

Uplink Sum-Rate Analysis of Massive MIMO System with Pilot Contamination and CSI Delay

Juan Cao · Dongming Wang · Jiamin Li ·
Shaohui Sun · Xiaohu You

Published online: 17 April 2014
© Springer Science+Business Media New York 2014

Abstract We consider multi-cell multi-user massive MIMO system under correlated Rayleigh fading channels. Taking pilot contamination and CSI delay into consideration, we derive the equivalent channel model with MMSE channel estimation and one-tap prediction. Employing this equivalent channel model, the lower bound of the uplink sum-rate is derived, and its asymptotical performance is studied when the base station antenna number goes without bound. We find that if we schedule the k -th user of all cells who have the same prediction coefficient, the uplink sum-rate is the same as the one with no CSI delay when the number of BS antennas goes without bound at a much greater rate than the number of users. Simulation results show that the asymptotic approximation has good performance for large M , and suggest that large antenna array can compensate for the decay due to CSI delay. Simulation results also verify our guess that CSI delay does not necessarily decrease the uplink sum-rate due to the impact of pilot contamination.

Keywords Massive MIMO · Pilot contamination · CSI delay · Uplink sum-rate

1 Introduction

Massive MIMO or large-scale MIMO systems are currently investigated with the remarkable potential of increasing the spectral efficiency and power efficiency even with very simple linear transmitter/receiver [1–3]. To avoid the per-antenna pilot training of the downlink, TDD (Time Division Multiplexing) is adopted and the downlink CSI can be obtained by

J. Cao (✉) · D. Wang · J. Li · X. You
National Mobile Communications Research Lab, Southeast University, Nanjing, China
e-mail: caojuan@seu.edu.cn

S. Sun
State Key Laboratory of Wireless Mobile Communications (CATT), Beijing, China

J. Cao
School of Electronics and Information, Nantong University, Nantong, China

exploiting channel reciprocity; Thus the overhead related to channel training scales linearly with the number of mobile users per cell (K) and is independent of the number of antennas per base station (M) [4]. In the multi-cell scenario, non-orthogonal training sequences must be used due to the limitation of the pilot resources. This non-orthogonal nature causes pilot contamination when the channel estimate at target cell becomes polluted by users from other cells. There're many papers studying the impact of pilot contamination; for example, the performances of the spectral efficiency for massive MIMO with the pilot contamination have been studied in [1, 5, 6] for both uplink and downlink transmission under the i.i.d Rayleigh fading channels; and [7–9] study the performances of the spectral efficiency under correlated channel.

The papers mentioned above all assume the channel state information (CSI) stay constant during the phase of pilot transmission and uplink data transmission. However, in practical TDD systems, due to increased user mobility, transmission delay and processing delay, the delayed CSI will be outdated. So the assumption of a locally time-invariant (block-fading) channel breaks down, and an effective approach is to predict the channel values at times when they will be used [10, 11]. Then how will the predicted CSI affect the spectral efficiency of practical massive MIMO system?

In this work, we study the performances of the uplink sum-rate of the massive MIMO under correlated channel taking both pilot contamination and CSI delay into consideration. Using correlation between the actual channel and the estimated one, and the channel's time-correlation, we derive an equivalent channel model with MMSE channel estimation and one-tap prediction. Employing this equivalent channel model, we then obtain the lower bound of the uplink sum-rate, and study its asymptotical performance when the base station antenna number goes without bound. We find that if we schedule the k -th user of all cells who have the same prediction coefficient, the uplink sum-rate is the same as the one with no CSI delay when the number of BS antennas goes without bound at a much greater rate than the number of users.

The notation adopted in this paper conforms to the following convention. Matrices are represented with uppercase boldface and vectors meaning column vector with lowercase boldface. $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ represent conjugate, transpose and Hermitian transpose respectively. $|\cdot|$ denotes the module operation and $\|\cdot\|$ denotes the spectral norm. $Tr(\mathbf{A})$ is the trace of \mathbf{A} and $\det(\mathbf{A})$ denotes the determinant of \mathbf{A} . $\text{Diag}(\mathbf{x})$ is a diagonal matrix with x on its diagonal. I_M denotes a $M \times M$ identity matrix. The operator $\mathcal{E}(\cdot)$ denotes expectation, and the covariance operator is given by $\text{cov}(\mathbf{x}, \mathbf{y}) \triangleq \mathcal{E}(\mathbf{x}\mathbf{y}^H) - \mathcal{E}(\mathbf{x})\mathcal{E}(\mathbf{y}^H)$. $\mathcal{CN}(\mathbf{0}, \mathbf{\Sigma})$ denotes complex Gaussian distribution with mean 0 and covariance matrix $\mathbf{\Sigma}$.

2 System Model of the Uplink Multi-cell Multi-user Massive MIMO

We consider the L cells system with one BS and K mobile users in each cell. Each BS is equipped with M antennas, and each user has single antenna. We assume that the system is operating on TDD protocol with full frequency reuse. Because of TDD operation and reciprocity the propagation is the same for either a downlink or an uplink transmission. Then at BS of cell l , the uplink received base-band signal vector at time t reads

$$\mathbf{y}_l(t) = \mathbf{G}_{l,l}(t)\mathbf{x}_l(t) + \sum_{i \neq l}^L \mathbf{G}_{l,i}(t)\mathbf{x}_i(t) + \mathbf{w}_l(t) \quad (1)$$

where $\mathbf{y}_l(t) = [y_{l,1}(t) \cdots y_{l,M}(t)]^T$ is the received signal vector at time t , and $\mathbf{x}_l(t) = [x_{l,1}(t) \cdots x_{l,K}(t)]^T \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_K)$ is the overall transmitted signal vector of the l -th cell is, $x_{l,k}(t)$ the transmit signal of user k in cell l , $\mathbf{w}_l(t) \sim \mathcal{CN}(\mathbf{0}, \gamma_{UL} \mathbf{I}_M)$ is the complex additive noise vector. $\mathbf{G}_{l,i}(t) = [\mathbf{g}_{l,i,1}(t) \cdots \mathbf{g}_{l,i,K}(t)]$ is the uplink channel matrix at time index t between all of the K users of cell i to the BS of cell l .

We assume all of the K users of each cell have the same antenna correlation, and model the channel vector $\mathbf{g}_{l,i,k}(t)$ as

$$\mathbf{g}_{l,i,k}(t) \triangleq \sqrt{cd_{l,i,k}^{-\alpha} s_{l,i,k}} \mathbf{R}_{l,i}^{\frac{1}{2}} \mathbf{h}_{l,i,k}(t) \tag{2}$$

where α is the path loss exponent, typically between 3.0 and 5.0, c is the median of the mean path gain at a reference distance $d_{l,i,k} = 1\text{km}$, and $s_{l,i,k}$ is a log-normal shadow fading variable. $\mathbf{h}_{l,i,k}(t) \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$ represents the small-scale fast fading. $\mathbf{R}_{l,i}$ is the deterministic receive correlation matrix of the K users of cell i to the BS of cell l .

Then the composite channel $\mathbf{G}_{l,i}(t)$ can be expressed as

$$\begin{aligned} \mathbf{G}_{l,i}(t) &= \mathbf{R}_{l,i}^{\frac{1}{2}} [\mathbf{h}_{l,i,1}(t) \cdots \mathbf{h}_{l,i,K}(t)] \mathbf{A}_{l,i}^{\frac{1}{2}} \\ &= \mathbf{R}_{l,i}^{\frac{1}{2}} \mathbf{H}_{l,i}(t) \mathbf{A}_{l,i}^{\frac{1}{2}} \end{aligned} \tag{3}$$

where $\lambda_{l,i,k} \triangleq cd_{l,i,k}^{-\alpha} s_{l,i,k}$, $\mathbf{A}_{l,i} = \text{diag} \{ [\lambda_{l,i,1} \cdots \lambda_{l,i,K}] \}$.

