How Many Signals Can Be Sent in a Multi-Cell Massive MIMO System

Haiquan Wang, Dakuan Yang, Xiao Li, and Peng Pan

***Abstract*—In this letter, we study how large the signal constellation can be under pilot contamination in a multi-cell massive multiple-input multiple-output uplink system. Specifically, based on an equivalent transmit–receive signal model, we observe that the transmitted signals by all users in different cells are compacted into a low-dimensional space when they arrive at the target base station. Then, the equivalent model reveals that the size of constellations used by all users is limited by a parameter, which is related to the large-scale fading factors only. We find that when the size of the signal constellation is smaller than the parameter, the error probability goes to zero. Otherwise, an error floor occurs regardless of the number of antennas and the pilot signal to noise ratio. Our simulations results validate the proposed analytical results.**

***Index Terms*—Pilot reuse, pilot contamination, multi-cell, massive MIMO.**

I. Introduction

MASSIVE multiple-input, multiple-output (MIMO) technology, where the base station (BS) is equipped with a large number of antennas and serves several single-antenna users at the same frequency band simultaneously, has attracted significant research interest recently. It has been demonstrated that the system can significantly increase the total throughput with simple signal processing algorithms, if the channel state information (CSI) is available for the receiver. However, when the CSI is not available, the performance will be greatly degraded, see [1]–[3]. Thus, the great benefits of the massive MIMO system are heavily based on the assumption that the receiver has the CSI.

Generally, the CSI is acquired from training-based channel estimation. For an uplink, each user transmits pilot signals to the BS, and then, the BS estimates the CSI based on the received pilots. Obviously, the channel estimation accuracy depends on the pilot signals. One of the pilot signal designs is that the pilot signals transmitted by different users are mutually orthogonal, but the orthogonality consumes a lot of radio resources. Hence, non-orthogonal pilot signals have been considered, where the users in different cells may use the same pilot signals, and the users in the same cell use orthogonal pilots, referred to as "*pilot reuse*" [1]. In the pilot reuse, channel estimation is interfered by pilot signals transmitted by the users in the neighboring cells, which degrades the performance of the estimator drastically. The phenomenon is called as *pilot contamination* [1].

To alleviate pilot contamination, a lot of works on channel estimation and signal detection were proposed, see [4]–[11]. In [4], it was shown that, under certain conditions on channel covariance, the pilot contamination effects can be removed when the number of transmit antennas goes to infinity. In [5], a new channel estimator was proposed based on subspace projection, and it was shown that the pilot contamination effects were mitigated under some conditions on the coherence time. An eigenvalue decomposition based channel estimation approach was proposed in [6]. More recently, the subspace based channel estimations were considered and analysed in [7] for a finite-dimensional physical channel model, where signals impinge on the base station from a finite number of angles of arrival. In [8], pilot optimization and channel estimation were formulated to an optimization problem and can be solved by using iterative algorithms. In [9], pilot beam pattern design was investigated and a new algorithm was proposed for the optimal channel estimation. Recently, partially decoded data were used to aid channel estimation in [10]. In [11], a multi-cell zero-forcing (ZF) detector which exploits and orthogonalizes all available directions to mitigate parts of the inter-cell interference was proposed. Also, in [12], an iterative detector was considered to reduce decoding complexity.

In this letter, for a multi-cell massive MIMO uplink system, we establish an equivalent transmit-receive signal transmission model. A detailed analysis on this model reveals the following fact: if the size of the signal constellations used by all users is smaller than a parameter, which only depends on the large-scale fading factors, the error floor does not exist when both the number of antenna in the BS and the signal-to noise (SNR) of the pilots go to infinity. Otherwise, the error floor exists.

The rest of this letter is organized as follows: in Section II, system model is introduced. In Section III, an equivalent transmit-receive signal model is given, and an impact of this model is analyzed. Simulation results are shown in Section IV. Finally, the last section concludes this letter.

Notations: For a matrix $\mathbf{A}$, $\mathbf{A}^t$ denotes its transpose, $\mathbf{A}^H$ denotes its conjugate transpose, $\det(\mathbf{A})$ denotes its determinant, $\mathrm{rank}(\mathbf{A})$ denotes its rank, and $\mathrm{tr}(\mathbf{A})$ denotes its trace.

