Interference-Aware RZF Precoding for Multi Cell Downlink Systems

Axel Müller, Romain Couillet, Emil Björnson, Sebastian Wagner, Mérouane Debbah

Interference-Aware RZF Precoding for Multi Cell Downlink Systems

Axel Müller, *Member, IEEE,* Romain Couillet, *Member, IEEE,* Emil Björnson, *Member, IEEE,* Sebastian Wagner, *Member, IEEE,* and Mérouane Debbah, *Senior Member, IEEE*

***Abstract*—Recently, a structure of an optimal linear precoder for multi cell downlink systems has been described in [1, Eq (3.33)]. Other references (e.g., [2], [3]) have used simplified versions of the precoder to obtain promising performance gains. These gains have been hypothesized to stem from the additional degrees of freedom that allow for interference mitigation through interference relegation to orthogonal subspaces. However, no conclusive or rigorous understanding has yet been developed.**

In this paper, we build on an intuitive interference induction trade-off and the aforementioned precoding structure to propose an interference aware RZF (iaRZF) precoding scheme for multi cell downlink systems and we analyze its rate performance. Special emphasis is placed on the induced interference mitigation mechanism of iaRZF. For example, we will verify the intuitive expectation that the precoder structure can either completely remove induced inter-cell or intra-cell interference. We state new results from large-scale random matrix theory that make it possible to give more intuitive and insightful explanations of the precoder behavior, also for cases involving imperfect channel state information (CSI). We remark especially that the interference-aware precoder makes use of all available information about interfering channels to improve performance. Even very poor CSI allows for significant sum-rate gains. Our obtained insights are then used to propose heuristic precoder parameters for arbitrary systems, whose effectiveness are shown in more involved system scenarios. Furthermore, calculation and implementation of these parameters does not require explicit inter base station cooperation.

***Index Terms*—Multi user MIMO, linear precoding, multi cell downlink, interference mitigation, random matrix theory.**

A. Müller was with Intel Mobile Communications, Sophia Antipolis, France and with the Alcatel-Lucent Chair on Flexible Radio, Supélec, Gif-sur-Yvette, France. He is currently with the Mathematical and Algorithmic Sciences Lab, France Research Center, Huawei Technologies Co. Ltd., Boulogne-Billancourt, France (email: axel.mueller@huawei.com).

R. Couillet is with Laboratoire des Signaux et Systmes (L2S, UMR CNRS 8506), CentraleSupelec - CNRS - Université Paris-Sud, Gif-sur-Yvette, France (email: romain.couillet@centralesupelec.fr).

E. Björnson was with the Alcatel-Lucent Chair on Flexible Radio, Supélec, Gif-sur-Yvette, France, and with the Department of Signal Processing, KTH Royal Institute of Technology, Stockholm, Sweden. He is currently with the Department of Electrical Engineering (ISY), Linköping University, Linköping, Sweden (email: emil.bjornson@liu.se).

S. Wagner is with Intel Mobile Communications, Sophia Antipolis, France (email: sebastian.wagner@intel.com).

M. Debbah is with Laboratoire des Signaux et Systèmes (L2S, UMR CNRS 8506), CentraleSupelec - CNRS - Université Paris-Sud, Gif-sur-Yvette, France, and with the Mathematical and Algorithmic Sciences Lab, France Research Center, Huawei Technologies Co. Ltd., Boulogne-Billancourt, France (email: merouane.debbah@centralesupelec.fr).

E. Björnson was funded by the International Postdoc Grant 2012-228 from The Swedish Research Council. This research has been supported by the ERC Starting Grant 305123 MORE (Advanced Mathematical Tools for Complex Network Engineering).

I. Introduction

The growth of data traffic and the number of user terminals (UTs) in cellular networks will likely persist for the foreseeable future [4]. In order to deal with the resulting demand, it is estimated [5] that a thousand-fold increase in network capacity is required over the next 10 years. Given that the available spectral resources are severely limited, the majority of the wireless community sees massive network densification as the most realistic approach to solving most pressing issues. Also historically, shrinking cell size has been the single most successful technique in satisfying demand for network capacity [6, Chapter 6.3.4]. In recent times, this technique has been named the *small cell* approach [7], [8]. A large body of research indicates that interference still is a major limiting factor for capacity in multi cell scenarios [9], [10], especially in modern cellular networks that serve a multitude of users within the same time/frequency resources. In general, we see a trend to using more and more antennas for interference mitigation, e.g., via the massive MIMO approach [11]. Here, the number of transmit antennas surpasses the number of served UTs by an order of magnitude. Independent of this specific approach, the surplus antennas can be used to mitigate interference by using spatial precoding [1], [12], [13], [10]. The interference problem is generally compounded by the effect of imperfect knowledge concerning the channel state information (CSI). Such imperfections are unavoidable, as imperfect estimation algorithms, limited number of orthogonal pilot sequences, mobile UTs, delays, etc. can not be avoided in practice. Hence, one is interested in employing precoding schemes that are robust to CSI estimation errors and exploit the available CSI as efficiently as possible.

Arguably, the most successful and practically applicable precoding scheme used today is RZF precoding [14] (also known as minimum mean square error (MMSE) precoding, transmit Wiener filter, generalized eigenvalue-based beamformer, etc.; see [1, Remark 3.2]). Classical RZF precoders are only defined for single cell systems and thus do not take inter cell interference into account. Disregarding available information about inter cell interference is particularly detrimental in high density scenarios, where interference is a main performance limiting factor. Early multi cell extensions of the RZF scheme do not take the quality of CSI into account [15] and later ones either rely on heuristic distributed optimization algorithms or on inter cell cooperation [16] to determine the precoding vector. Thus, they offer limited insight into the precoder structure, i.e., into how the precoder works and how

it could be improved.

An intuitive extension of the single cell RZF, with the goal of completely eliminating induced interference is to substitute the intra cell channel matrix $\mathbf{H}$ in the (qualitative, single antenna UTs) precoder formulation[1] $\mathbf{F} = \mathbf{H}(\mathbf{H}^{\mathsf{H}}\mathbf{H}+\xi\mathbf{I})^{-1}$ by a matrix $\breve{\mathbf{H}}$, which is $\mathbf{H}$ projected onto the space orthogonal to the inter cell channel matrices, i.e., $\breve{\mathbf{F}} = \breve{\mathbf{H}}\left(\breve{\mathbf{H}}^{\mathsf{H}}\breve{\mathbf{H}}+\xi\mathbf{I}\right)^{-1}$. Hence induced interference can be completely removed at the cost of reduced signal power, if the CSI is perfectly known. However, it is immediately clear that this is a very harsh requirement, since the projection negatively affects the amount of signal energy received at the served UTs (unless $\mathbf{H} = \breve{\mathbf{H}}$). Assuming the precoding objective is system wide sum-rate optimization, one realizes that single cell RZF is probably not optimal, since it reduces the rate in other cells due to induced interference. Thus, a trade-off between the two extremes is expected to be beneficial, especially when the channel matrices are estimated with dissimilar quality. In this paper we analyze the following class of precoders for multi-cell single antenna UT systems, which we will denote *interference-aware RZF* (iaRZF). This class allows for the desired trade-off, as will be shown later on:

$$\mathbf{F}_m^m = \left(\sum_{l=1}^{L} \alpha_l^m \hat{\mathbf{H}}_l^m (\hat{\mathbf{H}}_l^m)^{\mathsf{H}} + \xi_m \mathbf{I}_{N_m}\right)^{-1} \hat{\mathbf{H}}_m^m \nu_m^{\frac{1}{2}}. \qquad (1)$$

Here $\mathbf{F}_{m}^{m}$ is the linear precoder used by base station (BS) m and $\hat{\mathbf{H}}_l^m$ denotes the imperfect estimate of the channel matrix from BS m to the UTs in cell l. The factor ξ_m is a regularization parameter and ν_m normalizes the precoder. Each channel matrix is assigned a factor α_l^m, that can be interpreted as the importance placed on the respective estimated channel. It is easy to see how this structure can mimic single cell RZF under perfect CSI (choose $\alpha_l^m = 0$, $l \neq m$ and $\alpha_m^m = 1$). The weights α_l^m allow balancing signal power directed to the served users with interference induced to other cells. This can be used to optimize sum-rate performance in certain cases, as will be shown in Section II. In general the optimal weights are not known and the classical UL/DL duality approach (e.g., [17], [1]) cannot be applied to find these weights, outside of power minimization settings or, if imperfect CSI is considered. We note that every BS can try to estimate the interference from other cells without explicit inter cell cooperation or communication, by means of blind or known pilot based schemes, though the CSI quality might be rather poor. Such estimation might be considered as implicit coordination. In [3] a simplified version of iaRZF was discussed, where a single subset of UT channels was weighted with respect to an estimated receive covariance matrix of all interfering channels. Hoydis *et al.* argued that "large [weights] make the precoding vectors more orthogonal to the interference subspace", but they did not conclusively and rigorously show how or why this is achieved. The work in [3] builds upon results from [2, Theorem 6], which introduced the precoding structure offering a single common balance parameter weighting all inter cell interference. Differing from these works, we are more interested in increasing the sum-rate performance by giving the traditional RZF precoder even more additional degrees of freedom. Doing so enables it to separately take into account the interference induced to certain subspaces, associated with different cells. Increasing the degrees of freedom in the interference suppression was also shown to be effective in [18] for perfect CSI. Similar to the traditional RZF precoder our approach is empirical and based on several motivational aspects (see Section II). The iaRZF structure is also partially based on the work in [17] and [1, Eq (3.33)]. In the latter, one finds one of the most recent and general treatments of the multi cell RZF precoder, along with proof that the proposed structure is optimal w.r.t. many utility functions of practical interest (see also [19]).

This paper analyzes the proposed iaRZF scheme, showing that it can significantly improve sum-rate performance in high interference multi cellular scenarios. In particular, it is not necessary to have reliable estimations of interfering channels; even very poor CSI allow for significant gains. We facilitate intuitive understanding of the precoder through new methods of analysis in both finite and large dimensions. Special emphasis is placed on the induced interference mitigation mechanism of iaRZF. To obtain fundamental insights, we consider the large-system regime in which the number of transmit antennas and UTs are both large. Furthermore, new finite dimensional approaches for analyzing multi cell RZF precoding schemes are introduced and applied for limiting cases. We derive deterministic expressions for the asymptotic user rates, which also serve as accurate approximations in practical non-asymptotic regimes. Merely the channel statistics are needed for calculation and implementation of our deterministic expressions. These novel expressions generalize the prior work in [20] for single cell systems and in [21] for multi cell systems, where only deterministic channel covariance matrices can be used in the analysis pertaining to the suppression of inter cell interference. Then, these extensions are used to optimize the sum-rate of the iaRZF precoding scheme in limiting cases. Insights gathered from this lead us to propose and motivate an appropriate heuristic scaling of the precoder weights w.r.t. various system parameters, that offers attractive sum-rate performance; also in non-limiting cases.

The notation in this paper adheres to the following general rules. Boldface lower case is used for column vectors, and upper case for matrices. $\mathbf{X}^{\mathrm{T}}$ and $\mathbf{X}^{\mathrm{H}}$ denote the transpose, and conjugate transpose of $\mathbf{X}$, respectively, while $\mathrm{tr}(\mathbf{X})$ is the matrix trace function. The expectation operator is denoted $\mathbb{E}[\cdot]$. The spectral norm of $\mathbf{X}$ is denoted $\|\mathbf{X}\|_2$ and the Euclidean norm of $\mathbf{x}$ is denoted $\|\mathbf{x}\|_2$. Circularly symmetric complex Gaussian random vectors are denoted $\mathcal{CN}(\bar{\mathbf{x}}, \mathbf{Q})$, where $\bar{\mathbf{x}}$ is the mean and $\mathbf{Q}$ is the covariance matrix. The set of all complex numbers is denoted by $\mathbb{C}$, with $\mathbb{C}^{N\times 1}$ and $\mathbb{C}^{N\times M}$ being the generalizations to vectors and matrices, respectively. The $M \times M$ identity matrix is written as $\mathbf{I}_M$, the zero vector of length M is denoted $\mathbf{0}_{M\times 1}$ and the zero matrix $\mathbf{0}_M$. Throughout this paper, superscripts generally refer to the origin (e.g., cell m) and subscripts generally denote the destination (e.g., cell l or UT k of cell l), when both information are needed. We employ $\perp\!\!\!\perp$ and $\not\perp\!\!\!\perp$ to mean stochastic independence and dependence, respectively.

[1]Where ξ represents some positive regularization parameter.

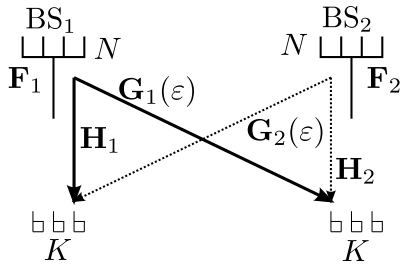

Fig. 1. Simple 2 BS downlink system.

II. Understanding iaRZF

In order to intuitively understand and motivate the iaRZF precoder this section first analyzes its behavior and impact in a relatively simple system.

A. Simple System

We start by examining a simple two cell downlink system, as depicted in Figure 1, which is a further simplification of the Wyner model [22], [23]. It features 2 BSs, BS_1 and BS_2, with N antennas each. Every BS serves one cell with K single antenna users. For convenience we introduce the notations $c = K/N$ and $\bar{x} = \mod(x,2)+1, x \in \{1,2\}$. In order to circumvent scheduling complications, we assume $N \geq K$. The aggregated channel matrix between BS_x and the affiliated users is denoted $\mathbf{H}_x = [\mathbf{h}_{x,1}, \ldots, \mathbf{h}_{x,K}] \in \mathbb{C}^{N\times K}$ and the matrix pertaining to the users of the other cell $\mathbf{G}_x(\varepsilon) = [\mathbf{g}_{x,1}, \ldots, \mathbf{g}_{x,K}] \in \mathbb{C}^{N\times K}$, which is usually abbreviated as $\mathbf{G}_x$. We treat ε as an arbitrary interference channel gain/path-loss factor. The precoding matrix used at BS_x is denoted $\mathbf{F}_x \in \mathbb{C}^{N\times K}$. For the channel realizations we choose a block-wise fast fading model, where $\mathbf{h}_{x,k} \sim \mathcal{CN}(0, \frac{1}{N}\mathbf{I}_N)$ and $\mathbf{g}_{x,k} \sim \mathcal{CN}(0, \varepsilon\frac{1}{N}\mathbf{I}_N)$ for $k = 1, \ldots, K$. The scaling factor $1/N$ is introduced for technical reasons and is further explained in Remark 1.

Denoting $\mathbf{f}_{x,k}$ the kth column of $\mathbf{F}_x$, $\mathbf{F}_{x[k]}$ as $\mathbf{F}_x$ with its kth column removed and $n_{x,k} \sim \mathcal{CN}(0,1)$ the received additive Gaussian noise at $\text{UT}_{x,k}$, we define the received signal at $\text{UT}_{x,k}$ as

$$y_{x,k} = \mathbf{h}_{x,k}^{\mathsf{H}}\mathbf{f}_{x,k}s_{x,k} + \underbrace{\mathbf{h}_{x,k}^{\mathsf{H}}\mathbf{F}_{x[k]}\mathbf{s}_{x[k]}}_{\text{intra cell interference}} + \underbrace{\mathbf{g}_{\bar{x},k}^{\mathsf{H}}\mathbf{F}_{\bar{x}}\mathbf{s}_{\bar{x}}}_{\text{inter cell interference}} + n_{x,k}$$

where $\mathbf{s}_x \sim \mathcal{CN}(0, \rho_x\mathbf{I}_N)$[2] is the vector of transmitted Gaussian symbols. It defines the average per UT transmit power of BS_x as ρ_x (normalized w.r.t. noise). The notations $\mathbf{s}_{x[k]}$ and $s_{x,k}$ designate the transmit vector without symbol k and the transmit symbol of $\text{UT}_{x,k}$.

