ICA Based Semi-Blind Decoding Method for a Multicell Multiuser Massive MIMO Uplink System in Rician/Rayleigh Fading Channels

Lei Shen, Yu-Dong Yao, Fellow, IEEE, Haiquan Wang, Member, IEEE, and Huaxia Wang, Student Member, IEEE

Abstract—A massive multiple-input multiple-output (MIMO) system, which utilizes a large number of antennas at base stations to communicate with multiple user terminals each with a single antenna, is one of the most promising techniques for future wireless communications systems. A successful massive MIMO implementation relies on accurate channel estimation, which is typically performed through pilot sequences. However, the channel estimation performance or massive MIMO performance is limited by the pilot contamination due to unavoidable reuse of pilot sequences from terminals in neighboring cells. In this paper, a semi-blind decoding method based on independent component analysis (ICA), channel energy levels, and reference bits is proposed. Specifically, the proposed decoding method uses ICA to separate and decode the received signals and to estimate the channels. The estimated channel energy is used to differentiate the in-cell signals and the neighboring cell signals, and reference bits are applied to identify a desired signal among signals within a cell. The analytical performance results of the proposed decoding method are derived. Numerical results show that the proposed semi-blind decoding method has better bit error rate (BER) performance and higher transmit efficiency than the traditional minimum mean-square error (MMSE) decoding method and the singular value decomposition (SVD) based decoding method.

Index Terms—Semi-blind decoding, massive multiple-input multiple-output (MIMO), independent component analysis (ICA), minimum mean-square error (MMSE), zero forcing (ZF).

I. INTRODUCTION

As one of the key technologies for future generation wireless communications systems, massive multiple-input multiple-output (MIMO) has emerged as a breakthrough technology for high data rate systems due to its high spectral efficiency [1]. In a massive MIMO system, a base station (BS) is equipped with a large number of antennas, which are able to serve multiple single-antenna users using the same time-frequency resource. If accurate channel state information (CSI) is known for a receiver, the system can achieve satisfactory performance by utilizing the conventional linear decoding method such as minimum mean-square error (MMSE) decoding or zero forcing (ZF) decoding [2]. However, if perfect CSI is not available, there will be significant performance degradation in massive MIMO [3]. For an uplink system, CSI is usually obtained through pilot sequences. In order to accurately estimate CSI, the pilot sequences sent by different users should be designed as mutually orthogonal. In a multi-cell scenario, neighboring cells might be set up in a close range and pilot sequences can be reused among neighboring cells. There will be interference among neighboring cells for pilot sequences, which is known as pilot contamination. In fact, pilot contamination has become a bottleneck for linear channel estimation methods in a massive MIMO wireless system [4]-[6].

In order to reduce the effect of pilot contamination, many techniques about pilot designs and channel estimation are proposed based on conventional MMSE decoding or ZF decoding. Paper [7] and [8] propose a time-shifted pilot-based scheme to address pilot contamination by rearranging the pilot transmission order for different cells. Based on this method, CSI can be accurately obtained by avoiding simultaneous transmissions in adjacent cells. However, an accurate time slot control is needed for this method, especially for systems with a large number of adjacent cells and high data rate transmissions. In [9], base stations use a low-rate coordination and additional second-order statistics about the user terminal channels to mitigate pilot contamination. Unfortunately, without perfect covariance information estimation, the performance of the proposed coordination method degrades. For a massive antenna system, the channels from different users are orthogonal when the number of BS antennas goes to infinity. [10]-[12] propose a singular value decomposition (SVD) based decoding method, which strongly depends on the asymptotic orthogonality. The performance of the proposed semi-blind decoding method degrades. For a massive antenna system, the channels from different users are orthogonal when the number of BS antennas goes to infinity. [10]-[12] propose a singular value decomposition (SVD) based decoding method, which strongly depends on the asymptotic orthogonality. In [10], the channels from different users are estimated by eigenvalue decomposition of the channel covariance and the residual scale multiplicative ambiguity are addressed by pilot sequences. [11] and [12] use a principal component analysis (PCA) technique to project the received signals on a signal subspace, while power-controlled transmissions and pilot sequences are used to identify a desired user. It should be noticed that, when the number of the antennas is very limited, the channel estimation accuracy using the SVD decoding method degrades due to pilot contamination. In other words, the SVD decoding method greatly depends on the number of antennas.

[13] and [14] investigated an independent component analysis (ICA) based decoding method, which uses ICA to separate and decode the received signals (from in-cell and neighboring cells). For the ICA based decoding method, the pair-wise asymptotic orthogonality of channels from different users is not required. In [13], only a single terminal in each cell is
Considered and, in [14], multiple terminals in each cell is considered. ICA and pilots are used to separate and decode the received signals in [14]. However, the design of pilot sequences should be coordinated among neighboring cells. With the number of neighboring cells increases, the length of pilot sequences becomes large and the performance of the ICA method degrades due to low efficiency.

This paper studies the ICA based decoding method for a multicell multiuser massive MIMO uplink system in Rician/Rayleigh fading channels. The main contributions of this paper are as follows. First, a semi-blind ICA decoding method is developed based on channel energy levels and reference bits. ICA is used to separate and decode the received signals, channel energy levels are used to differentiate in-cell and neighboring cell signals, and reference bits are used to identify a desired signal from other in-cell signals. Second, the desired signal identification probability and bit error rate (BER) performance of the proposed semi-blind decoding method in fading channels are analyzed.

