Large System Analysis of Power Normalization Techniques in Massive MIMO

Meysam Sadeghi, *Student Member, IEEE*, Luca Sanguinetti, *Senior Member, IEEE*, Romain Couillet, *Senior Member, IEEE*, and Chau Yuen, *Senior Member, IEEE*

***Abstract*—Linear precoding has been widely studied in the context of Massive MIMO together with the two common power normalization techniques, namely, matrix normalization (MN) and vector normalization (VN). However, the effect of both on the system performance has not been thoroughly studied. The aim of this paper is to address this problem using large system analysis. Considering a system model that accounts for channel estimation, pilot contamination, arbitrary pathloss, and per-user channel correlation, we compute tight approximations for the signal-to-interference-plus-noise ratio (SINR) and the rate of each user equipment (UE) in the system while employing maximum ratio transmission (MRT), zero forcing (ZF), and regularized ZF (RZF) precoding under both MN and VN techniques. Exploiting such results, we reveal the effect of power normalization on the performance of MRT and ZF, and determine how it affects noise, interference, pilot contamination, and signal powers of any arbitrary UE. We show that the power normalization can convey a notion of fairness or sum rate maximization for ZF. Numerical results are used to validate the accuracy of the asymptotic analysis and to show that in Massive MIMO, non-coherent interference and noise, rather than pilot contamination, are often the major limiting factors of the considered precoding schemes.**

***Index Terms*—Massive MIMO, linear precoding, power normalization techniques, large system analysis, pilot contamination.**

I. Introduction

Massive multiple-input multiple-output (MIMO) is a multiuser MIMO system that employs a large number of antennas at the base stations (BSs) to serve a relatively smaller number of user equipments (UEs) [1]–[4]. This large number of antennas enables each BS to focus the radiated energy into a specific location in space or to intercept the power of transmitted electromagnetic waves more efficiently. Therefore, Massive MIMO has higher spectral efficiency and energy efficiency compared to classical multiuser MIMO systems [3], [5]–[7]. Moreover, it has been recently shown that the capacity of Massive MIMO grows without bounds as the number of antennas increases, even under pilot contamination [8]. Due to the quasi-orthogonal nature of channels in Massive MIMO, linear precoding and detection schemes perform close-to-optimal [5], [6], [9]. Moreover, if the channel reciprocity is exploited, the overhead of the channel state information (CSI) acquisition is independent of the number of BS antennas [10]. These remarkable features candidate Massive MIMO as one of the most promising technologies for next generation of cellular networks [9], [11], [12].

Linear precoding has a central role in Massive MIMO and has been extensively studied in the past few years [5], [13]–[21]. The spectral efficiency and energy efficiency of maximum ratio transmission (MRT) and zero forcing (ZF) precoding in single-cell Massive MIMO systems are investigated in [13]. In [14], a multicell linear precoding is proposed to mitigate the effect of pilot contamination. The performance of MRT, ZF, and regularized ZF (RZF) precoding in single-cell large-scale MIMO systems is studied in [15], considering a per-user channel correlation model. A seminal treatment of MRT and RZF precoding schemes in multicell Massive MIMO systems is presented in [5]. The spectral efficiency of multicell Massive MIMO systems with downlink training and linear pilot contamination precoding is studied in [16]. In [17], closed-form approximations for the achievable downlink rates of MRT and ZF precoding schemes are presented for multi-cell Massive MIMO systems. A linear truncated polynomial expansion based precoding is proposed in [18], which reduces the complexity of RZF precoding. The effect of phase noise on the signal-to-interference-plus-noise (SINR) of MRT, ZF, and RZF precoding schemes is studied in [19]. A multicell MMSE precoder that improves the sum spectral efficiency of Massive MIMO systems is proposed in [20].

In order to utilize linear precoding, the power should be adjusted to meet the power constraint at the BS. This can be done either by optimized power allocation among the downlink data streams [20], [22]–[24], or simply by uniform power allocation among downlink data streams jointly with precoder power normalization [5], [13], [16]–[18]. Although the latter approach may provide a weaker performance compared to the former, it is the most used in the Massive MIMO literature [5], [10], [13], [16]–[18]. The reason for this is that power allocation presents the following major issues: (i) finding a global solution is a challenging task [23], [25], [26]; (ii) a certain level of coordination or cooperation among cells is required; and (iii) it should be performed very frequently, even for static users, as scheduling may change rapidly in practice.

M. Sadeghi (meysam@mymail.sutd.edu.sg) and C. Yuen (yuenchau@sutd.edu.sg) are with Singapore University of Technology and Design (SUTD), Singapore. L. Sanguinetti (luca.sanguinetti@unipi.it) is with the University of Pisa, Dipartimento di Ingegneria dell'Informazione, Italy and also with the Large Systems and Networks Group (LANEAS), CentraleSupélec, Université Paris-Saclay, 3 rue Joliot-Curie, 91192 Gif-sur-Yvette, France. R. Couillet (romain.couillet@centralesupelec.fr) is with the Telecommunication Department of CentraleSupélec, France.

This work was supported by A*Star SERC project number 142-02-00043. L. Sanguinetti and R. Couillet have been supported by the ERC Starting Grant 305123 MORE. L. Sanguinetti is also funded by the 5GIOTTO project from University of Pisa.

The two commonly used power normalization techniques in Massive MIMO are matrix normalization (MN) and vector normalization (VN) [25], [26]. In MN, the precoding matrix of each BS is adjusted by multiplying it with a scalar such that the power constraint at the BS is met [5], [10], [13], [16], [18]. On the other hand, with VN the precoding matrix is normalized such that equal amount of power is allocated to each UE while satisfying the power constraint [17], [25], [26]. Note that these two methods yield the same performance with optimal power allocation, but not with practical suboptimal power allocation [26], [27].

Although linear precoding has been largely studied in Massive MIMO, a detailed treatment of the impact of power normalization does not exist in the literature. The first attempt in this direction was carried out in [25] and extended in [26] wherein the authors study the impact of MN and VN on MRT and ZF precoding schemes. However, both works in [25], [26] do not grasp the essence of a practical Massive MIMO system since: 1) a single-cell network composed of three radio units is considered; 2) perfect CSI is assumed and thus CSI acquisition or pilot contamination are not accounted for; and 3) large-scale attenuation is neglected, though it has a fundamental impact on power normalization, as it will be detailed later.

The goal of this paper is to present a comprehensive picture on the effect of MN and VN on the performance of MRT, ZF, and RZF in Massive MIMO, in the simple and practical case of uniform power allocation over all downlink streams. Particularly, the following contributions are provided.

- We extend the analysis in [25], [26] to a multicell Massive MIMO system, which accounts for channel estimation, pilot contamination, an arbitrary pathloss model, and per-user channel correlation. Asymptotically tight approximations of the signal-to-interference-plus-noise ratio (SINR) and rate of each UE are provided and validated by numerical results for MRT, ZF, and RZF with VN and MN.
- Explicit asymptotic approximations for the SINR and rate of each UE are given for a Rayleigh fading channel model. These results are used: 1) to elaborate on how the two different power normalization techniques affect the signal, noise, and interference powers as well as the pilot contamination experienced by each UE in the system; 2) to prove that large-scale fading has a fundamental role on the performance provided by the two normalization techniques while both perform the same if neglected; 3) to show that ZF conveys a notion of sum rate maximization with VN and of fairness with MN.
- The asymptotic approximations of SINRs are used together with numerical results to study the main limiting factors of the investigated schemes in Massive MIMO. Particularly, we reveal that in Massive MIMO, non-coherent interference and noise, rather than pilot contamination, are often the major limiting factors for all schemes and also show how they are affected by power normalization.

The remainder of this paper is organized as follows. Section II introduces the network model, the channel estimation scheme, the precoding and power normalization methods, as well as the downlink achievable rates. The large system analysis is provided in Section III. The effect of power normalization techniques is elaborated in Section IV. The proposed asymptotic approximations are verified by means of numerical results in Section V. Some conclusions are drawn in Section VI.

Notations: The following notation is used throughout the paper. Scalars are denoted by lower case letters whereas boldface lower (upper) case letters are used for vectors (matrices). We denote by $\mathbf{I}_N$ the identity matrix of size N and represent the element on the ith row and kth column of $\mathbf{A}$ as $[\mathbf{A}]_{i,k}$. The symbol $\mathcal{CN}(.,.)$ denotes the circularly symmetric complex Gaussian distribution. The trace, transpose, conjugate transpose, real part, and expectation operators are denoted by $\mathsf{tr}(\cdot)$, $(\cdot)^T$, $(\cdot)^H$, $\mathsf{Re}(\cdot)$, and $\mathbb{E}[\cdot]$, respectively. The notations $\xrightarrow[N\to\infty]{a.s}$ and $\xrightarrow[N\to\infty]{}$ represent almost sure convergence.

II. Communication Scheme

Next, we introduce the system model, the channel estimation method, the precoding and power normalization techniques, and compute the downlink achievable rates.

A. System Model

We consider the downlink of a Massive MIMO system composed of L cells, where the set of all cells is denoted by $\mathcal{L}$. The BS of each cell has N antennas and serves K single-antenna UEs in the same time-frequency resource. The set of UEs belonging to cell l is denoted by $\mathcal{K}_l$. We assume transmissions over flat-fading channels. We employ a double index notation to refer to each UE as e.g., "user k in cell l". Under this convention, let $\mathbf{h}_{jlk} \in \mathbb{C}^N$ be the channel between BS j and UE k in cell l within a block and assume that

$$\mathbf{h}_{jlk} = \mathbf{\Theta}_{jlk}^{1/2}\mathbf{z}_{jlk} \tag{1}$$

where $\mathbf{z}_{jlk} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$ and $\mathbf{\Theta}_{jlk} \in \mathbb{C}^{N\times N}$ accounts for the corresponding channel correlation matrix. Note that (1) enables us to assign a unique correlation matrix between each user-BS pair and it includes many channel models in the literature as special cases [15].

B. Channel Estimation

We assume that BSs and UEs are perfectly synchronized and operate according to a time-division duplex (TDD) protocol. Then, the channels can be found by an uplink training phase and used in the downlink by exploiting channel reciprocity. Using orthogonal pilots in each cell while reusing them in all other cells, after correlating the received training signal with the pilot sequence of UE k, the observed channel of user k in cell j is

$$\mathbf{y}_{jk}^{\mathrm{tr}} = \mathbf{h}_{jjk} + \sum_{l=1, l\neq j}^{L} \mathbf{h}_{jlk} + \frac{1}{\sqrt{\rho_{\mathrm{tr}}}}\mathbf{n}_{jk} \tag{2}$$

where $\mathbf{n}_{jk} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_N)$ with σ^2 being the noise spectral density and ρ_{tr} is proportional to the training SNR. Applying

the MMSE estimation, the estimated channel can be computed as follows [5]

$$\hat{\mathbf{h}}_{jjk} = \mathbf{\Theta}_{jjk}\mathbf{Q}_{jk}\mathbf{y}^{\mathrm{tr}}_{jk} \tag{3}$$

where $\hat{\mathbf{h}}_{jjk} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Phi}_{jjk})$. Also, $\mathbf{Q}_{jk}$ and $\mathbf{\Phi}_{jlk}$ are given by

$$\mathbf{Q}_{jk} = \left(\sum_{l=1}^{L}\mathbf{\Theta}_{jlk} + \frac{\sigma^2}{\rho_{\mathrm{tr}}}\mathbf{I}_N\right)^{-1} \qquad \forall j,k \tag{4}$$

$$\mathbf{\Phi}_{jlk} = \mathbf{\Theta}_{jjk}\mathbf{Q}_{jk}\mathbf{\Theta}_{jlk} \qquad \forall j,l,k. \tag{5}$$

Note that due to the orthogonality principle of MMSE, the estimation error $\widetilde{\mathbf{h}}_{jjk} = \mathbf{h}_{jjk} - \hat{\mathbf{h}}_{jjk}$ is independent of $\hat{\mathbf{h}}_{jjk}$ and such that $\widetilde{\mathbf{h}}_{jjk} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Theta}_{jjk} - \mathbf{\Phi}_{jjk})$. For notational simplicity, we denote $\hat{\mathbf{H}}_{jj} = [\hat{\mathbf{h}}_{jj1}, \ldots, \hat{\mathbf{h}}_{jjK}]$ as the matrix collecting the estimated channels of cell j.

