Regularized ZF in Cooperative Broadcast Channels under Distributed CSIT: A Large System Analysis

Paul de Kerret*, David Gesbert*, and Umer Salim‡
* Mobile Communication Department, Eurecom
‡Intel Mobile Communications

***Abstract*—Obtaining accurate Channel State Information (CSI) at multiple Transmitters (TXs) is critical to the performance of many cooperative transmission schemes, including joint precoding in the context of network MIMO. Practical CSI feedback and limited backhaul-based sharing creates degradations of CSI which are *specific* to each TX, giving rise to a *Distributed* (D-) CSI configuration. In the D-CSI broadcast channel setting, each TX implements separate elements of the joint multi-user precoder based on its *own* individual CSI estimate. In this work, we presents a first finite-SNR regime rate analysis for a network-MIMO (broadcast channel) under a distributed CSI setting. Of particular importance is the notion of a "price of distributedness" which penalizes the D-CSI setting over the conventional centralized one with the same overall feedback quality. To tackle this problem, we apply tools from the field of Random Matrix Theory (RMT) to derive deterministic equivalents of the Signal to Interference plus Noise Ratio (SINR) for a popular class of precoders. Our key finding lies in the notion that the price of distributedness converges to a predictable value, bounded away from zero, as the number of antennas grows.** [1]

***Index Terms*—Multiuser channels, Cooperative communication, MIMO, Feedback Communications**

I. Introduction

Cooperative and coordinated transmission methods where multiple TXs exchange data and CSI related information in the hope of mitigating mutual interference are currently considered for next generation wireless networks [1]. With perfect message and CSI sharing, the different TXs can be seen as a unique virtual multiple-antenna array serving all RXs in a multiple-antenna broadcast channel (BC) fashion [2]. Existing joint precoding however requires global multi-user CSI at each TX in order to achieve near optimal sum rate performance [3].

The problem of CSI imperfections has been a central one in the literature on the BC. The cases of imperfect, noisy, or delayed CSI has been heavily investigated in the past (e.g. [3], [4]). Almost all of the past literature however typically assumes *centralized* CSIT, i.e., that the precoding is based on the basis of a *single* imperfect channel estimate which is common to every TX. Although meaningful in the case of a broadcast with a single transmitting device, this assumption can be challenged when the joint precoding is carried out across distant TXs linked by heterogeneous and imperfect backhaul links or having to communicate without backhaul (over the air) among each other, as in the case of direct device-to-device cooperation. In all these cases, it is expected that the CSI exchange will introduce further delay and quantization noise, thus making the CSI intrinsically TX-dependent. This setting is referred to as distributed CSI (D-CSI) in the rest of this paper.

From an information theoretic perspective, the study of TX cooperation in the D-CSI setting raises several intriguing and challenging open problems.

First, the capacity region of the broadcast channel under a general D-CSI setting is unknown. In [5], a rate characterization at high SNR is carried out using DoF analysis for the two TXs scenario. This study highlighted the penalty associated with the lack of a consistent CSI shared by the cooperating TXs from a DoF point of view, when using a conventional precoder. Interestingly, it was also shown that classical robust precoders (i.e. made robust with respect to centralized forms of CSI imperfections) [6] do not restore the DoF [5]. More importantly, the finite SNR performance analysis is unchartered territory. Although, the use of conventional linear precoders that are unaware of the D-CSI structure is expected to yield a loss with respect to a centralized (even imperfect) CSI setting, the quantifying of this loss in the finite SNR regime (dubbed here the "price of distributedness") has not been addressed previously.

The main goal of this paper is to study here comparatively the average rate achieved by popular precoders (namely regularized Zero Forcing (ZF)) in the centralized and distributed CSI settings. To render the problem amenable to closed form analysis, we consider the large number of antenna regime. Specifically we let the number of transmit antennas and the number of receive antennas jointly grow large with a fixed ratio, thus allowing to use efficient tools from the field of RMT. Although RMT has been applied in many works to the analysis of wireless communications [See [7]–[12] among others], its role in helping analyze cooperative systems with instantaneous distributed CSI has received little attention before.

Our main contribution consists in providing a deterministic equivalent for the average rate per user in a D-CSI setting where each TX receives its *own* estimate of the global multi-user channel matrix with the quality (in a statistical sense) of this estimate varying from TX to TX. A key finding is that although all SINR levels undergo classical hardening effect, there is a non vanishing price associated to distributed CSI feedback when compared with the centralized one.

[1]Pal de Kerret and David Gesbert acknowledge partial support of the European research project SHARING, which is partly funded by the European Union under its FP7 ICT Objective 1.1.

II. System Model

A. Transmission Model

We study a so-called network MIMO transmission where n TXs jointly serve K Receivers (RXs). We are interested in the finite-SNR rate performance at the RXs under linear precoding structures. Each TX is equipped with M_{TX} antennas and the total number of transmit antennas is denoted by $M \triangleq nM_{\mathrm{TX}}$ while every RX is equipped with a single-antenna. We assume that the ratio of transmit antennas to the number of users is fixed and given by $\beta \triangleq M/K \geq 1$.

