



Optimal Linear Precoding in Multi-User MIMO Systems: A Large System Analysis

Luca Sanguinetti, Emil Björnson, Mérouane Debbah, Aris Moustakas

► To cite this version:

Luca Sanguinetti, Emil Björnson, Mérouane Debbah, Aris Moustakas. Optimal Linear Precoding in Multi-User MIMO Systems: A Large System Analysis. 2014 IEEE Global Communications Conference (Globecom), Dec 2014, Texas, United States. 10.1109/glocom.2014.7037420 . hal-01098917

HAL Id: hal-01098917

<https://hal.science/hal-01098917>

Submitted on 17 Jan 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Optimal Linear Precoding in Multi-User MIMO Systems: A Large System Analysis

Luca Sanguinetti^{*†}, Emil Björnson^{††}, Mérouane Debbah[‡] and Aris L. Moustakas^{‡§}

^{*}Dipartimento di Ingegneria dell'Informazione, University of Pisa, Pisa, Italy

[‡]Alcatel-Lucent Chair, Ecole supérieure d'électricité (Supélec), Gif-sur-Yvette, France

[†]Dept. of Signal Processing, KTH, Stockholm, and Dept. of Electrical Engineering, Linköping University, Sweden

[§]Department of Physics, National & Capodistrian University of Athens, Athens, Greece

Abstract—We consider the downlink of a single-cell multi-user MIMO system in which the base station makes use of N antennas to communicate with K single-antenna user equipments (UEs) randomly positioned in the coverage area. In particular, we focus on the problem of designing the optimal linear precoding for minimizing the total power consumption while satisfying a set of target signal-to-interference-plus-noise ratios (SINRs). To gain insights into the structure of the optimal solution and reduce the computational complexity for its evaluation, we analyze the asymptotic regime where N and K grow large with a given ratio and make use of recent results from large system analysis to compute the asymptotic solution. Then, we concentrate on the asymptotically design of heuristic linear precoding techniques. Interestingly, it turns out that the regularized zero-forcing (RZF) precoder is equivalent to the optimal one when the ratio between the SINR requirement and the average channel attenuation is the same for all UEs. If this condition does not hold true but only the same SINR constraint is imposed for all UEs, then the RZF can be modified to still achieve optimality if statistical information of the UE positions is available at the BS. Numerical results are used to evaluate the performance gap in the finite system regime and to make comparisons among the precoding techniques.

I. INTRODUCTION

Multiple-Input Multiple-Output (MIMO) technologies are currently being adopted in many wireless communication standards such as fourth generation (4G) cellular networks [1]. The main limiting factor in multi-user MIMO systems is the multiple-access interference (MAI). In uplink transmissions, MAI mitigation is typically accomplished at the base station (BS) using linear multi-user detectors or non-linear techniques based on layered architectures. In downlink transmissions, MAI mitigation can only be accomplished at the BS using precoding techniques. As shown in [2], the capacity-achieving precoding strategy is dirty paper coding (DPC). Although optimal, the implementation of DPC requires a tremendous computational complexity at both BS and user equipments (UEs). On the other hand, a practical approach that has received considerable attention (due to its simplicity) is represented by linear precoding or beamforming [3], [4].

In this work, we focus on the problem of designing the optimal linear precoding for minimizing the total transmit power while satisfying a set of target signal-to-interference-plus-noise

ratios (SINRs) [5]–[8]. This problem is receiving renewed interest nowadays due to the emerging research area of green cellular networks [9]. In particular, we consider the downlink of a single-cell multi-user MIMO system in which the BS makes use of N antennas to communicate with K single-antenna UEs randomly positioned in the coverage area. Under the assumption of perfect channel state information (CSI), the solution to the power minimization problem in this context was originally computed in [10] and later extended to different scenarios in [6]–[8]. In particular, it turns out that the optimal linear precoder depends on some Lagrange multipliers whose computation can be performed using convex optimization tools or solving a fixed-point problem [4]. Although possible, both approaches do not provide any insights into the structure of the optimal values. Moreover, the computation must be performed for any new realization of the MIMO channel matrix.

To overcome these issues, we follow the same approach as in [11]–[14] and resort to the asymptotic regime where N and K grow large with a given ratio $c = K/N$. Differently from [11]–[14], the asymptotically optimal values of the Lagrange multipliers are computed using the approach adopted in [15], which provides us a much simpler means to overcome the technical difficulties arising with the application of standard random matrix theory tools (see for example [13]). As already pointed out in [11]–[14], in the asymptotic regime the optimal values can be computed in closed-form through a nice and simple expression, which depends only on the user positions and SINR requirements. The above results are then used to validate the optimality of different heuristic linear precoding techniques, which are inspired by the widely used regularized zero-forcing (RZF) concept [16]–[18] and its extensions to include arbitrary user priorities [4]. The optimal regularization parameter is provided in the asymptotic regime. To the best of authors' knowledge, this is the first time that such a result is found since most of the related works are focused on sum rate maximization. Comparisons are then made with two heuristic techniques. The former is the classical RZF precoder [16] while the latter is referred to as position-aware RZF (PA-RZF) precoder since it relies on knowledge of the UE positions [17]. Interestingly, it turns out that PA-RZF is equivalent to the optimal linear precoder when the same SINR constraint is imposed for all UEs. On the other hand, the commonly used RZF precoder becomes optimal only when the ratio between the SINR requirement and the average channel attenuation is the same for all UEs. Numerical results are used to evaluate

