SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

February 2017, Vol. 60 022311:1–022311:14
doi: 10.1007/s11432-016-0476-0

# On the use of H-inf criterion in channel estimation and precoding in massive MIMO systems

Peng XU[1], Dongming WANG[2*] & Feng QI[3*]

[1]*School of Computer Science and Engineering, Northeastern University, Shenyang* 110819*, China;*
[2]*National Mobile Communications Research Laboratory, Southeast University, Nanjing* 210096*, China;*
[3]*Key Laboratory of Opto-Electronic Information Processing, Institute of Automation, Chinese Academy of Science, Shenyang* 110016*, China*

Received July 12, 2016; accepted September 25, 2016; published online December 20, 2016

**Abstract** In this paper, a channel estimation (CE) and precoding scheme by using H-infinity (H-inf) criterion for mitigation of pilot contamination (PC) in massive multiple input multiple output (MIMO) orthogonal frequency division multiplexing (OFDM) systems is investigated. Firstly, different thresholds in H-inf CE and precoding are considered. Secondly, asymptotic analysis is presented to simplify the H-inf precoding, which shows that the complexity of an order of magnitude is reduced. Thirdly, approximate downlink achievable data rates per user are studied for different CE and precoding schemes, such as H-inf and minimum mean square error (MMSE) CE, MMSE, zero-forcing (ZF) and H-inf precoding. The analysis shows that the proposed scheme can provide dual mitigation to the PC. That is, the H-inf CE mitigates the PC by adjusting its thresholds, and the H-inf precoding is utilized to suppress the PC by considering inter-cell interference. The numerical results show that joint use of H-inf CE and H-inf precoding outperforms existing schemes in terms of mitigation to the PC.

**Keywords** massive MIMO, pilot contamination, precoding, channel estimation, H-inf

**Citation** Xu P, Wang D M, Qi F. On the use of H-inf criterion in channel estimation and precoding in massive MIMO systems. Sci China Inf Sci, 2017, 60(2): 022311, doi: 10.1007/s11432-016-0476-0

## 1 Introduction

Since multiple-input multiple-output (MIMO) technology is capable of significantly improving the capacity and reliability of wireless systems, it has been widely studied and applied to the existing wireless systems [1–3]. In order to achieve more gains as well as to simplify the required signal processing, massive MIMO has been proposed in [4–8] As a highly spectrally efficient technology, massive MIMO is indispensable for future wireless communication system or architecture [9–11]. With perfect channel state information (CSI) and zero inter-cell interference, massive MIMO provides multiple advantages compared to traditional MIMO. Theoretical results show that having more antennas at the BS can help increase the system capacity, simplify the signal processing methods, and reduce the transmit power in uplink/downlink [12–15]. Moreover, the effects of uncorrelated noise and small-scale fading can be eliminated [13]. To avoid the high overhead of pilots, channel estimation (CE) in massive MIMO is generally performed in time division duplex (TDD) mode [12, 13].

* Corresponding author (email: wangdm@seu.edu.cn, qifeng@sia.cn)

 

However, ideal assumptions do not generally hold in practice, and the CSI is severely contaminated by the interference. In a multi-cell scenario, accurate CSI is critical for achieving the advantages of massive MIMO. Due to the mobility of users and the limited bandwidth, it is difficult to allocate dedicated pilots to each cell for multi-user CSI estimation, thus the reuse of pilots is must. One of the problems of pilot reuse is pilot contamination (PC), which is caused by allocating non-orthogonal pilots to the users in different cells. Under the impact of PC, it has been shown that the achievable rate is saturated even if the number of antennas at the BS increases to infinite [4]. In recent years, CE and precoding have been respectively investigated to mitigate the PC in massive MIMO systems [16–26]. However, the existing researches generally addressed these two issues separately.

For CE methods, minimum mean square error (MMSE) criterion in the presence of PC was commonly used in [16–18]. A blind CE method based on eigenvalue decomposition (EVD) in [19, 20] was developed to mitigate the PC by exploiting the asymptotically orthogonal channels between different users. In [21], a CE technique was developed to mitigate the PC by utilizing partial decoded data. Besides, a compressive sensing CE method was proposed in [22] based on physical channel with finite scattering model. The existing researches are carried out mainly by considering single-carrier and flat-fading channels. However, when the multi-carrier and multi-path channels are taken into account, the MMSE based CE cannot be used directly, because the high overhead caused by computing the statistical property of channels in this scenario will be impractical, especially for the signals from multiple cells.

For precoding, without consideration of cooperation among the BSs, there are two main approaches to mitigate the PC: single-cell processing and coordinated (or called multi-cell) processing. Although there is no information exchange overhead in the first approach, this precoding cannot mitigate inter-cell interference (e.g., [23–25]). In order to suppress the PC effectively, the design of coordinated processing is essentially necessary. To this end, Ref. [26] proposed a multi-cell MMSE based precoding by minimizing the sum of its own squared error and the interference from other cells, which provides better performance than several single-cell precoding methods. However, the complexity of estimating each precoding matrix will be in the order of $\mathcal{O}(M^3)$, which leads to high complexity when the number of BS antennas is large.

H-infinity (H-inf) estimator, which shows an acceptable estimation performance without using accurate statistical information, has been applied to estimate the channels in single-antenna single-carrier systems [27]. To avoid high computational complexity, in [28], a simplified H-inf CE was designed for multiple-antenna multiple-carrier systems. In [29], the H-inf CE was designed in multi-cell multi-user MIMO orthogonal frequency division multiplexing (OFDM) systems [30–32] to render a substitute for traditional CE methods (e.g., [33–35]) and the PC impacts on mean square error (MSE) of CE were analyzed. H-inf criterion was applied again to design multi-cell precoding in massive MIMO OFDM systems [36]. It has been shown that the downlink achievable rate obtained by using H-inf precoding outperforms that by using single-cell precoding methods. However, in [29, 36], the identical thresholds in H-inf CE and precoding were assumed for simplicity, and the proposed H-inf precoding involves high complexity because of pseudo-inverse operations. Furthermore, the analysis of downlink achievable rate for joint use of H-inf criterion in CE and precoding was not given in [36].

In this paper, a jointly utilizing H-inf criterion in both CE and precoding is investigated for massive MIMO systems. Different from [29, 36], the scheme here is designed to take into account the effect of different thresholds in H-inf CE and precoding on downlink achievable rate. Also, a low-complexity approximate solution of H-inf precoding is obtained by applying the asymptotic analysis for large system dimension. In order to assess the performance of this joint scheme, approximations of downlink achievable rate for several precoding and CE methods (i.e., ZF, MMSE and H-inf precoding methods in the downlink, MMSE and H-inf CE methods in the uplink) are analyzed. By using the H-inf precoding, the downlink achievable rate can be enhanced because the PC is suppressed by considering the inter-cell interference. The performance of H-inf percoding can be improved by adjusting the thresholds of H-inf CE, which suppresses the PC and further enhances downlink achievable rate. Compared to the existing researches in multi-cell scenario, the complexity of the proposed scheme is reduced to the order of $\mathcal{O}(M^2)$. Numerical results show that joint use of H-inf criterion in CE and precoding illustrates significant rate gains compared to existing combinations of CE and precoding.

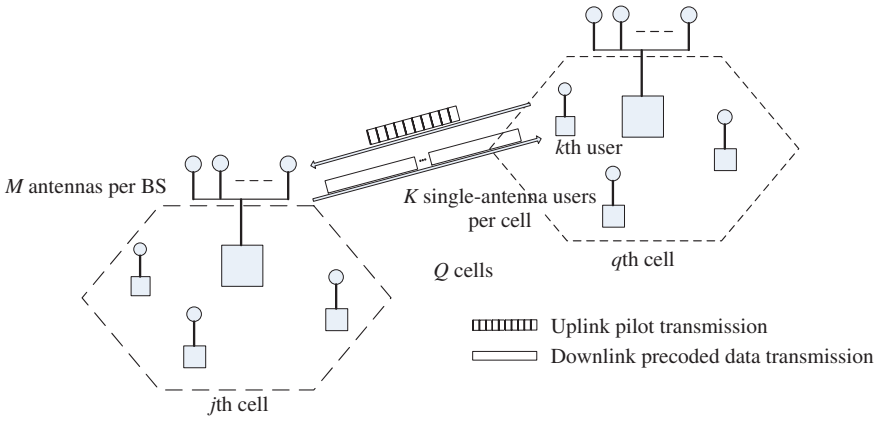


**Figure 1** A massive MIMO system model with TDD protocol.

This paper is structured as follows. In Section 2, the two-phase system models are firstly given. Then, the H-inf CE and multi-cell precoding are introduced in Sections 3. The performance of the proposed scheme is analyzed in Section 4. After the proposed scheme is evaluated via computer simulations in Section 5, the conclusions are drawn in Section 6.