Throughout the paper we adopt the following assumptions:

A1: $h_{l,i,m,k}(t)$ is the small-scale fast fading between k -th user of cell i to the m -th antenna of BS l , which is slowly time-varying according to Jakes model with Doppler spread f_d , and $\rho_{l,i,k}(\tau) \triangleq E \{ h_{l,i,m,k}^*(t) h_{l,i,m,k}(t + \tau) \} = J_0(2\pi f_d |\tau| T_s)$. T_s is the symbol period, and $J_0(\cdot)$ is the zeroth order Bessel function of the first kind.

A2: $\limsup_M \|\mathbf{R}_{l,i}\| < \infty$ for all l and i .

A3: $\liminf_M \frac{1}{M} \text{Tr}(\mathbf{R}_{l,i}) > 0$ for all l .

3 Uplink Sum-Rate Analysis of Massive MIMO with Pilot Contamination and CSI Delay

In this section, the channel estimation of massive MIMO system under correlated channel using minimum mean-square-error (MMSE) is reviewed first, and then prediction filter with one tap is used to predict the delayed channel. We derive the equivalent channel model which allows us to give further analysis of the achievable uplink sum-rate for a finite and an infinite number of BS antennas, and obtain some interesting and insightful observations. We also analyze these performances for i.i.d channel model.

3.1 Channel Estimation and Prediction

When TDD is adopted, the minimum required number of pilot symbols is equal to the number of mobile users per cell K and is independent of the number of antennas per base station M [4]. So without loss a generality, here we assume that the pilot matrix is an identity matrix of size K . Channel estimation is performed by sending pilot sequence of length K to BSs, and we assume that a total of L base stations share the same band of frequencies and the same set of K orthogonal pilot signals. Furthermore we assume synchronized transmissions and reception which constitutes a worst-case scenario from the standpoint of pilot contamination

[1]. There is no cooperation between the BSs. For slow varying channel, we assume that the channel is invariable during the phase of pilot transmission. For simplicity of writing, from here on we omit the time index (t) of channel matrix, and use subscript P to distinguish the phase of pilot transmission from the phase of data transmission. Then, from (1), the received pilot signal at BS of cell l can be given by

$$Y_{P,l} = G_{l,l}X_{P,l} + \sum_{i \neq l}^L G_{l,i}X_{P,i} + W_{P,l} \tag{4}$$

where $Y_{P,l}$ is an $M \times K$ received pilot signal matrix, here the dimension K refers to the length of pilot sequence. $X_{P,i} = I_K$ is a $K \times K$ pilot matrix. $W_{P,l}$ is an $M \times K$ noise matrix and each element is i.i.d ZMCSCG random variable with variance γ_P . Since $\mathcal{E}(\mathbf{g}_{l,i,k} \mathbf{g}_{l,i',j}^H) = 0$, for $j \neq k$, the estimation of $\mathbf{g}_{l,i,k}$ can be processed individually. We take k -th user for example,

$$y_{P,l,k} = g_{l,l,k} + \sum_{i \neq l} g_{l,i,k} + w_{P,l,k} \tag{5}$$

where $y_{P,l,k}$ and $w_{P,l,k}$ are the k -th column of $Y_{P,l}$ and $W_{P,l}$, respectively. Given (5), the MMSE estimation of channel vector $\mathbf{g}_{l,i,k}$ can be obtained by

$$\hat{\mathbf{g}}_{l,i,k} = \lambda_{l,i,k} \mathbf{R}_{l,i} \mathbf{Q}_{l,k} y_{P,l,k} \tag{6}$$

where

$$\mathbf{Q}_{l,k} = \left(\sum_{i=1}^L \lambda_{l,i,k} \mathbf{R}_{l,i} + \gamma_P \mathbf{I}_M \right)^{-1} \tag{7}$$

Invoking the orthogonality property of the MMSE estimate [12], the estimation error $\tilde{\mathbf{g}}_{l,i,k}$ is statistically independent of $\mathbf{g}_{l,i,k}$ due to the joint Gaussianity of both vectors. So we can decompose the channel as $\mathbf{g}_{l,i,k} = \tilde{\mathbf{g}}_{l,i,k} + \hat{\mathbf{g}}_{l,i,k}$. And the covariance matrix of the estimation error can be expressed as

$$\text{cov}(\tilde{\mathbf{g}}_{l,i,k}, \tilde{\mathbf{g}}_{l,i,k}) = \lambda_{l,i,k} \mathbf{R}_{l,i} - \lambda_{l,i,k}^2 \mathbf{R}_{l,i} \mathbf{Q}_{l,k} \mathbf{R}_{l,i} \tag{8}$$

See (6), we find that given k , the channel estimation $\hat{\mathbf{g}}_{l,i,k}$ are correlated for each i due to pilot contamination. As in [9], we define $\hat{\mathbf{h}}_{l,k} \triangleq \mathbf{Q}_{l,k} y_{P,l,k}$, which obeys $\hat{\mathbf{h}}_{l,k} \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q}_{l,k})$. So the estimated channel vector can be expressed by

$$\hat{\mathbf{g}}_{l,i,k} = \lambda_{l,i,k} \mathbf{R}_{l,i} \hat{\mathbf{h}}_{l,k} \tag{9}$$

where $\hat{\mathbf{h}}_{l,k}$ represents the Rayleigh fading part of the estimated channel.

Due to increased user mobility, transmission delay and processing delay, the CSI of the phase of data transmission will be different from that of the pilot transmission, which we call CSI delay. The time-varying characteristics of the small-scale fading cause estimation error also. We improve the initial channel estimation taking CSI delay into consideration.

Suppose that we use a prediction filter with one tap exploiting the time-domain correlation present according to A1 [10,11].

$$\begin{aligned} [\mathbf{h}_{l,i,1}^\tau \cdots \mathbf{h}_{l,i,K}^\tau] &= [\mathbf{h}_{l,i,1} \cdots \mathbf{h}_{l,i,K}] \text{diag} \{ [\rho_{l,i,1}^\tau \cdots \rho_{l,i,K}^\tau] \} \\ &+ [\mathbf{e}_{l,i,1}^\tau \cdots \mathbf{e}_{l,i,K}^\tau] \text{diag} \left\{ \left[\sqrt{1 - |\rho_{l,i,1}^\tau|^2} \cdots \sqrt{1 - |\rho_{l,i,K}^\tau|^2} \right] \right\} \end{aligned} \tag{10}$$

As mentioned before, for simplicity of writing, we use $\mathbf{h}_{l,i,k}^\tau$ for the channel delayed (the superscript τ indicates the CSI delay between channel estimation and uplink transmission). $\rho_{l,i,k}^\tau$ is the correlation coefficient between the delayed channel and current one, and it is decided according to the assumption of A1. $\mathbf{e}_{l,i,k}^\tau$ is i.i.d zero mean circularly symmetric complex Gaussian random variables of variance 1. And the second part of (10) stands for prediction error.

