Massive MIMO in the UL/DL of Cellular Networks: How Many Antennas Do We Need?

Jakob Hoydis, S. ten Brink, Mérouane Debbah

Massive MIMO in the UL/DL of Cellular Networks: How Many Antennas Do We Need?

Jakob Hoydis, *Member, IEEE*, Stephan ten Brink, *Senior Member, IEEE*,
and Mérouane Debbah, *Senior Member, IEEE*

***Abstract*—We consider the uplink (UL) and downlink (DL) of non-cooperative multi-cellular time-division duplexing (TDD) systems, assuming that the number N of antennas per base station (BS) and the number K of user terminals (UTs) per cell are large. Our system model accounts for channel estimation, pilot contamination, and an arbitrary path loss and antenna correlation for each link. We derive approximations of achievable rates with several linear precoders and detectors which are proven to be asymptotically tight, but accurate for realistic system dimensions, as shown by simulations. It is known from previous work assuming uncorrelated channels, that as $N \to \infty$ while K is fixed, the system performance is limited by pilot contamination, the simplest precoders/detectors, i.e., eigenbeamforming (BF) and matched filter (MF), are optimal, and the transmit power can be made arbitrarily small. We analyze to which extent these conclusions hold in the more realistic setting where N is not extremely large compared to K. In particular, we derive how many antennas per UT are needed to achieve η% of the ultimate performance limit with infinitely many antennas and how many more antennas are needed with MF and BF to achieve the performance of minimum mean-square error (MMSE) detection and regularized zero-forcing (RZF), respectively.**

***Index Terms*—massive MIMO, time-division duplexing, channel estimation, pilot contamination, large system analysis, large random matrix theory, linear precoding, linear detection**

I. Introduction

VERY large multiple-input multiple-output (MIMO) or "massive MIMO" time-division duplexing (TDD) systems [1], [2] are currently investigated as a novel cellular network architecture with several attractive features: First, the capacity can be theoretically increased by simply installing additional antennas to existing cell sites. Thus, massive MIMO provides an alternative to cell-size shrinking, the traditional way of increasing the network capacity [3]. Second, large antenna arrays can potentially reduce uplink (UL) and downlink (DL) transmit powers through coherent combining and an increased antenna aperture [4]. This aspect is not only relevant from a business point of view but also addresses environmental as well as health concerns related to mobile communications [5], [6]. Third, if channel reciprocity is exploited, the overhead related to channel training scales linearly with the number K of user terminals (UTs) per cell and is independent of the number N of antennas per base station (BS). Consequently, additional antennas do not increase the feedback overhead and, therefore, "always help" [7]. Fourth, if $N \gg K$, the simplest linear precoders and detectors are optimal, thermal noise, interference, and channel estimation errors vanish, and the only remaining performance limitation is pilot contamination [1], i.e., residual interference which is caused by the reuse of pilot sequences in adjacent cells.

The features described above are based on several crucial but optimistic assumptions about the propagation conditions, hardware implementations, and the number of antennas which can be deployed in practice. Therefore, recent papers study massive MIMO under more realistic assumptions, e.g., a physical channel model with a finite number of degrees of freedom (DoF) [8] or constant-envelope transmissions with per-antenna power constraints [9]. Also first channel measurements with large antenna arrays were reported in [10], [11], [12].

In this work, we provide a unified performance analysis of the UL and DL of non-cooperative multi-cell TDD systems. We consider a realistic system model which accounts for imperfect channel estimation, pilot contamination, antenna correlation, and path loss. Assuming that N and K are large, we derive asymptotically tight approximations of the achievable rates with several linear precoders/detectors, i.e., eigenbeamforming (BF) and regularized zero-forcing (RZF) in the DL, matched filter (MF) and minimum mean-square error detector (MMSE) in the UL. These approximations are easy to compute and shown to be accurate for realistic system dimensions. We then distinguish massive MIMO from "classical" MIMO as a particular operating condition of cellular networks where multiuser interference, channel estimation errors, and noise have a negligible impact compared to pilot contamination. If this condition is satisfied or not depends on several system parameters, such as the number of UTs per DoF the channel offers (we denote by DoF the rank of the antenna correlation matrices which might be smaller than N), the number of antennas per BS, the signal-to-noise ratio (SNR), and the path loss. We further study how many antennas per UT are needed to achieve η% of the ultimate performance limit with infinitely many antennas and how many more antennas are needed with BF/MF to achieve RZF/MMSE performance. Our simulations suggest that in certain scenarios, RZF/MMSE can perform as well as BF/MF with almost one order of magnitude fewer antennas.

Manuscript received January 28, 2012; revised June 7, 2012; accepted September 4, 2012. Communicated by Thomas L. Marzetta, Guest editor. Parts of this work have been presented at the Allerton Conference on Communication, Control, and Computing, Urbana-Champaign, IL, US, Sep. 2011, and the IEEE International Conference on Communications, Ottawa, Canada, Jun. 2012.

J. Hoydis is with Bell Laboratories, Alcatel-Lucent, Lorenzstr. 10, 70435 Stuttgart, Germany (email: jakob.hoydis@alcatel-lucent.com).

S. ten Brink is with Bell Laboratories, Alcatel-Lucent, Lorenzstr. 10, 70435 Stuttgart, Germany (email: stephan.tenbrink@alcatel-lucent.com).

M. Debbah is with the Alcatel Lucent Chair on Flexible Radio, Supélec, 91192 Gif-sur-Yvette, France (email: merouane.debbah@supelec.fr).

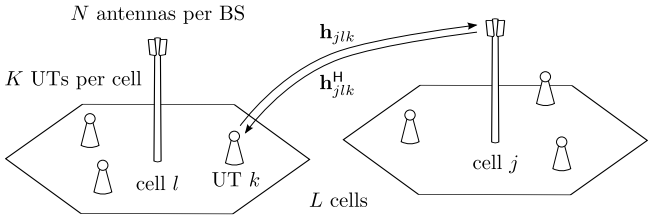

Fig. 1. In each of the L cells is one BS, equipped with N antennas, and K single-antenna UTs. We assume channel reciprocity, i.e., the downlink channel $\mathbf{h}_{jlk}^{\mathsf{H}}$ is the Hermitian transpose of the uplink channel $\mathbf{h}_{jlk}$.

The paper is organized as follows: In Section II, we describe the system model and derive achievable UL and DL rates with linear detectors and precoders. Section III contains our main technical results where we derive asymptotically tight approximations of these rates. In Section IV, we apply the asymptotic results to a simplified system model which leads to concise closed-form expressions of the achievable rates. This allows us to propose a precise definition of "massive" MIMO and to investigate if sub-optimal signal processing can be compensated for by the use of more antennas. We present some numerical results in Section V before we conclude the paper in Section VI. All proofs are deferred to the appendix.

Notations: Boldface lower and upper case symbols represent vectors and matrices, respectively ($\mathbf{I}_N$ is the size-N identity matrix). The trace, transpose, and Hermitian transpose operators are denoted by $\mathrm{tr}\,(\cdot)$, $(\cdot)^{\mathsf{T}}$, and $(\cdot)^{\mathsf{H}}$, respectively. The spectral norm of a matrix $\mathbf{A}$ is denoted by $\|\mathbf{A}\|$. We use $\mathcal{CN}(\mathbf{m},\mathbf{R})$ to denote the circular symmetric complex Gaussian distribution with mean $\mathbf{m}$ and covariance matrix $\mathbf{R}$. $\mathbb{E}[\cdot]$ denotes the expectation operator. $\lim_N$ stands for $\lim_{N\to\infty}$.

II. System model

Consider a multi-cellular system consisting of $L > 1$ cells with one BS and K UTs in each cell, as schematically shown in Fig. 1. The BSs are equipped with N antennas, the UTs have a single antenna. We assume that all BSs and UTs are perfectly synchronized and operate a TDD protocol with universal frequency reuse. We consider transmissions over flat-fading channels on a single frequency band or sub-carrier. Extensions to multiple sub-carriers, different numbers of antennas at the BSs, or different numbers of UTs in each cell are straightforward.

A. Uplink

The received base-band signal vector $\mathbf{y}_j^{\mathrm{ul}} \in \mathbb{C}^N$ at BS j at a given time instant reads

$$\mathbf{y}_j^{\mathrm{ul}} = \sqrt{\rho_{\mathrm{ul}}} \sum_{l=1}^{L} \mathbf{H}_{jl}\mathbf{x}_l^{\mathrm{ul}} + \mathbf{n}_j^{\mathrm{ul}} \tag{1}$$

where $\mathbf{H}_{jl} = [\mathbf{h}_{jl1} \cdots \mathbf{h}_{jlK}] \in \mathbb{C}^{N\times K}$, $\mathbf{h}_{jlk} \in \mathbb{C}^N$ is the channel from UT k in cell l to BS j, $\mathbf{x}_l^{\mathrm{ul}} = \left[x_{l1}^{\mathrm{ul}} \cdots x_{lK}^{\mathrm{ul}}\right]^{\mathsf{T}} \sim \mathcal{CN}(\mathbf{0},\mathbf{I}_K)$, with x_{lk}^{ul} the transmit signal of UT k in cell l, $\mathbf{n}_j^{\mathrm{ul}} \sim \mathcal{CN}(\mathbf{0},\mathbf{I}_N)$ is a noise vector, and $\rho_{\mathrm{ul}} > 0$ denotes the uplink SNR. We model the channel vectors $\mathbf{h}_{jlk}$ as

$$\mathbf{h}_{jlk} = \tilde{\mathbf{R}}_{jlk}\mathbf{v}_{jlk} \tag{2}$$

where $\mathbf{R}_{jlk} \triangleq \tilde{\mathbf{R}}_{jlk}\tilde{\mathbf{R}}_{jlk}^{\mathsf{H}} \in \mathbb{C}^{N\times N}$ are deterministic and $\mathbf{v}_{jlk} \sim \mathcal{CN}(\mathbf{0},\mathbf{I}_N)$ are independent fast-fading channel vectors. Our channel model is very versatile as it allows us to assign a different antenna correlation to each channel vector. This is especially important for large antenna arrays with a significant amount of antenna correlation due to either insufficient antenna spacing or a lack of scattering. The channel model is also valid for distributed antenna systems since we can assign a different path loss to each antenna. Moreover, (2) can represent a physical channel model with a fixed number of dimensions or angular bins P as in [8], by letting $\tilde{\mathbf{R}}_{jlk} = \sqrt{\ell_{jlk}}\,[\mathbf{A}\ \mathbf{0}_{N\times N-P}]$, where $\mathbf{A} \in \mathbb{C}^{N\times P}$, $\mathbf{0}_{N\times N-P}$ is the $N \times (N-P)$ zero matrix, and ℓ_{jlk} denotes the inverse path loss from UT k in cell l to BS j.

B. Downlink

The received signal $y_{jm}^{\mathrm{dl}} \in \mathbb{C}$ of the mth UT in the jth cell is given as

$$y_{jm}^{\mathrm{dl}} = \sqrt{\rho_{\mathrm{dl}}} \sum_{l=1}^{L} \mathbf{h}_{ljm}^{\mathsf{H}}\mathbf{s}_l + n_{jm}^{\mathrm{dl}} \tag{3}$$

where $\mathbf{s}_l \in \mathbb{C}^N$ is the transmit vector of BS l, $n_{jm}^{\mathrm{dl}} \sim \mathcal{CN}(0,1)$ is receiver noise, and $\rho_{\mathrm{dl}} > 0$ denotes the downlink SNR. We assume channel reciprocity, i.e., the downlink channel $\mathbf{h}_{ljm}^{\mathsf{H}}$ is the Hermitian transpose of the uplink channel $\mathbf{h}_{ljm}$. The transmit vector $\mathbf{s}_l$ is given as

$$\mathbf{s}_l = \sqrt{\lambda_l} \sum_{k=1}^{K} \mathbf{w}_{lk}x_{lk}^{\mathrm{dl}} = \sqrt{\lambda_l}\mathbf{W}_l\mathbf{x}_l^{\mathrm{dl}} \tag{4}$$

where $\mathbf{W}_l = [\mathbf{w}_{l1} \cdots \mathbf{w}_{lK}] \in \mathbb{C}^{N\times K}$ is a precoding matrix and $\mathbf{x}_l = \left[x_{l1}^{\mathrm{dl}} \cdots x_{lK}^{\mathrm{dl}}\right]^{\mathsf{T}} \in \mathbb{C}^K \sim \mathcal{CN}(\mathbf{0},\mathbf{I}_K)$ contains the data symbols for the K UTs in cell l. The parameter λ_l normalizes the average transmit power per UT of BS l to $\mathbb{E}\left[\frac{\rho_{\mathrm{dl}}}{K}\mathbf{s}_l^{\mathsf{H}}\mathbf{s}_l\right] = \rho_{\mathrm{dl}}$, i.e.,

$$\lambda_l = \frac{1}{\mathbb{E}\left[\frac{1}{K}\mathrm{tr}\,\mathbf{W}_l\mathbf{W}_l^{\mathsf{H}}\right]}. \tag{5}$$