C. Precoding and Power Normalization Techniques

As mentioned earlier, we consider MRT, ZF, and RZF with VN and MN [25], [26]. Denoting by $\mathbf{G}_j = [\mathbf{g}_{j1}, \ldots, \mathbf{g}_{jK}] \in \mathbb{C}^{N\times K}$ the precoding matrix of BS j, where $\mathbf{g}_{jk} \in \mathbb{C}^N$ is the precoding vector of UE k in cell j, we have

$$\mathbf{G}_j = \mathbf{F}_j\mathbf{D}_j^{1/2} \tag{6}$$

where $\mathbf{F}_j = [\mathbf{f}_{j1}, \ldots, \mathbf{f}_{jK}] \in \mathbb{C}^{N\times K}$ determines the precoding scheme and $\mathbf{D}_j \in \mathbb{C}^{K\times K}$ characterizes the power allocation strategy. Therefore, $\mathbf{F}_j$ takes one of the following forms:

$$\mathbf{F}_j = \begin{cases} \hat{\mathbf{H}}_{jj} & \text{MRT} \quad (7) \\ \hat{\mathbf{H}}_{jj}\left(\hat{\mathbf{H}}_{jj}^H\hat{\mathbf{H}}_{jj}\right)^{-1} & \text{ZF} \quad (8) \\ \left(\hat{\mathbf{H}}_{jj}\hat{\mathbf{H}}_{jj}^H + \mathbf{Z}_j + N\alpha_j\mathbf{I}_N\right)^{-1}\hat{\mathbf{H}}_{jj} & \text{RZF} \quad (9) \end{cases}$$

where $\alpha_j > 0$ is the regularization parameter and $\mathbf{Z}_j \in \mathbb{C}^{N\times N}$ is an arbitrary Hermitian nonnegative definite matrix that can be used to leverage the system performance [5].

As mentioned in the introduction, finding the optimal values for the elements of $\mathbf{D}_j$ is challenging in practice [23]. This is why VN or MN are usually employed [26]. In this case, $\mathbf{D}_j$ is diagonal with entries chosen so as to satisfy the following average power constraint $\mathbb{E}[\mathtt{tr}\mathbf{G}_j\mathbf{G}_j^H] = K\ \forall j$. If VN is used, then the kth diagonal element of $\mathbf{D}_j$ is computed as

$$[\mathbf{D}_j]_{k,k} = d_{jk} = \frac{1}{\mathbb{E}[\mathbf{f}_{jk}^H\mathbf{f}_{jk}]}. \tag{10}$$

On the other hand, if MN is employed, then $\mathbf{D}_j = \eta_j\mathbf{I}_K$ with

$$\eta_j = \frac{K}{\mathbb{E}[\mathtt{tr}\mathbf{F}_j\mathbf{F}_j^H]}. \tag{11}$$

D. Downlink Achievable Rate

The received signal of user k in cell j can be written as

$$y_{jk} = \mathbf{h}_{jjk}^H\mathbf{g}_{jk}s_{jk} + \sum_{i=1,i\neq k}^{K}\mathbf{h}_{jjk}^H\mathbf{g}_{ji}s_{ji} + \sum_{l=1,l\neq j}^{L}\sum_{i=1}^{K}\mathbf{h}_{ljk}^H\mathbf{g}_{li}s_{li} + n_{jk} \tag{12}$$

with $s_{li} \in \mathbb{C}$ being the signal intended to UE i in cell l, assumed independent across (l,i) pairs, of zero mean and unit variance, and $n_{jk} \sim \mathcal{CN}(0, \sigma^2/\rho_{\mathrm{dl}})$ where ρ_{dl} is proportional to the downlink signal power.

As in [1], [5], [6], [14] (among many others), we assume that there are no downlink pilots such that the UEs do not have knowledge of the current channels but can only learn the average channel gain $\mathbb{E}\{\mathbf{h}_{jjk}^H\mathbf{g}_{jk}\}$ and the total interference power. Note this is the common approach in Massive MIMO due to the channel hardening [28]. Using the same technique as in [29], an ergodic achievable information rate for UE k in cell j is obtained as $r_{jk} = \log_2(1+\gamma_{jk})$ where γ_{jk} is given by

$$\gamma_{jk} = \frac{|\mathbb{E}[\mathbf{h}_{jjk}^H\mathbf{g}_{jk}]|^2}{\frac{\sigma^2}{\rho_{\mathrm{dl}}} + \sum_{l=1}^{L}\sum_{i=1}^{K}\mathbb{E}[|\mathbf{h}_{ljk}^H\mathbf{g}_{li}|^2] - |\mathbb{E}[\mathbf{h}_{jjk}^H\mathbf{g}_{jk}]|^2} \tag{13}$$

where the expectation is taken with respect to the channel realizations. The above result holds true for any precoding scheme and is obtained by treating the interference (from the same and other cells) and channel uncertainty as worst-case Gaussian noise. By using VN and MN, i.e. (10) and (11), the SINR takes respectively the form in (16) and (17), given on the top of next page.

As for all precoding schemes, γ_{jk}^{MN} and γ_{jk}^{VN} depend on the statistical distribution of $\{\mathbf{h}_{jlk}\}$ and $\{\hat{\mathbf{h}}_{jlk}\}$. This makes hard to compute both in closed-form. To overcome this issue, a large system analysis is provided next to find tight asymptotic approximations (hereafter called deterministic equivalents) for γ_{jk}^{MN} and γ_{jk}^{VN} and their associated achievable rates.

III. Large System Analysis

We consider a regime in which N and K grow large with a non-trivial ratio N/K, where $1 < \liminf N/K \leq \limsup N/K < \infty$. We will represent it as $N \to \infty$. Under this assumption, we derive deterministic equivalents (DEs) for γ_{jk} when any of MRT, ZF, and RZF precoding schemes is used with either MN or VN. The DE is represented by $\overline{\gamma}_{jk}$, and it is such that $\gamma_{jk} - \overline{\gamma}_{jk} \xrightarrow[N\to\infty]{a.s} 0$. By applying the continuous mapping theorem [30], the almost sure convergence of the results illustrated below implies that $r_{jk} - \overline{r}_{jk} \xrightarrow[N\to\infty]{} 0$ with $\overline{r}_{jk} = \log_2(1+\overline{\gamma}_{jk})$, where $\overline{\gamma}_{jk}$ denotes one of the asymptotic approximations computed below.

As limiting cases are considered, the following conditions are needed. Note that they have been widely used in the literature as in [5], [15], [31], [32].

$$\text{A1}: \limsup||\mathbf{\Theta}_{jlk}^{1/2}|| < \infty \quad \text{and} \quad \liminf\frac{1}{N}\mathrm{tr}\left(\mathbf{\Theta}_{jlk}\right) > 0$$

$$\text{A2}: \exists\epsilon > 0 : \lambda_{\min}(\frac{1}{N}\mathbf{H}_{ll}^H\mathbf{H}_{ll}) > \epsilon$$

$$\text{A3}: \limsup_N||\frac{1}{N}\boldsymbol{Z}_l|| < \infty$$

$$\text{A4}: \mathrm{rank}(\hat{\mathbf{H}}_{ll}) \geq K.$$

A. Large System Results for Vector Normalization

In this subsection, we derive DEs for $\gamma_{jk}^{\mathrm{VN}}$, when any of MRT, ZF, and RZF precoding schemes are used.

$$\gamma_{jk}^{\mathrm{VN}} = \frac{d_{jk}\,|\mathbb{E}[\mathbf{h}_{jjk}^H \mathbf{f}_{jk}]|^2}{\frac{\sigma^2}{\rho_{\mathrm{dl}}} + d_{jk}\,\mathrm{var}(\mathbf{h}_{jjk}^H \mathbf{f}_{jk}) + \sum\limits_{l=1}^{L}\sum\limits_{i=1,i\neq k}^{K} d_{li}\,\mathbb{E}[|\mathbf{h}_{ljk}^H \mathbf{f}_{li}|^2] + \sum_{l=1,l\neq j}^{L} d_{lk}\,\mathbb{E}[|\mathbf{h}_{ljk}^H \mathbf{f}_{lk}|^2]} \tag{16}$$

$$\gamma_{jk}^{\mathrm{MN}} = \frac{\eta_j\,|\mathbb{E}[\mathbf{h}_{jjk}^H \mathbf{f}_{jk}]|^2}{\frac{\sigma^2}{\rho_{\mathrm{dl}}} + \eta_j\,\mathrm{var}(\mathbf{h}_{jjk}^H \mathbf{f}_{jk}) + \sum\limits_{l=1}^{L}\sum\limits_{i=1,i\neq k}^{K} \eta_l\,\mathbb{E}[|\mathbf{h}_{ljk}^H \mathbf{f}_{li}|^2] + \sum_{l=1,l\neq j}^{L} \eta_l\,\mathbb{E}[|\mathbf{h}_{ljk}^H \mathbf{f}_{lk}|^2]}. \tag{17}$$

Theorem 1. *Let A1 hold true. If MRT with VN is used, then* $\gamma_{jk}^{\mathrm{VN}} - \overline{\gamma}_{jk}^{(\mathrm{MRT-VN})} \xrightarrow[N\to\infty]{} 0$ *almost surely with*

$$\overline{\gamma}_{jk}^{(\mathrm{MRT-VN})} = \frac{d_{jk}^{\dagger}\left(\frac{1}{N}\mathtt{tr}\mathbf{\Phi}_{jjk}\right)^2}{\frac{\sigma^2}{N\rho_{\mathrm{dl}}} + \frac{1}{N}\sum\limits_{l=1}^{L}\sum\limits_{i=1}^{K} d_{li}^{\dagger} z_{li,jk} + \sum\limits_{l=1,l\neq j}^{L} d_{lk}^{\dagger}|\frac{1}{N}\mathtt{tr}\mathbf{\Phi}_{ljk}|^2} \tag{18}$$

where

$$d_{li}^{\dagger} = \left(\frac{1}{N}\mathtt{tr}\mathbf{\Phi}_{lli}\right)^{-1} \tag{19}$$

$$z_{li,jk} = \frac{1}{N}\mathtt{tr}\mathbf{\Theta}_{ljk}\mathbf{\Phi}_{lli}. \tag{20}$$

Proof. The proof is provided in Appendix A. □

Theorem 2. *Let A1 and A3 hold true. If RZF with VN is used, then* $\gamma_{jk}^{\mathrm{VN}} - \overline{\gamma}_{jk}^{(\mathrm{RZF-VN})} \xrightarrow[N\to\infty]{} 0$ *almost surely while*

$$\overline{\gamma}_{jk}^{(\mathrm{RZF-VN})} = \frac{d_{jk}^{\circ}\frac{u_{jk}^2}{(1+u_{jk})^2}}{\frac{\sigma^2}{N\rho_{dl}} + \frac{1}{N}\sum\limits_{l=1}^{L}\sum\limits_{i=1}^{K} d_{li}^{\circ}\frac{\epsilon_{li,jk}}{(1+u_{li})^2} + \sum\limits_{l=1,l\neq j}^{L} d_{lk}^{\circ}\frac{|u_{ljk}|^2}{(1+u_{lk})^2}} \tag{21}$$

with

$$d_{li}^{\circ} = \frac{(1+u_{li})^2}{\frac{1}{N}\mathtt{tr}\mathbf{\Phi}_{lli}\mathbf{T}'_{l,\mathbf{I}_N}} \tag{22}$$

$$u_{lk} = \frac{1}{N}\mathtt{tr}\mathbf{\Phi}_{llk}\mathbf{T}_l \tag{23}$$

$$u_{ljk} = \frac{1}{N}\mathtt{tr}\mathbf{\Phi}_{ljk}\mathbf{T}_l \tag{24}$$

$$\epsilon_{li,jk} = \frac{1}{N}\mathtt{tr}\mathbf{\Theta}_{ljk}\mathbf{T}'_{l,\mathbf{\Phi}_{lli}} + \frac{|u_{ljk}|^2}{(1+u_{lk})^2}\times\frac{1}{N}\mathtt{tr}\mathbf{\Phi}_{llk}\mathbf{T}'_{l,\mathbf{\Phi}_{lli}} - \frac{2}{1+u_{lk}}\,\mathrm{Re}\left(\frac{1}{N}\mathtt{tr}\mathbf{\Phi}_{ljk}\mathbf{T}'_{l,\mathbf{\Phi}_{lli}}\times u_{ljk}^*\right) \tag{25}$$

and $\mathbf{S}_l = \frac{\mathbf{Z}_l}{N}$. *Also,* $\mathbf{T}_l$, $\mathbf{T}'_{l,\mathbf{I}_N}$, *and* $\mathbf{T}'_{l,\mathbf{\Phi}_{lli}}$ *are given in Theorems 7 and 8 in Appendix E.*