The symbols used in this paper are as follows: Bold italic font variables denote matrices and vectors. $(\cdot)^*$, $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$ represent the conjugate, transpose and the Hermitian transpose operations. $(\cdot)^{\dagger}$ denote the inverse and pseudo inverse operations. $tr\{\cdot\}$ stands for the trace operation, $E[\cdot]$ and $var[\cdot]$ represent expectation and variance. $\boldsymbol{I}_N$ is $N$-dimensional identity matrix, and diag$\{\cdot\}$ stands for diagonal matrix. $\mathcal{CN}(\Gamma, \Upsilon)$ is complex Gaussian distribution with mean $\Gamma$ and covariance matrix $\Upsilon$.

## 2 System model

Consider a multi-cell multi-user MIMO OFDM system with $Q$ cells, each of which consists of one BS with $M$ antennas and $K$ single-antenna mobile terminals, as shown in Figure 1. The number of OFDM subcarriers is $N$. The frequency selective channel is modelled as a finite-duration channel impulse response (CIR) with $L$ paths. We assume all uplink pilots transmissions and BS receptions are synchronous, which provides a worst case scenario from a pilot contamination point of view. In TDD mode channel reciprocity is assumed, and in this case uplink channel is used as an estimation of downlink channel [4].

Two phases are included in this communication scheme: uplink pilot transmission and downlink precoded data transmission. In the first phase or time instant, users in each cell synchronously transmit mutually orthogonal pilot sequences, and the same set of orthogonal pilot sequences is reused in every cell so that the channel estimation is corrupted by pilot contamination from adjacent cells. In the second phase or time instant, after correlating the received pilot sequences at the same time, the BS estimate the uplink channels. Based on channel reciprocity, the estimated uplink channels are used for downlink precoding, and the precoded signals are transmitted from BS to the users.

### 2.1 Uplink pilots transmission

For each received OFDM symbol, the signal vector on all subcarriers at the $r$th antenna at the $j$th BS is given by[1)]

$$\boldsymbol{Y}_{rj} = \sum_{q=1}^{Q} \sum_{k=1}^{K} \boldsymbol{X}_{rqk} \tilde{\boldsymbol{H}}_{rqjk} + \boldsymbol{Z}_{rj}, \tag{1}$$

1) Seen from (1), since in this paper the different received parameters with respect to the $r$th BS antenna, such as $\boldsymbol{Y}_{rj}$, $\boldsymbol{X}_{rqk}$, $\tilde{\boldsymbol{H}}_{rqjk}$ and $\boldsymbol{Z}_{rj}$, are only considered, we will drop the antenna index $r$ for simplicity.

where $\boldsymbol{Y}_j = [y_{j0}, \ldots, y_{jN-1}]^{\mathrm{T}}$, $\boldsymbol{X}_{qk}$ is $N$-dimensional diagonal signal matrix from the $k$th user in the $q$th cell, its diagonal entries are $\mathcal{CN}(0,1)$[2)], $\boldsymbol{Z}_j = [z_{j0}, \ldots, z_{jN-1}]^{\mathrm{T}}$ is $\mathcal{CN}(0, \boldsymbol{I}_N)$ noise vector. $\tilde{\boldsymbol{H}}_{qjk}$ is $N \times 1$ channel frequency response vector between the $j$th BS and the $k$th user in the $q$th cell, $\tilde{\boldsymbol{H}}_{qjk} = \boldsymbol{F}_{NL}\boldsymbol{C}_{qjk}$, where $\boldsymbol{F}_{NL}$ is $1/\sqrt{N}$ times the first $L$ columns of Discrete Fourier Transform (DFT) matrix, $\boldsymbol{C}_{qjk}$ is the $L \times 1$ vector representing the CIR between the $j$th BS and the $k$th user in the $q$th cell, and is given by

$$\boldsymbol{C}_{qjk} = \boldsymbol{D}_{qjk}^{\frac{1}{2}}\boldsymbol{G}_{qjk}, \tag{2}$$

where $\boldsymbol{G}_{qjk}$ is the $L \times 1$ small-scale fading coefficients vector, and let $g_{qjkl}$ be the $l$th element of $\boldsymbol{G}_{qjk}$, $g_{qjkl} \sim \mathcal{CN}(0,1)$. $\boldsymbol{D}_{qjk}$ is $L$-dimensional diagonal matrix whose diagonal elements are $d_{qjkl}$, where $d_{qjkl}$ is assumed to be non-negative constant with respect to $l$ and $k$, corresponding to path-loss and shadow fading coefficients. Since $d_{qjkl}$ changes slowly, it is simplified as $d_{qj}$, and $d_{qj}$ is usually less than 1.

### 2.2 Downlink precoded data transmission

The received signal $\boldsymbol{U}_{jn}$ of all $K$ users at the $n$th subcarrier in the $j$th cell is expressed as (for simplicity, the subcarrier index $n$ is neglected)

$$\boldsymbol{U}_j = \sum_{q=1}^{Q} \sqrt{\lambda_q}\boldsymbol{H}_{qj}\boldsymbol{A}_q\boldsymbol{E}_q + \boldsymbol{V}_j, \tag{3}$$

where $\boldsymbol{A}_q$ is $M \times K$ precoding matrix, $\boldsymbol{E}_q = [e_{q1}, \ldots, e_{qk}]^{\mathrm{T}}$ is $K \times 1$ signal for the $K$ users in the $q$th cell, $\boldsymbol{E}_q \sim \mathcal{CN}(0, \boldsymbol{I}_K)$, and $\boldsymbol{V}_j = [v_{j1}, \ldots, v_{jK}]^{\mathrm{T}}$ is $\mathcal{CN}(0, \boldsymbol{I}_K)$ noise vector. $\lambda_q = \frac{1}{E[\frac{1}{K}\mathrm{tr}\{\boldsymbol{A}_q\boldsymbol{A}_q^{\mathrm{H}}\}]}$ is average power constraint factor. $\boldsymbol{H}_{qj} = [\boldsymbol{h}_{qj1}, \ldots, \boldsymbol{h}_{qjK}]^{\mathrm{T}}$ is $K \times M$ channel propagation matrix between the $K$ users in the $q$th cell and the BS in the $j$th cell, and $\boldsymbol{h}_{qjk} = \tilde{\boldsymbol{h}}_{qjk}d_{qj}^{\frac{1}{2}}$, $\tilde{\boldsymbol{h}}_{qjk}$ is fast fading coefficients vector, $\tilde{\boldsymbol{h}}_{qjk} = [h_{qjk1}, \ldots, h_{qjkM}]^{\mathrm{T}}$, $h_{qjkm} \sim \mathcal{CN}(0,1)$. The received signal from the $k$th user is given by

$$u_{jk} = \sum_{q=1}^{Q}\sum_{p=1}^{K} \sqrt{\lambda_q}\boldsymbol{h}_{qjk}\boldsymbol{a}_{qp}e_{qp} + v_{jk}, \tag{4}$$

where $\boldsymbol{a}_{qp}$ is the $p$th column of $\boldsymbol{A}_q$. The achievable rate $R_{jk}$ of the $k$th user in cell $j$ is expressed as [26]

$$R_{jk} = \log_2\left(1 + \rho_{jk}^{\mathrm{DL}}\right), \tag{5}$$

where $\rho_{jk}^{\mathrm{DL}}$ is the associated signal-to-interference-plus-noise ratio (SINR), and is given by

$$\rho_{jk}^{\mathrm{DL}} = \frac{\lambda_j|E[\boldsymbol{h}_{jjk}\boldsymbol{a}_{jk}]|^2}{1 + \lambda_j\mathrm{var}[\boldsymbol{h}_{jjk}\boldsymbol{a}_{jk}] + \sum_{(q,p)\neq(j,k)}\lambda_q E[|\boldsymbol{h}_{jqp}\boldsymbol{a}_{qp}|^2]}. \tag{6}$$

## 3 Design of multi-cell H-inf channel estimation and precoding with different thresholds

### 3.1 The scheme of H-inf channel estimator

As for CE in massive MIMO systems, the core concept of the H-inf criterion is to seek for a CE method that the ratio between the CE error and the input noise/interference is less than a given threshold. Considering all the users in the $j$th cell, given a positive scalar $s_{jj,\mathrm{UL}}$ for uplink, the objective function of the H-inf CE in the $j$th cell satisfies

$$\sup_{\boldsymbol{Z}_j} \frac{\|\hat{\boldsymbol{C}}_{jj} - \boldsymbol{C}_{jj}\|_{\boldsymbol{W}}^2}{\|\boldsymbol{Z}_j\|^2} < s_{jj,\mathrm{UL}}, \tag{7}$$