C. Channel estimation

During a dedicated uplink training phase, the UTs in each cell transmit mutually orthogonal pilot sequences which allow the BSs to compute estimates $\hat{\mathbf{H}}_{jj}$ of their local channels $\mathbf{H}_{jj}$. The same set of orthogonal pilot sequences is reused in every cell so that the channel estimate is corrupted by pilot contamination from adjacent cells [1]. After correlating the received training signal with the pilot sequence of UT k, the jth BS estimates the channel vector $\mathbf{h}_{jjk}$ based on the

observation $\mathbf{y}_{jk}^{\text{tr}} \in \mathbb{C}^N$, given as[1]

$$\mathbf{y}_{jk}^{\text{tr}} = \mathbf{h}_{jjk} + \sum_{l \neq j} \mathbf{h}_{jlk} + \frac{1}{\sqrt{\rho_{\text{tr}}}} \mathbf{n}_{jk}^{\text{tr}} \tag{6}$$

where $\mathbf{n}_{jk}^{\text{tr}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$ and $\rho_{\text{tr}} > 0$ is the effective training SNR. In general, ρ_{tr} depends on the pilot transmit power and the length of the pilot sequences. Here, we assume ρ_{tr} to be a given parameter. The MMSE estimate $\hat{\mathbf{h}}_{jjk}$ of $\mathbf{h}_{jjk}$ is given as [13]

$$\begin{aligned} \hat{\mathbf{h}}_{jjk} &= \mathbf{R}_{jjk} \mathbf{Q}_{jk} \mathbf{y}_{jk}^{\text{tr}} \\ &= \mathbf{R}_{jjk} \mathbf{Q}_{jk} \left(\sum_{l} \mathbf{h}_{jlk} + \frac{1}{\sqrt{\rho_{\text{tr}}}} \mathbf{n}_{jk}^{\text{tr}} \right) \end{aligned} \tag{7}$$

which can be shown to be distributed as $\hat{\mathbf{h}}_{jjk} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Phi}_{jjk})$, where we define

$$\mathbf{\Phi}_{jlk} = \mathbf{R}_{jjk} \mathbf{Q}_{jk} \mathbf{R}_{jlk} \quad , \quad \forall j, l, k \tag{8}$$

$$\mathbf{Q}_{jk} = \left(\frac{1}{\rho_{\text{tr}}} \mathbf{I}_N + \sum_{l} \mathbf{R}_{jlk} \right)^{-1} , \quad \forall j, k. \tag{9}$$

Invoking the orthogonality property of the MMSE estimate [13], we can decompose the channel $\mathbf{h}_{jjk}$ as $\mathbf{h}_{jjk} = \hat{\mathbf{h}}_{jjk} + \tilde{\mathbf{h}}_{jjk}$, where $\tilde{\mathbf{h}}_{jjk} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_{jjk} - \mathbf{\Phi}_{jjk})$ is the uncorrelated estimation error (which is also statistically independent of $\hat{\mathbf{h}}_{jjk}$ due to the joint Gaussianity of both vectors).

D. Achievable uplink rates with linear detection

We consider linear single-user detection, where the jth BS estimates the symbol x_{jm}^{ul} of UT m in its cell by computing the inner product between the received vector $\mathbf{y}_j^{\text{ul}}$ and a linear filter $\mathbf{r}_{jm} \in \mathbb{C}^N$. Two particular filters are of practical interest, namely the matched filter $\mathbf{r}_{jm}^{\text{MF}}$ and the MMSE detector $\mathbf{r}_{jm}^{\text{MMSE}}$, which we define respectively as

$$\mathbf{r}_{jm}^{\text{MF}} = \hat{\mathbf{h}}_{jjm} \tag{10}$$

$$\mathbf{r}_{jm}^{\text{MMSE}} = \left(\hat{\mathbf{H}}_{jj} \hat{\mathbf{H}}_{jj}^{\mathsf{H}} + \mathbf{Z}_j^{\text{ul}} + N \varphi_j^{\text{ul}} \mathbf{I}_N \right)^{-1} \hat{\mathbf{h}}_{jjm} \tag{11}$$

where $\varphi_j^{\text{ul}} > 0$ and $\mathbf{Z}_j^{\text{ul}} \in \mathbb{C}^{N \times N}$ is an arbitrary Hermitian nonnegative definite matrix. This formulation of $\mathbf{r}_{jm}^{\text{MMSE}}$ allows us to treat φ_j^{ul} and $\mathbf{Z}_j^{\text{ul}}$ as design parameters which could be optimized. One could choose for example $\varphi_j^{\text{ul}} = \frac{1}{\rho_{\text{ul}} N}$ and $\mathbf{Z}_j^{\text{ul}}$ to be the covariance matrix of the intercell interference and the channel estimation errors, i.e.,

$$\begin{aligned} \mathbf{Z}_j^{\text{ul}} &= \mathbb{E}\left[\tilde{\mathbf{H}}_{jj} \tilde{\mathbf{H}}_{jj}^{\mathsf{H}} + \sum_{l \neq j} \mathbf{H}_{jl} \mathbf{H}_{jl}^{\mathsf{H}} \right] \\ &= \sum_{k} (\mathbf{R}_{jjk} - \mathbf{\Phi}_{jjk}) + \sum_{l \neq j} \sum_{k} \mathbf{R}_{jlk}. \end{aligned} \tag{12}$$

[1]For an integer variable s taking values in a set $\mathcal{S}$, we use $\sum_s$ to denote the summation over all $s \in \mathcal{S}$ and $\sum_{s \neq j}$ to denote the summation over all $s \in \mathcal{S} \setminus \{j\}$. Similarly, let s' be another integer variable taking values in the set $\mathcal{S}'$, we denote by $\sum_{(s,s') \neq (j,j')}$ the summation over all tuples $(s, s') \in \mathcal{S} \times \mathcal{S}' \setminus \{(j, j')\}$.

Using a standard bound based on the worst-case uncorrelated additive noise [14] yields the ergodic achievable uplink rate R_{jm}^{ul} of UT m in cell j:

$$R_{jm}^{\text{ul}} = \mathbb{E}\left[\log_2 \left(1 + \gamma_{jm}^{\text{ul}} \right) \right] \tag{13}$$

where the associated signal-to-interference-plus-noise ratio (SINR) γ_{jm}^{ul} is given by (14) on the top of the next page and where we have used $\mathbb{E}[\cdot|\cdot]$ to denote the conditional expectation operator. We will denote by γ_{jm}^{MF} and $\gamma_{jm}^{\text{MMSE}}$ the SINR with MF and MMSE detection, respectively.

E. Achievable downlink rates with linear precoding

Since the UTs do not have any channel estimate, we provide an ergodic achievable rate based on the techniques developed in [15]. To this end, we decompose the received signal y_{jm}^{dl} as

$$\begin{aligned} y_{jm}^{\text{dl}} = \; & \sqrt{\rho_{\text{dl}} \lambda_j} \mathbb{E}\left[\mathbf{h}_{jjm}^{\mathsf{H}} \mathbf{w}_{jm} \right] x_{jm}^{\text{dl}} \\ & + \sqrt{\rho_{\text{dl}} \lambda_j} \left(\mathbf{h}_{jjm}^{\mathsf{H}} \mathbf{w}_{jm} - \mathbb{E}\left[\mathbf{h}_{jjm}^{\mathsf{H}} \mathbf{w}_{jm} \right] \right) x_{jm}^{\text{dl}} \\ & + \sum_{(l,k) \neq (j,m)} \sqrt{\rho_{\text{dl}} \lambda_l} \mathbf{h}_{ljm}^{\mathsf{H}} \mathbf{w}_{lk} x_{lk}^{\text{dl}} + n_{jm}^{\text{dl}} \end{aligned} \tag{16}$$

and assume that the average effective channels $\sqrt{\lambda_j} \mathbb{E}\left[\mathbf{h}_{jjm}^{\mathsf{H}} \mathbf{w}_{jm} \right]$ can be perfectly learned at the UTs. Thus, an ergodic achievable rate R_{jm}^{dl} of UT m in cell j is given as [15, Theorem 1]

$$R_{jm}^{\text{dl}} = \log_2 \left(1 + \gamma_{jm}^{\text{dl}} \right) \tag{17}$$

where the associated SINR γ_{jm}^{dl} is given by (15) on top of the next page.[2]

We consider two different linear precoders $\mathbf{W}_j$ of practical interest, namely eigenbeamforming (BF) $\mathbf{W}_j^{\text{BF}}$ and regularized zero-forcing (RZF) $\mathbf{W}_j^{\text{RZF}}$, which we define respectively as

$$\mathbf{W}_j^{\text{BF}} = \hat{\mathbf{H}}_{jj} \tag{18}$$

$$\mathbf{W}_j^{\text{RZF}} = \left(\hat{\mathbf{H}}_{jj} \hat{\mathbf{H}}_{jj}^{\mathsf{H}} + \mathbf{Z}_j^{\text{dl}} + N \varphi_j^{\text{dl}} \mathbf{I}_N \right)^{-1} \hat{\mathbf{H}}_{jj} \tag{19}$$

where $\varphi_j^{\text{dl}} > 0$ is a regularization parameter and $\mathbf{Z}_j^{\text{dl}} \in \mathbb{C}^{N \times N}$ is an arbitrary Hermitian nonnegative definite matrix. As the choice of $\mathbf{Z}_j^{\text{dl}}$ and φ_j^{dl} is arbitrary, they could be further optimized (see, e.g., [15, Theorem 6]). This is outside the scope of this paper and left to future work. We will denote by γ_{jm}^{BF} and γ_{jm}^{RZF} the SINR with BF and RZF, respectively.

Remark 2.1: Under a block-fading channel model with coherence time T, one could account for the rate loss due to channel training by considering the net ergodic achievable rates $\kappa(1 - \tau/T) R_{jm}^{\text{ul}}$ and $(1 - \kappa)(1 - \tau/T) R_{jm}^{\text{dl}}$ for a given training length $\tau \in [K, T]$ and some $\kappa \in [0, 1]$ which determines the fraction of the remaining time used for uplink transmissions.

[2]We denote by $\text{var}[x] \triangleq \mathbb{E}[(x - \mathbb{E}[x])(x - \mathbb{E}[x])^{\mathsf{H}}]$ for some random variable x.

$$\gamma_{jm}^{\text{ul}} = \frac{\left|\mathbf{r}_{jm}^{\mathsf{H}}\hat{\mathbf{h}}_{jjm}\right|^2}{\mathbb{E}\left[\mathbf{r}_{jm}^{\mathsf{H}}\left(\frac{1}{\rho_{\text{ul}}}\mathbf{I}_N + \tilde{\mathbf{h}}_{jjm}\tilde{\mathbf{h}}_{jjm}^{\mathsf{H}} - \mathbf{h}_{jjm}\mathbf{h}_{jjm}^{\mathsf{H}} + \sum_l \mathbf{H}_{jl}\mathbf{H}_{jl}^{\mathsf{H}}\right)\mathbf{r}_{jm}\middle|\hat{\mathbf{H}}_{jj}\right]} \tag{14}$$

$$\gamma_{jm}^{\text{dl}} = \frac{\lambda_j\left|\mathbb{E}\left[\mathbf{h}_{jjm}^{\mathsf{H}}\mathbf{w}_{jm}\right]\right|^2}{\frac{1}{\rho_{\text{dl}}} + \lambda_j\text{var}\left[\mathbf{h}_{jjm}^{\mathsf{H}}\mathbf{w}_{jm}\right] + \sum_{(l,k)\neq(j,m)}\lambda_l\mathbb{E}\left[\left|\mathbf{h}_{ljm}^{\mathsf{H}}\mathbf{w}_{lk}\right|^2\right]} \tag{15}$$

III. Asymptotic analysis

As the ergodic achievable rates R_{jm}^{ul} and R_{jm}^{dl} with both types of detectors and precoders are difficult to compute for finite system dimensions, we consider the large system limit, where N and K grow infinitely large while keeping a finite ratio K/N. This is in contrast to [1] where the authors assume that the number of UTs remains fixed while the number of antennas grows without bound. We will retrieve the results of [1] as a special case. In the following, the notation "$N \to \infty$" will refer to $K, N \to \infty$ such that $\limsup_N K/N < \infty$. From now on, all vectors and matrices must be understood as sequences of vectors and matrices of growing dimensions. For the sake of simplicity, their dependence on N and K is not explicitly shown. The large system limit implicitly assumes that the coherence time of the channel scales linearly with K (to allow for orthogonal pilot sequences of the UTs in a cell). However, as we use the asymptotic analysis only as a tool to provide tight approximations for finite N, K, this does not pose any problem.[3] In a realistic deployment, one could expect BSs equipped with several hundred antennas serving each tens of UTs simultaneously [1].

In what follows, we will derive deterministic approximations $\bar{\gamma}_{jm}^{\text{ul}}$ ($\bar{\gamma}_{jm}^{\text{dl}}$) of the SINR γ_{jm}^{ul} (γ_{jm}^{dl}) with the MF and the MMSE detector (BF and RZF precoder), respectively, such that

$$\gamma_{jm}^{\text{ul}} - \bar{\gamma}_{jm}^{\text{ul}} \xrightarrow[N\to\infty]{\text{a.s.}} 0, \qquad \gamma_{jm}^{\text{dl}} - \bar{\gamma}_{jm}^{\text{dl}} \xrightarrow[N\to\infty]{} 0 \tag{20}$$

where "$\xrightarrow[N\to\infty]{\text{a.s.}}$" denotes almost sure convergence. One can then show by the dominated convergence [18] and the continuous mapping theorem [19], respectively, that (20) implies that

$$\begin{aligned} R_{jm}^{\text{ul}} - \log_2\left(1+\bar{\gamma}_{jm}^{\text{ul}}\right) &\xrightarrow[N\to\infty]{} 0 \\ R_{jm}^{\text{dl}} - \log_2\left(1+\bar{\gamma}_{jm}^{\text{dl}}\right) &\xrightarrow[N\to\infty]{} 0. \end{aligned} \tag{21}$$

These results must be understood in the way that, for each given set of system parameters N and K, we provide approximations of the SINR and the associated rates which become increasingly tight as N and K grow. We will show later by simulations that these approximations are very accurate for realistic system dimensions. As we make limiting considerations, we assume that the following conditions hold:

A 1: $\limsup_N \|\mathbf{R}_{jlk}\| < \infty \ \forall j,l,k$

A 2: $\liminf_N \frac{1}{N}\text{tr}\,\mathbf{R}_{jlk} > 0 \ \forall j,l,k$

A 3: $\limsup_N \|\frac{1}{N}\mathbf{Z}_j^{\text{ul}}\| < \infty$, $\limsup_N \|\frac{1}{N}\mathbf{Z}_j^{\text{dl}}\| < \infty \ \forall j$

[3]Note that similar assumptions have been made in [16], [17].

