53rd Allerton Conference on Communication, Control and Computing Allerton, Iowa (USA), Sept. 2015

DOI: 10.1109/ALLERTON.2015.7447023

Robust Regularized ZF in decentralized Broadcast Channel with correlated CSIT

Q. Li
P. De Kerret
D. Gesbert
N. Gresset

Robust Regularized ZF in Cooperative Broadcast Channel under Distributed CSIT

Qianrui Li*‡, Paul de Kerret†, David Gesbert*, and Nicolas Gresset‡
* Mobile Communication Department, Eurecom
† Télécom Bretagne, IMT, UMR CNRS 3192 Lab-STICC
‡Mitsubishi Electric R&D Centre Europe

***Abstract*—We consider in this work the Distributed CSI (D-CSI) broadcast channel setting, in which the various TXs design elements of the precoder based on their individual estimates of the global multi-user channel matrix. Previous works relative to the D-CSI setting assume the estimation errors to be uncorrelated. In contrast, we generalize this model by allowing for correlation between the errors. This extended model then bridges the gap between the distributed and the centralized setting as choosing correlation coefficients equal to one corresponds to a centralized setting. In addition, we allow in our analysis for the TXs to use different regularization coefficients. Building upon random matrix theory tools, we are then able to obtain a deterministic equivalent for the rate achieved in the large system limit. In spite of the extreme generality of the setting studied, the expressions obtained remain easy to compute. Building upon the expression obtained, it becomes then possible to optimize the choice of the regularization coefficients at the different TXs. Interestingly, it is *not* always optimal to use the same regularization coefficient at each TX. We call this scheme where each TX applies the optimal regularization coefficient "D-CSI regularized ZF" and we show by numerical simulations that it allows to reduce the impact of the distributed CSI configuration.**

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

I. Introduction

Network (or Multi-cell) MIMO methods, whereby multiple interfering TXs share user messages and allow for joint precoding, 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, and well known precoding algorithms from the literature can be used [2]. Joint precoding however requires the feedback of an accurate multi-user CSI to each TX in order to achieve near optimal sum rate performance [3].

Although the case of imperfect, noisy or delayed, CSI has been considered in past work [3], [4], literature typically assumes *centralized* CSIT, i.e., that the precoding is done on the basis of a *single* imperfect channel estimate which is common at every TX. Although meaningful in the case of a broadcast with a single transmit 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 these cases, it is expected that the CSI exchange will introduce further delay and quantization noise. It is thus practically relevant for joint precoding across distant TXs to consider a CSI setting where each TX receives its *own* channel estimate. This setting is referred to as distributed CSI (D-CSI) in the rest of this paper.

From an information theoretic perspective, the study of transmitter cooperation in the D-CSI broadcast channel setting raises several intriguing and challenging questions.

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 transmitters scenario. This study highlighted the severe penalty caused by the lack of a consistent CSI shared by the cooperating TXs from a DoF point of view, when using a conventional precoder. It was also shown that classical (regularized) robust precoders [6] do not restore the DoF. Although a new DoF-restoring decentralized precoding strategy was presented in [5] for the two TXs case, the general case of more than 2 TXs remains open. At finite SNR, the problem of designing precoders that optimally tackle the D-CSI setting is open for any number of TXs. 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. Hence, an important question is how to reduce the the loss due to the D-CSI configuration, i.e., how to derive a D-CSI robust precoding scheme. This is exactly the question that we investigate in this work.

We study the setting where 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. RMT has been already applied in many works to the analysis of wireless communications [See [7]–[11] among others], and in a previous work, the authors have derived a deterministic equivalent for the sum rate achieved in the distributed CSI setting [12]. Note that such expression become more and more relevant and interesting with the developement of the so-called *massive MIMO* TXs having a large number of antennas [13].

Our main contributions are as described below:

- We introduce a novel CSI model where the estimation errors at the different TXs can be arbitraily correlated. This CSI model is specially adapted to model delay/imperfections in the links between the TXs and partially centralized networks.

Preprint submitted to 53rd Annual Allerton Conference on Communication, Control, and Computing. Received July 6, 2015.

- We allow each TX to use a different regularization coefficient. This extension is particularly interesting as it opens the door to the design of a robust precoding scheme taking into account the distributed structure of the CSI configuration.