Proof. The proof is provided in Appendix B. □

Theorem 3. *Let A1, A2 and A4 hold true. If ZF with VN is employed, then* $\gamma_{jk}^{\mathrm{VN}} - \overline{\gamma}_{jk}^{(\mathrm{ZF-VN})} \xrightarrow[N\to\infty]{} 0$ *almost surely with*

$$\overline{\gamma}_{jk}^{(\mathrm{ZF-VN})} = \frac{\underline{u_{jk}}}{\frac{\sigma^2}{N\rho_{\mathrm{dl}}} + \frac{1}{N}\sum\limits_{l=1}^{L}\sum\limits_{i=1}^{K}\frac{\epsilon_{li,jk}}{\underline{u_{li}}} + \sum\limits_{l=1,l\neq j}^{L}\frac{u_{ljk}^2}{\underline{u_{lk}}}} \tag{26}$$

where

$$\underline{u_{li}} = \frac{1}{N}\mathtt{tr}\left(\mathbf{\Phi}_{lli}\underline{\mathbf{T}_l}\right) \tag{27}$$

$$\underline{\mathbf{T}_l} = \left(\frac{1}{N}\sum_{i=1}^{K}\frac{\mathbf{\Phi}_{lli}}{\underline{u_{li}}} + \mathbf{I}_N\right)^{-1} \tag{28}$$

$$\underline{u_{ljk}} = \frac{1}{N}\mathtt{tr}\left(\mathbf{\Phi}_{ljk}\underline{\mathbf{T}_l}\right) \tag{29}$$

$$\underline{\epsilon_{li,jk}} = \frac{1}{N}\mathtt{tr}\mathbf{\Theta}_{ljk}\underline{\mathbf{T}'_{l,\mathbf{\Phi}_{lli}}} + \frac{|u_{ljk}|^2}{\underline{u_{lk}}^2}\frac{1}{N}\mathtt{tr}\mathbf{\Phi}_{llk}\underline{\mathbf{T}'_{l,\mathbf{\Phi}_{lli}}} - \frac{2}{\underline{u_{lk}}}\,\mathrm{Re}\left(\underline{u_{ljk}^*}\frac{1}{N}\mathtt{tr}\mathbf{\Phi}_{ljk}\underline{\mathbf{T}'_{l,\mathbf{\Phi}_{lli}}}\right) \tag{30}$$

$$\underline{\mathbf{T}'_{l,\mathbf{\Phi}_{llk}}} = \underline{\mathbf{T}_l}\left(\frac{1}{N}\sum_{i=1}^{K}\frac{\underline{u'_{li,\mathbf{\Phi}_{llk}}}\mathbf{\Phi}_{lli}}{\underline{u_{li}}^2} + \mathbf{\Phi}_{llk}\right)\underline{\mathbf{T}_l} \tag{31}$$

where $\underline{\mathbf{u}'_{l,\mathbf{\Phi}_{llk}}} = [\underline{u'_{l1,\mathbf{\Phi}_{llk}}},\ldots,\underline{u'_{lK,\mathbf{\Phi}_{llk}}}]^T \in \mathbb{C}^K$ *is computed as*

$$\underline{\mathbf{u}'_{l,\mathbf{\Phi}_{llk}}} = (\mathbf{I}_K - \underline{\mathbf{J}_l})^{-1}\underline{\mathbf{v}_{l,\mathbf{\Phi}_{llk}}} \tag{32}$$

with the entries of $\underline{\mathbf{J}_l} \in \mathbb{C}^{K\times K}$ *and* $\underline{\mathbf{v}_{l,\mathbf{\Phi}_{llk}}} \in \mathbb{C}^K$ *are given by:*

$$[\underline{\mathbf{J}_l}]_{n,i} = \frac{1}{N^2}\frac{\mathtt{tr}\left(\mathbf{\Phi}_{lln}\underline{\mathbf{T}_l}\mathbf{\Phi}_{lli}\underline{\mathbf{T}_l}\right)}{\underline{u_{li}}^2} \tag{33}$$

$$[\underline{\mathbf{v}_{l,\mathbf{\Phi}_{llk}}}]_i = \frac{1}{N}\mathtt{tr}\left(\mathbf{\Phi}_{lli}\underline{\mathbf{T}_l}\mathbf{\Phi}_{llk}\underline{\mathbf{T}_l}\right). \tag{34}$$

Proof. The proof is provided in the Appendix C. □

Notice that the computation of the DEs with ZF precoding (either VN or MN) for the considered multicell Massive MIMO system is more involved than with MRT or RZF precoding schemes. This is mainly due to the fact that it is not straightforward to start with ZF precoder and then compute the DEs by applying common techniques, e.g., matrix inversion lemma. Therefore, in proving Theorem 3 (and also Theorem 6) we start with the DE of RZF and then use a bounding and limiting technique to compute the DE for ZF.

B. Large System Results for Matrix Normalization

Next, the DEs of $\gamma_{jk}^{\mathrm{MN}}$ are given for MRT, ZF, and RZF. Note that the DEs of $\gamma_{jk}^{\mathrm{MN}}$ for MRT and RZF are obtained from [5].

Theorem 4. *[5, Theorem 4] Let A1 hold true. If MRT with MN is used, then* $\gamma_{jk}^{\mathrm{MN}} - \overline{\gamma}_{jk}^{(\mathrm{MRT-MN})} \xrightarrow[N\to\infty]{} 0$ *almost surely with*

$$\overline{\gamma}_{jk}^{(\mathrm{MRT-MN})} \overline{=} \frac{\lambda_j \left(\frac{1}{N}\mathtt{tr}\mathbf{\Phi}_{jjk}\right)^2}{\frac{\sigma^2}{N\rho_{\mathrm{dl}}} + \frac{1}{N}\sum\limits_{l=1}^{L}\sum\limits_{i=1}^{K}\lambda_l z_{li,jk} + \sum\limits_{l=1,l\neq j}^{L}\lambda_l |\frac{1}{N}\mathtt{tr}\mathbf{\Phi}_{ljk}|^2} \tag{35}$$

where $z_{li,jk}$ *is given in* (20) *and*

$$\lambda_j = \left(\frac{1}{K}\sum_{k=1}^{K}\frac{1}{N}\mathtt{tr}\mathbf{\Phi}_{jjk}\right)^{-1}. \tag{36}$$

Theorem 5. *[5, Theorem 6] Let A1 and A3 hold true. If RZF with MN is used, then* $\gamma_{jk}^{\mathrm{MN}} - \overline{\gamma}_{jk}^{(\mathrm{RZF-MN})} \xrightarrow[N\to\infty]{} 0$ *almost surely with*

$$\overline{\gamma}_{jk}^{(\mathrm{RZF-MN})} \overline{=} \frac{\lambda_j \frac{u_{jk}^2}{(1+u_{jk})^2}}{\frac{\sigma^2}{N\rho_{dl}} + \frac{1}{N}\sum\limits_{l=1}^{L}\sum\limits_{i=1}^{K}\lambda_l \frac{\epsilon_{li,jk}}{(1+u_{li})^2} + \sum\limits_{l=1,l\neq j}^{L}\lambda_l \frac{|u_{ljk}|^2}{(1+u_{lk})^2}} \tag{37}$$

with

$$\lambda_l = \frac{K}{N}\left(\frac{1}{N}\mathtt{tr}\mathbf{T}_l - \frac{1}{N}\mathtt{tr}(\frac{\mathbf{Z}_l}{N} + \alpha_l \mathbf{I}_N)\mathbf{T}'_{l,\mathbf{I}_N}\right)^{-1} \tag{38}$$

where $\mathbf{S}_l = \frac{\mathbf{Z}_l}{N}$ *and* $\mathbf{T}_l$ *and* $\mathbf{T}'_{l,\mathbf{I}_N}$ *are given by Theorem 7 and Theorem 8. Also* u_{li}, u_{ljk}, *and* $\epsilon_{li,jk}$ *are defined in Theorem 2.*

Theorem 6. *Let A1, A2 and A4 hold true. If ZF with MN is used, then* $\gamma_{jk}^{\mathrm{MN}} - \overline{\gamma}_{jk}^{(\mathrm{ZF-MN})} \xrightarrow[N\to\infty]{} 0$ *almost surely with*

$$\overline{\gamma}_{jk}^{(\mathrm{ZF-MN})} = \frac{\lambda_j}{\frac{\sigma^2}{N\rho_{\mathrm{dl}}} + \frac{1}{N}\sum\limits_{l=1}^{L}\sum\limits_{i=1}^{K}\lambda_l \frac{\epsilon_{li,jk}}{\underline{u_{li}}^2} + \sum\limits_{l=1,l\neq j}^{L}\lambda_l \frac{{u_{ljk}}^2}{\underline{u_{lk}}^2}} \tag{39}$$

with $\lambda_j = \left(\frac{1}{K}\sum\limits_{i=1}^{K}\frac{1}{\underline{u_{ji}}}\right)^{-1}$ *where* $\underline{u_{li}}$, $\underline{u_{ljk}}$, *and* $\underline{\epsilon_{li,jk}}$ *are given in Theorem 3.*

Proof sketch. The proof follows the same procedure as the proof of Theorem 3 presented in Appendix C. Start with the triangle equality and bound $|\gamma_{lk}^{(\mathrm{ZF-MN})} - \overline{\gamma}_{lk}^{(\mathrm{ZF-MN})}|$. Then find the DE of $\overline{\gamma}_{jk}^{(\mathrm{ZF-MN})}$ by letting $\alpha \to 0$ in $\overline{\gamma}_{jk}^{(\mathrm{RZF-MN})}$. □

The asymptotic expressions provided in Theorems 1, 2, 3, and 6 will be shown to be very tight, even for systems with finite dimensions, via numerical results in Section V. This allows us to use them for evaluating the performance of practical multicell Massive MIMO systems without the need for time-consuming Monte Carlo simulations. Moreover, they lay the foundation for further analysis of different configurations of Massive MIMO systems (e.g., distributed massive MIMO systems [33], [34]). Next, they are used to get further insights into the system under investigation for a simplified channel model.

IV. Effect of Power Normalization Technique

In this section, we use the asymptotic results provided above to gain novel insights into the interplay between the different system parameters and the power normalization techniques in Massive MIMO. To this end, we consider a special case of the general channel model of (1) in which $\mathbf{\Theta}_{jlk} = d_{jlk}\mathbf{I}_N$ such that

$$\mathbf{h}_{jlk} = \sqrt{d_{jlk}}\mathbf{z}_{jlk} \tag{40}$$

where $\mathbf{z}_{jlk} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$ and d_{jlk} accounts for an arbitrary large-scale fading coefficient including pathloss and shadowing. Note this is a quite popular model in Massive MIMO that allows us to capture the essence of the technology [1], [6]. Under the above circumstances, we have that:

Corollary 1. *Let* $\lambda_j = \bar{u}\big(\frac{1}{K}\sum_{i=1}^{K}\frac{\alpha_{ji}}{d_{jji}^2}\big)^{-1}$ *and* $\bar{u} = 1 - \frac{K}{N}$. *If the channel is modelled as in* (40)*, then*

$$\overline{\gamma}_{jk}^{(\mathrm{ZF-VN})} = \frac{\frac{d_{jjk}^2}{\alpha_{jk}}\bar{u}}{\nu_{jk} + \underbrace{\sum\limits_{l=1,l\neq j}^{L}\bar{u}\frac{d_{ljk}^2}{\alpha_{lk}}}_{\text{Pilot Contamination}}} \tag{41}$$

$$\overline{\gamma}_{jk}^{(\mathrm{ZF-MN})} = \frac{\lambda_j}{\nu_{jk} + \underbrace{\sum\limits_{l=1,l\neq j}^{L}\lambda_l\frac{d_{ljk}^2}{d_{llk}^2}}_{\text{Pilot Contamination}}} \tag{42}$$

where

$$\nu_{jk} = \underbrace{\frac{\sigma^2}{N\rho_{\mathrm{dl}}}}_{\text{Noise}} + \underbrace{\frac{K}{N}\sum_{l=1}^{L}d_{ljk}\left(1 - \frac{d_{ljk}}{\alpha_{lk}}\right)}_{\text{Interference}}. \tag{43}$$

with $\alpha_{lk} = \sum_{n=1}^{L}d_{lnk} + \frac{\sigma^2}{\rho_{\mathrm{tr}}}$.

