Area Spectral Efficiency and Area Energy Efficiency Analysis in Massive MIMO systems

Yuanxue Xin, Dongming Wang, Xiaohu You

National Mobile Communications Research Lab., Southeast University, China

Abstract—We consider the downlink multi-user multi-cell massive MIMO systems, and our system model accounts for channel estimation, pilot contamination, and uniformly random user location distribution. An equivalent channel model is investigated which indicates that the users in various cells come through the same 'fast fading' when they reuse the same training signal in uplink pilot transmission phase. Assuming that the number of antennas at base station (BS) is large, under the proposed equivalent channel model, we derive the asymptotic approximations of area spectral efficiency (ASE) with zero-forcing (ZF) and regularized zero-forcing (RZF) techniques which are proven to be accurate via simulation results. Furthermore, with a realistic power consumption model considering not only transmit power but also the precoding processing part, we analyze the performance of area energy efficiency (AEE). In particular, AEE shows quasiconcave with the number of antennas, and it indicates that we can obtain the optimal number of antennas with the existing convex methods.

I. Introduction

With the data traffic evolution over the last years, it is a common belief that a new fifth generation (5G) of mobile communication networks will be deployed to satisfy the 100 times higher typical user data rate [1]. In addition, the power consumption of communication techniques and the corresponding excessive carbon emissions become serious environmental and economical concerns. Therefore, academia and industry engage in innovating new technologies and system architectures to increase both spectral efficiency (SE) and energy efficiency (EE) [2] [3]. In recent years, massive multiple-input multiple-output (MIMO) system has emerged to become an essential technique as it can improve SE by more than tenfold while simultaneously increase the EE.

While SE and EE are treated as the basic metrics in network performance assessment [4], some novel measurements should be introduced in order to verify the new information technologies and various cell topologies in new commination systems. Area spectral efficiency (ASE) was introduced to evaluate the performance of transmission with the introduced concept 'affected area' [5]. Meanwhile, area power consumption (Watt/Km^2) was used as a measurement in optimizing the deployment of BS for cellular networks [6]. Similar to ASE, the area energy efficiency (AEE) in bit/Joule/Km^2 is considered in massive MIMO.

In this work, we analyze the ASE and AEE for downlink transmission in multi-cell multi-user massive MIMO systems with pilot contamination. Previous works [7] [8] have derived asymptotically tight approximations of the achievable rates with several linear precoders in non-cooperative TDD systems.
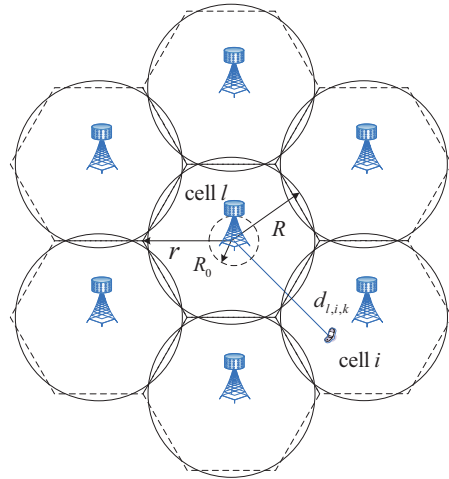

Fig. 1. Massive MIMO system.

Besides, [9] studied the closed-form expression for the achievable rates in downlink large-scale distributed antennas systems with maximum ratio transmission (MRT) precoder. In massive MIMO systems, as the number of antennas at BS tends to infinity, the asymptotic spectral efficiency in downlink with MRT has been studied [10] [11]. However, the zero-forcing (ZF) precoding and regularized zero-forcing (RZF) precoding need to compute the pseudoinverse of large matrices, which brings difficulties to achieve the asymptotic approximations easily or obtain closed-form expression of SE feasibly.

II. System Model

As depicted in Fig.1, we consider the downlink of a multi-user MIMO system composed of L hexagonal cells with radius r. Each cell consists of a central M antenna base station (BS) and K randomly and uniformly distributed single-antenna users. According to the same area assumption in [12], each hexagonal cell approximates to an equal area circle with radius R, where $R = \sqrt{3\sqrt{3}/(2\pi)}r$. R_0 denotes the closest distance of the user from BS. The users are assumed to be independently and uniformly distributed (i.i.d) in all cells.

A. Channel model

During a coherence block, we consider a flat-fading channel which is defined as

$$\boldsymbol{g}_{i,l,k} = \sqrt{\lambda_{i,l,k}}\boldsymbol{h}_{i,l,k}, \qquad (1)$$

where $\lambda_{i,l,k} = cd_{i,l,k}^{-\alpha}$ denotes the propagation loss, it depends on the distance between the k-th user in the l-th cell and the antennas of BS in the i-th cell. c is the median of mean path gain at a reference distance $d_{i,l,k} = 1\text{km}$, and the path loss exponent is α. $\boldsymbol{h}_{i,l,k}$ represents the small-scale fading vector whose elements are i.i.d zero mean circularly symmetric complex Gaussian (ZMCSCG) random variables with unit variance. For notation convenience, we define

$$\boldsymbol{G}_{i,l} = [\boldsymbol{g}_{i,l,1}, \cdots, \boldsymbol{g}_{i,l,K}]. \tag{2}$$

In order to acquire the channel state information (CSI) at the BS, we consider an uplink pilot sequence before transmitting downlink signal, and use the channel estimation to design precoding schemes by exploiting channel reciprocity property.