The remainder of this paper is organized as follows. In Section II, the signal model of a massive MIMO uplink system in Rician/Rayleigh fading channels is introduced. In Section III, a blind decoding method based on ICA and channel energy levels is presented. In Section IV, the desired signal identification probability as well as the BER performance of the proposed semi-blind decoding method is analyzed. In Section V, simulation results are presented and the proposed semi-blind decoding method is compared to methods using traditional MMSE decoding, SVD based decoding, and ICA decoding with cell coordinations. Finally, in Section VI, conclusions are drawn.

II. SIGNAL MODEL

In this paper, an $L$-cell massive MIMO uplink system will be considered in micro-cell Rician/Rayleigh fading channels. In each cell, there are one BS and $P$ independent terminal users. Each terminal user has a single antenna, while the BS is equipped with $M$ antennas.

For the micro-cell system, the signals from the desired cell are assumed to be Rician faded, which implies a direct line-of-sight signal component in in-cell transmissions. The interference signals from neighboring cells are assumed to be Rayleigh faded, since a direct line-of-sight between co-channel is unlikely to exist. This is called a Rician/Rayleigh fading environment [15]. Without loss of generality, the first cell is considered as the desired cell and the first user in the first cell is considered as the desired user. Thus, the base band observed matrix of the received signals at the desired BS (the base station of the desired user) can be written as

$$\mathbf{Y} = \sqrt{\rho} \mathbf{HBS} + \mathbf{N}_0$$ (1)

where $\mathbf{Y}$ is an $M$-by-$N$ observation matrix and it denotes the $N$ received samples at the base station with $M$ antennas. $\rho$ is the signal-to-noise ratio (SNR). $\mathbf{S} = (s_{1,1}, \ldots, s_{L,P}, \ldots, s_{L,P})^T$ is an $LP$-by-$N$ transmitted signal matrix, where $s_{l,p}$ is an $N$-by-$1$ signal vector sent by the $p$-th user in the $l$-th cell. For convenience, the transmitted signal is considered to be real. $\mathbf{B} = \text{diag}(\sqrt{\beta_{1,1}}, \ldots, \sqrt{\beta_{L,P}}, \ldots, \sqrt{\beta_{L,P}})$ is an $LP$-by-$LP$ diagonal matrix, and $\beta_{l,p}$ represents the path-loss and large-scale fading between the $p$-th transmitter in the $l$-th cell and the desired BS. In a cellular communications system, the mean value of a received signal power from in-cell are always larger than that from neighboring cells. Thus, for simplicity, $\beta_{1,1} = 1$. Without loss of generality, the first transmitter in the first cell is considered as the desired cell and the first user in the first cell is considered as the desired user. Finally, in Section VI, conclusions are drawn.

III. SEMI-BLIND DECODING BASED ON ICA

In a MIMO system, the MMSE decoding method and the ZF decoding method are widely used. In order to decode a desired signal, the channel information $\mathbf{H}$ should be known and, in a multicell multiuser massive MIMO uplink system, training sequences are always used to estimate the channel information. However, due to the pilot contamination, accurate channel information is difficult to be obtained. Without the accurate channel information, the decoding performance will be degraded significantly. In this paper, a semi-blind decoding method based on ICA channel energy levels and reference bits is proposed. The received signal is first processed using PCA to roughly estimate the channel and signals. The ICA process is then used to accurately estimate the channel and signals. The estimated channel energy levels and reference bits are used to identify the desired signal from the received signals (from the in-cell and neighboring cells). This semi-blind decoding method avoids the pilot contamination and achieves better decoding performance.

A. Transmitted Signal

In the context of the semi-blind decoding method, the signal transmission is divided into two phases. In the first phase, the reference bits are transmitted, which will be used to address
the ICA phase ambiguity issue and identify the desired signal.
In the second phase, data symbols are transmitted. Those two
phases can be explained as follows. According to (1), $s_{l;p}$ is
an $N$-by-$1$ signal vector sent by the $p$-th transmitter terminal
in the $l$-th cell. We assume $s_{l;p}(n) \in \{\pm 1\}$, which is the
$n$-th sample of $s_{l;p}$. $s_{l;p}(1) = 1$ is a phase reference bit.
$s_{l;p}(n) = \phi_p(n - 1), n = 2, \cdots, \log_2 P + 1$. $\phi_p(\cdot)$ are
additional reference bits for in-cell user separation.

B. PCA Processing of the Received Signal

PCA is a classical statistical method to extract the main
components from a set of observations, which has been widely
used in wireless communications. Recently, [11]-[12] use PCA
to estimate channels in a massive MIMO uplink system. In this
paper, PCA is proposed to extract the signal subspace from the
received signal matrix $Y$, given in (1). This PCA process is
known as pre-whitening, which can obtain a rough estimation
of the transmitted signals. The PCA process for the received
signal (pre-whitening) can be explained as follows.