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

II. System Model

A. Transmission Model

We study a communication system where n TXs serve jointly K receivers (RXs) over a network MIMO channel. Each TX is equipped with M_{TX} antennas such that the total number of transmit antennas is denoted by $M = nM_{\mathrm{TX}}$. Every RX is equipped with a single antenna. We assume that the ratio of transmit antennas with respect to the number of users is fixed and given by

$$\beta = \frac{M}{K} \geq 1. \tag{1}$$

We further assume that the RXs have perfect CSI so as to focus on the imperfect CSI feedback and exchange among the TXs. 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 $\mathbf{H} \in \mathbb{C}^{K\times M}$.

Considering linear precoding, the transmission is then described as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \boldsymbol{\eta} = \begin{bmatrix} \boldsymbol{h}_1^{\mathrm{H}} \\ \vdots \\ \boldsymbol{h}_K^{\mathrm{H}} \end{bmatrix} \mathbf{x} + \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_K \end{bmatrix} \tag{2}$$

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 $\boldsymbol{\eta} = [\eta_1, \ldots, \eta_K]^{\mathrm{T}} \in \mathbb{C}^{K\times 1}$ is the normalized Gaussian noise with its elements i.i.d. as $\mathcal{CN}(0,1)$.

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

$$\mathbf{x} = \mathbf{T}\boldsymbol{s} = \begin{bmatrix} \boldsymbol{t}_1, \ldots, \boldsymbol{t}_K \end{bmatrix} \begin{bmatrix} \boldsymbol{s}_1 \\ \vdots \\ \boldsymbol{s}_K \end{bmatrix} \tag{3}$$

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

Our main figure-of-merit is the average rate per user

$$\mathrm{R} = \frac{1}{K}\sum_{k=1}^{K} \mathbb{E}\left[\log_2\left(1 + \mathrm{SINR}_k\right)\right] \tag{4}$$

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

$$\mathrm{SINR}_k = \frac{\left|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{t}_k\right|^2}{1 + \sum_{\ell=1,\ell\neq k}^{K}\left|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{t}_\ell\right|^2}. \tag{5}$$

B. Distributed CSIT Model

In the distributed CSIT model studied here, each TX receives its own CSI based on which it designs its transmission parameters without any additional communication to the other TXs. Specifically, TX j receives the multi-user channel estimate $\hat{\mathbf{H}}^{(j)} \in \mathbb{C}^{K\times M}$ and designs its transmit coefficients $\mathbf{x}_j \in \mathbb{C}^{M_{\mathrm{TX}}\times 1}$ solely as a function of $\hat{\mathbf{H}}^{(j)}$. We model the imperfect channel estimate for RX k at TX j as

$$\hat{\boldsymbol{h}}_k^{(j)} \triangleq \sqrt{1 - (\sigma_k^{(j)})^2}\boldsymbol{h}_k + \sigma_k^{(j)}\boldsymbol{\delta}_k^{(j)} \tag{6}$$

with $\boldsymbol{\delta}_k^{(j)} \in \mathbb{C}^{M\times 1}$ having its elements i.i.d. $\mathcal{N}_{\mathbb{C}}(0,1)$. This model is widely used in the literature and is adapted to model a vector quantization [3], [9]. $(\hat{\boldsymbol{h}}_k^{(j)})^{\mathrm{H}}$ is the kth row of $\hat{\mathbf{H}}^{(j)}$, i.e., the estimate at TX j of the channel from all the transmit antennas to RX k. $\sigma_k^{(j)} \in [0,1]$ is a parameter indicating the accuracy or quality of the user channel of RX k at TX j. The estimation error is assumed to be independent of the channel. However, the estimations errors at the different TXs can be arbitrarily correlated. Hence, the error terms $\boldsymbol{\delta}_k^{(j)}$ and $\boldsymbol{\delta}_k^{(j')}$ satisfy

$$\mathbb{E}\left[\boldsymbol{\delta}_k^{(j)}(\boldsymbol{\delta}_k^{(j')})^{\mathrm{H}}\right] = (\rho_k^{(j,j')})^2\mathbf{I}_M \tag{7}$$

where parameter $\rho_k^{(j,j')}$ reflects the correlation between the two error terms. Note that $\rho_k^{(j,j')} = 1$ when $j = j'$.

This distributed CSI model allowing for correlation between the errors at the different TXs is very general and is one of the contribution of this paper. Indeed, this model allows to bridge the gap between the two extreme configuration: Distributed with *independent* CSI errors and centralized CSI. Indeed, the CSI configuration where

$$\sigma_k^{(j)} = \sigma_k^{(j')},\ \rho_k^{(j,j')} = 1, \qquad \forall j,j' = 1,\ldots,n,\ k = 1,\ldots,K \tag{8}$$

corresponds to the centralized CSI configuration while taking $\rho_k^{(j,j')} = 0$ is the distributed CSI configuration as previously studied in the literature.