Substituting (10) into (3), we obtain

$$\begin{aligned} \mathbf{G}_{l,i}^\tau &= \mathbf{R}_{l,i}^{\frac{1}{2}} [\mathbf{h}_{l,i,1} \cdots \mathbf{h}_{l,i,K}] \text{diag} \{ [\rho_{l,i,1}^\tau \cdots \rho_{l,i,K}^\tau] \} \Lambda_{l,i}^{\frac{1}{2}} \\ &\quad + \mathbf{R}_{l,i}^{\frac{1}{2}} [\mathbf{e}_{l,i,1}^\tau \cdots \mathbf{e}_{l,i,K}^\tau] \text{diag} \left\{ \left[\sqrt{1 - |\rho_{l,i,1}^\tau|^2} \cdots \sqrt{1 - |\rho_{l,i,K}^\tau|^2} \right] \right\} \Lambda_{l,i}^{\frac{1}{2}} \\ &= [\mathbf{g}_{l,i,1} \cdots \mathbf{g}_{l,i,K}] \text{diag} \{ [\rho_{l,i,1}^\tau \cdots \rho_{l,i,K}^\tau] \} + \mathbf{Z}_{l,i}^\tau \end{aligned} \tag{11}$$

where

$$\mathbf{Z}_{l,i}^\tau \triangleq \mathbf{R}_{l,i}^{\frac{1}{2}} [\mathbf{e}_{l,i,1}^\tau \cdots \mathbf{e}_{l,i,K}^\tau] \text{diag} \left\{ \left[\sqrt{1 - |\rho_{l,i,1}^\tau|^2} \cdots \sqrt{1 - |\rho_{l,i,K}^\tau|^2} \right] \right\} \Lambda_{l,i}^{\frac{1}{2}}$$

is the matrix of prediction error with covariance matrix $\mathcal{E} \left[\mathbf{Z}_{l,i}^\tau (\mathbf{Z}_{l,i}^\tau)^H \right] = \left(\sum_{k=1}^K \lambda_{l,i,k} \left(1 - |\rho_{l,i,k}^\tau|^2 \right) \right) \mathbf{R}_{l,i}$.

According to $\mathbf{g}_{l,i,k} = \tilde{\mathbf{g}}_{l,i,k} + \hat{\mathbf{g}}_{l,i,k}$, we get

$$\begin{aligned} \mathbf{G}_{l,i}^\tau &= [\hat{\mathbf{g}}_{l,i,1} \cdots \hat{\mathbf{g}}_{l,i,K}] \text{diag} \{ [\rho_{l,i,1}^\tau \cdots \rho_{l,i,K}^\tau] \} \\ &\quad + [\tilde{\mathbf{g}}_{l,i,1} \cdots \tilde{\mathbf{g}}_{l,i,K}] \text{diag} \{ [\rho_{l,i,1}^\tau \cdots \rho_{l,i,K}^\tau] \} + \mathbf{Z}_{l,i}^\tau \end{aligned} \tag{12}$$

Substituting (9) into (12), we obtain

$$\begin{aligned} \mathbf{G}_{l,i}^\tau &= \mathbf{R}_{l,i} [\hat{\mathbf{h}}_{l,1} \cdots \hat{\mathbf{h}}_{l,K}] \text{diag} \{ [\lambda_{l,i,1}^\tau \cdots \lambda_{l,i,K}^\tau] \} \\ &\quad + [\tilde{\mathbf{g}}_{l,i,1} \cdots \tilde{\mathbf{g}}_{l,i,K}] \text{diag} \{ [\rho_{l,i,1}^\tau \cdots \rho_{l,i,K}^\tau] \} + \mathbf{Z}_{l,i}^\tau \end{aligned} \tag{13}$$

where $\lambda_{l,i,k}^\tau \triangleq \lambda_{l,i,k} \rho_{l,i,k}^\tau$.

We define the following matrices

$$\begin{aligned} \mathbf{A}_{l,i}^\tau &\triangleq \text{diag} \{ [\lambda_{l,i,1}^\tau \cdots \lambda_{l,i,K}^\tau] \} \\ \hat{\mathbf{H}}_l &\triangleq [\hat{\mathbf{h}}_{l,1} \cdots \hat{\mathbf{h}}_{l,K}] \\ \hat{\mathbf{G}}_{l,i}^\tau &\triangleq [\hat{\mathbf{g}}_{l,i,1}^\tau \cdots \hat{\mathbf{g}}_{l,i,K}^\tau] \\ &= \mathbf{R}_{l,i} [\hat{\mathbf{h}}_{l,1} \cdots \hat{\mathbf{h}}_{l,K}] \text{diag} \{ [\lambda_{l,i,1}^\tau \cdots \lambda_{l,i,K}^\tau] \} \\ \tilde{\mathbf{G}}_{l,i}^\tau &\triangleq [\tilde{\mathbf{g}}_{l,i,1}^\tau \cdots \tilde{\mathbf{g}}_{l,i,K}^\tau] \\ &= [\tilde{\mathbf{g}}_{l,i,1} \cdots \tilde{\mathbf{g}}_{l,i,K}] \text{diag} \{ [\rho_{l,i,1}^\tau \cdots \rho_{l,i,K}^\tau] \} \end{aligned}$$

The predicted channel estimation can be expressed as

$$\hat{\mathbf{G}}_{l,i}^\tau = \mathbf{R}_{l,i} \hat{\mathbf{H}}_l \mathbf{A}_{l,i}^\tau \tag{14}$$

and the actual channel matrix can be expressed as

$$\mathbf{G}_{l,i}^\tau = \hat{\mathbf{G}}_{l,i}^\tau + \tilde{\mathbf{G}}_{l,i}^\tau + \mathbf{Z}_{l,i}^\tau \tag{15}$$

3.2 Lower Bound of the Sum-Rate and Its Asymptotical Performance Analysis for Large M

Consider the uplink transmission represented by (1), and according to (15), the overall uplink received signal at the BS of cell l can be written as

$$\mathbf{y}_l = \hat{\mathbf{G}}_{l,l}^\tau \mathbf{x}_l + \sum_{i \neq l}^L \hat{\mathbf{G}}_{l,i}^\tau \mathbf{x}_i + \sum_{i=1}^L \tilde{\mathbf{G}}_{l,i}^\tau \mathbf{x}_i + \sum_{i=1}^L \mathbf{Z}_{l,i}^\tau \mathbf{x}_i + \mathbf{w}_l^{\text{UL}} \tag{16}$$

Then the sum rate of the cell l in bits per second per channel use (bps/channel) (omitting the rate loss due to uplink channel training) is defined as

$$\hat{C}^\tau \triangleq I(\mathbf{x}_l, \mathbf{y}_l | \hat{\mathbf{G}}_{l,1}^\tau, \dots, \hat{\mathbf{G}}_{l,L}^\tau) \tag{17}$$

Theorem 1 For linear multi-user MMSE detection, while taking channel estimation error and CSI delay into consideration, the lower bound of (17) can be given by

$$\begin{aligned} \hat{C}^\tau \geq \hat{C}_{LB}^\tau &= \log_2 \det \left(\sum_{i=1}^L \hat{\mathbf{G}}_{l,i}^\tau (\hat{\mathbf{G}}_{l,i}^\tau)^H (\boldsymbol{\Sigma}_l^\tau)^{-1} + \mathbf{I}_M \right) \\ &\quad - \log_2 \det \left(\sum_{i \neq l}^L \hat{\mathbf{G}}_{l,i}^\tau (\hat{\mathbf{G}}_{l,i}^\tau)^H (\boldsymbol{\Sigma}_l^\tau)^{-1} + \mathbf{I}_M \right) \end{aligned} \tag{18}$$

where $\boldsymbol{\Sigma}_l^\tau$ is the covariance matrix of noise and interference caused by prediction and estimation error which is computed as

$$\begin{aligned} \boldsymbol{\Sigma}_l^\tau &= \text{cov} \left(\sum_{i=1}^L \tilde{\mathbf{G}}_{l,i}^\tau \mathbf{x}_i + \sum_{i=1}^L \mathbf{Z}_{l,i}^\tau \mathbf{x}_i + \mathbf{w}_l^{\text{UL}}, \sum_{i=1}^L \tilde{\mathbf{G}}_{l,i}^\tau \mathbf{x}_i + \sum_{i=1}^L \mathbf{Z}_{l,i}^\tau \mathbf{x}_i + \mathbf{w}_l^{\text{UL}} \right) \\ &= \sum_{i=1}^L \left(\sum_{k=1}^K \lambda_{l,i,k} \mathbf{R}_{l,i} - \mathbf{R}_{l,i} \left(\sum_{k=1}^K \lambda_{l,i,k}^2 |\rho_{l,i,k}^\tau|^2 \mathbf{Q}_{l,k} \right) \mathbf{R}_{l,i} \right) + \gamma_{\text{UL}} \mathbf{I}_M \end{aligned} \tag{19}$$