Considering (1), since $s_{l;p}$ is an independent input from $p$-th
user in the $l$-th cell, it is clear that $E(\text{SS}^T) = \rho I_{LP}$. Obviously,
$E(N_0 N_0^T) = \rho I_M$. The covariance matrix of the received signal
can be written as

$$ C = E(YY^T) = \rho HBB^T H^T + I_M $$

Applyi.. SVd to the covariance matrix, we have
$C = U \Sigma U^T$, where $D = \text{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_M\}$ is an $M$
by-$M$ singular value diagonal matrix and its main diagonal
elements are sorted in descending order. $U$ is an $M$-by-$M$
orthogonal matrix. For a large scale antenna system, the
antenna number, $M$, will always be much larger than the user
number, $LP$. The received signal from $LP$ users and the
Gaussian noise are statistically independent. As a result, the
orthogonal matrix, $U$, can be divided into a signal subspace
matrix and a noise subspace matrix. It can be written as
$U = (U_S, U_N)$, where $U_s = (U_1, U_2, \cdots, U_{LP})$ is the
extracted main signal subspace and $U_N = (U_{LP+1}, \cdots, U_M)$
is the noise subspace. The singular value matrix can be
accurately divided into two parts, which can be written as
$D = \begin{pmatrix} D_s & 0 \\ 0 & D_N \end{pmatrix}$, where $D_s = \text{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_{LP}\}$ and
$D_N = \text{diag}\{\lambda_{LP+1}, \lambda_{LP+2}, \cdots, \lambda_M\}$. Since the received signal
sample number is limited, the sample covariance matrix of the
limited observed data is always used to replace the covariance
matrix. The sample covariance matrix, which is denoted as $\Sigma$,
can be calculated by $\Sigma \approx \frac{1}{N} \sum_{n=1}^{N} (Y_n - \bar{Y})(Y_n - \bar{Y})^T$.
$Y_n$ is the $n$-th column vector of the observed data matrix $Y$,
and $\bar{Y} = \frac{1}{N} \sum_{n=1}^{N} Y_n$ is the sample average. Based on the
statistical theory, when the sample number, $N$, is large enough,
$C = \Sigma$.

The extracted signal subspace is an approximated channel
estimation and the received signal will be projected on the
extracted signal subspace. This PCA based pre-whitening can
proximately estimate the signal and reduce the computational
complexity. The whitening operation using PCA can be
performed as follows

$$ X = D_s^{-\frac{1}{2}} U_s^T Y = AS + D_s^{-\frac{1}{2}} U_s^T N_0 $$

where $A = \rho^2 D_s^{-\frac{1}{2}} U_s^T H$ is a mixing matrix. $D_s^{-\frac{1}{2}} U_s^T N_0$
can be viewed as noise. The whitened data $X$ is an $LP \times N$
matrix, while the received data before whitening, $Y$, is an
$M \times N$ matrix. In the massive antenna system, the number of
antennas, $M$, is always much larger than the number of
users, $LP$. Thus, the whitening operation based on PCA will
reduce the computation complexity for the following ICA
based decoding process.

C. Separation of the Received Signals Based on ICA

The whitened data, which is performed using PCA, is an
approximate estimation of the received signal (mixed signals
from in-cell and neighboring cells). Separations of the
mixed signals can not be done at this stage. In a multicell
multiuser massive MIMO uplink system with Rician/Rayleigh
fading surrounding, the channel $H_{l;p,m}, l = 1, \cdots, LP; p = 1, \cdots, P; m = 1, \cdots, M$ is i.i.d. taken from the Gaussian
random variables. The channel matrix $HB$ will be of full
column rank. Since $U_s^T U_s = I$, the mixing matrix, $A$, given in
(3), will also be of full column rank. The elements of source
matrix $S$, which are taken from the independent terminals,
are also statistical independent. It can be concluded that (3)
describes a noisy blind source separation model. We can use
an ordinary ICA [16] or a noisy ICA [17] to separate the signal
from the pre-whitened data.

For the noisy ICA, the noise variance of the antennas should
be known in order to remove the bias, which is difficult to
estimate accurately. In a massive MIMO system, the large
scale antennas provide great space diversity and the noise
will be greatly reduced by the pre-whitening process [18].
In this paper, an ordinary ICA based on kurtosis is used to
estimate the signals. The kurtosis object function is immune
to Gaussian noise and thus the noise influence will be further
limited [19].

The ICA method will search for an orthogonal matrix $W =
[w_1, \cdots, w_{LP}]$ to separate the received signals. $W$ is used to
linearly transform the pre-whitening data $X$, which can be
written as

$$ Z = WX $$

where $Z$ is the estimated signals from the $LP$ independent
terminals. A fast fixed point ICA based on kurtosis is proposed
to separate the $LP$ independent signals. The following iterative
algorithm will be implemented to get the orthogonal matrix
$W$. Assume that $w_1, \cdots, w_{p-1}, 2 \leq p \leq LP$ have been
obtained. Following [16], the $p$-th orthogonal vector $w_p$ can
be obtained as follows.

1. Select a random unit-norm as the initial vector $w_p(0)$.
2. Update $w_p$ by the following algorithm

$$ w_p(k + 1) = \frac{1}{N} \sum_{n=1}^{N} (X_n g(w_p(k)^T X_n)) $$

$$ + \frac{1}{N} \sum_{n=1}^{N} (g'(w_p(k)^T X_n)) w_p(k) $$

where $g(a) = a^3$ is a given kurtosis function. $g'(a)$ is the
derivative of $g(a)$. $X_n$ is the $n$-th column vector of $X$. 

