TDD reciprocity calibration for multi-user massive MIMO systems with iterative coordinate descent

Hao WEI[1], Dongming WANG[1*], Jiangzhou WANG[2] & Xiaohu YOU[1]

[1]*National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China;*
[2]*School of Engineering and Digital Arts, University of Kent, Canterbury, United Kingdom*

Abstract For massive multiple-input multiple-output (MIMO) antenna systems, time division duplexing (TDD) is preferred since the downlink precoding matrix can be obtained through the uplink channel estimation, thanks to the channel reciprocity. However, the mismatches of the transceiver radio frequency (RF) circuits at both sides of the link make the whole communication channel non-symmetric. This paper extends the total least square (TLS) method to the case of self-calibration, where only the antennas of the access points (APs) are involved to exchange the calibration signals with each other and the feedback from the user equipments (UEs) is not required. Then, the proof of the equivalence between the TLS method and the least square (LS) method is presented. Furthermore, to avoid the eigenvalue decomposition required by these two methods to obtain the calibration coefficients, a novel algorithm named as iterative coordinate descent (ICD) method is proposed. Theoretical analysis and simulation results show that the ICD method significantly reduces the complexity and achieves almost the same performance of the LS method.

Keywords TDD, reciprocity calibration, total least squares, coordinate descent, large-scale MIMO

1 Introduction

As promising techniques for the next generation mobile communication systems, massive multiple-input multiple-output (MIMO) [1, 2] and distributed MIMO [3–5] are proposed to greatly improve the spectral efficiency by simultaneously serving multiple users in the same time-frequency resource. The uplink achievable rate of a single-cell multi-user massive MIMO system has been analyzed in [6], which showed that circularly distributed massive MIMO system largely outperforms centralized massive MIMO system. In time division duplexing (TDD) operation, the transmitter can enable the multi-user joint precoding according to the estimation of the uplink channel state information (CSI) thanks to the channel reciprocity. Unfortunately, the transceiver radio frequency (RF) circuits at both sides of the link are not symmetric [7]. The mismatches disable the reciprocity of the whole communication and lead to a severe degradation in the system performance.

Recently, to avoid introducing extra calibration circuits, some methods have been proposed where the calibration procedure entirely takes place in the signal space. In [8], total least square (TLS) based

---

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

calibration was proposed for MIMO systems. However, this method is based on exchanging the calibration signals between the transmitter and the receiver, which causes large feedback overhead for CSI in massive MIMO systems [9]. Ref. [10] showed that we only need to perform calibration at the transmitter since the RF mismatches at the user equipment (UE) have a negligible impact on the system performance. In order to reduce the feedback, the over-the-air method was presented in [11], where only one or several UEs with good channel condition are chosen to estimate and feed back the downlink CSI to the base stations. Furthermore, to exclude UEs from the calibration procedure, a calibration method, referred to as Argos method, was presented in [12] for massive MIMO systems. The Argos method only involves the antennas of the base station and exchanges the calibration signals with a reference antenna. Similar to the Argos method, the master-slave protocol was proposed for distributed MIMO systems to realize the synchronization [13] and perform the calibration [14] in all access points (APs). All other APs are required to receive a beacon signal and exchange calibration signals with the master/reference AP. However, the Argos method and the master-slave protocol are very sensitive to the placement of the reference antenna/AP, and the system performance cannot keep stable unless all other antennas/APs have large signal-to-noise ratio (SNR) to the reference antenna/AP. To generalize the Argos method, an elegant solution named LS method was devised in [15], which defines the calibration problem as a LS cost function. The LS method makes use of the calibration pilot signals of all antennas and achieves essentially the performance of the perfect calibration.

In this paper, we extend the TLS method to the case of self-calibration. The feedback from UEs is not required, and only the APs are involved in exchanging calibration signals. Then, we give the proof that the TLS method is equivalent to the LS method. However, the optimal solutions of these two methods are based on singular value decomposition (SVD) or eigenvalue decomposition (EVD), and the computational complexity becomes very high for large-scale MIMO systems. Therefore, we propose an iterative algorithm named as iterative coordinate descent (ICD) method, of which the complexity is significantly reduced and the performance is very close to the perfect calibration.

The notation adopted in this paper conforms to the following convention. Vectors are denoted in lower case bold: $\boldsymbol{x}$. Matrices are upper case bold: $\boldsymbol{A}$. $[\boldsymbol{A}]_{ij}$ denotes the ith row jth column element of $\boldsymbol{A}$. $(\cdot)^*$, $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$ represent conjugate, transpose, and Hermitian transpose, respectively. $\mathrm{Tr}\,(\boldsymbol{A})$ denotes the trace of $\boldsymbol{A}$. $\mathrm{diag}\,(\boldsymbol{x})$ is a diagonal matrix with $\boldsymbol{x}$ on its diagonal, and $\mathrm{diag}\,(\boldsymbol{A})$ denotes a column vector with the main diagonal of $\boldsymbol{A}$.

2 System model and fundamentals

In this paper, we assume that APs and UEs are the transmitters and receivers respectively in the downlink transmission. The antennas involved in the calibration may be co-located at one AP or distributed in different APs, which are corresponding to co-located MIMO or distributed MIMO respectively.