We further assume that the RXs have perfect CSI so as to focus on the imperfectness of CSI feedback and exchange among the TXs (due to limited feedback and exchange capability). We consider that the RXs treat interference as noise. The channel from the n TXs to the K RXs is represented by the multi-user channel matrix $\mathbf{H} \in \mathbb{C}^{K\times M}$, whose elements are chosen as i.i.d. $\mathcal{N}_{\mathbb{C}}(0,1)$.

The transmission is then described as

$$\begin{bmatrix} y_1 \\ \vdots \\ y_K \end{bmatrix} = \mathbf{H}\mathbf{x} + \boldsymbol{\eta} = \begin{bmatrix} \boldsymbol{h}_1^{\mathrm{H}}\mathbf{x} \\ \vdots \\ \boldsymbol{h}_K^{\mathrm{H}}\mathbf{x} \end{bmatrix} + \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_K \end{bmatrix} \tag{1}$$

where $y_i \in \mathbb{C}$ is the signal received at the i-th RX, $\boldsymbol{h}_i^{\mathrm{H}} = \boldsymbol{e}_i^{\mathrm{H}}\mathbf{H} \in \mathbb{C}^{1\times M}$ is the channel from all transmit antennas to RX i, and $\boldsymbol{\eta} \triangleq [\eta_1,\dots,\eta_K]^{\mathrm{T}} \in \mathbb{C}^{K\times 1}$ is the normalized Gaussian noise with its elements i.i.d. as $\mathcal{N}_{\mathbb{C}}(0,1)$.

The transmitted multi-user signal $\mathbf{x} \in \mathbb{C}^{M\times 1}$ is obtained from the symbol vector $\boldsymbol{s} \triangleq [\boldsymbol{s}_1^{\mathrm{T}},\dots,\boldsymbol{s}_K^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{K\times 1}$ with its elements i.i.d. $\mathcal{N}_{\mathbb{C}}(0,1)$ as

$$\mathbf{x} = \mathbf{T}\boldsymbol{s} = \begin{bmatrix}\boldsymbol{t}_1,\dots,\boldsymbol{t}_K\end{bmatrix} \begin{bmatrix} \boldsymbol{s}_1 \\ \vdots \\ \boldsymbol{s}_K \end{bmatrix} \tag{2}$$

with $\mathbf{T} \in \mathbb{C}^{M\times K}$ being the *multi-user* precoder and $\boldsymbol{t}_i \triangleq \mathbf{T}\boldsymbol{e}_i \in \mathbb{C}^{M\times 1}$ being the beamforming vector used to transmit to RX i. We consider for simplification the sum power constraint $\|\mathbf{T}\|_{\mathrm{F}}^2 = P$.

Our main figure-of-merit is the average sum rate

$$\mathrm{R} \triangleq \frac{1}{K}\sum_{k=1}^{K}\mathbb{E}\left[\log_2\left(1+\mathrm{SINR}_k\right)\right] \tag{3}$$

where SINR_k denotes the Signal-to-Interference and Noise Ratio (SINR) at RX k and is defined as

$$\mathrm{SINR}_k \triangleq \frac{\left|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{t}_k\right|^2}{1+\sum_{\ell=1,\ell\neq k}^{K}\left|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{t}_\ell\right|^2}. \tag{4}$$

B. Distributed CSIT Model

General transmit cooperation scenarios rely on local CSI to be fedback to each TX, followed by an exchange mechanism over a wired/wireless backhaul. Backhaul links are subject to latency which cause TX-specific CSI degradation. Hence, a suitable and general CSIT model is one whereby each TX must make a precoding decision based on a TX-dependent estimate of the global channel matrix, a problem known in control theory as Team Decision [13]. Note that in our model, no further communication (or message passing) is allowed among TX. Specifically, TX j receives the multi-user channel estimate $\hat{\mathbf{H}}^{(j)} \in \mathbb{C}^{K\times M}$ and designs its transmit coefficient $\mathbf{x}_j \in \mathbb{C}^{M_{\mathrm{TX}}\times 1}$ solely as a function of $\hat{\mathbf{H}}^{(j)}$. As a first step, we assume in this work that the imperfect multi-user channel estimate is modeled by

$$\hat{\mathbf{H}}^{(j)} \triangleq \sqrt{1-(\sigma^{(j)})^2}\mathbf{H} + \sigma^{(j)}\boldsymbol{\Delta}^{(j)} \tag{5}$$

with $\boldsymbol{\Delta}^{(j)} \in \mathbb{C}^{K\times M}$ having its elements i.i.d. $\mathcal{N}_{\mathbb{C}}(0,1)$. The extension to a non uniform description quality at each TX and to correlated channel is carried out in the full version of this paper [14].

