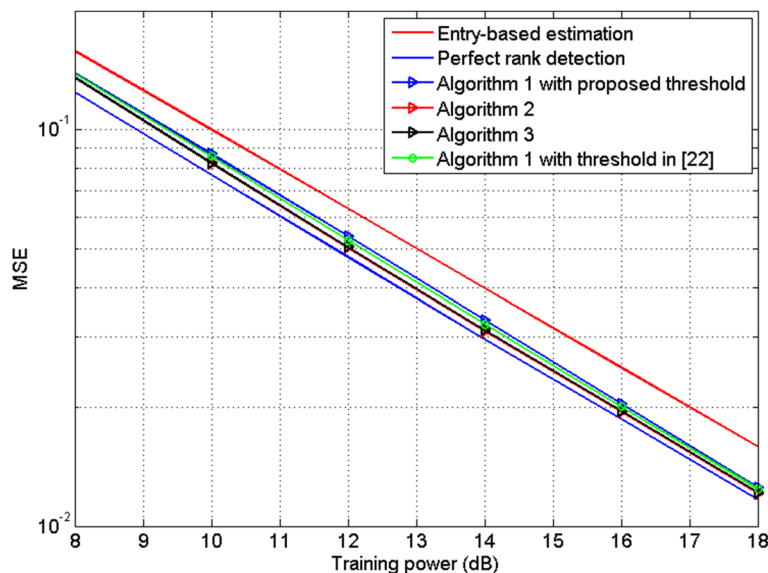

Fig. 8 MSE of 8 × 8 MIMO systems for different average training powers. The rank of the channel matrix is uniformly distributed between 1 and 8

will be served. Let **s** be the $\hat{r} \times 1$ vector of normalized information symbols, and P be the average total transmit power. By assuming equal power allocation, the transmitted signal vector is $\mathbf{x} = \sqrt{P/\hat{r}}\mathbf{s}\hat{\mathbf{U}}_{\hat{r}}^*$. At the receiver side, after right-multiplying the received signal vector **Y** by $\hat{\mathbf{V}}_{\hat{r}}$, the equivalent multi-stream transceiver equation can be written as

$$\tilde{\mathbf{y}} = \mathbf{y}\hat{\mathbf{V}}_{\hat{r}} = \sqrt{\frac{P}{\hat{r}}}\mathbf{s}\hat{\mathbf{U}}_{\hat{r}}^*\mathbf{H}\hat{\mathbf{V}}_{\hat{r}} + \tilde{\mathbf{w}}, \qquad (30)$$

where $\tilde{\mathbf{w}}$ is the equivalent noise vector. Define $\mathbf{Z} = \hat{\mathbf{U}}_{\hat{r}}^*\mathbf{H}\hat{\mathbf{V}}_{\hat{r}}$. The signal-to-noise-plus-interference ratio (SINR) of the ith data stream can be represented as

$$\text{SINR}_i = \frac{\frac{P}{\hat{r}}|z_{i,i}|^2}{\sigma_n^2 + \frac{P}{\hat{r}}\sum_{j=1, j\neq i}^{\hat{r}}|z_{j,i}|^2}, \qquad (31)$$

where σ_n^2 is the noise variance. In the following simulations, we assume that the noises have unit variance, i.e., $\sigma_n^2 = 1$.

For fair comparison of different rank detection schemes, the system outage probability is calculated as follows. We assume that there are K data streams/symbols in total. If the rank detection result is $\hat{r}$, only $\hat{r}$ symbols are transmitted and the communications of the rest symbols are seen to be in outage. For example, for a 8×8 channel, if the rank is detected as 5, the three symbols that are not served are in outage. This setup will make the outage probability value appear to be high especially when the channel rank is low. But it is fair since it avoids the preference of always having a lower rank detection.[4] For

the $\hat{r}$ symbols that are actually transmitted, a symbol is in outage if its SINR is below the pre-defined threshold. In Fig. 9, the outage probabilities for an 8×8 MIMO system are shown for different training powers. The rank of the channel matrix is uniformly distributed between 1 and 8. The training length is $T = 8$. The total transmit power is fixed and set as 8×10 dB. The SINR threshold is 3 dB. We compare our schemes with the cases of perfect channel estimation, entry-based estimation, and SVD-based estimation with perfect rank detection. The figure shows that perfect rank detection achieves the lowest outage probabilities at medium and high training power ranges, which verifies that better rank detection will improve the outage performance. It is also shown in the figure that the proposed multiple-threshold algorithms achieve almost the same outage probability as perfect rank detection and are better than other rank detection algorithms.