II. System Model

We consider uplink of a multi-cell, massive MIMO with L cells. Assume that the BS at each cell is equipped with M antennas, and supports K independent users, each with single antenna. The channel gains consist of two parts: one

Manuscript received October 1, 2017; revised November 16, 2017; accepted November 20, 2017. Date of publication November 29, 2017; date of current version June 19, 2018. This work was supported by the National Natural Science Foundations of China under Grant 61372093 and Grant 61728105. The associate editor coordinating the review of this paper and approving it for publication was Y. Gao. *(Corresponding author: Haiquan Wang.)*

The authors are with the School of Communications Engineering, Hangzhou Dianzi University, Hangzhou 300018, China (e-mail: tx_wang@hdu.edu.cn; yang_da_kuan@163.com; 15869004732@163.com; panpeng@hdu.edu.cn).

Digital Object Identifier 10.1109/LWC.2017.2778240

contains path loss and shadowing effects (also called as the large-scale fading factors), and the other is instantaneous fading gain. Thus, the channel gain between the k-th user of the l-th cell and the m-th antenna in the first BS can be expressed as $\sqrt{\beta_{mkl}} h_{mkl}$, where β_{mkl} denotes the large-scale fading based on the path-loss and shadowing, and h_{mkl} represents the small-scale fading. Also, we assume that $\beta_{mkl} = \beta_{kl}$ for all $m = 1, 2, \ldots, M$ without loss of generality. The fading factor h_{mkl} is an independent and identically distributed (i.i.d.) zero-mean circularly-symmetric complex Gaussian random variable with unit variance. All fading factors h_{mkl} are assumed to be unknown to all the users and the BS, while the large-scale fading factors are known to the BSs.

The input-output equation of the system can be written as

$$\mathbf{Y} = \sqrt{\rho}\mathbf{HBX} + \mathbf{W}, \tag{1}$$

where $\mathbf{Y}$ is the received signal vector in the BS, matrix $\mathbf{B}$ is a diagonal matrix, which can be expressed as

$$\mathbf{B} = \mathrm{diag}(\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_L)$$

and $\mathbf{B}_l = \mathrm{diag}(\sqrt{\beta_{1l}}, \sqrt{\beta_{2l}}, \ldots, \sqrt{\beta_{Kl}})$. The $M \times KL$ channel matrix $\mathbf{H}$ can be represented as follows:

$$\mathbf{H} = (\mathbf{H}_1\ \mathbf{H}_2\ \cdots\ \mathbf{H}_L) \tag{2}$$

where $\mathbf{H}_l = (h_{mkl})_{M\times K}$. Finally, ρ is the average received SNR, $\mathbf{X}$ is the transmitted signal vector, and $\mathbf{W}$ is the Gaussian noise matrix.

Consider that the system adopts the training-based transmission scheme, where the signal transmissions are divided into two phases. At the first phase, the k-th user in the l-th cell sends τ pilot signals $\mathbf{\Phi}_{kl} = (\phi_{kl,1}, \phi_{kl,2}, \ldots, \phi_{kl,\tau})$, $k = 1, 2, \ldots, K$ and $l = 1, 2, \ldots, L$. Denote

$$\mathbf{\Phi}_l = (\mathbf{\Phi}_{1l}^t, \mathbf{\Phi}_{2l}^t, \ldots, \mathbf{\Phi}_{Kl}^t)^t, \quad \mathbf{\Phi} = (\mathbf{\Phi}_1^t, \mathbf{\Phi}_2^t, \ldots, \mathbf{\Phi}_L^t)^t \tag{3}$$

Thus, matrix $\mathbf{\Phi}_l$ of size $K \times \tau$ and matrix $\mathbf{\Phi}$ of size $LK \times \tau$ denote the pilot signals transmitted by the users in the l-th cell and by users from all cells, respectively. We normalize the energy of the pilot signals as follows:

$$\mathbf{\Phi}_{kl}\mathbf{\Phi}_{kl}^H = 1, \quad \text{and hence} \quad \mathrm{tr}(\mathbf{\Phi}\mathbf{\Phi}^H) = KL. \tag{4}$$

The minimum mean square error (MMSE) estimator of channel matrix $\mathbf{H}$, denoted as $\hat{\mathbf{H}}$, can be written as [13]

$$\hat{\mathbf{H}} = \sqrt{\rho_0}\mathbf{Y}_0\Big(\mathbf{I}_\tau + \rho_0\mathbf{\Phi}^H\mathbf{B}^2\mathbf{\Phi}\Big)^{-1}\mathbf{\Phi}^H\mathbf{B}^H, \tag{5}$$

where ρ_0 is the average SNR.