Remark 1. Importantly, the D-CSI model encompasses the conventional (centralized) CSI model by taking $n = 1$. □

C. Regularized Zero Forcing with Distributed CSIT

We are interested in the impact of the D-CSIT model on the rate performance for a conventional precoding method, hence we focus on the example of the popular *regularized ZF* MISO broadcast precoder [6], [15]. The precoder designed at TX j is then assumed to take the form

$$\mathbf{T}_{\mathrm{rZF}}^{(j)} \triangleq \left(\hat{\mathbf{H}}^{(j)}(\hat{\mathbf{H}}^{(j)})^{\mathrm{H}} + M\alpha\mathbf{I}_M\right)^{-1}\hat{\mathbf{H}}^{(j)}\frac{\sqrt{P}}{\sqrt{\Psi}} \tag{6}$$

with regularization factor $\alpha > 0$. We also define

$$\mathbf{Q}^{(j)} \triangleq \left(\frac{\hat{\mathbf{H}}^{(j)}(\hat{\mathbf{H}}^{(j)})^{\mathrm{H}}}{M} + \alpha\mathbf{I}_M\right)^{-1} \tag{7}$$

such that the precoder at TX j can be rewritten as

$$\mathbf{T}_{\mathrm{rZF}}^{(j)} = \frac{1}{M}\mathbf{Q}^{(j)}\hat{\mathbf{H}}^{(j)}\frac{\sqrt{P}}{\sqrt{\Psi^{(j)}}}. \tag{8}$$

The scalar $\Psi^{(j)}$ corresponds to the power normalization at TX j. Hence, it holds that

$$\Psi^{(j)} = \|\left(\hat{\mathbf{H}}^{(j)}(\hat{\mathbf{H}}^{(j)})^{\mathrm{H}} + M\alpha\mathbf{I}_M\right)^{-1}\hat{\mathbf{H}}^{(j)}\|_{\mathrm{F}}^2. \tag{9}$$

Upon concatenation of all TX's precoding vectors, the effective global precoder denoted by $\mathbf{T}_{\mathrm{rZF}}^{\mathrm{DCSI}}$, is equal to

$$\mathbf{T}_{\mathrm{rZF}}^{\mathrm{DCSI}} \triangleq \begin{bmatrix} \mathbf{E}_1^{\mathrm{H}}\mathbf{T}_{\mathrm{rZF}}^{(1)} \\ \mathbf{E}_2^{\mathrm{H}}\mathbf{T}_{\mathrm{rZF}}^{(2)} \\ \vdots \\ \mathbf{E}_K^{\mathrm{H}}\mathbf{T}_{\mathrm{rZF}}^{(K)} \end{bmatrix} \tag{10}$$

where $\mathbf{E}_j^{\mathrm{H}} \in \mathbb{C}^{M_{\mathrm{TX}}\times M}$ is defined as

$$\mathbf{E}_j^{\mathrm{H}} \triangleq \begin{bmatrix}\mathbf{0}_{M_{\mathrm{TX}}\times(j-1)M_{\mathrm{TX}}} & \mathbf{I}_{M_{\mathrm{TX}}} & \mathbf{0}_{M_{\mathrm{TX}}\times(n-j)M_{\mathrm{TX}}}\end{bmatrix}. \tag{11}$$

We furthermore denote the kth column of $\mathbf{T}_{\mathrm{rZF}}^{\mathrm{DCSI}}$ (used to serve RX k) by $\boldsymbol{t}_{\mathrm{rZF},k}^{\mathrm{DCSI}}$.

Although the finite SNR rate analysis in the distributed CSI model in (5) is challenging in the general case because of the dependency of one user performance on all channel estimates, some useful insights can be obtained in the large antenna regime, as shown below.

III. Deterministic Equivalent of the SINR

Our approach relies on the use of the following theorem which is the basis of the RMT calculations based on the Stieltjes transform.

Theorem 1. *[10], [16] Consider the resolvent matrix* $\mathbf{Q} \triangleq \left(\frac{\mathbf{H}^{\mathrm{H}}\mathbf{H}}{M} + \alpha \mathbf{I}_M\right)^{-1}$ *with the matrix* $\mathbf{H}$ *defined according to Section II and* $\alpha > 0$*. Let the matrix* $\mathbf{U}$ *be any matrix with bounded spectral norm. Then,*

$$\frac{1}{M}\operatorname{tr}(\mathbf{UQ}) - \frac{\delta}{M}\operatorname{tr}(\mathbf{U}) \xrightarrow[K,M\to\infty]{a.s.} 0 \tag{12}$$

where δ *is the unique fixed point (and can be written trivially in closed form) of the equation*

$$x = \frac{1}{\alpha + \frac{1}{\beta(1+x)}}. \tag{13}$$

Using this theorem and the definition of δ, we can now state our main result.

Theorem 2. *Considering the D-CSI model described in Section II, then*

$$\mathrm{SINR}_k - \mathrm{SINR}_k^o \xrightarrow[K,M\to\infty]{a.s.} 0 \tag{14}$$

with SINR_k^o *defined as*

$$\mathrm{SINR}_k^o \triangleq \frac{\left(\frac{1}{n}\sum_{j=1}^{n}\sqrt{1-(\sigma^{(j)})^2}\right)^2 \frac{\delta^2}{(1+\delta)^2}}{I_k^o + \frac{\Gamma^o}{P}} \tag{15}$$