2) In most related literatures, the signal power of users are generally normalized, e.g., setting the same average transmit power. In this paper, the normalized power for $\boldsymbol{X}_{qk}$ is used in Section 6, however, before Section 6 we assume that the signal $\boldsymbol{X}_{qk}$ follows $\mathcal{CN}(0,1)$, like [25], for purpose of high-dimension matrix approximation and closed-from expression analysis.

where $\hat{\boldsymbol{C}}_{jj}$ is an $LK \times 1$ vector, denoting the CIR vector to be estimated between all the users in the $j$th cell and the BS in the $j$th cell, $\boldsymbol{C}_{jj} = [\boldsymbol{C}_{jj1}^{\mathrm{T}}, \ldots, \boldsymbol{C}_{jjK}^{\mathrm{T}}]^{\mathrm{T}}$, $\boldsymbol{W} > \boldsymbol{0}$ is a weighting matrix, and $\|\hat{\boldsymbol{C}}_{jj} - \boldsymbol{C}_{jj}\|_{\boldsymbol{W}}^2 = (\hat{\boldsymbol{C}}_{jj} - \boldsymbol{C}_{jj})^{\mathrm{H}}\boldsymbol{W}(\hat{\boldsymbol{C}}_{jj} - \boldsymbol{C}_{jj})$. It should be noted that the H-inf CE is generally a suboptimal solution to the problem because $s_{jj,\mathrm{UL}}$ is not necessarily minimum.

According to [29], the H-inf CE can be described as

$$\hat{\boldsymbol{C}}_{jj} = \boldsymbol{\beta}\boldsymbol{\alpha}^{-1}\boldsymbol{T}_j^{\dagger}\boldsymbol{Y}_j, \tag{8}$$

where $\boldsymbol{T}_j = [\boldsymbol{T}_{j1}, \ldots, \boldsymbol{T}_{jK}]$, $\boldsymbol{T}_{jk} = \boldsymbol{X}_{jk}\boldsymbol{F}_{N,L}$. $\boldsymbol{\beta} = \boldsymbol{M}_{21} + \boldsymbol{M}_{22}\boldsymbol{\zeta}_j$, $\boldsymbol{\alpha} = \boldsymbol{M}_{11} + \boldsymbol{M}_{12}\boldsymbol{\zeta}_j$, are both $LK \times LK$ matrices. $\boldsymbol{\zeta}_j$ is an $LK \times LK$ diagonal matrix, where its diagonal elements satisfy $\max(|\zeta_1|, \ldots, |\zeta_{LK}|) < 1$, and $\zeta_1 = \cdots = \zeta_{LK}$. $\boldsymbol{M}_{11}$, $\boldsymbol{M}_{12}$, $\boldsymbol{M}_{21}$, $\boldsymbol{M}_{22}$ can be expressed as

$$\begin{aligned} \boldsymbol{M}_{11} &= \boldsymbol{\Xi}\boldsymbol{\Psi}_j^{\frac{1}{2}} + \boldsymbol{\Psi}_j^{-\frac{1}{2}}, \quad \boldsymbol{M}_{12} = s_{jj,\mathrm{UL}}^{-\frac{1}{2}}\boldsymbol{\Xi}\boldsymbol{W}^{\frac{1}{2}}, \\ \boldsymbol{M}_{21} &= \boldsymbol{\Xi}\boldsymbol{\Psi}_j^{\frac{1}{2}}, \quad \boldsymbol{M}_{22} = s_{jj,\mathrm{UL}}^{-\frac{1}{2}}\boldsymbol{\Xi}\boldsymbol{W}^{\frac{1}{2}} - s_{jj,\mathrm{UL}}^{\frac{1}{2}}\boldsymbol{W}^{\frac{1}{2}}, \end{aligned} \tag{9}$$

where $\boldsymbol{\Psi}_j = \boldsymbol{T}_j^{\dagger}\boldsymbol{T}_j = \boldsymbol{I}_{LK}$ if quadrature phase shift keying (QPSK) is adopted, $\boldsymbol{\Xi}_2 = (\boldsymbol{\Psi}_j - s_{jj,\mathrm{UL}}^{-1}\boldsymbol{W})^{-1}$ and $\boldsymbol{\Xi}_1$ can be obtained by the canonical factorization of $\boldsymbol{I}_{LK} + \boldsymbol{\Xi}_2$.

Actually, $\boldsymbol{T}_j^{\dagger}\boldsymbol{Y}_j$ is the single-cell maximum likelihood (ML) CE. So, Eq. (9) can be rewritten as follows:

$$\hat{\boldsymbol{C}}_{jj} = \boldsymbol{\gamma}_{jj,\mathrm{UL}}\hat{\boldsymbol{C}}_{jj}^{ML}, \tag{10}$$

where $\boldsymbol{\gamma}_{jj,\mathrm{UL}} = \boldsymbol{\beta}\boldsymbol{\alpha}^{-1}$. Eq. (10) can be considered as a filter matrix $\boldsymbol{\gamma}_{jj,\mathrm{UL}}$ applied to the ML estimation. Thus, the performance of H-inf CE will depend on the design of $\boldsymbol{\gamma}_{jj,\mathrm{UL}}$.

The channels in the current cell ($\boldsymbol{C}_{jj}$) is estimated by (8), while the channels between cells ($\boldsymbol{C}_{qj}$) can be obtained directly via the similar process, just by using different signal from other cells ($\boldsymbol{T}_q$). With respect to the H-inf CE, the optimization problem of H-inf, to a great extent, depends on the thresholds. If the coordinated precoding is considered, the achievable rate in a multi-cell scenario does not depend only on the threshold of the current cell. Thus, it is assumed that the thresholds between cells ($s_{qj,\mathrm{UL}}$, $j \neq q$) are not the same as the threshold within the same cell ($s_{jj,\mathrm{UL}}$), i.e., $s_{jj,\mathrm{UL}} \neq s_{qj,\mathrm{UL}}$. The achievable rate of selecting different values of thresholds will be analyzed and evaluated in Sections 5 and 6.

### 3.2 The scheme of H-inf precoding

We formulate an optimization problem for determining the precoding matrices. Consider the signal and interference terms corresponding to the BS in the $j$th cell, and define two parts in the following objective function. The first part is the precoding error which is defined as the difference between the original signal generated by the $j$th BS and the precoded signal received at the users in the $j$th cell. The second part is the noise/interference from all the other cells experienced by the users in the $j$th cell.

Following the idea of H-inf criterion described above, the concept of the H-inf precoding is to find a precoding matrix $\boldsymbol{A}_j$, so that the ratio between the two parts is less than a prescribed positive threshold $s_{jj,\mathrm{DL}}$ for downlink. The objective function of the H-inf precoding is expressed as

$$\sup_{\sqrt{\lambda_j}\boldsymbol{H}_{qj}\boldsymbol{A}_j\boldsymbol{E}_j + \boldsymbol{V}_j} \frac{\|\sqrt{\lambda_j}\boldsymbol{H}_{jj}\boldsymbol{A}_j\boldsymbol{E}_j - \sqrt{\lambda_j}\boldsymbol{E}_j\|_{\boldsymbol{W}}^2}{\|\sum_{q\neq j}^{Q}\sqrt{\lambda_j}\boldsymbol{H}_{qj}\boldsymbol{A}_j\boldsymbol{E}_j + \boldsymbol{V}_j\|^2} < s_{jj,\mathrm{DL}}, \tag{11}$$

where $\|\sqrt{\lambda_j}\boldsymbol{H}_{jj}\boldsymbol{A}_j\boldsymbol{E}_j - \sqrt{\lambda_j}\boldsymbol{E}_j\|_{\boldsymbol{W}}^2 = \nabla^{\mathrm{H}}\boldsymbol{W}\nabla$, $\nabla = \sqrt{\lambda_j}\boldsymbol{H}_{jj}\boldsymbol{A}_j\boldsymbol{E}_j - \sqrt{\lambda_j}\boldsymbol{E}_j$. Solving (11), a closed-form solution to the precoding matrix $\boldsymbol{A}_j$ is given by [36]

$$\hat{\boldsymbol{A}}_j = \frac{\hat{\boldsymbol{H}}_{jj}^{\dagger}\left(r_{jj,\mathrm{DL}}\boldsymbol{\Delta}_1 + \boldsymbol{\Delta}_2\right)}{2}, \tag{12}$$

where $\boldsymbol{\Delta}_1 = \boldsymbol{I}_K + \sum_{q\neq j}^{Q}\hat{\boldsymbol{H}}_{qj}\hat{\boldsymbol{H}}_{jj}^{\dagger}$, $\boldsymbol{\Delta}_2 = \sqrt{r_{jj,\mathrm{DL}}^2\boldsymbol{\Delta}_1^2 + 4r_{jj,\mathrm{DL}}\lambda_j^{-\frac{1}{2}}\boldsymbol{V}_j\boldsymbol{E}_j^{\mathrm{H}}}$, $r_{jj,DL}$ is a positive scalar. Note that $\hat{\boldsymbol{H}}_{jj}$ is a $K \times M$ estimated channel matrix of $n$th subcarrier between the $K$ users in the $j$th cell and the $j$th BS, which is obtained by applying (8). The similar process for $\hat{\boldsymbol{H}}_{qj}$ is straightforward.