2.1 Mismatch model of transceiver RF circuits

In practice, the whole communication channel consists of not only the wireless propagation part, but also the transceiver RF circuits of the antennas at both sides of the link. As shown in Figure 1, each antenna has a transmitting RF and a receiving RF module. It is assumed that there are M antennas in total at the APs and K antennas at the UEs. Thus, the overall wireless channel matrix is $\boldsymbol{H} \in \mathbb{C}^{K \times M}$, and the uplink and downlink channel matrices are characterized as

$$\boldsymbol{G}_{\mathrm{UL}} = \boldsymbol{C}_{\mathrm{AP,r}} \boldsymbol{H}^{\mathrm{T}} \boldsymbol{C}_{\mathrm{UE,t}}, \tag{1}$$

$$\boldsymbol{G}_{\mathrm{DL}} = \boldsymbol{C}_{\mathrm{UE,r}} \boldsymbol{H} \boldsymbol{C}_{\mathrm{AP,t}}, \tag{2}$$

where $\boldsymbol{C}_{\mathrm{AP,t}}$ and $\boldsymbol{C}_{\mathrm{AP,r}}$ denote the transmitting and receiving RF gain matrices of the APs respectively. $\boldsymbol{C}_{\mathrm{UE,t}}$ and $\boldsymbol{C}_{\mathrm{UE,r}}$ denote the transmitting and receiving RF gain matrices of the UEs respectively. All of
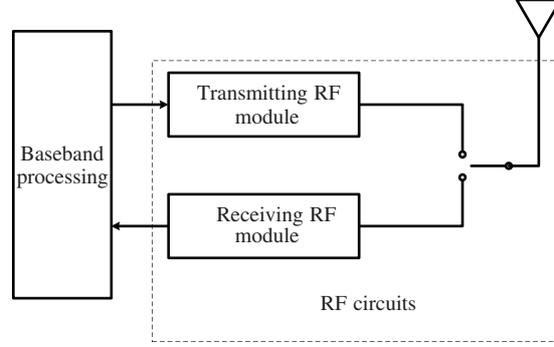
Figure 1 RF circuits include antennas, mixers, filters, analog to digital (A/D) converters, power amplifiers, etc., and are highly related to the temperature and humidity of the environment [8].

these matrices are diagonal. Define

$$\boldsymbol{C}_{\mathrm{AP,t}} = \mathrm{diag}\left(t_{\mathrm{AP},1}, \ldots, t_{\mathrm{AP},m}, \ldots, t_{\mathrm{AP},M}\right),$$
$$\boldsymbol{C}_{\mathrm{AP,r}} = \mathrm{diag}\left(r_{\mathrm{AP},1}, \ldots, r_{\mathrm{AP},m}, \ldots, r_{\mathrm{AP},M}\right),$$
$$\boldsymbol{C}_{\mathrm{UE,t}} = \mathrm{diag}\left(t_{\mathrm{UE},1}, \ldots, t_{\mathrm{UE},k}, \ldots, t_{\mathrm{UE},K}\right),$$
$$\boldsymbol{C}_{\mathrm{UE,r}} = \mathrm{diag}\left(r_{\mathrm{UE},1}, \ldots, r_{\mathrm{UE},k}, \ldots, r_{\mathrm{UE},K}\right),$$

where $t_{\mathrm{AP},m}$, $r_{\mathrm{AP},m}$ $(m = 1, \ldots, M)$ and $t_{\mathrm{UE},k}$, $r_{\mathrm{UE},k}$ $(k = 1, \ldots, K)$ are the RF gains of the APs and the UEs, respectively. Note that the whole communication channel becomes non-reciprocal due to the mismatches of RF gains, i.e. $\boldsymbol{G}_{\mathrm{DL}} \neq \boldsymbol{G}_{\mathrm{UL}}^{\mathrm{T}}$.

2.2 Downlink signal model

Given the uplink CSI for the TDD operation, the multi-user precoding matrix can be designed for downlink transmission. Considering zero-forcing precoding, the overall downlink received signals at the UEs are written as

$$\boldsymbol{y} = \beta \boldsymbol{G}_{\mathrm{DL}} \boldsymbol{G}_{\mathrm{UL}}^{*} \left(\boldsymbol{G}_{\mathrm{UL}}^{\mathrm{T}} \boldsymbol{G}_{\mathrm{UL}}^{*}\right)^{-1} \boldsymbol{x} + \boldsymbol{n}, \tag{3}$$

where $\beta = \sqrt{1/\mathrm{Tr}[(\boldsymbol{G}_{\mathrm{UL}}^{\mathrm{T}} \boldsymbol{G}_{\mathrm{UL}}^{*})^{-1}]}$ is the scaling factor to satisfy the transmit power constraint, $\boldsymbol{y} = [y_1, \ldots, y_K]^{\mathrm{T}}$ is the receiving signal vector, $\boldsymbol{x} = [x_1, \ldots, x_K]^{\mathrm{T}}$ is the signal vector transmitted to the UEs with the power constraint $\mathcal{E}[x_k x_k^*] = P$, and $\boldsymbol{n}$ is the complex additive white Gaussian noise (AWGN) vector, in which the elements are independent and identically distributed (i.i.d.) complex Gaussian random variables with zero mean and variance σ_{n}^2.

Substituting $\boldsymbol{G}_{\mathrm{UL}}$ and $\boldsymbol{G}_{\mathrm{DL}}$ by (1) and (2) respectively, Eq. (3) is given by

$$\boldsymbol{y} = \beta \boldsymbol{C}_{\mathrm{UE,r}} \boldsymbol{W} \boldsymbol{C}_{\mathrm{UE,t}}^{-1} \boldsymbol{x} + \boldsymbol{n}, \tag{4}$$

where

$$\boldsymbol{W} = \left(\boldsymbol{H} \boldsymbol{C}_{\mathrm{AP,t}} \boldsymbol{C}_{\mathrm{AP,r}}^{*} \boldsymbol{H}^{\mathrm{H}}\right) \left(\boldsymbol{H} \boldsymbol{C}_{\mathrm{AP,r}} \boldsymbol{C}_{\mathrm{AP,r}}^{*} \boldsymbol{H}^{\mathrm{H}}\right)^{-1}. \tag{5}$$