Proof. See Appendix D. □

Corollary 2. *Let* $\theta_l = (\frac{1}{K}\sum_{i=1}^{K}\frac{d_{lli}^2}{\alpha_{li}})^{-1}$. *If the channel is modelled as in* (40) *and MRT is used, then*

$$\overline{\gamma}_{jk}^{(\mathrm{MRT-VN})} = \frac{\frac{d_{jjk}^2}{\alpha_{jk}}}{\vartheta_{jk} + \underbrace{\sum\limits_{l=1,l\neq j}^{L}\frac{d_{ljk}^2}{\alpha_{lk}}}_{\text{Pilot Contamination}}} \tag{44}$$

$$\overline{\gamma}_{jk}^{(\mathrm{MRT-MN})} = \frac{\theta_j(\frac{d_{jjk}^2}{\alpha_{jk}})^2}{\vartheta_{jk} + \underbrace{\sum\limits_{l=1,l\neq j}^{L}\theta_l(\frac{d_{llk}d_{ljk}}{\alpha_{lk}})^2}_{\text{Pilot Contamination}}} \tag{45}$$

with

$$\vartheta_{jk} = \underbrace{\frac{\sigma^2}{N\rho_{\mathrm{dl}}}}_{\text{Noise}} + \underbrace{\frac{K}{N}\sum_{l=1}^{L}d_{ljk}}_{\text{Interference}}. \tag{46}$$

Proof. The proof follows a similar procedure as that of Corollary 1. □

The results of Corollaries 1 and 2 are instrumental in obtaining the following insights into MRT and ZF with either MN or VN.

Remark 1 (Effect of VN and MN)**.** The terms ν_{jk} and ϑ_{jk} in (43) and (46) are the same for both VN and MN. This means that both normalization techniques have exactly the same effect on the resulting noise and interference terms experienced by each UE in the system. On the other hand, they have different effects on the signal and pilot contamination powers. The expressions (41)-(46) explicitly state the relation between the SINR components (the signal power, the interference, the noise, and the pilot contamination), the propagation environment, and the two normalization techniques for ZF and MRT precoding schemes.

Remark 2 (On the mutual effect of UEs)**.** If VN is employed, then the signal power and the pilot contamination of UE k in cell j, for both MRT and ZF precoding, depends only on the coefficients d_{lnk} $\forall l, n \in \mathcal{L}$ through α_{lk}. This means that they are both affected only by the large-scale gains of the UEs in the network using the same pilot. On the other hand, under MN both terms depend on the coefficients λ_l $\forall l \in \mathcal{L}$ (or θ_l for MRT) and thus are influenced by all the UEs in the network - even though they make use of different pilot sequences.

Remark 3 (Large-Scale Fading and Power Normalization)**.** Assume that the large-scale fading is neglected such that it is the same for every UE in the network, i.e., $d_{ljk} = d$ $\forall l, j, k$. Then for ZF (41) and (42), and for MRT (44) and (45) become equal. Therefore, we can conclude that the large-scale fading has a fundamental effect on VN and MN and cannot be ignored.

Consider now, for further simplicity, a single-cell setup, i.e., $L = 1$. Dropping the cell index, α_{lk} reduces to $\alpha_k = d_k + \sigma^2/\rho_{\mathrm{tr}}$. Also assume that the UEs operate in the high training SNR regime such that $\rho_{\mathrm{tr}} \gg 1$. Under these conditions, we have that:

Lemma 1. *If $L = 1$ and $\rho_{\mathrm{tr}} \gg 1$, then for ZF precoding, VN outperforms MN in terms of sum rate and the sum rate gap $\Delta r \geq 0$ is given by*

$$\Delta r = \sum_{k=1}^{K} \log\left(1 + \frac{1}{\frac{\sigma^2}{N\rho_{\mathrm{dl}}\bar{u}}\frac{1}{d_k}}\right) - K\log\left(1 + \frac{1}{\frac{\sigma^2}{N\rho_{\mathrm{dl}}\bar{u}}\frac{1}{K}\sum_{i=1}^{K}\frac{1}{d_i}}\right). \quad (47)$$

Proof. From Corollary 1, setting $L = 1$ and assuming $\rho_{\mathrm{tr}} \gg 1$ we obtain that $\alpha_k \simeq d_k$ and $\nu_k \simeq \frac{\sigma^2}{N\rho_{\mathrm{dl}}}$. Then, the result follows by applying the Jensen's inequality (by the convexity of $\log(1 + 1/x)$). □

Notice that Lemma 1 extends the results of [25] and [26] to a system that accounts for CSI acquisition and arbitrary pathloss and UEs' distribution. Also, observe that (41) and (42) simplify as:

$$\overline{\gamma}_{jk}^{(\mathrm{ZF-VN})} = \frac{(N-K)\rho_{\mathrm{dl}}}{\sigma^2} d_k \quad (48)$$

$$\overline{\gamma}_{jk}^{(\mathrm{ZF-MN})} = \frac{(N-K)\rho_{\mathrm{dl}}}{\sigma^2}\Big(\frac{1}{K}\sum_{i=1}^{K}\frac{1}{d_i}\Big)^{-1} \quad (49)$$

from which it follows that VN provides higher SINR to the UEs that are closer to the BS and lower SINR for those that are far away from the BS (which resembles opportunistic resource allocation). On the other hand, MN provides a uniform quality of experience to all UEs. This proves evidence of the fact that ZF with VN resembles a sum rate maximizer. On the other hand, it provides a notion of fairness under MN. Notice that fairness means similar SINR (quality of experience) and it should not be confused with equal power allocation. The above results and observations will be validated below in Section V. Also, the DEs provided in Corollaries 1 and 2 will be used to investigate the main limiting factors of Massive MIMO.

V. Numerical results

Monte-Carlo (MC) simulations are now used to validate the asymptotic analysis for different values of N and K. We consider a multicell network composed of $L = 7$ cells, one in the center and six around. Each cell radius is 1000 meters. A 20 MHz channel is considered and the thermal noise power is assumed to be -174 dBm/Hz. The UEs are randomly and uniformly distributed within each cell excluding a circle of radius 100 meters. The channel is modeled as in [35]. In particular, we assume that the matrices $\mathbf{\Theta}_{ljk}^{1/2}$ are given by

$$\mathbf{\Theta}_{ljk}^{1/2} = \sqrt{d_{ljk}}\mathbf{A} \quad (50)$$

where $\mathbf{A} = [\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_N)] \in \mathbb{C}^N$ with $\mathbf{a}(\theta_i)$ given by

$$\mathbf{a}(\theta_i) = \frac{1}{\sqrt{N}}[1, e^{-i2\pi\omega\sin(\theta_i)}, \ldots, e^{-i2\pi\omega(N-1)\sin(\theta_i)}]^T \quad (51)$$

where $\omega = 0.3$ is the antenna spacing and $\theta_i = -\pi/2 + (i-1)\pi/N$. Also, d_{ljk} is the large-scale attenuation, which is modeled as $d_{ljk} = x_{ljk}^{-\beta}$ where x_{ljk} denotes the distance of UE k in cell j from BS l and $\beta = 3.7$ is the path-loss exponent. We let $\rho_{\mathrm{tr}} = 6$ dB and $\rho_{\mathrm{dl}} = 10$ dB, which corresponds to a practical setting [5]. The results are obtained for 100 different channel and UE distributions realizations.

Figs. 1 and 2 are presented to confirm the accuracy of the proposed DEs in Theorems 1, 2, 3, and 6. They illustrate the ergodic achievable sum rate of the center cell versus N for $K = 8$ and 16, respectively. The solid lines present the sum rate achieved based on the proposed deterministic equivalents, and the markers are achieved through Monte Carlo simulation. As it is depicted, the DEs match perfectly with numerical results. Notics that Figs. 1 and 2 (also Table 1), extend the results in [25] and [26] since CSI acquisition, pilot contamination, arbitrary pathloss and UEs' distribution are taken into account.

In Lemma 1, it is shown that ZF under VN conveys a notion of sum rate maximization, while ZF with MN resembles a fairness provisioning precoder. Now, we use Table I to

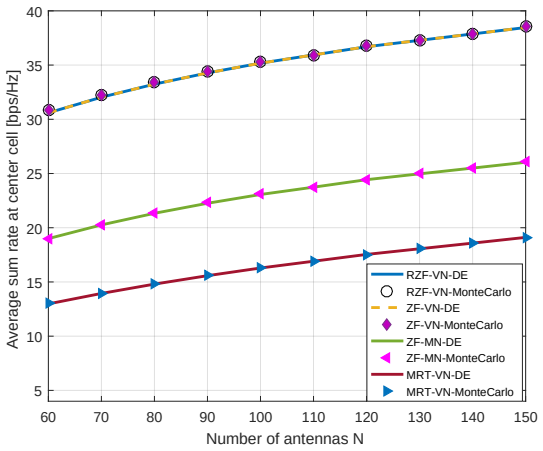


Fig. 1: Ergodic achievable sum rate of center cell for MRT, ZF, and RZF with VN and ZF with MN for K=8.

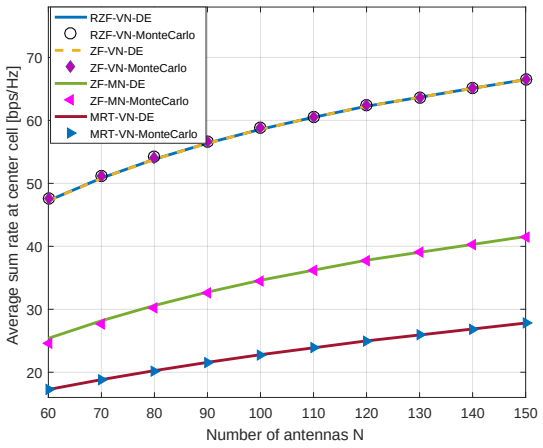


Fig. 2: Ergodic achievable sum rate of center cell for MRT, ZF, and RZF with VN and ZF with MN for K=16.

validate this observation and also to verify the accuracy of the computed DEs for the simplified channel model in (40). Table I presents the exact values of SINRs of all the UEs in the center cell for ZF with MN and VN. It also reports the estimated values of SINRs by the DEs proposed in Section IV. This helps the reader to have a better vision on the accuracy of the achieved DEs.

The first column of Table I reports the number of antennas, the second one is the UE index. The third and fourth columns are the SINR of each UE under MN, achieved by the proposed DE of (42) and by time-consuming Monte Carlo simulations, respectively. The sixth and seventh columns are the SINR of each UE under VN, achieved by the proposed DE of (41) and by intensive Monte Carlo simulations, respectively. The fifth and eighth columns present the percentage of the error while estimating a specific UE SINR via proposed DEs. As predicted by Lemma 1, ZF with MN provides a more uniform experience for all UEs, while ZF with VN provides very high SINR for specific UEs (UEs 2, 4, and 6) and much lower SINR for other UEs. More precisely, from Table I one can see that the variance of the SINR of the UEs with MN is equal to 0.8 (5.79) for $N = 40$ ($N = 80$), while for VN it is equal to 2627 (11550 for $N = 80$). Also note that the percentage of

TABLE I: SINR of each UE under ZF with VN and MN.