3. Gram-Schmidt-like decorrelation [20] is used to prevent the new iterations from converging to the same maximus. It means that the estimated outputs \(w_1, w_2, \ldots, w_{p-1}\) should be decorrelated. It can be expressed as

\[
w_p(k + 1) = w_p(k + 1) - \sum_{j=1}^{p-1} w_j^T w_p(k + 1)
\]

(6)

4. Divide \(w_p\) by its norm

\[
w_p(k + 1) = w_p(k + 1)/(w_p(k + 1)^T w_p(k + 1))
\]

(7)

5. If \(|w_p(k + 1)^T w_p(k)|\) is not close enough to 1, let \(k = k + 1\), and go back to Step 2. Otherwise, output the vector \(w_p\).

6. If \(p = LP\), the ICA separation process completes. Otherwise, let \(p = p + 1\) and go back to Step 1.

With both the mixing matrix and the signal sources are real valued, the convergence speed of fast fixed point ICA is known to be cubic [21]. It means a fast convergence for ICA iteration.

When the demixing matrix \(W\) is obtained by ICA, the signal \(S\) can be estimated through (4). Unfortunately, the ICA process may introduce a phase ambiguity [16]. For a BPSK modulation system, the \(\pi\) phase ambiguity issue can be addressed by a one-bit reference, with a negligible performance loss [13].

**D. In-Cell Signal Identification Based on Signal Power**

The basic ICA method can separate and decode the received mixed signals, which include the in-cells and the neighboring cells. However, the order of the estimated signals can not be determined. It means that it is difficulty to identify the desired signal from all the estimated signals. In this paper, the in-cell signals are first differentiated from the neighboring cell signals and then the desired signal is identified from the in-cell signals.

From the assumptions of the system, it is reasonable to consider that the signals from the in-cell always have more energy than the signals from the neighboring cells. The energy difference is due to the in-cell line-of-sight path gain as well as the larger path loss and large scale fading of signals from neighboring cells. This channel energy difference is significant for a system with massive scale antennas. Thus, the channel energy can be used to identify the in-cell signals from all the estimated signals.

Therefore, in the \(L\)-cell system, with \(P\) users in each cell, the signals from the in-cell can be obtained by selecting the \(P\) highest energy signals.

The channel of the desired signal can also be estimated by the above ICA process, which can be explained as follows. Following (3) and (4), when the ICA process is completed, we have \(WA = I\), which can be rewritten as

\[
\rho^2 W^T D_s^{-1/2} U_s^T \mathbf{H} B = I
\]

(8)

Following the SVD process, the observed signal covariance matrix is divided into the transmitted signal subspace and the noise subspace, which can be written as \(E(YY^T) = U_s D_s U_s^H + U_n D_n U_n^H\). In large scale antenna wireless communication systems, the signal subspace energy is always much larger than the noise subspace energy. Thus, the observed signal covariance matrix can be approximately expressed as

\[
\rho^2 \mathbf{H} B^T \mathbf{H} = U_s D_s U_s^T
\]

(9)

From (8) and (9), the channel matrix \((\mathbf{H} B)^T\) can be estimated. The estimated channel vector of the \(p\)-th user in the \(l\)-the cell can be written as

\[
\hat{H}_{l,p} \sqrt{\beta_{l,p}} = \rho^2 w_{i,p}^T D_s^{-1/2} U_s^T l = 1, \ldots, L, p = 1, \ldots, P
\]

(10)

where \(w_{i,p}\) is obtained in the ICA process, the signal subspace, \(U_s\), and the signal eigenvalue, \(D_s\), are obtained in the SVD process. In order to identify the in-cell users, the channels between the base station and the in-cell users, \(H_{1,p}, p = 1, \ldots, P\), are considered, which following Rician fading. The \(P\) channels with maximum energy levels among the \(LP\) estimated channels will be selected. Thus, the corresponding demixing vectors, \(w_1, p = 1, \ldots, P\), can also be obtained, and the \(P\) in-cell users signals can be estimated as

\[
s_{i,p} = s_{i,1} + w_{i,p}^T D_s^{-1/2} U_s^T N_0, p = 1, \ldots, P
\]

(11)

**E. Desired Signal Identification Using Reference Bits**

In order to identify the desired signal from the selected \(P\) in-cell signals, a few reference bits (denoted as \(\phi = [\phi_1, \ldots, \phi_p, \ldots, \phi_P]\)) can be used. \(\phi_p\) is assigned to the \(p\)-th in-cell user. In the reference bit vector, \(\phi_p(n)\) is \(\pm 1\), and reference vector \(\phi_p\) should not be the same with \(\phi_p\) when \(p \neq q\). With \(P\) independent users in one cell, only \([\log_2 P]\) bits are required to differentiate all the in-cell signals, where \([x]\) is a ceiling function. For example, with \(P = 4\), the reference bit matrix can be designed as \(\Phi = [\phi_1, \phi_2, \phi_3, \phi_4] =

\[
\begin{bmatrix}
1 & 1 & -1 & -1 \\
-1 & 1 & 1 & 1
\end{bmatrix}
\]

Suppose that \(\phi_1\) represents the reference bit vector of the desired signal, and \(\phi_p, p = 1, \ldots, P\), is the estimated \(p\)-th signal reference bit vector from (11). The desired signal, \(s_{1,1}\), can be identified as

\[
s_{1,1} = s_{i,1,k}, k = \arg \max\{|\phi_1^T \phi_p|, p = 1, \ldots, P\}
\]

(12)

According to (12), the reference bit vector design does not require cell coordination and the reference bit vector identification performance will not be affected by the neighboring cells. On the contrast, the ICA decoding method given in [14] needs cell coordination for pilot sequence designs and the length of a pilot sequence is \(4 \times [\log_2 P]\) (instead of \(\log_2 P\)).