Since $\boldsymbol{C}_{\mathrm{AP,t}} \neq \boldsymbol{C}_{\mathrm{AP,r}}$, $\boldsymbol{W}$ is not an identity matrix. From (4), it can be seen that the non-symmetric characteristic of the transceiver RF circuits at the APs will cause the multi-user interference. Hence, the received signal of the ith UE is

$$y_i = \beta [\boldsymbol{W}]_{ii} x_i + \beta \sum_{j=1, j \neq i}^{K} [\boldsymbol{W}]_{ij} x_j + n_i, \tag{6}$$

and the SINR of the ith UE is

$$\gamma_i^{\mathrm{mis}} = \frac{\rho \cdot \beta^2 \cdot |[\boldsymbol{W}]_{ii}|^2}{\rho \cdot \beta^2 \cdot \sum_{j=1, j \neq i}^{K} |[\boldsymbol{W}]_{ij}|^2 + 1}, \tag{7}$$

where $\rho = P/\sigma_{\rm n}^2$. Then the sum-rates of all UEs with RF mismatches are

$$R^{\rm mis} = \sum_{i=1}^{K} R_i^{\rm mis} = \sum_{i=1}^{K} \log\left(1 + \gamma_i^{\rm mis}\right). \tag{8}$$

2.3 Reciprocity calibration

Multiply the precoding matrix by a diagonal calibration matrix $\boldsymbol{C}_{\rm cal}$ on the left, which satisfies

$$\boldsymbol{C}_{\rm cal} = \alpha_{\rm cal} \boldsymbol{C}_{\rm AP,t}^{-1} \boldsymbol{C}_{\rm AP,r}, \tag{9}$$

and we can rewrite (3) as

$$\boldsymbol{y} = \beta_{\rm cal} \boldsymbol{G}_{\rm DL} \boldsymbol{C}_{\rm cal} \boldsymbol{G}_{\rm UL}^* \left(\boldsymbol{G}_{\rm UL}^{\rm T} \boldsymbol{G}_{\rm UL}^*\right)^{-1} \boldsymbol{x} + \boldsymbol{n}, \tag{10}$$

where $\beta_{\rm cal}$ is the scaling factor to satisfy the transmit power constraint with calibration. Therefore, the received signals at the UEs with perfect calibration are written as

$$\boldsymbol{y} = \beta_{\rm cal} \boldsymbol{C}_{\rm UE,r} \boldsymbol{C}_{\rm UE,t}^{-1} \boldsymbol{x} + \boldsymbol{n}. \tag{11}$$

According to (11), the multi-user interference caused by the RF mismatches at the APs can be eliminated through the calibration matrix $\boldsymbol{C}_{\rm cal}$.

3 TLS method for the self-calibration

TLS method was proposed in [8] to formalize the non-reciprocity model into a TLS problem. However, this method is based on exchanging calibration signals between the transmitters and receivers. To avoid involving the UEs in the calibration process, we extend TLS method to the case of self-calibration, where the feedback from UEs is not required, and only the APs are involved in exchanging calibration signals.

In the calibration procedure, each antenna transmits a time-orthogonal pilot signal with power $P_{\rm cal}$ in sequence to the other antennas of the APs. The observation matrix is given by

$$\boldsymbol{Y}_{\rm cal} = \boldsymbol{C}_{\rm AP,r} \boldsymbol{H}_{\rm cal} \boldsymbol{C}_{\rm AP,t} + \boldsymbol{N}, \tag{12}$$

where

$$[\boldsymbol{Y}_{\rm cal}]_{m,n} = \begin{cases} r_{{\rm AP},m} [\boldsymbol{H}_{\rm cal}]_{m,n} t_{{\rm AP},n} + [\boldsymbol{N}]_{m,n} & m \neq n, \\ 0 & m = n \end{cases} \tag{13}$$

is the calibration signal which is transmitted from the nth antenna of the APs and received at the mth antenna of the APs. $[\boldsymbol{N}]_{m,n}$ is the equivalent AWGN with zero mean and variance $1/\rho_{\rm cal}$, where $\rho_{\rm cal} = P_{\rm cal}/\sigma_{\rm n}^2$ and $\sigma_{\rm n}^2$ are the power of thermal noise. Since calibration pilots are time-orthogonal, increasing the length of the calibration pilot is equal to increasing the SNR of the calibration pilot. $\boldsymbol{H}_{\rm cal} \in \mathbb{C}^{M \times M}$ is the calibration channel matrix, $[\boldsymbol{H}_{\rm cal}]_{m,n}$ denotes the channel coefficient between the nth antenna of the APs and the mth antenna of the APs, and $[\boldsymbol{H}_{\rm cal}]_{m,n} = 0$ when $m = n$. We have $\boldsymbol{H}_{\rm cal} = \boldsymbol{H}_{\rm cal}^{\rm T}$ because of the wireless channel reciprocity. Then, if the observation matrix $\boldsymbol{Y}_{\rm cal}$ is noiseless, we have

$$\boldsymbol{C}_{\rm AP,r}^{-1} \boldsymbol{Y}_{\rm cal} \boldsymbol{C}_{\rm AP,t}^{-1} = \boldsymbol{C}_{\rm AP,t}^{-1} \boldsymbol{Y}_{\rm cal}^{\rm T} \boldsymbol{C}_{\rm AP,r}^{-1}. \tag{14}$$

From (9), we have $\boldsymbol{C}_{\rm cal} = \alpha_{\rm cal} \boldsymbol{C}_{\rm AP,t}^{-1} \boldsymbol{C}_{\rm AP,r}$. Eq. (14) reduces to

$$\boldsymbol{Y}_{\rm cal} \boldsymbol{C}_{\rm cal} = \boldsymbol{C}_{\rm cal} \boldsymbol{Y}_{\rm cal}^{\rm T}. \tag{15}$$