At the second phase, users in different cells transmit their own data signals simultaneously. Denote the signals transmitted by users in the l-th cell as $\mathbf{X}_l = (x_{1l}, x_{2l}, \ldots, x_{Kl})^t$, where x_{kl} represents the signal transmitted by the k-th user in the l-th cell. Denote $\mathbf{X} = (\mathbf{X}_1^t, \mathbf{X}_2^t, \ldots, \mathbf{X}_L^t)^t$. Then, the MMSE signal detection is [13]

$$\hat{x}_{k1} = \arg\min_{x_k} \|\mathbf{G}_{mmse,k}\tilde{\mathbf{Y}} - \sqrt{\tilde{\rho}}x_k\|, \tag{6}$$

where $\tilde{\mathbf{Y}} = \mathbf{Y}/\sigma$, and $\mathbf{G}_{mmse,k}$ denotes the k-th row of matrix $\mathbf{G}_{mmse}$ and

$$\mathbf{G}_{mmse} = \left(\frac{\mathbf{I}_{KL}}{\tilde{\rho}} + \mathbf{B}^H\hat{\mathbf{H}}^H\hat{\mathbf{H}}\mathbf{B}\right)^{-1}\mathbf{B}^H\hat{\mathbf{H}}^H, \tag{7}$$

and $\sigma^2 = 1 + \rho\mathrm{tr}(\mathbf{B}\mathbf{A}^{-1}\mathbf{B})$, $\mathbf{A} = \mathbf{I}_{KL} + \rho_0\,\mathbf{B}\mathbf{\Phi}\mathbf{\Phi}^H\mathbf{B}^H$, $\tilde{\rho} = \rho/\sigma^2$.

III. An Equivalent Model of the System

In this section, we derive an equivalent model of the system in Section II and then show some impacts on the system performance.

A. An Equivalent Model

Assume that $\Phi_l = \Phi_0$ for $l = 1, 2, \ldots, L$, where Φ_0 is a $K \times K$ unitary matrix. That is, the pilots used by users in the same cell are orthogonal but are reused in other cells. Let $\mathbf{U}_0\ \mathbf{D}_0\mathbf{V}_0^H$ be the Singular Value Decomposition (SVD) of matrix $\mathbf{B\Phi}$, that is, $\mathbf{B\Phi} = \mathbf{U}_0\mathbf{D}_0\mathbf{V}_0^H$, where $\mathbf{U}_0$ is a $KL \times K$ unitary matrix, $\mathbf{D}_0$ is a $K \times K$ diagonal matrix and $\mathbf{V}_0$ is a $K \times K$ unitary matrix. Then, we get

$$\begin{aligned}
\mathbf{U}_0 &= \begin{pmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \\ \vdots \\ \mathbf{B}_L \end{pmatrix}\Big(\mathbf{B}_1^2 + \mathbf{B}_2^2 + \cdots + \mathbf{B}_L^2\Big)^{-\frac{1}{2}}, \\
\mathbf{D}_0 &= \Big(\mathbf{B}_1^2 + \mathbf{B}_2^2 + \cdots + \mathbf{B}_L^2\Big)^{\frac{1}{2}}, \\
\mathbf{V}_0 &= \mathbf{\Phi}_0^H.
\end{aligned} \tag{8}$$

Define

$$\mathbf{P} \triangleq \mathbf{U}_0\mathbf{U}_0^H. \tag{9}$$

Obviously, $\mathbf{P}^2 = \mathbf{P}$, $\mathbf{P}^H = \mathbf{P}$ and $\mathrm{rank}(\mathbf{P}) = \mathrm{rank}(\mathbf{\Phi}) = K$. Thus, $\mathbf{P}$ is a projection matrix. It projects the KL-dimensional complex Euclidean vector space onto the subspace generated by columns of matrix $\mathbf{U}_0$, which is equal to the space generated by columns of matrix $\mathbf{B\Phi}$. Hence, $\mathbf{PB\Phi} = \mathbf{B\Phi}$.