The major interest of this model is to allow for all the intermediate CSI configuration where the CSI can then be seen as *partially centralized*. This is particularly adapted to model *imperfect* CSI exchange where delay and/or imperfections are introduced.

C. Regularized Zero Forcing with Distributed CSIT

We address the performance of a classical MISO broadcast precoder, namely *regularized ZF* [6], [14], when faced with distributed CSIT in the large system regime. Hence, the precoder designed at TX j is assumed to take the form

$$\mathbf{T}_{\mathrm{rZF}}^{(j)} = \left((\hat{\mathbf{H}}^{(j)})^{\mathrm{H}}\hat{\mathbf{H}}^{(j)} + M\alpha^{(j)}\mathbf{I}_M\right)^{-1}(\hat{\mathbf{H}}^{(j)})^{\mathrm{H}}\frac{\sqrt{P}}{\sqrt{\Psi^{(j)}}} \tag{9}$$

with regularization factor $\alpha^{(j)} > 0$. We also define

$$\mathbf{C}^{(j)} = \frac{(\hat{\mathbf{H}}^{(j)})^{\mathrm{H}}\hat{\mathbf{H}}^{(j)}}{M} + \alpha^{(j)}\mathbf{I}_M \tag{10}$$

Preprint submitted to 53rd Annual Allerton Conference on Communication,
Control, and Computing. Received July 6, 2015.

such that the precoder at TX j can be rewritten as

$$\mathbf{T}_{\mathrm{rZF}}^{(j)} = \frac{1}{M}(\mathbf{C}^{(j)})^{-1}(\hat{\mathbf{H}}^{(j)})^{\mathrm{H}}\frac{\sqrt{P}}{\sqrt{\Psi^{(j)}}}. \tag{11}$$

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

$$\Psi^{(j)} = \left\|\left((\hat{\mathbf{H}}^{(j)})^{\mathrm{H}}\hat{\mathbf{H}}^{(j)} + M\alpha^{(j)}\mathbf{I}_M\right)^{-1}(\hat{\mathbf{H}}^{(j)})^{\mathrm{H}}\right\|_{\mathrm{F}}^2. \tag{12}$$

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

$$\mathbf{T}_{\mathrm{rZF}}^{\mathrm{DCSI}} = \begin{bmatrix} \mathbf{E}_1^{\mathrm{H}}\mathbf{T}_{\mathrm{rZF}}^{(1)} \\ \mathbf{E}_2^{\mathrm{H}}\mathbf{T}_{\mathrm{rZF}}^{(2)} \\ \vdots \\ \mathbf{E}_n^{\mathrm{H}}\mathbf{T}_{\mathrm{rZF}}^{(n)} \end{bmatrix} \tag{13}$$

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

$$\mathbf{E}_j^{\mathrm{H}} = \begin{bmatrix}\mathbf{0}_{M_{\mathrm{TX}}\times(j-1)M_{\mathrm{TX}}} & \mathbf{I}_{M_{\mathrm{TX}}} & \mathbf{0}_{M_{\mathrm{TX}}\times(n-j)M_{\mathrm{TX}}}\end{bmatrix}. \tag{14}$$

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

Remark 1. It is important to note that TX j transmits using the regularization coefficient $\alpha^{(j)}$, which means that all the TXs do not use the same regularization coefficient. □

D. Naive Regularized ZF

When TX j is not aware of the imperfection of the sharing links between the TXs, it chooses its regularization parameter on the basis of its own CSI quality, which means choosing the optimal regularization coefficient corresponding to the centralized CSI configuration relative to the CSI at TX j. In general, this can be done using a linear search and the large system approximations given in [9].

In the case of same quality across all links, i.e., $\sigma_k^{(j)} = \sigma^{(j)}$, this can be done using the following closed formed expression given in [9]

$$\alpha^{(j),\mathrm{CCSI}} = \frac{1+(\sigma^{(j)})^2\rho}{1-(\sigma^{(j)})^2}\frac{1}{\beta\rho}. \tag{15}$$

E. DCSI-robust Regularized ZF

In the DCSI-configuration, each TX is aware of the statistics of the estimates at *all* the TXs. Hence, it takes this knowledge into account when choosing its regularization coefficient. As the performance depends on the regularization coefficient of all the TXs, this means that each TX effectively solves

$$(\alpha^{(1),\star},\dots,\alpha^{(n),\star}) = \underset{(\alpha^{(1)},\dots,\alpha^{(n)})}{\operatorname{argmax}} R. \tag{16}$$

One of our motivation for the derivation of the deterministic equivalent of the sum rate will be to use this deterministic equivalent to obtain numerical values for the optimal regularization parameters, i.e., we will solve

$$(\alpha^{(1),o},\dots,\alpha^{(n),o}) = \underset{(\alpha^{(1)},\dots,\alpha^{(n)})}{\operatorname{argmax}} R^o \tag{17}$$

where R^o is a deterministic equivalent for the ergodic sum rate R^o.