Then, in the presence of observation noise, $\boldsymbol{C}_{\rm cal}$ can be obtained by solving an optimization problem as follows

$$\underset{\{\boldsymbol{C}_{\rm cal}\}}{\arg\min} \left\| \boldsymbol{Y}_{\rm cal} \boldsymbol{C}_{\rm cal} - \boldsymbol{C}_{\rm cal} \boldsymbol{Y}_{\rm cal}^{\rm T} \right\|_{\rm F}^2. \tag{16}$$

Eq. (16) can be also shown as

$$\arg\min_{\{\boldsymbol{C}_{\text{cal}}\}} \left\| \text{vec}\left(\boldsymbol{Y}_{\text{cal}}\boldsymbol{C}_{\text{cal}}\right) - \text{vec}\left(\boldsymbol{C}_{\text{cal}}\boldsymbol{Y}_{\text{cal}}^{\text{T}}\right) \right\|^2. \tag{17}$$

According to the computation principle of matrix

$$\text{vec}\left(\boldsymbol{A}_{M\times N}\boldsymbol{B}_{N\times P}\right) = \left(\boldsymbol{I}_P \otimes \boldsymbol{A}\right)\text{vec}\left(\boldsymbol{B}\right) = \left(\boldsymbol{B}^{\text{T}} \otimes \boldsymbol{I}_M\right)\text{vec}\left(\boldsymbol{A}\right), \tag{18}$$

one obtains

$$\text{vec}\left(\boldsymbol{Y}_{\text{cal}}\boldsymbol{C}_{\text{cal}}\right) - \text{vec}\left(\boldsymbol{C}_{\text{cal}}\boldsymbol{Y}_{\text{cal}}^{\text{T}}\right) = \left(\boldsymbol{I}_M \otimes \boldsymbol{Y}_{\text{cal}} - \boldsymbol{Y}_{\text{cal}} \otimes \boldsymbol{I}_M\right)\text{vec}\left(\boldsymbol{C}_{\text{cal}}\right). \tag{19}$$

Defining the calibration coefficient vector $\boldsymbol{c}_{\text{cal}} = \text{diag}\left(\boldsymbol{C}_{\text{cal}}\right) = [c_1, \ldots, c_m, \ldots, c_M]^{\text{T}}$, and simplifying (19), Eq. (17) can be written as

$$\arg\min_{\{\boldsymbol{c}_{\text{cal}}\}} \left\| \boldsymbol{\Omega}\boldsymbol{c}_{\text{cal}} \right\|^2, \tag{20}$$

where $\boldsymbol{\Omega}$ is an $[M(M-1)/2] \times M$ matrix,

$$\boldsymbol{\Omega} = \left[\boldsymbol{\Phi}_1^{\text{T}} \cdots \boldsymbol{\Phi}_i^{\text{T}} \cdots \boldsymbol{\Phi}_{M-1}^{\text{T}}\right]^{\text{T}}, \tag{21}$$

and $\boldsymbol{\Phi}_i$ is an $i \times M$ matrix,

$$[\boldsymbol{\Phi}_i]_{u,v} = \begin{cases} -[\boldsymbol{Y}_{\text{cal}}]_{u,i+1} & v = i+1, \\ [\boldsymbol{Y}_{\text{cal}}]_{i,u} & u = v, \\ 0 & \text{others.} \end{cases} \tag{22}$$

Consequently, we introduce a model perturbation on $\boldsymbol{\Omega}$ because of the thermal noise. Then, we formulate (20) into the following TLS problem [8]

$$\arg\min_{\{\Delta\boldsymbol{\Omega}, \boldsymbol{c}_{\text{cal}}\}} \quad \left\| \Delta\boldsymbol{\Omega} \right\|_{\text{F}} \quad \text{s.t.} \quad \left(\boldsymbol{\Omega} + \Delta\boldsymbol{\Omega}\right)\boldsymbol{c}_{\text{cal}} = \boldsymbol{0}, \tag{23}$$

where $\Delta\boldsymbol{\Omega}$ is the correction term of the TLS optimization problem. Given the SVD of $\boldsymbol{\Omega}$,

$$\boldsymbol{\Omega} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^{\text{H}},$$

the estimated solution for $\boldsymbol{c}_{\text{cal}}$ lies in the last column of $\boldsymbol{V}$ corresponding to the smallest singular value in $\boldsymbol{D}$ [16]. Further, we define the matrix $\boldsymbol{\Psi} = \boldsymbol{\Omega}^{\text{H}}\boldsymbol{\Omega}$, the element of which is given by

$$[\boldsymbol{\Psi}]_{u,v} = \begin{cases} \sum_{i=1,i\neq u}^{M} \left|[\boldsymbol{Y}_{\text{cal}}]_{i,u}\right|^2 & u = v, \\ -[\boldsymbol{Y}_{\text{cal}}]_{v,u}^* [\boldsymbol{Y}_{\text{cal}}]_{u,v} & u \neq v. \end{cases} \tag{24}$$

Then, [16,17] proved that the objective function in (23) is also equivalent to the formula as follows:

$$\min \quad f\left(\boldsymbol{c}_{\text{cal}}\right) = \frac{\boldsymbol{c}_{\text{cal}}^{\text{H}}\boldsymbol{\Psi}\boldsymbol{c}_{\text{cal}}}{\boldsymbol{c}_{\text{cal}}^{\text{H}}\boldsymbol{c}_{\text{cal}}}. \tag{25}$$

According to the principle of Hermite matrix, we have

$$\lambda_{\min} \leqslant f\left(\boldsymbol{c}_{\text{cal}}\right) \leqslant \lambda_{\max}, \tag{26}$$

where $\lambda_{\min}$ and $\lambda_{\max}$ denote the minimum and maximum eigenvalue of $\boldsymbol{\Psi}$, respectively. Thus, the solution for $\boldsymbol{c}_{\text{cal}}$ is the eigenvector of $\boldsymbol{\Psi}$ corresponding to $\lambda_{\min}$.