To obtain an ideal solution in (12), $\boldsymbol{V}_j$ and $\boldsymbol{E}_j^{\mathrm{H}}$ should be given. When $M$ is close to infinity, the effects of small-scale fading will vanish [4], and the approximate expression ignoring $\boldsymbol{V}_j$ and $\boldsymbol{E}_j^{\mathrm{H}}$ can be given by

$$\hat{\boldsymbol{A}}_j^{\mathrm{app}} = r_{jj,\mathrm{DL}}\hat{\boldsymbol{H}}_{jj}^{\dagger} + \underbrace{r_{jj,\mathrm{DL}}\sum_{q\neq j}^{Q}\hat{\boldsymbol{H}}_{jj}^{\dagger}\hat{\boldsymbol{H}}_{qj}\hat{\boldsymbol{H}}_{jj}^{\dagger}}_{\text{considering inter-cell interference}}. \tag{13}$$

According to (13), the approximate solution of precoding matrix $\hat{\boldsymbol{A}}_j$ considers the inter-cell interference. In order to reduce the high complexity caused by pseudo-inverse, the following Lemma 1 is applied.

**Lemma 1.** Let $\boldsymbol{X}$ be $K \times M$ matrix, and the elements of $\boldsymbol{X}$ are $\mathcal{CN} \sim (0, 1)$, when $M$ is large and $M \gg K$ , then

$$\left(\frac{\boldsymbol{X}^{\mathrm{H}}\boldsymbol{X}}{M}\right)_{M\gg K} \approx \boldsymbol{I}_K. \tag{14}$$

By applying Lemma 1, Eq. (13) can be simplified as

$$\hat{\boldsymbol{A}}_j^{\mathrm{app}} = \frac{r_{jj,\mathrm{DL}}}{Md_{jj}r_{jj,\mathrm{UL}}^2}\hat{\boldsymbol{H}}_{jj}^{\mathrm{H}} + \frac{r_{jj,\mathrm{DL}}}{(Md_{jj}r_{jj,\mathrm{UL}}^2)^2}\sum_{q\neq j}^{Q}\hat{\boldsymbol{H}}_{jj}^{\mathrm{H}}\hat{\boldsymbol{H}}_{qj}\hat{\boldsymbol{H}}_{jj}^{\mathrm{H}}. \tag{15}$$

By using the approximation of large system dimension, the pseudo-inverse in (13) are vanished.

## 4 Performance analysis

### 4.1 Parameters analysis of matrix $\boldsymbol{W}$, $s_{jj,\mathrm{DL}}$ and $s_{jj,\mathrm{UL}}$

Because of different objective function (7) and (11), it is possible for the thresholds in the uplink and downlink to give different contribution to the performance. So, in this paper, the downlink threshold $s_{jj,\mathrm{DL}}$ for H-inf precoding is defined as a different value compared to the uplink threshold $s_{jj,\mathrm{UL}}$ for H-inf CE. The performance of setting different thresholds in the uplink and downlink will be assessed in the following theoretical analyses and simulation experiments.

According to [29, 36], $\boldsymbol{W}$ is a diagonal matrix. Define $\boldsymbol{W} = w\boldsymbol{I}_K$, where $w$ denotes the diagonal elements of $\boldsymbol{W}$, and then $w < s_{jj,\mathrm{DL}}$ is required. When $\boldsymbol{W}$ is fixed, the monotonic relations between $s_{jj,\mathrm{UL}}$ and $r_{jj,UL}$ is generally presented [29]. Besides, in this paper, $\boldsymbol{W}$ is used in the uplink and downlink, and it is not difficult to find $w < s_{jj,\mathrm{UL}}$.

### 4.2 Downlink achievable rate analysis of proposed scheme in finite system dimension

The existing researches characterized their proposed schemes by analyzing the achievable rate, mainly based on the assumption of infinite system dimension [25,26], however, it is not the case in practice. When finite system dimension is considered, although the exact achievable rate during downlink transmission is hard to derive, an approximate lower bound shown in the Theorem 1 can be derived.

**Theorem 1.** Considering finite system dimension, by using H-inf criterion in CE and precoding, the lower bound of the asymptotically achievable rate of the $k$th user in the $j$th cell is given by

$$R_{jk}^{\text{H-inf}} > \log_2\left(1 + \frac{d_{jj}r_{jj,\mathrm{UL}}^2}{\sum_{q\neq j}^{Q} d_{qj}r_{qj,\mathrm{UL}}^2}\right). \tag{16}$$

*Proof.* See Appendix A.

**Remark 1.** Since $r_{jj,\mathrm{UL}}(r_{qj,\mathrm{UL}})$ and $s_{jj,\mathrm{UL}}(s_{qj,\mathrm{UL}})$ have the monotonic relations [29], adjusting the value of $s_{jj,\mathrm{UL}}(s_{qj,\mathrm{UL}})$ is equivalent to adjusting $r_{jj,\mathrm{UL}}(r_{qj,\mathrm{UL}})$. Accordingly, the mitigation to PC can be implemented by adjusting the uplink thresholds $s_{jj,\mathrm{UL}}$ and $s_{qj,\mathrm{UL}}$. Furthermore, it seems that the performance is irrelevant to the downlink threshold $s_{jj,\mathrm{DL}}(s_{qj,\mathrm{DL}})$.

### 4.3 Downlink achievable rate analysis of different CE and precoding schemes

To assess the performance of the proposed scheme, the approximate achievable rate during downlink transmission by using different CE and precoding schemes will be analyzed in the following three cases.

#### 4.3.1 *The H-inf CE used in the uplink*

Single-cell ZF and MMSE precoding methods in the downlink are introduced [23, 24], and the case when SNR $\to \infty$ is considered. The following Theorem 2 can be obtained.

**Theorem 2.** By using H-inf CE and single-cell ZF (MMSE) precoding, considering SNR $\to \infty$, the asymptotically achievable rate of the $k$th user in the $j$th cell is given by

$$\lim_{\mathrm{SNR}\to\infty} R_{jk}^{\mathrm{ZF}}, R_{jk}^{\mathrm{MMSE}} = \log_2\left(1+\frac{d_{jj}r_{jj,\mathrm{UL}}^2}{\sum_{q\neq j}^{Q} d_{qj}r_{qj,\mathrm{UL}}^2}\right). \tag{17}$$

*Proof.* See Appendix B.

**Remark 2.** Similar to Theorem 2, when the H-inf CE is used, the PC can also be mitigated by adjusting the values of uplink thresholds, even if the H-inf precoding is not used. Since the inter-cell interference is neglected in this precoding methods, the achievable rate is worse than that of rate by using H-inf precoding.

Referring to Theorems 1 and 2, no matter what kind of precoding methods are used, the H-inf CE is an essential way to suppress the PC.

#### 4.3.2 *The H-inf precoding used in the downlink*

The MMSE CE is introduced, and for finite system dimension case the following Theorem 3 is obtained.

**Theorem 3.** By using MMSE CE and H-inf percoding, considering finite system dimension, the lower bound of the asymptotically achievable rate of the $k$th user in the $j$th cell is given by

$$R_{jk}^{\text{MMSE-CE}} > \log_2\left(1+\frac{d_{jj}}{\sum_{q\neq j}^{Q} d_{qj}}\right). \tag{18}$$

By using the similar derivation to Theorem 1, we can obtain Theorem 3. The detailed proof is neglected due to the limited space.

**Remark 3.** When MMSE CE is used, the achievable rate depends only on the values of direct gain $d_{jj}$ and cross gain $d_{qj}$, however, these two parameters cannot be controlled by the designer in practice.