L. Sanguinetti is funded by the People Programme (Marie Curie Actions) FP7 PIEF-GA-2012-330731 Dense4Green. E. Björnson is funded by an International Postdoc Grant from the Swedish Research Council. A. L. Moustakas is the holder of the DIGITEO "ASAPGONE" Chair. This research has also been supported by the FP7 NEWCOM# (Grant no. 318306), the ERC Starting MORE (Grant no. 305123), the French pôle de compétitivité SYSTEM@TIC within the project 4G in Vitro.

the performance gap in the finite system regime and to make comparisons among the different precoding techniques.

II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider the downlink of a single-cell multi-user MIMO system in which the BS makes use of N antennas to communicate with K single-antenna UEs. The K active UEs change over time and are randomly selected from a large set of UEs within the coverage area. The physical location of UE k is denoted by $\mathbf{x}_k \in \mathbb{R}^2$ (in meters) and it is computed with respect to the BS (assumed to be located in the origin). The function $l(\cdot): \mathbb{R}^2 \rightarrow \mathbb{R}_+$ describes the large-scale channel fading at different user locations; that is, $l(\mathbf{x}_k)$ is the average channel attenuation due to path-loss and shadowing at location \mathbf{x}_k . The large-scale fading between a UE and the BS is assumed to be the same for all BS antennas. This is reasonable since the distances between UEs and BS are much larger than the distance between the BS antennas. Since the forthcoming analysis does not depend on a particular choice of $l(\cdot)$, we keep it generic. Perfect CSI is assumed to be available at the BS for analytic tractability. The imperfect CSI case is left for future work.¹

The BS shall convey the information symbol s_k to UE k using linear precoding. The symbol vector $\mathbf{s} = [s_1, s_2, \dots, s_K]^T \in \mathbb{C}^{K \times 1}$ originates from a Gaussian codebook with zero mean and covariance matrix $\mathbb{E}[\mathbf{s}\mathbf{s}^H] = \mathbf{I}_K$. Denoting by $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_K] \in \mathbb{C}^{N \times K}$ the precoding matrix, the received sample $y_k \in \mathbb{C}$ at UE k takes the form

$$y_k = \mathbf{h}_k^H \mathbf{V} \mathbf{s} + n_k \quad (1)$$

where $n_k \sim \mathcal{CN}(0, \sigma^2)$ is the additive noise and the entry $h_{k,n}$ of the row vector $\mathbf{h}_k^H = [h_{k,1}, h_{k,2}, \dots, h_{k,N}] \in \mathbb{C}^{1 \times N}$ is the channel propagation coefficient between the n th antenna at the BS and the k th UE. We assume a Rayleigh fading channel model $\mathbf{h}_k = \sqrt{l(\mathbf{x}_k)} \mathbf{w}_k$ with $\mathbf{w}_k \sim \mathcal{CN}(0, \mathbf{I}_N)$ accounting for the small-scale fading channel. The SINR at the k th UE is easily written as [18]

$$\text{SINR}_k = \frac{|\mathbf{h}_k^H \mathbf{v}_k|^2}{\sum_{i=1, i \neq k}^K |\mathbf{h}_k^H \mathbf{v}_i|^2 + \sigma^2}. \quad (2)$$

As mentioned earlier, we consider the power minimization problem whose mathematical formulation is as follows:

$$\begin{aligned} \mathcal{P}: \quad & \underset{\mathbf{V}}{\text{minimize}} && P = \text{tr}(\mathbf{V}\mathbf{V}^H) \\ & \text{subject to} && \text{SINR}_k \geq \gamma_k \quad k = 1, 2, \dots, K \end{aligned} \quad (3)$$

where γ_k is the given SINR target of UE k obtained as (under the assumption of Gaussian codebooks) $\gamma_k = 2^{r_k} - 1$ with r_k being the target user rate in bit/s/Hz. For later convenience, we call $\boldsymbol{\gamma} = [\gamma_1, \gamma_2, \dots, \gamma_K]^T$.

¹We limit to observe that might in principle be included following the same approach adopted in [16]. See also [19].