**F. Steps in the Proposed Semi-Blind Decoding Method**

In summary, the proposed semi-blind decoding procedure are as follows.

1. Based on the user number in each cell, \(P\), \([\log_2 P]\) reference bits are designed.

2. For each user, \([\log_2 P] + 1\) reference bits and \(N - [\log_2 P] - 1\) signal data are transmitted.

3. For the received \(M\)-by-\(N\) signal matrix, calculate the received signal sample covariance matrix and project it on the extracted signal subspace as (3).
4. Implement ICA on the pre-whitening data, \( X \), and obtain the linear transform matrix \( W \). The LP user data, \( S \), and the channel, \( HB \), are also estimated using (10) and (11).

5. \( P \) signals with largest channel energy levels will be selected among LP signals and the in-cell signals are identified.

6. Use the one-bit phase reference to correct the phase ambiguity for the selected \( P \) in-cell signals.

7. Use the designed \( \left\lfloor \frac{\log_2 P}{5} \right\rfloor \) reference bit vector \( \phi_p \) to identify the desired \( p \)-th signal from the \( P \) in-cell signals.

G. Complexity and Efficiency Analysis

The computational complexity is discussed in this section and the number of addition and multiplication operations will be considered. According to [22], when the number of the antennas of the BS station is \( M \), the number of the cells is \( L \), and each cell has \( P \) users, the SVD process complexity is of the order of \( O(M^2LP) \). By (3), the pre-whitening of the observed matrix has the computational complexity of \( O(MNL^2P^2) \) when the block sample number is \( N \). Finally, from (6) and (7), the complexity separation of one signal from the mixed signals using fast ICA is of the order of \( O(NLP) \) [22]. When all \( LP \) users are separated using fast ICA, the computational complexity is of \( O(NL^2P^2) \).

Only \( \left\lfloor \frac{\log_2 P}{5} \right\rfloor \) bits are required to differentiate the in-cell signals. As a comparison, [14] uses ICA to separate the received signals and the desired signal should be selected from all the \( LP \) signals (from in-cell and neighboring cells). The pilot sequences are designed with cell coordination and \( \left\lfloor \frac{\log_2 LP}{5} \right\rfloor \) bits are required to differentiate the \( LP \) signals. The MMSE decoding method [2] requires \( P \) pilot bits and the performance will be affected by neighboring cells’ pilot bits.

The above two methods (ICA [14] and MMSE [2]) require more reference bits than the proposed semi-blind decoding method and the proposed semi-blind decoding method thus has a higher transmit efficiency.

IV. PERFORMANCE ANALYSIS OF THE PROPOSED SEMI-BLIND DECODING SCHEME

The proposed semi-blind decoding method uses ICA to separate the received signals. The energy levels of the estimated channels are used to identify the in-cell signals (from the mixed in-cell signals and neighboring cell signals). The reference bits are used to identify the desired signal from in-cell signals. Thus, the performance of the proposed semi-blind decoding method depends on the probabilities of the in-cell signal identification and the desired signal identification.

A. Probability of the In-Cell Signal Identification

The following presents the derivations of the probability of the in-cell signal identification, which implies that the \( P \) signals with highest energy levels are in-cell signals. Let \( E_{l,p}, l = 1, \cdots, L, p = 1, \cdots, P \), be the energy of the channel \( \sqrt{\beta_{l,p}} \mathbf{H}_{l,p}, l = 1, \cdots, L, p = 1, \cdots, P \), where \( \sqrt{\beta_{l,p}} \mathbf{H}_{l,p} \) represents the channel between the \( p \)-th user in the \( l \)-th cell and the multiple antennas of the in-cell base station. The channel energy, \( E_{l,p} \), can be written as

\[
E_{l,p} = \beta_{l,p} \mathbf{H}_{l,p}^T \mathbf{H}_{l,p}, l = 1, \ldots, L, p = 1, \cdots, P \tag{13}
\]

The \( P \) highest energy signals are selected from the estimated \( LP \) signals. Suppose that the first user in the first cell, \( s_{1,1} \), is the desired user. Let \( v \in \{1, \cdots, P\} \) be the number of in-cell signals included in the \( P \) selected signals and \( P - v \) is the number of neighboring cell signals included in the \( P \) highest energy signals. Define \( A \) as the event that the desired user is selected in the \( P \) highest energy signals. Let \( B_v \) be the event that \( v \) in-cell signals are selected in the \( P \) highest energy signals. Then, the remainder are from the neighboring cells and the desired signal is included in the \( v \) in-cell signals.