B. Equivalent channel model for estimated channel

Limited by the insufficient training resources, L BSs reuse the same set of mutually orthogonal pilot sequences $\boldsymbol{X}_\text{P}$. For analytical convenience, $\boldsymbol{X}_\text{P}$ is assumed to be an identity matrix by $K \times K$. During the training part, all users simultaneously transmit the pilot sequences, therefore, the $M \times K$ received pilot matrix at BS l is

$$\boldsymbol{Y}_{\text{P},l} = \boldsymbol{G}_{l,l}\boldsymbol{X}_\text{P} + \sum_{i \neq l} \boldsymbol{G}_{l,i}\boldsymbol{X}_\text{P} + \boldsymbol{Z}_\text{P}, \tag{3}$$

where $\boldsymbol{Z}_\text{P}$ is an $M \times K$ noise matrix and each element is an i.i.d. ZMCSCG random variable with variance γ_P. For $k \neq k'$, it is obvious $\mathcal{E}(\boldsymbol{g}_{l,l,k}\boldsymbol{g}_{l,i,k'}) = 0$, thus, the received pilot signal can be analyzed in columns separately as the follows

$$\boldsymbol{y}_{\text{P},l,k} = \boldsymbol{g}_{l,l,k} + \sum_{i \neq l} \boldsymbol{g}_{l,i,k} + \boldsymbol{z}_{\text{P},k},$$

where $\boldsymbol{y}_{\text{P},l,k}$ and $\boldsymbol{z}_{\text{P},k}$ are the k-th columns of $\boldsymbol{Y}_\text{P}$ and $\boldsymbol{Z}_\text{P}$, respectively. When joint multi-cell minimum-mean-squared-error (MMSE) channel estimation is used at the BS in all cells, the estimated channel vector $\boldsymbol{g}_{i,l,k}$ is modeled by [13]

$$\hat{\boldsymbol{g}}_{i,l,k} = \sqrt{\beta_{i,l,k}}\hat{\boldsymbol{h}}_{i,k},$$

where

$$\hat{\boldsymbol{h}}_{i,k} \sim \mathcal{CN}(0, \boldsymbol{I}_M),$$

and

$$\beta_{i,l,k} = \frac{\lambda_{i,l,k}^2}{\sum\limits_{l=1}^{L} \lambda_{i,l,k} + \gamma_\text{P}}.$$

Channel estimation error is defined as

$$\tilde{\boldsymbol{g}}_{i,l,k} \overset{\Delta}{=} \boldsymbol{g}_{i,l,k} - \hat{\boldsymbol{g}}_{i,l,k}.$$

Moreover, $\tilde{\boldsymbol{g}}_{i,l,k}$ and $\hat{\boldsymbol{g}}_{i,l,k}$ are in general uncorrelated as a property of linear MMSE estimator, and in this case the covariance matrix of $\tilde{\boldsymbol{g}}_{i,l,k}$ can be computed as

$$\boldsymbol{\Sigma}_{i,l,k} = (\lambda_{i,l,k} - \beta_{i,l,k})\boldsymbol{I}_M.$$

Once the channel is estimated, the channel matrix in (2) can be rewritten as

$$\boldsymbol{G}_{i,l} = \hat{\boldsymbol{G}}_{i,l} + \tilde{\boldsymbol{G}}_{i,l}, \tag{4}$$

where

$$\hat{\boldsymbol{G}}_{i,l} = [\hat{\boldsymbol{g}}_{i,l,1} \cdots \hat{\boldsymbol{g}}_{i,l,K}], \quad \tilde{\boldsymbol{G}}_{i,l} = [\tilde{\boldsymbol{g}}_{i,l,1} \cdots \tilde{\boldsymbol{g}}_{i,l,K}]. \tag{5}$$

Furthermore, the proposed equivalent estimated channel matrix $\hat{\boldsymbol{G}}_{l,i}$ in (5) can be expressed as

$$\hat{\boldsymbol{G}}_{l,i} = \hat{\boldsymbol{H}}_l \boldsymbol{\Lambda}_{l,i}, \tag{6}$$

where

$$\hat{\boldsymbol{H}}_l = \left[\hat{\boldsymbol{h}}_{l,1} \cdots \hat{\boldsymbol{h}}_{l,K} \right],$$

and

$$\boldsymbol{\Lambda}_{l,i} = \text{diag}\left[\sqrt{\beta_{l,i,1}}, \cdots, \sqrt{\beta_{l,i,K}} \right].$$

Note that the equivalent channel in (6) implies that the random part [1] of estimated channel is independent of the location of users. In the eyes of BS, users in different cells experience the same 'fast fading' when they share the same pilot sequence.