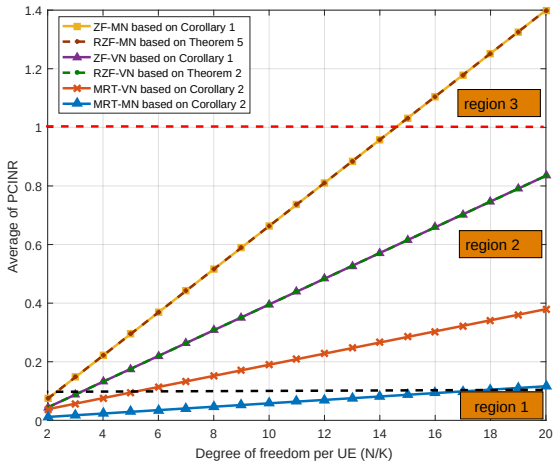
No. Ant.	UE	MN DE	MN MC	Er. %	VN DE	VN MC	Er. %
$N = 40$	1	2.25	2.19	2.6	1.85	1.84	0.5
	2	4.89	4.84	1.0	147	149	1.3
	3	3.34	3.29	1.5	3.61	3.53	2.2
	4	5.14	5.12	0.3	37.5	37.6	0.2
	5	4.09	4.02	1.7	1.97	1.96	0.5
	6	4.26	4.41	3.5	85	87	2.3
	7	3.30	3.33	0.9	2.14	2.2	2.8
	8	3.52	3.50	0.5	2.52	2.49	1.2
$N = 80$	1	3.20	3.15	1.5	3.1	2.98	3.8
	2	10.57	10.40	1.6	316	310	1.9
	3	5.94	6.03	1.5	6.40	6.36	0.6
	4	9.33	9.47	1.5	72.2	72.2	0
	5	8.41	8.62	2.5	3.87	3.90	0.7
	6	9.05	9.12	0.7	182	185	1.6
	7	5.02	4.90	2	3.50	3.43	2
	8	5.89	5.74	2.5	4.28	4.21	1.6

error is always less than 4%, which proves the high accuracy of the DEs. Therefore one can simply use the DEs to achieve insight on the performance of the network, instead of using time-consuming Monte Carlo simulations. Moreover, the DEs do not contain any randomness and are purely based on large-scale statistics of the system. Hence, they can be used for network optimization purposes.

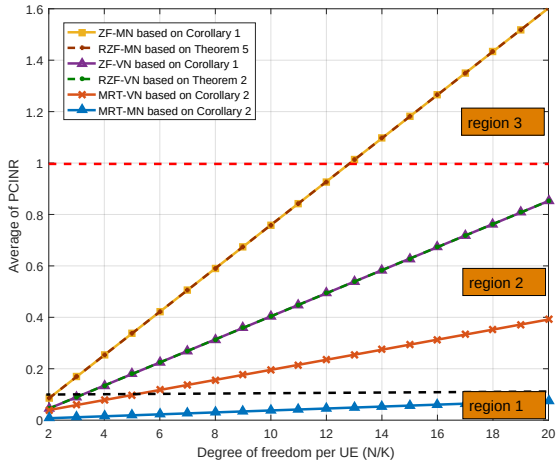
Now let us use our derived DEs, given in Corollaries 1 and 2 and Theorems 2 and 5, to investigate a common notion in the literature of massive MIMO systems. It is known that in massive MIMO systems, in the limit of an infinite number of antennas, the detrimental effects such as noise and interference vanish and pilot contamination become the main bottleneck of the system performance. This also can be seen from our results, e.g., corollaries 1 and 2, by letting N grow large while K is fixed. This has motivated a huge amount of researches on pilot decontamination. However, as has been shown in [28], it is usually desirable for massive MIMO systems to work in a regime where $\frac{N}{K} \leq 10$. Therefore, it is interesting to know what is the major drawback of massive MIMO systems under practical regimes, e.g., $\frac{N}{K} \leq 10$? Is it pilot contamination (coherent interference)? Or is the noise and interference (more exactly, non-coherent interference)? How is the answer to this questions related to the choice of power normalization and precoding scheme?

To answer these questions, we employ the so-called pilot contamination-to-interference-plus-noise ratio (PCINR) metric, which is computed by using the DEs provided in Corollaries 1 and 2. Fig. 3 plots the PCINR as a function of N/K, i.e., the number of degrees of freedom per-user in the system. Although, the optimal operating regime for maximal spectral efficiency is for $N/K < 10$ [28], we consider N/K up to 20 to cover a wider range of definition for massive MIMO. Moreover, as the interference increases by having more UEs in the system, we consider three different scenarios with $K = 5$, $K = 10$, and $K = 15$.

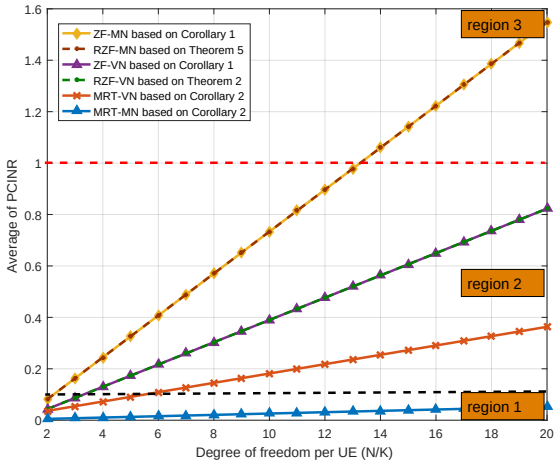
Fig. 3 is divided into 3 regions based on the significance of the PCINR term such that, as we move away from region 1 towards region 3, the importance of pilot contamination increases while that of the interference plus noise reduces. Region 1 is where the noise and interference are the dominant



(a) $K = 5$.



(b) $K = 10$.



(c) $K = 15$.

Fig. 3: PCINR versus the degree of freedom per-user, for different values of K.

limiting factors and pilot contamination has a negligible effect, less than 10% of the noise and interference. As it is depicted, MRT with MN operates within this regime, therefore pilot contamination is never a bottleneck for this scheme, which is mainly limited by noise and interference. Note that by adding more UEs in the system (bigger values of K), the PCINR reduces and pilot contamination becomes even less important. Hence, when MRT with MN is studied in Massive MIMO the effect of pilot contamination can be safely neglected.

Region 2 represents the regime where the noise and interference are the main limiting factors of the system performance, while pilot contamination is not negligible any more. It is interesting to observe that for the other schemes (other than MRT-MN), Massive MIMO often operates within this region. This shows that, although pilot contamination is a major challenge in Massive MIMO, the interference and noise have still the leading role in limiting the system performance.

Finally, region 3 presents the superiority of pilot contamination effect. If $K = 10$, then Fig. 3b shows the superiority of interference and noise over pilot contamination for ZF-MN and RZF-MN (ZF-VN and RZF-VN) up to $N = 130$ ($N = 233$) antennas at the BS. With MRT-VN, the system requires more than $N = 510$ to experience the superiority of pilot contamination over interference and noise. This increases to $N = 2650$ with MRT-MN. From Fig. 3, we see also that, for a given value of N/K, the value of PCINR for the considered schemes can be ordered as: ZF-MN = RZF-MN $\geq$ ZF-VN = RZF-VN $\geq$ MRT-VN $\geq$ MRT-MN. Based on the above discussion, it is clear that the choice of precoding scheme and normalization technique change the importance of pilot contamination, interference, and noise dramatically and it should be considered carefully when designing Massive MIMO systems.

VI. Conclusions

Linear precoding schemes, such as MRT and ZF, have a fundamental role in Massive MIMO. Although these precoding schemes can be employed with optimized power control policies, they are usually implemented by simple precoding power normalization techniques. This is due to the complexity of attaining optimal power control policies [23], as it requires coordination and cooperation among cells and computationally demanding algorithms. On the other hand, precoding power normalization techniques are simple and efficient [2], [3], [5], [6].

This work made use of large system analysis to show how the choice of precoding power normalization affects the experience of each individual user in the system. Particularly, we revealed that MN and VN treat the noise and interference in the same manner, but have different effects on pilot contamination and received signal power. We also revealed the key role played by large-scale fading, positions of UEs, and pilot assignment into power normalization. We explained how a simple change in power normalization can resemble two totally different behaviors, sum-rate maximization or fairness provisioning. Moreover, we showed numerically how the choice of the normalization technique can change the main bottleneck of massive MIMO systems.

VII. Appendices

The main strategy to find the DE of γ_{jk}^{VN} and γ_{jk}^{MN} for MRT, ZF, and RZF precoding schemes, is to replace their constitutive terms by their DEs, which will be detailed in the following. Also we frequently use the following notation.

$$\hat{\mathbf{H}}_{ll[k]}\hat{\mathbf{H}}_{ll[k]}^{H} = \hat{\mathbf{H}}_{ll}\hat{\mathbf{H}}_{ll}^{H} - \hat{\mathbf{h}}_{llk}\hat{\mathbf{h}}_{llk}^{H}$$

$$\mathbf{C}_l = \left(\hat{\mathbf{H}}_{ll}\hat{\mathbf{H}}_{ll}^{H} + \mathbf{Z}_l + \alpha_l N \mathbf{I}_N\right)^{-1} \qquad \overline{\mathbf{C}}_l = N\mathbf{C}_l$$

$$\mathbf{C}_{l[k]} = \left(\hat{\mathbf{H}}_{ll[k]}\hat{\mathbf{H}}_{ll[k]}^{H} + \mathbf{Z}_l + \alpha_l N \mathbf{I}_N\right)^{-1} \quad \overline{\mathbf{C}}_{l[k]} = N\mathbf{C}_{l[k]}.$$

Appendix A

We start from γ_{jk}^{VN} given in (16) while applying (7) and (10). Then we divide the numerator and denominator of $\gamma_{jk}^{(\mathrm{MRT-VN})}$ by N. Define $d_{li}^{\dagger} \triangleq N d_{li}$. Now by applying the continuous mapping theorem [30] and replacing each component of $\gamma_{jk}^{(\mathrm{MRT-VN})}$ by its DE, we have the DE of $\gamma_{jk}^{(\mathrm{MRT-VN})}$. The DE of signal power component, variance component, and interference components are given in [5], and we just need to find the DE of vector normalization coefficient, which is given as follows

$$d_{li} = \frac{1}{\mathbb{E}[\hat{\mathbf{h}}_{lli}^{H}\hat{\mathbf{h}}_{lli}]} = \frac{1}{\mathtt{tr}\mathbb{E}[\hat{\mathbf{h}}_{lli}\hat{\mathbf{h}}_{lli}^{H}]} \overset{a}{=} \frac{1}{\mathtt{tr}\boldsymbol{\Phi}_{lli}} \tag{52}$$

where in (a) we used the fact that $\hat{\mathbf{h}}_{lli} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Phi}_{lli})$.

Appendix B

We start from γ_{jk}^{VN} given in (16) while applying (9) and (10). Then we divide the numerator and denominator of $\gamma_{jk}^{(RZF-VN)}$ by N. Now we replace each component of $\gamma_{jk}^{(RZF-VN)}$ with its DE. The DE of signal power term, variance term, and interference terms are given in [5], and we just need to find the DE of VN coefficient. From (10) we have

$$d_{li}^{\circ} = \frac{1}{N} d_{li} = \frac{1}{N}(\mathbb{E}[\hat{\mathbf{h}}_{lli}^{H}\mathbf{C}_l^2\hat{\mathbf{h}}_{lli}])^{-1} \tag{53}$$

now consider the term $\hat{\mathbf{h}}_{lli}^{H}\mathbf{C}_l^2\hat{\mathbf{h}}_{lli}$, we can write

$$\begin{aligned}\hat{\mathbf{h}}_{lli}^{H}\mathbf{C}_l^2\hat{\mathbf{h}}_{lli} &\overset{a}{=} \frac{\hat{\mathbf{h}}_{lli}^{H}\mathbf{C}_{l[i]}^2\hat{\mathbf{h}}_{lli}}{(1+\hat{\mathbf{h}}_{lli}^{H}\mathbf{C}_{l[i]}\hat{\mathbf{h}}_{lli})^2} \overset{b}{\asymp} \frac{\frac{1}{N^2}\mathtt{tr}\boldsymbol{\Phi}_{lli}^{H}\overline{\mathbf{C}}_l^2}{(1+\frac{1}{N}\mathtt{tr}\boldsymbol{\Phi}_{lli}\overline{\mathbf{C}}_l)^2} \\ &\overset{c}{\asymp} \frac{\frac{1}{N^2}\mathtt{tr}\boldsymbol{\Phi}_{lli}^{H}\mathbf{T}'_{l,\mathbf{I}_N}}{(1+\frac{1}{N}\mathtt{tr}\boldsymbol{\Phi}_{lli}\mathbf{T}_l^{-1})^2}\end{aligned} \tag{54}$$

where (a) follows from Lemma 2, (b)[1] achieved by applying Lemma 3 and 4, and (c) follows from Theorem 7 and 8, while $\mathbf{S}_l = \frac{\mathbf{Z}_l}{N}$. Then due to continuous mapping and dominated convergence theorem we have

$$d_{li}^{\circ} \asymp \frac{(1+u_{li})^2}{\frac{1}{N}\mathtt{tr}\boldsymbol{\Phi}_{lli}^{H}\mathbf{T}'_{l,\mathbf{I}_N}}. \tag{55}$$