#### 4.3.3 *Without use of H-inf criterion in the uplink and downlink*

We assess the achievable rate when MMSE CE and single-cell ZF (MMSE) precoding are adopted, and the case when SNR $\to \infty$ is considered. The following Theorem 4 can be obtained.

**Theorem 4.** By using MMSE CE and single-cell ZF (MMSE) precoding, considering SNR $\to \infty$, the asymptotically achievable rate of the $k$th user in the $j$th cell is given by

$$\lim_{\mathrm{SNR}\to\infty} R_{jk}^{\mathrm{ZF}}, R_{jk}^{\mathrm{MMSE}} = \log_2\left(1+\frac{d_{jj}}{\sum_{q\neq j}^{Q} d_{qj}}\right). \tag{19}$$

By using the similar derivation to Theorem 2, we can obtain Theorem 4. The detailed proof is neglected due to the limited space.

**Remark 4.** Similar to Theorem 3, when MMSE CE is used, the achievable rate also depends only on the values of $d_{jj}$ and $d_{qj}$. Without considering the inter-cell interference from other cells in this precoding scheme, the worse achievable rate is obtained compared to that of rate by using H-inf precoding.

Referring to Theorems 1–4, no matter what kind of CE method is used, the H-inf precoding is another way to suppress the PC.

**Table 1** Achievable rate for different CE and precoding schemes

| CE | Precoding | Approximate achivable rate per user |
|---|---|---|
| H-inf | Single-cell ZF (MMSE) | $\log_2\left(1+\frac{d_{jj}r_{jj,\mathrm{UL}}^2}{\sum_{q\neq j}^{Q} d_{qj}r_{qj,\mathrm{UL}}^2}\right)$ |
| H-inf | Multi-cell H-inf | $>\log_2\left(1+\frac{d_{jj}r_{jj,\mathrm{UL}}^2}{\sum_{q\neq j}^{Q} d_{qj}r_{qj,\mathrm{UL}}^2}\right)$ |
| MMSE | Multi-cell H-inf | $>\log_2\left(1+\frac{d_{jj}}{\sum_{q\neq j}^{Q} d_{qj}}\right)$ |
| MMSE | Single-cell ZF (MMSE) | $\log_2\left(1+\frac{d_{jj}}{\sum_{q\neq j}^{Q} d_{qj}}\right)$ |

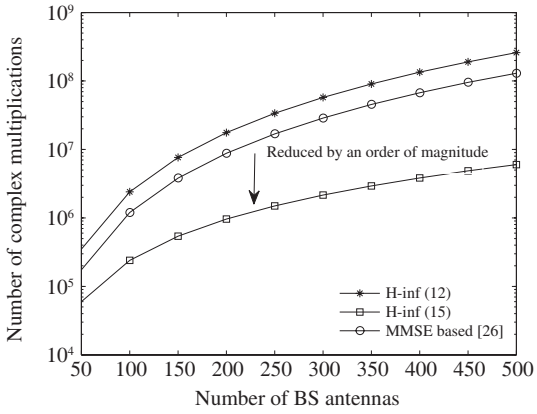


**Figure 2** Complexity comparisons between our proposed H-inf precoding schemes and MMSE based precoding in [26].

### 4.4 Summary of achievable rate for different CE and precoding schemes

To assess the achievable rate of the proposed H-inf criterion scheme, Table 1 is given for comparisons. According to Table 1, the following conclusions can be drawn:

(1) By using H-inf precoding, the achievable rate can be improved by mitigating the PC due to considering inter-cell interference.

(2) Introducing the H-inf CE can mitigate the PC, by adjusting the uplink thresholds within the same cell ($s_{jj,\mathrm{UL}}$) and between cells ($s_{qj,\mathrm{UL}}$).

(3) Joint use of H-inf criterion in CE and H-inf precoding, a dual suppression to the PC is provided, compared to other existing schemes. Moreover, the performance of precoding only depends on the uplink thresholds $s_{jj,\mathrm{UL}}$ and $s_{qj,\mathrm{UL}}$.

### 4.5 Complexity comparisons for different precoding methods

The number of complex multiplications for each estimated precoding matrix are considered as a complexity metric. Assuming $Q = K = 4$, a comparison of complexities between our proposed H-inf precoding in (12) and (15), and MMSE based precoding investigated in [26] is given in Figure 2.

Due to requiring two pseudo-inverse operations, the complexity of the H-inf precoding (12) will be the highest (i.e., $\mathcal{O}(M^3)$), compared with MMSE based precoding in [26] which requires one inverse operations. Thus the computational complexity of these two precoding methods grows remarkably as the number of antennas at the BS increases. With respect to (15), by using the approximation of large system dimensions, the pseudo-inverse operations of the H-inf precoding (12) are eliminated, and the computational complexity is reduced by an order of magnitude (i.e., $\mathcal{O}(M^2)$).

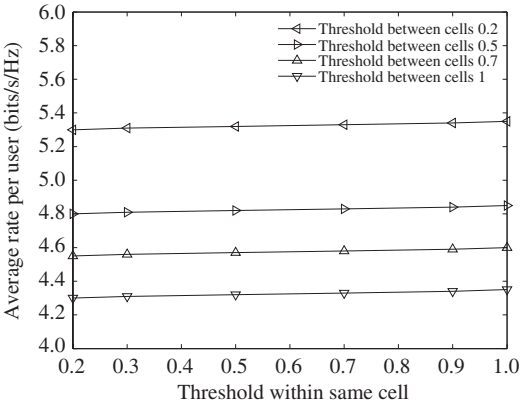


**Figure 3** Average per-user rate for different thresholds between cells.

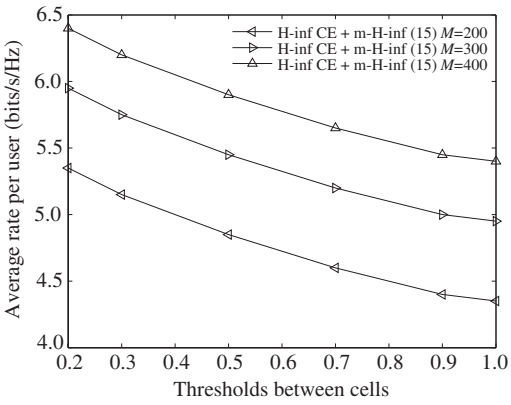


**Figure 4** Average per-user rate for different $M$ versus thresholds between cells at $s_{jj,\mathrm{UL}} = 1$.

## 5 Simulation results

We investigate the performance of the proposed schemes through simulations. It is assumed that there are $Q = 4$ cells, and $K = 4$ users in each cell. Phase shifted orthogonal pilot is adopted to estimate the channels [33] and is reused in all the cells. $L = 8$ multi-path channels are assumed. It is assumed the scale factor $w = 0.1$ for all simulations. The thresholds $s_{jj,\mathrm{DL}}$ of H-inf precoding is 1, and the thresholds of H-inf CE $s_{jj,\mathrm{UL}}$ and $s_{qj,\mathrm{UL}}$ are set to be not larger than 1[3)]. For all $k$, $d_{jqk} = 1$ when $j = q$, and $d_{jqk} = a$ when $j \neq q$ , and $a = 0.1$ from Figure 3 to Figure 6. The uplink SNR is 10 dB, and downlink SNR is 20 dB. OFDM subcarriers $N = 128$, cyclic prefix is 16, and normalized signal power for QPSK is used. In the following simulations, the single-cell OFDM based CE methods are used in the uplink. Each scheme simulated is simply expressed as a combination of "X CE + Y", which means that X CE and Y precoding are considered respectively in uplink and downlink. Specifically, for precoding methods, single-cell and coordinated processing (multi-cell) methods are expressed as "s-Y" and "m-Y" respectively.

### 5.1 Selection of $s_{jj,\mathrm{UL}}$ and $s_{qj,\mathrm{UL}}$

As analyzed in Section 4, the performance of our schemes is relevant to the selection of $s_{jj,\mathrm{UL}}$ and $s_{qj,\mathrm{UL}}$. Thus, the simulation experiments are performed in two cases: (1) For a fixed value of $s_{qj,\mathrm{UL}}$, the effect of varying $s_{jj,\mathrm{UL}}$ is evaluated. (2) For a fixed value of $s_{jj,\mathrm{UL}}$, the effect of varying $s_{qj,\mathrm{UL}}$ is evaluated.