III. ASE ANALYSIS

A. Achievable downlink sum-rates

With the equivalent channel in (6), we consider three conventional linear precoders MRT, ZF and RZF, and the $M \times K$ linear precoding matrix is,

$$\boldsymbol{W}_l = \begin{cases} \hat{\boldsymbol{H}}_l \boldsymbol{\Lambda}_{l,l} & \text{for MRT} \\ \hat{\boldsymbol{H}}_l \boldsymbol{\Lambda}_{l,l} \left(\boldsymbol{\Lambda}_{l,l} \hat{\boldsymbol{H}}_l^\text{H} \hat{\boldsymbol{H}}_l \boldsymbol{\Lambda}_{l,l} \right)^{-1} & \text{for ZF} \\ \left(\hat{\boldsymbol{H}}_l \sum_{i=1}^{L} \boldsymbol{\Lambda}_{l,i}^2 \hat{\boldsymbol{H}}_l^\text{H} + \sigma_l \boldsymbol{I}_M \right)^{-1} \hat{\boldsymbol{H}}_l \boldsymbol{\Lambda}_{l,l} & \text{for RZF} \end{cases} \tag{7}$$

Let the signal vector transmitted to users by the BS in the l-th cell be $\boldsymbol{x}_l$. The noisy signal vector received by all users in the l-th cell is

$$\boldsymbol{y}_l = \sqrt{\rho_l}\boldsymbol{G}_{l,l}^\text{H}\boldsymbol{W}_l\boldsymbol{x}_l + \sum_{i \neq l} \sqrt{\rho_i}\boldsymbol{G}_{i,l}^\text{H}\boldsymbol{W}_i\boldsymbol{x}_i + \boldsymbol{z}_l,$$

where $\boldsymbol{z}_l$ is the complex additive white Gaussian noise vector, and the entries of $\boldsymbol{z}_l$ are i.i.d ZMCSCG random variables with variance γ_DL. The normalization factor ρ_l is used to constrain the average transmit power in the l-th cell, i.e.,

$$\rho_l = \frac{P_\text{T}}{\mathcal{E}\left[\text{Tr}\left(\boldsymbol{W}_l \boldsymbol{W}_l^\text{H} \right) \right]}, \tag{8}$$

and P_T is the transmit power at the BS.

In practical systems with strict time-delay requirement and high level of accuracy needs, users can not get the exact

---

[1] Although users are randomly and uniformly distributed, we assume that the large-scale fading is known to the BS.

instantaneous CSI. Assuming that the statistical CSI is known to users as in [14], then

$$
\begin{aligned}
y_{l,k} = {} & \sqrt{\rho_l}\mathcal{E}\left(\boldsymbol{g}_{l,l,k}^{\mathrm{H}}\boldsymbol{W}_l\boldsymbol{e}_k\right)x_{l,k} \\
& + \sqrt{\rho_l}\left[\boldsymbol{g}_{l,l,k}^{\mathrm{H}}\boldsymbol{W}_l\boldsymbol{e}_k - \mathcal{E}\left(\boldsymbol{g}_{l,l,k}^{\mathrm{H}}\boldsymbol{W}_l\boldsymbol{e}_k\right)\right]x_{l,k} \\
& + \sum_{\substack{i,j \\ (i,j)\neq(l,k)}} \sqrt{\rho_i}\boldsymbol{g}_{i,l,k}^{\mathrm{H}}\boldsymbol{W}_i\boldsymbol{e}_j x_{i,j} + z_{l,k},
\end{aligned}
\tag{9}
$$

where $\boldsymbol{e}_k$ stands for an $K \times 1$ vector with the k-th element is 1 and others are 0, $y_{l,k}$ and $x_{l,k}$ denote the k-th entry of $\boldsymbol{y}_l$ and $\boldsymbol{x}_l$, respectively. In the first term of right-hand side (RHS) of (9), the expectation only depends on the property of channel distribution. Then, the following rate is achievable

$$
R_l = \sum_{k=1}^{K} \log_2\left(1 + \mathrm{SINR}_{l,k}\right), \tag{10}
$$

where

$$
\mathrm{SINR}_{l,k} =
$$
$$
\frac{\rho_l\left|\mathcal{E}\left(\boldsymbol{g}_{l,l,k}^{\mathrm{H}}\boldsymbol{W}_l\boldsymbol{e}_k\right)\right|^2}{\rho_l\mathrm{var}\left(\boldsymbol{g}_{l,l,k}^{\mathrm{H}}\boldsymbol{W}_l\boldsymbol{e}_k\right) + \sum_{\substack{i,j \\ (i,j)\neq(l,k)}} \rho_i\mathcal{E}\left(\left|\boldsymbol{g}_{i,l,k}^{\mathrm{H}}\boldsymbol{W}_i\boldsymbol{e}_j\right|^2\right) + \gamma_{\mathrm{DL}}}.
\tag{11}
$$

According to the definition of ASE, we obtain the ASE as follows

$$
C_l = \frac{R_l}{\pi\left(R^2 - R_0^2\right)}. \tag{12}
$$

B. ASE asymptotic analysis

Under the equivalent channel, we focus on the downlink transmission with ZF precoding and RZF precoding with unlimited number of antennas at BS.