Proof Because estimation error, prediction error and the additive noise are all uncorrelated with the signal, then according to the Worst Case Uncorrelated Additive Noise Theorem in [13],

$$\hat{C}_{LB}^\tau = \log_2 \det \left(\left(\sum_{i \neq l}^L \hat{\mathbf{G}}_{l,i}^\tau (\hat{\mathbf{G}}_{l,i}^\tau)^H + \boldsymbol{\Sigma}_l^\tau \right)^{-1} \hat{\mathbf{G}}_{l,l}^\tau (\hat{\mathbf{G}}_{l,l}^\tau)^H + \mathbf{I}_K \right) \tag{20}$$

Multiplying out $\left(\sum_{i \neq l}^L \hat{\mathbf{G}}_{l,i}^\tau (\hat{\mathbf{G}}_{l,i}^\tau)^H + \boldsymbol{\Sigma}_l^\tau \right)^{-1}$, we obtain (18).

Observe the similarity between our result and theorem 1 in [9]. And using the similar manipulation, we can deduce theorem 2 invoking the strong law of large number.

Theorem 2 As $M \rightarrow \infty$, \hat{C}_{LB}^τ obeys $\hat{C}_{LB}^\tau - \hat{C}_{LB,\text{inf}}^\tau \xrightarrow{M \rightarrow \infty} 0$, where

$$\begin{aligned} \hat{C}_{LB,\text{inf}}^\tau &= \sum_{k=1}^K \log_2 \left[\frac{\det(\mathbf{\Xi}_{l,k} + \mathbf{I}_L)}{\det(\mathbf{\Xi}'_{l,k} + \mathbf{I}_{L-1})} \right] \\ &= \sum_{k=1}^K \log_2 \left[\xi_{l,l,k} + 1 - [\xi_{l,2,k} \cdots \xi_{l,L,k}] (\mathbf{\Xi}'_{l,k} + \mathbf{I}_{L-1})^{-1} [\xi_{2,l,k} \cdots \xi_{L,l,k}]^T \right] \end{aligned} \tag{21}$$

where

$$\begin{aligned} \xi_{i,i',k} &= |\lambda_{l,i,k}^\tau|^2 \text{Tr} \left(\mathbf{Q}_{l,k} \mathbf{R}_{l,i} (\mathbf{\Sigma}_l^\tau)^{-1} \mathbf{R}_{l,i'} \right) \\ \mathbf{\Xi}_{l,k} &= \begin{bmatrix} \xi_{1,1,k} & \cdots & \xi_{1,L,k} \\ \vdots & \dots & \vdots \\ \xi_{L,1,k} & \cdots & \xi_{L,L,k} \end{bmatrix} \end{aligned}$$

and $\mathbf{\Xi}'_{l,k}$ is the submatrix of $\mathbf{\Xi}_{l,k}$ formed by deleting the l -th row and l -th column.

Proof According to Theorem 2 in [9], and using the property of the determinant of the block matrix, one can obtain the result.

If $\mathbf{R}_{l,i} = \mathbf{R}_l, \forall i$, $\hat{C}_{LB,\text{inf}}^\tau$ can be simplified as

$$\hat{C}_{LB,\text{inf}}^\tau = \sum_{k=1}^K \left(1 + \frac{|\lambda_{l,l,k} \rho_{l,l,k}^\tau|^2}{\sum_{i \neq l}^L |\lambda_{l,i,k} \rho_{l,i,k}^\tau|^2 + [\text{Tr}(\mathbf{Q}_{l,k} \mathbf{R}_l (\mathbf{\Sigma}_l^\tau)^{-1} \mathbf{R}_l)]^{-1}} \right) \tag{22}$$

The receive correlation matrix can be decomposed as $\mathbf{R}_l = \mathbf{U} \mathbf{V} \mathbf{U}^H$ where \mathbf{U} is a unitary matrix, then $\mathbf{Q}_{l,k}, (\mathbf{\Sigma}_l^\tau)^{-1}$ can be written in the similar way as $\mathbf{Q}_{l,k} = \mathbf{U} \mathbf{V}_{l,k}^Q \mathbf{U}^H, (\mathbf{\Sigma}_l^\tau)^{-1} = \mathbf{U} \mathbf{V}_\Sigma \mathbf{U}^H$, and with the assumption of A2 and A3, when $M/K \rightarrow \infty$, $[\text{Tr}(\mathbf{Q}_{l,k} \mathbf{R}_l (\mathbf{\Sigma}_l^\tau)^{-1} \mathbf{R}_l)]^{-1} \rightarrow 0$. We get following corollary.

Corollary 1 Let $\mathbf{R}_{l,i} = \mathbf{R}_l, \forall i$, when $M \rightarrow \infty$, and assume $M/K \rightarrow \infty$, (22) can be given by

$$\hat{C}_{LB,\text{inf}}^\tau = \sum_{k=1}^K \log_2 \left(1 + \frac{|\lambda_{l,l,k} \rho_{l,l,k}^\tau|^2}{\sum_{i \neq l}^L |\lambda_{l,i,k} \rho_{l,i,k}^\tau|^2} \right) \tag{23}$$

If there is no CSI delay, that is $\rho_{l,i,k}^\tau = 1, \forall i, k$, (23) degrades to

$$\hat{C}_{LB,\text{inf}}^\tau = \sum_{k=1}^K \log_2 \left(1 + \frac{(\lambda_{l,l,k})^2}{\sum_{i \neq l}^L (\lambda_{l,i,k})^2} \right) \tag{24}$$

which coincides with the conclusions derived by [1].

Remark Theorem 2 shows that for large M the approximation of \hat{C}_{LB}^τ can be given by $\hat{C}_{LB,\text{inf}}^\tau$. And setting $\rho_{l,i,k}^\tau = 1, \forall i, k$, (18) and (21) deduce to the conclusions with no CSI delay, which coincide with Theorem 1 and 2 of [9]. Corollary 1 tells us that as $M \rightarrow \infty$,

and $M/K \rightarrow \infty$, the effects of uncorrelated receiver noise, prediction error, estimation error and fast fading are eliminated completely. And transmissions from users within ones own cell do not interfere. However, transmission from users in other cells that use the same pilot sequence constitutes a residual interference called pilot contamination. And due to CSI delay, the direct gains and the cross gains are both propotional to the square of their respective prediction coefficients. It also tells us that as $M \rightarrow \infty$, and $M/K \rightarrow \infty$, the sum-rate will be up-bounded by the same infinite rate which is independent of the receive correlation matrix.

More interestingly, as $M \rightarrow \infty$, and $M/K \rightarrow \infty$, if we schedule the k -th user of target cell who move at lower speed than other $L - 1$ k -th users, his uplink rate will be greater than the one when all the users are immobile, and vice versa. It can be proven that the uplink rate of some of the k -th users will increase while others will decrease if all the k -th users move at different velocities. But if we schedule the k -th user of all cells with same mobile velocity, all of their uplink sum-rate will not degrade in spite of their mobility. Such observation is very useful for mobile communications. And the reason for such interesting observation is that the existence of pilot contamination makes CSI delay not only degrade the signal power, but also the interference power.

