SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

December 2018, Vol. 61 122302:1–122302:15
https://doi.org/10.1007/s11432-018-9591-8

Energy- and spectral-efficiency of zero-forcing beamforming in massive MIMO systems with imperfect reciprocity calibration: bound and optimization

Donglin LIU, Fei WU, Xin QUAN, Wanzhi MA, Ying LIU & Shihai SHAO*

National Key Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China, Chengdu 611731*, China*

Received 12 February 2018/Revised 7 June 2018/Accepted 7 September 2018/Published online 14 November 2018

Abstract In a time-division duplex (TDD) system with massive multiple input multiple output (MIMO), channel reciprocity calibration (RC) is generally required in order to cope with the reciprocity mismatch between the uplink and downlink channel state information. Currently, evaluating the achievable spectral efficiency (SE) and energy efficiency (EE) of TDD massive MIMO systems with imperfect RC (IRC) mainly relies on exhausting Monte Carlo simulations and it is infeasible to precisely and concisely quantify the achievable SE and EE with IRC. In this study, a novel method is presented for tightly bounding the achievable SE of massive MIMO systems with zero-forcing beamforming under IRC. On the basis of the analytical results, we demonstrate key insights for practical system design with IRC in three aspects: the scaling rule for interference power, saturation region of the SE, and the bound on the SE loss. Finally, the trade-off between spectral and energy efficiencies in the presence of IRC is determined with algorithms developed to optimize SE (EE) under a constrained EE (SE) value. The loss of optimal total SE and EE due to IRC is also quantified, which shows that the loss of optimal EE is more sensitive to IRC in a typical range of transmit power values.

Keywords massive MIMO, ZF beamforming, TDD reciprocity calibration, spectral and energy efficiencies

Citation Liu D L, Wu F, Quan X, et al. Energy- and spectral-efficiency of zero-forcing beamforming in massive MIMO systems with imperfect reciprocity calibration: bound and optimization. Sci China Inf Sci, 2018, 61(12): 122302, https://doi.org/10.1007/s11432-018-9591-8

1 Introduction

Driven by the ever-increasing data traffic over limited spectral resources with environmental concerns, spectral efficiency (SE) and energy efficiency (EE) have become fundamental metrics for the next generation of wireless communication systems [1–4]. Among various SE and EE enhancement technologies, the massive multiuser multiple-input multiple-output (MU-MIMO) system, where a base station equipped with a large number of antennas simultaneously schedules multiple users over the same frequency band, has drawn tremendous research interest from both industry and academia. The huge benefits of massive MU-MIMO systems in improving SE and EE have been recognized in theory [5–8].

The DL channel state information (CSI) must be obtained before implementing practical downlink (DL) massive MU-MIMO systems. TDD channel reciprocity allows one to perform DL beamforming in

* Corresponding author (email: ssh@uestc.eud.cn)

time-division duplexing (TDD) mode such that the downlink CSI can be extracted from the uplink (UL) estimate. Practically, the downlink or uplink CSI is the joint response of the wireless propagation channel and the radio-frequency (RF) circuits of the transmitter and receiver at both the base station (BS) and user sides. The TDD channel reciprocity generally holds between the UL and DL wireless propagation channels over air. Reciprocity cannot be guaranteed when one considers a hardware mismatch (HM) between the RF circuits in transceivers. The achievable rate and rate loss in MU-MIMO systems under uncalibrated reciprocity are derived in [9,10]. As shown in [9,10], reciprocity calibration is usually required for massive MU-MIMO systems. Recent research on this issue has experimentally and theoretically demonstrated various kinds of reciprocity calibration methods [10–13]. In particular, theoretical results and experimental trials have demonstrated that there is generally little need for calibration at the user side because the RF mismatch at the user side has negligible impact on the system performance [10–13].

Here, another concern regarding reciprocity calibration is DL performance after reciprocity calibration. The calibration process is inevitably imperfect in real applications. The achievable benefits of massive MU-MIMO can be significantly affected by imperfect reciprocity calibration (IRC) [10, 14]. As massive MU-MIMO systems have transitioned from a theoretical concept to one of the most compelling physical-layer technologies in the emerging 5G network, evaluating and understanding the impact of IRC becomes essential in the deployment of massive MU-MIMO systems. Currently, little research has focused on IRC. In [10, 14], different reciprocity calibration methods are analyzed and the impact of IRC was simulated. In [15], the authors provide a general analysis of the performance of MU-MIMO with imperfect calibration. This paper focuses on the relationship between system performance, the required calibration accuracy and the number of antennas to be calibrated. In [16], analytical expressions describing the achievable SINR for various beamforming schemes are derived under channel estimation errors and HM.

Motivated by [14–16], this paper studies the achievable SE and EE of MU-MIMO in zero-forcing (ZF) precoding systems with IRC. By adopting a singular value decomposition (SVD) method, a concise closed-form expression is given to approximate the capacity of TDD MU-MIMO in ZF precoding systems under IRC. The SE and EE tradeoff in the high SINR regime is presented. Algorithms are designed to achieve optimal SE values under a constrained EE value and vice versa. The contributions of this paper are summarized as follows:

(1) An SVD-based method for characterizing the capacity lower bound of IRC-impaired MU-MIMO in ZF precoding systems is presented. This capacity lower bound is given in closed form with respect to the mean square error of IRC (δ_{e}), the number of BS antennas (M), and the number of users (K).

(2) On the basis of the analytical results, key insights for practical system design with IRC are demonstrated and parameterized.

(a) Scaling of interference power: multiuser interference power resulting from IRC is scaled by $(K-1)/M$.

(b) Saturation region: The saturation region for the kth user, where the achievable SE is bounded and saturates with transmitted power due to IRC, can be feasibly characterized by the user SNR in the single-input and single-output (SISO) case (ξ_k), δ_{e}, M and K.

(c) Bound on the SE loss: The kth user suffers a maximum 1 bit/s/Hz SE loss due to IRC, as long as $\xi_k\delta_{\mathrm{e}} \leqslant 1$.

(3) In the presence of IRC, algorithms are designed to optimize SE under a constrained EE value or optimize EE under a constrained SE value. The impact of IRC on EE and SE is then quantified, which demonstrates that a higher total SE potentially leads to larger loss of EE due to IRC, and vice versa. It is also shown that the loss of total SE is less sensitive to IRC than that of EE within a dual range of EE and total SE.

2 System model

2.1 Uplink and downlink channel model

Consider a ZF-based TDD MU-MIMO system with M-antennas at a base station and K single-antenna users. Let $\boldsymbol{H}_{\mathrm{u}}$ and $\boldsymbol{H}_{\mathrm{d}}$ be the uplink and downlink matrices on the jth subcarrier, where the subcarrier index j is omitted for convenience. Similar to [7,12,17], $\boldsymbol{H}_{\mathrm{u}}$ and $\boldsymbol{H}_{\mathrm{d}}$ can be expressed as

$$\boldsymbol{H}_{\mathrm{u}} = \boldsymbol{H}_{\mathrm{ul}}\boldsymbol{D}_{\mathrm{pl}} = \boldsymbol{D}_{\mathrm{r}}\boldsymbol{H}\boldsymbol{D}_{\mathrm{ut}}\boldsymbol{D}_{\mathrm{pl}}, \tag{1}$$

$$\boldsymbol{H}_{\mathrm{d}} = \boldsymbol{D}_{\mathrm{pl}}\boldsymbol{H}_{\mathrm{dl}} = \boldsymbol{D}_{\mathrm{pl}}\boldsymbol{D}_{\mathrm{ur}}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{D}_{\mathrm{t}}, \tag{2}$$

where $\boldsymbol{H}_{\mathrm{ul}} = \boldsymbol{D}_{\mathrm{r}}\boldsymbol{H}\boldsymbol{D}_{\mathrm{ut}}$, $\boldsymbol{H}_{\mathrm{dl}} = \boldsymbol{D}_{\mathrm{ur}}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{D}_{\mathrm{t}}$, $\boldsymbol{D}_{\mathrm{r}} = \mathrm{diag}\left\{d_{\mathrm{r}}^1, \ldots, d_{\mathrm{r}}^M\right\}$ and $\boldsymbol{D}_{\mathrm{t}} = \mathrm{diag}\left\{d_{\mathrm{t}}^1, \ldots, d_{\mathrm{t}}^M\right\}$ are the HM coefficients at the BS side, d_{r}^m and d_{t}^m denote the HM coefficients of the receiver and transmitter circuits connected to the mth antenna, respectively, $\boldsymbol{D}_{\mathrm{ur}} = \mathrm{diag}\left\{d_{\mathrm{ur}}^1, \ldots, d_{\mathrm{ur}}^K\right\}$ and $\boldsymbol{D}_{\mathrm{ut}} = \mathrm{diag}\left\{d_{\mathrm{ut}}^1, \ldots, d_{\mathrm{ut}}^K\right\}$ indicate the HM coefficients at the user side with d_{ur}^k and d_{ut}^k being the HM coefficients induced by the kth users' receiver and transmitter, respectively, $\boldsymbol{D}_{\mathrm{pl}} = \mathrm{diag}\left\{\sqrt{\beta_k}\right\}$, $k = 1, 2, \ldots K$ denotes large-scale fading, and $\boldsymbol{H}$ and $\boldsymbol{H}^{\mathrm{T}}$ denote UL and DL small-scale fading of the wireless propagation channel, respectively, with entries of $\boldsymbol{H}$ being independent $\mathcal{CN}(0,1)$ random variables [12]. In more detail, the HM coefficients before calibration are i.i.d. random variables with uniformly distributed magnitudes over $[1-\rho, 1+\rho]$ and uniformly distributed phase over $[-\pi, \pi]$, as presented in [10,14,18]. Moreover, the users are uniformly distributed in an annular cell with radii $d_{\max}$ and $d_{\min}$, and $\{\beta_k\}$ is dominated by distance-dependent path-loss [19].