Proposition 1: When using ZF precoding and assuming that M is large, the downlink SINR can be approximated by the follows

$$
\overline{\mathrm{SINR}}_{l,k} = \frac{\overline{\rho}_l}{\sum_{i\neq l}\overline{\rho}_i\left(\frac{\lambda_{i,l,k}^2}{\lambda_{i,i,k}^2}\right) + \sum_{i,j}\overline{\rho}_i\left(\frac{\lambda_{i,l,k}-\beta_{i,l,k}}{M\beta_{i,i,j}}\right) + \gamma_{\mathrm{DL}}}, \tag{13}
$$

where

$$
\overline{\rho}_l = \frac{MP_{\mathrm{T}}}{\sum_{k=1}^{K}\beta_{l,l,k}^{-1}}. \tag{14}
$$

Proof: See Appendix A. ∎

Proposition 2: When using RZF precoding and assuming that M is large, the downlink SINR can be approximated by (15) where

$$
\mu_{l,k} = \left(M\sum_{i=1}^{L}\beta_{l,i,k} + \sigma_l\right)^{-1}, \tag{16}
$$

and

$$
\sigma_l = \sum_{i=1}^{L}\sum_{k=1}^{K}\left(\lambda_{l,i,k} - \beta_{l,i,k}\right). \tag{17}
$$

And the normalized power factor is

$$
\overline{\rho}_l = \frac{P_{\mathrm{T}}}{M\sum_{k=1}^{K}\beta_{l,l,k}\mu_{l,k}^2}. \tag{18}
$$

Proof: Applying Searle identity to the above RZF precoding matrix in (7) yields

$$
\boldsymbol{W}_l = \hat{\boldsymbol{H}}_l\left(\sum_{i=1}^{L}\boldsymbol{\Lambda}_{l,i}^2\hat{\boldsymbol{H}}_l^{\mathrm{H}}\hat{\boldsymbol{H}}_l + \sigma_l\boldsymbol{I}_K\right)^{-1}\boldsymbol{\Lambda}_{l,l},
$$

The following proofs are similar to the proofs of *Proposition 1*, and we omit the details here. ∎

Substituting the above asymptotic SINRs into (10), we can obtain the desired approximate ASE, which is

$$
\overline{C}_l = \frac{1}{\pi\left(R^2 - R_0^2\right)}\sum_{k=1}^{K}\log_2\left(1 + \overline{\mathrm{SINR}}_{l,k}\right). \tag{19}
$$

IV. AEE ANALYSIS

The AEE in a communication system is used to characterize the area data rate per power expenditure, which is measured in bits/Joule/km^2. AEE is defined as the follows

$$
\eta = \frac{BC_l}{P_{\mathrm{tot}}}, \tag{20}
$$

where P_{tot} denotes the overall power consumption, and B is the available bandwidth.

A. Power consumption model

Based on a practical power consumption model in [15], the power consumption model can be categorized as follows: (a)RF power at the BS(P_{T}), which has a reciprocal of drain efficiency of power amplifier (ξ), (b)internal non-RF power expended by each service-antenna (P_{BS}), (c)the power required by the circuit components of each terminal (P_{UE}),(d)the power consumption contributes to operating precoding (P_{PR}), (e)the basic power consumed at BS independent of M (P_0). Thus, the total system power consumption is given by

$$
P_{\mathrm{tot}} = \xi P_{\mathrm{T}} + MP_{\mathrm{BS}} + KP_{\mathrm{UE}} + P_{\mathrm{PR}} + P_0. \tag{21}
$$

$$
\overline{\mathrm{SINR}}_{l,k} = \frac{M^2\overline{\rho}_l\mu_{l,k}^2}{\sum_{i\neq l}M^2\overline{\rho}_i\beta_{i,l,k}\beta_{i,i,k}\mu_{i,k}^2 + \sum_{i,j}\overline{\rho}_iM\left(\lambda_{i,l,k} - \hat{\beta}_{i,l,k}\right)\beta_{i,i,j}\mu_{i,j}^2 + \gamma_{\mathrm{DL}}}, \tag{15}
$$

B. Precoding power consumption discussion

We assume the computational efficiency is L_{BS} operations per Joule at the BS (measured in Gflops/Watt), and the precoding is performed once per coherence block. During downlink transmission, there are U coherence blocks per second. Hence, the P_{PR} can be detailed as the follows,

$$P_{\text{PR}} = U \frac{C_{\text{PR}}}{L_{\text{BS}}}, \tag{22}$$

where C_{PR} denotes the floating-point operations required to carry out precoding, and it is obvious that the computation complexity strongly depends on the choice of precoding scheme. In [16], it analyzed the computational complexities of MRT, ZF, RZF precoding in a single cell MIMO system, and it approximated the RZF precoding operations as a simple result due to its high computational complexity. We give a brief summary of the specified C_{PR} in multi-cell multi-user MIMO systems.

During downlink data transmission, the precoding matrix is multiplied with the signal vector, hence it costs $3MK$ operations per coherence period when using MRT precoding. As for the inverse matrix computation, we consider the LU-based matrix inversion in [17]. Hence, we can infer that the total numbers of complex operations to be computed for implementing ZF and RZF precoding are $\frac{2K^3}{3} + 4K^2M$ and $\frac{2M^3}{3} + 4M^2K$, respectively.

C. AEE Analysis of Massive MIMO

Proposition 3: AEE function is strictly quasiconcave with M, and a unique global optimal M always exists. Moreover, η strictly increases and then decreases with M.