Next, we reconsider (1).

$$\begin{aligned}
\mathbf{Y} &= \sqrt{\rho}\mathbf{HPBX} + \sqrt{\rho}\mathbf{H}(\mathbf{I}_{KL} - \mathbf{P})\mathbf{BX} + \mathbf{W} \\
&= \sqrt{\rho}\mathbf{H}\mathbf{U}_0\mathbf{U}_0^H\mathbf{BX} + \mathbf{W}_\tau \\
&= \sqrt{\rho}\mathbf{H}_\tau\mathbf{U}_0^H\mathbf{BX} + \mathbf{W}_\tau,
\end{aligned} \tag{10}$$

where

$$\mathbf{H}_\tau = \mathbf{H}\mathbf{U}_0, \quad \mathbf{W}_\tau = \sqrt{\rho}\mathbf{H}(\mathbf{I}_{KL} - \mathbf{P})\mathbf{BX} + \mathbf{W}. \tag{11}$$

By using (8), the above can be written as

$$\mathbf{Y} = \sqrt{\rho}\mathbf{H}_\tau \begin{pmatrix} \frac{\beta_{11}x_{11} + \beta_{12}x_{12} + \cdots + \beta_{1L}x_{1L}}{\sqrt{\beta_{11} + \beta_{12} + \cdots + \beta_{1L}}} \\ \frac{\beta_{21}x_{21} + \beta_{22}x_{22} + \cdots + \beta_{2L}x_{2L}}{\sqrt{\beta_{21} + \beta_{22} + \cdots + \beta_{2L}}} \\ \vdots \\ \frac{\beta_{K1}x_{K1} + \beta_{K2}x_{K2} + \cdots + \beta_{KL}x_{KL}}{\sqrt{\beta_{K1} + \beta_{K2} + \cdots + \beta_{KL}}} \end{pmatrix} + \mathbf{W}_\tau. \tag{12}$$

Equation (12) shows that the transmitted signals of all users, which can be viewed as vectors in the KL-dimensional complex space, is compacted into a K-dimensional space.

B. On the Size of Constellations Used by All Users

Assume that each user in the considered system uses the following signal constellation:

$$\mathrm{QAM}_{kl} = \{z \mid z = n_1 d + \mathbf{j} n_2 d, n_1, n_2 \in \mathcal{N}\} \tag{13}$$

where $d = \sqrt{3/(2(4N^2 - 1))}$ and

$\mathcal{N} = \{-(2N-1), -(2N-3), \ldots, -1, 1, \ldots, 2N-3, 2N-1\}$.

Thus, QAM_{kl} is the standard $4N^2$-QAM with total energy $4N^2$. Define

$$\gamma_k \triangleq \frac{\beta_{k1}}{2(\beta_{k2} + \beta_{k3} + \cdots + \beta_{kL})} + \frac{1}{2}. \tag{14}$$

for $k = 1, 2, \ldots, K$. Then, we have the following main results.

Theorem 1: Assume that the first BS decodes the signals with the MMSE decoder and each user in cells uses $4N^2$-QAM constellation given in (13). If $N < \gamma_k$, then

$$\lim_{M \to +\infty} \lim_{\rho_0 \to +\infty} P_{e,k} = 0, \tag{15}$$

and if $N > \gamma_k$, then

$$\lim_{M \to +\infty} \lim_{\rho_0 \to +\infty} P_{e,k} \geq \delta > 0, \tag{16}$$

where $P_{e,k}$ is the error probability of user k, and δ is a positive constant.

A sketch proof of Theorem 1 is given in the Appendix.

Theorem 1 shows that if the k-th user in each cell uses $4N^2$-QAM and $N < \gamma_k$, the error floor does not exist when both ρ_0 and M go to infinity. Otherwise, the error floor exists.