If $\rho_{l,k}^\tau \triangleq \rho_{l,i,k}^\tau \neq 0, \forall i$, we define

$$R_\infty = \sum_{k=1}^K \log_2 \left(1 + \frac{(\lambda_{l,i,k})^2}{\sum_{i \neq l}^L (\lambda_{l,i,k})^2} \right) \tag{25}$$

3.3 Sum-Rate Analysis for i.i.d Channel Model

In order to deduce some insightful conclusions, we now consider the i.i.d Rayleigh fading channel model

$$\mathbf{G}_{l,i} = \mathbf{H}_{l,i} \mathbf{A}_{l,i}^{\frac{1}{2}} \tag{26}$$

The covariance matrix of the estimated channel vector $\hat{\mathbf{g}}_{l,i,k}$ can be simplified by

$$\begin{aligned} \text{cov}(\hat{\mathbf{g}}_{l,i,k}, \hat{\mathbf{g}}_{l,i,k}) &= \lambda_{l,i,k}^2 \mathbf{Q}_{l,k} \\ &= \lambda_{l,i,k}^2 \left[\left(\sum_{i=1}^L \lambda_{l,i,k} \right) + \gamma_P \right]^{-1} \mathbf{I}_M \end{aligned} \tag{27}$$

So in this case, the equivalent predicted channel estimation model (14) can be rewritten as

$$\hat{\mathbf{G}}_{l,i}^\tau = \mathbf{H} \mathbf{A}_{l,i} \left(\sum_{i=1}^L \mathbf{A}_{l,i} + \gamma_P \mathbf{I}_K \right)^{-\frac{1}{2}} \text{diag} \{ [\rho_{l,i,1}^\tau \cdots \rho_{l,i,K}^\tau] \} \tag{28}$$

where H is an $M \times K$ standard complex Gaussian matrices, and the covariance matrix of noise and interference caused by prediction and estimation reads

$$\Sigma_l^\tau = (\varepsilon^\tau + \gamma_{UL}) \mathbf{I}_M \tag{29}$$

where

$$\varepsilon^\tau \triangleq \sum_{k=1}^K \sum_{i=1}^L \lambda_{l,i,k} - \sum_{k=1}^K \left(\sum_{i=1}^L \lambda_{l,i,k}^2 |\rho_{l,i,k}^\tau|^2 \right) \left(\sum_{i=1}^L \lambda_{l,i,k} + \gamma_P \right)^{-1} \tag{30}$$

So for the physical channel model (26), the lower bound of the sum-rate defined in (18) can be given by

$$\hat{C}_{LB}^\tau = \log_2 \det \left(\varepsilon^{-\tau} \mathbf{H} \left(\sum_{i=1}^L \mathbf{A}_{l,i}^\tau (\mathbf{A}_{l,i}^\tau)^\mathbf{H} \right) \left(\sum_{i=1}^L \mathbf{A}_{l,i} + \gamma_P \mathbf{I}_K \right)^{-1} \mathbf{H}^\mathbf{H} + \mathbf{I}_M \right) - \log_2 \det \left(\varepsilon^{-\tau} \mathbf{H} \left(\sum_{i \neq l}^L \mathbf{A}_{l,i}^\tau (\mathbf{A}_{l,i}^\tau)^\mathbf{H} \right) \left(\sum_{i=1}^L \mathbf{A}_{l,i} + \gamma_P \mathbf{I}_K \right)^{-1} \mathbf{H}^\mathbf{H} + \mathbf{I}_M \right) \quad (31)$$

where $\varepsilon^{-\tau} \triangleq (\varepsilon^\tau + \gamma_{UL})^{-1}$.

Theorem 3 For the physical channel model (26), as $M \rightarrow \infty$, \hat{C}_{LB}^τ obeys $\hat{C}_{LB}^\tau - \hat{C}_{LB,\text{inf}}^\tau \xrightarrow{M \rightarrow \infty} 0$, where the infinite sum-rate is expressed as

$$\hat{C}_{LB,\text{inf}}^\tau = \sum_{k=1}^K \log_2 \left(1 + \frac{|\lambda_{l,l,k} \rho_{l,l,k}^\tau|^2}{\sum_{i \neq l}^L |\lambda_{l,i,k} \rho_{l,i,k}^\tau|^2 + \frac{(\varepsilon^\tau + \gamma_{UL})}{M} \left(\sum_{i=1}^L \lambda_{l,i,k} + \gamma_P \right)} \right) \quad (32)$$

Proof See ‘‘Appendix’’.

We can make several observations from (32). We consider a group of users with the same CSI delay in all cells, so their channel prediction coefficients only depend on their mobile velocity. If $\rho_{l,i,k}^\tau \triangleq \rho_{l,i,k}^\tau \neq 0$, for all i, k , substituting (30) into (32), and make some algebraic manipulations, the second part of the logarithm becomes

$$\frac{\lambda_{l,l,k}^2}{\sum_{i \neq l}^L \lambda_{l,i,k}^2 + \lambda'_{l,k} \left(\frac{\left(\sum_{k=1}^K \sum_{i=1}^L \lambda_{l,i,k} + \gamma_{UL} \right)}{M |\rho_{l,i,k}^\tau|^2} - \frac{\sum_{k=1}^K \left(\sum_{i=1}^L \lambda_{l,i,k}^2 \right) \lambda'_{l,k}^{-1}}{M} \right)}$$

where $\lambda'_{l,k} = \left(\sum_{i=1}^L \lambda_{l,i,k} + \gamma_P \right)$.

Since $|\rho_{l,i,k}^\tau|^2 \leq 1$, $\hat{C}_{LB,\text{inf}}^\tau \leq \hat{C}_{LB,\text{inf}}^0$. That is to say, if the scheduled users are all with the same mobile velocity, the uplink sum-rate will degrade due to mobility. But if we schedule the users of all cells with different mobile velocity, how will the sum-rate become? It seems as if the sum-rate may increase as well as decrease. To verify our guess, we make approximation as follows.

According to (29), we know $\varepsilon^\tau \mathbf{I}_M$ is the covariance matrix of channel error caused by prediction and estimation, so $\varepsilon^\tau \geq 0$. For each k ,

$$\varepsilon_k^\tau = \sum_{i=1}^L \lambda_{l,i,k} - \left(\sum_{i=1}^L \lambda_{l,i,k}^2 |\rho_{l,i,k}^\tau|^2 \right) \left(\sum_{i=1}^L \lambda_{l,i,k} + \gamma_P \right)^{-1} = \frac{\sum_{i=1}^L \lambda_{l,i,k}^2 - \sum_{i=1}^L \lambda_{l,i,k}^2 |\rho_{l,i,k}^\tau|^2 + \sum_{j \neq i}^L \sum_{i=1}^L \lambda_{l,i,k} \lambda_{l,j,k} + \gamma_P \sum_{i=1}^L \lambda_{l,i,k}}{\sum_{i=1}^L \lambda_{l,i,k} + \gamma_P}$$

since $|\rho_{l,i,k}^\tau|^2 \leq 1$, $\sum_{i=1}^L \lambda_{l,i,k}^2 - \sum_{i=1}^L \lambda_{l,i,k}^2 |\rho_{l,i,k}^\tau|^2 \geq 0$, so $\varepsilon_k^\tau > 0$.