When calculating the precoder $\mathbf{F}_x$, we assume that the channel $\mathbf{H}_x$ can be arbitrarily well estimated, however, we allow for mis-estimation of the "inter cell interference channel" $\mathbf{G}_x$ by adopting the generic Gauss-Markov formulation

$$\hat{\mathbf{G}}_x = \sqrt{1-\tau^2}\mathbf{G}_x + \tau\tilde{\mathbf{G}}_x\,.$$

[2]We remark that ρ_x is of order 1.

Choosing $\tilde{\mathbf{g}}_{x,k} \sim \mathcal{CN}(0, \varepsilon\frac{1}{N}\mathbf{I}_N)$, we can vary the available CSI quality by adjusting $0 \leq \tau \leq 1$ appropriately.

In this section we use the previously introduced iaRZF precoding scheme. Given our simple system the unnormalized precoder reads

$$\mathbf{M}_x = \left(\alpha_x\mathbf{H}_x\mathbf{H}_x^{\mathsf{H}} + \beta_x\hat{\mathbf{G}}_x\hat{\mathbf{G}}_x^{\mathsf{H}} + \xi_x\mathbf{I}\right)^{-1}\mathbf{H}_x\,. \tag{2}$$

One remarks that the regularization of the identity matrix can also be controlled by scaling α_x and β_x at the same time, if ξ_x is fixed to an arbitrary value (e.g., 1). We usually keep α_x, β_x and ξ_x in our formulas to allow easy comparison to traditional RZF formulations. We assume the following normalization of the precoder:

$$\mathbf{F}_x = \sqrt{K}\frac{\mathbf{M}_x}{\sqrt{\text{tr}\left(\mathbf{M}_x^{\mathsf{H}}\mathbf{M}_x\right)}} \tag{3}$$

i.e., the sum energy of the precoder $\text{tr}\left(\mathbf{F}_x^{\mathsf{H}}\mathbf{F}_x\right)$ is K.[3]

Remark 1 (Channel Scaling $1/N$)**.** *The statistics of the channel matrices in this section incorporate the factor $1/N$, which simplifies comparisons with the later, more general, large-scale results (see Section III). This can also be interpreted as transferring a scaling of the transmit power into the channel itself. The precoder formulations presented in the current section can be simply rewritten to fit the more traditional statistics of $\mathbf{h}_k \sim \mathcal{CN}(0, \mathbf{I}_N)$ and $\mathbf{g}_k \sim \mathcal{CN}(0, \varepsilon\mathbf{I}_N)$, by using*

$$\widetilde{\mathbf{M}}_x = \left(\alpha_x\mathbf{H}_x\mathbf{H}_x^{\mathsf{H}} + \beta_x\hat{\mathbf{G}}_x\hat{\mathbf{G}}_x^{\mathsf{H}} + N\xi_x\mathbf{I}\right)^{-1}\mathbf{H}_x$$

instead of $\mathbf{M}_x$. This equation shows that, under the chosen model, the regularization implicitly scales with N. However, one can either chose ξ or α, β appropriately, to achieve any scaling.

B. Performance of Simple System

First, we compare the general performance of the proposed iaRZF scheme with classical approaches, i.e., single cell zero-forcing (ZF), maximum-ratio transmission (MRT) and RZF. The rate of $\text{UT}_{x,k}$ can be defined as

$$r_{x,k} = \log_2\left(1 + \frac{\text{Sig}_{x,k}}{\text{Int}_{x,k}^a + \text{Int}_{x,k}^r + 1}\right)$$

where $\text{Sig}_{x,k} = \rho_x\mathbf{h}_{x,k}^{\mathsf{H}}\mathbf{f}_{x,k}\mathbf{f}_{x,k}^{\mathsf{H}}\mathbf{h}_{x,k}$, $\text{Int}_{x,k}^a = \rho_x\mathbf{h}_{x,k}^{\mathsf{H}}\mathbf{F}_{x[k]}\mathbf{F}_{x[k]}^{\mathsf{H}}\mathbf{h}_{x,k}$ and $\text{Int}_{x,k}^r = \rho_{\bar{x}}\mathbf{g}_{\bar{x},k}^{\mathsf{H}}\mathbf{F}_{\bar{x}}\mathbf{F}_{\bar{x}}^{\mathsf{H}}\mathbf{g}_{\bar{x},k}$ denote the received signal power, received intra cell interference and received inter cell interference, respectively.

For comparison we used the following precoder formulations: $\mathbf{M}_x^{\text{MRT}} = \mathbf{H}_x$, $\mathbf{M}_x^{\text{ZF}} = \mathbf{H}_x(\mathbf{H}_x^{\mathsf{H}}\mathbf{H}_x)^{-1}$, $\mathbf{M}_x^{\text{RZF}} = \mathbf{H}_x(\mathbf{H}_x^{\mathsf{H}}\mathbf{H}_x + \frac{K}{N\rho_x}\mathbf{I})^{-1}$, where the regularization in $\mathbf{M}_x^{RZF}$ is chosen according to [19], [20] and all precoders are normalized as in (3). The iaRZF weights have been chosen to be $\alpha = \beta = N\cdot\rho_x$ and $\xi = 1$, to simplify comparison with traditional RZF precoding. The corresponding performance graphs, obtained by extensive Monte-Carlo (MC) simulations

[3]It can be shown, using results from Appendix C-A by taking $\chi_i = 1\ \forall i$, that this implies $\|\mathbf{f}_{x,k}\|_2^2 \to 1$, almost surely, under Assumption 1 for the given simplified system.

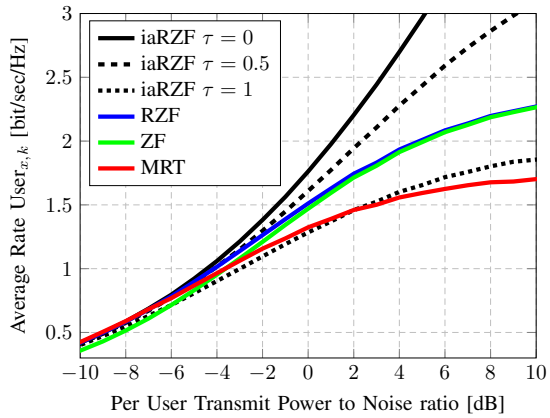

Fig. 2. Average user rate vs. transmit power to noise ratio ($N = 160$, $K = 40$, $\varepsilon = 0.7$, $\rho_1 = \rho_2 = \rho$).

in the simplified system (for $N = 160$, $K = 40$, $\varepsilon = 0.7$, $\rho_1 = \rho_2 = \rho$), are depicted in Figure 2.

We observe that iaRZF generally outperforms the other schemes. This is not surprising, as the classical schemes do not take information about the interfered UTs into account. What is surprising, however, is the gain in performance even for very bad channel estimates (see curve for $\tau = 0.5$). Only for extremely bad CSI we observe that iaRZF wastes energy due to a an improper choice of α, β. Thus, it performs worse than the other schemes, that do not take τ into account for precoding. This problem can easily be circumvented by choosing proper weights that let $\beta \to 0$ for $\tau \to 1$; as will be elaborated later on.

C. Properties of iaRZF for $\boldsymbol{\alpha_x, \beta_x \to \infty}$

As has been briefly remarked by Hoydis *et al.* in [3], the iaRZF weights α_x and β_x should, intuitively, allow to project the transmitted signal in subspaces orthogonal to the UT$_x$'s ("own users") and UT$_{\bar{x}}$'s ("other users") channels, respectively. This behavior, in the limit cases of α_x or $\beta_x \to \infty$, is treated analytically in this subsection.

1) Finite Dimensional Analysis: Limiting ourselves to finite dimensional approaches and to perfect CSI ($\tau = 0$), we can obtain the following insights.

First, we introduce the notation $\mathbf{P}^{\perp}_{\mathbf{X}}$ as a projection matrix on the space orthogonal to the column space of $\mathbf{X}$ and we remind ourselves that $\xi = 1$ is still assumed. Following the path outlined in Appendix B-A and assuming $\mathbf{H}_x^{\mathsf{H}}\mathbf{H}_x$ invertible (true with probability 1), one finds for $\alpha_x \to \infty$:

$$\alpha_x \mathbf{M}_x \overset{\alpha_x \to \infty}{\longrightarrow} \mathbf{H}_x \left(\mathbf{H}_x^{\mathsf{H}}\mathbf{H}_x\right)^{-1} - \mathbf{P}^{\perp}_{\mathbf{H}_x}\mathbf{G}_x \left(\beta_x^{-1}\mathbf{I} + \mathbf{G}_x^{\mathsf{H}}\mathbf{P}^{\perp}_{\mathbf{H}_x}\mathbf{G}_x\right)^{-1} \mathbf{G}_x^{\mathsf{H}}\mathbf{H}_x \left(\mathbf{H}_x^{\mathsf{H}}\mathbf{H}_x\right)^{-1}. \tag{4}$$

Recall that the received signal at the UTs of BS$_x$ in our simple model, due to (only) the intra cell users, is given as $\mathbf{y}_x^{\text{intra}} = \mathbf{H}_x^{\mathsf{H}}\mathbf{F}_x\mathbf{s}_x$. Thus,

$$\mathbf{y}_x^{\text{intra}} \overset{\text{Lem 2}}{=} \upsilon \mathbf{H}_x^{\mathsf{H}}\mathbf{H}_x \left(\mathbf{H}_x^{\mathsf{H}}\mathbf{H}_x\right)^{-1} \mathbf{s}_x = \upsilon \mathbf{s}_x$$

where the precoder normalization leaves a scaling factor υ that is independent of α_x. The Lemma 2 used here can be found in Appendix A. Hence, we see that for $\alpha_x \to \infty$ and β_x bounded, the precoder acts similar to a traditional ZF precoder, i.e., the intra cell interference is completely suppressed in our system. The scaling factor includes the loss of energy caused by the alignment constraint, i.e., υ decreases the more $\operatorname{span}\{\mathbf{H}_x\}$ and $\operatorname{span}\{\mathbf{G}_x\}$ intersect. It remains to mention that due to the iaRZF definition, exact ZF can only be achieved in the limit for $N = K$, where $\mathbf{H}_x(\mathbf{H}_x^{\mathsf{H}}\mathbf{H}_x)^{-1} = (\mathbf{H}_x\mathbf{H}_x^{\mathsf{H}})^{-1}\mathbf{H}_x$; assuming the inverses exist.

Looking now at the limit $\beta_x \to \infty$ and following Appendix B-A, one arrives at

$$\begin{aligned}\mathbf{M}_x &\overset{\beta_x \to \infty}{\longrightarrow} \left[\mathbf{P}^{\perp}_{\mathbf{G}_x} - \mathbf{P}^{\perp}_{\mathbf{G}_x}\mathbf{H}_x \left(\alpha_x^{-1}\mathbf{I} + \mathbf{H}_x^{\mathsf{H}}\mathbf{P}^{\perp}_{\mathbf{G}_x}\mathbf{H}_x\right)^{-1} \mathbf{H}_x^{\mathsf{H}}\mathbf{P}^{\perp}_{\mathbf{G}_x}\right]\mathbf{H}_x \\ &= \check{\mathbf{H}}\left(\mathbf{I} + \alpha\check{\mathbf{H}}^{\mathsf{H}}\check{\mathbf{H}}\right)^{-1}\end{aligned} \tag{5}$$

where we introduced $\check{\mathbf{H}} = \mathbf{P}^{\perp}_{\mathbf{G}}\mathbf{H}$, as the channel matrix $\mathbf{H}$ projected on the space orthogonal to the channels of $\mathbf{G}$. One remembers that the received signal due to inter cell interference in our simple model is given as

$$\mathbf{y}_x^{\text{inter}} = \mathbf{G}_{\bar{x}}^{\mathsf{H}}\mathbf{F}_{\bar{x}}\mathbf{s}_{\bar{x}}$$

which, via (5) in the case of $\mathbf{M}_{\bar{x}}$ and Lemma 2, directly gives $\mathbf{y}_x^{\text{inter}} = 0$. I.e., we see that for $\beta_{\bar{x}} \to \infty$ and $\alpha_{\bar{x}}$ bounded, the induced inter cell interference vanishes.

In (5), we finally see one of the main motivators for defining iaRZF, in the chosen form. Choosing $\beta_x = 0$ gives the standard single cell RZF solution; choosing $\beta_x \to \infty$ gives an intuitively reasonable RZF precoder on projected channels that makes sure no interference is induced in the other cell. It stands to reason that a sum-rate optimal solution can be found as a trade-off between these two extremes, by balancing induced interference and received signal power.

2) Large-Scale Analysis: In order to find an appropriate expression of the sum-rate that does not rely on random quantities, we anticipate results from Subsection III-E. There we find a deterministic limit to which the random values of SINR_x almost surely converge, when $N, K \to \infty$; assuming $0 < c < \infty$. We can adapt the results from Theorem 1 to fit our the current simplified model (by choosing the parameters of the generalized Theorem 1 as $L = 2, K_x = K, N_x = N, \chi_x^x = 1, \chi_{\bar{x}}^x = \varepsilon, \tau_{\bar{x}}^x = \tau, \tau_x^x = 0, \alpha_x^x = \alpha_x, \alpha_{\bar{x}}^x = \beta_x, \xi = 1$, $P_x = \rho_x$, for $x \in \{1, 2\}$). Doing so ultimately results in the following performance indicators $\text{Sig}_x \xrightarrow[N,K\to+\infty]{\text{a.s.}} \overline{\text{Sig}}_x$ and $\text{Int}_x \xrightarrow[N,K\to+\infty]{\text{a.s.}} \overline{\text{Int}}_x$, where

$$\begin{aligned}\overline{\text{Sig}}_x &= P_x\left(1 - \frac{c\alpha_x^2 e_x^2}{(1+\alpha_x e_x)^2} - \frac{c\beta_x^2\varepsilon^2 e_x^2}{(1+\beta_x\varepsilon e_x)^2}\right) \\ \overline{\text{Int}}_x &= \underbrace{P_x c \frac{1}{(1+\alpha_x e_x)^2}}_{\text{from BS } x} + \underbrace{P_{\bar{x}} c\varepsilon \frac{1+2\beta_{\bar{x}}\varepsilon\tau^2 e_{\bar{x}} + \beta_{\bar{x}}^2\varepsilon^2\tau^2 e_{\bar{x}}^2}{(1+\beta_{\bar{x}}\varepsilon e_{\bar{x}})^2}}_{\text{from BS } \bar{x}}\end{aligned} \tag{6}$$

$$\triangleq \overline{\text{Int}}_x^{\text{BS}x} + \overline{\text{Int}}_x^{\text{BS}\bar{x}}$$

$$e_x = \left(1 + \frac{c\alpha_x}{1+\alpha_x e_x} + \frac{c\beta_x\varepsilon}{1+\beta_x\varepsilon e_x}\right)^{-1} \tag{7}$$

where e_x is the unique non negative solution to the fixed point equation (7). These expressions are precise in the large-scale regime ($N, K \to \infty$, $0 < K/N < \infty$) and good approximations for finite dimensions. As a consequence of the continuous mapping theorem (e.g., [24]) the above finally implies $\mathrm{SINR}_x \xrightarrow[N,K\to+\infty]{\text{a.s.}} \overline{\mathrm{SINR}}_x = \overline{\mathrm{Sig}}_x(\overline{\mathrm{Int}}_x+1)^{-1}$.