Although the finite SNR rate analysis under the precoding structure (13) and the distributed CSI model in (6) is challenging in the general case because of the dependency of one user performance on all the channel estimates, some useful insights can be obtained in the large antenna regime, as shown below.

III. Deterministic Equivalent of the SINR

The Stieltjes transform has proven very useful many times in the analysis of wireless networks [9], [10] and we will also follow this approach. Hence, our approach will be based on the following fundamental result.

Theorem 1. *[10], [15] Consider the resolvent matrix* $\mathbf{Q} = \left(\frac{\mathbf{H}^{\mathrm{H}}\mathbf{H}}{M} + \alpha\mathbf{I}_M\right)^{-1}$ *with the matrix* $\mathbf{H}$ *defined according to Section II and* $\alpha > 0$*. Then the equation*

$$x = \frac{1}{M}\operatorname{tr}\left(\left(\alpha\mathbf{I}_M + \frac{\mathbf{I}_M}{\beta(1+x)}\right)^{-1}\right) \tag{18}$$

admits a unique fixed point which we will denote by δ *in the following and can be obtained in closed form as*

$$\delta = \frac{\beta - 1 - \alpha\beta + \sqrt{(\alpha\beta-\beta+1)^2+4\alpha\beta^2}}{2\alpha\beta}. \tag{19}$$

Let

$$\mathbf{Q}_o = \left(\alpha\mathbf{I}_M + \frac{\mathbf{I}_M}{\beta(1+\delta)}\right)^{-1} \tag{20}$$

and let the matrix $\mathbf{U}$ *be any matrix with bounded spectral norm. Then,*

$$\frac{1}{M}\operatorname{tr}(\mathbf{U}\mathbf{Q}) - \frac{1}{M}\operatorname{tr}(\mathbf{U}\mathbf{Q}_o) \xrightarrow[K,M\to\infty]{a.s.} 0. \tag{21}$$

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

Theorem 2. *Considering the D-CSI network MIMO channel described in Section II, then* $\mathrm{SINR}_k - \mathrm{SINR}_k^o \to 0$ *with* SINR_k^o *defined as*

$$\mathrm{SINR}_k^o = \frac{P\left(\frac{1}{n}\sum_{j=1}^n\sqrt{\frac{1-(\sigma_k^{(j)})^2}{\Gamma_o^{(j)}}}\frac{\delta^{(j)}}{1+\delta^{(j)}}\right)^2}{1+I_k^o} \tag{22}$$

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

$$I_k^o = P - P\sum_{j=1}^n\sum_{j'=1}^n\frac{\Gamma_{j,j'}^o}{\sqrt{\Gamma_o^{(j)}\Gamma_o^{(j')}}}\left[\frac{2c_{0,k}^{(j)}}{n^2}\frac{\delta^{(j)}}{1+\delta^{(j)}} - \frac{\left((\rho_k^{(j,j')})^2c_{2,k}^{(j)}c_{2,k}^{(j')} + c_{0,k}^{(j)}c_{0,k}^{(j')}\right)\delta^{(j)}\delta^{(j')}}{n^2\left(1+\delta^{(j)}\right)\left(1+\delta^{(j')}\right)}\right] \tag{23}$$

while $\Gamma_o^{(j)} \in \mathbb{R}$, $\Gamma_{j,j'}^o \in \mathbb{R}$, $\delta^{(j)}$ *and* $c_{0,k}^{(j)}, c_{2,k}^{(j)}$ *are respectively defined as*

$$\Gamma_o^{(j)} = \frac{(\delta^{(j)})^2}{\beta\left(1+\delta^{(j)}\right)^2 - (\delta^{(j)})^2} \tag{24}$$

$$\Gamma_{j,j'}^o = \frac{\delta^{(j)}\delta^{(j')}\frac{1}{M}\sum_{\ell=1}^K \eta_\ell^{(j,j')}}{(1+\delta^{(j)})(1+\delta^{(j')}) - \delta^{(j)}\delta^{(j')}\frac{1}{M}\sum_{\ell=1}^K \left(\eta_\ell^{(j,j')}\right)^2} \tag{25}$$