with $I_k^o \in \mathbb{R}$ *given by*

$$I_k^o \triangleq \sum_{j=1}^{n} \frac{\Gamma^o}{(1+\delta)^2 n^2}\left[1-(\sigma^{(j)})^2+(1+\delta)^2\left(-1+n+(\sigma^{(j)})^2\right)\right]$$
$$+\sum_{j=1}^{n}\sum_{\substack{j'=1\\ j'\neq j}}^{n} \frac{\Gamma^o_{j,j'}\delta}{(1+\delta)^2 n^2}\left[-1+(1+\delta)\left(-1+(\sigma^{(j)})^2+(\sigma^{(j')})^2\right)\right] \tag{16}$$

while $\Gamma^o \in \mathbb{R}$ *and* $\Gamma^o_{j,j'} \in \mathbb{R}$ *are respectively defined as*

$$\Gamma^o \triangleq \frac{\delta^2}{\beta(1+\delta)^2 - \delta^2} \tag{17}$$

$$\Gamma^o_{j,j'} \triangleq \frac{\sqrt{\left(1-(\sigma^{(j)})^2\right)\left(1-(\sigma^{(j')})^2\right)}\delta^2}{\beta(1+\delta)^2 - \left(1-(\sigma^{(j)})^2\right)\left(1-(\sigma^{(j')})^2\right)\delta^2}. \tag{18}$$

As expected, taking $\sigma^{(j)} = \sigma^{(j')} = 0$ in $\Gamma^o_{j,j'}$ gives the expression for Γ^o. The deterministic equivalent of the SINR then simplifies to

$$\mathrm{SINR}_k^o = \frac{\delta^2}{\Gamma^o\left(1 + \frac{(1+\delta)^2}{P}\right)}. \tag{19}$$

Remark 2. Equation (19) does not depend on n. This follows from the fact that the distributed aspect of the CSI becomes trivial when the CSI is perfect. Hence, only the total number of transmit antennas (i.e., $\beta = M/K$) becomes a significant parameter. □

Furthermore, taking $n = 1$ corresponds to the centralized CSI configuration and the deterministic equivalent simplifies to:

$$\mathrm{SINR}_k^o = \frac{(1-(\sigma^{(1)})^2\delta^2}{\Gamma^o\left((1-(\sigma^{(1)})^2+(1+\delta)^2(\sigma^{(1)})^2+\frac{(1+\delta)^2}{P}\right)}. \tag{20}$$

These two particular cases correspond to a centralized CSI configuration and it can be verified that the expressions obtained above match with the results for the centralized CSI configuration given in [11, Theorem 14.1].

IV. Proof of Theorem 2

Due to space limitations, only a sketch of the proof is given in the main body of this paper. In particular, the detailed steps of the derivation of a deterministic equivalent for the interference term are relegated to the extended version available online in [17]. Our calculation is built upon results from both [9] and [10]. We also make extensive use of classical RMT lemmas recalled in the Appendix. Note that Lemma 5 and Lemma 6 are novel and the proofs can be found in [17]. In particular, Lemma 5 extends [10, Lemma 15] and is an interesting result in itself.

During the calculation we use the notation $x \asymp y$ to denote that $x - y \xrightarrow[K,M\to\infty]{a.s.} 0$.

A. Deterministic Equivalent for $\Psi^{(j)}$

We start by finding a deterministic equivalent for $\Psi^{(j)}$. A deterministic equivalent for $\Psi^{(j)}$ can be found in [9], but it can also be obtained using Lemma 5 with $\sigma^{(j)} = \sigma^{(j')} = 0$, which gives

$$\Psi^{(j)} \asymp \Gamma^o. \tag{21}$$

It can be noted that, as expected from the definition of $\Psi^{(j)}$, this deterministic equivalent does not depend on $\sigma^{(j)}$.

B. Deterministic Equivalent for $\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{t}_{\mathrm{rZF},k}^{\mathrm{DCSI}}$:

Turning to the desired signal at RX k, we can write

$$\begin{aligned}
\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{t}_{\mathrm{rZF},k}^{\mathrm{DCSI}} &= \sum_{j=1}^{n}\frac{1}{M}\frac{\sqrt{P}}{\sqrt{\Psi^{(j)}}}\boldsymbol{h}_k^{\mathrm{H}}\mathbf{E}_j\mathbf{E}_j^{\mathrm{H}}\mathbf{Q}^{(j)}\hat{\boldsymbol{h}}_k^{(j)} \\
&\overset{(a)}{\asymp} \sqrt{\frac{P}{\Gamma^o}}\sum_{j=1}^{n}\frac{\frac{1}{M}\boldsymbol{h}_k^{\mathrm{H}}\mathbf{E}_j\mathbf{E}_j^{\mathrm{H}}\mathbf{Q}_{[k]}^{(j)}\hat{\boldsymbol{h}}_k^{(j)}}{1+\frac{1}{M}\boldsymbol{h}_k^{\mathrm{H}}\mathbf{Q}_{[k]}^{(j)}\boldsymbol{h}_k} \\
&\overset{(b)}{\asymp} \sqrt{\frac{P}{\Gamma^o}}\sum_{j=1}^{n}\sqrt{1-(\sigma^{(j)})^2}\frac{\frac{1}{M}\boldsymbol{h}_k^{\mathrm{H}}\mathbf{E}_j\mathbf{E}_j^{\mathrm{H}}\mathbf{Q}_{[k]}^{(j)}\boldsymbol{h}_k}{1+\frac{1}{M}\boldsymbol{h}_k^{\mathrm{H}}\mathbf{Q}_{[k]}^{(j)}\boldsymbol{h}_k} \\
&\overset{(c)}{\asymp} \sqrt{\frac{P}{\Gamma^o}}\sum_{j=1}^{n}\sqrt{1-(\sigma^{(j)})^2}\frac{\frac{1}{M}\operatorname{tr}\left(\mathbf{E}_j\mathbf{E}_j^{\mathrm{H}}\mathbf{Q}_{[k]}^{(j)}\right)}{1+\frac{1}{M}\operatorname{tr}\left(\mathbf{Q}_{[k]}^{(j)}\right)} \\
&\overset{(d)}{\asymp} \sqrt{\frac{P}{\Gamma^o}}\frac{1}{n}\sum_{j=1}^{n}\sqrt{1-(\sigma^{(j)})^2}\frac{\delta}{1+\delta}
\end{aligned} \tag{22}$$