After realizing that $0 < \liminf e_x < \limsup e_x < \infty$ for $K, N \to \infty$ (see Lemma 6), the large-scale formulations give the insights we already obtained from the finite dimensional analysis (see previous subsection). Slightly simplifying (6) to reflect the perfect CSI case ($\tau = 0$), one obtains

$$\lim_{\alpha_x\to\infty} \overline{\mathrm{Int}}_x^{\mathrm{BS}x} = \lim_{\alpha_x\to\infty} P_x c \frac{1}{(1+\alpha_x e_x)^2} = 0$$
$$\lim_{\beta_{\bar{x}}\to\infty} \overline{\mathrm{Int}}_x^{\mathrm{BS}\bar{x}} = \lim_{\beta_{\bar{x}}\to\infty} P_{\bar{x}} c \frac{\varepsilon}{(1+\beta_x\varepsilon e_x)^2} = 0$$

i.e., for $\alpha_x \to \infty$ the intra cell interference vanishes and for $\beta_{\bar{x}} \to \infty$ the induced inter cell interference vanishes. Hence, at this point we have re-obtained the results from the previous subsection, where only finite dimensional techniques were used.

The large system formulation can now also be used to look at the important practical case of mis-estimation of the channels to the other cell's users. Employing again $0 < \liminf e_x < \limsup e_x < \infty$ and (6) leads to

$$\lim_{\alpha_x\to\infty} P_x c \frac{1}{(1+\alpha_x e_x)^2} = 0$$
$$\lim_{\beta_{\bar{x}}\to\infty} P_{\bar{x}} c \frac{(\beta_x^{-2}+2\varepsilon\tau^2 e_x\beta_x^{-1}+\varepsilon^2\tau^2 e_x^2)\varepsilon}{\left(\beta_x^{-1}+\varepsilon e_x\right)^2} = P_{\bar{x}} c\tau^2\varepsilon$$

i.e., for $\alpha_x \to \infty$ the intra cell interference still vanishes, under the assumption of perfect intra cell CSI. However, for $\beta_{\bar{x}} \to \infty$ the induced inter cell interference converges to $P_{\bar{x}} c\tau^2\varepsilon$. Unsurprisingly we see that, due to imperfect CSI, the induced inter cell interference cannot be completely canceled any more.

3) Large Scale Optimization: One advantage of the large-scale approximation is the possibility to find asymptotically optimal weights for the limit behavior of iaRZF. However, to keep the calculations within reasonable effort, we limit the model to the case $P_1 = P_2 = P$ and $\alpha_1 = \alpha_2 = \alpha$, $\beta_1 = \beta_2 = \beta$. Proceeding similar to the previous subsection, we obtain a formulation for the large-scale approximation of the (now equal) SINR values, when $\alpha \to \infty$. This is denoted $\overline{\mathrm{SINR}}^{\alpha\to\infty} = \overline{\mathrm{Sig}}^{\alpha\to\infty}\left(1+\overline{\mathrm{Int}}^{\alpha\to\infty}\right)^{-1}$, where

$$\overline{\mathrm{Sig}}^{\alpha\to\infty} = P\left(1-c-\frac{c\beta^2\varepsilon^2 e^2}{(1+\beta\varepsilon e)^2}\right)$$
$$\overline{\mathrm{Int}}^{\alpha\to\infty} = Pc\varepsilon\frac{1+2\beta\varepsilon\tau^2 e+\beta^2\varepsilon^2\tau^2 e^2}{(1+\beta\varepsilon e)^2}$$

and

$$e \triangleq e^{\alpha\to\infty} = \left(1+\frac{c}{e}+\frac{c\beta\varepsilon}{1+\beta\varepsilon e}\right)^{-1}. \tag{8}$$

The optimal values of the weight β in limit case $\alpha \to \infty$ can be found by solving $\partial\overline{\mathrm{SINR}}^{\alpha\to\infty}/\partial\beta = 0$. This leads to (see Appendix B-C)

$$\beta_{opt}^{\alpha\to\infty} = \frac{P(1-\tau^2)}{Pc\varepsilon\tau^2+1}. \tag{9}$$

In other words, in the perfect CSI case ($\tau = 0$), one should choose β equal to the transmit power of the BSs. It also shows how one should scale β in between the two obvious solutions, i.e., full weight on the interfering channel information for perfect CSI and no weight under random CSI ($\tau = 1$). We remark that the interference channel gain factor ε is also implicitly included in the precoder. Thus for $\varepsilon \to 0$, we have $\beta(\mathrm{tr}\,\hat{\mathbf{G}}_x^{\mathsf{H}}\hat{\mathbf{G}}_x)^2 \to 0$, while β remains bounded. Hence no energy is wasted to precode for non-existent interference, as one would expect.

The same large-scale optimization can also be carried out for the limit $\beta \to \infty$. The SINR optimal weight for α can be found as (similar to Appendix B-C)

$$\alpha_{opt}^{\beta\to\infty} = \frac{P}{Pc\varepsilon\tau^2+1} = \frac{1}{c\varepsilon\tau^2+1/P}. \tag{10}$$

Thus, like in the perfect CSI case ($\tau = 0$), one should choose α equal to the transmit power of the BSs. The implications for other limit-cases are not so clear. We see that increasing the transmit power also increases the weight α, up to the maximum value of $1/(c\varepsilon\tau^2)$. The weight reduces as the interference worsens, i.e., when τ^2, ε grow. This makes sense, as the precoder would give more importance on the interfering channel (by indirectly increasing β via normalization). The weight is also reduced, if the cell performance is expected to be bad, i.e., c approaches 1, which makes sense from a sum-rate optimization point of view.

Finally, we can easily calculate the SINR in the limit of both α and β independently tending to infinity, as

$$\overline{\mathrm{SINR}}^{\alpha,\beta\to\infty} = \frac{P(1-2c)}{Pc\varepsilon\tau^2+1}.$$

The rationale behind all analyses in this section is that optimal weights in the limiting case often make for good heuristic approximations in more general cases. For instance, one can re-introduce the weights, found under the large-scale assumption, into the finite dimensional limit formulations. Particularly interesting for this approach is combining (9) with (4) to find an "heuristic interference aware zeroforcing" precoder:

$$\mathbf{M}_x^{iaZF} = \mathbf{H}_x\left(\mathbf{H}_x^{\mathsf{H}}\mathbf{H}_x\right)^{-1} - \mathbf{P}_{\mathbf{H}_x}^{\perp}\hat{\mathbf{G}}_x \times\left(\frac{Pc\varepsilon\tau^2+1}{P(1-\tau^2)}\mathbf{I}+\hat{\mathbf{G}}_x^{\mathsf{H}}\mathbf{P}_{\mathbf{H}_x}^{\perp}\hat{\mathbf{G}}_x\right)^{-1}\hat{\mathbf{G}}_x^{\mathsf{H}}\mathbf{H}_x\left(\mathbf{H}_x^{\mathsf{H}}\mathbf{H}_x\right)^{-1}.$$

To finish this section, we remark that doing these optimizations without the assumption of arbitrarily well estimated intra cell channels, is still an open problem.

4) Graphical Interpretation of the Results: Here we show the implications of the previous subsection on the system performance of our simple model. Of particular interest to us are comparisons of the heuristic scheme with, numerically obtained, sum-rate optimal weights.

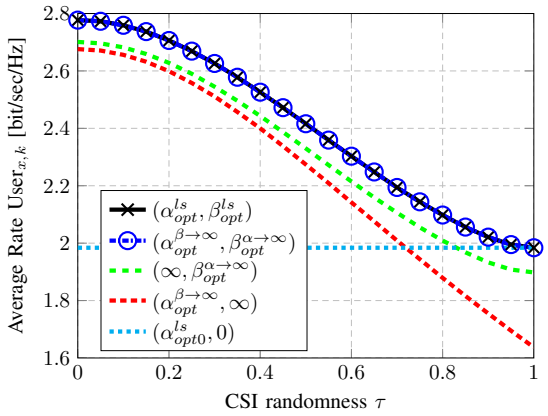

Fig. 3. Average user rate vs. CSI quality for adaptive precoder weights ($N = 160$, $K = 40$, $\varepsilon = 0.7$, $P = 10$dB).

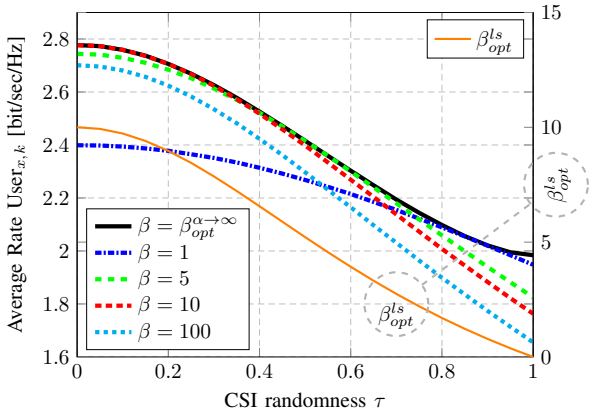

Fig. 4. Average user rate vs. CSI quality for constant precoder weights ($N = 160$, $K = 40$, $\varepsilon = 0.7$, $\alpha = \alpha_{opt}^{\beta\to\infty}$, $P = 10$dB).

In Figure 3, we analyze the average UT rate with respect to CSI randomness (τ), for different sets of precoder weights (α, β), that (mostly) adapt to the available CSI quality. The values $(\alpha_{opt}^{ls}$ and $\beta_{opt}^{ls})$ are obtained using 4D grid search, thereby renouncing the $\alpha_1 = \alpha_2 = \alpha$ and $\beta_1 = \beta_2 = \beta$ restrictions from before. Crucially, we see that the performance under $(\alpha_{opt}^{ls}, \beta_{opt}^{ls})$ and $(\alpha_{opt}^{\beta\to\infty}, \beta_{opt}^{\alpha\to\infty})$ is virtually the same. The plot also contains the pair $(\alpha_{opt0}^{ls}, 0)$, which corresponds to single cell RZF precoding. The weight α_{opt0}^{ls} is again found by grid search. The performance is constant, as the precoder does not take the interfering channel into account. However, we see that the optimally weighted iaRZF is equal to single cell RZF, when the channel estimation is purely random.

In Figure 4, we illustrate the effect of (sub-optimally, but conveniently) choosing a constant value for β. We set $\alpha = \alpha_{opt}^{\beta\to\infty}$ for all curves and also give the familiar $(\alpha_{opt}^{\beta\to\infty}, \beta_{opt}^{\alpha\to\infty})$ curve as a benchmark. Furthermore, the actual value of β_{opt}^{ls} is given on a second axis to illustrate how one would need to adapt β for optimal average rate performance. Overall one observes that a constant value for β is (unsurprisingly) only acceptable for a limited region of the CSI quality spectrum. Small values of β fit well for large τ, middle values fit well for small τ. Overly large (or small) βs do not reach optimal performance in any region.

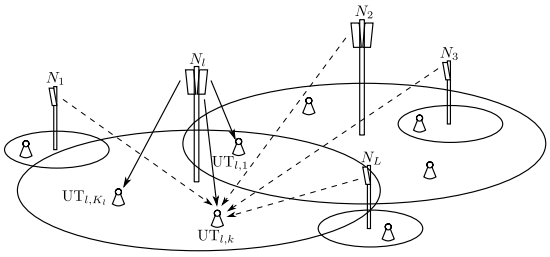

Fig. 5. Illustration of a general heterogeneous downlink system.

The encouraging performance of iaRZF using the optimal weights derived under limit assumptions, paired with the promise of simple and intuitive insights, provides motivation for the next section. There we apply the iaRZF scheme to a more general system.

III. General System for iaRZF Analysis

A. System Model

In the following, we analyze cellular downlink multi-user MIMO systems, of the more general type illustrated in Fig. 5. Each of the L cells consists of one BS associated with a number of single antenna UTs. In more detail, the lth BS is equipped with N_l transmit antennas and serves K_l UTs. We generally set $N_l \geq K_l$ in order to avoid scheduling complications. We assume transmission on a single narrow-band carrier, full transmit-buffers, and universal frequency reuse among the cells.

The lth BS transmits a data symbol vector $\mathbf{s}_l = [s_{l,1}, \ldots, s_{l,K_l}]^{\mathsf{T}}$ intended for its K_l uniquely associated UTs. This BS uses the linear precoding matrix $\mathbf{F}_l^l \in \mathbb{C}^{N_l \times K_l}$, where the columns $\mathbf{f}_{l,k}^l \in \mathbb{C}^{N_l}$ constitute the precoding vectors for each UT. We note that BSs do not directly interact with each other and users from other cells are explicitly not served. Thus, the received signal $y_{l,k} \in \mathbb{C}$ at the kth UT in cell l is

$$y_{l,k} = \sqrt{\chi_{l,k}^l}(\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{f}_{l,k}^l s_{l,k} + \sum_{k' \neq k} \sqrt{\chi_{l,k}^l}(\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{f}_{l,k'}^l s_{l,k'} + \sum_{m \neq l} \sqrt{\chi_{l,k}^m}(\mathbf{h}_{l,k}^m)^{\mathsf{H}}\mathbf{F}_m^m \mathbf{s}_m + n_{l,k}$$

where $n_{l,k} \sim \mathcal{CN}(0,1)$ is an additive noise term. The transmission symbols are chosen from a Gaussian codebook, i.e., $s_{l,k} \sim \mathcal{CN}(0,1)$. We assume block-wise small scale Rayleigh fading, thus the channel vectors are modeled as $\mathbf{h}_{l,k}^m \sim \mathcal{CN}(\mathbf{0}, \frac{1}{N_m}\mathbf{I}_{N_m})$. The path-loss and other large-scale fading effects are incorporated in the $\chi_{l,k}^m$ factors. The scaling factor $\frac{1}{N_m}$ in the fading variances is of technical nature and utilized in the asymptotic analysis. It can be canceled for a given arbitrarily sized system by modifying the transmission power accordingly; similar to Remark 1. Our setting here assumes that interference from other cells dominates w.r.t. other types of rate limitations, such as pilot contamination, which is thus not accounted for in our system model. According to recent works, this assumption is sensible if one considers practical ranges of antennas (on the order of 100s) [25], in systems

with optimized pilot-reuse [26], and also when using non-ideal hardware [27, Figure 14].

B. Imperfect Channel State Information

The UTs are assumed to perfectly estimate the respective channels to their serving BS, which enables coherent reception. This is reasonable, even for moderately fast traveling users, if proper downlink reference signals are alternated with data symbols. Generally, downlink CSI can be obtained using either a time-division duplex protocol where the BS acquires channel knowledge from uplink pilot signaling [21] and using channel reciprocity or a frequency-division duplex protocol, where temporal correlation is exploited as in [28]. In both cases, the transmitter usually has imperfect knowledge of the instantaneous channel realizations, e.g., due to imperfect pilot-based channel estimation, delays in the acquisition protocols, or user mobility. To model imperfect CSI without making explicit assumptions on the acquisition protocol, we employ the generic Gauss-Markov formulation (see, e.g., [20], [29], [30]) and we define the estimated channel vectors $\hat{\mathbf{h}}_{l,k}^m \in \mathbb{C}^{N_m}$ to be

$$\hat{\mathbf{h}}_{l,k}^m = \sqrt{\chi_{l,k}^m}\left[\sqrt{(1-(\tau_l^m)^2)}\mathbf{h}_{l,k}^m + \tau_l^m \tilde{\mathbf{h}}_{l,k}^m\right] \tag{11}$$

where $\tilde{\mathbf{h}}_{l,k}^m \sim \mathcal{CN}(0, \frac{1}{N_m}\mathbf{I}_{N_m})$ is the normalized independent estimation error. Using this formulation, we can set the accuracy of the channel acquisition between the UTs of cell l and the BS of cell m by selecting $\tau_l^m \in [0,1]$; a small value for τ_l^m implies a good estimate. Furthermore, we remark that these choices imply $\hat{\mathbf{h}}_{l,k}^m \sim \mathcal{CN}(0, \chi_{l,k}^m \frac{1}{N_m}\mathbf{I}_{N_m})$. For convenience later on, we define the aggregated estimated channel matrices as $\hat{\mathbf{H}}_l^m = [\hat{\mathbf{h}}_{l,1}^m, \ldots, \hat{\mathbf{h}}_{l,K_l}^m] \in \mathbb{C}^{N_m \times K_l}$.