Before we continue, we recall two related results of large random matrix theory which will be required for the asymptotic performance analysis of the MMSE detector and the RZF precoder.

Theorem 1 ([20, Theorem 1]): Let $\mathbf{D} \in \mathbb{C}^{N\times N}$ and $\mathbf{S} \in \mathbb{C}^{N\times N}$ be Hermitian nonnegative definite and let $\mathbf{H} \in \mathbb{C}^{N\times K}$ be random with independent column vectors $\mathbf{h}_k \sim \mathcal{CN}\left(\mathbf{0}, \frac{1}{N}\mathbf{R}_k\right)$. Assume that $\mathbf{D}$ and the matrices $\mathbf{R}_k$, $k = 1,\ldots,K$, have uniformly bounded spectral norms (with respect to N). Then, for any $\rho > 0$,

$$\frac{1}{N}\text{tr}\,\mathbf{D}\left(\mathbf{H}\mathbf{H}^{\mathsf{H}} + \mathbf{S} + \rho\mathbf{I}_N\right)^{-1} - \frac{1}{N}\text{tr}\,\mathbf{D}\mathbf{T}(\rho) \xrightarrow[N\to\infty]{\text{a.s.}} 0$$

where $\mathbf{T}(\rho) \in \mathbb{C}^{N\times N}$ is defined as

$$\mathbf{T}(\rho) = \left(\frac{1}{N}\sum_{k=1}^{K}\frac{\mathbf{R}_k}{1+\delta_k(\rho)} + \mathbf{S} + \rho\mathbf{I}_N\right)^{-1}$$

and the elements of $\boldsymbol{\delta}(\rho) \triangleq [\delta_1(\rho)\cdots\delta_K(\rho)]^{\mathsf{T}}$ are defined as $\delta_k(\rho) = \lim_{t\to\infty}\delta_k^{(t)}(\rho)$, where for $t = 1,2,\ldots$

$$\delta_k^{(t)}(\rho) = \frac{1}{N}\text{tr}\,\mathbf{R}_k\left(\frac{1}{N}\sum_{j=1}^{K}\frac{\mathbf{R}_j}{1+\delta_j^{(t-1)}(\rho)} + \mathbf{S} + \rho\mathbf{I}_N\right)^{-1}$$

with initial values $\delta_k^{(0)}(\rho) = 1/\rho$ for all k.

Remark 3.1: The fixed-point algorithm in Theorem 1 to compute the quantities $\delta_k(\rho)$ can be efficiently numerically solved and is proved to converge. In some cases, closed-form solutions for $\boldsymbol{\delta}(\rho)$ exists. An example will be shown later in Corollary 3.

Theorem 2 ([21], see also [20]): Let $\boldsymbol{\Theta} \in \mathbb{C}^{N\times N}$ be Hermitian nonnegative definite with uniformly bounded spectral norm (with respect to N). Under the conditions of Theorem 1,

$$\begin{aligned}\frac{1}{N}\text{tr}\,\mathbf{D}\left(\mathbf{H}\mathbf{H}^{\mathsf{H}} + \mathbf{S} + \rho\mathbf{I}_N\right)^{-1}\boldsymbol{\Theta}\left(\mathbf{H}\mathbf{H}^{\mathsf{H}} + \mathbf{S} + \rho\mathbf{I}_N\right)^{-1} \\ - \frac{1}{N}\text{tr}\,\mathbf{D}\mathbf{T}'(\rho) \xrightarrow[N\to\infty]{\text{a.s.}} 0\end{aligned} \tag{22}$$

where $\mathbf{T}'(\rho) \in \mathbb{C}^{N\times N}$ is defined as

$$\mathbf{T}'(\rho) = \mathbf{T}(\rho)\boldsymbol{\Theta}\mathbf{T}(\rho) + \mathbf{T}(\rho)\frac{1}{N}\sum_{k=1}^{K}\frac{\mathbf{R}_k\delta_k'(\rho)}{(1+\delta_k(\rho))^2}\mathbf{T}(\rho)$$

$\mathbf{T}(\rho)$ and $\boldsymbol{\delta}(\rho)$ are given by Theorem 1, and $\boldsymbol{\delta}'(\rho) = [\delta'_1(\rho) \cdots \delta'_K(\rho)]^{\mathsf{T}}$ is calculated as

$$\boldsymbol{\delta}'(\rho) = \left(\mathbf{I}_K - \mathbf{J}(\rho)\right)^{-1} \mathbf{v}(\rho)$$

where $\mathbf{J}(\rho) \in \mathbb{C}^{K \times K}$ and $\mathbf{v}(\rho) \in \mathbb{C}^K$ are defined as

$$[\mathbf{J}(\rho)]_{kl} = \frac{\frac{1}{N}\mathrm{tr}\,\mathbf{R}_k \mathbf{T}(\rho)\mathbf{R}_l \mathbf{T}(\rho)}{N\left(1+\delta_l(\rho)\right)^2}, \quad 1 \le k,l \le K$$
$$[\mathbf{v}(\rho)]_k = \frac{1}{N}\mathrm{tr}\,\mathbf{R}_k \mathbf{T}(\rho)\boldsymbol{\Theta}\mathbf{T}(\rho), \quad 1 \le k \le K.$$

Next, we provide SINR approximations in the sense of (20) for MF and MMSE detection in the UL and for BF and RZF precoding in the DL. These are our main results. Due to the similarity of the SINR expressions for the UL and DL, we only provide the proofs for BF and RZF in the appendix. The proofs for MF and MMSE are very similar and omitted due to space constraints.

Theorem 3 (Matched filter): Assume that **A 1–3** hold. Then, $\gamma_{jm}^{\mathrm{MF}} - \bar{\gamma}_{jm}^{\mathrm{MF}} \xrightarrow[N\to\infty]{\text{a.s.}} 0$, where $\bar{\gamma}_{jm}^{\mathrm{MF}}$ is given in (23).

Theorem 4 (Eigenbeamforming): Assume that **A 1–3** hold. Then, $\gamma_{jm}^{\mathrm{BF}} - \bar{\gamma}_{jm}^{\mathrm{BF}} \xrightarrow[N\to\infty]{} 0$, where $\bar{\gamma}_{jm}^{\mathrm{BF}}$ is given in (24) with $\bar{\lambda}_j = \left(\frac{1}{K}\sum_{k=1}^{K}\frac{1}{N}\mathrm{tr}\,\boldsymbol{\Phi}_{jjk}\right)^{-1} \ \forall j$.

Theorem 5 (MMSE detector): Assume that **A 1–3** hold. Then, $\gamma_{jm}^{\mathrm{MMSE}} - \bar{\gamma}_{jm}^{\mathrm{MMSE}} \xrightarrow[N\to\infty]{\text{a.s.}} 0$, where $\bar{\gamma}_{jm}^{\mathrm{MMSE}}$ is given in (25) with

$$\mu_{jlkm} = \frac{1}{N}\mathrm{tr}\,\mathbf{R}_{jlk}\mathbf{T}'_{jm} - \frac{2\mathrm{Re}\left(\vartheta^*_{jlk}\vartheta'_{jlkm}\right)(1+\delta_{jk}) - |\vartheta_{jlk}|^2\delta'_{jkm}}{(1+\delta_{jk})^2}$$
$$\vartheta_{jlk} = \frac{1}{N}\mathrm{tr}\,\boldsymbol{\Phi}_{jlk}\mathbf{T}_j$$
$$\vartheta'_{jlkm} = \frac{1}{N}\mathrm{tr}\,\boldsymbol{\Phi}_{jlk}\mathbf{T}'_{jm}$$

where

(i) $\mathbf{T}_j = \mathbf{T}(\varphi_j^{\mathrm{ul}})$ and $\boldsymbol{\delta}_j = [\delta_{j1}\cdots\delta_{jK}]^{\mathsf{T}} = \boldsymbol{\delta}(\varphi_j^{\mathrm{ul}})$ are given by Theorem 1 for $\mathbf{S} = \mathbf{Z}_j^{\mathrm{ul}}/N$, $\mathbf{D} = \mathbf{I}_N$, and $\mathbf{R}_k = \boldsymbol{\Phi}_{jjk}\ \forall k$,
(ii) $\bar{\mathbf{T}}'_j = \mathbf{T}'(\varphi_j^{\mathrm{ul}})$ is given by Theorem 2 for $\mathbf{S} = \mathbf{Z}_j^{\mathrm{ul}}/N$, $\boldsymbol{\Theta} = \mathbf{I}_N$, $\mathbf{D} = \mathbf{I}_N$, and $\mathbf{R}_k = \boldsymbol{\Phi}_{jjk}\ \forall k$,
(iii) $\mathbf{T}'_{jm} = \mathbf{T}'(\varphi_j^{\mathrm{ul}})$ and $\boldsymbol{\delta}'_{jm} = \left[\delta'_{j1m}\cdots\delta'_{jKm}\right]^{\mathsf{T}} = \boldsymbol{\delta}'(\varphi_j^{\mathrm{ul}})$ are given by Theorem 2 for $\mathbf{S} = \mathbf{Z}_j^{\mathrm{ul}}/N$, $\boldsymbol{\Theta} = \boldsymbol{\Phi}_{jjm}$, $\mathbf{D} = \mathbf{I}_N$, and $\mathbf{R}_k = \boldsymbol{\Phi}_{jjk}\ \forall k$.

Theorem 6 (Regularized Zero-Forcing): Assume that **A 1–3** hold. Then, $\gamma_{jm}^{\mathrm{RZF}} - \bar{\gamma}_{jm}^{\mathrm{RZF}} \xrightarrow[N\to\infty]{} 0$, where $\bar{\gamma}_{jm}^{\mathrm{RZF}}$ is given in (26) with

$$\mu_{ljmk} = \frac{1}{N}\mathrm{tr}\,\mathbf{R}_{ljm}\mathbf{T}'_{lk} - \frac{2\mathrm{Re}\left(\vartheta^*_{ljm}\vartheta'_{ljmk}\right)(1+\delta_{lm}) - |\vartheta_{ljm}|^2\delta'_{lmk}}{(1+\delta_{lm})^2}$$
$$\vartheta_{ljm} = \frac{1}{N}\mathrm{tr}\,\boldsymbol{\Phi}_{ljm}\mathbf{T}_l$$
$$\vartheta'_{ljmk} = \frac{1}{N}\mathrm{tr}\,\boldsymbol{\Phi}_{ljm}\mathbf{T}'_{lk}$$
$$\bar{\lambda}_l = \frac{K}{N}\left(\frac{1}{N}\mathrm{tr}\,\mathbf{T}_l - \frac{1}{N}\mathrm{tr}\left(\frac{\mathbf{Z}_l^{\mathrm{dl}}}{N} + \varphi_l^{\mathrm{dl}}\mathbf{I}_N\right)\bar{\mathbf{T}}'_l\right)^{-1}$$

where

(i) $\mathbf{T}_l = \mathbf{T}(\varphi_l^{\mathrm{dl}})$ and $\boldsymbol{\delta}_l = [\delta_{l1}\cdots\delta_{lK}]^{\mathsf{T}} = \boldsymbol{\delta}(\varphi_l^{\mathrm{dl}})$ are given by Theorem 1 for $\mathbf{S} = \mathbf{Z}_l^{\mathrm{dl}}/N$, $\mathbf{D} = \mathbf{I}_N$, and $\mathbf{R}_k = \boldsymbol{\Phi}_{llk}\ \forall k$,
(ii) $\bar{\mathbf{T}}'_l = \mathbf{T}'(\varphi_l^{\mathrm{dl}})$ is given by Theorem 2 for $\mathbf{S} = \mathbf{Z}_l^{\mathrm{dl}}/N$, $\boldsymbol{\Theta} = \mathbf{I}_N$, $\mathbf{D} = \mathbf{I}_N$, and $\mathbf{R}_k = \boldsymbol{\Phi}_{llk}\ \forall k$,
(iii) $\mathbf{T}'_{lk} = \mathbf{T}'(\varphi_l^{\mathrm{dl}})$ and $\boldsymbol{\delta}'_{lk} = [\delta'_{l1k}\cdots\delta'_{lKk}]^{\mathsf{T}} = \boldsymbol{\delta}'(\varphi_l^{\mathrm{dl}})$ are given by Theorem 2 for $\mathbf{S} = \mathbf{Z}_l^{\mathrm{dl}}/N$, $\boldsymbol{\Theta} = \boldsymbol{\Phi}_{llk}$, $\mathbf{D} = \mathbf{I}_N$, and $\mathbf{R}_k = \boldsymbol{\Phi}_{llk}\ \forall k$.

Remark 3.2: Observe the similarity between the results for MF and BF (MMSE and RZF). The main difference is that in the downlink, all transmit powers are multiplied by the power normalization factors $\bar{\lambda}_j$ and the indices j,l and k,m are swapped for the interference terms.