III. OPTIMAL LINEAR PRECODING

As originally shown in [10], the non-convex optimization problem \mathcal{P} in (3) can be put in a convex form by reformulating the SINR constraints as second-order cone constraints. In doing so, the optimal \mathbf{V}^* is found to be [6]–[8]

$$\mathbf{V}^* = \left(\sum_{i=1}^K \lambda_i^* \mathbf{h}_i \mathbf{h}_i^H + N \mathbf{I}_N \right)^{-1} \mathbf{H} \sqrt{\mathbf{P}^*} \quad (5)$$

where $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_K] \in \mathbb{C}^{N \times K}$ and $\boldsymbol{\lambda}^* = [\lambda_1^*, \lambda_2^*, \dots, \lambda_K^*]^T$ is the positive unique fixed point of the following equations [4]–[6]:

$$\left(1 + \frac{1}{\gamma_k} \right) \lambda_k^* = \frac{1}{\mathbf{h}_k^H \left(\sum_{i=1}^K \lambda_i^* \mathbf{h}_i \mathbf{h}_i^H + N \mathbf{I}_N \right)^{-1} \mathbf{h}_k} \quad (6)$$

for $k = 1, 2, \dots, K$. Also, $\mathbf{P}^* = \text{diag}\{p_1^*, p_2^*, \dots, p_K^*\}$ is a diagonal matrix whose entries are such that the SINR constraints in \mathcal{P} are all satisfied with equality when $\mathbf{V} = \mathbf{V}^*$. Plugging (2) into (4), the optimal vector $\mathbf{p}^* = [p_1^*, p_2^*, \dots, p_K^*]^T$ is computed as [4]

$$\mathbf{p}^* = \sigma^2 \mathbf{D}^{-1} \mathbf{1}_K \quad (7)$$

where the (k, i) th element of $\mathbf{D} \in \mathbb{C}^{K \times K}$ is

$$[\mathbf{D}]_{k,i} = \begin{cases} \frac{1}{\gamma_k} |\mathbf{h}_k^H \mathbf{a}_k^*|^2 & \text{for } k = i \\ -|\mathbf{h}_k^H \mathbf{a}_i^*|^2 & \text{for } k \neq i \end{cases} \quad (8)$$

with \mathbf{a}_k^* being the k th column of $\mathbf{A}^* = (\sum_{i=1}^K \lambda_i^* \mathbf{h}_i \mathbf{h}_i^H + N \mathbf{I}_N)^{-1} \mathbf{H}$.

As seen, \mathbf{V}^* in (5) is parameterized by $\boldsymbol{\lambda}^*$ and \mathbf{p}^* , where $\boldsymbol{\lambda}^*$ needs to be evaluated by an iterative procedure due to the fixed-point equations in (6). This is a computationally demanding task when N and K are large since the matrix inversion operation in (6) must be recomputed at every iteration and its computational complexity scales proportionally to $N^2 K$. Moreover, computing $\boldsymbol{\lambda}^*$ as the fixed point of (6) does not provide any insights into the optimal structure of both $\boldsymbol{\lambda}^*$ and \mathbf{p}^* . In addition, the parameter values depend directly on the channel vectors $\{\mathbf{h}_i\}$ and change at the same pace as the small-scale fading (i.e., at the order of milliseconds).

To overcome the above issues, we assume that $N, K \rightarrow \infty$ with $K/N = c \in (0, 1]$ and use some recent tools in large system analysis to compute the so-called deterministic equivalents of $\boldsymbol{\lambda}^*$ and \mathbf{p}^* . For later convenience, we call

$$\xi = 1 - \frac{1}{N} \sum_{i=1}^K \frac{\gamma_i}{1 + \gamma_i} \quad \text{and} \quad A = \frac{1}{K} \sum_{i=1}^K \frac{\gamma_i}{l(\mathbf{x}_i)}. \quad (9)$$

A. Asymptotically Optimal Linear Precoding

The following theorem provides the solution to the optimization problem in (3) in the asymptotic regime.

Theorem 1. *If $N, K \rightarrow \infty$ with $c \in (0, 1]$, then*

$$\max_{k=1,2,\dots,K} |\lambda_k^* - \bar{\lambda}_k| \xrightarrow{a.s.} 0 \quad (10)$$

and

$$\max_{k=1,2,\dots,K} |p_k^* - \bar{p}_k| \xrightarrow{a.s.} 0 \quad (11)$$

where $\bar{\lambda}_k$ and \bar{p}_k are the deterministic equivalents of λ_k^* and p_k^* , respectively, and are given by

$$\bar{\lambda}_k = \frac{\gamma_k}{l(\mathbf{x}_k)\xi} \quad (12)$$

and

$$\bar{p}_k = \frac{\gamma_k}{l(\mathbf{x}_k)\xi^2} \left(\bar{P} + \frac{\sigma^2}{l(\mathbf{x}_k)} (1 + \gamma_k)^2 \right) \quad (13)$$

with

$$\bar{P} = \frac{cA\sigma^2}{\xi} \quad (14)$$

being the deterministic equivalent of the transmit power P .

Proof: Similar results have previously been derived by applying standard random matrix theory tools to the right-hand-side of (6). However, the application of these tools to the problem at hand is not analytically correct since the Lagrange multipliers in (6) are a function of the channel vectors $\{\mathbf{h}_k\}$. To overcome this issue, we make use of the same approach adopted in [15] whose main steps are sketched in [20]. On the other hand, (13) is proved using standard random matrix theory results (omitted for space limitations). ■

The following remarks elaborate on some of the insights that are obtained from Theorem 1.