Proof: For ZF precoding, from (13) we can conclude that $\overline{C}_l$ is a strictly concave function w.r.t M. For RZF precoding, when M is unlimited, we have the follows from (16),

$$\lim_{M \to \infty} M\mu_{l,k} = \left(\sum_{i=1}^{L} \beta_{l,i,k} \right)^{-1}. \tag{23}$$

Substituting (23) into (18) and (15), we can easily conclude that $\overline{C}_l$ for RZF precoding is also a strictly concave function w.r.t M. Moreover, the power consumption function is a linear function with M. As a result, η for different precoding schemes are strictly quasiconcave functions. For any strictly quasiconcave function, a unique global maximum always exists, and a local maximum is also globally optimal. To obtain the monotonicity of η, it is easy to show that $\lim_{M \to 0} \eta = 0$, and $\lim_{M \to \infty} \eta = 0$. As η is supposed to be positive and strictly quasiconcave, we can conclude that η either first strictly increases and then strictly decreases with M. ∎

V. NUMERICAL ILLUSTRATIONS

As shown in Fig.1, we consider a hexagonal system with $L = 7$, $r = 1$km, and assume that users are uniformly random distributed. We use the modified COST231 Hata urban propagation model, and we neglect the shadowing fading for

analytical simplicity. The corresponding simulation parameters given in Table I and are inspired by a variety of prior works [18] [19] [20].

TABLE I
SYSTEM PARAMETERS

Parameters	Value
Median of mean path gain:c	-140.7dB
Path loss exponent: α	3.5
Thermal noise: $\gamma_{\text{P}}, \gamma_{\text{DL}}$	-174dBm/Hz
power amplifier efficiency: ξ	1/0.39
Radius of cell: r	1km
Closest distance to BS:R_0	30m
Power consumption of a antenna at BS: P_{BS}	32mW
Power consumption of each user: P_{UE}	0.3W
Basic power consumption at BS:P_0	10W
Transmit power :$P_{\max}$	30W
Number of users :K	36
Transmission bandwidth :B	20MHz
Computational efficiency :L_{BS}	7.4 Gflops/W
Number of coherence blocks per sec : U	1800

Fig.2 shows the set of achievable ASE values with different precoding schemes, and verifies the accuracy of the asymptotic approximation results in Section III for ZF and RZF precoding techniques. For the MRT approximation in massive MIMO systems, we refer to [9]. We can observe that RZF leads to significant performance gains over the other two precoding technologies as it reduces the multiuser interference effectively. We can see that the approximations by (19) for both ZF and RZF precoding match well with the Monte Carlo simulations as the number of antennas grows. Note that the gaps between C_l and $\overline{C}_l$ for both RZF and ZF precodings are more obvious than for MRT precoding, as the approximate method for MRT precoding is different from we used in this paper. Unfortunately the more accurate approximate method in [9] can not be extended to RZF and ZF precodings.
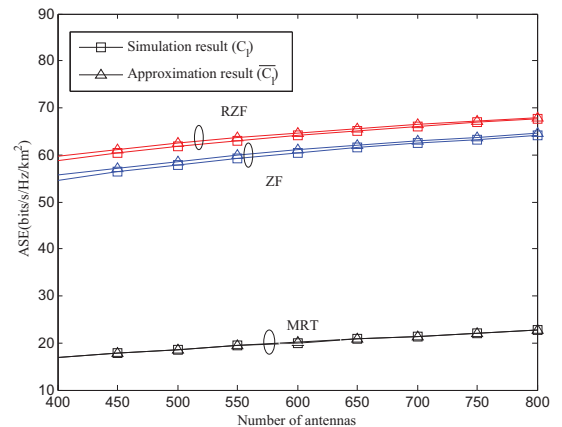
Fig. 2. Area spectral efficiency versus number of antennas.

Fig.3. describes the AEE of various precoding strategies in massive MIMO systems and validates the tendency of AEE as discussed in Section IV. While RZF leads to the superior ASE as plotted in Fig.1, RZF precoding performs
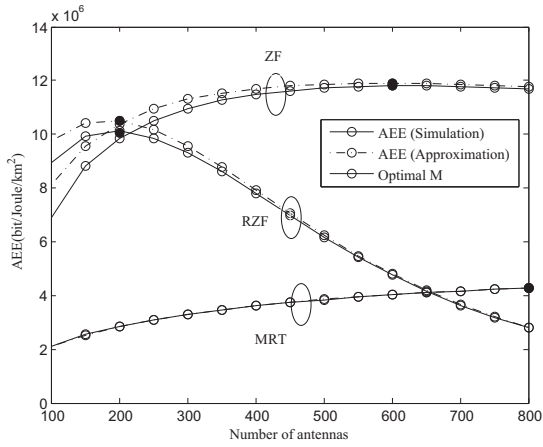
Fig. 3. Area spectral efficiency versus number of antennas.

best in AEE comparing with MRT and ZF when BS deploys relatively few antennas. However, AEE for RZF decreases rapidly when the number of antennas at BS increases. It is reasonable as the expenditure of power for RZF precoding operations is proportional to M^3 and non-RF power expended by service antennas grows dramatically, and consequently the ASE gain benefited from the growing M can not compensate the extra power consumption. With large number of antennas, ZF precoding shows better performance in AEE as a result of the moderate power consumption and relatively favorable performance in ASE. In comparison, we also depict AEE curves when using MRT precoding. Over the range of the M plotted in Fig.3, MRT precoding shows the smallest value of AEE which dues to the the fact that MRT brings the worst performance in ASE. In addition, Fig.3 confirms that there exists an optimal M for maximum AEE. Since we use the different asymptotic analysis methods, the approximation curves for RZF and ZF percodings are distinguishable from the simulation results when M is small, but MRT approximation line is closed to the simulation one. However, the optimal numbers of antennas for both approximation and simulation results appear the same.