Remark 3.3: The expressions of $\bar{\gamma}_{jm}^{\mathrm{MMSE}}$ and $\bar{\gamma}_{jm}^{\mathrm{RZF}}$ can be greatly simplified under a less general channel model, e.g., no antenna correlation or Wyner-type models with the same path loss for all interfering UTs [22]. We provide later a special case for which $\bar{\gamma}_{jm}^{\mathrm{MMSE}}$ and $\bar{\gamma}_{jm}^{\mathrm{RZF}}$ are given in closed form.

Next, we consider the case when the number of antennas per BS is much larger than the number of UTs per cell, i.e., $N \gg K$.

Corollary 1: Let $N \to \infty$, such that $K/N \to 0$. Denote $\beta_{jlk} = \lim_N \frac{1}{N}\mathrm{tr}\,\boldsymbol{\Phi}_{jlk}$ whenever the limit exists, and define

$$\bar{\lambda}_j^{\infty,\mathrm{BF}} = \left(\frac{1}{K}\sum_{k=1}^{K}\beta_{jjk}\right)^{-1}$$

$$\bar{\lambda}_j^{\infty,\mathrm{RZF}} = \left(\frac{1}{K}\sum_{k=1}^{K}\frac{\beta_{jjk}}{\left(\varphi_j^{\mathrm{dl}}+\beta_{jjk}\right)^2}\right)^{-1}.$$

Then,

$$\bar{\gamma}_{jm}^{\mathrm{MF}} \to \frac{\beta_{jjm}^2}{\sum_{l\neq j}|\beta_{jlm}|^2}$$
$$\bar{\gamma}_{jm}^{\mathrm{BF}} \to \frac{\bar{\lambda}_j^{\infty,\mathrm{BF}}\beta_{jjm}^2}{\sum_{l\neq j}\bar{\lambda}_l^{\infty,\mathrm{BF}}|\beta_{ljm}|^2}$$
$$\bar{\gamma}_{jm}^{\mathrm{MMSE}} \to \frac{\beta_{jjm}^2}{\sum_{l\neq j}\left(\frac{\varphi_j^{\mathrm{ul}}}{\varphi_l^{\mathrm{ul}}}\right)^2|\beta_{jlm}|^2}$$
$$\bar{\gamma}_{jm}^{\mathrm{RZF}} \to \frac{\bar{\lambda}_j^{\infty,\mathrm{RZF}}\beta_{jjm}^2}{\sum_{l\neq j}\left(\frac{\varphi_l^{\mathrm{dl}}\varphi_j^{\mathrm{dl}}+\varphi_l^{\mathrm{dl}}\beta_{jjm}}{\varphi_l^{\mathrm{dl}}\varphi_j^{\mathrm{dl}}+\varphi_j^{\mathrm{dl}}\beta_{llm}}\right)^2\bar{\lambda}_l^{\infty,\mathrm{RZF}}|\beta_{ljm}|^2}.$$

$$\bar{\gamma}_{jm}^{\text{MF}} = \frac{\left(\frac{1}{N}\text{tr}\,\boldsymbol{\Phi}_{jjm}\right)^2}{\frac{1}{\rho_{\text{ul}}N}\frac{1}{N}\text{tr}\,\boldsymbol{\Phi}_{jjm} + \frac{1}{N}\sum_{l,k}\frac{1}{N}\text{tr}\,\mathbf{R}_{jlk}\boldsymbol{\Phi}_{jjm} + \sum_{l\neq j}\left|\frac{1}{N}\text{tr}\,\boldsymbol{\Phi}_{jlm}\right|^2} \tag{23}$$

$$\bar{\gamma}_{jm}^{\text{BF}} = \frac{\bar{\lambda}_j\left(\frac{1}{N}\text{tr}\,\boldsymbol{\Phi}_{jjm}\right)^2}{\frac{1}{\rho_{\text{dl}}N} + \frac{1}{N}\sum_{l,k}\bar{\lambda}_l\frac{1}{N}\text{tr}\,\mathbf{R}_{ljm}\boldsymbol{\Phi}_{llk} + \sum_{l\neq j}\bar{\lambda}_l\left|\frac{1}{N}\text{tr}\,\boldsymbol{\Phi}_{ljm}\right|^2} \tag{24}$$

$$\bar{\gamma}_{jm}^{\text{MMSE}} = \frac{\delta_{jm}^2}{\frac{1}{\rho_{\text{ul}}N}\frac{1}{N}\text{tr}\,\boldsymbol{\Phi}_{jjm}\bar{\mathbf{T}}_j' + \frac{1}{N}\sum_{l,k}\mu_{jlkm} + \sum_{l\neq j}|\vartheta_{jlm}|^2} \tag{25}$$

$$\bar{\gamma}_{jm}^{\text{RZF}} = \frac{\bar{\lambda}_j\delta_{jm}^2}{\frac{(1+\delta_{jm})^2}{\rho_{\text{dl}}N} + \frac{1}{N}\sum_{l,k}\bar{\lambda}_l\left(\frac{1+\delta_{jm}}{1+\delta_{lk}}\right)^2\mu_{ljmk} + \sum_{l\neq j}\bar{\lambda}_l\left(\frac{1+\delta_{jm}}{1+\delta_{lm}}\right)^2|\vartheta_{ljm}|^2} \tag{26}$$

Proof: Note that the first and the second term in the denominator of the asymptotic SINR expressions in Theorems 3–6 vanish as $N \to \infty$ while $K/N \to 0$. For the remaining terms, further note that $\mathbf{T}_j(\varphi) \to \varphi^{-1}$ and $\bar{\mathbf{T}}_j'(\varphi) \to \varphi^{-2}$. Lastly, for RZF, we can write $\bar{\lambda}_j$ equivalently as $\bar{\lambda}_j = \left(\frac{1}{K}\sum_k \frac{\frac{1}{N}\text{tr}\,\boldsymbol{\Phi}_{jjk}\mathbf{T}_j'}{(1+\frac{1}{N}\text{tr}\,\boldsymbol{\Phi}_{jjk}\mathbf{T}_j)^2}\right)^{-1}$. Replacing these quantities in the corresponding SINR expressions leads to the desired result. ∎

Remark 3.4: As already observed in [1, Eq. (13)], the performance of the MF and the MMSE detector coincide with an infinite number of BS-antennas per UT if $\varphi_j^{\text{ul}} = \varphi_l^{\text{ul}}\ \forall l$. However, even for $\bar{\lambda}_j^{\infty} = \bar{\lambda}_l^{\infty}$ and $\varphi_j^{\text{dl}} = \varphi_l^{\text{dl}}\ \forall l$, the SINR under RZF and BF are not necessarily identical. This is because the received interference power depends on the correlation matrices $\boldsymbol{\Phi}_{llm}$.

IV. On the massive MIMO effect

Let us now consider the simplified channel model

$$\mathbf{H}_{jj} = \sqrt{\frac{N}{P}}\mathbf{A}\mathbf{V}_{jj}, \quad \mathbf{H}_{jl} = \sqrt{\alpha\frac{N}{P}}\mathbf{A}\mathbf{V}_{jl}, \quad l \neq j \tag{27}$$

where $\mathbf{A} \in \mathbb{C}^{N\times P}$ is composed of $P \leq N$ columns of an arbitrary unitary $N \times N$ matrix, $\mathbf{V}_{jl} \in \mathbb{C}^{P\times K}$ are standard complex Gaussian matrices and $\alpha \in (0,1]$ is an intercell interference factor. Note that this is a special case of (2). Under this model, the total energy of the channel grows linearly with N and K, since $\mathbb{E}\left[\text{tr}\,\mathbf{H}_{jj}\mathbf{H}_{jj}^{\mathsf{H}}\right] = \frac{KN}{P}\text{tr}\,\mathbf{A}\mathbf{A}^{\mathsf{H}} = KN$. The motivation behind this channel model is twofold. First, we assume that the antenna aperture increases with each additional antenna element. Thus, the captured energy increases linearly with N. This is in contrast to existing works which assume that more and more antenna elements are packed into a fixed volume, see, e.g., [23]. An insufficiency of this channel model is that the captured energy grows without bounds as $N \to \infty$. However, we believe that linear energy gains can be achieved up to very large numbers of antennas if the size of the antenna array is scaled accordingly. For example, at a carrier frequency of 2.6 GHz (i.e., wavelength $\lambda \approx 12$ cm), a 16×16 antenna array with $\lambda/2$-spacing would occupy an area of roughly 1 m^2. Second, the number of DoF P offered by the channel does not need to be equal to N [8]. One could either assume P to be large but constant[4] or to scale with N, e.g., $P = cN$, where $c \in (0,1]$. In general, P depends on the amount of scattering in the channel and, therefore, on the radio environment.[5] Let us further assume that the transmit powers per UT in the uplink and downlink are equal, i.e., $\rho_{\text{ul}} = \rho_{\text{dl}} = \rho$, and that the matrices $\mathbf{Z}_j^{\text{dl}}$ and $\mathbf{Z}_j^{\text{ul}}$ used for precoding and detection are also equal and given by (12). Under these assumptions, the performance of MF and BF (MMSE and RZF) coincides and Theorems 3–6 can be given in closed form:

Corollary 2: For the channel model (27) and $\rho_{\text{ul}} = \rho_{\text{dl}} = \rho$, $\bar{\gamma}_{jm}^{\text{MF}}$ and $\bar{\gamma}_{jm}^{\text{BF}}\ \forall j,m$, are given as

$$\bar{\gamma}^{\text{MF}} = \bar{\gamma}^{\text{BF}} = \frac{1}{\frac{1}{\nu\rho N} + \frac{K}{P}\frac{\bar{L}}{\nu} + \alpha(\bar{L}-1)} \tag{28}$$

$$= \frac{1}{\underbrace{\frac{\bar{L}}{\rho N}}_{\text{noise}} + \underbrace{\frac{1}{\rho_{\text{tr}}}\left(\frac{P/N}{\rho N} + \frac{K}{N}\bar{L}\right)}_{\text{imperfect CSI}} + \underbrace{\frac{K}{P}\bar{L}^2}_{\text{interference}} + \underbrace{\alpha\left(\bar{L}-1\right)}_{\text{pilot contamination}}} \tag{29}$$

where $\bar{L} = 1 + \alpha(L-1)$ and $\nu = \frac{\rho_{\text{tr}}\frac{N}{P}}{1+\rho_{\text{tr}}\frac{N}{P}\bar{L}}$.

Corollary 3: For the channel model (27), $\rho_{\text{ul}} = \rho_{\text{dl}}$, $\varphi_j^{\text{ul}} = \varphi_j^{\text{dl}} = \varphi$, and $\mathbf{Z}_j^{\text{ul}} = \mathbf{Z}_j^{\text{dl}} = \mathbf{Z}_j = \sum_{l,k}\mathbf{R}_{jlk} - \sum_k \boldsymbol{\Phi}_{jjk}\ \forall j$, $\bar{\gamma}_{jm}^{\text{MMSE}}$ and $\bar{\gamma}_{jm}^{\text{RZF}}\ \forall j,m$, are given as

$$\bar{\gamma}^{\text{MMSE}} = \bar{\gamma}^{\text{RZF}} = \frac{1}{\frac{1}{\nu\rho N}X + \frac{K}{P}\frac{\bar{L}}{\nu}Y + \alpha(\bar{L}-1)} \tag{30}$$

where

$$Y = X + \frac{\nu(1+\alpha^2(L-1))(1-2Z)}{\bar{L}(Z^2 - K/P)}$$

$$\delta = \frac{1 - S + \sqrt{(1+S)^2 - 4K/P}}{2(S - K/P)}$$

[4]Note that if P is assumed to be fixed, the spectral norm of the matrix $\sqrt{N/P}\mathbf{A}$ grows without bound as $N \to \infty$. Thus, assumption **A 1** is violated and the asymptotic analysis in Section III is not valid anymore.

[5]See also [2], [10], [11] for a discussion of the issue of "favorable propagation conditions" which is closely related to the important connection between P and N.

$\bar{L} = 1 + \alpha(L-1)$, $\nu = \frac{\rho_{\text{tr}} \frac{N}{P}}{1+\rho_{\text{tr}} \frac{N}{P} \bar{L}}$, $X = \frac{Z^2}{Z^2 - \frac{K}{P}}$, $Z = \frac{1+\delta}{\delta}$, and $S = \frac{\varphi}{\nu} + \frac{K\bar{L}}{P\nu}$.

Sketch: Notice that $\mathbf{\Phi}_{jlk} = \max(\mathbb{1}\{l=j\}, \alpha)\frac{N}{P}\mathbf{A}\mathbf{A}^{\mathsf{H}}$, where $\mathbb{1}\{\cdot\}$ is the indicator function, and $\mathbf{Z}_j = K(\bar{L}-\nu)\frac{N}{P}\mathbf{A}\mathbf{A}^{\mathsf{H}}$ $\forall j, l, k$. Using these expressions, one can show after some straight-forward but tedious calculus that Theorems 1 and 2 can be given in closed form and that the SINR expressions in Theorems 3–6 can be greatly simplified. ■

One can make several observations from (28) and (30). First, the asymptotic SINR depends on the transmit SNR ρ only through the term $\frac{1}{\nu\rho N}$. Thus, the "effective SNR" ρN increases linearly with N. If the number of antennas is doubled, the transmit power can consequently be reduced by a factor two to achieve the same performance. However, if the transmit *and* the training power are reduced as N grows, this conclusions fails to hold, as can been seen from the term $\nu\rho N$. The product of transmit and training power must satisfy $\liminf_N \rho\rho_{\text{tr}} N > 0$ (if $\liminf_N P/N > 0$). Otherwise the SINR converges to zero as $N \to \infty$. As already observed in [4], if $\rho_{\text{tr}} = \rho$, the transmit power can be made only inversely proportional to $\sqrt{N}$. Second, the interference depends mainly on the ratio P/K (number of DoF per UT) and not directly on N. Thus, interference can only be reduced by the use of additional antennas if the environment provides sufficient scattering. Third, noise, channel estimation errors, and interference vanish if $N, P \to \infty$ at the same speed, while pilot contamination remains the only performance limitation:

$$\bar{\gamma}^{\text{MF}}, \bar{\gamma}^{\text{BF}}, \bar{\gamma}^{\text{MMSE}}, \bar{\gamma}^{\text{RZF}} \xrightarrow[N,P\to\infty,\ K/N\to 0]{} \gamma_\infty = \frac{1}{\alpha(\bar{L}-1)}. \tag{31}$$

We denote by R_∞ the ultimately achievable rate, defined as

$$R_\infty = \log_2(1+\gamma_\infty) = \log_2\left(1 + \frac{1}{\alpha(\bar{L}-1)}\right). \tag{32}$$

It is interesting that all precoders and detectors achieve the same asymptotic performance limit γ_∞. Note that without pilot contamination, i.e., for $L = 1$ or $\alpha = 0$, the SINR grows without bounds as $P, N \to \infty$. If P is fixed but large, the SINR saturates at a smaller value than γ_∞. In this case, adding antennas only improves the SNR but does not reduce the multiuser interference. Also MMSE/RZF have a superior performance than MF/BF as $N \to \infty$.