2.2 Downlink beamforming with reciprocity calibration

From (1) and (2), $\boldsymbol{D}_{\mathrm{t}}$, $\boldsymbol{D}_{\mathrm{r}}$, $\boldsymbol{D}_{\mathrm{ut}}$, and $\boldsymbol{D}_{\mathrm{ur}}$ destroy the reciprocity between the UL and DL channels. Generally, it is necessary to perform reciprocity calibration. Previous results have shown that the calibration is only required at the BS side, and $\boldsymbol{D}_{\mathrm{ut}}$ and $\boldsymbol{D}_{\mathrm{ur}}$ exert little effect on the beamforming performance [10–13]. Hence, this paper only considers calibration with respect to $\boldsymbol{D}_{\mathrm{t}}$ and $\boldsymbol{D}_{\mathrm{r}}$.

As thoroughly explained in [12,14], after calibration at the BS side, the downlink signal can be written as

$$\boldsymbol{s} = \frac{\sqrt{E_s}\boldsymbol{W}\boldsymbol{x}}{\|\boldsymbol{W}\|_F}, \tag{3}$$

where $\boldsymbol{x} = \{x_1, \ldots, x_K\}^{\mathrm{T}}$ is a $K \times 1$ vector, x_k is the signal intended for the kth user, $\boldsymbol{W} = \hat{\boldsymbol{D}}_{\mathrm{J}}\tilde{\boldsymbol{W}}$ with $\tilde{\boldsymbol{W}} = (\boldsymbol{H}_{\mathrm{ul}}^{\mathrm{T}})^{\mathrm{H}}[\boldsymbol{H}_{\mathrm{ul}}^{\mathrm{T}}(\boldsymbol{H}_{\mathrm{ul}}^{\mathrm{T}})^{\mathrm{H}}]^{-1}$ and $\hat{\boldsymbol{D}}_{\mathrm{J}}$ is the reciprocity calibration matrix, and $\|\boldsymbol{W}\|_F$ is the normalization factor for $\boldsymbol{s}$ in order to satisfy the power constraint E_s. It is worth noting that the $\hat{\boldsymbol{D}}_{\mathrm{J}}$ model depends on the calibration method. For the sake of analytical tractability, this paper considers the calibration scheme in [11,14]. The reciprocity calibration matrix can be written as [14]

$$\hat{\boldsymbol{D}}_{\mathrm{J}} = \boldsymbol{D}_{\mathrm{J}} - \boldsymbol{D}_{\mathrm{J}}\hat{\boldsymbol{D}}_{\mathrm{J}}^{\mathrm{n}}, \tag{4}$$

where $\boldsymbol{D}_{\mathrm{J}} = \alpha\boldsymbol{D}_{\mathrm{t}}^{-1}\boldsymbol{D}_{\mathrm{r}}$, α is an unknown but deterministic constant, and $\hat{\boldsymbol{D}}_{\mathrm{J}}^{\mathrm{n}} = \mathrm{diag}\left\{N_m/(d_{\mathrm{t}}^m X_m)\right\}$, where $m = 1, 2, \ldots, M$, N_m is circularly symmetric complex Gaussian (CSCG) noise, and X_m is the training signal for calibration. It is worth mentioning that $\hat{\boldsymbol{D}}_{\mathrm{J}}^{\mathrm{n}} = \mathrm{diag}\left\{N_m/(d_{\mathrm{t}}^m X_m)\right\}$ means that transmitter noise is dominant since calibration occurs in short-range communication scenarios, and transmitter noise is much more significant than receiver noise [14,20,21]. Thus, the variance of the mth entry in $\hat{\boldsymbol{D}}_{\mathrm{J}}^{\mathrm{n}}$ can be written in the form of $\delta_{\mathrm{e}}(m) = 1/\mathrm{SNR}_{\mathrm{J}}(m)$, where $\mathrm{SNR}_{\mathrm{J}}(m)$ is the SNR in the mth transmitter. In this paper, the SNR in all transmitters is considered identical for analytical convenience, i.e., $\delta_{\mathrm{e}}(m) = \delta_{\mathrm{e}}$, $m = 1, 2, \ldots, M$ [14].

The received signal is written as

$$\boldsymbol{Y} = \boldsymbol{H}_{\mathrm{d}}\boldsymbol{s} + \boldsymbol{n}, \tag{5}$$

where $\boldsymbol{n}$ denotes the CSCG noise with variance σ^2. By simple algebraic manipulation, the received signal can be rewritten as [14]

$$\boldsymbol{Y} = \frac{\alpha\sqrt{E_s}}{\|\boldsymbol{W}\|_F}\boldsymbol{D}_{\mathrm{ur}}\boldsymbol{D}_{\mathrm{ut}}^{-1}\boldsymbol{D}_{\mathrm{PL}}\left[\boldsymbol{x} + \boldsymbol{H}_{\mathrm{ul}}^{\mathrm{T}}\hat{\boldsymbol{D}}_{\mathrm{J}}^{\mathrm{n}}\tilde{\boldsymbol{W}}\boldsymbol{x}\right] + \boldsymbol{n}. \tag{6}$$

The instantaneous SINR for the kth user is given by

$$\gamma_k = \frac{|\mu_k|^2\beta_k E_s/\|\boldsymbol{W}\|_F^2}{|\mu_k|^2\frac{\beta_k E_s}{\|\boldsymbol{W}\|_F^2}\sum_{n=1,n\neq k}^{K}\sum_{m=1}^{M}\delta_{\mathrm{e}}|h_{nm}\tilde{w}_{mn}|^2 + \sigma^2}, \tag{7}$$

where $h_{nm} = (\boldsymbol{H}_{\mathrm{ul}}^{\mathrm{T}})_{nm}$, $\tilde{w}_{kn} = (\tilde{\boldsymbol{W}})_{kn}$ and $\mu_k = d_{\mathrm{ur}}^k/d_{\mathrm{ut}}^k$.

Denote $\xi_k = \beta_k E_s/\sigma^2$, which is the achievable SNR in SISO systems conditioned on β_k. For the sake of analytical tractability, the large-scale fading β_k is treated as a constant in this section, as was also the case in [7]. The effect of β_k will be presented in Section 4.

From Jensen's inequality [7] and after simple algebraic manipulation, the lower bound on the rate of user k can be written as

$$\begin{aligned} R_k &= \mathrm{E}\left[\log_2\left(1+\gamma_k\right)\right] \geqslant \log_2\left(1+\frac{1}{\mathrm{E}\{1/\gamma_k\}}\right) \\ &= \log_2\left(1+\frac{\xi_k}{\xi_k\mathrm{E}\left[\sum_{n=1,n\neq k}^{K}\sum_{m=1}^{M}\delta_{\mathrm{e}}|h_{nm}\tilde{w}_{mn}|^2 + \frac{\sigma^2}{|\mu_k|^2\beta_k E_s/\|\boldsymbol{W}\|_F^2}\right]}\right) \\ &= \log_2\left(1+\frac{\xi_k}{\mathrm{E}\left(\xi_k\overline{\mathcal{I}}_k + \overline{\mathcal{N}}_k\right)}\right), \end{aligned} \tag{8}$$

where $\overline{\mathcal{I}}_k = \sum_{n=1,n\neq k}^{K}\sum_{m=1}^{M}\delta_{\mathrm{e}}|h_{nm}\tilde{w}_{mn}|^2$ and $\overline{\mathcal{N}}_k = \|\boldsymbol{W}\|_F^2/|\mu_k|^2$ are used for concise notation. In (8), $\mathrm{E}\left(\xi_k\overline{\mathcal{I}}_k\right)$ denotes the average power of multi-user interference resulting from imperfect calibration and $\mathrm{E}\left(\overline{\mathcal{N}}_k\right)$ is the average power of the equivalent CSCG noise.

3 Lower capacity bound with imperfect calibration

In this section, we present the capacity lower bound for a MU-MIMO system with residual reciprocity calibration errors.

3.1 Average power of multi-user interference

Define a $K\times 1$ vector $\boldsymbol{u} = \boldsymbol{H}_{\mathrm{ul}}^{\mathrm{T}}\hat{\boldsymbol{D}}_{\mathrm{J}}^{\mathrm{n}}\tilde{\boldsymbol{W}}\boldsymbol{x}$. It is noted that $\boldsymbol{u}_k$ is the sum of the weighted $\{x_k\}$ components, of which $K-1$ terms are related to the interference power resulting from imperfect calibration. As $\boldsymbol{H}_{\mathrm{ul}}^{\mathrm{T}}$ and $\hat{\boldsymbol{D}}_{\mathrm{J}}^{\mathrm{n}}$ are isotropic (independent of index k), we propose treating these as isotropic terms, i.e., they contribute equally to $\mathrm{E}[(\boldsymbol{u}\boldsymbol{u}^{\mathrm{H}})_{k,k}]$. Hence, the average interference power can be written as

$$\mathrm{E}\left(\xi_k\overline{\mathcal{I}}_k\right) \approx \xi_k\frac{K-1}{K}\mathrm{E}\left[\left(\boldsymbol{u}\boldsymbol{u}^{\mathrm{H}}\right)_{k,k}\right]. \tag{9}$$

Again, from the system model, we have $\mathrm{E}[(\boldsymbol{u}\boldsymbol{u}^{\mathrm{H}})_{k,k}] = \mathrm{E}[(\boldsymbol{u}\boldsymbol{u}^{\mathrm{H}})_{j,j}]$. As such,

$$\mathrm{E}\left[\left(\boldsymbol{u}\boldsymbol{u}^{\mathrm{H}}\right)_{k,k}\right] = \frac{1}{K}\mathrm{E}\left\{\mathrm{tr}\left(\boldsymbol{u}\boldsymbol{u}^{\mathrm{H}}\right)\right\}, \tag{10}$$

where tr() denotes the trace of a matrix. Based on (9) and (10), the following proposition is given.

Proposition 1 (Scaling of interference power)**.** With $M \gg K$, the average power of multi-user interference resulting from imperfect calibration can be approximated by

$$\mathrm{E}\left(\xi_k \overline{\mathcal{I}}_k\right) \approx \xi_k \frac{K-1}{M} \delta_{\mathrm{e}}. \tag{11}$$

Proof. See Appendix A.