Remark 1. In sharp contrast to (6), the computation of $\bar{\lambda}_k$ in (12) only requires knowledge of the user position through $l(\mathbf{x}_k)$. This information can be easily observed and estimated accurately at the BS because it changes slowly with time (relative to the small-scale fading). The Lagrange multiplier λ_k is known to act as a user priority parameter that implicitly determines how much interference the other UEs may cause to UE k [4]. Interestingly, its asymptotic value $\bar{\lambda}_k$ is proportional to the SINR γ_k and inversely proportional to $l(\mathbf{x}_k)$ such that users with weak channels have larger values. Higher priority is thus given to users that require high performance and/or have weak propagation conditions [4].

Remark 2. A known problem with using the asymptotically optimal power allocation in Lemma 1 is that the target SINRs are not guaranteed to be achieved at finite numbers of antennas (see for example [14]). This is because the approximation errors are translated into fluctuations in the resulting SINR values. However, these errors rapidly vanish also in the finite regime when N is larger than K , which is the regime envisioned for massive MIMO systems [21]. It can also be avoided by using only the deterministic equivalents $\bar{\lambda}_k$ of the Lagrange multipliers and computing the power allocation coefficients according to (7). This approach retains most of the complexity benefits of the asymptotic analysis.

The following corollaries can be easily obtained from Theorem 1 and will be useful later on.

Corollary 1. *If the ratio between the SINR requirement and the average channel attenuation is the same for all UEs and*

equal to some $\zeta \geq 0$, i.e.,

$$\frac{\gamma_k}{l(\mathbf{x}_k)} = \zeta \quad k = 1, 2, \dots, K, \quad (15)$$

then $\bar{\lambda}_k$ in (12) takes the form

$$\bar{\lambda}_k = \frac{\zeta}{\xi} = \zeta \left(1 - \frac{1}{N} \sum_{i=1}^K \frac{\gamma_i}{1 + \gamma_i} \right)^{-1}. \quad (16)$$

Corollary 2 ([13]). *If the same target SINR is imposed for each user, i.e.,*

$$\gamma = \gamma \mathbf{1}_K, \quad (17)$$

then $\bar{\lambda}_k$ in (12) reduces to

$$\bar{\lambda}_k = \frac{\gamma}{l(\mathbf{x}_k)\xi} = \frac{\gamma}{l(\mathbf{x}_k)} \left(1 - c \frac{\gamma}{1 + \gamma} \right)^{-1} \quad (18)$$

and \bar{P} becomes

$$\bar{P} = cA\sigma^2 \left(1 - c \frac{\gamma}{1 + \gamma} \right)^{-1}. \quad (19)$$

IV. HEURISTIC LINEAR PRECODING

Inspired by the optimal linear precoding in (5), we now consider suboptimal precoding techniques that builds on heuristics [4]. To this end, we let \mathbf{V} take the following general form

$$\mathbf{V} = \left(\sum_{i=1}^K \alpha_i \mathbf{h}_i \mathbf{h}_i^H + N\rho \mathbf{I}_N \right)^{-1} \mathbf{H} \sqrt{\mathbf{P}} \quad (20)$$

where $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_K]^T$ is now a given vector with positive scalars and ρ is a design parameter to be optimized. Note that (20) is basically obtained from (5) by setting $\lambda_k = \alpha_k/\rho$ for all k . As before, the power allocation \mathbf{P} is computed according to (7) and satisfies all the SINR constraints with equality.

Observe that if $\boldsymbol{\alpha}$ is set to $\mathbf{1}_K$, then \mathbf{V} in (5) reduces to the well-known RZF precoder [16]:

$$\mathbf{V}_{\text{RZF}} = \left(\sum_{i=1}^K \mathbf{h}_i \mathbf{h}_i^H + N\rho \mathbf{I}_N \right)^{-1} \mathbf{H} \sqrt{\mathbf{P}}. \quad (21)$$

This particular precoding matrix is also known as the transmit Wiener filter and signal-to-leakage-and-noise ratio (SLNR) maximizing beamforming (see Remark 3.2 in [18] for a historical exposition).

On the other hand, if the BS makes use of knowledge of the user positions and let $\boldsymbol{\alpha}$ be equal to $\mathbf{L}^{-1} \mathbf{1}_K$ with $\mathbf{L} = \text{diag}\{l(\mathbf{x}_1), l(\mathbf{x}_2), \dots, l(\mathbf{x}_K)\}$, then the processing matrix \mathbf{V} in (5) reduces to (see also [17], [19], [20])

$$\mathbf{V}_{\text{PA-RZF}} = \left(\sum_{i=1}^K \mathbf{w}_i \mathbf{w}_i^H + N\rho \mathbf{I}_N \right)^{-1} \mathbf{H} \sqrt{\mathbf{P}} \quad (22)$$

which we refer to as PA-RZF precoder in the sequel.

Differently from the optimal linear precoding that requires to compute the fixed point of a set of equations, the optimization of a linear precoder in the form of (20) requires only to look for the value of ρ minimizing the transmit power. This can

generally not be done in closed-form but requires a numerical optimization procedure [4]. To overcome this problem, the asymptotic regime is analyzed in the sequel.

A. Asymptotic Analysis of the Heuristic Linear Precoding

We keep α generic and look for the value of ρ that minimizes the total transmit power P in (3) when $N, K \rightarrow \infty$ with $c \in (0, 1]$. In doing so, the following result is obtained.