where (a) follows from Lemma 1 and the use of the deterministic equivalent derived for $\Psi^{(j)}$, (b) from Lemma 3, (c) from Lemma 2, (d) from Lemma 4, the fundamental Theorem 1 and the unitary invariance of the distribution of $\mathbf{H}$. It follows then directly that

$$\left|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{t}_{\mathrm{rZF},k}^{\mathrm{DCSI}}\right|^2 \asymp \frac{P}{\Gamma^o}\left(\frac{1}{n}\sum_{j=1}^{n}\sqrt{1-(\sigma^{(j)})^2}\right)^2 \frac{\delta^2}{(1+\delta)^2}. \tag{23}$$

C. Deterministic Equivalent for the Interference Term

Our first step is to write explicitly the interference term using the definition of $\mathbf{T}^{\mathrm{DCSI}}$ in Subsection II-C and replace $\Psi^{(j)}$ by its deterministic equivalent Γ^o

$$\begin{aligned}
\mathcal{I}_k &\triangleq \sum_{\ell=1,\ell\neq k}^{K} |\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{t}_{\mathrm{rZF},\ell}^{\mathrm{DCSI}}|^2 \\
&= \boldsymbol{h}_k^{\mathrm{H}}\mathbf{T}_{\mathrm{rZF}}^{\mathrm{DCSI}}(\mathbf{T}_{\mathrm{rZF}}^{\mathrm{DCSI}})^{\mathrm{H}}\boldsymbol{h}_k - \boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{t}_{\mathrm{rZF},k}^{\mathrm{DCSI}}(\boldsymbol{t}_{\mathrm{rZF},k}^{\mathrm{DCSI}})^{\mathrm{H}}\boldsymbol{h}_k \\
&= \frac{1}{M^2}\sum_{j=1}^{n}\sum_{j'=1}^{n}\frac{P}{\sqrt{\Psi^{(j)}}\sqrt{\Psi^{(j')}}}\boldsymbol{h}_k^{\mathrm{H}}\mathbf{E}_j\mathbf{E}_j^{\mathrm{H}}\mathbf{Q}^{(j)}(\mathbf{H}_{[k]}^{(j)})^{\mathrm{H}} \\
&\qquad\qquad \cdot \mathbf{H}_{[k]}^{(j')}\mathbf{Q}^{(j')}\mathbf{E}_{j'}\mathbf{E}_{j'}^{\mathrm{H}}\boldsymbol{h}_k \\
&\asymp \frac{P}{\Gamma^o}\frac{1}{M^2}\sum_{j=1}^{n}\sum_{j'=1}^{n}\boldsymbol{h}_k^{\mathrm{H}}\mathbf{E}_j\mathbf{E}_j^{\mathrm{H}}\mathbf{Q}_{[k]}^{(j)}(\mathbf{H}_{[k]}^{(j)})^{\mathrm{H}} \\
&\qquad\qquad \cdot \mathbf{H}_{[k]}^{(j')}\mathbf{Q}^{(j')}\mathbf{E}_{j'}\mathbf{E}_{j'}^{\mathrm{H}}\boldsymbol{h}_k \\
&\quad + \frac{P}{\Gamma^o}\frac{1}{M^2}\sum_{j=1}^{n}\sum_{j'=1}^{n}\boldsymbol{h}_k^{\mathrm{H}}\mathbf{E}_j\mathbf{E}_j^{\mathrm{H}}\left(\mathbf{Q}^{(j)}-\mathbf{Q}_{[k]}^{(j)}\right)(\mathbf{H}_{[k]}^{(j)})^{\mathrm{H}} \\
&\qquad\qquad \cdot \mathbf{H}_{[k]}^{(j')}\mathbf{Q}^{(j')}\mathbf{E}_{j'}\mathbf{E}_{j'}^{\mathrm{H}}\boldsymbol{h}_k.
\end{aligned} \tag{24}$$