By expanding $\boldsymbol{c}_{\text{cal}}^{\text{H}}\boldsymbol{\Psi}\boldsymbol{c}_{\text{cal}}$, we can turn (25) into the following equivalent constrained optimization problem as

$$\min \quad g\left(\boldsymbol{c}_{\text{cal}}\right) = \boldsymbol{c}_{\text{cal}}^{\text{H}}\boldsymbol{\Psi}\boldsymbol{c}_{\text{cal}} = \sum_{i,j=1}^{M} \left| c_i[\boldsymbol{Y}_{\text{cal}}]_{j,i} - c_j[\boldsymbol{Y}_{\text{cal}}]_{i,j} \right|^2 \quad \text{s.t.} \quad \boldsymbol{c}_{\text{cal}}^{\text{H}}\boldsymbol{c}_{\text{cal}} = 1. \tag{27}$$

Recently, a novel calibration algorithm named as LS method was presented in [15]. Coincidentally, the optimization problem in [15] is the same as (27), and the solution of $\boldsymbol{c}_{\text{cal}}$ is also the eigenvector corresponding to $\lambda_{\min}$ of $\boldsymbol{\Psi}$. Thus, the TLS method for the self-calibration is consistent with the LS method.

4 Iterative coordinate descent method

Since the TLS method and the LS method need to perform SVD or EVD operation, the computational complexity will become unacceptable for large-scale MIMO systems. Some iterative algorithms, including inverse power method and Rayleigh-Ritz method [16], were proposed to find the minimum eigenvalue $\lambda_{\min}$ and its corresponding eigenvector. However, the complexity of these algorithms is still very high, which is in the order of $\mathcal{O}\left(M^3\right)$ because of the matrix inversion operation in each iteration.

In this section, to avoid the matrix inversion, we propose a new iterative algorithm to reduce the computation complexity, which is named as ICD method. Differentiating $g\left(\boldsymbol{c}_{\mathrm{cal}}\right)$ with respect to c_i^* in (27) and treating c_i^* as a variable independent on c_i, one obtains

$$\frac{\partial g\left(\boldsymbol{c}_{\mathrm{cal}}\right)}{\partial c_i^*} = c_i \sum_{j=1,j\neq i}^{M} \left|[\boldsymbol{Y}_{\mathrm{cal}}]_{j,i}\right|^2 - \sum_{j=1,j\neq i}^{M} c_j\,[\boldsymbol{Y}_{\mathrm{cal}}]_{j,i}^*\,[\boldsymbol{Y}_{\mathrm{cal}}]_{i,j}. \tag{28}$$

Then, by setting the partial derivatives to zero, we have

$$c_i = \frac{\sum_{j=1,j\neq i}^{M} c_j\,[\boldsymbol{Y}_{\mathrm{cal}}]_{j,i}^*\,[\boldsymbol{Y}_{\mathrm{cal}}]_{i,j}}{\sum_{j=1,j\neq i}^{M} \left|[\boldsymbol{Y}_{\mathrm{cal}}]_{j,i}\right|^2}. \tag{29}$$

The ICD method is described as Algorithm 1. According to the principle of coordinate descent [18] and (29), by fixing other $(M-1)$ coefficients, the optimal value of the current coefficient c_i makes the objective function $g\left(\boldsymbol{c}_{\mathrm{cal}}\right)$ achieve a local minimum. Then, this operation is carried out sequently for every calibration coefficient in each iteration. The complexity of each iteration is only in the order of $\mathcal{O}\left(M^2\right)$. Finally, since $g\left(\boldsymbol{c}_{\mathrm{cal}}\right)$ is a convex function, the optimal $\boldsymbol{c}_{\mathrm{cal}}$ can be obtained when the method reaches convergence [18, 19]. However, without the norm constraint $\boldsymbol{c}_{\mathrm{cal}}^{\mathrm{H}}\boldsymbol{c}_{\mathrm{cal}} = 1$, the minimization of the objective function $g\left(\boldsymbol{c}_{\mathrm{cal}}\right)$ will only give a solution of all-zero vector to $\boldsymbol{c}_{\mathrm{cal}}$. Thus, in order to avoid the convergence to the all-zero solution, we use the value of $f\left(\boldsymbol{c}_{\mathrm{cal}}\right)$ in (25) as the condition to decide when to break the iteration. According to the simulation results in the next section, the value of $f\left(\boldsymbol{c}_{\mathrm{cal}}\right)$ converges to a limit which is larger than but very close to $\lambda_{\min}$ when ρ_{cal} is in the moderate and high region. Therefore, although the value of $g\left(\boldsymbol{c}_{\mathrm{cal}}\right)$ declines to zero with the iterations, we can break the iteration according to the converged value of $f\left(\boldsymbol{c}_{\mathrm{cal}}\right)$ to obtain the approximated solution of $\boldsymbol{c}_{\mathrm{cal}}$.

Note that, besides the low complexity, ICD method has another advantage. If the number of antennas at the APs decreases or increases (which can be seen in dynamical clustering for distributed antenna systems [20]), we can obtain the new calibration coefficients rapidly by using the old ones as the initial values. Then, the following two cases are presented.

Case 1: The number of antennas at the APs decreases. We assume that the Mth antenna is removed from the APs. Then, after eliminating c_M and its related calibration pilot signals $[\boldsymbol{Y}_{\mathrm{cal}}]_{M,i}, [\boldsymbol{Y}_{\mathrm{cal}}]_{i,M}$, $(i = 1 \ldots M-1)$, we can obtain the rest $(M-1)$ new coefficients by taking the current $c_i\,(i = 1, \ldots, M-1)$ into the iteration.