C. iaRZF and Power Constraints

Following the promising results observed in Section II, we continue our analysis of the iaRZF precoding matrices $\mathbf{F}_m^m$, $m = 1, \ldots, L$, introduced in (1). For some derivations, it will turn out to be useful to restate this precoder as

$$\mathbf{F}_m^m = \left(\alpha_m^m \hat{\mathbf{H}}_m^m (\hat{\mathbf{H}}_m^m)^{\mathsf{H}} + \mathbf{Z}^m + \xi_m \mathbf{I}_{N_m}\right)^{-1} \hat{\mathbf{H}}_m^m \nu_m^{\frac{1}{2}}$$

where $\mathbf{Z}^m = \sum_{l \neq m} \alpha_l^m \hat{\mathbf{H}}_l^m (\hat{\mathbf{H}}_l^m)^{\mathsf{H}}$. The α_l^m can be considered as weights pertaining to the importance one wishes to attribute to the respective estimated channel. We remark that the regularization parameter ξ_m is usually chosen to be the number of users over the total transmit power [19] in classical RZF. The factors ν_m are used to fulfill the average per UT transmit power constraint P_m[4], pertaining to BS m:

$$\frac{1}{K_m}\mathrm{tr}\left[\mathbf{F}_m^m (\mathbf{F}_m^m)^{\mathsf{H}}\right] = P_m \,. \tag{12}$$

[4]We remark that choosing P_m of order 1 will assure proper scaling of all terms of the SINR in the following (see (16)).

D. Performance Measure

Most performance measures in cellular systems are functions of the SINRs at each UT; e.g., (weighted) sum-rate and outage probability. Under the treated system model, the expected (w.r.t. the transmitted symbols $s_{l,k}^{(l)}$) received signal power at the kth UT of cell l, i.e., $\mathrm{UT}_{l,k}$, is

$$\mathrm{Sig}_{l,k}^{(l)} = \chi_{l,k}^l (\mathbf{h}_{l,k}^l)^{\mathsf{H}} \mathbf{f}_{l,k}^l (\mathbf{f}_{l,k}^l)^{\mathsf{H}} \mathbf{h}_{l,k}^l \,. \tag{13}$$

Similarly, the interference power is

$$\begin{aligned}\mathrm{Int}_{l,k}^{(l)} = &\sum_{m \neq l} \chi_{l,k}^m (\mathbf{h}_{l,k}^m)^{\mathsf{H}} \mathbf{F}_m^m (\mathbf{F}_m^m)^{\mathsf{H}} \mathbf{h}_{l,k}^m \\ &+ \chi_{l,k}^l (\mathbf{h}_{l,k}^l)^{\mathsf{H}} \mathbf{F}_{l[k]}^l (\mathbf{F}_{l[k]}^l)^{\mathsf{H}} \mathbf{h}_{l,k}^l \end{aligned} \tag{14}$$

where

$$\mathbf{F}_{l[k]}^l = \left(\alpha_l^l \hat{\mathbf{H}}_l^l (\hat{\mathbf{H}}_l^l)^{\mathsf{H}} + \mathbf{Z}^l + \xi_l \mathbf{I}_{N_l}\right)^{-1} \hat{\mathbf{H}}_{l[k]}^l \nu_l^{\frac{1}{2}} \tag{15}$$

and $\hat{\mathbf{H}}_{l[k]}^l$ is $\hat{\mathbf{H}}_l^l$ with its kth column removed. Hence, the SINR at $\mathrm{UT}_{l,k}$ can be expressed as

$$\mathrm{SINR}_{l,k} = \mathrm{Sig}_{l,k}^{(l)} \left(\mathrm{Int}_{l,k} + 1\right)^{-1} \,. \tag{16}$$

In the following, we focus on the sum-rate, which is a commonly used performance measure utilizing the SINR values and straightforward to interpret. Under the assumption that interference is treated as noise, the ergodic sum-rate is expressed as

$$R_{sum} = \mathbb{E}\sum_{l,k} r_{l,k} = \mathbb{E}\sum_{l,k} \log(1 + \mathrm{SINR}_{l,k})$$

where SINRs are random quantities defined by the system model.

E. Deterministic Equivalent of the SINR

In order to obtain tractable and insightful expressions of the system performance, we propose a large scale approximation. This allows us to state the sum-rate expression in a deterministic and compact form that can readily be interpreted and optimized. Also, the large system approximations are accurate in both massive MIMO systems and conventional small-scale MIMO of tractable size, as will be evidenced later via simulations (see Subsection IV-B). In certain special cases, optimizations of such approximations w.r.t. many performance measures, can be carried out analytically (see for example [20]). In almost all cases, optimizations can be done numerically. We will derive a deterministic equivalent (DE) of the SINR values that allows for a large scale approximation of the sum-rate expression in (16). DEs are preferable to standard limit calculations, as they are precise in the limit case, they are also defined for finite dimensions and they provably approach the random quantity for increasing dimensions. Introducing the ratio $c_i = K_i/N_i$, we make the following technical assumption in order to obtain a DE.

Assumption A-1. *$N_i, K_i \to \infty$, such that for all i we have*

$$0 < \liminf c_i \leq \limsup c_i < \infty \,.$$

This asymptotic regime is denoted $N \to \infty$ for brevity.

In other words, we require for N_i and K_i to grow large at the same speed. By extending the analytical approach in [20] and [21] to the SINR expression in (16), we obtain a DE of the SINR, which is denoted $\overline{\mathrm{SINR}}_{l,k}$ in the following.

Theorem 1 (Deterministic Equivalent of the SINR)**.** *Under* **A-1***, we have*

$$\mathrm{SINR}_{l,k} - \overline{\mathrm{SINR}}_{l,k} \xrightarrow[N\to\infty]{\text{a.s.}} 0\,.$$

Here

$$\overline{\mathrm{SINR}}_{l,k} = \overline{\mathrm{Sig}}_{l,k}^{(l)} \left(\overline{\mathrm{Int}}_{l,k}+1\right)^{-1}$$

with

$$\overline{\mathrm{Sig}}_{l,k}^{(l)} = \overline{\nu}_l (\chi_{l,k}^l)^2 e_l^2 \left(1-(\tau_l^l)^2\right) (y_{l,k}^l)^2$$

and

$$\overline{\mathrm{Int}}_{l,k} = \sum_{m=1}^{L} \overline{\nu}_m \left(1+2x_{l,k}^m e_m + \alpha_l^m \chi_{l,k}^m x_{l,k}^m e_m^2\right) \chi_{l,k}^m g_m (y_{l,k}^m)^2$$

given $x_{l,k}^m = \alpha_l^m \chi_{l,k}^m (\tau_l^m)^2$*. The parameter* $\overline{\nu}_m$*, the abbreviations* g_m *and* $y_{l,k}^m$*, as well as the corresponding fixed-point equation* e_m *and* e'_m *are given in the following.*

First, we define e_m *to be the unique positive solution of the fixed-point equation*

$$e_m = \left(\xi_m + \frac{1}{N_m}\sum_{j=1}^{K_m} \alpha_m^m \chi_{m,j}^m y_{m,j}^m + \frac{1}{N_m}\sum_{l\neq m}\sum_{k=1}^{K_l} \alpha_l^m \chi_{l,k}^m y_{l,k}^m\right)^{-1} \tag{17}$$

where $y_{l,k}^m = \left(1+\alpha_l^m \chi_{l,k}^m e_m\right)^{-1}$*. We also have* $\overline{\nu}_m = P_m K_m / (N_m g_m)$ *with*

$$g_m = -\frac{1}{N_m}\sum_{j=1}^{K_m} \chi_{m,j}^m e'_m (y_{m,k}^m)^2$$

and e'_m *can be found directly, once* e_m *is known:*

$$e'_m = \left[\frac{1}{N_m}\sum_{j=1}^{K_m} (\alpha_m^m)^2 (\chi_{m,j}^m)^2 (y_{m,j}^m)^2 + \frac{1}{N_m}\sum_{l\neq m}\sum_{k=1}^{K_l} (\alpha_l^m)^2 (\chi_{l,k}^m)^2 (y_{l,k}^m)^2 - e_m^{-2}\right]^{-1}. \tag{18}$$

Proof. See Appendix C. □

By employing dominated convergence arguments and the continuous mapping theorem (e.g., [24]), we see that Theorem 1 implies, for each UT(l,k),

$$r_{l,k} - \log_2(1+\overline{\mathrm{SINR}}_{l,k}) \xrightarrow[N\to\infty]{\text{a.s.}} 0\,. \tag{19}$$

These results have already been used in Section II and will also serve as the basis in the following.

IV. Numerical Results

In this section we will, first, introduce a heuristic generalization of the previously found (see Paragraph II-C3) "limit-optimal" iaRZF precoder weights. Furthermore, we provide simulations that corroborate the viability of the proposed precoder, also in systems that are substantially different to the idealized system used in Section II.

A. Heuristic Generalization of Optimal Weights

Following from the encouraging sum-rate performance results in Paragraph II-C3, it seems promising to intuitively generalize the heuristic weights to systems with arbitrary many BSs, transmit powers, CSI randomness and user/antenna ratios. Adapting the structures obtained in (9) and (10), we define the general heuristic precoder weights as

$$\tilde{\alpha}_b^a = \frac{P_a(1-(\tau_b^a)^2)}{P_b c_a \varepsilon_b^a (\tau_b^a)^2 + 1}\,. \tag{20}$$

Here, we introduced the new notation ε_b^a, which is the average gain factor between BS a and the UTs of cell b, i.e., $\varepsilon_b^a = \frac{1}{K_b}\sum_k \chi_{b,k}^a$. One can intuitively understand (20) by remembering that α_b^a represents the "importance" of the associated channels (from BSa to UTs in cell b). The numerator increases the weight, i.e., makes orthogonality a priority, if the interfering BS uses a large transmit power (P_a). Also, importance is lowered for poorly estimated channels. The denominator reduces orthogonality to cells whose performance is expected to be bad, i.e., c_b approaches 1, which makes sense from a sum-rate optimization point of view. However, this aspect should be revisited, if interference mitigation is deemed more important than throughput. Weights are also lowered for cells that tolerate interference better due to large own transmit power (P_b). Poor channel estimates reduce importance again. The intuitive reason for having ε_b^a in the denominator is not immediately evident, since one would expect to place lower importance on UTs that are very far away. This behavior becomes clear over the realization, that the estimated channels in our model are not normalized (see (11)). Thus, the approximate effective weight of the precoder with respect to a normalized channel is $w_b^a = \tilde{\alpha}_b^a \varepsilon_b^a$. Hence, for $\varepsilon_b^a \to 0$, we have $w_b^a \to 0$, i.e., no importance is placed on very weak channels. Using the same deliberation for $\varepsilon_b^a \to \infty$ we have w_b^a tending to a constant value and for $\tau_b^a \to 0$ we have $w_b^a \to P_a \varepsilon_b^a$. Thus no energy is wasted on far away interferers/weak channels, even if one has perfect CSI of those channels.

We remind ourselves that in order to arrive at (20), we assumed $\xi = 1$. Furthermore, systems serving one cluster of closely located UTs per BS, reproduce the initial simplified system closely and, thus, should respond particularly well to the heuristic weights.

B. Performance

In order to verify viability of the heuristic approach, we introduce two models (see Figure 6). In the first one, two BSs are distanced 1.5 units, have a height of 0.1 units and use 160 antennas each. Around each BS, 40 single antenna

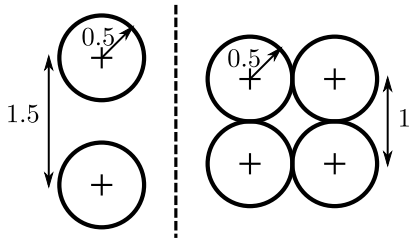

Fig. 6. Geometries of the 2 BS and 4 BS downlink models.

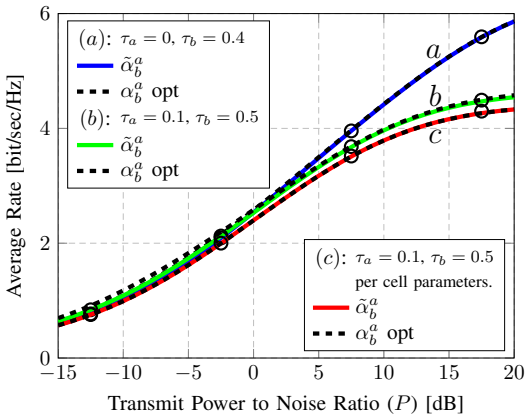

Fig. 7. 2 BSs: Average rate vs. transmit power to noise ratio ($N_x = 160$, $K_x = 40$, $P_x = P$, $(\tau_a, \tau_b) \in \{(0, 0.4), (0.1, 0.5)\}$, i.e., cases a, b, c).

UTs of height 0, are randomly (uniformly) distributed within a radius of 1 unit. Hence, one obtains clear non-overlapping clusters that are closely related to the Wyner-like simplified model in Section II. The pathloss between each BS and all UTs is defined as the inverse of the distance to the power of 2.8. The quality of CSI estimation between a BS and its associated UTs is denoted by $\tau_1^1 = \tau_2^2 = \tau_a$ and inter cell wise by $\tau_2^1 = \tau_1^2 = \tau_b$. Given the new more general 2 BS model with randomly distributed UTs for each realization, one has to consider 4 different channel weights (α_1^1, α_2^2, α_2^1 and α_1^2) in the whole system. The transmit power to noise ratio (per UT) at each BS is taken to be equal, i.e., $P_1 = P_2 = P$. For this system we obtain the average UT rate performance, shown in Figure 7. The markers denote results of MC simulations that randomize over UT placement scenarios and channel realizations, when the precoding weights are chosen as in (20). The main intention for this graph is to compare the performance under heuristic weights ($\tilde{\alpha}_b^a$) and numerically optimal weights (α_b^a opt), found via 4D grid search. We observe that the performance of both approaches is virtually the same. We also see that even when diverging from the simple system ($\tau_a = 0$, case a) by choosing $\tau_a = 0.1$, i.e., case b, the heuristic weights still perform practically the same as exhaustive numerical optimization. The same holds true for the more extreme case c, where we chose to modify the numbers of BS antennas, UTs and the CSI quality on a per cell basis ($N_1 = 160$, $N_2 = 60$, $K_1 = 40$, $N_2 = 20$, $P_x = P$, $\tau_a = 0.1$, $\tau_b = 0.5$).

Also, we look at a more complex system of 4 BSs (see

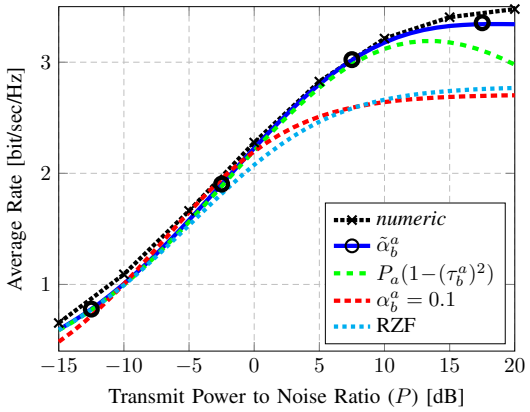

Fig. 8. 4 BSs: Average rate vs. transmit power to noise ratio ($N_x = 160$, $K_x = 40$, $P_x = P$, $(\tau_a, \tau_b, \tau_c) = (0.1, 0.3, 0.4)$).