Appendix C

The main idea is to find the DE of ZF precoding from RZF precoding scheme by letting $\alpha_l = \alpha \ \ \forall l \in \mathcal{L}$ and $\alpha \to 0$. First we use the triangle inequality to bound $|\gamma_{lk}^{(\mathrm{ZF-VN})} - \overline{\gamma}_{lk}^{(\mathrm{ZF-VN})}|$ as follows

$$\begin{aligned}|\gamma_{lk}^{(\mathrm{ZF-VN})} - \overline{\gamma}_{lk}^{(\mathrm{ZF-VN})}| \leq\ & |\gamma_{lk}^{(\mathrm{ZF-VN})} - \gamma_{lk}^{(\mathrm{RZF-VN})}| \\ &+ |\gamma_{lk}^{(\mathrm{RZF-VN})} - \overline{\gamma}_{lk}^{(\mathrm{RZF-VN})}| \\ &+ |\overline{\gamma}_{lk}^{(\mathrm{RZF-VN})} - \overline{\gamma}_{lk}^{(\mathrm{ZF-VN})}|.\end{aligned} \tag{56}$$

Now we show each term in the right hand side of (56) can be made arbitrarily small (i.e. smaller than any given $\varepsilon > 0$) as long as $\alpha > 0$ becomes small enough. Let us consider the term $|\gamma_{lk}^{(\mathrm{ZF-VN})} - \gamma_{lk}^{(\mathrm{RZF-VN})}|$. Note that the difference between $\gamma_{lk}^{(\mathrm{ZF-VN})}$ and $\gamma_{lk}^{(\mathrm{RZF-VN})}$ is due to the different format of $\mathbf{F}_l$ in (8) and (9). As $\alpha \to 0$ and for $\mathbf{Z}_l = \mathbf{0}$ we have

$$\begin{aligned}\lim_{\mathbf{Z}_l=\mathbf{0},\alpha\to 0} \mathbf{f}_{lk}^{RZF} &= \lim_{\alpha\to 0}\left(\hat{\mathbf{H}}_{ll}\hat{\mathbf{H}}_{ll}^{H} + N\alpha\mathbf{I}_{\mathbf{N}}\right)^{-1}\hat{\mathbf{H}}_{ll}\mathbf{e}_k \\ &= \hat{\mathbf{H}}_{ll}\left(\hat{\mathbf{H}}_{ll}^{H}\hat{\mathbf{H}}_{ll}\right)^{-1}\mathbf{e}_k = \mathbf{f}_{lk}^{ZF}.\end{aligned} \tag{57}$$

Therefore the term $|\gamma_{lk}^{(\mathrm{ZF-VN})} - \gamma_{lk}^{(\mathrm{RZF-VN})}|$ can be made as small as we want as α goes to zero. For the second term on the right hand side of (56), from Theorem 2 we have for any $\alpha > 0$, $|\gamma_{lk}^{(\mathrm{RZF-VN})} - \overline{\gamma}_{lk}^{(\mathrm{RZF-VN})}| \xrightarrow[N\to\infty]{a.s} 0$. Now consider

[1] $a_N \asymp b_N$ is equivalent to $a_N - b_N \xrightarrow[N\to\infty]{a.s} 0$.

the third term, $|\overline{\gamma}_{lk}^{(\mathrm{RZF-VN})} - \overline{\gamma}_{lk}^{(\mathrm{ZF-VN})}|$. Define $\overline{\gamma}_{lk}^{(\mathrm{ZF-VN})} := \lim_{\alpha\to 0} \overline{\gamma}_{lk}^{(\mathrm{RZF-VN})}$, we need to find its value. We have

$$\lim_{\alpha\to 0} \overline{\gamma}_{jk}^{(\mathrm{RZF-VN})} \tag{58}$$

$$= \lim_{\alpha\to 0} \frac{d_{jk}^{\circ} \frac{u_{jk}^2}{(1+u_{jk})^2}}{\frac{\sigma^2}{N\rho_{dl}} + \frac{1}{N}\sum_{l=1}^{L}\sum_{i=1}^{K} d_{li}^{\circ} \frac{\epsilon_{li,jk}}{(1+u_{li})^2} + \sum_{l=1,l\neq j}^{L} d_{lk}^{\circ} \frac{|u_{ljk}|^2}{(1+u_{lk})^2}}$$

$$= \lim_{\alpha\to 0} \frac{d_{jk}^{\circ} \frac{\alpha^2 u_{jk}^2}{(\alpha+\alpha u_{jk})^2}}{\frac{\sigma^2}{N\rho_{dl}} + \frac{1}{N}\sum_{l=1}^{L}\sum_{i=1}^{K} d_{li}^{\circ} \frac{\alpha^2\epsilon_{li,jk}}{(\alpha+\alpha u_{li})^2} + \sum_{l=1,l\neq j}^{L} d_{lk}^{\circ} \frac{|\alpha u_{ljk}|^2}{(\alpha+\alpha u_{lk})^2}}$$

Define $\underline{u_{lk}} := \lim_{\alpha\to 0} \alpha u_{lk}$ for every l and k. Based on [15] and by replacing u_{lk} from Theorem 2 we have

$$\underline{u_{lk}} = \lim_{\alpha\to 0} \alpha u_{lk} = \lim_{\alpha\to 0} \frac{1}{N} tr\boldsymbol{\Phi}_{llk} \Big(\frac{1}{N}\sum_{i=1}^{K} \frac{\boldsymbol{\Phi}_{lli}}{\alpha u_{li}} + \mathbf{I}_N\Big)^{-1}$$

$$= \frac{1}{N}\mathtt{tr}\left(\boldsymbol{\Phi}_{llk}\underline{\mathbf{T}_l}\right) \tag{59}$$

with

$$\underline{\mathbf{T}_l} = \left(\frac{1}{N}\sum_{i=1}^{K} \frac{\boldsymbol{\Phi}_{lli}}{\underline{u_{li}}} + \mathbf{I}_N\right)^{-1}. \tag{60}$$

Also defining $\underline{u_{ljk}} \triangleq \lim_{\alpha\to 0} \alpha u_{ljk}$ for every l, j, and k we have

$$\underline{u_{ljk}} = \lim_{\alpha\to 0} \alpha_l \frac{1}{N} tr\boldsymbol{\Phi}_{ljk}\mathbf{T}_l = \frac{1}{N} tr\boldsymbol{\Phi}_{ljk}\underline{\mathbf{T}_l}. \tag{61}$$

For the term $\lim_{\alpha\to 0} d_{li}^{\circ}$ we have

$$\lim_{\alpha\to 0} d_{li}^{\circ} = \lim_{\alpha\to 0} \frac{(1+u_{li})^2}{\frac{1}{N} tr\boldsymbol{\Phi}_{lli}\mathbf{T}'_{l,\mathbf{I}_N}} \tag{62}$$

$$= \lim_{\alpha\to 0} \frac{\underline{u_{li}}^2}{\alpha^2 \frac{1}{N} tr\boldsymbol{\Phi}_{lli}\mathbf{T}'_{l,\mathbf{I}_N}} = \lim_{\alpha\to 0} \frac{\underline{u_{li}}^2}{\alpha^2 u'_{li,\mathbf{I}_N}}$$

where $\mathbf{T}'_{l,\mathbf{I}_N}$ and $u'_{li,\mathbf{I}_N}$ are given in Theorem 8. Now note that

$$\lim_{\alpha\to 0} \alpha^2 u'_{li,\mathbf{I}_N} = \alpha^2 \frac{1}{N} tr\boldsymbol{\Phi}_{lli}\mathbf{T}'_{l,\mathbf{I}_N} \tag{63}$$

$$= \lim_{\alpha\to 0} \alpha^2 \frac{1}{N} tr\boldsymbol{\Phi}_{lli}\mathbf{T}_l \left(\frac{1}{N}\sum_{t=1}^{K} \frac{u'_{lt,\mathbf{I}_N}\boldsymbol{\Phi}_{llt}}{(1+u_{lt})^2} + \mathbf{I}_M\right)\mathbf{T}_l$$

$$= \frac{1}{N} tr\boldsymbol{\Phi}_{lli}\underline{\mathbf{T}_l} \left(\frac{1}{N}\sum_{t=1}^{K} \frac{(\lim_{\alpha\to 0}\alpha^2 u'_{lt,\mathbf{I}_N})\boldsymbol{\Phi}_{llt}}{\underline{u_{lt}}^2} + \mathbf{I}_N\right)\underline{\mathbf{T}_l}$$

replacing $\lim_{\alpha\to 0}\alpha^2 u'_{lt,\mathbf{I}_N}$ with $\underline{u_{lt}}$ for every t, (63) reduces to $\underline{u_{li}} = \frac{1}{N}\mathtt{tr}\boldsymbol{\Phi}_{lli}\underline{\mathbf{T}_l}$. Therefore $\lim_{\alpha\to 0}\alpha^2 u'_{lk,\mathbf{I}_N} = \underline{u_{lk}}$. Hence from (62) we can conclude that $d_{li}^{\circ} = \underline{u_{li}}$. For $\underline{\epsilon_{li,jk}}$ we have

$$\underline{\epsilon_{li,jk}} = \lim_{\alpha\to 0} \alpha^2 \epsilon_{li,jk} \tag{64}$$

$$= \frac{1}{N}\mathtt{tr}\boldsymbol{\Theta}_{ljk}\underline{\mathbf{T}'_{l,\boldsymbol{\Phi}_{lli}}} - \frac{2}{\underline{u_{lk}}}\,\mathrm{Re}\left(\underline{u_{ljk}^*}\frac{1}{N}\mathtt{tr}\boldsymbol{\Phi}_{ljk}\underline{\mathbf{T}'_{l,\boldsymbol{\Phi}_{lli}}}\right)$$

$$+ \frac{\underline{|u_{ljk}|^2}}{\underline{u_{lk}}^2}\frac{1}{N}\mathtt{tr}\boldsymbol{\Phi}_{llk}\underline{\mathbf{T}'_{l,\boldsymbol{\Phi}_{lli}}}$$

where $\underline{\mathbf{T}'_{l,\boldsymbol{\Phi}_{lli}}} \triangleq \lim_{\alpha\to 0}\alpha^2\mathbf{T}'_{l,\boldsymbol{\Phi}_{lli}}$ is

$$\underline{\mathbf{T}'_{l,\boldsymbol{\Phi}_{lli}}} = \lim_{\alpha\to 0}\alpha^2\mathbf{T}_l\left[\frac{1}{N}\sum_{t=1}^{K}\frac{u'_{lt,\boldsymbol{\Phi}_{lli}}\boldsymbol{\Phi}_{llt}}{(1+u_{lt})^2} + \boldsymbol{\Phi}_{lli}\right]\mathbf{T}_l$$

$$= \underline{\mathbf{T}_l}\left[\frac{1}{N}\sum_{t=1}^{K}\frac{\underline{u'_{lt,\boldsymbol{\Phi}_{lli}}}\boldsymbol{\Phi}_{llt}}{\underline{u_{lt}}^2} + \boldsymbol{\Phi}_{lli}\right]\underline{\mathbf{T}_l} \tag{65}$$

where $\underline{u'_{lt,\boldsymbol{\Phi}_{lli}}} = \lim_{\alpha\to 0}\alpha^2 u'_{lt,\boldsymbol{\Phi}_{lli}}$. From Theorem 8 we have

$$\underline{\mathbf{u}'_{l,\boldsymbol{\Phi}_{lli}}} = \lim_{\alpha\to 0}(\mathbf{I}_K - \mathbf{J}_l)^{-1}\alpha^2\mathbf{v}_{l,\boldsymbol{\Phi}_{lli}} = (\mathbf{I}_K - \underline{\mathbf{J}_l})^{-1}\underline{\mathbf{v}_{l,\boldsymbol{\Phi}_{lli}}}$$

where $\underline{\mathbf{J}_l}$ and $\underline{\mathbf{v}_{l,\boldsymbol{\Phi}_{lli}}}$ are given by (33) and (34), respectively. Therefore $\underline{\epsilon_{lk,jn}} = \lim_{\alpha\to 0}\alpha^2\epsilon_{lk,jn}$ follows (30). Replacing all these limits in (58) completes the proof.