We can apply Lemma 2 for the first term to obtain that

$$\begin{aligned}
&\frac{1}{M^2}\boldsymbol{h}_k^{\mathrm{H}}\mathbf{E}_j\mathbf{E}_j^{\mathrm{H}}\mathbf{Q}_{[k]}^{(j)}(\mathbf{H}_{[k]}^{(j)})^{\mathrm{H}}\mathbf{H}_{[k]}^{(j')}\mathbf{Q}^{(j')}\mathbf{E}_{j'}\mathbf{E}_{j'}^{\mathrm{H}}\boldsymbol{h}_k \\
&\asymp \frac{1}{M^2}\operatorname{tr}\left(\mathbf{E}_{j'}\mathbf{E}_{j'}^{\mathrm{H}}\mathbf{E}_j\mathbf{E}_j^{\mathrm{H}}\mathbf{Q}_{[k]}^{(j)}(\mathbf{H}_{[k]}^{(j)})^{\mathrm{H}}\mathbf{H}_{[k]}^{(j')}\mathbf{Q}^{(j')}\right).
\end{aligned} \tag{25}$$

It is then possible to apply Lemma 5 to obtain a deterministic of the expression in (25). To obtain a deterministic equivalent for the second term of (24), we use the following relation

$$\mathbf{Q}^{(j)} - \mathbf{Q}_{[k]}^{(j)} = \mathbf{Q}^{(j)}\left((\mathbf{Q}_{[k]}^{(j)})^{-1} - (\mathbf{Q}^{(j)})^{-1}\right)\mathbf{Q}_{[k]}^{(j)}. \tag{26}$$

Inserting (26) in the second term of (24) and using Lemma 6 provides expressions for which it is possible to apply Lemma 5 as in (25). Putting all the terms together and simplifying concludes the proof.

V. Simulation Results

We now verify using Monte-Carlo simulations the accuracy of the asymptotic expression derived in Theorem 2. We consider a network consisting of $n = 3$ TXs with a sum power constraint given by $P = 10$ dB and we assume that $(\sigma^{(j)})^2 = \sigma^2 = 0.1, \forall j = 1, \ldots, n$.

Remark 3. The fact that error variances are equally distributed does not imply a centralized CSIT as CSI errors at the various TXs remain *independent* of each other. □

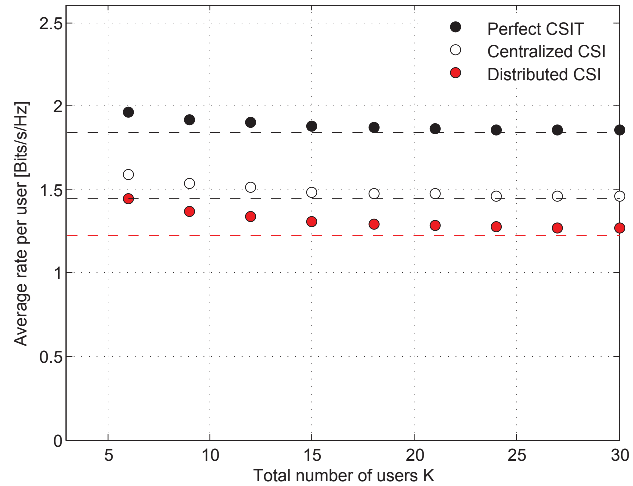

Fig. 1: Average rate per user as a function of the number of users K with $(\sigma^{(j)})^2 = 0.1, \forall j$.

We use for α its optimal value for the centralized case, given by [9, eq. (53)]

$$\alpha^{\mathrm{CCSI}} = \frac{1+\sigma^2 P}{1-\sigma^2}\frac{1}{\beta P}. \tag{27}$$

In Fig. 1, we show the rate per user as a function of the number of users for a square setting where $M = nM_{\mathrm{TX}} = K$ (i.e., $\beta = 1$). For comparison purpose, we also show the rate per user obtained in the case of centralized CSIT with $(\sigma^{\mathrm{CCSI}})^2 = 0.1$ and with perfect CSIT (obtained using $n = 1$ in Theorem 2). The large system deterministic equivalents are shown to be accurate with just 20 to 30 users and antennas. The cost of having distributed information is also highlighted by the losses compared to the centralized configuration *for the same average feedback quality*.

VI. Conclusion

We have studied regularized ZF joint precoding in a distributed CSI configuration. Using RMT tools, an analytical expression has been derived to approximate the average rate per user in the large system limits. This new deterministic equivalent reveals the cost related not just to CSI feedback limitation, but also to backhaul sharing limitations, and can be helpful in terms of robust system design. The extension to more general channel and CSI models are subject to ongoing work [14]. The price of distributedness is evaluated here for a conventional precoder, which further motivates the development of novel precoding schemes being more suitable to the distributed CSI setting.