The probability that the desired signal is in the selected \( P \) in-cell signals can be calculated as

\[
P_d = P(A) = \sum_{v=1}^{P} P(A|B_v) P(B_v) \tag{14}
\]

where \( P(B_v) \) can be calculated as

\[
P(B_v) = \left( \frac{P-v}{P} \right)^v = \frac{v!}{P^v} \tag{15}
\]

The probability \( P(A|B_v) \) can be calculated as

\[
P(A|B_v) = \sum_{n=1}^{K_1} \sum_{m=1}^{K_2} \Pr\{E_{\min}(v, m, n) > E_{\max}(v, m, n)\}
\]

\[
K_1 = \left( \frac{(L-1)P}{P-v} \right), K_2 = \left( \frac{P}{v} \right) \tag{16}
\]

and

\[
E_{\min}(v, m, n) = \min\{E_{\in}(v, m), E_{\nb}(P-v, m)\}
\]

\[
E_{\max}(v, m, n) = \max\{E_{\in}(P-v, m), E_{\nb}((L-2)P+v, n)\} \tag{17}
\]

where \( E_{\in}(v, m) \) denotes the channel energy of \( v \) in-cell signals, which is the \( m \)-th in-cell combination selecting \( v \) signals from \( P \) in-cell signals. It can be expressed as \( E_{\in}(v, m) = \{E_{\in,m,1}, \cdots, E_{\in,m,k}, \cdots, E_{\in,m,v}\} \), \( E_{\nb}(P-v, m) \) is the energy of the remaining \( P-v \) in-cell signals. \( E_{\nb}(P-v, n) \) denotes the channel energy from the \( P-v \) neighboring cells, which is the \( n \)-th neighboring cell combination selecting \( P-v \) signals from \( (L-1)P \) signal in neighboring cells. It can be expressed as \( E_{\nb}(P-v, n) = \{E_{n,1}, \cdots, E_{n,k}, \cdots, E_{n,P-v}\} \), \( E_{\nb}((L-2)P+v, n) \) is the channel energy of the remaining \( (L-2)P+v \) neighboring cells.

\[
E_{\min}(v, m, n) \text{ is the minimum channel energy from the selected } P \text{ highest energy signals with } m \text{-th in-cell combination and } n \text{-th neighboring cells combination, and } E_{\max}(m, n) \text{ is the maximum channel energy from the remaining } (L-1)P \text{ signals. For the multicell multiuser wireless system, the transmission channels, } \sqrt{\beta_{l,p}} \mathbf{H}_{l,p} \text{ are statistically independent. The probability } P(A|B_v)(m, n) = \Pr\{E_{\min}(v, m, n) > E_{\max}(v, m, n)\} \text{ can be calculated as}
\]
\[
P(A|B_e)(m,n) = \int_{-\infty}^{\infty} \Pr\{E_{\min}(v,m,n) > E_{\max}(v,m,n)\} f_{E_{\max}}(y,v,m,n)dy
\]
\[
= \int_{-\infty}^{\infty} \int_y f_{E_{\min}}(x,v,m,n)dx f_{E_{\max}}(y,v,m,n)dy
\]
\[
= \int_{-\infty}^{\infty} f_{E_{\min}}(x,v,m,n)dx f_{E_{\max}}(y,v,m,n)dy
\]
where \(f_{E_{\min}}(x,v,m,n)\) is the probability density function (PDF) of \(E_{\min}(v,m,n)\), and \(f_{E_{\max}}(x,v,m,n)\) is the PDF of \(E_{\max}(v,m,n)\).

Considering (17) and (18), the distribution of \(E_{l,p}, l = 1,\cdots,L, p = 1,\cdots,P\), will be analyzed to obtain the closed form expression of \(P_d\).

In a Rician/Rayleigh fading channel, the channel energy from the in-cell signals, \(E_{1,p} = 1,\cdots,P\), can be rewritten as
\[
E_{1,p} = \beta_{1,p} \sum_{m=1}^{M} h_{1,p,m}^2
\]
where the in-cell channels are Rician faded, and \(h_{1,p,m} \sim N(\mu_{up}, 1)\). Thus \(E_{l,p} = \beta_{l,p} \sum_{m=1}^{M} h_{l,p,m}^2\) is noncentral chi-squared distributed with mean value, \(\mu_{up}\), and degrees of freedom \(M\). The PDF and cumulative distribution function (CDF) of \(E_{l,p}\) are denoted as \(f_{nc}(x, \mu_{up}, M)\) and \(F_{nc}(x, \mu_{up}, M)\). The PDF and CDF of \(E_{1,p} = 1,\cdots,P\), can be expressed as \(f_{E_{1,p}}(x) = f_{nc}(x, \mu_{up}, M)\) and \(F_{E_{1,p}}(x) = F_{nc}(x, \mu_{up}, M)\) respectively.

\[
f_{nc}(x, \mu_{up}, M) = \frac{1}{I_{M/2-1}(\sqrt{\mu_{up}})} \frac{1}{2} e^{-(x+\mu_{up})/2} \frac{x^{(M/2)-1}}{\Gamma(M/2)}
\]
where \(I_{M/2-1}(x)\) is a modified Bessel function of the first kind.