VI. CONCLUSIONS

We have explored the equivalent channel model which can be viewed as a diagonal matrix multiplies by another matrix whose elements are i.i.d complex Gaussian with zero mean and unit variance. When Assuming a large number of antennas at BS, we have deduced asymptotically tight approximations of achievable downlink-rates by ZF and RZF precoding strategies under the proposed equivalent channel model. These approximations are shown to be accurate, especially at large number of BS antennas. Besides, RZF outperforms ZF and MRT from the point of view of ASE. With a realistic power model who takes the precoding process complexity into account, it shows that there always exists optimal number of antennas at BS to guarantee maximum AEE. The results bring insights on

designing an energy-effective MIMO system by considering AEE as a main factor.

APPENDIX A
USEFUL LEMMAS

Lemma 1:
Let $\boldsymbol{A} \in \mathbb{C}^{M \times M}$, and assume that $\boldsymbol{A}$ has uniformly bounded spectral norm. Let $\boldsymbol{x} \sim \mathcal{CN}\left(0, \frac{1}{M}\boldsymbol{I}_M\right)$, and it is independent of $\boldsymbol{A}$. According to [21, Theorem 3.4], we obtain

$$\boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x} - \frac{1}{M}\mathrm{tr}\boldsymbol{A} \xrightarrow[M \to \infty]{\text{a.s.}} 0.$$

Lemma 2: Let $\boldsymbol{A} \in \mathbb{C}^{M \times K}$ be a standard complex Gaussian matrix, then

$$\frac{1}{M}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A} - \boldsymbol{I}_{\mathrm{K}} \xrightarrow[M \to \infty]{\text{a.s.}} 0.$$

APPENDIX B
PROOFS OF PROPOSITION 1

Proof: From the ZF precoding matrix in (7), we have

$$M\mathcal{E}\left[\mathrm{Tr}\left(\boldsymbol{W}_l\boldsymbol{W}_l^{\mathrm{H}}\right)\right]$$
$$= M\mathcal{E}\left\{\mathrm{Tr}\left[\hat{\boldsymbol{H}}_l\boldsymbol{\Lambda}_{l,l}\left(\boldsymbol{\Lambda}_{l,l}^2\hat{\boldsymbol{H}}_l^{\mathrm{H}}\hat{\boldsymbol{H}}_l\right)^{-1}\left(\boldsymbol{\Lambda}_{l,l}^2\hat{\boldsymbol{H}}_l^{\mathrm{H}}\hat{\boldsymbol{H}}_l\right)^{-1}\boldsymbol{\Lambda}_{l,l}\hat{\boldsymbol{H}}_l^{\mathrm{H}}\right]\right\}$$
$$\stackrel{(a)}{=} M\mathcal{E}\left\{\mathrm{Tr}\left[\boldsymbol{\Lambda}_{l,l}^2\hat{\boldsymbol{H}}_l^{\mathrm{H}}\hat{\boldsymbol{H}}_l\left(\boldsymbol{\Lambda}_{l,l}^2\hat{\boldsymbol{H}}_l^{\mathrm{H}}\hat{\boldsymbol{H}}_l\right)^{-1}\left(\boldsymbol{\Lambda}_{l,l}^2\hat{\boldsymbol{H}}_l^{\mathrm{H}}\hat{\boldsymbol{H}}_l\right)^{-1}\right]\right\}$$
$$= \mathcal{E}\left\{\mathrm{Tr}\left[\left(\frac{1}{M}\boldsymbol{\Lambda}_{l,l}^2\hat{\boldsymbol{H}}_l^{\mathrm{H}}\hat{\boldsymbol{H}}_l\right)^{-1}\right]\right\}, \tag{B1}$$

where (a) follows from the property that the matrices in a trace of a product can be switched. By *Lemma 1*, straight-forward computation yields to

$$\mathcal{E}\left\{\mathrm{Tr}\left[\left(\frac{1}{M}\boldsymbol{\Lambda}_{l,l}^2\hat{\boldsymbol{H}}_l^{\mathrm{H}}\hat{\boldsymbol{H}}_l\right)^{-1}\right]\right\} - \sum_{k=1}^{K}\frac{1}{\beta_{l,l,k}} \xrightarrow[M \to \infty]{} 0. \tag{B2}$$

Hence, average transmit power in (8) tends to be $\bar{\rho}_l$ as shown in (14).

1) Interference power: Note that $\boldsymbol{W}_l$ is independent of the channel estimation error, thus,

$$\text{var}\left(\boldsymbol{g}_{l,l,k}^{\text{H}}\boldsymbol{W}_{l,k}\boldsymbol{e}_k\right) = \mathcal{E}\left(\left|\hat{\boldsymbol{g}}_{l,l,k}^{\text{H}}\boldsymbol{W}_l\boldsymbol{e}_k\right|^2\right) - \mathcal{E}^2\left(\hat{\boldsymbol{g}}_{l,l,k}^{\text{H}}\boldsymbol{W}_l\boldsymbol{e}_k\right)$$
$$+ \mathcal{E}\left(\left|\tilde{\boldsymbol{g}}_{l,l,k}^{\text{H}}\boldsymbol{W}_l\boldsymbol{e}_k\right|^2\right). \tag{B3}$$

$$\mathcal{E}\left(\left|\hat{\boldsymbol{g}}_{i,l,k}^{\text{H}}\boldsymbol{W}_i\boldsymbol{e}_j\right|^2\right) = \mathcal{E}\left(\left|\hat{\boldsymbol{g}}_{i,l,k}^{\text{H}}\boldsymbol{W}_i\boldsymbol{e}_j\right|^2\right) + \mathcal{E}\left(\left|\tilde{\boldsymbol{g}}_{i,l,k}^{\text{H}}\boldsymbol{W}_i\boldsymbol{e}_j\right|^2\right). \tag{B4}$$