APPENDIX D

For brevity we only consider ZF with VN. The same steps can be used for ZF with MN. If the channel is modelled as in (40), then $\boldsymbol{\Theta}_{ljk} = d_{ljk}\mathbf{I}_N$ and

$$\boldsymbol{\Phi}_{ljk} = \frac{d_{llk}d_{ljk}}{\alpha_{lk}}\mathbf{I}_N \tag{66}$$

with $\alpha_{lk} = \sum_{n=1}^{L} d_{lnk} + \frac{\sigma^2}{\rho_{\mathrm{tr}}}$. Plugging (66) into (27) and (28) yields $\underline{u_{lk}} = \frac{d_{llk}^2}{\alpha_{lk}}\frac{1}{N}\mathtt{tr}\left(\underline{\mathbf{T}_l}\right)$ with

$$\underline{\mathbf{T}_l} = \left(\frac{1}{N}\sum_{i=1}^{K}\frac{1}{\frac{1}{N}\mathtt{tr}\left(\underline{\mathbf{T}_l}\right)} + 1\right)^{-1}\mathbf{I}_N. \tag{67}$$

Call $\bar{u} = \frac{1}{N}\mathtt{tr}\left(\underline{\mathbf{T}_l}\right)$. Therefore, we have that

$$\bar{u} = \frac{1}{N}\mathtt{tr}\left(\underline{\mathbf{T}_l}\right) = \left(\frac{K}{N}\frac{1}{\bar{u}} + 1\right)^{-1}. \tag{68}$$

Solving with respect to $\bar{u}$ yields $\bar{u} = 1 - \frac{K}{N}$. Then, we eventually have that

$$\underline{u_{lk}} = \frac{d_{llk}^2}{\alpha_{lk}}\bar{u} \tag{69}$$

and also $u_{ljk} = \frac{d_{llk}d_{ljk}}{\alpha_{lk}}\bar{u}$. Therefore, the pilot contamination term in $\overline{\gamma}_{jk}^{(\mathrm{ZF-VN})}$ reduces to

$$\sum_{l=1,l\neq j}^{L}\frac{\underline{u_{ljk}^2}}{\underline{u_{lk}}} = \sum_{l=1,l\neq j}^{L}\frac{d_{ljk}^2}{\alpha_{lk}}\bar{u}. \tag{70}$$

Let's now compute $[\underline{\mathbf{J}_l}]_{n,i}$ defined as in (33). Using the above results yields

$$[\underline{\mathbf{J}_l}]_{n,i} = \frac{1}{N^2}\frac{d_{lln}^2}{\alpha_{ln}}\frac{d_{lli}^2}{\alpha_{li}}\frac{1}{\underline{u_{li}}^2}\mathtt{tr}\left(\underline{\mathbf{T}}^2\right) = \frac{1}{N}\frac{d_{lln}^2}{\alpha_{ln}}\frac{\alpha_{li}}{d_{lli}^2}. \tag{71}$$

Similarly, we have that

$$[\underline{\mathbf{v}_{k,l}}]_i = \frac{d_{lli}^2}{\alpha_{li}}\frac{d_{llk}^2}{\alpha_{lk}}\bar{u}^2. \tag{72}$$

In compact form, we may write $\underline{\mathbf{J}_l}$ and $\underline{\mathbf{v}_{l,k}}$ as

$$\underline{\mathbf{J}_l} = \frac{1}{N}\mathbf{a}_l\mathbf{b}_l^T \qquad \underline{\mathbf{v}_{l,k}} = \frac{d_{llk}^2}{\alpha_{lk}}\bar{u}^2\mathbf{a}_l \tag{73}$$

with $[\mathbf{a}_l]_i = d_{lli}^2/\alpha_{li}$ and $[\mathbf{b}_l]_i = 1/[\mathbf{a}_l]_i$. Then, we have that (applying matrix inversion Lemma)

$$\underline{\mathbf{u}'_{k,l}} = \frac{d_{llk}^2}{\alpha_{lk}}\bar{u}^2\left(\mathbf{I}_K - \frac{1}{N}\mathbf{a}_l\mathbf{b}_l^T\right)^{-1}\mathbf{a}_l = \frac{d_{llk}^2}{\alpha_{lk}}\bar{u}\mathbf{a}_l = u_{lk}\mathbf{a}_l. \quad (74)$$

Plugging the above result into (31) produces

$$\underline{\mathbf{T}'_{l,\mathbf{\Phi}_{lli}}} = \frac{d_{lli}^2}{\alpha_{li}}\underline{\mathbf{T}_l}\left(\frac{K}{N}\frac{1}{\bar{u}} + 1\right)\underline{\mathbf{T}_l} = \frac{d_{lli}^2}{\alpha_{li}}\bar{u}\mathbf{I}_N = \underline{u_{li}}\mathbf{I}_N. \quad (75)$$

We are thus left with evaluating (30). Using the above results yields

$$\begin{aligned}\underline{\epsilon_{li,jk}} = &\frac{d_{ljk}}{N}\mathtt{tr}\left(\mathbf{T}'_{l,\mathbf{\Phi}_{lli}}\right) - 2\frac{d_{ljk}}{d_{llk}}\frac{1}{N}\mathtt{tr}\left(\mathbf{\Phi}_{ljk}\mathbf{T}'_{l,\mathbf{\Phi}_{lli}}\right) + \\ &+ \frac{d_{ljk}^2}{d_{llk}^2}\frac{1}{N}\mathtt{tr}\left(\mathbf{\Phi}_{llk}\mathbf{T}'_{l,\mathbf{\Phi}_{lli}}\right) \quad (76)\end{aligned}$$

from which, using (66) and (75), we obtain

$$\underline{\epsilon_{li,jk}} = d_{ljk}\underline{u_{li}} - \frac{d_{ljk}^2}{\alpha_{lk}}\underline{u_{li}}.$$

Therefore, we have that

$$\frac{1}{N}\frac{\underline{\epsilon_{li,jk}}}{\underline{u_{li}}} = \frac{1}{N}d_{ljk}\left(1 - \frac{d_{ljk}}{\alpha_{lk}}\right). \quad (77)$$

Plugging (69), (70) and (77) into (26) produces

$$\frac{1}{N}\sum_{l=1}^{L}\sum_{i=1}^{K}\frac{\underline{\epsilon_{li,jk}}}{\underline{u_{li}}} = \frac{K}{N}\sum_{l=1}^{L}d_{ljk}\left(1 - \frac{d_{ljk}}{\alpha_{lk}}\right). \quad (78)$$

Collecting all the above results together completes the proof.

APPENDIX E

A. Required Theorems

Theorem 7. *[15, Theorem 1] Let $\mathbf{B}_l = \frac{1}{N}\hat{\mathbf{H}}_{ll}\hat{\mathbf{H}}_{ll}^H + \mathbf{S}_l$ with $\hat{\mathbf{H}}_{ll} \in \mathbb{C}^{N\times K}$ be random with independent column vectors $\hat{\mathbf{h}}_{llk} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Phi}_{llk})$ for $k \in \{1,\ldots,K\}$, $\mathbf{S}_l \in \mathbb{C}^{N\times N}$ and $\mathbf{Q}_l \in \mathbb{C}^{N\times N}$ be Hermitian nonnegative definite. Assume that $\mathbf{Q}_l$ and the matrices $\mathbf{\Phi}_{llk}$ for $k \in \{1,\ldots,K\}$ have uniformly bounded spectral norms (with respect to N). Define*

$$m_{\mathbf{B}_l,\mathbf{Q}_l}(-\alpha_l) \triangleq \frac{1}{N}\mathtt{tr}\mathbf{Q}_l(\mathbf{B}_l + \alpha_l\mathbf{I}_N)^{-1} \quad (79)$$

Then, for any $\alpha_l > 0$, as N and K grow large with $\beta = \frac{N}{K}$ such that $0 < \liminf\beta \le \limsup\beta < \infty$ we have that

$$m_{\mathbf{B}_l,\mathbf{Q}_l}(-\alpha_l) - m^{\circ}_{\mathbf{B}_l,\mathbf{Q}_l}(-\alpha_l) \xrightarrow[N\to\infty]{a.s} 0 \quad (80)$$

where $m^{\circ}_{\mathbf{B}_l,\mathbf{Q}_l}(-\alpha_l)$ is given by

$$m^{\circ}_{\mathbf{B}_l,\mathbf{Q}_l}(-\alpha_l) = \frac{1}{N}\mathtt{tr}\mathbf{Q}_l\mathbf{T}_l \quad (81)$$

with $\mathbf{T}_l$ is given by

$$\mathbf{T}_l = \left(\frac{1}{N}\sum_{i=1}^{K}\frac{\mathbf{\Phi}_{lli}}{1 + u_{li}(-\alpha_l)} + \mathbf{S}_l + \alpha_l\mathbf{I}_N\right)^{-1} \quad (82)$$

where the elements of $\mathbf{u}_l(-\alpha_l) = [u_{l1}(-\alpha_l),\ldots,u_{lK}(-\alpha_l)]^T$ are defined as $u_{li}(-\alpha_l) = \lim_{t\to\infty}u_{li}^{(t)}(-\alpha_l)$, where for $t \in \{1,2,\ldots\}$

$$\begin{aligned}&u_{lk}^{(t)}(-\alpha_l) = \quad (83)\\ &\frac{1}{N}\mathtt{tr}\mathbf{\Phi}_{llk}\left(\frac{1}{N}\sum_{i=1}^{K}\frac{\mathbf{\Phi}_{lli}}{1 + u_{li}^{(t-1)}(-\alpha_l)} + \mathbf{S}_l + \alpha_l\mathbf{I}_N\right)^{-1}\end{aligned}$$

with initial values $u_{lk}^{(0)}(-\alpha) = \frac{1}{\alpha}$ for all k.

Also we need the following theorem.

Theorem 8. *[15] Let $\mathbf{\Omega}_l \in \mathbb{C}^{N\times N}$ be Hermitian nonnegative definite with uniformly bounded spectral norm (with respect to N). Under the conditions of Theorem 1*

$$\begin{aligned}&\frac{1}{N}tr\mathbf{Q}_l(\frac{1}{N}\hat{\mathbf{H}}_{ll}\hat{\mathbf{H}}_{ll}^H + \mathbf{S}_l + \alpha_l\mathbf{I}_N)^{-1}\mathbf{\Omega}_l... \quad (84)\\ &(\frac{1}{N}\hat{\mathbf{H}}_{ll}\hat{\mathbf{H}}_{ll}^H + \mathbf{S}_l + \alpha_l\mathbf{I}_N)^{-1} - \frac{1}{N}tr\mathbf{Q}_l\mathbf{T}'_{l,\mathbf{\Omega}_l} \xrightarrow[N\to\infty]{a.s} 0\end{aligned}$$

where $\mathbf{T}'_{l,\mathbf{\Omega}_l} \in \mathbb{C}^{N\times N}$ is defined as

$$\mathbf{T}'_{l,\mathbf{\Omega}_l} = \mathbf{T}_l \times \left(\frac{1}{N}\sum_{j=1}^{K}\frac{u'_{lj,\mathbf{\Omega}_l}(-\alpha_l)\mathbf{\Phi}_{llj}}{(1 + u_{lj}(-\alpha_l))^2} + \mathbf{\Omega}_l\right) \times \mathbf{T}_l \quad (85)$$

where $\mathbf{T}_l$ and $\mathbf{u}_l(-\alpha)$ are given by theorem 1, and $\mathbf{u}'_{l,\mathbf{\Omega}_l}(-\alpha) = [u'_{l1,,\mathbf{\Omega}_l}(-\alpha),\ldots,u'_{lK,,\mathbf{\Omega}_l}(-\alpha)]^T$ is computed from

$$\mathbf{u}'_{l,\mathbf{\Omega}_l}(-\alpha) = (\mathbf{I}_K - \mathbf{J}_l)^{-1}\mathbf{v}_{l,\mathbf{\Omega}_l} \quad (86)$$

where $\mathbf{J}_l \in \mathbb{C}^{K\times K}$ and $\mathbf{v}_l \in \mathbb{C}^K$ are:

$$[\mathbf{J}_l]_{mn} = \frac{tr\mathbf{\Phi}_{llm}\mathbf{T}_l\mathbf{\Phi}_{lln}\mathbf{T}_l}{N^2(1 + u_{ln}(-\alpha))^2} \quad 1 \le m,n \le K \quad (87)$$

$$[\mathbf{v}_{l,\mathbf{\Omega}_l}]_{t1} = \frac{1}{N}tr\mathbf{\Phi}_{llt}\mathbf{T}_l\mathbf{\Omega}_l\mathbf{T}_l \quad 1 \le t \le K \quad (88)$$