Case 2: The number of antennas at the APs increases. We assume that the $(M+1)$th antenna is added into the APs. Then, the initial value of c_{M+1} can be set by the added observation signals and the current calibration coefficients $c_i\,(i = 1, \ldots, M)$, which is written as

$$c_{M+1} = \frac{\sum_{j=1}^{M} c_j\,[\boldsymbol{Y}_{\mathrm{cal}}]_{j,M+1}^*\,[\boldsymbol{Y}_{\mathrm{cal}}]_{M+1,j}}{\sum_{j=1}^{M} \left|[\boldsymbol{Y}_{\mathrm{cal}}]_{j,M+1}\right|^2}. \tag{30}$$

Then, we can obtain all the $(M+1)$ new coefficients by taking the current $c_i\,(i = 1, \ldots, M+1)$ into the iteration.

Algorithm 1 ICD (iterative coordinate descent) method

Requre:
   1. $\boldsymbol{Y}_{\text{cal}}$

Initialization:
   1. $\boldsymbol{c}_{\text{curr}} = [c_1, \ldots c_i, \ldots c_M]^{\text{T}}, c_i = 1, (i = 1 \ldots M)$
   To record the current calibration coefficients in each iteration.
   2. $\boldsymbol{c}_{\text{prev}} = \boldsymbol{c}_{\text{curr}}$
   To record the previous calibration coefficients in each iteration.
   3. $e = +\infty, e_{\text{thres}}$
   To calculate the calibration error between the current and previous step iteration. To determine whether to break the iteration.
   4. $n_{\text{iter}} = 0, n_{\text{iter\_max}}$
   To count the number of the iteration times. To determine whether to break the iteration.

Iteration:
   while $e > e_{\text{thres}}$ **and** $n_{\text{iter}} < n_{\text{itr\_max}}$ **do**
   1. $\lambda_{\text{prev}} = f(\boldsymbol{c}_{\text{prev}})$ according to (25)
   2. **for** $i = 1 : M$
$$c_i = \frac{\sum_{j=1, j \neq i}^{M} c_j [\boldsymbol{Y}_{\text{cal}}]_{j,i}^* [\boldsymbol{Y}_{\text{cal}}]_{i,j}}{\sum_{j=1, j \neq i}^{M} \left| [\boldsymbol{Y}_{\text{cal}}]_{j,i} \right|^2}$$
   end for
   3. $\lambda_{\text{curr}} = f(\boldsymbol{c}_{\text{curr}})$ according to (25)
   4. $e = |\lambda_{\text{prev}} - \lambda_{\text{curr}}|^2 / |\lambda_{\text{curr}}|^2$
   5. $\boldsymbol{c}_{\text{prev}} = \boldsymbol{c}_{\text{curr}}$
   6. $n_{\text{iter}} = n_{\text{iter}} + 1$
   end while

Ending:
   1. $\boldsymbol{c}_{\text{cal}} = \boldsymbol{c}_{\text{curr}}$
   To obtain the calibration coefficients vector.

5　Simulation results

In this section, system simulations have been carried out to illustrate the performance of our proposed calibration method. A large-scale MIMO system is considered, where there are M antennas in total at the APs and K antennas at the UEs. For clarity and brevity of discussion, both the transmission and calibration channel are supposed to be with a fading following a Rayleigh distribution, which is a zero mean circularly symmetric complex Gaussian random variables of variance $1/2$ per dimension. The amplitudes of the RF gains are assumed to be of log-normal distribution, and the phases are assumed to be of uniform distribution. The variance of the amplitude mismatches and the range of the phase mismatches are set to be $\delta_{\text{AP}}^2 = 2$ dB and $\theta_{\text{AP}} = \pi/2$, respectively.

5.1　Convergence analysis

According to (25), λ is defined as the value of $f(\boldsymbol{c}_{\text{cal}})$. Further, λ_{Argos}, λ_{LS} and λ_{ICD} are defined as the values of $f(\boldsymbol{c}_{\text{cal}})$ corresponding to the Argos, LS and ICD method, respectively. Actually, the goal of the calibration methods is to minimize λ. Thus, we compare λ for these three methods. Figure 2 shows convergence behavior of the calibration methods for the minimum λ. It can be seen that, ICD method converges quickly and λ_{ICD} reaches close to λ_{LS} in less than five iterations for all the configurations ($M = 8, 64, 128$). Figure 3 illustrates the impact of M on λ. It can be seen that basically the convergence λ increases as M increases. λ_{LS} and λ_{ICD} are much smaller than λ_{Argos}. In the Argos method, the calibration coefficients are given by [12]

$$c_i = \begin{cases} 1 & i = n, \\ \dfrac{[\boldsymbol{Y}_{\text{cal}}]_{i,n}}{[\boldsymbol{Y}_{\text{cal}}]_{n,i}} & i \neq n, \end{cases} \tag{31}$$