6 Conclusions

In this paper, we proposed novel threshold-based rank detection algorithms for reduced-rank MIMO systems. Different from previous work, we consider a MIMO system with a random channel matrix model, a general training length, and a unitary training matrix. Lower bounds on the probability of correct rank detection were derived using the distribution of the channel matrix and noise matrix, based on which the rank detection thresholds can be optimized. In addition to the traditional single-threshold detection algorithm, we further proposed two low-complexity multiple-threshold algorithms.
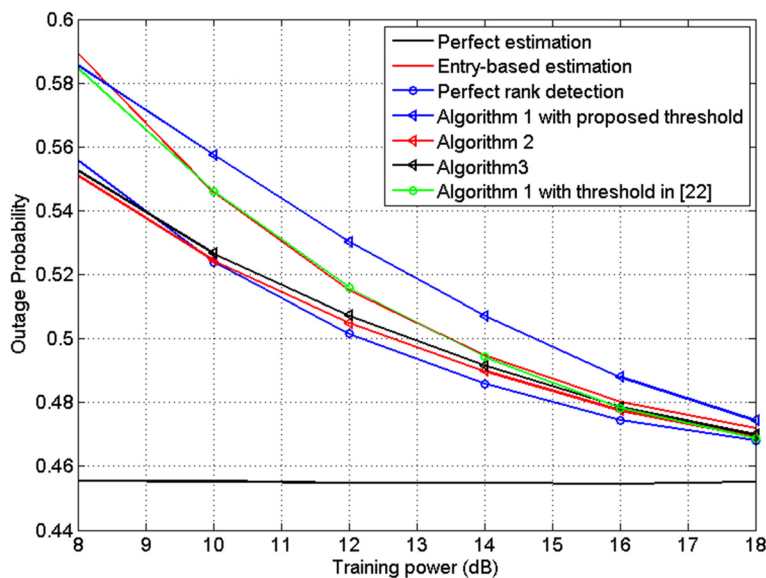

Fig. 9 Outage probability of 8×8 MIMO systems for different average training powers. The rank of the channel matrix is uniformly distributed between 1 and 8

Compared with the existing schemes, our proposed schemes can achieve higher rank detection rate for various scenarios. Our simulation results also show that better rank detection can improve the channel estimation quality and the outage performance of the MIMO system.

Endnotes
[1]In simulations, the thresholds we numerically obtain from (25) are in non-increasing order, except when the MIMO channel dimension gets large. When the values of M and N are large, violation of the non-increasing order occasionally happens due to the limited precision of computer calculation.

[2]The cases where the channel rank is 7 and 8 are not considered, since based on the the experimental results on several scenarios in [33], the rank of the 8×8 channel matrix is always less than 7.

[3]For very small rank values, e.g., 1,2,3,4, our schemes are inferior to existing ones.

[4]For example, for a 8×8 MIMO system whose channel rank is uniformly distributed between 1 and 8, under this setup, a lower bound on the outage probability will be $\sum_{i=0}^{7} i/8/8 = 43.75\,\%$. Another way to show the comparison is to draw the simulated outage probability values subtracted by the lower bound. But either way, the same comparison result can be obtained.