Figure 6). The BSs, of height 0.1 units, are placed on the corners of a square with edge length 1 units. The UT distribution and pathloss are chosen as before. Figure 8 shows the performance of the 4 BS system, assuming that each BS has 160 antennas with a power constraint of P per UT and serves 40 UTs. We assume that the CSI randomness is overwhelmingly determined by inter-BS distance, i.e., we have τ_a for each BS to the adherent UTs, τ_b for each BS to UTs of BSs 1 unit away and τ_c for each BS to UTs of BSs $\sqrt{2}$ units away. It is, thus, reasonable to choose $\tau_a < \tau_b < \tau_c$. In the graph we compare the heuristic weights ($\tilde{\alpha}_b^a$) with various other weighting approaches. Round markers stem from Monte-Carlo simulations of the performance pertaining to the heuristic weights DEs and confirm the viability of the large scale approximation. The benchmark *numeric* result in this figure is obtained from optimizing the 16 precoder weights via extensive numerical search, using $\tilde{\alpha}_b^a$ as a starting point. Of practical interest is the performance of the simplified heuristics $\alpha_b^a = P_a(1-(\tau_b^a)^2)$. This choice means that no interference is taken into account, i.e., $\varepsilon_b^a = 0$. We observe that most of the viability of the heuristic method comes from this part. Only at low SNR, where interference is the dominant problem, and very high SNR does the $P_a(1-(\tau_b^a)^2)$ approach become noticeably suboptimal. The constant weight approach ($\alpha_b^a = 0.1$) behaves like in Section II, in that it is only a good match for a limited part of the curve. For comparison purposes, we also compare with standard non-cooperative RZF, as defined in Subsection II-B.

In general, employing $\tilde{\alpha}_b^a$ is most advantageous in high interference scenarios, as would be expected due to the "interference aware" conception of the precoder. Carrying out the same simulations for different levels of CSI randomness, one observes that the gain of using the heuristic variant of iaRZF is substantial as long as the estimations of the interfering channels are not too bad. For extremely bad CSI, standard non-cooperative RZF can outperform iaRZF with $\tilde{\alpha}_b^a$. We also note that better CSI widens the gap between the $\tilde{\alpha}_b^a$ and $\alpha_b^a = P_a(1-(\tau_b^a)^2)$ weighted iaRZF approaches.

V. Conclusion

In this paper, we analyzed a linear precoder structure for multi cell systems, based on an intuitive interference induction trade-off and recent results on multi cell RZF, denoted iaRZF. It was shown that the relegation of interference into orthogonal subspaces by iaRZF can be explained rigorously and intuitively, even without assuming large scale systems. For example, one can indeed observe that the precoder can either completely get rid of inter cell or intra cell interference (assuming perfect channel knowledge).

Stating and proving new results from large-scale random matrix theory, allowed us to give more conclusive and intuitive insights into the behavior of the precoder, especially with respect to imperfect CSI knowledge and induced interference mitigation. The effectiveness of these large-scale results has been demonstrated in practical finite dimensional systems. Most importantly, we concluded that iaRZF can use all available (also very bad) interference channel knowledge to obtain significant performance gains, while not requiring explicit inter base station cooperation.

Moreover, it is possible to analytically optimize the iaRZF precoder weights in certain limit scenarios using our large-scale results. Insights from this were used to propose a heuristic generalization of the limit optimal iaRZF weighting for arbitrary systems. The efficacy of the heuristic iaRZF approach has been demonstrated by achieving a sum-rate close to the numerically optimally weighted iaRZF, for a wide range of general and practical systems. The effectiveness of our heuristic approach has been intuitively explained by mainly balancing the importance of available knowledge about various channel and system variables.

Appendix A
Useful Notation and Lemmas

In this appendix we give some frequently used lemmas and definitions to facilitate exposition in the following.

Lemma 1 (Common Matrix Identities)**.** *Let* $\mathbf{A}$, $\mathbf{B}$ *be complex invertible matrices and* $\mathbf{C}$ *a rectangular complex matrix, all of proper size. We restate the following, well known, relationships:*
Woodbury Identity:

$$\left(\mathbf{A}+\mathbf{C}\mathbf{B}\mathbf{C}^{\mathsf{H}}\right)^{-1} = \mathbf{A}^{-1}-\mathbf{A}^{-1}\mathbf{C}\left(\mathbf{B}^{-1}+\mathbf{C}^{\mathsf{H}}\mathbf{A}^{-1}\mathbf{C}\right)^{-1}\mathbf{C}^{\mathsf{H}}\mathbf{A}^{-1}. \quad (21)$$

Searl Identity:

$$\left(\mathbf{I}+\mathbf{A}\mathbf{B}\right)^{-1}\mathbf{A} = \mathbf{A}\left(\mathbf{I}+\mathbf{B}\mathbf{A}\right)^{-1}. \quad (22)$$

Resolvent Identity:

$$\mathbf{A}^{-1}+\mathbf{B}^{-1} = -\mathbf{A}^{-1}\left(\mathbf{A}-\mathbf{B}\right)\mathbf{B}^{-1}. \quad (23)$$

Lemma 2 (Unitary Projection Matrices)**.** *Let* $\mathbf{X}$ *be an* $N\times K$ *complex matrix, where* $N \geq K$ *and rank*$(\mathbf{X}) = K$*. We define* $\mathbf{P}_{\mathbf{X}} = \mathbf{X}\left(\mathbf{X}^{\mathsf{H}}\mathbf{X}\right)^{-1}\mathbf{X}^{\mathsf{H}}$ *and* $\mathbf{P}_{\mathbf{X}}^{\perp} = \mathbf{I}-\mathbf{P}_{\mathbf{X}}$*. It follows (see e.g., [31, Chapter 5.13])*

$$\mathbf{P} = \mathbf{P}^2 \Leftrightarrow \mathbf{P} = \mathbf{P}^{\mathsf{H}}$$
$$\mathbf{P}_{\mathbf{X}}^{\perp}\mathbf{X} = 0 \Leftrightarrow \mathbf{X}^{\mathsf{H}}\mathbf{P}_{\mathbf{X}}^{\perp} = 0\,.$$

Generally one denotes $\mathbf{P}_{\mathbf{X}}$ *as the projection matrix onto the column space of* $\mathbf{X}$ *and* $\mathbf{P}_{\mathbf{X}}^{\perp}$ *as the projection matrix onto the orthogonal space of the column space of* $\mathbf{X}$*.*

Definition 1 (Notation of Resolvents)**.** *Given the notations from Section III, we define resolvent matrices of* $\hat{\mathbf{H}}_a^a$ *as:*

$$\mathbf{Q}_a \triangleq \left(\alpha_a^a\hat{\mathbf{H}}_a^a(\hat{\mathbf{H}}_a^a)^{\mathsf{H}}+\mathbf{Z}^a+\xi_a\mathbf{I}_{N_a}\right)^{-1}.$$

We will also use of the following modified versions

$$\mathbf{Q}_{a[bc]} \triangleq \left(\alpha_a^a\hat{\mathbf{H}}_a^a(\hat{\mathbf{H}}_a^a)^{\mathsf{H}}+\mathbf{Z}^a-\alpha_b^a\hat{\mathbf{h}}_{b,c}^a(\hat{\mathbf{h}}_{b,c}^a)^{\mathsf{H}}+\xi_a\mathbf{I}_{N_a}\right)^{-1}$$
$$\mathbf{Q}_{a[b]} \triangleq \left(\alpha_a^a\hat{\mathbf{H}}_a^a(\hat{\mathbf{H}}_a^a)^{\mathsf{H}}+\mathbf{Z}^a-\alpha_a^a\hat{\mathbf{h}}_{a,b}^a(\hat{\mathbf{h}}_{a,b}^a)^{\mathsf{H}}+\xi_a\mathbf{I}_{N_a}\right)^{-1}$$
$$= \left(\alpha_a^a\hat{\mathbf{H}}_{a[b]}^a(\hat{\mathbf{H}}_{a[b]}^a)^{\mathsf{H}}+\mathbf{Z}^a+\xi_a\mathbf{I}_{N_a}\right)^{-1}.$$

Lemma 3 (Matrix Inversion Lemma [32, Lemma 2.2])**.** *Let* $\mathbf{A}$ *be an* $M\times M$ *invertible matrix and* $\mathbf{x}\in\mathbb{C}^M, c\in\mathbb{C}$ *for which* $\mathbf{A}+c\mathbf{x}\mathbf{x}^{\mathsf{H}}$ *is invertible. Then, as an application of* (21)*, we have*

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

For the previously defined resolvent matrices, we have in particular

$$\mathbf{Q}_a\hat{\mathbf{h}}_{a,b}^a = \frac{\mathbf{Q}_{a[b]}\hat{\mathbf{h}}_{a,b}^a}{1+\alpha_a^a(\hat{\mathbf{h}}_{a,b}^a)^{\mathsf{H}}\mathbf{Q}_{a[b]}\hat{\mathbf{h}}_{a,b}^a}.$$

Lemma 4 (Convergence of Quadratic Forms [33])**.** *Let* $\mathbf{x}_M = [X_1,\ldots,X_M]^T$ *be an* $M\times 1$ *vector where the* X_n *are i.i.d. Gaussian complex random variables with unit variance. Let* $\mathbf{A}_M$ *be an* $M\times M$ *matrix independent of* $\mathbf{x}_M$*. If in addition* $\limsup_M \|\mathbf{A}\|_2 < \infty$ *then we have that*

$$\frac{1}{M}\mathbf{x}_M^H\mathbf{A}_M\mathbf{x}_M - \frac{1}{M}\operatorname{tr}(\mathbf{A}_M) \xrightarrow[M\to+\infty]{a.s.} 0\,.$$

Corollary 1. *Let* $\mathbf{A}_M$ *be as in Lemma 4, i.e.,* $\limsup_M \|\mathbf{A}\|_2 < \infty$*, and* $\mathbf{x}_M, \mathbf{y}_M$ *be random, mutually independent with complex Gaussian entries of zero mean and variance* 1*. Then we have*

$$\frac{1}{M}\mathbf{y}_M^H\mathbf{A}_M\mathbf{x}_M \xrightarrow[M\to+\infty]{a.s.} 0\,.$$

Lemma 5. *[Rank-One Perturbation Lemma [24, Lemma 14.3]] Let* $\mathbf{Q}_a$ *and* $\mathbf{Q}_{a[b]}$ *be the resolvent matrices as defined in Definition 1. Then, for any matrix* $\mathbf{A}$ *we have:*

$$\operatorname{tr}\left[\mathbf{A}\left(\mathbf{Q}_a-\mathbf{Q}_{a[b]}\right)\right] \leq \frac{1}{\xi_a}\|\mathbf{A}\|_2\,.$$

Appendix B
Simple System Limit Behavior Proofs

In this section, we provide the proofs pertaining to the limit behavior of the simple system in Section II.

A. Finite Dimensions

In order to simplify the notation we will not explicitly state the index x in the following, unless needed, hence the normalized precoder $\mathbf{F}$ for each of the two cells is $\mathbf{F} = \sqrt{K}\mathbf{M}/\sqrt{\operatorname{tr}\mathbf{M}\mathbf{M}^{\mathsf{H}}}$ for $\mathbf{M} = \left(\alpha\mathbf{H}\mathbf{H}^{\mathsf{H}}+\beta\mathbf{G}\mathbf{G}^{\mathsf{H}}+\xi\mathbf{I}\right)^{-1}\mathbf{H}$.

1) $\boldsymbol{\beta \to \infty}$: For the limit when $\beta \to \infty$ we use (21) with $\mathbf{A} = \beta\mathbf{G}\mathbf{G}^{\mathsf{H}} + \xi\mathbf{I}$ and $\mathbf{C}\mathbf{B}\mathbf{C}^{\mathsf{H}} = \mathbf{H}\alpha\mathbf{I}\mathbf{H}^{\mathsf{H}}$ to reformulate the matrix $\mathbf{M}$

$$\begin{aligned}\mathbf{M} &= \left(\alpha\mathbf{H}\mathbf{H}^{\mathsf{H}} + \beta\mathbf{G}\mathbf{G}^{\mathsf{H}} + \xi\mathbf{I}\right)^{-1}\mathbf{H} \\ &= \left[\mathbf{Q}_{\mathbf{G}} - \mathbf{Q}_{\mathbf{G}}\mathbf{H}\left(\alpha^{-1}\mathbf{I} + \mathbf{H}^{\mathsf{H}}\mathbf{Q}_{\mathbf{G}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathsf{H}}\mathbf{Q}_{\mathbf{G}}\right]\mathbf{H}\end{aligned}$$

where

$$\begin{aligned}\mathbf{Q}_{\mathbf{G}} &= \left(\beta\mathbf{G}\mathbf{G}^{\mathsf{H}} + \xi\mathbf{I}\right)^{-1} \\ &\overset{(21)}{=} \xi^{-1}\mathbf{I} - \xi^{-1}\mathbf{G}\left(\frac{\xi}{\beta}\mathbf{I} + \mathbf{G}^{\mathsf{H}}\mathbf{G}\right)^{-1}\mathbf{G}^{\mathsf{H}}.\end{aligned}$$

We now let $\beta \to \infty$, assuming $\mathbf{G}^{\mathsf{H}}\mathbf{G}$ is invertible (which is true with probability 1) and ξ bounded. In this regime, we remember Lemma 2, and rewrite $\mathbf{Q}_{\mathbf{G}} = \xi^{-1}\mathbf{P}^{\perp}_{\mathbf{G}}$. One finally arrives at

$$\begin{aligned}&\mathbf{M} \overset{\beta\to\infty}{\longrightarrow} \\ &\left[\xi^{-1}\mathbf{P}^{\perp}_{\mathbf{G}} - \xi^{-2}\mathbf{P}^{\perp}_{\mathbf{G}}\mathbf{H}\left(\alpha^{-1}\mathbf{I} + \xi^{-1}\mathbf{H}^{\mathsf{H}}\mathbf{P}^{\perp}_{\mathbf{G}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathsf{H}}\mathbf{P}^{\perp}_{\mathbf{G}}\right]\mathbf{H}.\end{aligned}$$

Relying further on properties of projection matrices ($\mathbf{P}^{\perp}_{\mathbf{G}} = \mathbf{P}^{\perp}_{\mathbf{G}}\mathbf{P}^{\perp}_{\mathbf{G}}$, $(\mathbf{P}^{\perp}_{\mathbf{G}})^{\mathsf{H}} = \mathbf{P}^{\perp}_{\mathbf{G}}$) and introducing the matrix $\check{\mathbf{H}} = \mathbf{P}^{\perp}_{\mathbf{G}}\mathbf{H}$, as the channel matrix $\mathbf{H}$ projected on the space orthogonal to the channels of $\mathbf{G}$, we get

$$\begin{aligned}&\mathbf{M} \overset{\beta\to\infty}{\longrightarrow} \\ &\xi^{-1}\left[\mathbf{P}^{\perp}_{\mathbf{G}}\mathbf{H} - \mathbf{P}^{\perp}_{\mathbf{G}}\mathbf{H}\left(\frac{\xi}{\alpha}\mathbf{I} + \mathbf{H}^{\mathsf{H}}\mathbf{P}^{\perp}_{\mathbf{G}}\mathbf{P}^{\perp}_{\mathbf{G}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathsf{H}}\mathbf{P}^{\perp}_{\mathbf{G}}\mathbf{P}^{\perp}_{\mathbf{G}}\mathbf{H}\right] \\ &= \xi^{-1}\left[\check{\mathbf{H}} - \check{\mathbf{H}}\left(\mathbf{I} - \frac{\xi}{\alpha}\left(\frac{\xi}{\alpha}\mathbf{I} + \check{\mathbf{H}}^{\mathsf{H}}\check{\mathbf{H}}\right)^{-1}\right)\right] \\ &= \check{\mathbf{H}}\left(\xi\mathbf{I} + \alpha\check{\mathbf{H}}^{\mathsf{H}}\check{\mathbf{H}}\right)^{-1}.\end{aligned}$$