Appendix

A. Classical Lemmas from the Literature

Lemma 1 (Resolvent Identities [10], [11])**.** *Given any matrix $\mathbf{H} \in \mathbb{C}^{K\times M}$, let $\boldsymbol{h}_k^{\mathrm{H}}$ denote its kth row and $\mathbf{H}_k \in \mathbb{C}^{(K-1)\times M}$ denote the matrix obtained after removing the kth* *row from* $\mathbf{H}$. *The resolvent matrices of* $\mathbf{H}$ *and* $\mathbf{H}_k$ *are denoted by* $\mathbf{Q} \triangleq \left(\mathbf{H}^{\mathrm{H}}\mathbf{H}+\alpha\mathbf{I}_M\right)^{-1}$ *and* $\mathbf{Q}_k \triangleq \left(\mathbf{H}_k^{\mathrm{H}}\mathbf{H}_k+\alpha\mathbf{I}_M\right)^{-1}$ *with* $\alpha>0$ *respectively. It then holds that*

$$\mathbf{Q}=\mathbf{Q}_k-\frac{1}{M}\frac{\mathbf{Q}_k\boldsymbol{h}_k\boldsymbol{h}_k^{\mathrm{H}}\mathbf{Q}_k}{1+\frac{1}{M}\boldsymbol{h}_k^{\mathrm{H}}\mathbf{Q}_k\boldsymbol{h}_k} \tag{28}$$

and

$$\boldsymbol{h}_k^{\mathrm{H}}\mathbf{Q}=\frac{\boldsymbol{h}_k^{\mathrm{H}}\mathbf{Q}_k}{1+\frac{1}{M}\boldsymbol{h}_k^{\mathrm{H}}\mathbf{Q}_k\boldsymbol{h}_k}. \tag{29}$$

Lemma 2 ([10], [11])**.** *Let* $(\mathbf{A}_N)_{N\geq1},\mathbf{A}_N\in\mathbb{C}^{N\times N}$ *be a sequence of matrices such that* $\limsup\|\mathbf{A}_N\|<\infty$*, and* $(\mathbf{x}_N)_{N\geq1},\mathbf{x}_N\in\mathbb{C}^{N\times1}$ *be a sequence of random vectors of i.i.d. entries of zero mean, unit variance, and finite 8th order moment independent of* $\mathbf{A}_N$*. Then,*

$$\frac{1}{N}\mathbf{x}_N^{\mathrm{H}}\mathbf{A}_N\mathbf{x}_N-\frac{1}{N}\operatorname{tr}\left(\mathbf{A}_N\right)\xrightarrow[N\to\infty]{a.s.}0. \tag{30}$$

Lemma 3 ([10], [11])**.** *Let* $(\mathbf{A}_N)_{N\geq1},\mathbf{A}_N\in\mathbb{C}^{N\times N}$ *be a sequence of matrices such that* $\limsup\|\mathbf{A}_N\|<\infty$*, and* $\mathbf{x}_N,\mathbf{y}_N$ *be random, mutually independent with i.i.d. entries of zero mean, unit variance, finite 8th order moment, and independent of* $\mathbf{A}_N$*. Then,*

$$\frac{1}{N}\mathbf{x}_N^{\mathrm{H}}\mathbf{A}_N\mathbf{y}_N\xrightarrow[N\to\infty]{a.s.}0. \tag{31}$$

Lemma 4 ([9], [11])**.** *Let* $\mathbf{Q}$ *and* $\mathbf{Q}_k$ *be as given in Lemma 1. Then, for any matrix* $\mathbf{A}$*, we have*

$$\operatorname{tr}\left(\mathbf{A}\left(\mathbf{Q}-\mathbf{Q}_k\right)\right)\leq\|\mathbf{A}\|_2. \tag{32}$$

B. New Lemmas

Lemma 5. *Let* $\hat{\mathbf{H}}^{(j)}$ *(resp.* $\hat{\mathbf{H}}^{(j')}$*) be the imperfect multi-user channel estimate at TX* j *(resp. TX* j'*) as described in Section II. Let* $\mathbf{Q}^{(j)}\triangleq\left(\frac{(\mathbf{H}^{(j)})^{\mathrm{H}}\mathbf{H}^{(j)}}{M}+\alpha\mathbf{I}_M\right)^{-1}$ *and* $\mathbf{Q}^{(j')}\triangleq\left(\frac{(\mathbf{H}^{(j')})^{\mathrm{H}}\mathbf{H}^{(j')}}{M}+\alpha\mathbf{I}_M\right)^{-1}$ *with* $\alpha>0$ *and* $j\neq j'$*. Then,*

$$\frac{1}{M^2}\operatorname{tr}\left(\mathbf{A}\mathbf{Q}^{(j)}(\mathbf{H}^{(j)})^{\mathrm{H}}\mathbf{H}^{(j')}\mathbf{Q}^{(j')}\right)-\frac{\frac{1}{M}\operatorname{tr}(\mathbf{A})\,\delta^2\sqrt{c_0^{(j)}c_0^{(j')}}}{\beta(1+\delta)^2}\left(1+\sqrt{c_0^{(j)}c_0^{(j')}}\mathrm{Y}_0\right)\xrightarrow[N\to\infty]{a.s.}0 \tag{33}$$

with $c_0^{(j)}\triangleq1-(\sigma^{(j)})^2$*,* $c_0^{(j')}\triangleq1-(\sigma^{(j')})^2$*,* δ *defined in Theorem 1, and* Y_0 *defined as*

$$Y_0\triangleq\frac{\sqrt{c_0^{(j)}c_0^{(j')}}\delta^2}{\left(\beta\left(1+\delta\right)^2-c_0^{(j)}c_0^{(j')}\delta^2\right)}. \tag{34}$$