IV. Simulation Results

In this section, we show some simulation results to verify analysis presented in Section III. In all following simulations, we assume that the system has three cells, and three users in each cell. That is, $L = K = 3$. Let $\mathbf{\Phi} = (\mathbf{I}_3\ \mathbf{I}_3\ \mathbf{I}_3)^T$. Matrix $\mathbf{B}$ in (1) is selected as follows:

$$\mathbf{B} = \mathrm{diag}(0.9511,\ 0.7371,\ 0.6860,\ 0.6171,\ 0.3100,\ 0.1537,\ 0.1457,\ 0.1379,\ 0.0833).$$

Bases on (14), we have

$$\gamma_1 = 1.6250, \quad \gamma_2 = 2.8599, \quad \gamma_3 = 8.1989.$$

We perform the simulations of three cases. The first case is that $N = 1$, that is, the constellation is 4-QAM, and the SNR ρ is fixed to -10dB. The second case is that $N = 2$ (the constellation is 16-QAM), and $\rho = 4$dB. The last case is that $N = 4$ (the constellation is 64-QAM), and $\rho = 4$dB. In the three cases, antenna number M varies from 50 to 500, and $\rho_0 = 40$dB. The simulation results are shown in Fig. 1, Fig. 2 and Fig. 3, respectively.

In the first case, condition $N = 1 < \gamma_k$ is satisfied for $k = 1, 2, 3$. From Theorem 1, we deduce that the error floor does not exist for all three users. It is confirmed in Fig. 1. When $N = 2$, condition $N < \gamma_k$ is satisfied for user 2 and user 3. Thus, for these two users, the error floors do not exist, but it does exist for user 1. Fig. 2 shows these phenomena. Finally, when N is equal to 4, condition $N < \gamma_k$ is satisfied for user 3 only. Thus, the error floor does not exist for user 3 only. This is again confirmed by Fig. 3.

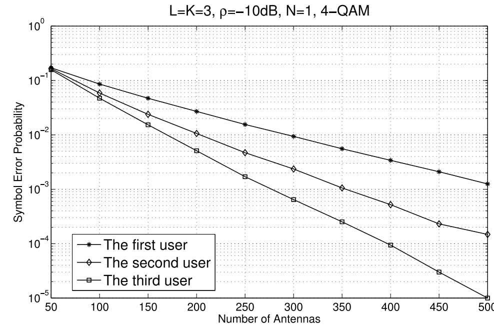

Fig. 1. Performances on large scale factors and antenna numbers (I).

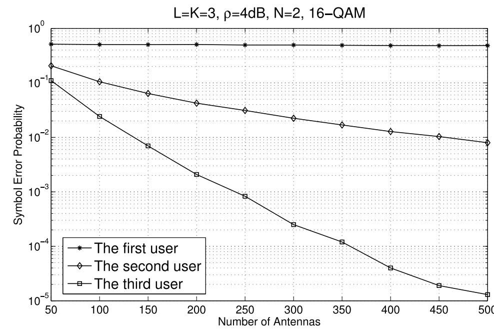

Fig. 2. Performances on large scale factors and antenna numbers (II).

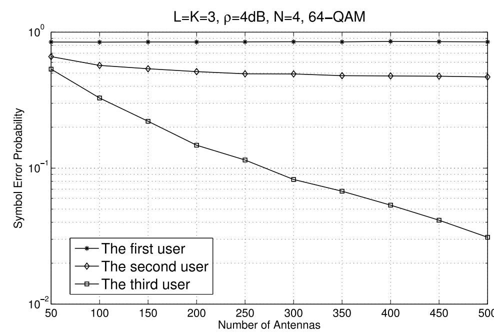

Fig. 3. Performances on large scale factors and antenna numbers (III).

V. Conclusion

In this letter, we studied the maximum number of signals that can be supported in a multi-cell massive MIMO system. By deriving an equivalent transmit-receive signal model, we found that if the size of the signal constellation is smaller than a parameter, which is related to the large-scale fading factors only, the error probability goes to zero. Otherwise, an error floor occurs regardless of the number of antennas and the pilot SNR.

Appendix

In this Appendix, we give a sketch of proof of Theorem 1. First, in the pilot training phase, from (10), we see that when $\mathbf{X} = \mathbf{\Phi}$ and $\rho_0 = \rho$, the received pilot signal $\mathbf{Y}_0$ can be written as

$$\mathbf{Y}_0 = \sqrt{\rho_0}\mathbf{H}_\tau \mathbf{D}_0 \Phi_0 + \mathbf{W}. \tag{17}$$