Remark 1. As implied by [14], multi-user interference resulting from imperfect calibration takes the form of (truncated) Gaussian white noise weighted by the beam-forming vector and transmitted signal. From this perspective, this paper further shows that the multi-user interference power, which is approximated as white noise, increases linearly with $K-1$ and decreases with M.

3.2 Average power of Gaussian white noise

Following the notation in this paper, the equivalent noise power is $\overline{\mathcal{N}}_k = \|\boldsymbol{W}\|_F^2 / |\mu_k|^2$. Considering the magnitudes of HM at the BS side are i.i.d. random variables, we can obtain the average of $\bar{\mathcal{N}}_k$ as follows.

Proposition 2. The average of the equivalent noise power $\overline{\mathcal{N}}_k = \|\boldsymbol{W}\|_F^2 / |\mu_k|^2$ can be written as

$$\mathrm{E}\left(\bar{\mathcal{N}}_k\right) = \frac{\mathrm{E}\left|\left(\hat{\boldsymbol{D}}_{\mathrm{J}}\right)_{k,k}\right|^2 \mathrm{E}\left[\left|\frac{d_{\mathrm{ut}}^k}{d_{\mathrm{ur}}^k}\right|^2 \sum_{k=1}^{K} \left|\frac{1}{d_{\mathrm{ut}}^k}\right|^2\right]}{(M-K)}. \tag{12}$$

Proof. See Appendix B.

The calculation of $\mathrm{E}|(\hat{\boldsymbol{D}}_{\mathrm{J}})_{k,k}|^2 \mathrm{E}[|\frac{d_{\mathrm{ut}}^k}{d_{\mathrm{ur}}^k}|^2 \sum_{k=1}^{K} |\frac{1}{d_{\mathrm{ut}}^k}|^2]$ depends on the HM model ($\boldsymbol{D}_{\mathrm{ut}}$) and calibration method, which determines $\hat{\boldsymbol{D}}_{\mathrm{J}}$. One should note that the HM at the user side has negligible impact on the system performance [10,11,14]. Thus, for simplicity, we neglect $\boldsymbol{D}_{\mathrm{ut}}$ and $\boldsymbol{D}_{\mathrm{ur}}$, i.e., $\boldsymbol{D}_{\mathrm{ut}} = \boldsymbol{D}_{\mathrm{ur}} = \mathbf{I}$.

Therefore, from the HM model at the BS side in Section 2, Eq. (12) can be modified as

$$\mathrm{E}\left(\bar{\mathcal{N}}_k\right) = \frac{K\left(1+\delta_{\mathrm{e}}\right)}{(M-K)} \left[\frac{1+\rho^2/3}{(1-\rho^2)}\right]^2, \tag{13}$$

where $\mathrm{E}|(\hat{\boldsymbol{D}}_{\mathrm{J}})_{i,i}|^2$ can be calculated as follows: $\mathrm{E}|(\hat{\boldsymbol{D}}_{\mathrm{J}})_{i,i}|^2 = [\frac{1+\rho^2/3}{(1-\rho^2)}]^2(1+\delta_{\mathrm{e}})$. Hence, substituting (11) and (13) into (8), the following lower bound can be obtained.

Proposition 3. Conditioned on $M \gg K$, the lower bound for the rate of MU-MIMO ZF beamforming is

$$R_k^{\mathrm{LB}} \approx \log_2\left[1 + \frac{\xi_k}{\xi_k \frac{K-1}{M}\delta_{\mathrm{e}} + \frac{K(1+\delta_{\mathrm{e}})}{M-K}\left(\frac{1+\rho^2/3}{1-\rho^2}\right)^2}\right] \approx \log_2\left[1 + \frac{\xi_k}{\xi_k \frac{K-1}{M}\delta_{\mathrm{e}} + \frac{K(1+\delta_{\mathrm{e}})}{M-K}}\right], \tag{14}$$

where the second approximation holds when ρ is small.

Remark 2. Let $\delta_{\mathrm{e}} \to 0$ and omit the HM at the BS side, i.e., $\rho \to 0$. The rate in (14) subsequently reduces to the classical form of massive MIMO [7], which is widely used to approximate the achievable rate in massive MIMO systems. As shown in Section 4, the derived results in (14) are also a tight bound of (8). Eq. (14) is used to approximate the achievable rate of massive MIMO ZF beam-forming with imperfect calibration in the rest of this paper.

As (14) provides a concise form for the achievable rate for ZF massive MIMO, several theoretical insights can be obtained. From (14), the following corollaries are given.

Corollary 1 (Saturation region)**.** The rate in (14) saturates with respect to the transmitted power E_s if

$$\xi_k \gg \left(\frac{1}{\delta_{\mathrm{e}}} + 1\right) \frac{M}{M-K} \frac{K}{K-1}. \tag{15}$$

Under IRC, if the transmit power E_s grows without bound, the achievable rate is bounded by

$$R_k^{\mathrm{LB}} \leqslant \log_2\left[1+\frac{M}{K-1}\times\frac{1}{\delta_{\mathrm{e}}}\right]. \tag{16}$$

Proof. One should note that if

$$\xi_k\frac{K-1}{M}\delta_{\mathrm{e}} \gg \frac{K\left(1+\delta_{\mathrm{e}}\right)}{M-K}, \tag{17}$$

R_k^{LB} can be simplified as

$$R_k^{\mathrm{LB}} \to \log_2\left[1+\frac{M}{K-1}\times\frac{1}{\delta_{\mathrm{e}}}\right]. \tag{18}$$

In this case, as shown in (18), the rate R_k^{LB} saturates with respect to the transmitted power E_s. Eq. (15) is obtained by simple algebraic manipulation of (17). As R_k^{LB} increases with E_s monotonously, it is known that $R_k^{\mathrm{LB}} \leqslant R_k^{\mathrm{LB}}\big|_{E_s\to\infty}$, which yields (16).

Remark 3. Eq. (15) indicates the interference-limited massive MIMO system, in which the achievable rate is bounded by (16) due to IRC. When Eq. (15) is satisfied, increasing E_s causes energy to be wasted. It is noted that similar observations from simulations were reported for regularized ZF precoding in [10] when the HM is not calibrated. This paper provides a detailed criterion to determine the saturation region of E_s for ZF precoding with imperfect calibration.

Remark 4. From the asymptotic perspective, let $K \gg 1$ and $M/K \gg 1$, and Eq. (15) can be rewritten as

$$\xi_k \gg \left(1/\delta_{\mathrm{e}}+1\right). \tag{19}$$

This provides a concise estimation of the saturation region of E_s with imperfect calibration. Note that δ_{e} is the MSE of IRC in general, and $1/\delta_{\mathrm{e}}$ can be interpreted as the calibration SNR for the "Argos" calibration scheme [14]. Thus, one can say that the instantaneous SISO SNR is much larger than the calibration SNR in (19).

In practice, Eq. (15) or (19) usually means calibration is very poor such that the MU-MIMO system wastes energy. Thus, E_{s} should be controlled such that Eq. (15) or (19) cannot be satisfied. This observation is reflected in the SE and EE trade-off in Section 4. Moreover, another issue is the performance loss due to imperfect calibration.

Corollary 2. Conditioned on high SINR for user k, the loss of user k's rate due to imperfect calibration can be calculated as follows:

$$L = \log_2\left(1+\xi_k\delta_{\mathrm{e}}\frac{M-K}{M}\frac{K-1}{K}+\delta_{\mathrm{e}}\right) \text{ bits/s/Hz}. \tag{20}$$

Proof. The performance loss is written as $L = R_k^{\mathrm{LB}}|_{\delta_{\mathrm{e}}\to 0} - R_k^{\mathrm{LB}}$, where $R_k^{\mathrm{LB}}|_{\delta_{\mathrm{e}}\to 0}$ denotes the case of perfect calibration. Eq. (20) is produced after simple algebraic manipulation of L.

Remark 5. As shown in (20), a larger ξ_k means we potentially lower the achievable rates due to imperfect calibration. Mathematically, $L \to \infty$ if ξ_k grows without bound. This observation is implicitly reflected in Corollary 3.1: R_k^{LB} converges to $\log_2[1+\frac{M}{K-1}\frac{1}{\delta_{\mathrm{e}}}]$, while $R_k^{\mathrm{LB}}|_{\delta_{\mathrm{e}}\to 0}$ increases with ξ_k without bound. Thus, $R_k^{\mathrm{LB}}|_{\delta_{\mathrm{e}}\to 0} - R_k^{\mathrm{LB}}$ can be extremely large with a very large ξ_k.

Corollary 3 (Bound of SE Loss)**.** It is straightforward to see that L is bounded by

$$L < L|_{M\gg K\to\infty} = \log_2\left[1+\delta_{\mathrm{e}}\left(1+\xi_k\right)\right] \approx \log_2\left[1+\delta_{\mathrm{e}}\xi_k\right] \tag{21}$$

since L is an increasing function of K and M.

Remark 6. Note that as $M \gg K \gg 1$, the bound in (21) becomes tight. In this paper, we focus on the case where efficient calibration is deployed such that neither (15) nor (19) are true. In more detail, δ_{e} is determined from the calibration SNR for the typical massive MIMO calibration method [11,14], i.e.,

$\delta_{\mathrm{e}} = 1/\mathrm{SNR}_{\mathrm{J}}$. Thus, $\xi_k\delta_{\mathrm{e}} \leqslant 1$ is interpreted as the user SNR in the SISO case and would not exceed the calibration SNR. Hence, if $\xi_k\delta_{\mathrm{e}} \leqslant 1$ is considered, the bound becomes

$$L < L|_{M\gg K\to\infty} \approx 1. \tag{22}$$

This means that when efficient calibration ($\xi_k\delta_{\mathrm{e}} \leqslant 1$) is performed, user k has 1 bit/s/Hz maximum rate loss due to imperfect calibration.