This proves that ε^τ is monotonic increasing of K . So the interference depends mainly on the ration K/M . As a result, when $M/K \rightarrow \infty$,

$$\frac{(\varepsilon^\tau + \gamma_{UL})}{M} \left(\sum_{i=1}^L \lambda_{l,i,k} + \gamma_P \right) \rightarrow 0$$

and (32) can be further simplified as

$$\hat{C}_{LB,inf}^\tau = \sum_{k=1}^K \log_2 \left(1 + \frac{|\lambda_{l,l,k} \rho_{l,l,k}^\tau|^2}{\sum_{i \neq l} |\lambda_{l,i,k} \rho_{l,i,k}^\tau|^2} \right) \tag{33}$$

which coincides with (23).

4 Numerical Results

In this section, we validate the analyses presented above through a set of Monte-Carlo simulations. Same as [9], a 7-cell hexagonal system layout is adopted. The inner cell radius is normalized to one, the distance between two adjacent cells is normalized to 2, and we assume a distance-based path loss model with path loss exponent $\alpha = 3.7$. To allow for reproducibility of our results, we distribute $K = 30$ UTs uniformly on a circle of radius $2/3$ around each BS and do not consider shadowing. The entries of the correlation matrix are modeled via the common exponential correlation model $[R_{l,i}]_{m,n} = \kappa^{|m-n|}$ with κ being the transmit correlation coefficient [14]. The carrier frequency is 2.3 GHz. Symbol interval is $\frac{1}{14}$ ms. We assume the CSI delay is invariable, and equals 4 in symbol. So the prediction coefficient is only related to mobile velocity v . We further assume $\gamma_P = \gamma_{UP}$.

Attention: for Figs. 1, 2, 3 and 5, users in all cells move at the same given velocity. Figure 1 plots the achievable sum-rate against the number of antennas M while uses move at different

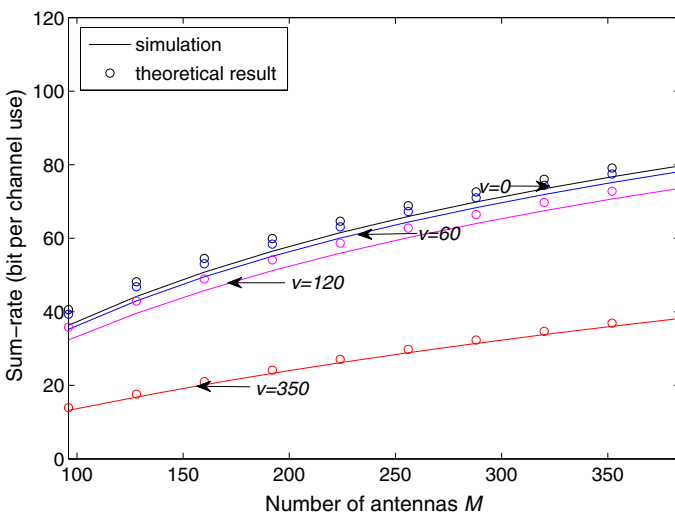


Fig. 1 Ergodic sum-rate for different v , $\kappa = 0.9$

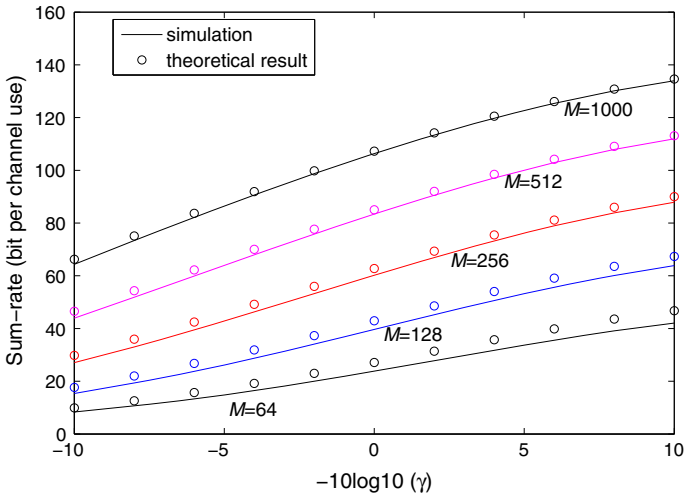


Fig. 2 Ergodic sum-rate for different M , $\kappa = 0.9$

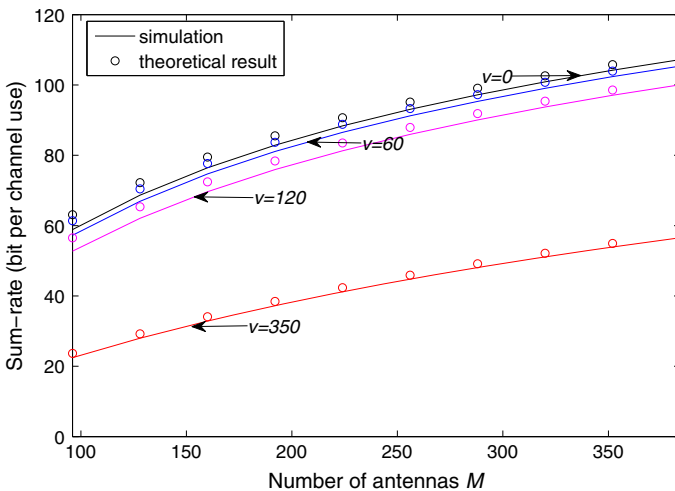


Fig. 3 Ergodic sum-rate for different v , $\kappa = 0$

velocity, and SNR equals 0, κ equals 0.9. We can see that for not so large number and low velocity, our approximation is not very accurate. So we plots Fig. 2 which plots the achievable sum-rate against the signal-to-noise ratio (SNR) with different numbers of antenna M and the mobile velocity is set to 120 km/h. As expected, it can be seen that the approximation for large M is very accurate. More important, we can see that as the mobile velocity increases, the sum-rate degrades more deeply for finite M . Interestingly, our approximation is more accurate for higher velocity for finite M . As Fig. 1 shows for users with different mobile velocity, the faster the user moves, the more antennas he needs to achieve the same given sum-rate; and as Fig. 2 shows, the lower the SNR is (which means the lower the transmit power is when we assume the power of the noise is invariable), the more antennas he needs to achieve the same

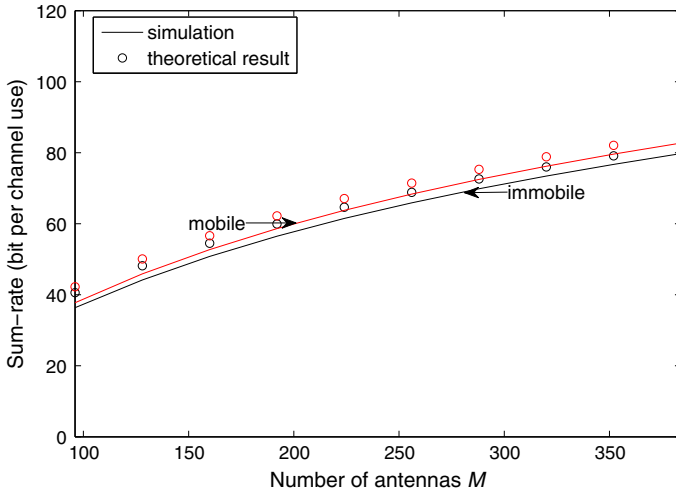


Fig. 4 Ergodic sum-rate for immobile and mobile users, $\kappa = 0.9$

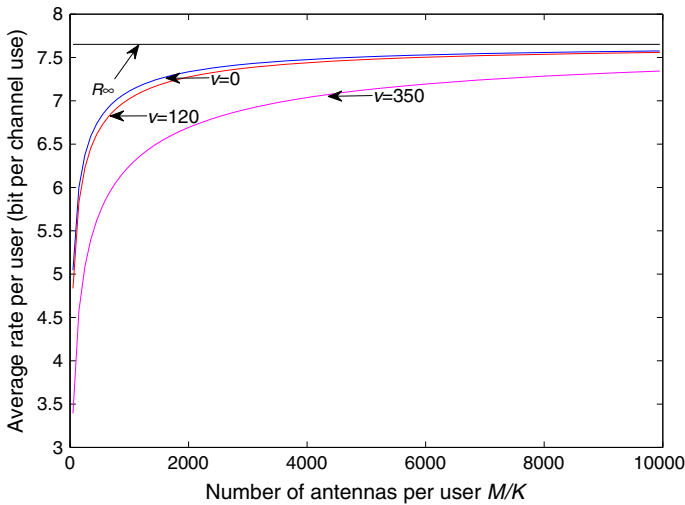


Fig. 5 Average rate per user for different v , $\kappa = 0$

given sum-rate. In other words, large antenna array can not only potentially reduce uplink transmit power, but also compensate for the decay due to user mobility.