Theorem 2. *If $N, K \rightarrow \infty$ with $c \in (0, 1]$, then the parameter ρ minimizing the deterministic equivalent of P with \mathbf{V} given by (5) is*

$$\rho^* = \frac{1}{\mu^*} - \frac{1}{N} \sum_{i=1}^K \frac{\alpha_i l(\mathbf{x}_i)}{1 + \alpha_i l(\mathbf{x}_i) \mu^*} \quad (23)$$

where μ^* is the solution of the following fixed point equation:

$$\mu^* = \left(\sum_{i=1}^K \frac{\alpha_i l(\mathbf{x}_i) \gamma_i}{(1 + \alpha_i l(\mathbf{x}_i) \mu^*)^3} \right) \left(\sum_{i=1}^K \frac{(\alpha_i l(\mathbf{x}_i))^2}{(1 + \alpha_i l(\mathbf{x}_i) \mu^*)^3} \right)^{-1} \quad (24)$$

In addition, the deterministic equivalent of p_k takes the form

$$\bar{p}_k = \frac{\gamma_k}{l(\mathbf{x}_k)(\mu^*)^2} \left[\bar{P} + \frac{\sigma^2}{l(\mathbf{x}_k)} (1 + \alpha_k l(\mathbf{x}_k) \mu^*)^2 \right] \quad (25)$$

where

$$\bar{P} = \frac{cA\sigma^2}{1 - (\mu^*)^2 F_2 - cB} \quad (26)$$

is the deterministic equivalent of transmit power with A given by (9) and

$$B = \frac{1}{K} \sum_{i=1}^K \frac{\gamma_i}{(1 + \alpha_i l(\mathbf{x}_i) \mu^*)^2} \quad (27)$$

$$F_2 = \frac{1}{N} \sum_{i=1}^K \frac{(\alpha_i l(\mathbf{x}_i))^2}{(1 + \alpha_i l(\mathbf{x}_i) \mu^*)^2}. \quad (28)$$

Proof: The proof is detailed in Appendix and operates in two steps. In the first step, we use the results of Theorem 1 in [16] to compute SINR_k and \bar{P} , i.e., the deterministic equivalents of SINR_k and P , respectively. In the second step, we set $\bar{\text{SINR}}_k = \gamma_k$ for $k = 1, 2, \dots, K$ and compute the corresponding powers $\{\bar{p}_k\}$, which are eventually used to obtain \bar{P} . The latter takes the form in (26) from which taking the derivative with respect to ρ we obtain (23) and (24). ■

As mentioned earlier, this is the first time that the optimal value of ρ minimizing the power consumption is given in explicit form for a generic heuristic precoding matrix \mathbf{V} defined as in (20). Most of the existing works have only looked for the value of ρ that maximizes the sum rate of the network (see for example [16]).

From the results of Theorem 2, the optimal value of ρ for RZF or PA-RZF easily follows.

Corollary 3. *If a RZF precoder is used and $N, K \rightarrow \infty$ with $c \in (0, 1]$, then*

$$\rho_{\text{RZF}}^* = \frac{1}{\mu^*} - \frac{1}{N} \sum_{i=1}^K \frac{l(\mathbf{x}_i)}{1 + l(\mathbf{x}_i) \mu^*} \quad (29)$$

with μ^* being solution of the following fixed point equation:

$$\mu^* = \left(\sum_{i=1}^K \frac{l(\mathbf{x}_i) \gamma_i}{(1 + l(\mathbf{x}_i) \mu^*)^3} \right) \left(\sum_{i=1}^K \frac{(l(\mathbf{x}_i))^2}{(1 + l(\mathbf{x}_i) \mu^*)^3} \right)^{-1}. \quad (30)$$

Proof: The proof easily follows from the results of Theorem 1 setting $\alpha_i = 1$ for $i = 1, 2, \dots, K$. ■

Corollary 4. *If a PA-RZF precoder is used and $N, K \rightarrow \infty$ with $c \in (0, 1]$, then*

$$\rho_{\text{PA-RZF}}^* = \frac{1}{\beta} - \frac{c}{1 + \beta} \quad (31)$$

with $\beta > 0$ being the average target SINR given by

$$\beta = \frac{1}{K} \sum_{k=1}^K \gamma_k. \quad (32)$$

The deterministic equivalent of the minimum transmit power reduces to

$$\bar{P}_{\text{PA-RZF}} = cA\sigma^2 \left(1 - c \frac{\beta}{1 + \beta} \right)^{-1}. \quad (33)$$

Proof: The result follows directly from Theorem 1 setting $\alpha_i = 1/l(\mathbf{x}_i)$ for $i = 1, 2, \dots, K$. ■

Interestingly, the above results can be used to prove under which conditions RZF and PA-RZF are optimal.