Figure 3 shows the achievable rates per user for different thresholds $s_{qj,\mathrm{UL}}$ as a function of $s_{jj,\mathrm{UL}}$ at $M = 200$. It can be seen that when $s_{qj,\mathrm{UL}}$ is 0.2, 0.5, 0.7 or 1, the achievable rates are not sensitive to the value of $s_{jj,\mathrm{UL}}$ which ranges from 0.2 to 1. On the other hand, the performance of our scheme could be improved by decreasing the value of $s_{qj,\mathrm{UL}}$. In order to characterize the effect of $s_{qj,\mathrm{UL}}$ in detail, the following experiment is undertaken to give a clear assessment. Figure 4 shows the achievable rates for different $M$ as a function of $s_{qj,\mathrm{UL}}$ at $s_{jj,\mathrm{UL}} = 1$. The thresholds $s_{qj,\mathrm{UL}}$ are assumed to be gradually increased from 0.2 to 1. As depicted in the figure, for different values of $M$, when $s_{jj,\mathrm{UL}}$ is fixed, the achievable rate per user decreases as the value of $s_{qj,\mathrm{UL}}$ increases, according to Theorem 1.

In Figures 3 and 4, $s_{qj,\mathrm{UL}}$ can be considered as a filter performance level of the H-inf CE towards the channels between cells. Since the coordinated precoding mitigates the PC by utilizing the channel information between cells, when $s_{qj,\mathrm{UL}}$ is decreased, the performance of H-inf precoding will be improved, and thus for the system performance. Therefore, in a multi-cell scenario, the performance, to a great extent, depends on the mitigation to the inter-cell interference. Furthermore, it is observed that the H-inf CE is a good way for multi-cell scenario to mitigate the PC.

3) In our simulations, we observe the performance of H-inf CE and precoding are not very sensitive when the thresholds is larger than 1.

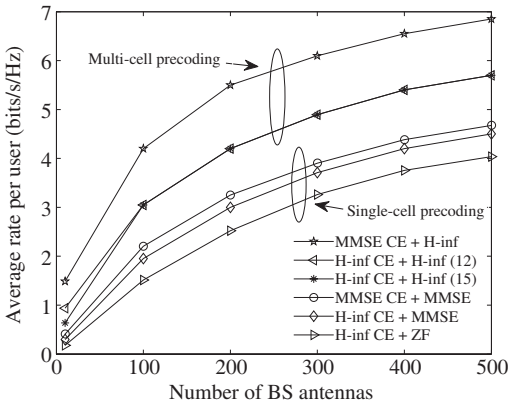


**Figure 5** Average per-user rate for different CE and precoding schemes versus the number of BS antennas at $s_{qj,\mathrm{UL}} = 1$.

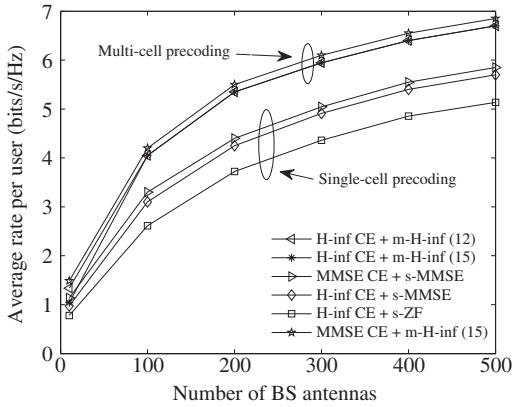


**Figure 6** Average per-user rate for different joint CE and precoding schemes versus the number of BS antennas at $s_{qj,\mathrm{UL}} = 0.2$.

### 5.2 Performance comparisons between different schemes

From the above, the performance of our scheme is enhanced by decreasing the value of $s_{qj,\mathrm{UL}}$, and the minimum and maximum can be obtained respectively in the given range. For purposed of comparing our scheme to the existing schemes, the experiments will be performed in the following two cases: $s_{qj,\mathrm{UL}} = 1$ and $s_{qj,\mathrm{UL}} = 0.2$.

Figure 5 illustrates the achievable rates by adopting different CE and precoding schemes with the number of BS antennas at $s_{qj,\mathrm{UL}} = 1$. It can be seen that the H-inf percoding (15) can provide the same performance as the H-inf precoding (12) when $M \geqslant 100$, which is because the small-scale fading is averaged out for large $M$. Thus, a substitute for (12) by approximately utilizing (15) is reasonable in large-scale MIMO case. Compared to the schemes which use single-cell precoding (ZF or MMSE) cases, the multi-cell H-inf precoding exhibits significant advantage by considering inter-cell interference, which suppresses the PC and improve the achievable rate. It is also observed in those multi-cell precoding designs that when $s_{qj,\mathrm{UL}} = 1$ the performance of using H-inf CE is much worse than that of using MMSE CE which provides the optimal filter performance. This phenomenon means that the H-inf CE is not always good for any values of $s_{qj,\mathrm{UL}}$. Since the performance of H-inf CE depends on the value of thresholds and the signals from other cells are assumed to be sent synchronously, the filter performance level to the channels between cells should be enhanced, that is, the values of $s_{qj,\mathrm{UL}}$ should be decreased.

Figure 6 shows the achievable rates for different CE and precoding schemes as a function of the number of BS antennas at $s_{qj,\mathrm{UL}} = 0.2$. As expected, the scheme which adopts single-cell CE and single-cell precoding gets worse performance without considering the mitigation to inter-cell interference. Similar to Figure 5, when $M \geqslant 100$, the H-inf precoding (15) can obtain the same performance as the H-inf precoding (12). It is shown in those multi-cell precoding designs that when $s_{qj,\mathrm{UL}} = 0.2$, the filter performance level to the channels between cells is enhanced, the performance of using H-inf CE is improved obviously compared to Figure 5, which indicates that the PC can be suppressed dramatically by adjusting the value of $s_{qj,\mathrm{UL}}$. In this case, the dual mitigation by utilizing the H-inf criterion in both uplink and downlink exhibits significant gains compared to nearly all the schemes. Furthermore, its performance is almost the same as the scheme of using optimal MMSE CE. All the results shown in this figure validate the analysis in Section 4, indicating that the use of H-inf criterion in both uplink and downlink provide a better mitigation to the PC than other schemes.

Figure 7 illustrates the performance of four CE and precoding schemes versus different cross gains ($a$) at $s_{qj,\mathrm{UL}} = 0.2$ and $M = 200$. Seen from the figure, the achievable rates per user for all schemes show gradual degradation as cross gain increases. The PC impact is very significant if cross gain is in the same order of direct gain. Even so, it can also be seen that significant advantage by using H-inf criterion in both uplink and downlink for a wide range of values of cross gain. For example, when $a = 0.1$, the

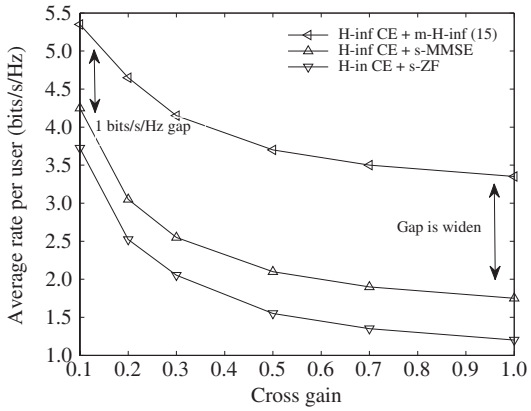


**Figure 7** Average per-user rate for different precoding methods versus different values of $a$, at $s_{qj,\mathrm{UL}} = 0.2$ and $M = 200$.

proposed dual mitigation scheme to the PC has about 1 bits/s/Hz gain than the other three schemes. Furthermore, as the cross gain is increased up to 1, that is, cross gain is equal to direct gain, the gap is further widen. Therefore, the proposed H-inf CE and H-inf precoding scheme performs well in the worst case where the cross gain is of the same order of direct gain.

## 6 Conclusion

This paper investigated the CE and precoding scheme by applying H-inf criterion in massive MIMO systems. The effect of different thresholds on downlink achievable rate for H-inf CE and precoding is considered. The asymptotic analysis for large system dimension was utilized to simplify the high complexity of H-inf precoding. The achievable rate per user for different CE and precoding schemes were assessed. The conclusions are as follows:

(1) The H-inf CE is capable of mitigating the PC by adjusting the thresholds for different cells. Specifically, the H-inf CE is not sensitive to the threshold within same cell ($s_{jj,\mathrm{UL}}$) and depends mainly on the decreasing value of thresholds between cells ($s_{qj,\mathrm{UL}}$). In a multi-cell scenario, the mitigation to the inter-cell interference plays an important role in improving the performance.

(2) By using the approximation, the computational complexity of the solution of precoding is reduced by an order of magnitude. Moreover, when the number of antennas at the BS is large, the performance loss of the approximate solution is negligible.