Note that in the case where $\mathbf{A}=\mathbf{I}_M$*, the result simplifies to*

$$\frac{1}{M^2}\operatorname{tr}\left(\mathbf{Q}^{(j)}(\mathbf{H}^{(j)})^{\mathrm{H}}\mathbf{H}^{(j')}\mathbf{Q}^{(j')}\right)-Y_0\xrightarrow[N\to\infty]{a.s.}0. \tag{35}$$

Lemma 6. *Let* $\mathbf{L},\mathbf{R},\bar{\mathbf{A}}\in\mathbb{C}^{M\times M}$ *be of uniformly bounded spectral norm with respect to* M *and let* $\bar{\mathbf{A}}$ *be invertible. Let* $\mathbf{x},\mathbf{y}$ *have i.i.d. complex entries of zero mean, variance* $1/M$ *and finite 8th order moment and be mutually independent as well as independent of* $\mathbf{L},\mathbf{R},\bar{\mathbf{A}}$*. Then we have:*

$$\mathbf{x}^{\mathrm{H}}\mathbf{L}\mathbf{A}^{-1}\mathbf{R}\mathbf{x}\asymp u_{\mathrm{LR}}-c_0u_{\mathrm{L}}u_{\mathrm{R}}\frac{1+c_1u}{1+u}+c_2^2u_{\mathrm{L}}u_{\mathrm{R}}\frac{u}{1+u}$$
$$\mathbf{x}^{\mathrm{H}}\mathbf{L}\mathbf{A}^{-1}\mathbf{R}\mathbf{y}\asymp u_{\mathrm{LR}}+c_1c_2u_{\mathrm{L}}u_{\mathrm{R}}\frac{u}{1+u}-c_2u_{\mathrm{L}}u_{\mathrm{R}}\frac{1+c_1u}{1+u}$$

with

$$\mathbf{A}=\bar{\mathbf{A}}+c_0\mathbf{x}\mathbf{x}^{\mathrm{H}}+c_1\mathbf{y}\mathbf{y}^{\mathrm{H}}+c_2\mathbf{x}\mathbf{y}^{\mathrm{H}}+c_2\mathbf{y}\mathbf{x}^{\mathrm{H}}$$

with $c_0+c_1=1$ *and* $c_0c_1-c_2^2=0$*, and*

$$u\triangleq\frac{\operatorname{tr}(\bar{\mathbf{A}}^{-1})}{M},\qquad u_{\mathrm{L}}\triangleq\frac{\operatorname{tr}(\mathbf{L}\bar{\mathbf{A}}^{-1})}{M},$$
$$u_{\mathrm{R}}\triangleq\frac{\operatorname{tr}(\bar{\mathbf{A}}^{-1}\mathbf{R})}{M},\qquad u_{\mathrm{LR}}\triangleq\frac{\operatorname{tr}(\mathbf{L}\bar{\mathbf{A}}^{-1}\mathbf{R})}{M}.$$

REFERENCES

[1] 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 J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1380–1408, Dec. 2010.

[2] M. K. Karakayali, G. J. Foschini, and R. A. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *IEEE Wireless Communications*, vol. 13, no. 4, pp. 56–61, Aug. 2006.

[3] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.

[4] M. Maddah-Ali and D. Tse, "Completely stale transmitter channel state information is still very useful," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4418–4431, Jul. 2012.

[5] P. de Kerret and D. Gesbert, "Degrees of freedom of the network MIMO channel with distributed CSI," *IEEE Trans. Inf. Theory*, vol. 58, no. 11, pp. 6806–6824, Nov. 2012.

[6] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication-part I: channel inversion and regularization," *IEEE Trans. on Commun.*, vol. 53, no. 1, pp. 195–202, 2005.

[7] B. Hochwald and S. Vishwanath, "Space-time multiple access: Linear growth in the sum rate," in *Proc. Allerton Conference on Communication, Control, and Computing (Allerton)*, 2002.

[8] A. Tulino and S. Verdu, *Random matrix theory and wireless communications*. Now Publisher Inc., 2004.

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

[10] A. Müller, A. Kammoun, E. Björnson, and M. Debbah, "Linear precoding based on polynomial expansion: reducing complexity in massive MIMO," 2013. [Online]. Available: http://arxiv.org/abs/1310.1806

[11] R. Couillet and M. Debbah, *Random matrix methods for wireless Communications*. Cambridge University Press, 2011.

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

[13] R. Radner, "Team decision problems," *The Annals of Mathematical Statistics*, 1962.

[14] P. de Kerret and D. Gesbert, "Performance of regularized Zero Forcing with distributed CSIT: A large system analysis," 2015, in preparation.

[15] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO Channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, Feb. 2004.

[16] W. Hachem, O. Khorunzhiy, P. Loubaton, J. Najim, and L. Pastur, "A new approach for mutual information analysis of larger dimensional multi-antenna channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 3987–4004, Sep. 2008.

[17] P. de Kerret, D. Gesbert, and U. Salim, "Large system analysis of joint regularized Zero Forcing precoding with distributed CSIT," 2015. [Online]. Available: http://arxiv.org/abs/1502.03654