where n denotes the reference antenna. This is equivalent to minimizing $g(\boldsymbol{c}_{\text{cal}})$ subject to the constraint $c_n = 1$ [15]. When the calibration channels from some antennas to the reference antenna are not good
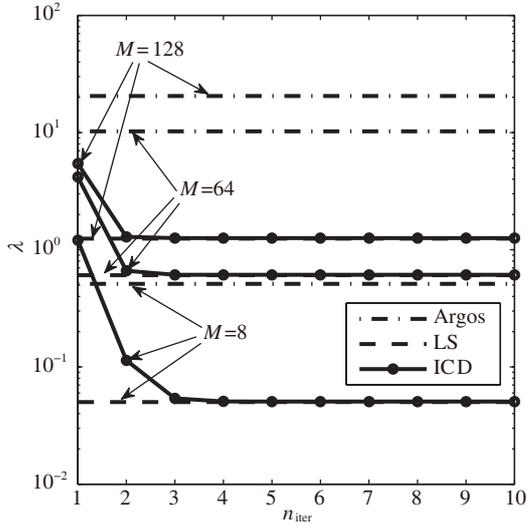
Figure 2 Convergence behavior for the minimum λ. The calibration SNR is set to be $\rho_{\text{cal}} = 20$ dB, and the number of the antennas at APs is set to be $M = 8$, $M = 64$ and $M = 128$, respectively.
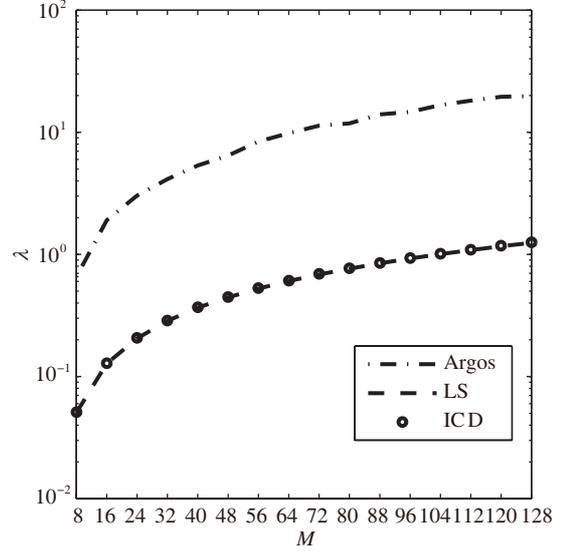
Figure 3 Impact of M on λ. The calibration SNR is set to be $\rho_{\text{cal}} = 20$ dB, and the threshold to break the iteration is $e_{\text{thres}} = 10^{-4}$.
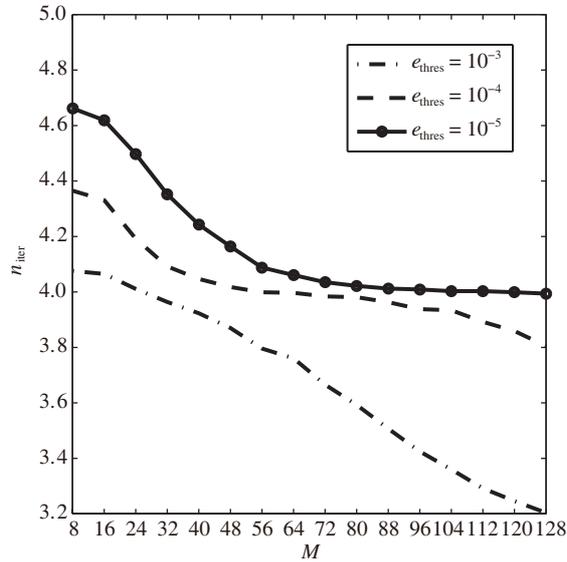
Figure 4 Impact of M on n_{iter}. The calibration SNR is set to be $\rho_{\text{cal}} = 20$ dB, and the threshold to break the iteration is set to be $e_{\text{thres}} = 10^{-3}, 10^{-4}, 10^{-5}$, respectively.

enough, the estimation error of their calibration coefficients may be very large, which will cause significant degradation on the system performance.

Figure 4 depicts the relation between n_{iter} and M. According to the description in Algorithm 1, the total complexity of ICD method is $\left[n_{\text{iter}} \cdot \mathcal{O} \left(M^2 \right) \right]$. It can be seen from Figure 4 that n_{iter} decreases slowly when M increases, and smaller e_{thres} leads to larger n_{iter}. Since n_{iter} is basically less than five, the ICD method reaches convergence very quickly, and the total complexity is much lower than the LS method especially when M is very large.

5.2 Sum-rates of the system for downlink transmission

Figure 5 shows the relation between sum-rates and ρ_{cal} when $M = 128$ and $K = 16$. It can be seen from
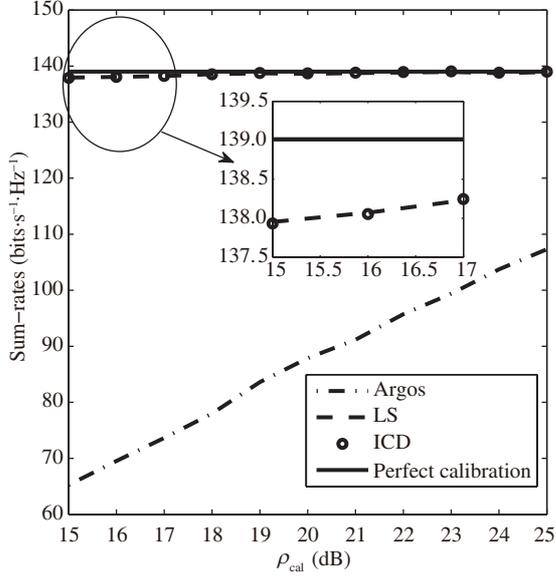
Figure 5 The relation between sum-rates and ρ_{cal}. The average SNR of transmission signals is set to be $\rho = 20$ dB. The number of antennas at APs is $M = 128$, and the number of antennas at UEs is $K = 16$.
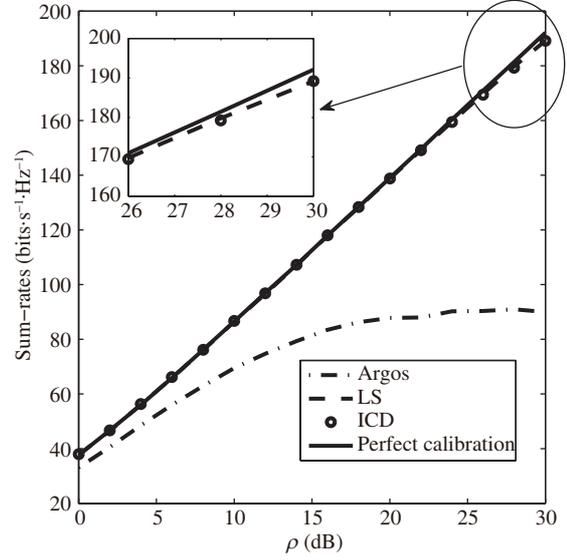
Figure 6 The relation between sum-rates and ρ. The average SNR of calibration pilot signals is set to be $\rho_{\text{cal}} = 20$ dB. The number of antennas at APs is $M = 128$, and the number of antennas at UEs is $K = 16$.