2) $\boldsymbol{\alpha \to \infty}$: Introducing the abbreviations $\mathbf{Q}_{\mathbf{H}} = \left(\mathbf{H}\mathbf{H}^{\mathsf{H}} + \frac{\xi}{\alpha}\mathbf{I}\right)^{-1}$ and $\bar{\mathbf{Q}}_{\mathbf{H}} = \left(\mathbf{H}^{\mathsf{H}}\mathbf{H} + \frac{\xi}{\alpha}\mathbf{I}\right)^{-1}$, we can rewrite the matrix $\mathbf{M}$ as follows.

$$\begin{aligned}\alpha\mathbf{M} &= \left(\mathbf{H}\mathbf{H}^{\mathsf{H}} + \frac{\beta}{\alpha}\mathbf{G}\mathbf{G}^{\mathsf{H}} + \frac{\xi}{\alpha}\mathbf{I}\right)^{-1}\mathbf{H} \\ &\overset{(21)}{=} \left[\mathbf{Q}_{\mathbf{H}} - \mathbf{Q}_{\mathbf{H}}\mathbf{G}\left(\frac{\alpha}{\beta}\mathbf{I} + \mathbf{G}^{\mathsf{H}}\mathbf{Q}_{\mathbf{H}}\mathbf{G}\right)^{-1}\mathbf{G}^{\mathsf{H}}\mathbf{Q}_{\mathbf{H}}\right]\mathbf{H} \\ &\overset{(22)}{=} \mathbf{H}\bar{\mathbf{Q}}_{\mathbf{H}} - \mathbf{Q}_{\mathbf{H}}\mathbf{G}\left(\frac{\alpha}{\beta}\mathbf{I} + \mathbf{G}^{\mathsf{H}}\mathbf{Q}_{\mathbf{H}}\mathbf{G}\right)^{-1}\mathbf{G}^{\mathsf{H}}\mathbf{H}\bar{\mathbf{Q}}_{\mathbf{H}}.\end{aligned}$$

Applying (23) to the expression $\left(\mathbf{H}\mathbf{H}^{\mathsf{H}} + \frac{\xi}{\alpha}\mathbf{I}\right)^{-1} + \left(-\frac{\xi}{\alpha}\mathbf{I}\right)^{-1}$, one eventually finds the relationship $\mathbf{Q}_{\mathbf{H}} = \alpha\xi^{-1}\left(\mathbf{I} - \mathbf{H}\bar{\mathbf{Q}}_{\mathbf{H}}\mathbf{H}^{\mathsf{H}}\right)$. Hence,

$$\begin{aligned}\alpha\mathbf{M} =& \mathbf{H}\bar{\mathbf{Q}}_{\mathbf{H}} - \xi^{-1}\left(\mathbf{I} - \mathbf{H}\bar{\mathbf{Q}}_{\mathbf{H}}\mathbf{H}^{\mathsf{H}}\right) \\ &\times\mathbf{G}\left[\frac{1}{\beta}\mathbf{I} + \xi^{-1}\mathbf{G}^{\mathsf{H}}\left(\mathbf{I} - \mathbf{H}\bar{\mathbf{Q}}_{\mathbf{H}}\mathbf{H}^{\mathsf{H}}\right)\mathbf{G}\right]^{-1}\mathbf{G}^{\mathsf{H}}\mathbf{H}\bar{\mathbf{Q}}_{\mathbf{H}}.\end{aligned}$$

Now, taking the limit of $\alpha \to \infty$, assuming $\mathbf{H}^{\mathsf{H}}\mathbf{H}$ invertible (true with probability 1), and recognizing $\mathbf{P}^{\perp}_{\mathbf{H}} = \mathbf{I} - \mathbf{H}\left(\mathbf{H}^{\mathsf{H}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathsf{H}}$ we arrive at

$$\begin{aligned}\alpha\mathbf{M} &\overset{\alpha\to\infty}{\longrightarrow} \mathbf{H}\left(\mathbf{H}^{\mathsf{H}}\mathbf{H}\right)^{-1} - \gamma^{-1}\left[\mathbf{I} - \mathbf{H}\left(\mathbf{H}^{\mathsf{H}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathsf{H}}\right]\mathbf{G} \\ &\quad\left\{\beta^{-1}\mathbf{I} + \gamma^{-1}\mathbf{G}^{\mathsf{H}}\left[\mathbf{I} - \mathbf{H}\left(\mathbf{H}^{\mathsf{H}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathsf{H}}\right]\mathbf{G}\right\}^{-1} \\ &\quad\mathbf{G}^{\mathsf{H}}\mathbf{H}\left(\mathbf{H}^{\mathsf{H}}\mathbf{H}\right)^{-1} \\ &= \mathbf{H}\left(\mathbf{H}^{\mathsf{H}}\mathbf{H}\right)^{-1} \\ &\quad - \gamma^{-1}\mathbf{P}^{\perp}_{\mathbf{H}}\mathbf{G}\left\{\beta^{-1}\mathbf{I} + \gamma^{-1}\mathbf{G}^{\mathsf{H}}\mathbf{P}^{\perp}_{\mathbf{H}}\mathbf{G}\right\}^{-1}\mathbf{G}^{\mathsf{H}}\mathbf{H}\left(\mathbf{H}^{\mathsf{H}}\mathbf{H}\right)^{-1}.\end{aligned}$$

B. Large-Scale Approximation

In this subsection, we primarily show that the fixed point equation e is bounded in the sense of $0 < \liminf e < \limsup e < \infty$. This knowledge simplifies the limit calculations in Subsection II-C to simple operations. We remind ourselves, that for perfect and imperfect CSI the resulting fixed point equations are equivalent:

$$e = \left(1 + \frac{c}{\alpha^{-1} + e} + \frac{c\varepsilon}{\beta^{-1} + \varepsilon e}\right)^{-1} \tag{24}$$

where we abbreviated e_α with e for notational convenience.

Lemma 6 (e is Bounded)**.** *For either $\alpha \to \infty$ and β, ε bounded or $\beta \to \infty$ and α, ε bounded, we have*

$$0 < \liminf e < \limsup e < \infty.$$

Proof. 1) $e < \infty$ when α or $\beta \to \infty$, follows immediately from contradiction, when one takes $e \to \infty$ in (24).

2) To see that e positive when α or $\beta \to \infty$, we take either $\alpha \to \infty$ and β, ε bounded or $\beta \to \infty$ and α, ε bounded. For the case $\alpha \to \infty$, we first denote $\upsilon = \alpha e$ and we look at

$$\upsilon = \left(\frac{1}{\alpha} + \frac{c}{1+\upsilon} + \frac{c\beta\varepsilon}{\alpha + \beta\varepsilon\upsilon}\right)^{-1}.$$

Now we assume υ to be bounded for $\alpha \to \infty$

$$\upsilon = \lim_{\alpha\to\infty}\left(\frac{1}{\alpha} + \frac{c}{1+\upsilon} + \frac{c\beta\varepsilon}{\alpha + \beta\varepsilon\upsilon}\right)^{-1} = \left(\frac{c}{1+\upsilon}\right)^{-1}$$

thus implying $\upsilon = \frac{1}{c-1} < 0$, as $c < 1$. Case 1 directly contradicts the assumption and case 2 is contradicting, as e can not be negative for positive values of α, β, c and ε. Thus, υ is not bounded for $\alpha \to \infty$, hence e can neither be zero nor negative. For the case of $\beta \to \infty$, we denote $\upsilon = \beta e$ and proceed analogously. □

C. Large-Scale Optimization $\alpha \to \infty$

Continuing from Appendix B-B, we see that in the limit $\alpha \to \infty$ the large-scale approximation of the SINR values pertaining to the users of each cell, i.e., $\overline{\mathrm{SINR}}^{\alpha\to\infty}$, is indeed as stated in Paragraph II-C3.

Differentiating $\overline{\mathrm{SINR}}^{\alpha\to\infty}$ w.r.t. β, while taking into account that e is an abbreviation for $e_\beta^{\alpha\to\infty}$ leads us to

$$\begin{aligned}&\frac{\partial\overline{\mathrm{SINR}}^{\alpha\to\infty}}{\partial\beta} = -2Pc\varepsilon^2\left[e + \beta e'\right] \\ &\times\frac{t_1}{\left[P\left(c\beta^2e^2\varepsilon^3\tau^2 + 2c\beta e\varepsilon^2\tau^2 + c\varepsilon\right) + \beta^2e^2\varepsilon^2 + 2\beta e\varepsilon + 1\right]^2}\end{aligned} \tag{25}$$

where we used e' as shorthand for $\frac{\partial e^{\alpha\to\infty}(\beta)}{\partial\beta}$ and

$$t_1 = P\left[c-1-\beta\varepsilon e+2\beta c\varepsilon e\right]+\beta e+\beta^2e^2\varepsilon \\ -P_{\bar{x}}\tau^2\left[c-1-\beta\varepsilon e+\beta c\varepsilon e-\beta^2ce^2\varepsilon^2\right].$$

Realizing that the denominator of (25) can not become zero, we have two possible solutions for $\partial\overline{\mathrm{SINR}}^{\alpha\to\infty}/\partial\beta = 0$. In Lemma 7 we show that $e+\beta e' > 0$, hence we only need to deal with the term t_1. We remember from (8) that $c-1-\beta\varepsilon e+2\beta c\varepsilon e+e+\beta\varepsilon e^2 = 0$. Thus,

$$c-1-\beta\varepsilon e+\beta c\varepsilon e-\beta^2ce^2\varepsilon^2 = -\beta c\varepsilon e-e-\beta\varepsilon e^2-\beta^2ce^2\varepsilon^2$$

and similarly

$$P\left[c-1-\beta\varepsilon e+2\beta c\varepsilon e\right]+\beta e+\beta^2e^2\varepsilon \\ = -Pe-P\beta\varepsilon e^2+\beta e+\beta^2\varepsilon e^2.$$

Hence,

$$t_1 = \left(\varepsilon e^2+P\tau^2ce^2\varepsilon^2\right)\left(\beta-\frac{P(1-\tau^2)}{Pc\varepsilon\tau^2+1}\right)\left(\beta+\frac{1}{e\varepsilon}\right).$$

Given that only the middle term can become zero, we find β_{opt} to be

$$\beta_{opt} = \frac{P(1-\tau^2)}{Pc\varepsilon\tau^2+1} \tag{26}$$

as stated in (9). The physical interpretation of the SINR guarantees this point to be the maximum.

We used the assumption $e+\beta e' > 0$ to arrive at the previous result. This claim is proved by the following lemma.

Lemma 7. *Given the notation and definitions from Appendix B-C, we have that* $e+\beta e' > 0$.

Proof Sketch. From [32] we know that an object of the form

$$m(z) = \left[-z+c\int\frac{t}{1+tm(z)}d\nu(t)\right]^{-1}$$

where ν is a non negative finite measure, is a so-called Stieltjes transform of a measure ν, defined $\forall z \notin \mathrm{Supp}(\nu)$. Adapting (24) by re-naming $\tilde{e}\triangleq\beta\varepsilon e$ we see that it is indeed a valid Stieltjes transform for an appropriately chosen measure. Finally, one recognizes $\beta e'+e$ as the derivative of a Stieltjes transform, which is always positive. □

Appendix C
Proof of Theorem 1

The objective of this section is to find a DE for the SINR term (16). A broad outline of the required steps is as follows. In the beginning of the proof we condition that $\mathbf{Z}^m$ is fixed to some realization and we follow the steps given in [20, Appendix II] for the power normalization ν_m. Invoking [20, Theorem 1] we obtain the fundamental equations for e_m. We, then, allow $\mathbf{Z}^m$ to be random and apply [24, Theorem 3.13] to obtain (17). Invoking Tonelli's theorem, it is admissible to apply the two theorems one after the other, as $\mathbf{Z}^m$ is a bounded sequence with probability one. The DEs of all required terms are found by following [20, Appendix II] again. This is true for the terms from Subsection III-D, as well. However here the interference terms ask for a slightly more generalized version of [20, Lemma 7].

A. Power Normalization Term

We start by finding a DE of the term ν_m, which will turn out to be a frequently reoccurring object throughout this Section. From (12), we see that the power normalization term ν_m is defined by the relationship

$$\frac{P_m}{\nu_m}\frac{K_m}{N_m} = \frac{1}{N_m}\mathrm{tr}\left[\hat{\mathbf{H}}_m^m(\hat{\mathbf{H}}_m^m)^{\mathsf{H}}\mathbf{Q}_m^2\right] \\ = \frac{\partial}{\partial\xi_m}\left\{\frac{1}{\alpha_m^mN_m}\mathrm{tr}\left[\left(\mathbf{Z}^m+\xi_m\mathbf{I}_{N_m}\right)\mathbf{Q}_m\right]\right\} \tag{27}$$

where we used the general identities $\frac{\partial}{\partial y}\left\{-\mathrm{tr}\left[\mathbf{A}\left(\mathbf{A}+\mathbf{B}+y\mathbf{I}\right)^{-1}\right]\right\} = \mathrm{tr}\left[\mathbf{A}\left(\mathbf{A}+\mathbf{B}+y\mathbf{I}\right)^{-2}\right]$ and $\mathbf{A}\left(\mathbf{A}+\mathbf{B}+y\mathbf{I}\right)^{-1} = \mathbf{I}-\left(\mathbf{B}+y\mathbf{I}\right)\left(\mathbf{A}+\mathbf{B}+y\mathbf{I}\right)^{-1}$. The goal now is to find a deterministic object $\bar{X}_m$ that satisfies

$$\frac{1}{N_m}\mathrm{tr}\left[\hat{\mathbf{H}}_m^m(\hat{\mathbf{H}}_m^m)^{\mathsf{H}}\mathbf{Q}_m^2\right]-\bar{X}_m \xrightarrow[N\to\infty]{\text{a.s.}} 0$$

for the regime defined in A-1.

To do this, we apply [20, Theorem 1] to (27), where we set the respective variables to be $\mathbf{\Psi}_i = \chi_{m,i}^m\mathbf{I}$, $\mathbf{Q}_N = \mathbf{Z}^m+\xi_m\mathbf{I}_{N_m}$, $\mathbf{B}_N = \alpha_m^m\hat{\mathbf{H}}_m^m(\hat{\mathbf{H}}_m^m)^{\mathsf{H}}+\mathbf{Z}^m$ and $z = -\xi_m$. Thus, we find the (partially deterministic) quantity

$$\bar{X}_m = \frac{\partial}{\partial\xi_m}\frac{1}{\alpha_m^mN_m}\mathrm{tr}\Bigg[\left(\mathbf{Z}^m+\xi_m\mathbf{I}_{N_m}\right) \\ \left(\frac{1}{N_m}\sum_{j=1}^{K_m}\frac{\alpha_m^m\chi_{m,j}^m\mathbf{I}_{N_m}}{1+e_m^j}+\mathbf{Z}^m+\xi_m\mathbf{I}_{N_m}\right)^{-1}\Bigg]$$

where $e_m^j = \alpha_m^m\chi_{m,j}^me_m$ and

$$e_m = \frac{1}{N_m}\mathrm{tr}\left(\frac{1}{N_m}\sum_{j=1}^{K_m}\frac{\alpha_m^m\chi_{m,j}^m\mathbf{I}_{N_m}}{1+\alpha_m^m\chi_{m,j}^me_m}+\mathbf{Z}^m+\xi_m\mathbf{I}_{N_m}\right)^{-1}.$$

Remark 2. *In order to reuse the results from this section later on, it will turn out to be useful to realize the following relationship involving* e_m.

$$\frac{1}{N_m}\mathrm{tr}\,\mathbf{Q}_m-e_m \xrightarrow[N\to\infty]{\text{a.s.}} 0. \tag{28}$$

This can be quickly verified by using [20, Theorem 1], via choosing $\tilde{\mathbf{R}}_i = \chi_{m,i}^m\mathbf{I}$, $\mathbf{D}_N = \mathbf{I}$, $\mathbf{B}_N = \alpha_m^m\hat{\mathbf{H}}_m^m(\hat{\mathbf{H}}_m^m)^{\mathsf{H}}+\mathbf{Z}^m$ *and* $z = -\xi_m$.