We focus on the first term in the RHS of (B4), which is

$$\mathcal{E}\left(\left|\hat{\boldsymbol{g}}_{i,l,k}^{\text{H}}\boldsymbol{W}_i\boldsymbol{e}_j\right|^2\right)$$
$$= \mathcal{E}\left[\left|\boldsymbol{e}_k^{\text{H}}\boldsymbol{\Lambda}_{i,l}\hat{\boldsymbol{H}}_i^{\text{H}}\hat{\boldsymbol{H}}_i\boldsymbol{\Lambda}_{i,i}\left(\boldsymbol{\Lambda}_{i,i}\hat{\boldsymbol{H}}_i^{\text{H}}\hat{\boldsymbol{H}}_i\boldsymbol{\Lambda}_{i,i}\right)^{-1}\boldsymbol{e}_j\right|^2\right],$$
$$= \mathcal{E}\left[\left|\boldsymbol{e}_k^{\text{H}}\boldsymbol{\Lambda}_{i,l}\boldsymbol{\Lambda}_{i,i}^{-1}\boldsymbol{e}_j\right|^2\right] \tag{B5}$$

Thus, we obtain

$$\mathcal{E}\left(\left|\hat{\boldsymbol{g}}_{i,l,k}^{\text{H}}\boldsymbol{W}_i\boldsymbol{e}_j\right|^2\right) = \begin{cases} \frac{\lambda_{i,l,k}^2}{\lambda_{i,i,k}^2}, & k = j \\ 0. & k \neq j \end{cases} \tag{B6}$$

As for the second term of RHS in (B3), we have

$$\mathcal{E}\left(\hat{\boldsymbol{g}}_{l,l,k}^{\text{H}}\boldsymbol{W}_l\boldsymbol{e}_k\right)$$
$$= \mathcal{E}\left[\boldsymbol{e}_k^{\text{H}}\boldsymbol{\Lambda}_{l,l}\hat{\boldsymbol{H}}_l^{\text{H}}\hat{\boldsymbol{H}}_l\boldsymbol{\Lambda}_{l,l}\left(\boldsymbol{\Lambda}_{l,l}\hat{\boldsymbol{H}}_l^{\text{H}}\hat{\boldsymbol{H}}_l\boldsymbol{\Lambda}_{l,l}\right)^{-1}\boldsymbol{e}_k\right],$$
$$= 1. \tag{B7}$$

Hence

$$\text{var}\left(\boldsymbol{g}_{l,l,k}^{\text{H}}\boldsymbol{W}_{l,k}\boldsymbol{e}_k\right) = \mathcal{E}\left(\left|\tilde{\boldsymbol{g}}_{l,l,k}^{\text{H}}\boldsymbol{W}_l\boldsymbol{e}_k\right|^2\right). \tag{B8}$$

For the second term in the RHS of (B4), we have

$$\mathcal{E}\left(\left|\tilde{\boldsymbol{g}}_{i,l,k}^{\text{H}}\boldsymbol{W}_i\boldsymbol{e}_j\right|^2\right)$$
$$= \mathcal{E}\left(\tilde{\boldsymbol{g}}_{i,l,k}^{\text{H}}\boldsymbol{W}_i\boldsymbol{e}_j\boldsymbol{e}_j^{\text{H}}\boldsymbol{W}_i^{\text{H}}\tilde{\boldsymbol{g}}_{i,l,k}\right)$$
$$\overset{(a)}{\rightarrow} \left(\lambda_{i,l,k} - \beta_{i,l,k}\right)\text{Tr}\left[\mathcal{E}\left(\boldsymbol{W}_i\boldsymbol{e}_j\boldsymbol{e}_j^{\text{H}}\boldsymbol{W}_i^{\text{H}}\right)\right]$$
$$= \left(\lambda_{i,l,k} - \beta_{i,l,k}\right)\mathcal{E}\left(\boldsymbol{e}_j^{\text{H}}\boldsymbol{W}_i^{\text{H}}\boldsymbol{W}_i\boldsymbol{e}_j\right),$$
$$= \left(\lambda_{i,l,k} - \beta_{i,l,k}\right)\mathcal{E}\left[\boldsymbol{e}_j^{\text{H}}\left(\boldsymbol{\Lambda}_{i,i}\hat{\boldsymbol{H}}_i^{\text{H}}\hat{\boldsymbol{H}}_i\boldsymbol{\Lambda}_{i,i}\right)^{-1}\boldsymbol{e}_j\right] \tag{B9}$$

where (a) follows from *Lemma 1*. Implying *Lemma 2*, we have

$$M\mathcal{E}\left(\left|\tilde{\boldsymbol{g}}_{i,l,k}^{\text{H}}\boldsymbol{W}_i\boldsymbol{e}_j\right|^2\right) - \frac{\lambda_{i,l,k} - \beta_{i,l,k}}{\beta_{i,i,j}} \xrightarrow[M\rightarrow\infty]{} 0. \tag{B10}$$