Since the rank of channel matrix $\mathbf{H}_\tau$ is equal to the rank of the pilot matrix, according to the results given in [13], we have

$$E_{\mathbf{H}_\tau}[\|\mathbf{H}_\tau - \hat{\mathbf{H}}_\tau\|] = M\sum_{k=1}^{K}\frac{1}{1+\rho_0\lambda_k}, \tag{18}$$

where $\hat{\mathbf{H}}_\tau$ is the MMSE estimation of $\mathbf{H}_\tau$, and $\lambda_k = \sum_{l=1}^{L}\beta_{kl}$. Thus, if ρ_0 goes to infinity, $E_{\mathbf{H}_\tau}[\| \mathbf{H}_\tau - \hat{\mathbf{H}}_\tau\|]$ goes to zero, and channel $\mathbf{H}_\tau$ can be perfectly estimated. Hence, in the following, we assume that $\mathbf{H}_\tau$ is perfectly known to the first BS when ρ_0 goes to infinity. Since

$$E\big(\mathbf{H}_\tau\mathbf{H}_\tau^H\big) = E\big(\mathbf{H}\mathbf{U}_0\mathbf{U}_0^H\mathbf{H}^H\big) = \tau\mathbf{I}_M, \tag{19}$$

and

$$E\big(\mathbf{H}_\tau^H\mathbf{H}_\tau\big) = E\big(\mathbf{U}_0^H\mathbf{H}^H\mathbf{H}\mathbf{U}_0\big) = M\mathbf{I}_\tau, \tag{20}$$

we know that components of $\mathbf{H}_\tau$ are i.i.d. Gaussian random variables with zero mean and unit variance. By the large number theorem, we have

$$\lim_{M\to+\infty}\frac{\mathbf{H}_\tau^H\mathbf{H}_\tau}{M} \overset{a.e.}{=} \mathbf{I}_\tau, \tag{21}$$

where $a.e.$ means that the equality holds with probability one.

In the signal transmission phase, (12) indicates that the considered system can be viewed as a "virtual system", where a BS with M receiving antennas receives signals sent from K different users with single antenna, and the k-th user uses the following constellation:

$$\mathcal{C}_k \triangleq \bigcup_{x\in\mathrm{QAM}_{k1}} \mathcal{C}_{k,x}, \tag{22}$$

where

$$\mathcal{C}_{k,x} \triangleq \left\{\frac{\beta_{k1}x + \beta_{k2}x_{k2} + \cdots + \beta_{kL}x_{kL}}{\sqrt{\beta_{k1}+\beta_{k2}+\cdots+\beta_{kL}}} \mid z_{kl} \in \mathrm{QAM}_{kl}\right\}. \tag{23}$$

Assume that $N < \gamma_k$. Then

$$\frac{\beta_{k1} - (2N-1)(\beta_{k2}+\cdots+\beta_{kL})}{\sqrt{\beta_{k1}+\beta_{k2}+\cdots+\beta_{kL}}} > 0. \tag{24}$$

Thus, set $Cov(\mathcal{C}_{k,x})$ and set $Cov(\mathcal{C}_{k,x'})$ are separated with the minimal distance at least $2\bar{\gamma}_k d$ if $x \neq x'$, where $Cov(\mathcal{A})$ denotes the convex set generated by points in set $\mathcal{A}$. Thus, according to results presented in [14], we have

$$\lim_{M\to+\infty}\lim_{\rho_0\to+\infty} P_{e,k} = 0.$$

Let us consider the converse part of the theorem. From (11), it is easy to get that the covariance of $\mathbf{W}_\tau$ is

$$\sigma_\tau^2 \triangleq 1 + \rho\mathrm{tr}(\mathbf{B}^H(\mathbf{I}_{KL} - \mathbf{P})\mathbf{B}). \tag{25}$$

Thus, the MMSE decoder, denoted as $\mathbf{G}_m$, is

$$\mathbf{G}_m = \left(\frac{\mathbf{I}_\tau}{\rho_\tau} + \mathbf{H}_\tau^H\mathbf{H}_\tau\right)^{-1}\mathbf{H}_\tau^H, \tag{26}$$

and the decoder is

$$\hat{x}_{k1} = \mathrm{argmin}_{x\in QAM_{k1}}\left\|\frac{\mathbf{G}_{m,k}\mathbf{Y}}{\sigma_\tau} - \frac{\sqrt{\rho_\tau}\beta_{k1}}{\sqrt{\beta_{k1}+\beta_{k2}+\cdots+\beta_{kL}}}x\right\|, \tag{27}$$

where $\mathbf{Y}$ is given in (12), and $\rho_\tau = \rho/\sigma_\tau^2$.