Figure 5 that the performances of the LS method and the ICD method are very close to the perfect calibration, and are much better than the Argos method.

Figure 6 illustrates sum-rates performance as a function of ρ when $M = 128$ and $K = 16$. It can be seen that due to the residual interference introduced by the imperfect calibration, there is a ceiling effect with the Argos method. However, the performances of the LS method and the ICD method are very close to the perfect calibration, for all values of ρ.

Note that the Argos method has very low complexity ($\mathcal{O}(M)$). However, to achieve better performance of the large-scale MIMO systems, both the LS method and the ICD method are very attractive. Thus, compared with the LS and the Argos, the ICD method is a favorable choice in multi-user large-scale MIMO systems due to the low complexity and the high performance.

6   Conclusion

In this paper, we extended the TLS method to the case of self-calibration for TDD systems. With this method, UEs are not required to feed back downlink CSI. We proved that the TLS method is equivalent to the LS method proposed by [15]. Furthermore, in order to greatly reduce the complexity of the LS method, we proposed a novel algorithm named as the ICD method. Theoretical analysis and simulation results showed that, the complexity of the ICD method is only in the order of $\mathcal{O}(M^2)$ for each iteration, and it converges close to the performance of the LS method in less than five iterations. When the number of antennas in all APs is very large (in the case of multi-user large-scale MIMO systems), the performance of the ICD method achieves almost the perfect calibration, and is much better than the Argos method.

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

References

1 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: 40–60

2 Ma Z, Zhang Z, Ding Z, et al. Key techniques for 5G wireless communications: network architecture, physical layer, and MAC layer perspectives. Sci China Inf Sci, 2015, 58: 041301

3 Wang D M, Wang J Z, You X H, et al. Spectral efficiency of distributed MIMO systems. IEEE J Sel Areas Commun, 2013, 31: 2112–2127

4 Wang J Z, Zhu H, Gomes N J. Distributed antenna systems for mobile communications in high speed trains. IEEE J Sel Areas Commun, 2012, 30: 675–683

5 Zhu H. Performance comparison between distributed antenna and microcellular systems. IEEE J Sel Areas Commun, 2011, 29: 1151–1163

6 Yang A, Jing Y D, Xing C W, et al. Performance analysis and location optimization for massive MIMO systems with circularly distributed antennas. arXiv preprint, 2014, arXiv:1408.1468

7 Bourdoux A, Come B, Khaled N. Non-reciprocal transceivers in OFDM/SDMA systems: impact and mitigation. In: Proceedings of IEEE Radio and Wireless Conference (RAWCON 03), Boston, 2003. 183–186

8 Kaltenberger F, Jiang H, Guillaud M. Relative channel reciprocity calibration in MIMO/TDD systems. In: Proceedings of IEEE Future Network and Mobile Summit, Florence, 2010. 1–10

9 Kouassi B, Ghauri I, Deneire L. Reciprocity-based cognitive transmissions using a MU massive MIMO approach. In: Proceedings of IEEE International Conference on Communications (ICC), Budapest, 2013. 2738–2742

10 Huang F, Geng J, Wang Y. Performance analysis of antenna calibration in coordinated multi-point transmission system. In: Proceedings of IEEE Vehicular Technology Conference (VTC-Spring), Taipei, 2010. 1–5

11 Su L Y, Yang C Y, Wang G, et al. Retrieving channel reciprocity for coordinated multi-point transmission with joint processing. IEEE Trans Commun, 2014, 62: 1541–1553

12 Shepard C, Yu H, Anand N. Argos: practical many-antenna base stations. In: Proceedings of the 18th annual International Conference on Mobile Computing and Networking, Istanbul, 2012. 53–64

13 Rahul H S, Kumar S, Katabi D. JMB: scaling wireless capacity with user demands. In: Proceedings of the ACM SIGCOMM Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication, Helsinki, 2012. 235–246

14 Rogalin R, Bursalioglu O Y, Papadopoulos H C. Hardware-impairment compensation for enabling distributed large-scale MIMO. In: Proceedings of IEEE Information Theory and Applications Workshop (ITA), San Diego, 2013. 1–10

15 Rogalin R, Bursalioglu O Y, Papadopoulos H, et al. Scalable synchronization and reciprocity calibration for distributed multiuser MIMO. IEEE Trans Wirel Commun, 2014, 13: 1815–1831

16 Golub G H, van C F. Matrix Computations. Baltimore: Johns Hopkins University Press, 2012

17 Golub G H. Some modified matrix eigenvalue problems. Siam Rev, 1973, 15: 318–334

18 Bertsekas D P, Tsitsiklis J N. Parallel and Distributed Computation: Numerical Methods. Upper Saddle River: Prentice-Hall, 1989

19 Luo Z Q, Tseng P. On the convergence of the coordinate descent method for convex differentiable minimization. J Optimiz Theory Appl, 1992, 72: 7–35

20 Liu J X, Wang D M. An improved dynamic clustering algorithm for multi-user distributed antenna system. In: Proceedings of IEEE International Conference on Wireless Communications and Signal Processing (WCSP), Nanjing, 2009. 1–5