Corollary 5. *If condition (15) holds true, then RZF becomes the optimal linear precoder in the asymptotic regime.*

Proof: From (15), it follows that $l(\mathbf{x}_k) \gamma_k = (l(\mathbf{x}_k))^2 \zeta$. Plugging this result into (30) yields $\mu^* = \zeta$ from which we get

$$\rho_{\text{RZF}}^* = \frac{1}{\zeta} \left(1 - \frac{1}{N} \sum_{i=1}^K \frac{\gamma_i}{1 + \gamma_i} \right) \quad (34)$$

by simple manipulations. Plugging this result into (21) we obtain

$$\mathbf{V}_{\text{RZF}}^* = \frac{1}{\rho_{\text{RZF}}^*} \left(\frac{1}{\rho_{\text{RZF}}^*} \sum_{i=1}^K \mathbf{w}_i \mathbf{w}_i^H + N \mathbf{I}_N \right)^{-1} \mathbf{H} \sqrt{\bar{P}} \quad (35)$$

which is equal to (5) after replacing λ_k^* with $\bar{\lambda}_k$ in (16). ■

Corollary 6. *If condition (17) holds true, then PA-RZF becomes the optimal linear precoder in the asymptotic regime*

Proof: If $\gamma = \gamma \mathbf{1}_K$, then (31) reduces to $\rho_{\text{PA-RZF}}^* = \frac{1}{\gamma} - \frac{c}{1 + \gamma}$ and $\mathbf{V}_{\text{PA-RZF}}^*$ in (22) becomes equivalent to (5) after replacing λ_k^* with $\bar{\lambda}_k$ given by (18). ■

V. NUMERICAL RESULTS

In this section, Monte Carlo simulations are used to validate the analysis in the asymptotic regime and to make comparisons between optimal linear precoding and different heuristic precoding techniques. We assume that the UEs are uniformly distributed in a circular cell with radius $D = 250$ m and minimum distance $D_{\min} = 15$ m. Moreover, we consider

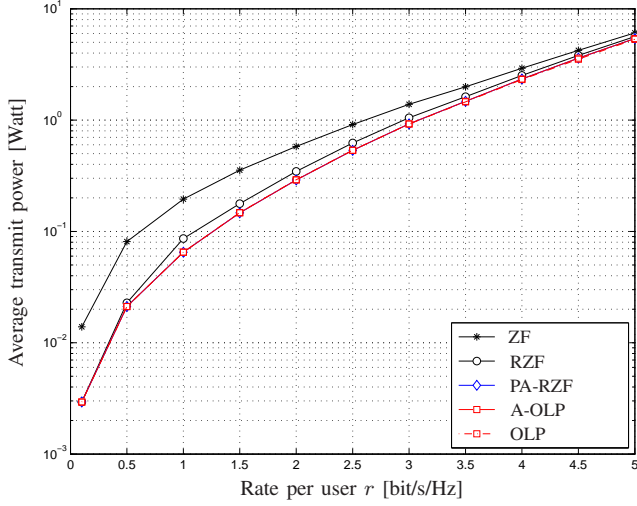


Fig. 1. Average transmit power in Watt vs. the rate per user r in bit/s/Hz when $K = 8$ and $N = 10$.

a system in which the large-scale fading is dominated by the path-loss [22]. This is modelled as $l(\mathbf{x}) = d_0/\|\mathbf{x}\|^\kappa$ for $\|\mathbf{x}\| \geq D_{\min}$ where $\kappa \geq 2$ is the path-loss exponent and the constant $d_0 > 0$ regulates the channel attenuation at distance D_{\min} . In all subsequent simulations, we set $\kappa = 3.76$ and $d_0 = 10^{-3.53}$. In addition, the transmission bandwidth is $W = 10$ MHz and the total noise power $W\sigma^2$ is -104 dBm.

We begin by considering a cellular network in which the same rate r in bit/s/Hz must be guaranteed to each UE. This amounts to saying that $\gamma_k = \gamma = 2^r - 1$ for $k = 1, 2, \dots, K$. Fig. 1 illustrates the average transmit power in Watt with $K = 8$ and $N = 10$ when r spans the interval from 0.1 to 5 bit/s/Hz. The curves labelled OLP and A-OLP refer to the performance of the optimal and asymptotically optimal linear precoders, respectively. On the other hand, ZF refers to the classical zero-forcing precoder. From the results of Fig. 1, it follows that OLP and A-OLP have substantially the same performance. As pointed out in Remark 3, PA-RZF provides the same performance of A-OLP. While PA-RZF achieves only a marginal gain compared to RZF, a substantial power reduction is obtained with respect to ZF for moderate values of r . The mean-square-error of the effective user rates (not reported here for space limitations) is found to be smaller than 2% meaning that the performance loss is reasonably negligible.

Fig. 2 plots the average transmit power in Watt vs. N when $K = 8$ and the user rates $\{r_k\}$ are randomly taken within the interval $[2, 3]$ bit/s/Hz. Although different rates are requested by the UEs, PA-RZF has substantially the same performance of A-OLP for any value of N . A significant gap is observed with respect to ZF for values of N in the order of K , while all the schemes guarantee basically the same performance when N becomes larger.