When M goes to infinity, using (21), we have

$$\lim_{M\to+\infty}\mathbf{G}_m\mathbf{H}_\tau \overset{a.e.}{=} \mathbf{I}_\tau. \tag{28}$$

Consequently, we get

$$\hat{y}_k = \sqrt{\rho_\tau}\frac{\beta_{k1}x_{k1} + \beta_{k2}x_{k2} + \cdots + \beta_{kL}x_{kL}}{\sqrt{\beta_{k1}+\beta_{k2}+\cdots+\beta_{kL}}} + \hat{w}_k, \tag{29}$$

where $\hat{y}_k = \frac{\mathbf{G}_{m,k}\mathbf{Y}}{\sigma_\tau}$, and $\hat{w_k}$ is the corresponding noise. Hence, the decoder (27) is reduced into

$$\hat{x}_{k1} = \mathrm{argmin}_{x\in QAM_{k1}}\left\|\hat{y}_k - \frac{\sqrt{\rho_\tau}\beta_{k1}x}{\sqrt{\beta_{k1}+\beta_{k2}+\cdots+\beta_{kL}}}\right\|. \tag{30}$$

Assume that $N > \gamma_k$. Then, (24) does not hold, and $Cov(\mathcal{C}_{k,x}) \cap Cov(\mathcal{C}_{k,x'}) \neq \emptyset$ if x and x' are neighboring points in QAM_{k1}. Thus, when the k-th user in the first cell transmits signal $x \in \mathrm{QAM}_{k1}$, there always exist linear combinations

$$\sqrt{\rho_\tau}\frac{\beta_{k1}x + \beta_{k2}x_{k2} + \cdots + \beta_{kL}x_{kL}}{\sqrt{\beta_{k1}+\beta_{k2}+\cdots+\beta_{kL}}}$$

with $x_{k2} \in \mathrm{QAM}_{k2}, \ldots, x_{kL} \in \mathrm{QAM}_{kL}$, such that these points lie outside the decision region of x. Thus, decoding errors occur. Moreover, these error events are independent of M and ρ_0, we have proved the Theorem 1.

References

[1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[2] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[3] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[4] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.

[5] R. R. Müller, L. Cottatellucci, and M. Vehkaperä, "Blind pilot decontamination," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 773–786, Oct. 2014.

[6] H. Q. Ngo and E. G. Larsson, "EVD-based channel estimation in multicell multiuser MIMO systems with very large antenna arrays," in *Proc. ICASSP*, Kyoto, Japan, 2012, pp. 3249–3252.

[7] M. Teeti, J. Sun, D. Gesbert, and Y. Liu, "The impact of physical channel on performance of subspace-based channel estimation in massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 9, pp. 4743–4756, Sep. 2015.

[8] T. E. Bogale and L. B. Le, "Pilot optimization and channel estimation for multiuser massive MIMO systems," in *Proc. 48th Annu. Conf. Inf. Sci. Syst. (CISS)*, Princeton, NJ, USA, Mar. 2014, pp. 1–6.

[9] S. Noh, M. D. Zoltowski, Y. Sung, and D. J. Love, "Pilot beam pattern design for channel estimation in massive MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 787–801, Oct. 2014.

[10] J. Ma and L. Ping, "Data-aided channel estimation in large antenna systems," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3111–3124, Jun. 2014.

[11] E. Björnson, E. G. Larsson, and M. Debbah, "Massive MIMO for maximal spectral efficiency: How many users and pilots should be allocated?" *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1293–1308, Feb. 2016.

[12] J. Choi and L. Bai, "MMSE-based random sampling for iterative detection for large-scale MIMO systems," *J. Commun. Inf. Netw.*, vol. 1, no. 2, pp. 29–36, Aug. 2016.

[13] H. Wang, W. Zhang, Y. Liu, Q. Xu, and P. Pan, "On design of nonorthogonal pilot signals for a multi-cell massive MIMO system," *IEEE Wireless Commun. Lett.*, vol. 4, no. 2, pp. 129–132, Apr. 2015.

[14] H. Wang, M. Zhou, R. Chen, and W. Zhang, "On error rate performance of multi-cell massive MIMO systems with linear receivers," *Phys. Commun.*, vol. 20, pp. 123–132, Sep. 2016.