4 Spectral efficiency v.s. energy efficiency

Spectral and energy efficiencies are key theoretical concepts related to evaluating the performance of massive MIMO. As indicated in Corollary 3.1, IRC causes energy waste, especially when E_s is large, or equivalently when the SE is large. Hence, it is rewarding to consider the tradeoff between SE and EE in a MU-MIMO system with IRC. This section maximizes the achievable SE (EE) under a constrained EE (SE) value.

Mathematically the spectral efficiency is considered as the maximum achievable total rate averaged over all fading states and user locations. That is, $f_{\mathrm{se}}(K, E_s) = \sum_{k=1}^{K} \mathrm{E}_{\beta_k}\left(R_k^{\mathrm{LB}}\right)$, where E_{β_k} denotes an average over β_k. The energy efficiency is defined as $\eta_{\mathrm{ee}}(K, E_s) = \frac{1}{E_s}\sum_{k=1}^{K} \mathrm{E}_{\beta_k}\left(R_k^{\mathrm{LB}}\right)$.

Then, the following problems are formulated:

P1:

$$\max_{K,E_s}\ f_{\mathrm{se}}(K, E_s) \tag{23}$$

$$\text{s.t.}\ \ 2 \leqslant K \leqslant K_{\max}, \quad \eta_{\mathrm{ee}}(K, E_s) \geqslant \eta_0, \quad E_{\min} \leqslant E_s \leqslant E_{\max}, \tag{24}$$

and P2:

$$\max_{K,E_s}\ \eta_{\mathrm{se}}(K, E_s) \tag{25}$$

$$\text{s.t.}\ \ 2 \leqslant K \leqslant K_{\max}, \quad f_{\mathrm{se}}(K, E_s) \geqslant f_0, \quad E_{\min} \leqslant E_s \leqslant E_{\max}, \tag{26}$$

where $K_{\max}$ is the maximum number of users to be scheduled, $E_{\min}$ denotes the minimum power that the downlink should provide to meet the quality of service (QOS) threshold. f_0 and η_0 are the required minimum SE and EE threshold, respectively, and $E_{\max}$ is the maximum power available for downlink signal transmission.

Note that it is intractable to obtain an explicit form for $\mathrm{E}_{\beta_k}\left(R_k^{\mathrm{LB}}\right)$. This paper implicitly solves P1 and P2 without obtaining an explicit form of $\mathrm{E}_{\beta_k}\left(R_k^{\mathrm{LB}}\right)$.

One can verify that R_k^{LB} is a monotonically increasing function of E_s, $\forall K \in [2, K_{\max}]$ for any instantaneous R_k^{LB} with respect to β_k. Hence, the convexity of $f_{\mathrm{se}}(K, E_s)$ is only determined by K. For simplicity, K is treated as a continuous variable since higher order derivatives of $f_{\mathrm{se}}(K, E_s)$ exist with respect to continuous K.

As analyzed in Corollaries 3.1 and 3.2, the rate loss and energy waste due to IRC become more significant in the high SINR regime. To simplify the analysis, consider that the cell has satisfying coverage such that γ_k is in the high SINR region within the cell and $R_k^{\mathrm{LB}} \approx \log_2(\gamma_k)$, thanks to the promising array gains and better coverage guaranteed by massive MIMO [22, 23]. This condition facilitates the notation and analysis in this paper. Hence, the following propositions are offered.

Proposition 4. $f_{\mathrm{se}}(K, E_s)$ is concave with respect to K.

Proof. See Appendix C.

Proposition 5. η_{ee} is a monotonically decreasing function of E_s, regardless of K.

Proof. See Appendix D.

From Propositions 4 and 5, it is known that (1) $f_{\mathrm{se}}(K, E_s)$ and $\eta_{\mathrm{ee}}(K, E_s)$ are unimodal functions of K; (2) $f_{\mathrm{se}}(K, E_s)$ increases with E_s and $\eta_{\mathrm{ee}}(K, E_s)$ decreases with E_s; (3) if $E_{s1} > E_{s2}$, then

$$f_{\mathrm{se}}\left(\tilde{K}_1, E_{s1}\right) \geqslant f_{\mathrm{se}}\left(\tilde{K}_2, E_{s1}\right) > f_{\mathrm{se}}\left(\tilde{K}_2, E_{s2}\right), \tag{27}$$

$$\eta_{\mathrm{ee}}\left(\tilde{K}_1, E_{s1}\right) \leqslant \eta_{\mathrm{ee}}\left(\tilde{K}_1, E_{s2}\right) < \eta_{\mathrm{ee}}\left(\tilde{K}_2, E_{s2}\right), \tag{28}$$

where $\tilde{K}_i$ denotes the optimal number of active users that maximize $f_{\rm se}$ conditioned on E_{si}. It is worth noting that $f_{\rm se}(K, E_s)$ and $\eta_{\rm ee}(K, E_s)$ have identical convexity with respect to K conditioned on fixed E_s.

Therefore, Eqs. (27) and (28) imply that (1) the solution to P1 is the maximum E_s value that satisfies (24); (2) the solution to P2 is the minimum E_s value that satisfies (26). Then, P1 and P2 can be solved using a combination of the low-complexity ternary search algorithm and bisection method. Let us take P1 as an example. The detailed algorithms are given as Algorithms 1 and 2.

Algorithm 1 Optimal $E_{\rm s}^*$ and K^* for P1

Initialization $E_1 = E_{\min}, E_2 = E_{\max}$, tolerance $\varepsilon > 0$;
Let $E_s = E_2$ and compute K_2 and $\eta_{\rm ee}(K_2, E_2)$ by Algorithm 2;
Let $E_s = E_1$ and compute K_1 and $\eta_{\rm ee}(K_1, E_1)$ by Algorithm 2;
if $\eta_{\rm ee}(K_2, E_2) < \eta_0$ **then**
    no solution and stop;
end if
if $\eta_{\rm ee}(K_1, E_1) \geqslant \eta_0$ **then**
    return $E_s^* = E_2$ and $K^* = K_2$;
end if
repeat
    $E_m = (E_1 + E_2)/2$; compute K_m and $\eta_{\rm ee}(K_m, E_m)$ by Algorithm 2;
    if $\eta_{\rm ee}(K_m, E_m) > \eta_0$ **then**
        $E_1 = E_m$ and $K_1 = K_m$;
    else
        $E_2 = E_m$ and $K_2 = K_m$;
    end if
until $|\eta_{\rm ee}(K_1, E_1) - \eta_0| < \varepsilon$;
return $E_s^* = E_1$ and $K^* = K_1$;

Algorithm 2 Optimal number of active users

Initialization $l = 2, r = K_{\max}, \Delta = 1$;
while $l + \Delta < r$ **do**
    $m = (l + r)/2, mm = (m + r)/2$;
    compute $R(K = m) = \frac{1}{N}\sum_{n=1}^{N} KR_n^{\rm LB}$ and $R(K = mm) = \frac{1}{N}\sum_{n=1}^{N} KR_n^{\rm LB}$;
    if $R(K = m) \geqslant R(K = mm)$ **then**
        $r = m$;
    else
        $l = m$;
    end if
end while
$K_1 = \lfloor l \rfloor$, $K_2 = \lceil l \rceil$, and $K_3 = \lceil r \rceil$;
compute $R(K = K_i) = \frac{1}{N}\sum_{n=1}^{N} KR_n^{\rm LB}$, $i = 1, 2, 3$;
$f_{\rm se}(K_0, E_s) = \max\{R(K = K_i)\}$, where $K_0 = \text{argmax}\{R(K = K_i)\}$;
$\eta_{\rm ee}(K_0, E_s) = f_{\rm se}(K_0, E_s)/E_s$;
return $\eta_{\rm ee}(K_0, E_s)$ and K_0;

Algorithm 2 (A2) is invoked in Algorithm 1 (A1), which is the ternary search algorithm used to compute the optimal number of active users conditioned on a given E_s value. In A2, $\lceil X \rceil$ denotes rounding X to the nearest integer greater than or equal to X, and $\lfloor X \rfloor$ denotes rounding X to the nearest integer less than or equal to X. Considering that K should be an integer, Δ can be large, e.g., $\Delta = 1$ for the sake of reducing iterations.

Accordingly, A1 can be simply modified for P2 since P2 and P1 are dual problems. Thus, the algorithm for P2 is not listed specifically.

Moreover, it is worth mentioning that steps 4 and 5 in A2 calculate the average achievable rate numerically using the Monte Carlo method with respect to the path loss. An alternative method is to

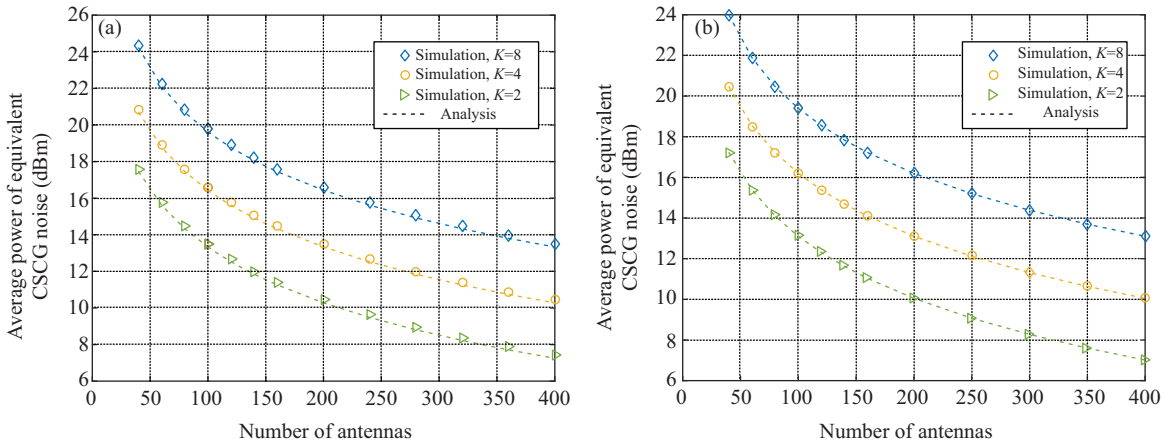

Figure 1 (Color online) Average equivalent CSCG noise power, where σ^2 is normalized to 1 (0 dBm). (a) $\rho = 0.2$ and $\delta_{\rm e} = 0.001$; (b) $\rho = 0.03$ and $\delta_{\rm e} = 0.001$.