(3) Joint use of H-inf criterion in CE and precoding can provide dual mitigation to the PC. The performance of precoding depends only on uplink thresholds $s_{jj,\mathrm{UL}}$ and $s_{qj,\mathrm{UL}}$.

**Acknowledgements** This work was supported by Basic Scientific Research Business Expenses of China (Grant No. N152304009), National Natural Science Foundation of China (Grant Nos. 61271205, 61374097, 61300195, 61473066, 61403069).

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

1 Gesbert D, Shafi M, Shiu D, et al. From theory to practice: an overview of MIMO space-time coded wireless systems. IEEE J Sel Areas Commun, 2003, 21: 281–302
2 Wang J Z, Zhu H L, Gomes N. Distributed antenna systems for mobile communications in high speed trains. IEEE J Sel Areas Commun, 2012, 30: 675–683
3 Zhu H L. Performance comparison between distributed antenna and microcellular systems. IEEE J Sel Areas Commun, 2011, 29: 1151–1163
4 Marzetta T L. Noncooperative cellular wireless with unlimited numbers of base station antennas. IEEE Trans Wirel Commun, 2010, 9: 3590–3600
5 Wang D M, Zhang Y, Wei H, et al. An overview of transmission theory and techniques of large-scale antenna systems for 5G wireless communications. Sci China Inf Sci, 2016, 59: 081301

6 Xin Y X, Wang D M, Li J M, et al. Area spectral efficiency and area energy efficiency of massive MIMO cellular systems. IEEE Trans Veh Technol, 2016, 65: 3243–3254
7 Wei H, Wang D M, Wang J Z, et al. Impact of RF mismatches on the performance of massive MIMO systems with ZF precoding. Sci China Inf Sci, 2016, 59: 022302
8 Wei H, Wang D M, Zhu H L, et al. Mutual coupling calibration for multiuser massive MIMO systems. IEEE Trans Wirel Commun, 2016, 15: 606–619
9 Garcia V, Zhou Y Q, Shi J L. Coordinated multipoint transmission in dense cellular networks with user-centric adaptive clustering. IEEE Trans Wirel Commun, 2014, 13: 4297–4308
10 Zhou Y Q, Liu H, Pan Z G, et al. Spectral and energy efficient two-stage cooperative multicast for LTE-A and beyond. IEEE Wirel Commun, 2014, 21: 34–41
11 Gao F F, Zhang K Q. Enhanced multi-parameter cognitive architecture for future wireless communications. IEEE Commun Mag, 2015, 53: 86–92
12 Rusek F, Persson D, Lau B K, et al. Scaling up MIMO: opportunities and challenges with very large arrays. IEEE Signal Process Mag, 2013, 30: 40–46
13 Ngo H Q, Larsson E G, Marzetta T L. Energy and spectral efficiency of very large multiuser MIMO systems. IEEE Trans Commun, 2013, 61: 1436–1449
14 Larsson E G, Edfors O, Tufvesson F, et al. Massive MIMO for next generation wireless systems. IEEE Commun Mag, 2014, 52: 186–195
15 Ngo H Q, Matthaiou M, Duong T Q, et al. Uplink performance analysis of multiuser MU-SIMO systems with ZF receivers. IEEE Trans Veh Technol, 2013, 62: 4471–4483
16 Wang D M, Ji C, Gao X Q, et al. Uplink sum-rate analysis of multi-cell multi-user massive MIMO system. In: Proceedings of IEEE International Conference on Communications, Budapest, 2013. 5404–5408
17 Yin H, Gesbert D, Filippou M, et al. A coordinated approach to channel estimation in large-scale multiple-antenna systems. IEEE J Sel Areas Commun, 2013, 31: 264–273
18 Shariati N, Björnson E, Bengtsson M, et al. Low-complexity channel estimation in large-scale MIMO using polynomial expansion. In: Proceedings of IEEE Annual International Symposium on Personal, Indoor, and Mobile Radio Communications, London, 2013. 1158–1162
19 Ngo H Q, Larsson E G. EVD-based channel estimation in multicell multiuser MIMO systems with very large antenna array. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, Kyoto, 2012. 3249–3252
20 Guo K F, Guo Y, Ascheid G. On the performance of EVD-based channel estimations in MU-massive-MIMO Systems. In: Proceedings of IEEE Annual International Symposium on Personal, Indoor, and Mobile Radio Communications, London, 2013. 1376–1380
21 Ma J J, Li P. Data-aided channel estimation in large antenna system. IEEE Trans Signal Process, 2014, 62: 3111–3124
22 Nguyen S, Ghrayeb A. Compressive sensing-based channel estimation for massive multiuser MIMO systems. In: Proceedings of IEEE Wireless Communications and Networking Conference, Shanghai, 2013. 2890–2895
23 Peel C, Hochwald B, Swindlehurst A. A vector-perturbation technique for near-capacity multiantenna multiuser communication—Part I: channel inversion and regularization. IEEE Trans Commun, 2005, 53: 195–202
24 Gao X, Edfors O, Rusek F, et al. Linear pre-coding performance in measured very-large MIMO channels. In: Proceedings of IEEE Vehicular Technology Conference, San Francisco, 2011. 1–5
25 Hoydis J, Brink S T, Debbah M. Massive MIMO in the UL/DL of cellular networks: how many antennas do we need. IEEE J Sel Areas Commun, 2013, 31: 160–171
26 Jose J, Ashikhmin A, Marzetta T L, et al. Pilot contamination and precoding in multi-cell TDD systems. IEEE Trans Wirel Commun, 2011, 10: 2640–2651
27 Cao C T, Xie L H, Zhang L S. $H_\infty$ channel estimator design for DS-CDMA systems: a polynomial approach in Krein space. IEEE Trans Veh Technol, 2008, 57: 819–827
28 Xu P, Wang J K, Qi F. EM-based H-inf channel estimation in MIMO-OFDM systems. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, Kyoto, 2012. 3189–3192
29 Xu P, Wang J Z, Wang J K, et al. Analysis and design of channel estimation in multi-cell multi-user MIMO OFDM systems. IEEE Trans Veh Technol, 2015, 64: 610–620
30 Zhu H L, Wang J Z. Chunk-based resource allocation in OFDMA systems—Part I: chunk allocation. IEEE Trans Commun, 2009, 57: 2734–2744
31 Zhu H L, Wang J Z. Chunk-based resource allocation in OFDMA systems—Part II: joint chunk, power and bit allocation. IEEE Trans Commun, 2012, 60: 499–509
32 Zhu H L. Radio resource allocation for OFDMA systems in high speed environments. IEEE J Sel Areas Commun, 2012, 30: 748–759
33 Barhumi I, Leus G, Moonen M. Optimal training design for MIMO OFDM systems in mobile wireless channels. IEEE Trans Signal Process, 2003, 51: 1615–1624
34 Xie Y Z, Georghiades C N. Two EM-type channel estimation algorithms for OFDM with transmitter diversity. IEEE Trans Commun, 2003, 51: 106–115
35 Gao J, Liu H P. Low-complexity MAP channel estimation for mobile MIMO-OFDM systems. IEEE Trans Wirel Commun, 2008, 7: 774–780
36 Xu P, Wang J Z, Wang J K. Multi-cell H-inf precoding in massive MIMO systems. In: Proceedings of IEEE International Conference on Communications, Sydney, 2014. 4483–4487

## Appendix A Proof of Theorem 1

For convenient analysis, (13) is divided into two parts

$$\hat{\boldsymbol{A}}_j^{\text{app}} = \hat{\boldsymbol{A}}_1 + \hat{\boldsymbol{A}}_2, \tag{A1}$$

where $\hat{\boldsymbol{a}}_{jk1}$ is defined as the $k$th column of $\hat{\boldsymbol{A}}_1$, $\hat{\boldsymbol{A}}_1 = \frac{r_{jj,\text{DL}}}{Md_{jj}r_{jj,\text{UL}}^2}\hat{\boldsymbol{H}}_{jj}^{\text{H}}$, and $\hat{\boldsymbol{a}}_{jk2}$ is defined as the $k$th column of $\hat{\boldsymbol{A}}_2$, $\hat{\boldsymbol{A}}_2 = \frac{r_{jj,\text{DL}}}{(Md_{jj}r_{jj,\text{UL}}^2)^2}\sum_{q\neq j}^{Q}\hat{\boldsymbol{H}}_{jj}^{\text{H}}\hat{\boldsymbol{H}}_{qj}\hat{\boldsymbol{H}}_{jj}^{\text{H}}$.