One notices, that the fixed-point equation e_m contains the term $\mathbf{Z}^m$, which is not deterministic. Thus, the derived objects are not yet DEs. In order to resolve this we need to condition $\mathbf{Z}^m$ to be fixed, for now. Under this assumption we now find the DE of e_m. To do this, it is necessary to realize that e_m contains another Stieltjes transform:

$$e_m = \frac{1}{N_m}\mathrm{tr}\left[\left(\mathbf{Z}^m+\beta_m\mathbf{I}_{N_m}\right)^{-1}\right]$$

where

$$\beta_m = \frac{1}{N_m}\sum_{j=1}^{K_m}\frac{\alpha_m^m\chi_{m,j}^m}{1+\alpha_m^m\chi_{m,j}^me_m}+\xi_m. \tag{29}$$

The solution becomes immediate once we rephrase $\mathbf{Z}^m$ as

$$\mathbf{Z}^m = \sum_{l\neq m}\sum_{k=1}^{K_l} \alpha_l^m \hat{\mathbf{h}}_{l,k}^m (\hat{\mathbf{h}}_{l,k}^m)^{\mathsf{H}} = \check{\mathbf{H}}_{[m]}^m \mathbf{A}_{[m]}^m \left(\check{\mathbf{H}}_{[m]}^m\right)^{\mathsf{H}}$$

where $\check{\mathbf{H}}_{[m]}^m \in \mathbb{C}^{N_m \times K_{[m]}}$, with $K_{[m]} = \sum_{l\neq m} K_l$, is the aggregated matrix of the vectors $\check{\mathbf{h}}_{l,k}^m \sim \mathcal{CN}(0, \frac{1}{N_m}\mathbf{I}_{N_m}), \forall\, l \neq m$ and

$$\begin{aligned}\mathbf{A}_{[m]}^m = \operatorname{diag}\Big[&\alpha_1^m\chi_{1,1}^m, \ldots, \alpha_1^m\chi_{1,K_1}^m, \alpha_2^m\chi_{2,1}^m, \ldots,\\ &\alpha_2^m\chi_{2,K_2}^m, \cdots, \alpha_{m-1}^m\chi_{m-1,K_{m-1}}^m,\\ &\alpha_{m+1}^m\chi_{m+1,1}^m, \cdots, \alpha_L^m\chi_{L,K_L}^m\Big]\end{aligned}$$

i.e., a diagonal matrix with the terms pertaining to α_m^m removed.

One can directly apply [32] or [24][Theorem 3.13, Eq 3.23] with $\mathbf{T} = \mathbf{A}_{[m]}^m$ and $\mathbf{X} = (\check{\mathbf{H}}_{[m]}^m)^{\mathsf{H}}$. Being careful with the notation ($\mathbf{XTX}^{\mathsf{H}}$ instead of $(\check{\mathbf{H}}_{[m]}^m)^{\mathsf{H}}\mathbf{A}_{[m]}^m\check{\mathbf{H}}_{[m]}^m$), we arrive at:

$$e_m = \frac{1}{N_m}\operatorname{tr}\left\{\left[\check{\mathbf{H}}_{[m]}^m \mathbf{A}_{[m]}^m (\check{\mathbf{H}}_{[m]}^m)^{\mathsf{H}} + \beta_m \mathbf{I}_{N_m}\right]^{-1}\right\}$$

where

$$e_m - \frac{1}{N_m}\left[\beta_m + \frac{1}{N_m}\sum_{l\neq m}^{L}\sum_{k}^{K_l}\frac{\alpha_l^m\chi_{l,k}^m}{1+\alpha_l^m\chi_{l,k}^m e_m}\right]^{-1} \xrightarrow[N\to\infty]{\text{a.s.}} 0\,.$$

Here we used Remark 2 and β_m is given in (29).

Combining the intermediate results, using Remark 2 and the relationship $\operatorname{tr}\mathbf{A}\left(\mathbf{A}+x\mathbf{I}\right)^{-1} = \operatorname{tr}\mathbf{I} - x\operatorname{tr}\left(\mathbf{A}+x\mathbf{I}\right)^{-1}$ with $\mathbf{A} = \mathbf{Z}^m + \xi_m\mathbf{I}_{N_m}$, we arrive at

$$\bar{X}_m = -\frac{1}{\alpha_m^m N_m}\sum_{j=1}^{K_m}\frac{\alpha_m^m\chi_{m,j}^m e_m'}{(1+\alpha_m^m\chi_{m,j}^m e_m)^2}$$

where e_m' is shorthand for $\partial/\partial\xi_m e_m$ and can found (by prolonged calculus) to be as stated in (18), which concludes this part of the proof.

B. Signal Power Term

The important part of finding the DE of the signal power term (13) to find a DE of $(\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{Q}_l\hat{\mathbf{h}}_{l,k}^l$, which will now first be done. Before proceeding, we remind ourselves that our chosen model of the estimated channel (11) entails the following relationships: $\mathbf{h}_{l,k}^l \perp\!\!\!\perp \tilde{\mathbf{h}}_{l,k}^l$, $\hat{\mathbf{h}}_{l,k}^l \not\perp\!\!\!\perp \mathbf{h}_{l,k}^l$, $\hat{\mathbf{h}}_{l,k}^l \not\perp\!\!\!\perp \tilde{\mathbf{h}}_{l,k}^l$, $\mathbf{Q}_{l[k]} \perp\!\!\!\perp \hat{\mathbf{h}}_{l,k}^l$, $\mathbf{Q}_{l[k]} \perp\!\!\!\perp \mathbf{h}_{l,k}^l$. Also, formulations containing $\hat{\mathbf{h}}_{l,k}^l$ can often be split into two terms comprising $\mathbf{h}_{l,k}^l$ and $\tilde{\mathbf{h}}_{l,k}^l$. Hence, the application of Lemmas 3, 4, 5 and Corollary 1, in the following is well justified. Employing (28) one sees

$$(\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{Q}_l\hat{\mathbf{h}}_{l,k}^l - \frac{\sqrt{\chi_{l,k}^l}\sqrt{(1-(\tau_l^l)^2)}e_{(l)}}{1+\alpha_l^l\chi_{l,k}^l e_{(l)}} \xrightarrow[N\to\infty]{\text{a.s.}} 0\,.$$

Finally, applying this result to the complete formulation (13), we arrive at the familiar term from Theorem 1:

$$\overline{\operatorname{Sig}}_{l,k}^{(l)} = \overline{\nu}_l(\chi_{l,k}^l)^2 e_{(l)}^2\left(1-(\tau_l^l)^2\right)(f_{l,k}^l)^2\,.$$

C. Preparation for Interference Terms

In this subsection we derive the deterministic equivalents of the two terms $(\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{B}\mathbf{Q}_l\mathbf{h}_{l,k}^l$ and $(\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{B}\mathbf{Q}_l\tilde{\mathbf{h}}_{l,k}^l$, where $\mathbf{B} \in \mathbb{C}^{N_l\times N_l}$ has uniformly bounded spectral norm w.r.t. N_l and is independent of $\mathbf{h}_{l,k}^l$ and $\tilde{\mathbf{h}}_{l,k}^l$. The following approach is based on and slightly generalizes [20, Lemma 7]. First, it is helpful to realize an implication of our resolvent notation (Definition 1) and channel estimation model (11):

$$\begin{aligned}\mathbf{Q}_a^{-1} - \mathbf{Q}_{a[bc]}^{-1} = {}& c_0\mathbf{h}_{b,c}^a(\mathbf{h}_{b,c}^a)^{\mathsf{H}} + c_2\mathbf{h}_{b,c}^a(\tilde{\mathbf{h}}_{b,c}^a)^{\mathsf{H}} +\\ & c_2\tilde{\mathbf{h}}_{b,c}^a(\mathbf{h}_{b,c}^a)^{\mathsf{H}} + c_1\tilde{\mathbf{h}}_{b,c}^a(\tilde{\mathbf{h}}_{b,c}^a)^{\mathsf{H}}\end{aligned} \tag{30}$$

where $c_0 = \alpha_b^a\chi_{b,c}^a\left(1-(\tau_b^a)^2\right)$, $c_1 = \alpha_b^a\chi_{b,c}^a(\tau_b^a)^2$ and $c_2 = \alpha_b^a\chi_{b,c}^a\sqrt{(1-(\tau_b^a)^2)}\tau_b^a$. We omitted designating the dependencies of c on a and b, as this is always clear from the context. To ease the exposition, we also introduce the following abbreviations

$$\begin{aligned} Y_1 &\triangleq (\tilde{\mathbf{h}}_{l,k}^l)^{\mathsf{H}}\mathbf{Q}_{l[k]}\mathbf{h}_{l,k}^l & Y_4 &\triangleq (\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{B}\mathbf{Q}_{l[k]}\mathbf{h}_{l,k}^l\\ Y_2 &\triangleq (\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{Q}_{l[k]}\tilde{\mathbf{h}}_{l,k}^l & Y_5 &\triangleq (\tilde{\mathbf{h}}_{l,k}^l)^{\mathsf{H}}\mathbf{Q}_{l[k]}\tilde{\mathbf{h}}_{l,k}^l\\ Y_3 &\triangleq (\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{B}\mathbf{Q}_{l[k]}\tilde{\mathbf{h}}_{l,k}^l & Y_6 &\triangleq (\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{Q}_{l[k]}\mathbf{h}_{l,k}^l\,.\end{aligned}$$

Finally, we begin with the term $(\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{B}\mathbf{Q}_l\tilde{\mathbf{h}}_{l,k}^l$:

$$\begin{aligned}(\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{B}\mathbf{Q}_l\tilde{\mathbf{h}}_{l,k}^l - (\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{B}\mathbf{Q}_{l[k]}\tilde{\mathbf{h}}_{l,k}^l &\overset{(23)}{=}\\ -(\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{B}\mathbf{Q}_l\left(\mathbf{Q}_l^{-1} - \mathbf{Q}_{l[k]}^{-1}\right)\mathbf{Q}_{l[k]}\tilde{\mathbf{h}}_{l,k}^l\end{aligned}$$

and, using (30), we find

$$(\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{B}\mathbf{Q}_l\tilde{\mathbf{h}}_{l,k}^l = \frac{Y_3 - (\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{B}\mathbf{Q}_l\mathbf{h}_{l,k}^l\,(c_0Y_2 + c_2Y_5)}{1+c_2Y_2+c_1Y_5}\,. \tag{31}$$

Similarly, for the term $(\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{B}\mathbf{Q}_l\mathbf{h}_{l,k}^l$ we arrive at

$$\begin{aligned}&(\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{B}\mathbf{Q}_l\mathbf{h}_{l,k}^l\,(1+c_0Y_6+c_2Y_1)\\ &\quad = Y_4 - (\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{B}\mathbf{Q}_l\tilde{\mathbf{h}}_{l,k}^l\,(c_2Y_5+c_1Y_1)\,.\end{aligned} \tag{32}$$

Now, applying (31) to (32), one arrives at

$$\begin{aligned}&(\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{B}\mathbf{Q}_l\mathbf{h}_{l,k}^l\\ &\times\left[(1+c_0Y_6+c_2Y_1) - \frac{(c_0Y_2+c_2Y_5)\,(c_2Y_6+c_1Y_1)}{1+c_2Y_2+c_1Y_5}\right]\\ &\quad = Y_4 - \frac{(\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{B}\mathbf{Q}_l\tilde{\mathbf{h}}_{l,k}^l\,(c_2Y_6+c_1Y_1)}{1+c_2Y_2+c_1Y_5}\,.\end{aligned} \tag{33}$$

Similar to Appendix C-B, we notice that Y_1, Y_2 and Y_3, converge almost surely to 0 in the large system limit:

$$Y_1, Y_2, Y_3 \xrightarrow[N\to\infty]{\text{a.s.}} 0\,.$$

We also foresee that

$$Y_4 - u' \xrightarrow[N\to\infty]{\text{a.s.}} 0\,,\quad Y_5 - u_1 \xrightarrow[N\to\infty]{\text{a.s.}} 0\,,\quad Y_6 - u_2 \xrightarrow[N\to\infty]{\text{a.s.}} 0$$

where the values for u', u_1 and u_2 are not yet of concern. Thus, (33) finally leads to

$$(\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{B}\mathbf{Q}_l\mathbf{h}_{l,k}^l\left[(1+c_0u_2) - \frac{(c_2u_1)\,(c_2u_2)}{1+c_1u_1}\right] - u' \xrightarrow[N\to\infty]{\text{a.s.}} 0$$

and we finally find the expression we were looking for

$$(\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{B}\mathbf{Q}_l\mathbf{h}_{l,k}^l - \frac{u'\,(1+c_1u_1)}{1+c_1u_1+c_0u_2+(c_0c_1-c_2^2)\,u_1u_2} \xrightarrow[N\to\infty]{\text{a.s.}} 0\,. \tag{34}$$

In order to find the second original term ($(\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{B}\mathbf{Q}_l\tilde{\mathbf{h}}_{l,k}^l$), we reform and plug (32) into (31) and follow analogously the path we took to arrive at (34). We finally find

$$(\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{B}\mathbf{Q}_l\tilde{\mathbf{h}}_{l,k}^l - \frac{-c_2u_1u'}{1+c_1u_1+c_0u_2+(c_0c_1-c_2^2)\,u_1u_2} \xrightarrow[N\to\infty]{\text{a.s.}} 0\,. \tag{35}$$

D. Interference Power Terms

Having obtained the preparation results in Appendix C-C we can now continue to find the DEs for different parts of the interference power term. From (14) we arrive at

$$\begin{aligned}\mathrm{Int}_{l,k}^{(l)} = \sum_{m\neq l}\nu_m\chi_{l,k}^m \underbrace{(\mathbf{h}_{l,k}^m)^{\mathsf{H}}\mathbf{Q}_m\hat{\mathbf{H}}_m^m(\hat{\mathbf{H}}_m^m)^{\mathsf{H}}\mathbf{Q}_m\mathbf{h}_{l,k}^m}_{\text{Part A}_m} \\ +\nu_l\chi_{l,k}^l \underbrace{(\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{Q}_l\hat{\mathbf{H}}_{l[k]}^l(\hat{\mathbf{H}}_{l[k]}^l)^{\mathsf{H}}\mathbf{Q}_l\mathbf{h}_{l,k}^l}_{\text{Part B}}\,.\end{aligned} \tag{36}$$

We start by treating (36) Part B first. Employing the relationships $\mathbf{ABD}=\mathbf{ACD}+\mathbf{A}(\mathbf{B}-\mathbf{C})\mathbf{D}$ and (23) one finds

$$\begin{aligned}\text{Part B} = (\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{Q}_{l[k]}\hat{\mathbf{H}}_{l[k]}^l(\hat{\mathbf{H}}_{l[k]}^l)^{\mathsf{H}}\mathbf{Q}_l\hat{\mathbf{H}}_{l[k]}^l\mathbf{h}_{l,k}^l \\ -(\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{Q}_l\left[\mathbf{Q}_l^{-1}-\mathbf{Q}_{l[k]}^{-1}\right]\mathbf{Q}_{l[k]}\hat{\mathbf{H}}_{l[k]}^l(\hat{\mathbf{H}}_{l[k]}^l)^{\mathsf{H}}\mathbf{Q}_l\mathbf{h}_{l,k}^l\,.\end{aligned}$$