calculate the lower-bound of $f_{\rm se}(K, E_s)$ and $\eta_{\rm ee}(K, E_s)$ by

$$f_{\rm se}^{\rm L}(K, E_s) = K\log_2\left(1 + \frac{E_s/\sigma^2}{\frac{E_s}{\sigma^2}\frac{K-1}{M}\delta_{\rm e} + \bar{\beta}\frac{K}{M-K}(1+\delta_{\rm e})}\right) \tag{29}$$

and

$$\eta_{\rm ee}^{\rm L}(K, E_s) = f_{\rm se}^{\rm L}(K, E_s)/E_s, \tag{30}$$

where $\bar{\beta} = E\{\beta_k^{-1}\}$. It is straightforward to see that Propositions 4 and 5 are also applicable to (29) and (30). Thus, A1 and A2 can also be adopted for the case of the lower bounds given in (29) and (30). However, the lower-bounds given in (29) and (30), which are averaged with respect to the path loss, may be somewhat loose since the path loss can vary over a large range in dB. Therefore, this paper uses the Monte Carlo method to obtain more accurate results.

5 Simulation and numerical results

This section presents the numerical results that support the analysis presented in the preceding sections. The parameterizations in all simulations are set as follows. The calibration SNR is less than or equal to 30 dB, i.e., $\delta_{\rm e} \geqslant 0.001$. We assume ρ is small with $\rho \leqslant 0.2$ [10]. Note that $\rho = 0.2$ leads to a maximum power mismatch of 3.5 dB and denotes a case with very severe HM. Moreover, the users are uniformly distributed in an annular micro-cell with radii $d_{\max} = 300$ m and $d_{\min} = 50$ m [24]. Large-scale fading is described with the COST231-WI model [25] for simplicity, where the path loss is given in dB as a function of distance d in km and carrier frequency $f_{\rm c}$ in MHz as $\mathrm{PL}_k(d) = 10\log_{10}\beta_k = 42 + 20\log_{10}f_{\rm c} + 26\log_{10}d$.

5.1 Validation of derived analytical expressions

First, the derived average equivalent CSCG noise power and average multi-user interference power due to imperfect calibration are verified.

The average equivalent noise power, which is a scaled version of CSCG noise, is shown in Figure 1 with $\rho = 0.2$ and $\rho = 0.03$. As shown in Figure 1, the analytical approximation agrees fairly well with the simulation results, even when $\rho = 0.2$. It is worth mentioning that the average equivalent CSCG noise power is presented in dBm with σ^2 normalized to 1 for concise illustration in Figure 1. A small number of users was chosen as an example since complete simulation of imperfect calibration with a large number of antennas is indeed difficult and time consuming.

The multiuser interference power resulting from imperfect calibration is shown in Figure 2, in which markers ($*, \diamond, \circ,$ and $\square$) denote simulation results, and dashed lines denote a theoretical approximation

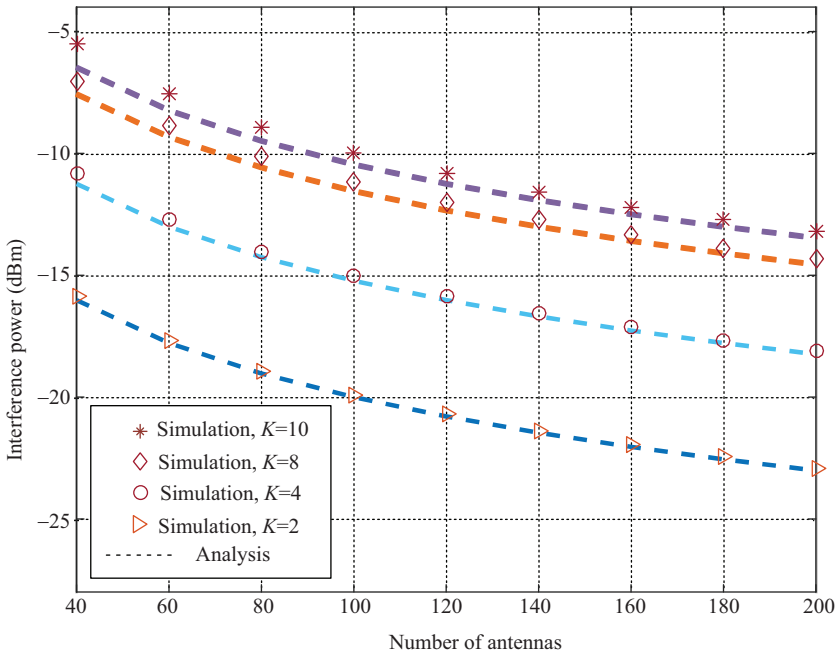

Figure 2 (Color online) Average power of multi-user interference resulting from imperfect calibration, where the number of users is 2, 4, 8, and 10; $\rho = 0.2$; $\delta_{\mathrm{e}} = 0.001$; $\xi_k = 0$ dB.

from (11). We set a magnitude mismatch with $\rho = 0.2$, and ξ_k is normalized to 0 dB as an example. Figure 2 shows that the theoretical results provide a good approximation for the average multiuser interference power, especially when $M \gg K$. As analyzed below Proposition 1, the interference power increases linearly with $K-1$ and decreases with M.

Then, the total SE is presented in Figure 3. One should note that the markers in Figure 3 denote complete simulation results, and that "Analysis" denotes the analytical results based on (14) with numerically averaged large-scale fading. As shown in Figure 3, theoretical results provide a good approximation for the achievable total SE with imperfect calibration. Moreover, Figure 3 also implies that the effect of imperfect calibration on the total spectral efficiency becomes significant when the transmit power is high, or when the users are scheduled in the high SNR regime. In particular, IRC should be considered in the high SNR regime.

The performance gap resulting from IRC is shown in what follows. Here, the achievable spectral efficiency under perfect reciprocity is adopted as a benchmark, as was the case in [5, 7]. The single-user spectral efficiency loss due to imperfect calibration is presented in Figure 4. As shown in Figure 4, as long as the instantaneous ξ_k is not larger than $1/\delta_{\mathrm{e}}$, which has the interpretation of achievable SISO SNR, a single scheduled user has at most 1 bit/s/Hz rate loss due to imperfect calibration. This observation agrees with Corollary 3.3. Again, both Figures 3 and 4 show that the effect of imperfect calibration should be particularly considered in the high SNR regime since IRC leads to non-negligible SE loss and subsequent EE loss in this regime.

5.2 Energy and spectral efficiencies with IRC

As Eq. (14) offers an accurate approximation for the achievable total SE with imperfect calibration, it is feasible to determine the trade-off between EE and SE. Figure 5(a) shows the optimal EE under a required total SE value. Here, optimal EE can be interpreted as the achievable maximum energy efficiency under constrained K, E_{s}, and required total SE values. Figure 5(a) shows the optimal EE value calculated with the proposed algorithms. For simplicity, the results obtained via exhaustive search are only provided for $M = 200$. As a comparison, the achievable EE value under perfect TDD channel reciprocity from [7] is also presented in Figure 5(a). Figure 5(b) quantifies the reduction in optimal EE due to IRC and compares with the reduction with the results in [7].

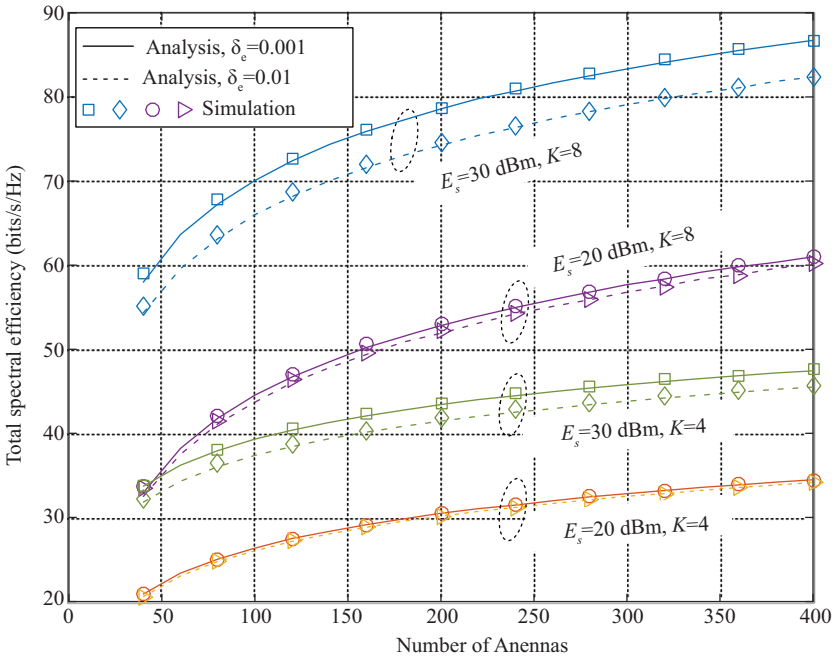

Figure 3 (Color online) Total spectral efficiency versus the number of antennas, where $\rho = 0.1$ and $f_{\rm c} = 1.8$ GHz.