References
1. GJ Foschini, Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas. Bell. Labs Tech. J. **1**(2), 41–59 (1996)
2. IE Telatar, Capacity of multi-antenna gaussian channel. Eur. Trans. Telecom. **10**(6), 585–595 (1999)
3. AG Burr, Capacity bounds and estimates for the finite scatterers mimo wireless channel. IEEE J. Sel. Areas. Commun. **21**(5), 812–818 (2003)
4. HQ Ngo, EG Larsson, TL Thomas, The multi-cell multi-user mimo uplink with very large antenna arrays and a finite-dimensional channel. IEEE Trans. Commun. **61**(6), 2350–2361 (2013)
5. F Rusek, D Persson, BK Lau, EG Larsson, TL Marzetta, O Edfors, F Tufvesson, Scaling up mimo: Opportunities and challenges with very large arrays. IEEE Signal Process. Mag. **30**(1), 40–60 (2013)
6. JS Goldstein, IS Reed, LL Scharf, A multistage representation of the wiener filter based on orthogonal projections. IEEE Trans. Inf. Theory. **44**(7), 2943–2959 (1998)
7. ML Honig, JS Goldstein, Adaptive reduced-rank interference suppression based on the multistage wiener filter. IEEE Trans. Commun. **50**(6), 986–994 (2002)
8. RC de Lamare, M Haardt, R Sampaio-Neto, Blind adaptive constrained reduced-rank parameter estimation based on constant modulus design for CDMA interference suppression. IEEE Trans. Signal Process. **56**(6), 2470–2482 (2008)
9. Y Sun, V Tripathi, ML Honig, Adaptive turbo reduced-rank equalization for mimo channels. IEEE Trans. Wireless Commun. **4**(6), 2789–2800 (2005)
10. H Qian, SN Batalama, Data record-based criteria for the selection of an auxiliary vector estimator of the MMSE/MVDR filter. IEEE Trans. Commun. **51**(10), 1700–1708 (2003)
11. RC de Lamare, R Sampaio-Neto, Adaptive reduced-rank processing based on joint and iterative interpolation, decimation, and filtering. IEEE Trans. Signal Process. **57**(7), 2503–2514 (2009)
12. RC de Lamare, R Sampaio-Neto, Adaptive reduced-rank equalization algorithms based on alternating optimization design techniques for mimo systems. IEEE Trans. Veh. Tech. **60**(6), 2482–2494 (2011)
13. Y Hua, M Nikpour, P Stoica, Optimal reduced-rank estimation and filtering. IEEE Trans. Signal Process. **49**(3), 457–469 (2001)
14. P Stoica, M Viberg, Maximum likelihood parameter and rank estimation in reduced-rank multivariate linear regressions. IEEE Trans. Signal Process. **44**(12), 3069–3078 (1996)
15. Y Jing, X Yu, Ml-based channel estimations for non-regenerative relay networks with multiple transmit and receive antennas. IEEE J. Sel. Areas Commun. **30**(8), 1428–1439 (2012)
16. E Lindskog, C Tidestav, in *Proc. IEEE 49th Veh. Tech. Conf.: Jul. 1999*. Reduced rank channel estimation, vol. 2 (IEE, Houston, TX, 1999), pp. 1126–1130
17. M Nicoli, U Spagnolini, Reduced-rank channel estimation for time-slotted mobile communication systems. IEEE Trans. Signal Process. **53**(3), 926–944 (2005)
18. AA Shabalin, AB Nobel, Reconstruction of a low-rank matrix in the presence of gaussian noise. J. Multivariate Anal. **118**, 67–76 (2013)
19. M Gavish, DL Donoho, Optimal shrinkage of singular values. (arXiv preprint, 2015). http://arxiv.org/abs/1405.7511. Accessed 14 Oct. 2015
20. DL Donoho, M Gavish, The optimal hard threshold for singular values is 4/sqrt(3). IEEE Trans. Inf. Theory. **60**(8), 5040–5053 (2014)
21. R Nadakuditi, Optshrink: An algorithm for improved low-rank signal matrix denoising by optimal, data-driven singular value shrinkage. IEEE Trans. Inf. Theory. **60**(5), 3002–3018 (2014)
22. J Josse, S Sardy, Adaptive shrinkage of singular values. Statistics and Computing (2015). doi:10.1007/s1122201595549
23. K Konstantinides, K Yao, Statistical analysis of effective singular values in matrix rank determination. IEEE Trans. Acoust., Speech, Signal Process. **36**(5), 757–763 (1988)
24. B Hassibi, BM Hochwald, How much training is needed in multiple-antenna wireless links? IEEE Trans. Inf. Theory. **49**, 951–961 (2003)
25. N Shariati, J Wang, M Bengtsson, Robust training sequence design for correlated mimo channel estimation. IEEE Trans. Signal Process. **61**(1), 107–120 (2014)
26. C Huang, T Chang, X Zhou, Y-WP Hong, Two-way training for discriminatory channel estimation in wireless mimo systems. IEEE Trans. Signal Process. **61**(10), 2724–2738 (2013)
27. M Biguesh, AB Gershman, Training-based mimo channel estimation: A study of estimator tradeoffs and optimal training signals. IEEE Trans. Signal Process. **54**(3), 884–893 (2006)
28. C Eckart, G Young, The approximation of one matrix by another of lower rank. Psychometrika. **1**(3), 211–218 (1936)
29. M Nicoli, *Multi-User Reduced Rank Receivers for TD/CDMA Systems, Dissertation,* PhD thesis, (Politecnico di Milano, Milan, Italy, 2001)
30. A Ordonez, D Palomar, J Fonollosa, Ordered eigenvalues of a general class of Hermitian matrices with application to the performance analysis of mimo systems. IEEE Trans. Signal Process. **57**(2), 672–689 (2009)
31. S Boyd, L Vandenberghe, *Convex Optimization*. (Cambridge University Press, Cambridge, 2004), pp. 29–48
32. Q Li, G Li, W Lee, M Lee, D Mazzarese, B Clerckx, Z Li, Mimo techniques in wimax and LTE: a feature overview. IEEE Commun. Mag. **48**(5), 86–92 (2010)
33. A Molisch, H Asplund, R Heddergott, M Steinbauer, T Zwick, The cost259 directional channel model-part i: Overview and methodology. IEEE Trans. Wireless Commun. **5**(12), 3421–3433 (2006)