Before we proceed, let us verify the accuracy of the rate-approximations $\bar{R}^{\text{MF}} = \log(1+\bar{\gamma}^{\text{MF}})$ and $\bar{R}^{\text{MMSE}} = \log(1+\bar{\gamma}^{\text{MMMSE}})$ as given by Corollaries 2 and 3, respectively, for finite N and K. In Fig. 2, we depict the ergodic achievable rate R_{jm} (13) of an arbitrary UT with MF and MMSE detection as a function of the number of antennas N for $K = 10$ UTs, $L = 4$ cells, $\rho_{\text{tr}} = 6\,\text{dB}$, $\rho = 10\,\text{dB}$, $\varphi_l^{\text{ul}} = 1/(\rho N)$, and intercell interference factor $\alpha = 0.1$. We compare two different cases: $P = N$ and $P = N/3$. As expected, the performance in the latter scenario is worse due to stronger multiuser interference. Most importantly, our closed-form approximations are almost indistinguishable from the simulation results over the entire range of N. The results for the downlink with MF and RZF look similar and are omitted due to space limitations.

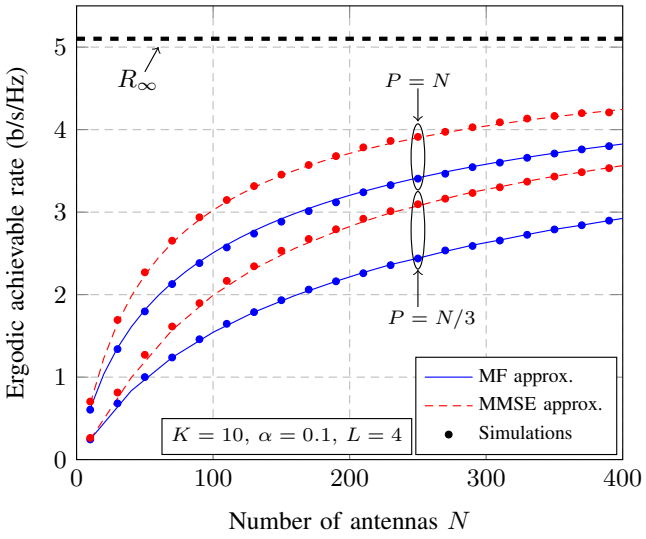

Fig. 2. Ergodic achievable rate with MF and MMSE detection versus number of antennas N for $P \in \{N, N/3\}$, $\rho_{\text{tr}} = 6\,\text{dB}$ and $\rho = 10\,\text{dB}$.

Based on our previous observations, it is justified to speak about a *massive MIMO effect* whenever the SINR γ_{jm} (in the UL or DL) is close to γ_∞, or in other words, whenever noise, channel estimation errors, and interference are small compared to pilot contamination. It becomes evident from (29) and (30) that the number of antennas needed to achieve this effect depends strongly on the system parameters P, K, L, α, ρ_{tr}, and ρ. In particular, there is no massive MIMO effect without pilot contamination since $\gamma_\infty \to \infty$. Thus, massive MIMO can be seen as a particular operating condition in multi-cellular systems where the performance is ultimately limited by pilot contamination and MF/BF achieve a performance close to this ultimate limit. To make this definition more precise, we say that we operate under massive MIMO conditions if, for some desired "massive MIMO efficiency" $\eta \in (0,1)$,

$$R = \log(1+\gamma) \ \geq \ \eta R_\infty \tag{33}$$

where γ is the SINR in the UL/DL with any detection/precoding scheme. This condition implies that we achieve at least the fraction η of the ultimate performance limit. If we assume that $\rho_{\text{tr}} \gg 1$, i.e., $\nu \approx \bar{L}^{-1}$, the expressions of $\bar{\gamma}^{\text{MF}}$, $\bar{\gamma}^{\text{BF}}$, $\bar{\gamma}^{\text{MMSE}}$, and $\bar{\gamma}^{\text{RZF}}$ in Corollaries 2 and 3 depend on P, K, ρ, and N only through the ratio $\frac{P}{K}$ and the effective SNR ρN. Thus, for a given set of parameters ρ, N, α, L, and φ, we can easily find the fraction $\frac{P}{K}$ necessary to satisfy (33).

Figs. 3 and 4 show the necessary DoF per UT $\frac{P}{K}$ for a given effective SNR ρN to achieve a spectral efficiency of ηR_∞ with either MF/BF (solid lines) or MMSE/RZF (dashed lines). We consider $L = 4$ cells, $\varphi = 1/(\rho N)$, and an intercell interference factor $\alpha = 0.3$ and $\alpha = 0.1$, respectively. The plots must be understood in the following way: Each curve corresponds to a particular value of η. In the region above each curve, the condition (33) is satisfied.

Let us first focus on Fig. 3 with $\alpha = 0.3$. For an effective SNR $\rho N = 20\,\text{dB}$ (e.g., $\rho = 0\,\text{dB}$ and $N = 100 = 20\,\text{dB}$),

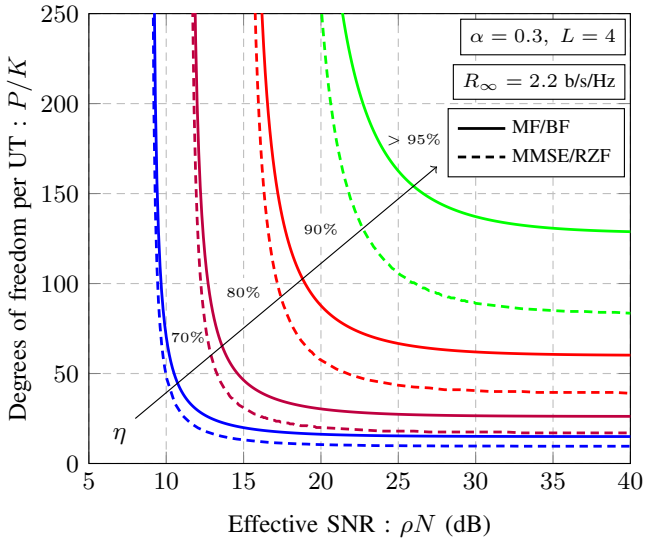

Fig. 3. Degrees of freedom per UT P/K necessary to achieve ηR_∞ versus effective SNR ρN for $L = 4$ and $\alpha = 0.3$.

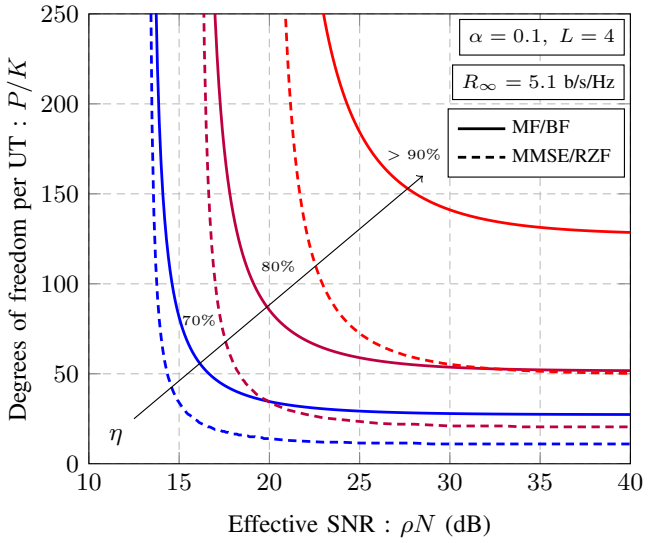

Fig. 4. Degrees of freedom per UT P/K necessary to achieve ηR_∞ versus effective SNR ρN for $L = 4$ and $\alpha = 0.1$.

we need about $P/K = 90$ DoF per UT with MF/BF to achieve 90 % of the ultimate performance R_∞, i.e., $0.9 \times 2.2 \approx 2$ b/s/Hz. If $P \approx N$, only a single UT could be served (Note that this is a simplifying example. Our analysis assumes $K \gg 1$.). However, if we had $N = 1000 = 30$ dB antennas, the transmit power ρ could be decreased by 10 dB and 10 UTs could be served with the same performance. At the same operating point, the MMSE/RZF requires only ~ 60 DoF per UT to achieve 90 % of the ultimate performance. Thus, the use of MMSE/RZF would allow us to increase the number of simultaneously served UTs by a factor $\frac{90}{60} = 1.5$. This example also demonstrates the importance of the relation between N and P. In particular, if P saturates for some N, adding more antennas increases the effective SNR but does not reduce the multiuser interference. Thus, the number of UTs which can be simultaneously supported depends significantly on the radio environment. We can further see that adding antennas shows diminishing returns. This is because the distances between the curves for different values of η grow exponentially fast. Remember that for $\eta = 1$, a ratio of $P/K = \infty$ would be needed. A last observation we can make is that the absolute difference between MF/BF and MMSE/RZF is marginal for small values of η but gets quickly pronounced as $\eta \to 1$.

Moving to Fig. 4 for $\alpha = 0.1$, we can see that for the same effective SNR $\rho N = 20$ dB and the same number of DoF per UT $P/K = 90$ as in the previous example, only 80 % of the ultimate performance are achieved by MF/BF. However, since the intercell interference is significantly smaller compared to the previous example, this corresponds to $0.9 \times 5.1 \approx 4.6$ b/s/Hz. Thus, although we operate further away from the ultimate performance limit, the resulting spectral efficiency is still higher. With MMSE/RZF, only 35 DoF per UT are necessary to achieve the same performance and, consequently, $90/35 \approx 2.5$ times more UTs could be simultaneously served. With decreasing intercell interference (and hence decreasing pilot contamination) the advantages of MMSE/RZF become more and more important.

V. Numerical results

Let us now validate the accuracy of Theorems 4 and 6 for finite N and K in a more realistic downlink scenario. Simulations for the uplink, i.e., Theorems 3 and 5, are omitted due to space constraints but provide similar results (see Remark 3.2). We consider a hexagonal system with $L = 7$ cells as shown in Fig. 5. The inner cell radius is normalized to one and we assume a distance-based path loss model with path loss exponent $\beta = 3.7$. To allow for reproducibility of our results, we distribute $K = 10$ UTs uniformly on a circle of radius $2/3$ around each BS and do not consider shadowing. We further assume a training SNR $\rho_{\text{tr}} = 6$ dB and transmit SNR $\rho_{\text{dl}} = 10$ dB. For RZF, we use a regularization factor $\varphi_j^{\text{dl}} = 1/\rho_{\text{dl}}$ and $\mathbf{Z}_j^{\text{dl}} = 0$. Average rates are then calculated for the UTs in the center cell.

First, we consider a simple channel model without antenna correlation, i.e., $\tilde{\mathbf{R}}_{jlk} = d_{jlk}^{-\beta/2}\mathbf{I}_N$, where d_{jlk} is the distance between BS j and the kth UTs in cell l (cf. (2)). For an unlimited number of antennas per UT, the precoding schemes lead respectively to the ultimate average rates 7.2 b/s/Hz (BF) and 7.08 b/s/Hz (RZF) (see Remark 3.4). In Fig. 6, we show the achievable rates under both precoding techniques and their approximations by Theorems 4 and 6 as a function of the number of antennas N. Both results match very well, even for small N. We can observe that RZF leads to significant performance gains over BF as it reduces multiuser interference. For $N = 400$, RZF achieves 82 % of the ultimate limit while BF achieves only 65 %.

Second, we consider a physical channel model with a fixed number of dimensions P as in [8]. For a uniform linear array, the matrices $\tilde{\mathbf{R}}_{jlk}$ are given as $\tilde{\mathbf{R}}_{jlk} = d_{jlk}^{-\beta/2}\,[\mathbf{A}\ \mathbf{0}_{N\times N-P}]$, where $\mathbf{A} = [\mathbf{a}(\phi_1)\cdots\mathbf{a}(\phi_P)] \in \mathbb{C}^{N\times P}$ is composed of the

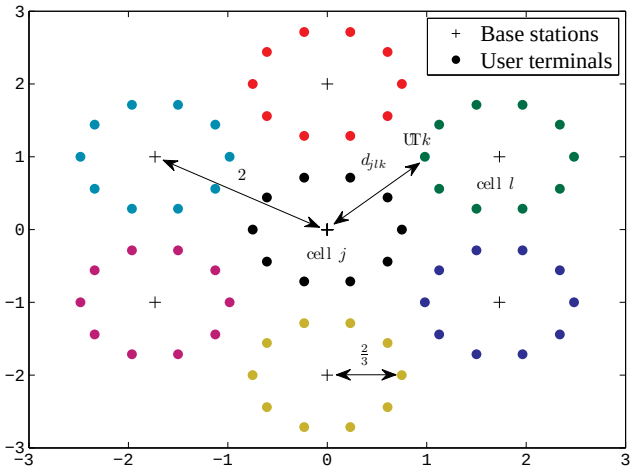

Fig. 5. 7–cell hexagonal system layout. The distance between two adjacent cells is normalized to 2. There are $K = 10$ UTs uniformly distributed on a circle of radius $2/3$ around each BS.