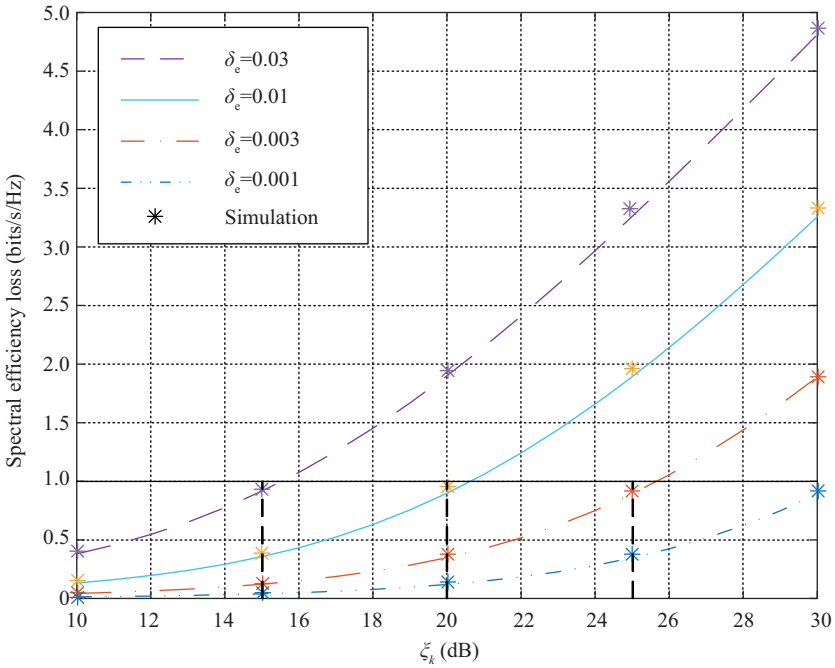

Figure 4 (Color online) Spectral efficiency loss from a single user side due to imperfect calibration, where ξ_k is interpreted as the achievable SISO SNR, $1/\delta_{\rm e}$ is interpreted as the calibration SNR, K=10, and M=100.

It is noted that the available transmission power is bounded in this paper, i.e., $0 < E_{\min} \leqslant E_s \leqslant E_{\max}$. From the monotonicity of SE with respect to E_s, $E_s = E_{\min}$ can provide a total SE that is larger than the required value as long as the required SE is low enough. Then, $E_s^* = E_{\min}$ is the optimal solution according to Propositions 4 and 5 and the analysis below (27) and (28). Hence, at this stage, the optimal transmitted power always converges to $E_{\min}$, which implies the optimal EE is bounded and unchanged in the low SE regime in Figure 5(a). Moreover, as indicated in (14), interference resulting from IRC is negligible compared to Gaussian white noise in the low SE regime, where the required transmit power

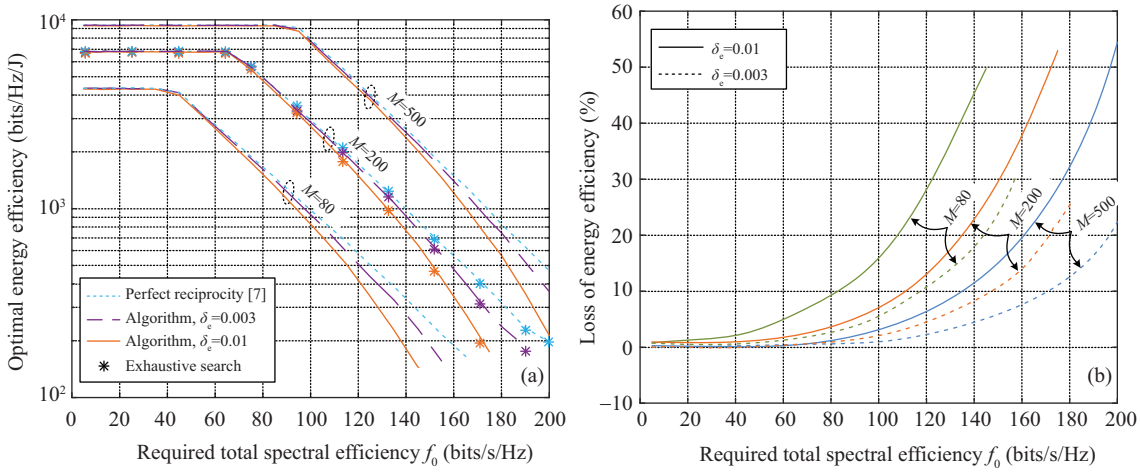

Figure 5 (Color online) Optimal EE (a) and Loss of EE (b) under a required SE value, where $2 \leqslant K \leqslant 20$ and $0.01 \leqslant E_s \leqslant 1$.

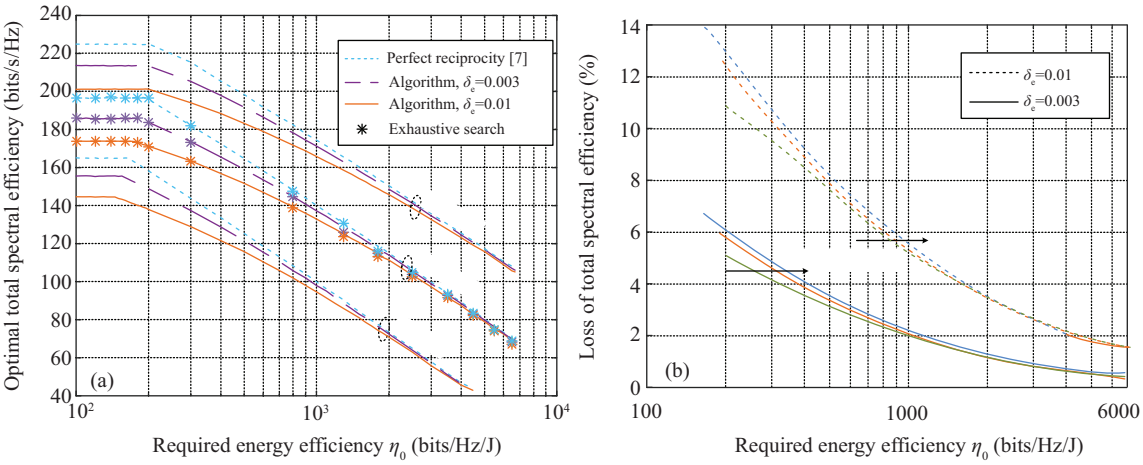

Figure 6 (Color online) Optimal SE (a) and Loss of SE (b) under a required EE value, where $2 \leqslant K \leqslant 20$ and $0.01 \leqslant E_s \leqslant 1$.

is low. Thus, at this stage, the effect of IRC is not significant. In the high SE regime, higher total SE potentially leads to larger EE loss due to imperfect calibration. This observation is reasoned by Corollary 3.1: SE in the presence of IRC can saturate when E_s becomes larger, implying more power is consumed by interference and the value of EE loss grows without bound. In addition, deploying more antennas can potentially increase the achievable SINR [5–7], thus reducing the minimum required power in order to achieve a required SE value. Hence, it is straightforward to see from Figure 5 that: (1) The global maximum of the optimal energy efficiency increases with the number of deployed antennas, and (2) deploying a larger number of antennas widens the required SE range, where the optimal EE suffers almost no efficiency loss due to IRC. The results in Figure 5(b) can also be explained in a similar way.

The optimal total SE for a given required EE is shown in Figure 6(a). As a comparison, the achievable total SE under perfect TDD channel reciprocity is also presented, as in [7]. Then, Figure 6(b) quantifies the loss of the optimal total SE due to IRC and compares the loss with that reported in [7]. As P1 and P2 are dual problems, the impact of imperfect calibration on the optimal total SE in Figure 6(a) is similar. As quantified in Figure 6(a) and (b), the efficiency loss due to imperfect calibration becomes severe when the required total SE is low (i.e., the optimal total SE is high). However, by comparing Figures 5(b) and 6(b), the loss of total SE is less sensitive to imperfect calibration than the loss of EE in Figure 6 within a dual range of EE and total SE. Hence, the reciprocity calibration should be more carefully performed in EE-limited systems.

6 Conclusion

The impact of imperfect reciprocity calibration was analyzed in this paper. A concise closed-form expression for the achievable lower bound of SE for zero-forcing massive MIMO systems is derived. Simulation results agreed well with the analytical expression we provide, and this expression was used to gain theoretical insights into the impact of IRC. The scaling rule for interference power, SE saturation region, and bound of SE loss were demonstrated for practical system design. Finally, the trade-off between the spectral and energy efficiencies in the high SINR regime was determined in the presence of IRC using algorithms designed to achieve optimal SE under a constrained EE value, and vice versa. We demonstrated that higher total SE potentially leads to a larger EE loss due to IRC, and vice versa, and that the loss of total SE was less sensitive to imperfect calibration than that of EE within a dual range of EE and total SE values.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 61531009, 61471108, 61771107, 61701075), the National Major Projects (Grant No. 2016ZX03001009), the Fund from the China Scholarship Council (Grant No. 201706070084), and the Fundamental Research Funds for the Central Universities.