VI. CONCLUSIONS

In this work, we have focused on a single-cell multi-user MIMO system and have studied the problem of designing linear precoding techniques for minimizing the total power

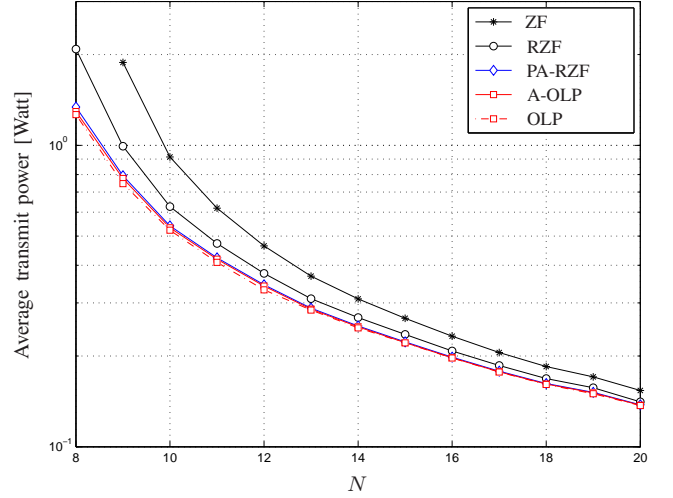


Fig. 2. Average transmit power in Watt vs. N when $K = 8$ and the user rates take values within the interval $[2, 3]$ bit/s/Hz.

consumption while satisfying a set of target SINRs. The solution to this problem is generally given by solving a set of fixed-point equations, which is cumbersome in large-scale MIMO systems. To simplify the analysis and overcome complexity issues, we have resorted to the asymptotic regime in which the number of antennas and users grow large with a given ratio. The asymptotic solutions to the fixed-point equations have been given in closed form, thereby providing insights on the optimal precoding structure. In particular, we have used these results to prove that the conventional RZF precoding technique is the optimal one in the asymptotic regime when the ratio between the SINR requirement and the average channel attenuation is the same for all UEs. A position-aware RZF (PA-RZF) precoding that exploits statistical knowledge of the UE positions has been shown to be asymptotically optimal in realistic scenarios where the SINR constraints are the same but the path-losses are different.

APPENDIX

If \mathbf{V} takes the generic heuristic form in (20), then for any given $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_K]^T$ and ρ the following lemma can be proved using the results of Theorem 1 in [16].

Lemma 1. *If $N, K \rightarrow \infty$ with $c \in (0, 1]$, then*

$$P - \overline{P} \xrightarrow{a.s.} 0 \quad (36)$$

$$\text{SINR}_k - \overline{\text{SINR}}_k \xrightarrow{a.s.} 0 \quad (37)$$

where \overline{P} and $\overline{\text{SINR}}_k$ are given by

$$\overline{P} = \frac{c\mu'}{K} \sum_{i=1}^K \frac{p_i l(\mathbf{x}_i)}{(1 + \alpha_i l(\mathbf{x}_i)\mu)^2} \quad (38)$$

$$\overline{\text{SINR}}_k = \frac{p_k l(\mathbf{x}_k)\mu^2}{\overline{P} + \frac{\sigma^2}{l(\mathbf{x}_k)} [1 + \alpha_k l(\mathbf{x}_k)\mu]^2} \quad (39)$$

where μ is the solution of the following fixed point equation

$$\mu = \left(\frac{1}{N} \sum_{i=1}^K \frac{\alpha_i l(\mathbf{x}_i)}{1 + \alpha_i l(\mathbf{x}_i)\mu} + \rho \right)^{-1} \quad (40)$$

and μ' in (38) is its derivative with respect to ρ .

To proceed further, we set $\overline{\text{SINR}}_k = \gamma_k$ for $k = 1, 2, \dots, K$ and compute the corresponding power \overline{P} .

Lemma 2. *If $\overline{\text{SINR}}_k$ is set equal to γ_k for $k = 1, 2, \dots, K$, then \overline{P} is found to be*

$$\overline{P} = \frac{cA\sigma^2}{1 - \mu^2 F_2 - cB} \quad (41)$$

with A and μ being given by (9) and (40) whereas B and F_2 takes the form in (27) and (28).

Proof: Setting $\overline{\text{SINR}}_k$ in (39) equal to γ_k leads to

$$p_k = \frac{\gamma_k}{f_k(\rho)} = \frac{\gamma_k \left\{ \overline{P} + \frac{\sigma^2}{l(\mathbf{x}_k)} [1 + \alpha_k l(\mathbf{x}_k) \mu]^2 \right\}}{l(\mathbf{x}_k) \mu^2} \quad (42)$$

which used in (38) yields

$$\overline{P} = -\frac{c\mu'}{\mu^2} \frac{1}{K} \sum_{i=1}^K \frac{\gamma_i \left\{ \overline{P} + \frac{\sigma^2}{l(\mathbf{x}_i)} [1 + \alpha_i l(\mathbf{x}_i) \mu]^2 \right\}}{(1 + \alpha_i l(\mathbf{x}_i) \mu)^2}. \quad (43)$$

Solving with respect to \overline{P} produces

$$\overline{P} = -\frac{c\mu'}{\mu^2} \frac{A\sigma^2}{1 + c\frac{\mu'}{\mu^2} B}. \quad (44)$$

Observing that the derivative of μ in (40) with respect to ρ is $\mu' = \frac{\mu}{1 - \mu^2 F_2}$ the result in (41) follows from (44). ■