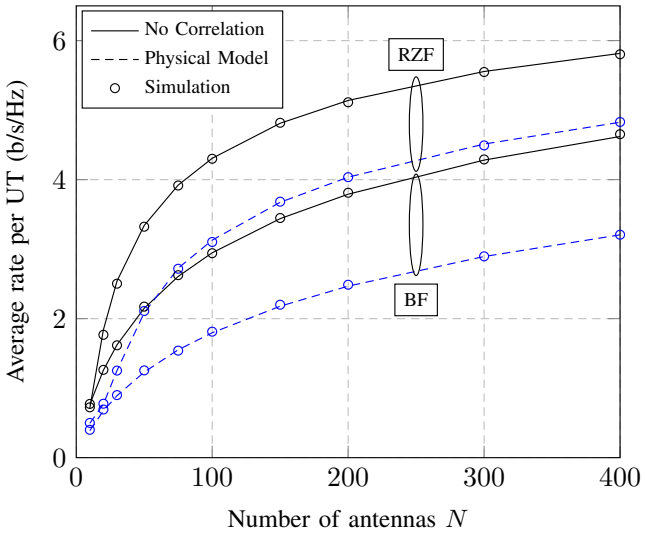

Fig. 6. Average per-user rate with BF and RZF precoding versus the number of antennas N. Solid and dashed lines depict the asymptotic approximations, markers the simulation results.

steering vectors $\mathbf{a}(\phi) \in \mathbb{C}^N$ defined as

$$\mathbf{a}(\phi) = \frac{1}{\sqrt{P}} \left[1, e^{-\mathbf{i}2\pi\omega \sin(\phi)}, \ldots, e^{-\mathbf{i}2\pi\omega(N-1)\sin(\phi)}\right]^{\mathsf{T}} \quad (34)$$

where ω is the antenna spacing in multiples of the wavelength and $\phi_p = -\pi/2 + (p-1)\pi/P$, $p = 1, \ldots, P$, are the uniformly distributed angles of transmission. We assume that the physical dimensions P scale with the number of antennas as $P = N/2$ and let $\omega = 0.3$. Since $\frac{1}{N}\mathrm{tr}\,\mathbf{A}\mathbf{A}^{\mathsf{H}} = 1$, the ultimately achievable rates under this channel model are equal to those of the previous channel model without antenna correlation. For comparison, we also depict in Fig. 6 the achievable rates and their approximations for the physical channel model. Interestingly, while the shapes of the curves for both precoders are similar to those without antenna correlation, it becomes clear that low rank correlation matrices severely degrade the performance. Note that we have assumed the same correlation matrix $\mathbf{A}$ for all UTs. In a practical system, however, different UTs will have different correlation matrices, possibly spanning different subspaces. In such a scenario, antenna correlation might have also some positive effects. For a further discussion of this topic, we refer to the very recent work [24].

VI. Conclusions

We have provided a unified analysis of the UL/DL performance of linear detectors/precoders in non-cooperative multi-cell multi-user TDD systems. Assuming a large system limit, we have derived asymptotically tight approximations of achievable UL/DL-rates under a very general channel model which accounts for imperfect channel estimation, pilot contamination, path loss, and terminal-specific antenna correlation. These approximations were shown to be accurate for realistic system dimensions and enable, consequently, future studies of realistic effects, such as antenna correlation, spacing and aperture, without the need for simulations. Our results are also directly applicable in the context of large distributed antenna systems. For a simplified channel model, we have observed that the performance depends mainly on the physical DoF per UT the channel offers and the effective SNR. Moreover, we have determined how many antennas are needed to achieve $\eta\,\%$ of the ultimate performance limit with infinitely many antennas and how many more antennas are needed with MF/BF to achieve MMSE/RZF performance. Simulations for a more realistic system model suggest that MMSE/RZF can achieve the performance of the simple MF/BF schemes with a significantly reduced number of antennas. Since massive MIMO TDD systems are a promising network architecture, it seems necessary to verify the theoretical performance predictions by channel measurements and prototypes.

Appendix A
Useful Lemmas

Lemma 1 (Matrix inversion lemma (I) [25, Eq. (2.2)]): Let $\mathbf{A} \in \mathbb{C}^{N\times N}$ be Hermitian invertible. Then, for any vector $\mathbf{x} \in \mathbb{C}^N$ and any scalar $\tau \in \mathbb{C}$ such that $\mathbf{A} + \tau\mathbf{x}\mathbf{x}^{\mathsf{H}}$ is invertible,

$$\mathbf{x}^{\mathsf{H}}(\mathbf{A} + \tau\mathbf{x}\mathbf{x}^{\mathsf{H}})^{-1} = \frac{\mathbf{x}^{\mathsf{H}}\mathbf{A}^{-1}}{1 + \tau\mathbf{x}^{\mathsf{H}}\mathbf{A}^{-1}\mathbf{x}}.$$

Lemma 2 (Matrix inversion lemma (II)): Let $\mathbf{A} \in \mathbb{C}^{N\times N}$ be Hermitian invertible. Then, for any vector $\mathbf{x} \in \mathbb{C}^N$ and any scalar $\tau \in \mathbb{C}$ such that $\mathbf{A} + \tau\mathbf{x}\mathbf{x}^{\mathsf{H}}$ is invertible,

$$(\mathbf{A} + \tau\mathbf{x}\mathbf{x}^{\mathsf{H}})^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\tau\mathbf{x}\mathbf{x}^{\mathsf{H}}\mathbf{A}^{-1}}{1 + \tau\mathbf{x}^{\mathsf{H}}\mathbf{A}^{-1}\mathbf{x}}.$$

Lemma 3 (Rank-1 perturbation lemma [25]): Let $z < 0$, $\mathbf{A} \in \mathbb{C}^{N\times N}$, $\mathbf{B} \in \mathbb{C}^{N\times N}$ with $\mathbf{B}$ Hermitian nonnegative definite, and $\mathbf{v} \in \mathbb{C}^N$. Then,

$$\left|\mathrm{tr}\left((\mathbf{B} - z\mathbf{I}_N)^{-1} - (\mathbf{B} + \mathbf{v}\mathbf{v}^{\mathsf{H}} - z\mathbf{I}_N)^{-1}\right)\mathbf{A}\right| \leq \frac{\|\mathbf{A}\|}{|z|}.$$

Lemma 4 ([26, Lem. B.26], [27, Thm. 3.7],[21, Lem. 12]): 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 $p \geq 1$,

$$\begin{aligned}
&(i) \quad \mathbb{E}\left[\left|\mathbf{x}^{\mathsf{H}}\mathbf{A}\mathbf{x} - \frac{1}{N}\mathrm{tr}\,\mathbf{A}\right|^{p}\right] = \mathcal{O}\left(\frac{1}{N^{\frac{p}{2}}}\right)\\
&(ii) \quad \mathbf{x}^{\mathsf{H}}\mathbf{A}\mathbf{x} - \frac{1}{N}\mathrm{tr}\,\mathbf{A} \xrightarrow[N\to\infty]{\text{a.s.}} 0\\
&(iii) \quad \mathbf{x}^{\mathsf{H}}\mathbf{A}\mathbf{y} \xrightarrow[N\to\infty]{\text{a.s.}} 0\\
&(iv) \quad \mathbb{E}\left[\left|\left(\mathbf{x}^{\mathsf{H}}\mathbf{A}\mathbf{x}\right)^{2} - \left(\frac{1}{N}\mathrm{tr}\,\mathbf{A}\right)^{2}\right|\right] \xrightarrow[N\to\infty]{} 0.
\end{aligned}$$

Appendix B
Proofs

Proof of Theorem 4: We start by dividing the denominator and numerator of $\gamma_{jm}^{\mathrm{dl}}$ by $\frac{1}{N}$.

1) Signal power: Straight-forward computations yield to

$$\begin{aligned}
&\frac{1}{N}\lambda_j\left|\mathbb{E}\left[\mathbf{h}_{jjm}^{\mathsf{H}}\hat{\mathbf{h}}_{jjm}\right]\right|^{2}\\
&= \frac{1}{\mathbb{E}\left[\frac{1}{K}\sum_{k=1}^{K}\frac{1}{N}\hat{\mathbf{h}}_{jjk}^{\mathsf{H}}\hat{\mathbf{h}}_{jjk}\right]}\left|\mathbb{E}\left[\frac{1}{N}\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}\hat{\mathbf{h}}_{jjm}\right]\right|^{2}\\
&= \bar{\lambda}_j\left(\frac{1}{N}\mathrm{tr}\,\mathbf{\Phi}_{jjm}\right)^{2} \qquad (35)
\end{aligned}$$

where $\bar{\lambda}_j = \left(\frac{1}{K}\sum_{k=1}^{K}\frac{1}{N}\mathrm{tr}\,\mathbf{\Phi}_{jjk}\right)^{-1}$. By **A 1** and **A 2**, $0 < \liminf_N \bar{\lambda}_j \leq \limsup_N \bar{\lambda}_j < \infty$ holds.

2) Interference power: As a direct consequence of Lemma 4 (i) and the independence of $\hat{\mathbf{h}}_{jjm}$ and $\tilde{\mathbf{h}}_{jjm}$,

$$\begin{aligned}
&\frac{\lambda_l}{N}\mathrm{var}\left[\mathbf{h}_{jjm}^{\mathsf{H}}\hat{\mathbf{h}}_{jjm}\right]\\
&= \bar{\lambda}_j\mathbb{E}\left[\left|\frac{1}{N}\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}\hat{\mathbf{h}}_{jjm} - \frac{1}{N}\mathrm{tr}\,\mathbf{\Phi}_{jjm}\right|^{2}\right]\\
&\quad + \frac{\bar{\lambda}_j}{N^2}\mathbb{E}\left[\left|\tilde{\mathbf{h}}_{jjm}^{\mathsf{H}}\hat{\mathbf{h}}_{jjm}\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}\tilde{\mathbf{h}}_{jjm}\right|\right] \xrightarrow[N\to\infty]{} 0. \qquad (36)
\end{aligned}$$

For the remaining terms, we have by (7)

$$\begin{aligned}
&\frac{1}{N}\lambda_l\mathbb{E}\left[\left|\mathbf{h}_{ljm}^{\mathsf{H}}\hat{\mathbf{h}}_{llk}\right|^{2}\right]\\
&= \bar{\lambda}_l\mathbb{E}\left[\frac{1}{N^2}\left|\mathbf{h}_{ljm}^{\mathsf{H}}\mathbf{R}_{llk}\mathbf{Q}_{lk}\left(\sum_{i=1}^{L}\mathbf{h}_{lik} + \frac{1}{\sqrt{\rho_{\mathrm{tr}}}}\mathbf{n}_{lk}^{\mathrm{tr}}\right)\right|^{2}\right]\\
&= \bar{\lambda}_l\begin{cases}\frac{1}{N^2}\mathrm{tr}\,\mathbf{R}_{ljm}\mathbf{\Phi}_{llk} & ,\ k\neq m\\ \mathbb{E}\left[\left|\frac{1}{N}\mathbf{h}_{ljm}^{\mathsf{H}}\mathbf{R}_{llm}\mathbf{Q}_{lm}\mathbf{h}_{ljm}\right|^{2}\right] & \\ +\frac{1}{N^2}\mathrm{tr}\,\mathbf{R}_{ljm}\mathbf{\Phi}_{llm} & \\ -\frac{1}{N^2}\mathrm{tr}\,\mathbf{R}_{ljm}\mathbf{\Phi}_{ljm}\mathbf{Q}_{lm}\mathbf{R}_{llm} & ,\ k=m\end{cases}. \qquad (37)
\end{aligned}$$

By Lemma 4 (iv),

$$\mathbb{E}\left[\left|\frac{1}{N}\mathbf{h}_{ljm}^{\mathsf{H}}\mathbf{R}_{llm}\mathbf{Q}_{lm}\mathbf{h}_{ljm}\right|^{2}\right] - \left|\frac{1}{N}\mathrm{tr}\,\mathbf{\Phi}_{ljm}\right|^{2} \xrightarrow[N\to\infty]{} 0.$$

Combining all results yields

$$\begin{aligned}
&\sum_{(l,k)\neq(j,m)}\frac{1}{N}\lambda_l\mathbb{E}\left[\left|\mathbf{h}_{ljm}^{\mathsf{H}}\hat{\mathbf{h}}_{llk}\right|^{2}\right] -\\
&\frac{1}{N}\sum_{l,k}\bar{\lambda}_l\frac{1}{N}\mathrm{tr}\,\mathbf{R}_{ljm}\mathbf{\Phi}_{llk} - \sum_{l\neq j}\bar{\lambda}_l\left|\frac{1}{N}\mathrm{tr}\,\mathbf{\Phi}_{ljm}\right|^{2} \xrightarrow[N\to\infty]{} 0. \qquad (38)
\end{aligned}$$