B. Required Lemmas

Lemma 2. *(Matrix Inversion Lemma)*
Let $\mathbf{U}$ be an $N\times N$ invertible matrix and $x \in \mathbb{C}^N$, $c \in \mathbb{C}$ for which $\mathbf{U} + c\,\mathbf{x}\mathbf{x}^H$ is invertible. Then

$$\mathbf{x}^H\left(\mathbf{U} + c\,\mathbf{x}\mathbf{x}^H\right)^{-1} = \frac{\mathbf{x}^H\mathbf{U}^{-1}}{1 + c\,\mathbf{x}^H\mathbf{U}^{-1}\mathbf{x}} \quad (89)$$

Lemma 3. *(Trace Lemma)*
Let $\mathbf{A} \in \mathbb{C}^{N\times N}$ and $\mathbf{x},\mathbf{y} \sim \mathcal{CN}(\mathbf{0}, \frac{1}{N}\mathbf{I}_N)$. Assume that $\mathbf{A}$ has uniformly bounded spectral norm (with respect to N) and that $\mathbf{x}$ and $\mathbf{y}$ are mutually independent and independent of $\mathbf{A}$. Then, for all $p \ge 1$,

$$\mathbf{x}^H\mathbf{A}\mathbf{x} - \frac{1}{N}tr\mathbf{A} \xrightarrow[N\to\infty]{a.s} 0 \quad \text{and} \quad \mathbf{x}^H\mathbf{A}\mathbf{y} \xrightarrow[N\to\infty]{a.s} 0 \quad (90)$$

Lemma 4. *(Rank-1 perturbation lemma)*
Let $\mathbf{A}_1$, $\mathbf{A}_2$, ..., with $\mathbf{A}_N \in \mathbb{C}^{N\times N}$, be deterministic with uniformly bounded spectral norm and $\mathbf{B}_1$, $\mathbf{B}_2$, ..., with $\mathbf{B}_N \in \mathbb{C}^{N\times N}$, be random Hermitian, with eigenvalues $\lambda_1^{\mathbf{B}_N} \le \lambda_2^{\mathbf{B}_N} \le \ldots \le \lambda_N^{\mathbf{B}_N}$ such that, with probability 1, there exist $\epsilon > 0$ for which $\lambda_1^{\mathbf{B}_N} > \epsilon$ for all large N. Then for $\mathbf{v} \in \mathbb{C}^N$

$$\frac{1}{N}tr\mathbf{A}_N\mathbf{B}_N^{-1} - \frac{1}{N}tr\mathbf{A}_N(\mathbf{B}_N^{-1} + \mathbf{v}\mathbf{v}^H)^{-1} \xrightarrow[N\to\infty]{a.s} 0 \quad (91)$$

where $\mathbf{B}_N^{-1}$ and $(\mathbf{B}_N^{-1} + \mathbf{v}\mathbf{v}^H)^{-1}$ exist with probability 1.

REFERENCES

[1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. on Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, November 2010.

[2] F. Rusek, D. Persson, B. K. Lau, E. Larsson, T. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Proc. Mag.*, vol. 30, no. 1, pp. 40 – 60, Jan. 2013.

[3] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, 2014.

[4] L. Liu, C. Yuen, Y. L. Guan, Y. Li, and Y. Su, "Convergence analysis and assurance for gaussian message passing iterative detector in massive MU-MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6487–6501, Sep. 2016.

[5] J. Hoydis, S. Ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, 2013.

[6] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, 2013.

[7] J. Zuo, J. Zhang, C. Yuen, W. Jiang, and W. Luo, "Energy-efficient downlink transmission for multicell massive DAS with pilot contamination," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 1209–1221, Feb. 2017.

[8] E. Björnson, J. Hoydis, and L. Sanguinetti, "Pilot contamination is not a fundamental asymptotic limitation in massive MIMO," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Paris, France, May 2017, to appear. [Online]. Available: https://arxiv.org/abs/1611.09152

[9] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, 2014.

[10] T. L. Marzetta, "How much training is required for multiuser MIMO?" in *IEEE Asilomar Conf. Signals, Sys. and Comput.*, Pacific Grove, CA, Nov. 2006, pp. 359–363.

[11] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186 – 195, Feb. 2014.

[12] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[13] H. Yang and T. L. Marzetta, "Performance of conjugate and zero-forcing beamforming in large-scale antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 172–179, Feb. 2013.

[14] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, 2011.

[15] S. Wagner, R. Couillet, M. Debbah, and D. Slock, "Large system analysis of linear precoding in correlated miso broadcast channels under limited feedback," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4509–4537, Jul. 2012.

[16] J. Zuo, J. Zhang, C. Yuen, W. Jiang, and W. Luo, "Multi-cell multi-user massive MIMO transmission with downlink training and pilot contamination precoding," *IEEE Trans. Veh. Technol.*, Aug. 2015.

[17] A. Khansefid and H. Minn, "Achievable downlink rates of MRC and ZF precoders in massive MIMO with uplink and downlink pilot contamination," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4849–4864, Dec. 2015.

[18] A. Kammoun, A. Müller, E. Björnson, and M. Debbah, "Linear precoding based on polynomial expansion: Large-scale multi-cell MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 861–875, Oct. 2014.

[19] R. Krishnan, M. R. Khanzadi, N. Krishnan, Y. Wu, A. G. i Amat, T. Eriksson, and R. Schober, "Linear massive MIMO precoders in the presence of phase noise - a large-scale analysis," *IEEE Trans. Veh. Technol.*, vol. 65, no. 5, pp. 3057–3071, May 2016.

[20] X. Li, E. Bjornson, E. G. Larsson, S. Zhou, and J. Wang, "A multi-cell MMSE precoder for massive MIMO systems and new large system analysis," in *IEEE Int. Conf. Global Commun. (GLOBECOM)*, San Diego, CA, Dec 2015, pp. 1–6.

[21] M. Sadeghi and C. Yuen, "Multi-cell multi-group massive MIMO multicasting: An asymptotic analysis," in *Proc. IEEE Global Commun. Conf. (Globecom)*, San Diego, CA, Dec. 2015, pp. 1–6.

[22] T. V. Chien, E. Bjornson, and E. G. Larsson, "Joint power allocation and user association optimization for massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6384–6399, Sep. 2016.

[23] Q. Zhang, S. Jin, M. McKay, D. Morales-Jimenez, and H. Zhu, "Power allocation schemes for multicell massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 5941–5955, Nov. 2015.

[24] M. Sadeghi, L. Sanguinetti, R. Couillet, and C. Yuen, "Reducing the computational complexity of multicasting in large-scale antenna systems," *IEEE Trans. Wireless Commun.*, 2017, to appear. [Online]. Available: https://arxiv.org/abs/1702.05901

[25] C. Lee, C. B. Chae, T. Kim, S. Choi, and J. Lee, "Network massive MIMO for cell-boundary users: From a precoding normalization perspective," in *IEEE Globecom Workshops (GC Wkshps)*, Anaheim, CA, Dec. 2012, pp. 233–237.

[26] Y.-G. Lim, C.-B. Chae, and G. Caire, "Performance analysis of massive MIMO for cell-boundary users," *IEEE Trans. Wireless Commun.*, vol. 14, no. 12, pp. 6827–6842, Dec. 2015.

[27] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2845–2866, Jun. 2010.

[28] E. Bjornson, E. G. Larsson, T. L. Marzetta *et al.*, "Massive MIMO: Ten myths and one critical question," *IEEE Commun. Mag.*, vol. 54, no. 2, pp. 114–123, Feb. 2016.

[29] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 933–946, May 2000.

[30] P. Billingsley, *Probability and measure*. John Wiley & Sons, 2008.

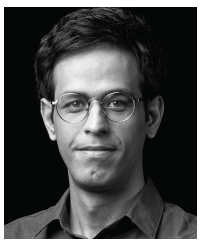
[31] L. Sanguinetti, A. L. Moustakas, and M. Debbah, "Interference management in 5G reverse TDD hetnets with wireless backhaul: A large system analysis," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 6, pp. 1187–1200, Jun. 2015.

[32] L. Sanguinetti, A. L. Moustakas, E. Björnson, and M. Debbah, "Large system analysis of the energy consumption distribution in multi-user MIMO systems with mobility," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1730–1745, Mar. 2015.

[33] K. T. Truong and R. W. Heath, "The viability of distributed antennas for massive MIMO systems," in *2013 Asilomar Conf. on Signals, Syst. and Comput.*, Pacific Grove, CA, Nov. 2013, pp. 1318–1323.

[34] M. Sadeghi, C. Yuen, and Y. H. Chew, "Sum rate maximization for uplink distributed massive MIMO systems with limited backhaul capacity," in *2014 IEEE Globecom Workshops (GC Wkshps)*, Austin, TX, Dec. 2014, pp. 308–313.

[35] H. Q. Ngo, T. L. Marzetta, and E. G. Larsson, "Analysis of the pilot contamination effect in very large multicell multiuser MIMO systems for physical channel models," in *2011 IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP)*, Prague, Czech, May 2011, pp. 3464–3467.



Meysam Sadeghi received his B.S. and M.S. degrees in Electrical Engineering from Shiraz University of Technology and K. N. Toosi University of Technology, Iran. He is currently pursuing his Ph.D. degree at Singapore University of Technology and Design, Singapore. He has been a visiting research scholar at Linköping university, CentraleSupélec, and university of Pisa. He is the recipient of Merlion, SPGF, and SINGA awards. His research interest includes wireless communication and signal processing, and in particular Massive MIMO systems.

Luca Sanguinetti (SM'15) received the Laurea Telecommunications Engineer degree (cum laude) and the Ph.D. degree in information engineering from the University of Pisa, Italy, in 2002 and 2005, respectively. Since 2005 he has been with the Dipartimento di Ingegneria dell'Informazione of the University of Pisa. During the period June 2007 - June 2008, he was a postdoctoral associate in the Dept. Electrical Engineering at Princeton. During the period June 2010 - Sept. 2010, he was selected for a research assistantship at the Technische Universität Munchen. From July 2013 to July 2015, he was with the Alcatel-Lucent Chair on Flexible Radio, Supélec, Gif-sur-Yvette, France. He is an Assistant Professor at the Dipartimento di Ingegneria dell'Informazione of the University of Pisa. L. Sanguinetti is currently serving as an Associate Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE SIGNAL PROCESSING LETTERS. He served as Lead Guest Editor of IEEE JOURNAL ON SELECTED AREAS OF COMMUNICATIONS Special Issue on "Game Theory for Networks and as an Associate Editor for IEEE JOURNAL ON SELECTED AREAS OF COMMUNICATIONS (series on Green Communications and Networking). His expertise and general interests span the areas of communications and signal processing, game theory and random matrix theory for wireless communications. He was the co-recipient of two best paper awards: *IEEE Wireless Commun. and Networking Conference (WCNC) 2013* and *IEEE Wireless Commun. and Networking Conference (WCNC) 2014*. He was also the recipient of the FP7 Marie Curie IEF 2013 "Dense deployments for green cellular networks".

Romain Couillet received his MSc in Mobile Communications at the Eurecom Institute and his MSc in Communication Systems in Telecom ParisTech, France in 2007. From 2007 to 2010, he worked with ST-Ericsson as an Algorithm Development Engineer on the Long Term Evolution Advanced project, where he prepared his PhD with Supelec, France, which he graduated in November 2010. He is currently a full professor in the LSS laboratory at CentraleSupélec, France. His research topics are in random matrix theory applied to statistics, machine learning, signal processing, and wireless communications. He is an IEEE Senior Member. In 2015, he received the HDR title from University ParisSud. He is the recipient of the 2013 CNRS Bronze Medal in the section "science of information and its interactions", of the 2013 IEEE ComSoc Outstanding Young Researcher Award (EMEA Region), of the 2011 EEA/GdR ISIS/GRETSI best PhD thesis award, and of the Valuetools 2008 best student paper award.

Chau Yuen (S'02M'08SM'12) received the B. Eng and PhD degree from Nanyang Technological University, Singapore in 2000 and 2004 respectively. Dr Yuen was a Post Doc Fellow in Lucent Technologies Bell Labs, Murray Hill during 2005. He was a Visiting Assistant Professor of Hong Kong Polytechnic University in 2008. During the period of 2006 - 2010, he worked at the Institute for Infocomm Research (Singapore) as a Senior Research Engineer. He joined Singapore University of Technology and Design from June 2010. He serves as an Associate Editor for IEEE Transactions on Vehicular Technology. On 2012, he received IEEE Asia-Pacific Outstanding Young Researcher Award.