2)Desired signal power: As the independence of $\boldsymbol{W}_l\boldsymbol{e}_k$ and $\tilde{\boldsymbol{g}}_{l,l,k}$,

$$\left|\mathcal{E}\left(\boldsymbol{g}_{l,l,k}^{\text{H}}\boldsymbol{W}_l\boldsymbol{e}_k\right)\right|^2 = \left|\mathcal{E}\left(\hat{\boldsymbol{g}}_{l,l,k}^{\text{H}}\boldsymbol{W}_l\boldsymbol{e}_k\right)\right|^2. \tag{B11}$$

From (B7), we can obtain

$$\left|\mathcal{E}\left(\boldsymbol{g}_{l,l,k}^{\text{H}}\boldsymbol{W}_l\boldsymbol{e}_k\right)\right|^2 = 1. \tag{B12}$$

Replacing the asymptotic approximations for the desired signal power and the interference power in (11) concludes the proof. ∎

REFERENCES

[1] Z. Ma, Z. Zhang, Z. Ding, P. Fan, and H. Li, "Key techniques for 5G wireless communications:network architecture, physical layer, and MAC layer perspectives," *SCIENCE CHINA Information Sciences*, vol. 58, no. 4, p. 41301, 2015.
[2] D. Wang, J. Wang, X. You, Y. Wang, M. Chen, and X. Hou, "Spectral efficiency of distributed MIMO systems," *Selected Areas in Communications, IEEE Journal on*, vol. 31, no. 10, pp. 2112–2127, Oct. 2013.
[3] J. Wang, H. Zhu, and N. Gomes, "Distributed antenna systems for mobile communications in high speed trains," *Selected Areas in Communications, IEEE Journal on*, vol. 30, no. 4, pp. 675–683, May 2012.
[4] H. Zhu, "Performance comparison between distributed antenna and microcellular systems," *Selected Areas in Communications, IEEE Journal on*, vol. 29, no. 6, pp. 1151–1163, Jun. 2011.
[5] L. Zhang, H.-C. Yang, and M. Hasna, "Generalized area spectral efficiency: An effective performance metric for green wireless communications," *IEEE Trans. Communications*, vol. 62, no. 2, pp. 747–757, Feb. 2014.
[6] F. Richter, A. J. Fehske, and G. P. Fettweis, "Energy efficiency aspects of base station deployment strategies for cellular networks," in *Proc. IEEE Vehicular Technology Conference Fall (VTC 2009-Fall)*, 2009, pp. 1–5.
[7] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *Selected Areas in Communications, IEEE Journal on*, vol. 31, no. 2, pp. 160–171, Feb. 2013.
[8] H. Q. Ngo and E. G. Larsson, "Blind estimation of effective downlink channel gains in massive MIMO," *to appear on IEEE Trans. on Vehicular Technology*, 2015 Online Available: http://arxiv.org/abs/1503.09059.
[9] J. Li, D. Wang, P. Zhu, and X. You, "Spectral efficiency analysis of large-scale distributed antenna system in a composite correlated rayleigh fading channel," *Communications, IET*, vol. 9, no. 5, pp. 681–688, 2015.
[10] ——, "Spectral efficiency analysis of large-scale distributed antenna system in a composite correlated rayleigh fading channel," *Communications, IET*, vol. 9, no. 5, pp. 681–688, Jul. 2015.
[11] A. Ashikhmin and T. Marzetta, "Pilot contamination precoding in multi-cell large scale antenna systems," in *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on*, Jul. 2012, pp. 1137–1141.
[12] H. Zhu, "Performance comparison between distributed antenna and microcellular systems," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 6, pp. 1151–1163, Jun. 2011.
[13] D. Wang, C. Ji, X. Gao, S. Sun, and X. You, "Uplink sum-rate analysis of multi-cell multi-user massive MIMO system," in *Proc. IEEE International Conference on Communications (ICC'13)*, Budpest, Hungary, 2013.
[14] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *Selected Areas in Communications, IEEE Journal on*, vol. 31, no. 2, pp. 160–171, Feb. 2013.
[15] E. Bjornson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?" *Wireless Communications, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2015.
[16] E. Bjornson, E. Jorswieck, M. Debbah, and B. Ottersten, "Multiobjective signal processing optimization: The way to balance conflicting metrics in 5G systems," *Signal Processing Magazine, IEEE*, vol. 31, no. 6, pp. 14–23, Nov. 2014.
[17] S. Boyd and L. Vandenberghe, *Convex optimization.* Cambridge university press, 2004.
[18] E. Bjornson, E. Jorswieck, M. Debbah, and B. Ottersten, "Multiobjective signal processing optimization: The way to balance conflicting metrics in 5G systems," *Signal Processing Magazine, IEEE*, vol. 31, no. 6, pp. 14–23, Nov. 2014.

[19] D. Schneider, "Could supercomputing turn to signal processors (again)?" *Spectrum, IEEE*, pp. 13–14, Oct. 2012.

[20] E. Bjornson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?" *Wireless Communications, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2015.

[21] R.Couillet and M. Debbah, *Random matrix methods for wireless communications*, 1st ed.   New York: Cambridge University Press, 2011.