Note that we have neglected the terms $\frac{1}{N^2}\mathrm{tr}\,\mathbf{R}_{ljm}\mathbf{\Phi}_{ljm}\mathbf{Q}_{lm}\mathbf{R}_{llm}$ which appear only $L-1$ times and therefore vanish asymptotically. Moreover, we have added the single term $\bar{\lambda}_j\frac{1}{N^2}\mathrm{tr}\,\mathbf{R}_{jjm}\mathbf{\Phi}_{jjk}$ which is also negligible for large N. Replacing the asymptotic approximations for the useful signal power and the interference power in (15) finally yields

$$\begin{aligned}
&\gamma_{jm}^{\mathrm{dl}} -\\
&\frac{\bar{\lambda}_j\left(\frac{1}{N}\mathrm{tr}\,\mathbf{\Phi}_{jjm}\right)^{2}}{\frac{1}{\rho_{\mathrm{dl}}N} + \frac{1}{N}\sum_{l,k}\bar{\lambda}_l\frac{1}{N}\mathrm{tr}\,\mathbf{R}_{ljm}\mathbf{\Phi}_{llk} + \sum_{l\neq j}\bar{\lambda}_l\left|\frac{1}{N}\mathrm{tr}\,\mathbf{\Phi}_{ljm}\right|^{2}}\\
&\xrightarrow[N\to\infty]{} 0. \qquad (39)
\end{aligned}$$

■

Proof of Theorem 6: Define the following matrices for $j = 1,\ldots,L$ and $k = 1,\ldots,K$:

$$\mathbf{\Sigma}_j = \left(\hat{\mathbf{H}}_{jj}\hat{\mathbf{H}}_{jj}^{\mathsf{H}} + \mathbf{Z}_j^{\mathrm{dl}} + N\varphi_j^{\mathrm{dl}}\mathbf{I}_N\right)^{-1} \qquad (40)$$

$$\mathbf{\Sigma}_{jk} = \left(\hat{\mathbf{H}}_{jj}\hat{\mathbf{H}}_{jj}^{\mathsf{H}} - \hat{\mathbf{h}}_{jjk}\hat{\mathbf{h}}_{jjk}^{\mathsf{H}} + \mathbf{Z}_j^{\mathrm{dl}} + N\varphi_j^{\mathrm{dl}}\mathbf{I}_N\right)^{-1}. \qquad (41)$$

1) Signal power: We divide the denominator and numerator of $\gamma_{jm}^{\mathrm{dl}}$ by $\frac{1}{N}$. Thus,[6]

$$\begin{aligned}
&\sqrt{\frac{\lambda_j}{N}}\mathbf{h}_{jjm}^{\mathsf{H}}\mathbf{w}_{jm}\\
&= \sqrt{\frac{\lambda_j}{N}}\mathbf{h}_{jjm}^{\mathsf{H}}\mathbf{\Sigma}_j\hat{\mathbf{h}}_{jjm}\\
&\overset{(a)}{=} \sqrt{\frac{K}{N\mathbb{E}\left[\mathrm{tr}\,\mathbf{\Sigma}_j\hat{\mathbf{H}}_{jj}\hat{\mathbf{H}}_{jj}^{\mathsf{H}}\mathbf{\Sigma}_j\right]}}\frac{\mathbf{h}_{jjm}^{\mathsf{H}}\mathbf{\Sigma}_{jm}\hat{\mathbf{h}}_{jjm}}{1+\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}\mathbf{\Sigma}_{jm}\hat{\mathbf{h}}_{jjm}}\\
&\overset{(b)}{\asymp} \sqrt{\frac{K}{N\mathbb{E}\left[\mathrm{tr}\,\mathbf{\Sigma}_j - \mathrm{tr}\,\left(\mathbf{Z}_j^{\mathrm{dl}} + N\varphi_j^{\mathrm{dl}}\mathbf{I}_N\right)\mathbf{\Sigma}_j^{2}\right]}}\frac{\frac{1}{N}\mathrm{tr}\,\mathbf{\Phi}_{jjm}\mathbf{T}_j}{1+\frac{1}{N}\mathrm{tr}\,\mathbf{\Phi}_{jjm}\mathbf{T}_j}\\
&\overset{(c)}{\asymp} \sqrt{\frac{K}{N}}\sqrt{\frac{1}{\frac{1}{N}\mathrm{tr}\,\mathbf{T}_j - \frac{1}{N}\mathrm{tr}\left(\frac{\mathbf{Z}_j^{\mathrm{dl}}}{N} + \varphi_j^{\mathrm{dl}}\mathbf{I}_N\right)\bar{\mathbf{T}}'}}\frac{\frac{1}{N}\mathrm{tr}\,\mathbf{\Phi}_{jjm}\mathbf{T}_j}{1+\frac{1}{N}\mathrm{tr}\,\mathbf{\Phi}_{jjm}\mathbf{T}_j}\\
&\overset{(d)}{=} \sqrt{\bar{\lambda}_j}\frac{\delta_{jm}}{1+\delta_{jm}} \qquad (42)
\end{aligned}$$

where (a) follows from Lemma 1, (b) follows from Lemma 4 (ii), Lemma 3, and Theorem 1 applied to the term $\mathbf{h}_{jjm}^{\mathsf{H}}\mathbf{\Sigma}_{jm}\hat{\mathbf{h}}_{jjm}$,[7] and (c) results from Theorem 1 applied to $\mathrm{tr}\,\mathbf{\Sigma}_j$ and Theorem 2 applied to $\mathrm{tr}\left(\mathbf{Z}_j^{\mathrm{dl}} + N\varphi_j^{\mathrm{dl}}\mathbf{I}_N\right)\mathbf{\Sigma}_j^{2}$. Note that both theorems do not only imply almost sure

[6] We denote $a_n \asymp b_n$ the equivalence relation $a_n - b_n \xrightarrow[n\to\infty]{\text{a.s.}} 0$ for two infinite sequences a_n and b_n.

[7] Note that these are standard steps in proofs using large random matrix theory. Details can be found, e.g., in [20] or [27].

convergence but also convergence in the mean. In the last step, we have used the definitions $\delta_{jm} = \frac{1}{N}\mathrm{tr}\,\mathbf{R}_{jjm}\mathbf{T}_j$ and $\bar{\lambda}_j = \frac{K}{N}\left(\frac{1}{N}\mathrm{tr}\,\mathbf{T}_j - \frac{1}{N}\mathrm{tr}\left(\frac{\mathbf{Z}_j^{\mathrm{dl}}}{N} + \varphi_j^{\mathrm{dl}}\mathbf{I}_N\right)\bar{\mathbf{T}}'\right)^{-1}$. By the dominated convergence theorem [18] and the continuous mapping theorem [19], it is straight-forward to show that

$$\frac{\lambda_j}{N}\left|\mathbb{E}\left[\mathbf{h}_{jjm}^{\mathsf{H}}\boldsymbol{\Sigma}_j\hat{\mathbf{h}}_{jjm}\right]\right|^2 - \bar{\lambda}_j\frac{\delta_{jm}^2}{(1+\delta_{jm})^2} \xrightarrow[N\to\infty]{} 0. \quad (43)$$

2) Interference power: Define the following quantities:

$$a = \hat{\mathbf{h}}_{jjm}^{\mathsf{H}}\boldsymbol{\Sigma}_j\hat{\mathbf{h}}_{jjm} \quad (44)$$

$$\bar{a} = \mathbb{E}\left[\mathbf{h}_{jjm}^{\mathsf{H}}\boldsymbol{\Sigma}_j\hat{\mathbf{h}}_{jjm}\right] \quad (45)$$

$$b = \tilde{\mathbf{h}}_{jjm}^{\mathsf{H}}\boldsymbol{\Sigma}_j\hat{\mathbf{h}}_{jjm}. \quad (46)$$

By Lemma 1, we have $0 \le a, \bar{a} \le 1$. Moreover, $\mathbb{E}[b] = 0$ and $\mathbb{E}[ab] = \mathbb{E}[ab^*] = 0$. Thus,

$$\begin{aligned}\mathrm{var}\left[\mathbf{h}_{jjm}^{\mathsf{H}}\boldsymbol{\Sigma}_j\hat{\mathbf{h}}_{jjm}\right] &= \mathbb{E}\left[|a+b-\bar{a}|^2\right] \\ &= \mathbb{E}\left[(a-\bar{a})(a+\bar{a})\right] + \mathbb{E}\left[|b|^2\right] \\ &\le 2\,\mathbb{E}\left[|a-\bar{a}|\right) + \mathbb{E}\left[|b|^2\right]. \quad (47)\end{aligned}$$

We have shown in (42) (b) that $a - \frac{\delta_{jm}}{1+\delta_{jm}} \xrightarrow[N\to\infty]{\text{a.s}} 0$. Since $a, \bar{a}$ are bounded, this implies by the dominated convergence theorem that also $\mathbb{E}[|a-\bar{a}|] \xrightarrow[N\to\infty]{} 0$. Moreover, one can show that $\mathbb{E}\left[|b|^2\right] \le \frac{1}{N}{\varphi_j^{\mathrm{dl}}}^{-2}\|\mathbf{R}_{jjm}\|^2 \xrightarrow[N\to\infty]{} 0$. Thus, we have from (47)

$$\frac{1}{N}\lambda_j\mathrm{var}\left[\mathbf{h}_{jjm}^{\mathsf{H}}\boldsymbol{\Sigma}_j\hat{\mathbf{h}}_{jjm}\right] \xrightarrow[N\to\infty]{} 0. \quad (48)$$

Consider now the terms $|\mathbf{h}_{jlm}^{\mathsf{H}}\mathbf{w}_{lk}|^2$:

$$\begin{aligned}|\mathbf{h}_{jlm}^{\mathsf{H}}\mathbf{w}_{lk}|^2 &\overset{(a)}{=} \frac{\hat{\mathbf{h}}_{llk}^{\mathsf{H}}\boldsymbol{\Sigma}_{lk}\mathbf{h}_{jlm}\mathbf{h}_{jlm}^{\mathsf{H}}\boldsymbol{\Sigma}_{lk}\hat{\mathbf{h}}_{llk}}{\left(1+\hat{\mathbf{h}}_{llk}^{\mathsf{H}}\boldsymbol{\Sigma}_{lk}\hat{\mathbf{h}}_{llk}\right)^2} \\ &\overset{(b)}{\asymp} \frac{1}{(1+\delta_{lk})^2}\begin{cases}\mathbf{h}_{ljm}^{\mathsf{H}}\boldsymbol{\Sigma}_{lk}\boldsymbol{\Phi}_{llk}\boldsymbol{\Sigma}_{lk}\mathbf{h}_{ljm} & ,\ k\ne m \\ |\vartheta_{ljm}|^2 & ,\ k=m\end{cases} \quad (49)\end{aligned}$$

where (a) is due to Lemma 1, (b) follows from Lemmas 4, 3, Theorem 1, and where we have used the definitions $\delta_{lk} = \frac{1}{N}\mathrm{tr}\,\boldsymbol{\Phi}_{llk}\mathbf{T}_l$ and $\vartheta_{ljm} = \frac{1}{N}\mathrm{tr}\,\boldsymbol{\Phi}_{ljm}\mathbf{T}_l$. In order to treat the terms for $k \ne m$ further, we need the following identity from Lemma 2:

$$\boldsymbol{\Sigma}_{lk} = \boldsymbol{\Sigma}_{lkm} - \frac{\boldsymbol{\Sigma}_{lkm}\hat{\mathbf{h}}_{llm}\hat{\mathbf{h}}_{llm}^{\mathsf{H}}\boldsymbol{\Sigma}_{lkm}}{1+\hat{\mathbf{h}}_{llm}^{\mathsf{H}}\boldsymbol{\Sigma}_{lkm}\hat{\mathbf{h}}_{llm}}, \qquad k \ne m \quad (50)$$

where

$$\boldsymbol{\Sigma}_{lkm} = \left(\hat{\mathbf{H}}_{jj}\hat{\mathbf{H}}_{jj}^{\mathsf{H}} - \hat{\mathbf{h}}_{jjk}\hat{\mathbf{h}}_{jjk}^{\mathsf{H}} - \hat{\mathbf{h}}_{jjm}\hat{\mathbf{h}}_{jjm}^{\mathsf{H}} + \mathbf{Z}_j^{\mathrm{dl}} + N\varphi_j^{\mathrm{dl}}\mathbf{I}_N\right)^{-1}. \quad (51)$$

Note that $\boldsymbol{\Sigma}_{lkm}$ is independent of $\mathbf{h}_{jlm}$ while $\boldsymbol{\Sigma}_{lk}$ is not. Using (50), we can write

$$\begin{aligned}&\mathbf{h}_{ljm}^{\mathsf{H}}\boldsymbol{\Sigma}_{lk}\boldsymbol{\Phi}_{llk}\boldsymbol{\Sigma}_{lk}\mathbf{h}_{ljm} \\ &= \mathbf{h}_{ljm}^{\mathsf{H}}\boldsymbol{\Sigma}_{lkm}\boldsymbol{\Phi}_{llk}\boldsymbol{\Sigma}_{lkm}\mathbf{h}_{ljm} \\ &\quad + \frac{\left|\mathbf{h}_{ljm}^{\mathsf{H}}\boldsymbol{\Sigma}_{lkm}\hat{\mathbf{h}}_{llm}\right|^2\hat{\mathbf{h}}_{llm}^{\mathsf{H}}\boldsymbol{\Sigma}_{lkm}\boldsymbol{\Phi}_{llk}\boldsymbol{\Sigma}_{lkm}\hat{\mathbf{h}}_{llm}}{\left(1+\hat{\mathbf{h}}_{llm}^{\mathsf{H}}\boldsymbol{\Sigma}_{lkm}\hat{\mathbf{h}}_{llm}\right)^2} \\ &\quad - 2\mathrm{Re}\left\{\frac{\hat{\mathbf{h}}_{llm}^{\mathsf{H}}\boldsymbol{\Sigma}_{lkm}\mathbf{h}_{ljm}\mathbf{h}_{ljm}^{\mathsf{H}}\boldsymbol{\Sigma}_{lkm}\boldsymbol{\Phi}_{llk}\boldsymbol{\Sigma}_{lk}\hat{\mathbf{h}}_{llm}}{1+\hat{\mathbf{h}}_{llm}^{\mathsf{H}}\boldsymbol{\Sigma}_{lkm}\hat{\mathbf{h}}_{llm}}\right\}. \quad (52)\end{aligned}$$