References

1 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
2 Papazafeiropoulos A K, Ngo H Q, Ratnarajah T. Performance of massive MIMO uplink with zero-forcing receivers under delayed channels. IEEE Trans Veh Technol, 2017, 66: 3158–3169
3 Cao J, Wang D M, Li J M, et al. Uplink spectral efficiency analysis of multi-cell multi-user massive MIMO over correlated Ricean channel. Sci China Inf Sci, 2018, 61: 082305
4 Zhang Z S, Wang X, Zhang C Y, et al. Massive MIMO technology and challenges (in Chinese). Sci Sin Inform, 2015, 45: 1095–1110
5 Zhang Q, Jin S, Wong K K, et al. Power scaling of uplink massive MIMO systems with arbitrary-rank channel means. IEEE J Sel Top Signal Process, 2014, 8: 966–981
6 Dai J X, Wang J, Cheng C H, et al. Linear precoding based on a non-ideal Nakagami-m channel in a massive MIMO system. Sci China Inf Sci, 2017, 60: 069301
7 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
8 Xu S, Zhang H, Tian J, et al. Pilot reuse and power control of D2D underlaying massive MIMO systems for energy efficiency optimization. Sci China Inf Sci, 2017, 60: 100303
9 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
10 Zhang W C, Ren H, Pan C H, et al. Large-scale antenna systems with UL/DL hardware mismatch: achievable rates analysis and calibration. IEEE Trans Commun, 2015, 63: 1216–1229
11 Shepard C, Yu H, Anand N, et al. Argos: practical many-antenna base stations. In: Proceedings of ACM Conference MOBICOM, Turkey, 2012. 53–64
12 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
13 Vieira J, Rusek F, Tufvesson F. Reciprocity calibration methods for massive MIMO based on antenna coupling. In: Proceedings of IEEE GLOBECOM, Austin, 2014. 3708–3712
14 Liu D L, Ma W Z, Shao S H, et al. Performance analysis of TDD reciprocity calibration for massive MU-MIMO systems with ZF beamforming. IEEE Commun Lett, 2016, 20: 113–116
15 Luo X. Multiuser massive MIMO performance with calibration errors. IEEE Trans Wirel Commun, 2016, 15: 4521–4534
16 Mi D, Dianati M, Zhang L, et al. Massive MIMO performance with imperfect channel reciprocity and channel estimation error. IEEE Trans Commun, 2017, 65: 3734–3749
17 Ma W Z, Liu D L, Liu Y, et al. On the capacity of ZF beamforming in massive MIMO systems with imperfect reciprocity calibration. In: Proceedings of IEEE GLOBECOM, Singapore, 2017. 3708–3712
18 Minasian A, Adve R S, Shahbazpanahi S. The impact of hardware calibration errors on the performance of massive MIMO systems. In: Proceedings of IEEE GLOBECOM, Washington, 2016. 1–6
19 Bjornson E, Sanguinetti L, Hoydis J, et al. Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer? IEEE Trans Wirel Commun, 2015, 14: 3059–3075
20 Liu D L, Zhao B, Wu F, et al. Semi-blind SI cancellation for in-band full-duplex wireless communications. IEEE Commun Lett, 2018, 22: 1078–1081

21 Suzuki H, Tran T V A, Collings I B, et al. Transmitter noise effect on the performance of a MIMO-OFDM hardware implementation achieving improved coverage. IEEE J Sel Areas Commun, 2008, 26: 867–876
22 Jungnickel V, Manolakis K, Zirwas W, et al. The role of small cells, coordinated multipoint, and massive MIMO in 5G. IEEE Commun Mag, 2014, 52: 44–51
23 Jin S, Wang J, Sun Q, et al. Cell coverage optimization for the multicell massive MIMO uplink. IEEE Trans Veh Technol, 2015, 64: 5713–5727
24 Wei H X, Li Y Z, Xiao L M, et al. Queue-aware energy-efficient scheduling and power allocation with feedback reduction in small-cell networks. Sci China Inf Sci, 2018, 61: 048301
25 Choi J, Oh H, Jeon H C. Propagation prediction for LTE small cells with antenna beam tilt. In: Proceedings of VTC-Fall, Vancouver, 2014. 1–5

Appendix A

Applying singular value decomposition to $\boldsymbol{H}_{\rm ul}^{\rm T}$ yields $\boldsymbol{H}_{\rm ul}^{\rm T} = \tilde{\boldsymbol{Y}}\left(\boldsymbol{\Sigma}, \boldsymbol{0}\right)\tilde{\boldsymbol{X}}^{\rm H}$, where $\boldsymbol{\Sigma} = {\rm diag}\left(\chi_1, \ldots, \chi_K\right)$ is a diagonal matrix containing the non-negative real singular values of $\boldsymbol{H}_{\rm ul}^{\rm T}$. $\tilde{\boldsymbol{Y}}$ and $\tilde{\boldsymbol{X}}$ denote the left and right singular matrices for the corresponding singular values. Substituting $\boldsymbol{H}_{\rm ul}^{\rm T}$, we obtain $\boldsymbol{H}_{\rm ul}^{\rm T} = \tilde{\boldsymbol{Y}}\left(\boldsymbol{\Sigma}, \boldsymbol{0}\right)\tilde{\boldsymbol{X}}^{\rm H}$ into (10) yields

$$E\left[\left(\boldsymbol{u}\boldsymbol{u}^{\rm H}\right)_{k,k}\right] = \tfrac{1}{K}{\rm E}\left\{{\rm tr}\left[\tilde{\boldsymbol{X}}^{\rm H}\hat{\boldsymbol{D}}_{\rm J}^{\rm n}\tilde{\boldsymbol{X}}\begin{pmatrix}\Sigma^{-2} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\tilde{\boldsymbol{X}}^{\rm H}\left(\hat{\boldsymbol{D}}_{\rm J}^{\rm n}\right)^{\rm H}\tilde{\boldsymbol{X}}\begin{pmatrix}\Sigma^{2} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\right]\right\}. \tag{A1}$$

Define an $M \times M$ matrix

$$\boldsymbol{A} = \tilde{\boldsymbol{X}}^{\rm H}\hat{\boldsymbol{D}}_{\rm J}^{\rm n}\tilde{\boldsymbol{X}}\begin{pmatrix}\Sigma^{-2} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\tilde{\boldsymbol{X}}^{\rm H}\left(\hat{\boldsymbol{D}}_{\rm J}^{\rm n}\right)^{\rm H}\tilde{\boldsymbol{X}}.$$

One should note that $\boldsymbol{H}_{\rm ul}^{\rm T}\left(\boldsymbol{H}_{\rm ul}^{\rm T}\right)^{\rm H} \approx K\mathbf{I}_K$ for massive MIMO with $M \to \infty$ and small HM magnitudes, i.e., $\chi_k^2 \approx K$. Thus,

$${\rm E}\left[\left(\boldsymbol{u}\boldsymbol{u}^{\rm H}\right)_{k,k}\right] \approx \frac{1}{K}{\rm E}\left\{{\rm tr}\left[\boldsymbol{A}\begin{pmatrix}K\mathbf{I}_K & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\right]\right\} = \frac{1}{K}\times K{\rm E}\left\{\sum_{i=1}^{K}\boldsymbol{A}_{i,i}\right\} = {\rm E}\left\{\sum_{i=1}^{K}\boldsymbol{A}_{i,i}\right\}, \tag{A2}$$

where $\boldsymbol{A}_{i,i}$ denotes the ith diagonal element of $\boldsymbol{A}$.

Again, from $\chi_k^2 \approx K$, it is known that ${\rm tr}\left(\boldsymbol{A}\right) \approx \delta_{\rm e}$. As the elements of $\hat{\boldsymbol{D}}_{\rm J}^{\rm n}$ are ergodic and permutations of χ_k^2 can be random with equal probability, $\boldsymbol{A}_{i,i}$ contributes equally to ${\rm tr}\left(\boldsymbol{A}\right)$. Hence,

$${\rm E}\left\{\sum_{i=1}^{K}\boldsymbol{A}_{i,i}\right\} \approx \frac{K}{M}{\rm tr}\left(\boldsymbol{A}\right) = \frac{K\delta_{\rm e}}{M}, \tag{A3}$$

since $\boldsymbol{A}$ has M elements on the diagonal. By substituting (A3) and (A2) into (9), one obtains ${\rm E}\left(\xi_k\overline{\mathcal{I}}_k\right) \approx \xi_k\frac{K-1}{M}\delta_{\rm e}$, which concludes the proof.

Appendix B

From $\overline{\mathcal{N}}_k = \|\boldsymbol{W}\|_F^2\big/|\mu_k|^2$, we have

$${\rm E}\left(\overline{\mathcal{N}}_k\right) = {\rm E}\left[\left|\frac{d_{\rm ut}^k}{d_{\rm ur}^k}\right|^2{\rm tr}\left(\hat{\boldsymbol{D}}_{\rm J}\tilde{\boldsymbol{W}}\tilde{\boldsymbol{W}}^{\rm H}\hat{\boldsymbol{D}}_{\rm J}^{\rm H}\right)\right] = {\rm E}\left[\left|\frac{d_{\rm ut}^k}{d_{\rm ur}^k}\right|^2{\rm tr}\left(\hat{\boldsymbol{D}}_{\rm J}^{\rm H}\hat{\boldsymbol{D}}_{\rm J}\tilde{\boldsymbol{W}}\tilde{\boldsymbol{W}}^{\rm H}\right)\right]. \tag{B1}$$

One should note that ${\rm tr}(\hat{\boldsymbol{D}}_{\rm J}^{\rm H}\hat{\boldsymbol{D}}_{\rm J}\tilde{\boldsymbol{W}}\tilde{\boldsymbol{W}}^{\rm H})$ is a weighted summation of diagonal elements of $\tilde{\boldsymbol{W}}\tilde{\boldsymbol{W}}^{\rm H}$ with weighting factor given in $\hat{\boldsymbol{D}}_{\rm J}^{\rm H}\hat{\boldsymbol{D}}_{\rm J}$. The complicated uplink channel model indicated by (1) makes obtaining the mathematical expectations of (B1) demanding. In order to obtain explicit analytical results, we consider $\boldsymbol{H}_{\rm ul} \approx \boldsymbol{H}\boldsymbol{D}_{\rm ut}$ when calculating ${\rm E}\left(\overline{\mathcal{N}}_k\right)$ since the HM should be small due to the EVM requirements of modern transmitters. Thus,

$${\rm E}\left(\overline{\mathcal{N}}_k\right) \approx {\rm E}\left[\left|\left(\hat{\boldsymbol{D}}_{\rm J}\right)_{k,k}\right|^2\right]{\rm E}\left[\left|\frac{d_{\rm ut}^k}{d_{\rm ur}^k}\right|^2{\rm tr}\left(\tilde{\boldsymbol{W}}\tilde{\boldsymbol{W}}^{\rm H}\right)\right] = \frac{{\rm E}\left|\left(\hat{\boldsymbol{D}}_{\rm J}\right)_{k,k}\right|^2{\rm E}\left[\left|\frac{d_{\rm ut}^k}{d_{\rm ur}^k}\right|^2\sum_{k=1}^{K}\left|\frac{1}{d_{\rm ut}^k}\right|^2\right]}{(M-K)}, \tag{B2}$$

where Eq. (B2) is obtained from the fact that the elements of $\hat{\boldsymbol{D}}_{\rm J}$ are ergodic from [7]

$${\rm E}_{\rm H}\left[{\rm tr}\left(\tilde{\boldsymbol{W}}\tilde{\boldsymbol{W}}^{\rm H}\right)\right] = {\rm E}_{\rm H}\left[{\rm tr}\left(\tilde{\boldsymbol{W}}^{\rm H}\tilde{\boldsymbol{W}}\right)\right] = \frac{1}{(M-K)}\sum_{k=1}^{K}\left|\frac{1}{d_{\rm ut}^k}\right|^2 \tag{B3}$$

with ${\rm E}_{\rm H}$ denoting the expectation over small-scale fading in $\boldsymbol{H}$. It is worth mentioning that similar results can also be obtained for ${\rm E}\left(\overline{\mathcal{N}}_k\right)$ by following the approximation strategy [15, 18] for calculating the average of $\|\boldsymbol{W}\|_F^2$. The key of these approximations relies on the fact that HM has small values. The simulation results in this paper also show that the calculation in this appendix provides a good approximation.