Taking the derivative of \overline{P} in (41) with respect to ρ yields (the mathematical details are omitted for space limitations)

$$\overline{P}' = 2c^2 A \sigma^2 \mu' \frac{\mu A - B}{(1 - \mu^2 F_2 - cB)^2} \quad (45)$$

with

$$A = \frac{1}{K} \sum_{i=1}^K \frac{(\alpha_i l(\mathbf{x}_i))^2}{(1 + \alpha_i l(\mathbf{x}_i) \mu)^3} \quad B = \frac{1}{K} \sum_{i=1}^K \frac{\alpha_i l(\mathbf{x}_i) \gamma_i}{(1 + \alpha_i l(\mathbf{x}_i) \mu)^3}.$$

From (45), it turns out that the optimal μ is such that $\mu^* = B/A$ from which using (40) the optimal ρ is found to be in the form of (23) in the text.

ACKNOWLEDGMENT

The authors thank Dr. Romain Couillet for helpful discussions on the large system analysis of the optimal linear precoding and in particular for the results of Theorem 1.

REFERENCES

- [1] Q. Li, G. Li, W. Lee, M. il Lee, D. Mazzarese, B. Clerckx, and Z. Li, "MIMO techniques in WiMAX and LTE: a feature overview," *IEEE Commun. Mag.*, vol. 48, no. 5, pp. 86–92, May 2010.
- [2] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sept 2006.
- [3] A. Gershman, N. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, "Convex optimization-based beamforming," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 62–75, May 2010.
- [4] E. Björnson, M. Bengtsson, and B. Ottersten, "Optimal multiuser transmit beamforming: A difficult problem with a simple solution structure [lecture notes]," *IEEE Signal Processing Magazine*, vol. 31, no. 4, pp. 142–148, July 2014.

- [5] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Veh. Tech.*, vol. 53, no. 1, pp. 18–28, Jan. 2004.
- [6] A. Wiesel, Y. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2006.
- [7] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2646–2660, Jun. 2007.
- [8] E. Björnson, G. Zheng, M. Bengtsson, and B. Ottersten, "Robust monotonic optimization framework for multicell MISO systems," *IEEE Trans. Signal Process.*, vol. 60, no. 5, pp. 2508–2523, May 2012.
- [9] Y. Chen, S. Zhang, S. Xu, and G. Li, "Fundamental trade-offs on green wireless networks," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 30–37, June 2011.
- [10] M. Bengtsson and B. Ottersten, "Optimal and suboptimal transmit beamforming," in *Handbook of Antennas in Wireless Communications*, L. C. Godara, Ed. CRC Press, 2001.
- [11] S. Lakshminarayana, J. Hoydis, M. Debbah, and M. Assaad, "Asymptotic analysis of distributed multi-cell beamforming," in *IEEE 21st International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, Sept 2010, pp. 2105–2110.
- [12] Y. Huang, C. W. Tan, and B. Rao, "Large system analysis of power minimization in multiuser MISO downlink with transmit-side channel correlation," in *International Symposium on Information Theory and its Applications (ISITA)*, Oct. 2012, pp. 240–244.
- [13] R. Zakhour and S. Hanly, "Base station cooperation on the downlink: Large system analysis," *IEEE Trans. Inf. Theory*, vol. 58, no. 4, pp. 2079–2106, Apr. 2012.
- [14] A. T. H. Asgharimoghaddam and N. Rajatheva, "Decentralizing the optimal multi-cell beamforming via large system analysis," in *Proceedings of the IEEE International Conference on Communications*, Sydney, Australia, June 2014. [Online]. Available: <http://arxiv.org/abs/1310.3843>
- [15] R. Couillet and M. McKay, "Large dimensional analysis and optimization of robust shrinkage covariance matrix estimators," *Journal of Multivariate Analysis*, vol. 131, no. 0, pp. 99 – 120, 2014.
- [16] S. Wagner, R. Couillet, M. Debbah, and D. T. M. Slock, "Large system analysis of linear precoding in correlated MISO broadcast channels under limited feedback," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4509–4537, July 2012.
- [17] R. Muharar, R. Zakhour, and J. Evans, "Optimal power allocation and user loading for multiuser MISO channels with regularized channel inversion," *IEEE Trans. Commun.*, vol. 61, no. 12, pp. 5030–5041, Dec. 2013.
- [18] 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.
- [19] L. Sanguinetti, A. L. Moustakas, and M. Debbah, "Interference management in 5G reverse TDD HetNets: A large system analysis," submitted to *IEEE J. Sel. Areas Commun.*, July 2014. [Online]. Available: <http://arxiv.org/abs/1407.6481>
- [20] L. Sanguinetti, A. L. Moustakas, E. Björnson, and M. Debbah, "Large system analysis of the energy consumption distribution in multi-user MIMO systems with mobility," submitted to *IEEE Trans. Wireless Commun.*, June 2014. [Online]. Available: <http://arxiv.org/abs/1406.5988>
- [21] F. Rusek, D. Persson, B. Lau, E. Larsson, T. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.
- [22] *Further advancements for E-UTRA physical layer aspects (Release 9)*. 3GPP TS 36.814, Mar. 2010.