Using the relationship (30) pertaining to $\left[\mathbf{Q}_l^{-1}-\mathbf{Q}_{l[k]}^{-1}\right]$, we can split Part B as

$$\text{Part B} = X_1 - c_0X_3X_1 - c_2X_3X_2 - c_2X_4X_1 - c_1X_4X_2\,.$$

Where we have found and abbreviated the 4 quadratic forms,

$$\begin{aligned}X_1 &= (\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{Q}_{l[k]}\hat{\mathbf{H}}_{l[k]}^l(\hat{\mathbf{H}}_{l[k]}^l)^{\mathsf{H}}\mathbf{Q}_l\mathbf{h}_{l,k}^l\\ X_2 &= (\tilde{\mathbf{h}}_{l,k}^l)^{\mathsf{H}}\mathbf{Q}_{l[k]}\hat{\mathbf{H}}_{l[k]}^l(\hat{\mathbf{H}}_{l[k]}^l)^{\mathsf{H}}\mathbf{Q}_l\mathbf{h}_{l,k}^l\\ X_3 &= (\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{Q}_l\mathbf{h}_{l,k}^l\\ X_4 &= (\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{Q}_l\tilde{\mathbf{h}}_{l,k}^l\,.\end{aligned}$$

To find the deterministic equivalents for X_1 and X_2, we can use (34) and (35), respectively, where $\mathbf{B}=\mathbf{Q}_{l[k]}\hat{\mathbf{H}}_{l[k]}^l(\hat{\mathbf{H}}_{l[k]}^l)^{\mathsf{H}}$. The respective variables u_1, u_2 and u' for this choice of $\mathbf{B}$ are found (using the same standard techniques as in Appendix C-B) to be

$$u_1 = (\tilde{\mathbf{h}}_{l,k}^l)^{\mathsf{H}}\mathbf{Q}_{l[k]}\tilde{\mathbf{h}}_{l,k}^l \qquad \Rightarrow \quad u_1 - e_l \xrightarrow[N\to\infty]{\text{a.s.}} 0\,.$$

Analogously,

$$u_1 - e_{(l)} \xrightarrow[N\to\infty]{\text{a.s.}} 0\,.$$

Hence, we see that u_1 and u_2 converge to the same value and we will abbreviate them henceforth as u. For the still missing term u' we arrive at

$$\begin{aligned}u' &= (\mathbf{h}_{l,k}^l)^{\mathsf{H}}\mathbf{Q}_{l[k]}\hat{\mathbf{H}}_{l[k]}^l(\hat{\mathbf{H}}_{l[k]}^l)^{\mathsf{H}}\mathbf{Q}_{l[k]}\mathbf{h}_{l,k}^l\\ &\Rightarrow u' - g_l \xrightarrow[N\to\infty]{\text{a.s.}} 0\end{aligned}$$

where the last step makes have use of the results in Appendix C-A. Also, we remind ourselves that we have $c_0 = \alpha_l^l\chi_{l,k}^l\left(1-(\tau_l^l)^2\right)$, $c_1 = \alpha_l^l\chi_{l,k}^l(\tau_l^l)^2$ and $c_2 = \alpha_l^l\chi_{l,k}^l\sqrt{(1-(\tau_l^l)^2)}\tau_l^l$, hence $c_0+c_1=\alpha_l^l\chi_{l,k}^l$ and $c_0c_1-c_2^2=0$. So, finally, we have

$$X_1 - \frac{u'\,(1+c_1u)}{1+(c_1+c_0)\,u} \xrightarrow[N\to\infty]{\text{a.s.}} 0$$

and similarly

$$\text{and } X_2 - \frac{-c_2uu'}{1+(c_1+c_0)\,u} \xrightarrow[N\to\infty]{\text{a.s.}} 0\,.$$

To find the DEs for X_3 and X_4, we can again use (34) and (35), respectively. This time $\mathbf{B}=\mathbf{I}$ and hence the variables simplify to $u'=u_1=u_2\triangleq u$, where $u-e_l \xrightarrow[N\to\infty]{\text{a.s.}} 0$. Thus,

$$\begin{aligned}X_3 - \frac{u\,(1+c_1u)}{1+(c_1+c_0)\,u} &\xrightarrow[N\to\infty]{\text{a.s.}} 0\\ X_4 - \frac{-c_2u^2}{1+(c_1+c_0)\,u} &\xrightarrow[N\to\infty]{\text{a.s.}} 0\,.\end{aligned}$$

Combining all results after further simplifications, we can express the DE of Part B, i.e., $\overline{\text{Part B}}$, as

$$\overline{\text{Part B}} = g_l\frac{1-(\tau_l^l)^2}{\left(1+\alpha_l^l\chi_{l,k}^le_l\right)^2} + g_l(\tau_l^l)^2\,.$$

The next step is to derive the DE of (36) Part A_m, i.e., $\overline{\text{Part A}}_m$. Fortunately, the sum obliges $m\neq l$ and, thus, the same derivation like for Part B applies. Hence, we arrive at

$$\overline{\text{Part A}}_m = g_m\frac{1-(\tau_l^m)^2}{\left(1+\alpha_l^m\chi_{l,k}^me_m\right)^2} + g_m(\tau_l^m)^2\,.$$

Combing Part B and the sum of Part A_m with our original expression of the interference power, we arrive at the familiar expression from Theorem 1.

REFERENCES

[1] E. Björnson and E. Jorswieck, "Optimal Resource Allocation in Coordinated Multi-Cell Systems," *Foundations and Trends in Communications and Information Theory*, vol. 9, no. 2-3, pp. 113–381, 2013.

[2] J. Jose, A. Ashikhmin, T. Marzetta, and S. Vishwanath, "Pilot Contamination and Precoding in Multi-Cell TDD Systems," *IEEE Transactions on Communications*, vol. 10, no. 8, pp. 2640–2651, 2011.

[3] J. Hoydis, K. Hosseini, S. ten Brink, and M. Debbah, "Making Smart Use of Excess Antennas: Massive MIMO, Small Cells, and TDD," *Bell Labs Technical Journal*, vol. 18, no. 2, pp. 5–21, 2013.

[4] Cisco, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2012-2017," *White Paper*, 2013.

[5] G. Americas, "Meeting the 1000x Challenge: The Need for Spectrum, Technology and Policy Innovation," 4G Americas, Tech. Rep., October 2013.

[6] W. Webb, *Wireless Communications: The Future*. John Wiley & Sons, 2007.

[7] J. G. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. C. Reed, "Femtocells: Past, Present, and Future," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, pp. 497–508, 2012.

[8] J. Hoydis, M. Kobayashi, and M. Debbah, "Green Small-Cell Networks," *IEEE Vehicular Technology Magazine*, vol. 6, no. 1, pp. 37–43, Mar. 2011.

[9] H. Dai, A. F. Molisch, and H. V. Poor, "Downlink Capacity of Interference-limited MIMO Systems with Joint Detection," *IEEE Transactions on Wireless Communications*, vol. 3, no. 2, pp. 442–453, 2004.

[10] D. Gesbert, S. Hanly, H. Huang, S. Shamai Shitz, O. Simeone, and W. Yu, "Multi-Cell MIMO Cooperative Networks: A New Look at Interference," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 9, pp. 1380–1408, 2010.

[11] T. Marzetta, "Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, November 2010.

[12] R. Irmer, H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck, H.-P. Mayer, L. Thiele, and V. Jungnickel, "Coordinated Multipoint: Concepts, Performance, and Field Trial Results," *IEEE Communications Magazine*, vol. 49, no. 2, pp. 102–111, 2011.

[13] D. Gesbert, M. Kountouris, R. W. Heath, C.-B. Chae, and T. Sälzer, "Shifting the MIMO Paradigm," *IEEE Signal Processing Magazine*, vol. 24, no. 5, pp. 36–46, 2007.

[14] C. Peel, B. Hochwald, and A. Swindlehurst, "A Vector-Perturbation Technique for Near-Capacity Multiantenna Multiuser Communication—Part I: Channel Inversion and Regularization," *IEEE Transactions on Communications*, vol. 53, no. 1, pp. 195–202, 2005.

[15] R. Zakhour and D. Gesbert, "Distributed Multicell-MISO Precoding Using the Layered Virtual SINR Framework," *IEEE Transactions on Wireless Communications*, vol. 9, no. 8, pp. 2444–2448, 2010.

[16] E. Björnson and B. Ottersten, "On the Principles of Multicell Precoding with Centralized and Distributed Cooperation," in *Proc. WCSP*, 2009.

[17] H. Dahrouj and W. Yu, "Coordinated Beamforming for the Multicell Multi-Antenna Wireless System," *IEEE Transactions on Wireless Communications*, vol. 9, no. 5, pp. 1748–1759, 2010.

[18] E. Björnson, M. Bengtsson, and B. Ottersten, "Pareto Characterization of the Multicell MIMO Performance Region With Simple Receivers," *IEEE Transactions on Signal Processing*, vol. 60, no. 8, pp. 4464–4469, 2012.

[19] E. Björnson, M. Bengtsson, and B. Ottersten, "Optimal Multiuser Transmit Beamforming: A Difficult Problem with a Simple Solution Structure," *IEEE Signal Processing Magazine*, vol. 31, no. 4, pp. 142–148, 2014.

[20] S. Wagner, R. Couillet, M. Debbah, and D. Slock, "Large System Analysis of Linear Precoding in MISO Broadcast Channels with Limited Feedback," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4509–4537, July 2012.

[21] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of Cellular Networks: How Many Antennas Do We Need?" *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 160–171, Feb. 2013.

[22] A. Wyner, "Shannon-Theoretic Approach to a Gaussian Cellular Multiple-Access Channel," *IEEE Transactions on Information Theory*, vol. 40, no. 6, pp. 1713–1727, 1994.

[23] J. Xu, J. Zhang, and J. Andrews, "On the Accuracy of the Wyner Model in Cellular Networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 9, pp. 3098–3109, 2011.

[24] R. Couillet and M. Debbah, *Random Matrix Methods for Wireless Communications*. New York, NY, USA: Cambridge University Press, 2011, first Edition.

[25] E. Björnson, E. G. Larsson, and M. Debbah, "Massive MIMO for Maximal Spectral Efficiency: How Many Users and Pilots Should Be Allocated?" *IEEE Transactions on Wireless Communications*, 2014, submitted.

[26] H. Yang and T. Marzetta, "Total Energy Efficiency of Cellular Large Scale Antenna System Multiple Access Mobile Networks," in *Proc. IEEE Online Conference on Green Communiactions (OnlineGreenComm)*, 2013.

[27] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah, "Massive MIMO Systems with Non-Ideal Hardware: Energy Efficiency, Estimation, and Capacity Limits," *IEEE Transactions on Information Theory*, July 2013, arXiv:1307.2584.

[28] J. Choi, D. Love, and P. Bidigare, "Downlink Training Techniques for FDD Massive MIMO Systems: Open-Loop and Closed-Loop Training with Memory," *IEEE Journal of Selected Topics in Signal Processing*, Sept. 2013, submitted, arXiv:1309.7712.

[29] C. Wang and R. Murch, "Adaptive Downlink Multi-User MIMO Wireless Systems for Correlated Channels with Imperfect CSI," *IEEE Transactions on Wireless Communications*, vol. 5, no. 9, pp. 2435–2436, Sept. 2006.

[30] B. Nosrat-Makouei, J. Andrews, and R. Heath, "MIMO Interference Alignment Over Correlated Channels With Imperfect CSI," *IEEE Transactions on Signal Processing*, vol. 59, no. 6, pp. 2783–2794, Jun. 2011.

[31] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*. Siam, 2000, vol. 2, http://www.matrixanalysis.com/DownloadChapters.html.

[32] 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.

[33] Z. D. Bai and J. W. Silverstein, "No Eigenvalues Outside the Support of the Limiting Spectral Distribution of Large Dimensional Sample Covariance Matrices," *Annals of Probability*, vol. 26, no. 1, pp. 316–345, Jan. 1998.

Axel Müller received his B.Sc.(hons) from Napier University, Edinburgh as well as his Dipl.-Ing.(FH) from the University of Applied Sciences, Ulm, in 2008. In 2010, he graduated from Ulm University with a M.Sc. in communications engineering. In November 2014, Axel Mueller obtained his Ph.D. degree in Telecommunications at Centrale-Supélec, Paris, on the topic "Random Matrix Analysis of Future Multicell MU-MIMO Networks". His thesis advisor was Prof. Mérouane Debbah and his co-advisor was Prof. Romain Couillet. He is the recipient of a SAM 2014 best student paper award. Currently, Axel Mueller is working as a research engineer at Huawei Technologies (France Research Center, Paris), where his main tasks center around 4.5G and 5G cellular technologies.

Romain Couillet received his M.Sc. 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 Supélec, France, which he graduated in November 2010. He is currently an assistant professor in the Telecommunication department of SUPÉLEC, France. His research topics are in information theory, signal processing, complex systems and random matrix theory. 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.

Emil Björnson received the M.S. degree in Engineering Mathematics from Lund University, Sweden, in 2007. He received the Ph.D. degree in Telecommunications from the KTH Royal Institute of Technology, Stockholm, Sweden, in 2011. From 2012 to July 2014, he was a joint postdoc at Supélec, Gif-sur-Yvette, France, and at KTH Royal Institute of Technology. He is currently an Assistant Professor and Docent at the Department of Electrical Engineering (ISY) at Linköping University, Sweden.

His research interests include multi-antenna cellular communications, radio resource allocation, energy efficiency, massive MIMO, and network topology design. He is the first author of the textbook "Optimal Resource Allocation in Coordinated Multi-Cell System" published in Foundations and Trends in Communications and Information Theory, 2013. He is also dedicated to reproducible research and has made a large amount of simulation code publicly available. Dr. Björnson received the 2014 Outstanding Young Researcher Award from IEEE ComSoc EMEA and has received 4 best paper awards for novel research on multi-cell multi-antenna communications: WCNC 2014, SAM 2014, CAMSAP 2011, and WCSP 2009.

Sebastian Wagner received the M.Sc. degree in Electrical Engineering from the Technische Universität Dresden, Germany in 2008 and the Ph.D. degree from TELECOM ParisTech (EURECOM), Sophia Antipolis, France in 2011. From 2008 to 2011, he was an research engineer at ST-ERICSSON, Sophia Antipolis, France, where he worked toward the Ph.D. degree. In 2012 he joined EURECOM as a post-doctoral fellow to work on the OpenAirInterface project. Since 2013 he is with Intel Mobile Communications developing and implementing algorithms for LTE products. His current research interests include signal processing for multiple antenna systems and the application of random matrix theory to wireless communications.

Mérouane Debbah entered the Ecole Normale Supérieure de Cachan (France) in 1996 where he received his M.Sc and Ph.D. degrees respectively. He worked for Motorola Labs (Saclay, France) from 1999-2002 and the Vienna Research Center for Telecommunications (Vienna, Austria) until 2003. From 2003 to 2007, he joined the Mobile Communications department of the Institut Eurecom (Sophia Antipolis, France) as an Assistant Professor. Since 2007, he is a Full Professor at Supélec (Gif-sur-Yvette, France). From 2007 to 2014, he was director of the Alcatel-Lucent Chair on Flexible Radio. Since 2014, he is Vice-President of the Huawei France R&D center and director of the Mathematical and Algorithmic Sciences Lab. His research interests are in information theory, signal processing and wireless communications. He is an Associate Editor in Chief of the journal Random Matrix: Theory and Applications and was an associate and senior area editor for IEEE Transactions on Signal Processing respectively in 2011-2013 and 2013-2014. Mérouane Debbah is a recipient of the ERC grant MORE (Advanced Mathematical Tools for Complex Network Engineering). He is an IEEE Fellow, a WWRF Fellow and a member of the academic senate of Paris-Saclay. He is the recipient of the Mario Boella award in 2005, the 2007 IEEE GLOBECOM best paper award, the Wi-Opt 2009 best paper award, the 2010 Newcom++ best paper award, the WUN CogCom Best Paper 2012 and 2013 Award, the 2014 WCNC best paper award as well as the Valuetools 2007, Valuetools 2008, CrownCom2009, Valuetools 2012 and SAM 2014 best student paper awards. In 2011, he received the IEEE Glavieux Prize Award and in 2012, the Qualcomm Innovation Prize Award.