Appendix C

Although deriving an explicit form of $\mathrm{E}_{\beta_k}\left(R_k^{\mathrm{LB}}\right)$ is usually infeasible, f_{se} can be numerically calculated using the Monte Carlo method:

$$f_{\mathrm{se}} \approx \frac{1}{N}\sum_{n=1}^{N} K \times R_n^{\mathrm{LB}}, \tag{C1}$$

where β_n in R_n^{LB} denotes the nth path loss realization from the distribution of user locations.

$$\begin{aligned}\frac{\partial K R_k^{\mathrm{LB}}}{\partial K} &= R_k^{\mathrm{LB}} + K\frac{A(K)}{B(K)} \\ &= R_k^{\mathrm{LB}} + \frac{\log_2 e\left[-\xi_k\delta_{\mathrm{e}}K^3 + 2\xi_k\delta_{\mathrm{e}}MK^2 - M^2\left(1+\xi_k\delta_{\mathrm{e}}\right)K\right]}{\xi_k\delta_{\mathrm{e}}K^3 - \left(2\xi_k\delta_{\mathrm{e}}M + M + \xi_k\delta_{\mathrm{e}}\right)K^2 + \left(\xi_k\delta_{\mathrm{e}}M^2 + M^2 + 2\xi_k\delta_{\mathrm{e}}M\right)K - \xi_k\delta_{\mathrm{e}}M^2}.\end{aligned} \tag{C2}$$

Hence, the convexity of $f_{\mathrm{se}}(K, E_s)$ can be observed from (C1). In this appendix, we show that $\frac{\partial^2 K R_k^{\mathrm{LB}}}{\partial K^2} < 0$ if $M \gg K$. First, $\frac{\partial K R_k^{\mathrm{LB}}}{\partial K}$ is given at the top of the next page, where $\frac{\partial R_k^{\mathrm{LB}}}{\partial K} = \frac{A(K)}{B(K)}$. Thus, $\frac{\partial^2 K R_k^{\mathrm{LB}}}{\partial K^2}$ can be written as

$$\begin{aligned}\frac{\mathrm{d}^2 R_k^{\mathrm{LB}}}{\mathrm{d}K^2} &= \frac{2A(K)B(K) + KA'(K)B(K) - KA(K)B'(K)}{B^2(K)} \\ &= \frac{B(K)\left[A(K) + KA'(K)\right] + A(K)\left[B(K) - KB'(K)\right]}{B^2(K)},\end{aligned} \tag{C3}$$

where $A'(K) = \frac{\partial A(K)}{\partial K}$ and $B'(K) = \frac{\partial B(K)}{\partial K}$.

Note that this paper considers large-scale fading to be dominated by path loss, as was the case in [19]. Thus, the instantaneous value of ξ_k is bounded. Based on the condition $M \gg K$, it is known that

$$\xi_k\delta_{\mathrm{e}}K^3 \ll \left(2\xi_k\delta_{\mathrm{e}}M + M + \xi_k\delta_{\mathrm{e}}\right)K^2 \ \left(2\xi_k\delta_{\mathrm{e}}M + M + \xi_k\delta_{\mathrm{e}}\right)K^2 \ll \left(\xi_k\delta_{\mathrm{e}}M^2 + M^2 + 2\xi_k\delta_{\mathrm{e}}M\right)K, \tag{C4}$$

$$\xi_k\delta_{\mathrm{e}}M^2 \ll \left(\xi_k\delta_{\mathrm{e}}M^2 + M^2 + 2\xi_k\delta_{\mathrm{e}}M\right)K,\ \xi_k\delta_{\mathrm{e}}K^2 \ll 2\xi_k\delta_{\mathrm{e}}MK \text{ and } 2\xi_{\mathrm{k}}\delta_{\mathrm{e}}\mathrm{MK} \ll \mathrm{M}^2\left(1+\xi_{\mathrm{k}}\delta_{\mathrm{e}}\right). \tag{C5}$$

Thus,

$$\begin{aligned}&A(K) + KA'(K) \approx -M^2\left(1+\xi_k\delta_{\mathrm{e}}\right),\ A(K) \approx -M^2\left(1+\xi_k\delta_{\mathrm{e}}\right), \\ &B(K) - KB'(K) = -2\xi_k\delta_{\mathrm{e}}K^3 - \xi_k\delta_{\mathrm{e}}M^2 + \left(2\xi_k\delta_{\mathrm{e}}M + M + \xi_k\delta\right)K^2 \text{ and } B(K) \\ &\qquad\qquad \approx \left(\xi_k\delta_{\mathrm{e}}M^2 + M^2 + 2\xi_k\delta_{\mathrm{e}}M\right)K.\end{aligned} \tag{C6}$$

It is worth mentioning that the exact forms of $A(K)$ and $B(K)$ are used when calculating $A'(K)$ and $B'(K)$.

One can verify that $B(K)\left[A(K) + KA'(K)\right] + A(K)\left[B(K) - KB'(K)\right] < 0$ conditioned on $M \gg K$. Hence, $\frac{\partial^2 K R_k^{\mathrm{LB}}}{\partial K^2} < 0$, which implies $f_{\mathrm{se}} \approx \frac{1}{N}\sum_{n=1}^{N} K \times R_n^{\mathrm{LB}}$ is concave.

Appendix D

Similarly, the convexity of η_{ee} can be observed from $\eta_{\mathrm{ee}} \approx \frac{1}{E_s}\frac{1}{N}\sum_{n=1}^{N} K \times R_n^{\mathrm{LB}}$. By calculating $\frac{\partial\left(K R_n^{\mathrm{LB}}/E_s\right)}{\partial E_s}$, we have

$$\frac{\partial\left(R_n^{\mathrm{LB}}/E_s\right)}{\partial E_s} = \frac{1}{E_s^2}\left[\frac{\frac{K}{M-K}\times\log_2 e}{\frac{\beta_n E_s}{\sigma^2}\frac{K-1}{M}\delta_{\mathrm{e}} + \frac{K(1+\delta_{\mathrm{e}})}{M-K}} - R_n^{\mathrm{LB}}\right]. \tag{D1}$$

Letting $\frac{\partial\left(R_n^{\mathrm{LB}}/E_s\right)}{\partial E_s} < 0$, i.e.,

$$\log_2\left[\frac{\exp\left(\frac{\frac{K}{M-K}}{\frac{\beta_n E_s}{\sigma^2}\frac{K-1}{M}\delta_{\mathrm{e}} + \frac{K}{M-K}(1+\delta_{\mathrm{e}})}\right)}{1 + \frac{\beta_n E_s}{\sigma^2}\frac{K-1}{M}\delta_{\mathrm{e}} + \frac{K}{M-K}\left(1+\delta_{\mathrm{e}}\right)}\right] < 0, \tag{D2}$$

we have $E_s \in \mathcal{U}_0 = \{E_s | E_s > \frac{\sigma^2}{\beta_n} \times Z_0\}$, where $Z_0 = \frac{MK}{2(K-1)(M-K)}\frac{\sqrt{(1+\delta_{\mathrm{e}})^2 + \frac{4\delta_{\mathrm{e}}(K-1)}{M}} - (1+\delta_{\mathrm{e}})}{\delta_{\mathrm{e}}}$ and the Taylor Series $e^x = 1 + x + 0.5x^2 + o\left(x^2\right)$ is used to obtain the explicit solution to (D2). One can verify that Z_0 is a monotonically decreasing function of δ_e for $\delta_{\mathrm{e}} > 0$. Thus, $\mathcal{U}_1 = \{E_s | E_s > \frac{\sigma^2}{\beta_n} Z_0|_{\delta_{\mathrm{e}}\to 0} = \frac{\sigma^2}{\beta_n}\frac{M}{(M-K)}\}$ is a subset of $\mathcal{U}_0$ since $Z_0 < \max(Z_0) = Z_0|_{\delta_{\mathrm{e}}\to 0}$.

Note that this paper considers $\gamma_n \gg 1$, which implies

$$\mathcal{U}_2 = \left\{ E_s |\, E_s \gg \frac{\sigma^2}{\beta_n}\frac{M}{M-K}\frac{M\left(1+\delta_{\mathrm{e}}\right)}{M - (K-1)\,\delta_{\mathrm{e}}} \right\}. \tag{D3}$$

Hence, $\mathcal{U}_2 \subseteq \mathcal{U}_1 \subseteq \mathcal{U}_0$, implying $\frac{R_n^{\mathrm{LB}}}{E_s}$ is decreasing with respect to E_s. As $\mathcal{U}_2 \subseteq \mathcal{U}_1 \subseteq \mathcal{U}_0$ holds and is independent of K, $\frac{R_n^{\mathrm{LB}}}{E_s}$ is decreasing with respect to E_s, regardless of K. In fact, in a reasonable range of $\frac{\beta_n E_s}{\sigma^2}$ values, $M \gg K$, $E_s \in \mathcal{U}_1$ can always be satisfied without requiring (D3).