In the Rician/Rayleigh fading channel, the channel energy from the neighboring cells, \(E_{l,p}, l = 2,\cdots,L, p = 1,\cdots,P\), can be rewritten as
\[
E_{l,p} = \beta_{l,p} \sum_{m=1}^{M} h_{l,p,m}^2\]
where the neighboring cell channels are Rayleigh faded, and \(h_{l,p,m} \sim N(0,1)\). Thus \(E_{l,p} = \beta_{l,p} \sum_{m=1}^{M} h_{l,p,m}^2\) are central chi-squared distributed with degrees of freedom \(M\). The PDF and CDF of \(E_{l,p} = 2,\cdots,L, p = 1,\cdots,P\), are denoted as \(f_{c}(x,M)\) and \(F_{c}(x,M)\). The PDF and CDF of \(E_{l,p} = 2,\cdots,L, p = 1,\cdots,P\), can be expressed as \(f_{E_{l,p}}(x) = f_{c}(x,M)\) and \(F_{E_{l,p}}(x) = F_{c}(x,M)\) respectively.

\[
f_{c}(x,M) = \frac{x^{M/2-1} e^{-x/2}}{2^{M/2} \Gamma(M/2)}, x > 0
\]
where \(\Gamma(M/2)\) denotes the Gamma function. The PDF of \(E_{\min}(v,m,n)\) can be expressed as [24]
maximum likelihood (ML) decoding method, which can be written as [25]

\[ P_e(\rho|h_{1,1}) \geq \frac{1}{\sqrt{\pi}} \int_0^\infty \int_{\rho \beta_{1,1} \Sigma_{m=1}^M h_{1,1,m}^2} e^{-x^2} \, dx \, dy \]  

(26)

where \( \Sigma_{m=1}^M h_{1,1,m}^2 \) is noncentral chi-squared distributed with mean value, \( M \mu_1 \), and degrees of freedom \( M \). The average BER of the desired signal can be calculated as

\[ P_e(\rho) \geq \frac{1}{\sqrt{\pi}} \int_0^\infty \int_{\rho \beta_{1,1} y} e^{-x^2} \, dx \, \frac{1}{2} e^{-(y+M \mu_1)/2} \]  

\[ \left( \frac{y}{M \mu_1} \right)^{\frac{M}{2} - 1} I_{M/2 - 1} \left( \sqrt{M \mu_1 y} \right) dy \]  

(27)

When all \( \left[ \log_2 P \right] \) bits are correctly decided, the desired signal is identified from the selected in-cell signals. Thus, the identification probability with reference bits can be written as

\[ P_{d,bits}(\rho) \leq \left( 1 - \frac{1}{\sqrt{\pi}} \int_0^\infty \int_{\rho \beta_{1,1} y} e^{-x^2} \, dx \, \frac{1}{2} e^{-(y+M \mu_1)/2} \right) \]  

\[ \left( \frac{y}{M \mu_1} \right)^{\frac{M}{2} - 1} I_{M/2 - 1} \left( \sqrt{M \mu_1 y} \right) dy \]  

\[ \left[ \log_2 P \right] \]  

(28)

According to (28), the probability of the desired signal identification with reference bits is related to large scale fading, \( \beta_{1,1} \), and Rice factor, \( K_1 \). It is also related to the number of signals in each cell, \( P \), and the number of antennas at BS, \( M \). The desired signal identification performance with reference bits improves with increasing \( M \), \( K_1 \) and \( \beta_{1,1} \), and the identification performance degrades with increasing \( \left[ \log_2 P \right] \).

C. BER Performance of the Proposed Semi-Blind Decoding

The following presents derivations of the BER performance of the proposed semi-blind decoding. The BER performance evaluation can be considered in three scenarios. In the first scenario, the \( P \) in-cell signals are not selected correctly. In the second scenario, the \( P \) in-cell signals are selected correctly, but the desired signal identification is incorrect. The final scenario is that the in-cell signal identification and the desired signal identification are both correct. The BER performance of the proposed semi-blind decoding can be calculated as

\[ P_{E,LB}(\rho) \geq \frac{1 - P_d}{2} + P_d \left( 1 - P_{d,bits}(\rho) \right) + P_d P_{d,bits}(\rho) P_e(\rho) \]  

(29)

(29) is a lower bound (LB) BER performance. When SNR, \( \rho \), approaches infinity, \( P_e(\rho) \) goes to 0 and the desired signal identification with reference bits, \( P_{d,bits}(\rho) \) becomes 1. Thus, the LB BER performance \( P_{E,LB}(\rho) \) becomes 1 - \( P_d \). It can be written as

\[ P_{E,LB}(\infty) \geq 1 - \frac{P_d}{2}. \]  

(30)

According to (30), when \( \rho \) goes to infinity, the LB BER performance is decided by in-cell signal identification probability, \( P_d \).

Fig. 1. \( P_d \) versus \( K_p \), with different number of cells \( (\text{SNR}=0\text{dB}, \beta_{1,p} = 1, \beta_{l,p} = 0.8, l = 2, \ldots, L, p = 1, \ldots, P, L = 3, M = 50) \).