As already shown above, we have $\hat{\mathbf{h}}_{llm}^{\mathsf{H}}\boldsymbol{\Sigma}_{lkm}\hat{\mathbf{h}}_{llm} \asymp \delta_{lm}$ and $\hat{\mathbf{h}}_{llm}^{\mathsf{H}}\boldsymbol{\Sigma}_{lkm}\mathbf{h}_{ljm} \asymp \vartheta_{ljm}^*$. From Lemmas 4, 3 and Theorem 2, we can similarly obtain

$$\mathbf{h}_{ljm}^{\mathsf{H}}\boldsymbol{\Sigma}_{lkm}\boldsymbol{\Phi}_{llk}\boldsymbol{\Sigma}_{lkm}\mathbf{h}_{ljm} \asymp \frac{1}{N^2}\mathrm{tr}\,\mathbf{R}_{ljm}\mathbf{T}'_{lk} \quad (53)$$

$$\hat{\mathbf{h}}_{llm}^{\mathsf{H}}\boldsymbol{\Sigma}_{lkm}\boldsymbol{\Phi}_{llk}\boldsymbol{\Sigma}_{lkm}\hat{\mathbf{h}}_{llm} \asymp \frac{1}{N^2}\mathrm{tr}\,\boldsymbol{\Phi}_{llm}\mathbf{T}'_{lk} = \frac{\delta'_{lmk}}{N} \quad (54)$$

$$\mathbf{h}_{ljm}^{\mathsf{H}}\boldsymbol{\Sigma}_{lkm}\boldsymbol{\Phi}_{llk}\boldsymbol{\Sigma}_{lkm}\hat{\mathbf{h}}_{llm} \asymp \frac{1}{N^2}\mathrm{tr}\,\boldsymbol{\Phi}_{ljm}\mathbf{T}'_{lk} = \frac{\vartheta'_{ljmk}}{N} \quad (55)$$

where $\mathbf{T}'_{lk} = \mathbf{T}'(\varphi_l^{\mathrm{dl}})$ and $\boldsymbol{\delta}'_{lk} = [\delta'_{l1k}\dots\delta'_{lKk}]^{\mathsf{T}} = \boldsymbol{\delta}'(\varphi_l^{\mathrm{dl}})$ are given by Theorem 2 for $\mathbf{S} = \mathbf{Z}_l^{\mathrm{dl}}/N$, $\boldsymbol{\Theta} = \boldsymbol{\Phi}_{llk}$, $\mathbf{D} = \mathbf{I}_N$, and $\mathbf{R}_k = \boldsymbol{\Phi}_{llk}$ for all k. Combining the last results yields to

$$\begin{aligned}&\mathbf{h}_{ljm}^{\mathsf{H}}\boldsymbol{\Sigma}_{lk}\boldsymbol{\Phi}_{llk}\boldsymbol{\Sigma}_{lk}\mathbf{h}_{ljm} \asymp \frac{\mathrm{tr}\,\mathbf{R}_{ljm}\mathbf{T}'_{lk}}{N^2} \\ &- \frac{2\mathrm{Re}\left\{\vartheta_{ljm}^*\vartheta'_{ljmk}\right\}(1+\delta_{jm}) - |\vartheta_{ljm}|^2\delta'_{lmk}}{N(1+\delta_{lm})^2} = \frac{\mu_{ljmk}}{N}. \quad (56)\end{aligned}$$

Note now that

$$\begin{aligned}\sum_{(l,k)\ne(j,m)}\frac{\lambda_l}{N}|\mathbf{h}_{jlm}^{\mathsf{H}}\mathbf{w}_{lk}|^2 &\le \sum_l \bar{\lambda}_l\mathbf{h}_{jlm}^{\mathsf{H}}\boldsymbol{\Sigma}_l\hat{\mathbf{H}}_{ll}\hat{\mathbf{H}}_{ll}^{\mathsf{H}}\boldsymbol{\Sigma}_l^{\mathsf{H}}\mathbf{h}_{jlm} \\ &\le \sum_l \bar{\lambda}\mathbf{h}_{jlm}^{\mathsf{H}}\boldsymbol{\Sigma}_{lm}\mathbf{h}_{ljm}. \quad (57)\end{aligned}$$

Since $\mathbf{h}_{jlm}^{\mathsf{H}}\boldsymbol{\Sigma}_{lm}\mathbf{h}_{ljm} \asymp \frac{1}{N}\mathrm{tr}\,\mathbf{R}_{ljm}\mathbf{T}_l$, $\mathbb{E}\left[\mathbf{h}_{jlm}^{\mathsf{H}}\boldsymbol{\Sigma}_{lm}\mathbf{h}_{ljm}\right] - \frac{1}{N}\mathrm{tr}\,\mathbf{R}_{ljm}\mathbf{T}_l \to 0$ by Lemmas 4, 3 and Theorem 1, and $\frac{1}{N}\mathrm{tr}\,\mathbf{R}_{ljm}\mathbf{T}_l \le \frac{1}{\varphi_l^{\mathrm{dl}}}\|\mathbf{R}_{ljm}\|$, we have by dominated convergence arguments

$$\begin{aligned}\sum_{(l,k)\ne(j,m)}\frac{\lambda_l}{N}\mathbb{E}\left[|\mathbf{h}_{jlm}^{\mathsf{H}}\mathbf{w}_{lk}|^2\right] - \sum_{l,k}\frac{\bar{\lambda}_l}{N}\frac{\mu_{ljmk}}{(1+\delta_{lk})^2} \\ - \sum_{l\ne j}\bar{\lambda}_j\frac{|\vartheta_{ljm}|^2}{(1+\delta_{lm})^2} \xrightarrow[N\to\infty]{} 0 \quad (58)\end{aligned}$$

where we have also subtracted the asymptotically negligible term $\frac{\bar{\lambda}_j}{N}\frac{\mu_{jjmm}}{(1+\delta_{jm})^2}$.

Combining (43), (48), and (58) concludes the proof. ■

REFERENCES

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

[2] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, 2012, to appear. [Online]. Available: http://arxiv.org/abs/1201.3210

[3] J. G. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. Reed, "Femtocells: Past, present, and future," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 497–508, Apr. 2012.

[4] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, 2012, submitted. [Online]. Available: http://arxiv.org/abs/1112.3810

[5] A. Fehske, G. Fettweis, J. Malmodin, and G. Biczok, "The global footprint of mobile communications: The ecological and economic perspective," *IEEE Commun. Mag.*, vol. 49, no. 8, pp. 55–62, Aug. 2011.

[6] International Agency for Research on Cancer (IARC), "IARC classifies radiofrequency electromagnetic fields as possibly carcinogenic to humans," May 2011. [Online]. Available: http://www.iarc.fr/en/media-centre/pr/2011/pdfs/pr208_E.pdf

[7] T. L. Marzetta, "How much training is required for multiuser MIMO?" in *Proc. IEEE Asilomar Conference on Signals, Systems and Computers (ACSSC'06)*, Pacific Grove, CA, US, Nov. 2006, pp. 359–363.

[8] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Analysis of the pilot contamination effect in very large multicell multiuser MIMO systems for physical channel models," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'11)*, Prague, Czech Repulic, May 22–27, 2011, pp. 3464–3467.

[9] S. K. Mohammed and E. G. Larsson, "Single-user beamforming in large-scale MISO systems with per-antenna constant-envelope constraints: The doughnut channel," *IEEE Trans. Wireless Commun.*, 2011, to appear. [Online]. Available: http://arxiv.org/abs/1111.3752

[10] X. Gao, O. Edfors, F. Rusek, and F. Tufvesson, "Linear pre-coding performance in measured very-large MIMO channels," in *Proc. IEEE Vehicular Technology Conf. (VTC Fall)*, San Francisco, CA, US, Sep. 2011, pp. 1–5.

[11] S. Payami and F. Tufvesson, "Channel measurements and analysis for very large array systems at 2.6 Ghz," in *Proc. 6th European Conference on Antennas and Propagation (EuCAP'12)*, Prague, Czech Republic, Mar. 2012, pp. 433–437.

[12] J. Hoydis, C. Hoek, T. Wild, and S. ten Brink, "Channel measurements for large antenna arrays," in *Proc. IEEE International Symposium on Wireless Communication Systems (ISWCS'12)*, Paris, France, Aug. 2012, pp. 811–815.

[13] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, Inc. Upper Saddle River, NJ, USA, 1993.

[14] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.

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

[16] H. Huh, G. Caire, H. C. Papadopoulos, and S. A. Ramprashad, "Achieving massive MIMO spectral efficiency with a not-so-large number of antennas," *IEEE Trans. Wireless Commun.*, 2012, to appear. [Online]. Available: http://arxiv.org/abs/1107.3862

[17] H. Huh, A. M. Tulino, and G. Caire, "Network MIMO with linear zero-forcing beamforming: Large system analysis, impact of channel estimation, and reduced-complexity scheduling," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 2911–2934, May 2012.

[18] P. Billingsley, *Probability and Measure*, 3rd ed. John Wiley & Sons, Inc., 1995.

[19] A. W. van der Vaart, *Asymptotic Statistics (Cambridge Series in Statistical and Probabilistic Mathematics)*. Cambridge University Press, New York, 2000.

[20] S. Wagner, R. Couillet, M. Debbah, and D. T. M. 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.

[21] J. Hoydis, "Random matrix methods for advanced communication systems," Ph.D. dissertation, Supélec, Gif-Sur-Yvette, France, Apr. 2012. [Online]. Available: http://www.flexible-radio.com/sites/default/files/publications/1/jakob_thesis.pdf

[22] A. D. Wyner, "Shannon-theoretic approach to a Gaussian cellular multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 40, no. 6, pp. 1713–1727, 1994.

[23] S. Wei, D. Goeckel, and R. Janaswamy, "On the asymptotic capacity of MIMO systems with antenna arrays of fixed length," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1608–1621, Jul. 2005.

[24] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing," 2012, submitted. [Online]. Available: http://arxiv.org/abs/1209.1402

[25] J. W. Silverstein and Z. D. Bai, "On the empirical distribution of eigenvalues of a class of large dimensional random matrices," *Journal of Multivariate Analysis*, vol. 54, no. 2, pp. 175–192, 1995.

[26] Z. D. Bai and J. W. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices*, 2nd ed. Springer Series in Statistics, New York, NY, USA, 2009.

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

Jakob Hoydis (S'08–M'12) received the diploma degree (Dipl.-Ing.) in electrical engineering and information technology from RWTH Aachen University, Germany, and the Ph.D. degree from Supélec, Gif-sur-Yvette, France, in 2008 and 2012, respectively. Since May 2012, he is a research engineer at Bell Laboratories, Alcatel-Lucent, Stuttgart, Germany. His research interests are in the area of large random matrix theory and information theory and their applications to wireless communications.

Stephan ten Brink (M'97–SM'11) received a Dipl.-Ing. degree in electrical engineering and a Dr.-Ing. degree from the University of Stuttgart, Germany, in 1997 and 2000, respectively. From 2000 to 2003, he was with the Wireless Research Laboratory, Bell Laboratories, Lucent Technologies, Holmdel, New Jersey, conducting research on channel coding for multiple-antenna systems. From 2003 to 2010, he was with Realtek Semiconductor Corp., Irvine, California, developing ASIC solutions for wireless LAN and ultra-wideband systems. Since April 2010, he has been with Bell Laboratories, Alcatel-Lucent, Stuttgart, Germany, where he heads the Wireless Physical Layer Research Department. His research interests include multiple-antenna communications and channel coding for wireless and optical systems.

Mérouane Debbah (S'01–A'03–M'04–SM'08) was born in Madrid, Spain. He entered the Ecole Normale Supérieure de Cachan (France) in 1996 where he received his M.Sc. and Ph.D. degrees respectively in 1999 and 2002. From 1999 to 2002, he worked for Motorola Labs on Wireless Local Area Networks and prospective fourth generation systems. From 2002 until 2003, he was appointed Senior Researcher at the Vienna Research Center for Telecommunications (FTW) (Vienna, Austria). From 2003 until 2007, he joined the Mobile Communications department of the Institut Eurecom (Sophia Antipolis, France) as an Assistant Professor. He is presently a Professor at Supélec (Gif-sur-Yvette, France), holder of the Alcatel-Lucent Chair on Flexible Radio. His research interests are in information theory, signal processing, and wireless communications. Mérouane Debbah is the recipient of the "Mario Boella" prize award in 2005, the 2007 General Symposium IEEE GLOBECOM best paper award, the WiOpt 2009 best paper award, the 2010 Newcom++ best paper award, as well as the Valuetools 2007, Valuetools 2008, and CrownCom 2009 best student paper awards. He is a WWRF fellow. In 2011, he received the IEEE Glavieux Prize Award. He is the co-founder of the Live Video Network Ximinds (http://www.ximinds.com).