Since the exact closed-form solution cannot be obtained for finite system dimension, we investigate the lower bound of SINR, and thus the achievable rate. The column-vectors of $\hat{\boldsymbol{H}}_{qj}$ and $\hat{\boldsymbol{H}}_{jj}$ are asymptotically orthogonal, and it is not difficult to derive $E[\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk2}] = 0$, and then $\text{var}[\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk2}] = E[|\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk2}|^2]$.

From the denominator of (6), one obtains

$$E\left[\left|\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk1}+\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk2}\right|^2\right] < E\left[\left|\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk1}\right|^2+\left|\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk2}\right|^2\right] < E\left[\left|\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk1}\right|^2\right]+E\left[\left|\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk2}\right|^2\right]. \tag{A2}$$

In (A2), the term $E[|\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk1}|^2]$ is expressed as

$$E\left[\left|\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk1}\right|^2\right] = E^2\left[\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk1}\right]+\text{var}\left[\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk1}\right], \tag{A3}$$

where the expectation and variance of $\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk1}$ are given by

$$E\left[\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk1}\right] = \frac{r_{jj,\text{DL}}}{M}E\left[\sum_{m=1}^{M}\left|\hat{h}_{jjkm}\right|^2\right], \tag{A4}$$

$$\text{var}\left[\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk1}\right] = \frac{r_{jj,\text{DL}}^2}{M^2}\text{var}\left[\sum_{m=1}^{M}\left|\hat{h}_{jjkm}\right|^2\right]. \tag{A5}$$

Since $h_{jjkm}$ is standard normal distribution and $\sum_{m=1}^{M}|\hat{h}_{jjkm}|^2$ is generalized Chi-square distribution, we get

$$E\left[\sum_{m=1}^{M}|\hat{h}_{jjkm}|^2\right] = \text{var}\left[\sum_{m=1}^{M}|\hat{h}_{jjkm}|^2\right] = M,$$

and

$$E\left[\left|\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk1}\right|^2\right] = r_{jj,\text{DL}}^2+\frac{r_{jj,\text{DL}}^2}{M}. \tag{A6}$$

In (A2), the term $E[|\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk2}|^2]$ is expressed as

$$\begin{aligned}E\left[\left|\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk2}\right|^2\right] &< \frac{r_{jj,\text{DL}}^2}{M^4}E\left(\left|\sum_{m=1}^{M}\left(h_{jjkm}\sum_{l=1}^{M}\left(\sum_{i=1}^{K}\|h_{jjim}\|^2\right)h_{jjkl}^*\right)\right|^2\right)\\ &= \frac{r_{jj,\text{DL}}^2}{M^4}E\left(\left|K\sum_{m=1}^{M}\left(h_{jjkm}\sum_{l=1}^{M}h_{jjkl}^*\right)\right|^2\right) = \frac{r_{jj,\text{DL}}^2K^2}{M^2}.\end{aligned} \tag{A7}$$

From (A6) and (A7), (A2) is derived further as

$$E\left[\left|\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk1}+\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk2}\right|^2\right] < r_{jj,\text{DL}}^2+\frac{r_{jj,\text{DL}}^2}{M}+\frac{r_{jj,\text{DL}}^2K^2}{M^2}. \tag{A8}$$

Referring to (6), the variance is given by

$$\text{var}\left[\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk1}+\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk2}\right] < \frac{r_{jj,\text{DL}}^2}{M}+\frac{r_{jj,\text{DL}}^2K^2}{M^2}. \tag{A9}$$

The power normalization factor $\lambda_j$ is given by

$$\begin{aligned}\lambda_j &= \frac{1}{E[\frac{1}{K}\text{tr}\{(\hat{\boldsymbol{A}}_1+\hat{\boldsymbol{A}}_2)(\hat{\boldsymbol{A}}_1+\hat{\boldsymbol{A}}_2)^{\text{H}}\}]} = \frac{1}{E[\frac{1}{K}\text{tr}\{\hat{\boldsymbol{A}}_1\hat{\boldsymbol{A}}_1^{\text{H}}\}]}\\ &= \frac{1}{\frac{1}{K}(\frac{r_{jj,\text{DL}}}{Md_{jj}r_{jj,\text{UL}}^2})^2r_{jj,\text{UL}}^2d_{jj}\sum_{m=1}^{M}E[\sum_{m=1}^{K}\|\hat{h}_{jjkm}\|^2]} = \frac{Md_{jj}r_{jj,\text{UL}}^2}{r_{jj,\text{DL}}^2}.\end{aligned} \tag{A10}$$

The approximate lower bound of SINR for the proposed scheme is given by

$$\begin{aligned}\rho_{jk}^{\text{DL}} &> \lim_{M arrow \infty}\frac{\frac{Md_{jj}r_{jj,\text{UL}}^2}{r_{jj,\text{DL}}^2}r_{jj,\text{DL}}^2}{1+\frac{Md_{jj}r_{jj,\text{UL}}^2}{r_{jj,\text{DL}}^2}(\frac{r_{jj,\text{DL}}^2}{M}+\frac{r_{jj,\text{DL}}^4K^2}{M^2})+\sum_{q\neq j}^{Q}\frac{Md_{qj}r_{qj,\text{UL}}^2}{r_{qj,\text{DL}}^2}(r_{qj,\text{DL}}^2+\frac{r_{qj,\text{DL}}^2}{M}+\frac{r_{qj,\text{DL}}^4K^2}{M^2})}\\ &= \frac{d_{jj}r_{jj,\text{UL}}^2}{\sum_{q\neq j}^{Q}d_{qj}r_{qj,\text{UL}}^2}.\end{aligned} \tag{A11}$$

Thus, Theorem 1 is proved.

## Appendix B Proof of Theorem 2

Since the single-cell MMSE precoding will approach the single-cell ZF precoding when SNR $\to \infty$ [23, 24], we evaluate the performance of ZF. $\hat{\boldsymbol{a}}_{jk}$ is defined as the $k$th column of $\hat{\boldsymbol{A}}_j^{\mathrm{ZF}}$. Considering H-inf CE in the uplink, the expectation and variance of $\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk}$ are given by

$$E\left[\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk}\right] = \frac{1}{Md_{jj}r_{jj,\mathrm{UL}}^2} d_{jj}r_{jj,\mathrm{UL}}^2 E\left[\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{h}}_{jjk}^{\mathrm{H}}\right] = 1, \tag{B1}$$

$$\mathrm{var}\left[\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk}\right] = \left(\frac{1}{Md_{jj}r_{jj,\mathrm{UL}}^2}\right)^2 d_{jj}^2 r_{jj,\mathrm{UL}}^4 \mathrm{var}\left[\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{h}}_{jjk}^{\mathrm{H}}\right] = \frac{1}{M}, \tag{B2}$$

and

$$E\left[\left|\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk}\right|^2\right] = E^2\left[\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk}\right] + \mathrm{var}\left[\hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{a}}_{jk}\right] = 1 + \frac{1}{M}. \tag{B3}$$

In addition, the power normalization factor $\lambda_j$ is given by

$$\begin{aligned}\lambda_j &= \frac{1}{E[\frac{1}{K}tr\{\hat{\boldsymbol{A}}_j^{\mathrm{ZF}}(\hat{\boldsymbol{A}}_j^{\mathrm{ZF}})^{\mathrm{H}}\}]} = \frac{1}{\frac{1}{K}(\frac{1}{Md_{jj}r_{jj,\mathrm{UL}}^2})^2 r_{jj,\mathrm{UL}}^2 d_{jj}\sum_{m=1}^{M} E[\sum_{m=1}^{K}\|\hat{h}_{jjkm}\|^2]} \\ &= Md_{jj}r_{jj,\mathrm{UL}}^2.\end{aligned} \tag{B4}$$

The asymptotically lower bound of SINR by using single-cell ZF and MMSE precoding is given by

$$\begin{aligned}\lim_{\mathrm{SNR}\to\infty} \rho_{jk}^{\text{ZF-DL}}, \rho_{jk}^{\text{MMSE-DL}} &= \lim_{\mathrm{SNR}\to\infty} \frac{Md_{jj}r_{jj,\mathrm{UL}}^2}{1 + Md_{jj}r_{jj,\mathrm{UL}}^2\frac{1}{M} + \sum_{q\neq j}^{Q} Md_{qj}r_{qj,\mathrm{UL}}^2(1+\frac{1}{M})} \\ &= \frac{d_{jj}r_{jj,\mathrm{UL}}^2}{\sum_{q\neq j}^{Q} d_{qj}r_{qj,\mathrm{UL}}^2}.\end{aligned} \tag{B5}$$

Thus, Theorem 2 is proved.