V. Simulation Results

In this section, both simulation and theoretical results are shown to illustrate the performance of the proposed semi-blind decoding method based on ICA. The performance of the widely used MMSE decoding method [2], the SVD based decoding method [12] and the decoding method based on ICA with cell coordination [14] are also shown for comparison. MMSE channel estimation is used in the MMSE decoding [6] when the channel is unknown. The channel is assumed to be block faded and the transmitted signal is BPSK modulated in the simulation.

Figs. 1, 2, and 3 present the simulation and theoretical results of the in-cell signal identification probability, \( P_t \), considering the number of cells, \( L = 3 \), the antenna number at BS, \( M = 50 \), and SNR = 0dB. It can be seen that the simulation results match well with the analytical results (25).

In Fig. 1 and Fig. 2, the large scale fading of the in-cell signals are set to be \( \beta_{1,p} = 1, p = 1, \ldots, P \), and the large scale fading of the neighboring cell signals are set to be \( \beta_{1,p} = 0.8, p = 1, \ldots, P, l = 2, \ldots, L \). From Fig. 1 and Fig. 2, it can be concluded that the in-cell signal identification probability improves with the increase of the number of antennas, \( M \), and degrades with increasing user number in each cell, \( P \). With the increase of Rician factor, the energy difference between the in-cell signals and the neighboring cell signals increases. Thus, the identification probability improves.

In Fig. 3, the number of antennas at BS, \( M \), is 50. The number of cells, \( L \), is 3, and each cell has two users. The large scale fading of the in-cell signals are set to be \( \beta_{1,p} = 1, p = 1, \ldots, P \). It is shown that the in-cell signal identification probability improves with decreasing \( \beta_{1,p} \). It is also shown that the signal energy difference between the in-cell signals and neighboring cell signals increases with decreasing \( \beta_{1,p} \). It is also shown that the in-cell signal identification probability will be close to 1, even if the Rician fading factor becomes 0 (Rayleigh fading.
blind ICA decoding, with block size \( N \). Semi-blind ICA decoding outperforms MMSE decoding with cell coordination. For MMSE decoding, pilot sequences are reused in neighboring cells. With the proposed semi-blind ICA decoding method improvements. The MMSE decoding method does not depend on the orthogonality. In Fig. 4 and Fig. 5, the large scale fading of the in-cell signals are set to be \( \beta_{1,p} = 1, p = 1, \cdots, P \). The number of antennas at BS, \( M \), is 50. The number of cells, \( L \), is 3, and each cell has 4 users.

Fig. 4. The proposed semi-blind ICA decoding compared with MMSE decoding (\( \beta_{1,p} = 1, \beta_{l,p} = 0.6, l = 2, \cdots, L, p = 1, \cdots, P, L = 3, P = 4, M = 50 \)).

Fig. 5. The proposed semi-blind ICA decoding compared with SVD decoding (\( \beta_{1,p} = 1, K_p = 0.6, p = 1, \cdots, P, L = 3, P = 4, M = 50 \)).
In the neighboring cells, the SVD decoding performance degrades severely, and the performance of the proposed semi-blind ICA decoding method is almost the same with different neighboring cells in large scale fading. 

Fig. 6 shows the BER performance of the proposed semi-blind decoding method compared with ICA decoding with cell coordination [14]. In the simulation, the large scale fading of the in-cell signals are set to be $\beta_{1,p} = 1, p = 1, \cdots, P$, and the large scale fading of the neighboring cell signals are set to be $\beta_{l,p} = 0.6, p = 1, \cdots, P, l = 2, \cdots, L$. The number of the BS antennas, $M$, is 50. The Rician factor $K_p = 0.6, p = 1, \cdots, P$. The proposed semi-blind ICA decoding method uses the channel energy levels to differentiate the in-cell signals and neighboring cell signals. Reference bits are designed without cell coordination, and only $\lceil \log_2 P \rceil$ reference bits are needed to identify the desired signal in the in-cell signals. The ICA decoding method with cell coordination uses reference bits to identify the desired signal from all the signals (in-cell signals and neighboring cell signals), and $\lceil \log_2 L P \rceil$ reference bits are needed. It can be concluded that the proposed semi-blind decoding method has better BER performance and transmit efficiency than the ICA decoding method with cell coordination [14].

Following (29) and (30), Fig. 7 shows the analytical lower bound BER performance of the proposed semi-blind ICA decoding method. In the simulation, the large scale fading of the in-cell signals are set to be $\beta_{1,p} = 1, p = 1, \cdots, P$, and the large scale fading of the neighboring cell signals are set to be $\beta_{l,p} = 0.6, p = 1, \cdots, P, l = 2, \cdots, L$. The number of BS antennas, $M$, is 50. It can be seen that the lower bound is tight under high SNR. The lower bound is loose under low SNR due to the unstable ICA iteration.

VI. CONCLUSIONS

In this paper, a semi-blind decoding scheme based on ICA and reference bits is proposed for a multicell multiuser massive MIMO uplink system, in Rician-Rayleigh fading channels. The proposed semi-blind decoding scheme operates with only a few reference bits and avoids pilot contaminations caused by pilot sequence reuses in neighboring cells. Both analytical and simulation results show that the proposed blind decoding method outperforms MMSE decoding, SVD based decoding and ICA decoding with cells coordination. The proposed semi-blind decoding method also improves the transmission efficiency due to fewer reference bits.

REFERENCES