Figure 3 plots the achievable sum-rate against the number of antennas M for i.i.d channel model while users move at different given velocity, and SNR equals 0. Compared to Fig. 1, we see that high correlation degrades the sum-rate deeply, and low correlation makes our approximation more accurate.

In order to verify our guess that the sum-rate may increase as well as decrease if we schedule the users of all cells with different mobile velocity, we plot Fig. 4. In Fig. 4, the curve marked immobile is the case with pilot contamination only; the curve marked mobile is the case with both pilot contamination and CSI delay while all users in target cell move at 60km/h, and all users in the interfering cells move at 240km/h. SNR all equal 0. As

is analyzed, due to the impact of pilot contamination the sum-rate even increase when we schedule the users in target cell who move at lower velocity than users in the interfering cells.

Figure 5 plots the average rate per user against large number of antennas M for i.i.d channel model while users move at different given velocity, and SNR equals 0. We use average rate per user instead of sum-rate for more distinct exhibition. As is expected, while we schedule the k -th user of all cells with same mobile velocity, the average rate per user with different mobile velocity will both approach R_∞ as M goes to infinity.

5 Conclusions

We have analyzed the spectral efficiency for massive MIMO taking the practical problems such as the antenna correlation, pilot contamination and CSI delay into consideration. The equivalent channel model for massive MIMO with pilot contamination and CSI delay is derived. With this new model, the lower bound of the sum-rate is obtained, and the asymptotic performance of the sum-rate is analyzed for $M \rightarrow \infty$. The results are general and it covers the conclusions of [9]. Simulation results show that the asymptotic approximation has good performance for large M . We also find that large antenna array can not only potentially reduce uplink transmit power, but also compensate for the decay due to user mobility. Simulation results also verify our guess that CSI delay may not decrease the uplink sum-rate due to the impact of pilot contamination. When the number of BS antennas goes without bound at a much greater rate than the number of users, if we schedule the k -th user of all cells who have the same prediction coefficient, the uplink sum-rate is the same as that with no CSI delay, and simulations verify our analysis.

Acknowledgments This work was supported in part by the Natural Science Foundation of China (NSFC) under Grant 61221002 and 61371113, the National Basic Research Program of China (973 Program 2013CB336600), National Key Special Program No. 2012ZX03 003005-003, 2013ZX03003003-005, the NSFC under Grants 61271205, and State Key Laboratory of Wireless Mobile Communications (CATT).

Appendix: Proof of Theorem 3

Proof Using the matrix identity

$$\det(\mathbf{I} + \mathbf{A}\mathbf{B}) = \det(\mathbf{I} + \mathbf{B}\mathbf{A})$$

(31) can be rewritten as

$$\begin{aligned} \hat{C}_{LB}^\tau &= \log_2 \det \left(\varepsilon^{-\tau} \left(\sum_{i=1}^L \mathbf{A}_{l,i}^\tau (\mathbf{A}_{l,i}^\tau)^\mathbf{H} \right) \left(\sum_{i=1}^L \mathbf{A}_{l,i} + \gamma_P \mathbf{I}_K \right)^{-1} \mathbf{H}^\mathbf{H} \mathbf{H} + \mathbf{I}_K \right) \\ &\quad - \log_2 \det \left(\varepsilon^{-\tau} \left(\sum_{i \neq l}^L \mathbf{A}_{l,i}^\tau (\mathbf{A}_{l,i}^\tau)^\mathbf{H} \right) \left(\sum_{i=1}^L \mathbf{A}_{l,i} + \gamma_P \mathbf{I}_K \right)^{-1} \mathbf{H}^\mathbf{H} \mathbf{H} + \mathbf{I}_K \right) \end{aligned} \quad (34)$$

According to [15, Theorem 3.4, Theorem 3.7], we know $\frac{1}{M} \mathbf{H}^\mathbf{H} \mathbf{H} \xrightarrow{M \rightarrow \infty} \mathbf{I}_K$, so we obtain

$$\frac{1}{M} \varepsilon^{-\tau} \left(\sum_{i=1}^L \mathbf{A}_{l,i}^\tau (\mathbf{A}_{l,i}^\tau)^\mathbf{H} \right) \left(\sum_{i=1}^L \mathbf{A}_{l,i} + \gamma_P \mathbf{I}_K \right)^{-1} \mathbf{H}^\mathbf{H} \mathbf{H}$$

$$-\varepsilon^{-\tau} \left(\sum_{i=1}^L \mathbf{A}_{l,i}^\tau (\mathbf{A}_{l,i}^\tau)^\text{H} \right) \left(\sum_{i=1}^L \mathbf{A}_{l,i} + \gamma_P \mathbf{I}_K \right)^{-1} \mathbf{I}_K \xrightarrow{M \rightarrow \infty} 0 \tag{35}$$

And because $\log_2 \det(\cdot)$ is continuous function, we have

$$\begin{aligned} & \log_2 \det \left(\varepsilon^{-\tau} \left(\sum_{i=1}^L \mathbf{A}_{l,i}^\tau (\mathbf{A}_{l,i}^\tau)^\text{H} \right) \left(\sum_{i=1}^L \mathbf{A}_{l,i} + \gamma_P \mathbf{I}_K \right)^{-1} \mathbf{H}^\text{H} \mathbf{H} + \mathbf{I}_K \right) \\ & - \sum_{k=1}^K \log_2 \left(M \varepsilon^{-\tau} \frac{\sum_{i=1}^L |\lambda_{l,i,k}^\tau|^2}{\sum_{i=1}^L \lambda_{l,i,k} + \gamma_P} + 1 \right) \xrightarrow{M \rightarrow \infty} 0 \end{aligned} \tag{36}$$

Similarly, as $M \rightarrow \infty$, for the second term of the RHS of (34), we have

$$\begin{aligned} & \log_2 \det \left(\varepsilon^{-\tau} \left(\sum_{i \neq l}^L \mathbf{A}_{l,i}^\tau (\mathbf{A}_{l,i}^\tau)^\text{H} \right) \left(\sum_{i=1}^L \mathbf{A}_{l,i} + \gamma_P \mathbf{I}_K \right)^{-1} \mathbf{H}^\text{H} \mathbf{H} + \mathbf{I}_K \right) \\ & - \sum_{k=1}^K \log_2 \left(M \varepsilon^{-\tau} \frac{\sum_{i \neq l}^L |\lambda_{l,i,k}^\tau|^2}{\sum_{i=1}^L \lambda_{l,i,k} + \gamma_P} + 1 \right) \xrightarrow{M \rightarrow \infty} 0 \end{aligned} \tag{37}$$

Defining

$$\begin{aligned} \hat{C}_{\text{LB,inf}}^\tau & \triangleq \sum_{k=1}^K \log_2 \left(M \varepsilon^\tau \frac{\sum_{i=1}^L |\lambda_{l,i,k}^\tau|^2}{\sum_{i=1}^L \lambda_{l,i,k} + \gamma_P} + 1 \right) - \sum_{k=1}^K \log_2 \left(M \varepsilon^{-\tau} \frac{\sum_{i \neq l}^L |\lambda_{l,i,k}^\tau|^2}{\sum_{i=1}^L \lambda_{l,i,k} + \gamma_P} + 1 \right) \\ & = \sum_{k=1}^K \log_2 \left(1 + \frac{|\lambda_{l,l,k}^\tau|^2}{\sum_{i \neq l}^L |\lambda_{l,i,k}^\tau|^2 + \frac{(\varepsilon^\tau + \gamma_{\text{UL}})}{M} (\sum_{i=1}^L \lambda_{l,i,k} + \gamma_P)} \right) \end{aligned} \tag{38}$$

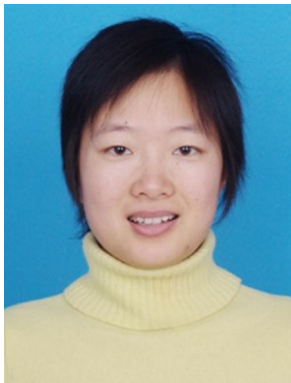
Using (36) and (37), we see that $\hat{C}_{\text{LB}}^\tau - \hat{C}_{\text{LB,inf}}^\tau \xrightarrow{M \rightarrow \infty} 0$.

Substituting $\lambda_{l,i,k}^\tau = \lambda_{l,i,k} \rho_{l,i,k}^\tau$ into (38), we can derive (32).

References

1. Marzetta, T. L. (2010). Noncooperative cellular wireless with unlimited numbers of base station antennas. *IEEE Transactions on Wireless Communications*, 9(11), 3590–3600.
2. Rusek, F., Persson, D., Lau, B. K., Larsson, E. G., Marzetta, T. L., Edfors, O., et al. (2013). Scaling up MIMO: Opportunities and challenges with very large arrays. *IEEE Signal Processing Magazine*, 30(1), 40–60.
3. Ngo, H. Q., & Larsson, E. G. (2012). Very large MIMO, massive MIMO, large-scale antenna systems. <http://www.commsys.isy.liu.se/egl/vlm/vlm.html>.
4. Marzetta, T. L. (2006). How much training is required for multiuser MIMO? In *Proceedings of IEEE asilomar conference on signals, systems and computers (ACSSC06)*, Pacific Grove, CA, USA (pp. 359–363).
5. Jose, J., Ashikhmin, A., Marzetta, T. L., & Vishwanath, S. (2011). Pilot contamination and precoding in multi-cell TDD systems. *IEEE Transactions on Wireless Communications*, 10(8), 2640–2651.
6. Ngo, H. Q., Larsson, E. G., & Marzetta, T. L. (2013). Energy and spectral efficiency of very large multiuser MIMO systems. *IEEE Transactions on Communications*, 61(4), 1436–1449.

7. Ngo, H. Q., Larsson, E. G., & Marzetta, T. L. (2011). Analysis of the pilot contamination effect in very large multicell multiuser MIMO systems for physical channel models. In *Proceedings of IEEE international conference on acoustics, speech and signal processing (ICASSP11)*, Prague, Czech Republic (pp. 3464–3467).
8. Hoydis, J., Brinkz, S., & Debbah, M. (2013). Massive MIMO in the UL/DL of cellular networks: How many antennas do we need? *IEEE Journal on Selected Areas in Communications*, 31(2), 160–171.
9. Wang, D., Ji, C., Gao, X. Q., Sun, S., & You, X. (2013). Uplink sum-rate analysis of multi-cell multi-user massive MIMO system. In *Proceedings of IEEE international conference on communications (ICC13)*, Budapest, Hungary (pp. 3997–4001).
10. Zhou, B., Jiang, L., Zhao, S., & He, C. (2011). BER analysis of TDD downlink multiuser MIMO systems with imperfect channel state information. *EURASIP Journal on Advances in Signal Processing*, 2011, 104.
11. Zhou, S., & Giannakis, G. B. (2004). How accurate channel prediction needs to be for transmit-beamforming with adaptive modulation over Rayleigh MIMO channels? *IEEE Transactions on Wireless Communications*, 3(4), 1285–1294.
12. Kay, S. (1993). *Fundamental of statistical signal processing: Estimation theory*. Englewood Cliffs: Prentice Hall.
13. Hassibi, B., & Hochwald, B. M. (2003). How much training is needed in multiple-antenna wireless links? *IEEE Transactions on Information Theory*, 49(4), 951–963.
14. Loyka, S. L. (2011). Channel capacity of MIMO architecture using the exponential correlation matrix. *IEEE Communications Letters*, 5(9), 369–371.
15. Couillet, R., & Debbah, M. (2011). *Random matrix methods for wireless communications* (1st ed.). New York, NY: Cambridge University Press.



Juan Cao received the B.S. degree in Electronic and Information Engineering and M.S. degree in Communication and Information Systems from Harbin Engineering University in 2003 and 2006 respectively. From 2006 to 2009 she was a teacher in school of electronics and information of Nantong University. She is currently a Ph.D. student in the national mobile communications research laboratory, Southeast University, Jiangsu Province, China. Her research interests include multi-user MIMO precoding, channel estimation, capacity analysis of MIMO systems.



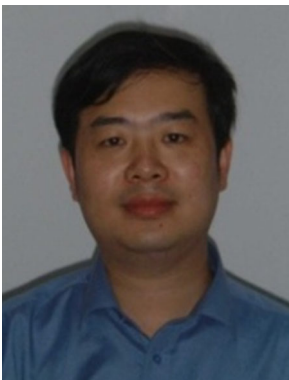
Dongming Wang received the B.S. degree from Chongqing University of Posts and Telecommunications in 1999, the M.S. degree from Nanjing University of Posts and Telecommunications in 2002, and the Ph.D. degree from Southeast University in 2006. He is currently an Associate Professor with the national mobile communication research laboratory, Southeast University. His research interests include turbo detection, channel estimation, capacity analysis of the distributed MIMO channels.



Jiamin Li was born in Hebei province, China, in 1983. He received his B.S. and M.S. degrees in communication and information systems from Hohai University, Jiangsu Province, China, in 2006 and 2009, respectively. He is currently a Ph.D. student in the national mobile communications research laboratory, Southeast University, Jiangsu Province, China. His research interests include cooperative communications, precoding strategies, distributed antenna systems.



Shaohui Sun received the M.S. and Ph.D. degrees in communication and information systems from Xidian University, Xi'an, China, in 1999 and 2003, respectively. From March 2003 to June 2006, he was a Postdoctoral Fellow with the China Academy of Telecommunication Technology, Beijing, China. From June 2006 to December 2010, he was with the Datang Mobile Communications Equipment Company Ltd., Beijing, where he led the standardization activities on the Third-Generation Partnership Project Long-Term Evolution (3GPP LTE) time-division duplex mode. Since December 2010, he has been the Chief Technical Officer with Datang Wireless Mobile Innovation Center of the China Academy of Telecommunication Technology.



Xiaohu You received his B.S. and M.S. degrees from Southeast University, Jiangsu Province, China. He is a ChangJiang Professor and doctoral supervisor. Now he is the director of the National Mobile Communications Research Laboratory in the Southeast University, Jiangsu Province, China. His research interests lie in the areas of communication and signal processing, including future wideband mobile communication theory and technology, intelligent signal processing and communications.