$$\eta_\ell^{(j,j')} = \sqrt{c_{0,\ell}^{(j)}c_{0,\ell}^{(j')}} + \sigma_\ell^{(j')}\sigma_\ell^{(j')}(\rho_\ell^{(j,j')})^2 \tag{26}$$

$$\delta^{(j)} = \frac{\beta - 1 - \alpha^{(j)}\beta + \sqrt{(\alpha^{(j)}\beta - \beta + 1)^2 + 4\alpha^{(j)}\beta^2}}{2\alpha^{(j)}\beta} \tag{27}$$

$$c_{0,k}^{(j)} = 1 - (\sigma_k^{(j)})^2 \tag{28}$$

$$c_{2,k}^{(j)} = \sigma_k^{(j)}\sqrt{1-(\sigma_k^{(j)})^2}. \tag{29}$$

The above deterministic SINR expression is very generic and encompasses many well-known subcases as listed below.

Corollary 1. (Deterministic SINR for centralized CSI) *Let* $n = 1$ *the D-CSI model degenerates to the Centralized CSI model: Upon removing subscript for the TX, the deterministic SINR becomes:*

$$\mathrm{SINR}_k^{o(CCSI)} = \frac{(1-\sigma_k^2)\delta^2}{\Gamma^o\left(1-\sigma_k^2 + (1+\delta)^2\sigma_k^2 + \frac{(1+\delta)^2}{P}\right)}$$

Corollary 2. (Deterministic SINR for distributed identical CSI) *Let* $n > 1,\ \sigma_k^{(j)} = \sigma_k^{(j')} = \sigma_k,\ \rho_k^{(j,j')} = 1, \forall j, j' = 1,\dots,n,\ k = 1,\dots,K$, *which corresponds to the D-CSI setting with Identical channel estimate at each TX. Let* $\alpha^{(j)} = \alpha^{(j')} = \alpha, \forall j, j' = 1,\dots,n$, *which indicates that each TX implements the RZF precoder with the same regularization parameter, therefore* $\delta^{(j)} = \delta^{(j')} = \delta, \forall j, j' = 1,\dots,n$. *The deterministic SINR becomes:*

$$\mathrm{SINR}_k^{o(DICSI)} = \frac{(1-\sigma_k^2)\delta^2}{\Gamma^o\left(1-\sigma_k^2 + (1+\delta)^2\sigma_k^2 + \frac{(1+\delta)^2}{P}\right)}$$

The above expression matches with the results for the C-CSI expression just as expected.

Corollary 3. (Deterministic SINR for distributed CSI with the same regularization parameter) *Let* $n > 1, \alpha^{(j)} = \alpha^{(j')} = \alpha, \forall j, j' = 1,\dots,n$, *which indicates that each TX implements the RZF precoder with the same regularization parameter, therefore* $\delta^{(j)} = \delta^{(j')} = \delta, \forall j, j' = 1,\dots,n$. *The deterministic SINR becomes:*

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

$$I_k^{o(DCSI)} = P - P\sum_{j=1}^n\sum_{j'=1}^n \frac{\delta\Gamma_{j,j'}^o}{n^2(1+\delta)^2\Gamma^o}\cdot\Big[2c_{0,k}^{(j)} + \delta\left(2c_{0,k}^{(j)} - c_{0,k}^{(j)}c_{0,k}^{(j')} - (\rho_k^{(j,j')})^2 c_{2,k}^{(j)}c_{2,k}^{(j')}\right)\Big]$$

If $\rho_k^{(j,j')} = 0, \forall j \neq j',\ j, j' = 1,\dots,n,\ k = 1,\dots,K$, the above expression matches with the results for the D-CSI expression in [12].

IV. Proof of Theorem 2

Due to space limitations, only a sketch of the proof is provided in this paper, the detailed steps of the derivation are in a longer version journal paper [16]. 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 [16]. In particular, Lemma 5 extends [10, Lemma 15] and is an interesting result in itself.

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

A deterministic equivalent for $\Psi^{(j)}$ can be found in [9], or be obtained using Lemma 5 with $\sigma^{(j)} = \sigma^{(j')} = 0$, which gives

$$\Psi^{(j)} \asymp \Gamma_o^{(j)}.$$

This deterministic equivalent $\Psi^{(j)}$ does not depend on $\sigma^{(j)}$.

It should be noted that when the system becomes large, the effective global precoder $\mathbf{T}_{\mathrm{rZF}}^{\mathrm{DCSI}}$ satisfies the total power constraint, since

$$\begin{aligned}\|\mathbf{T}_{\mathrm{rZF}}^{\mathrm{DCSI}}\|_F^2 &= \sum_{i=1}^n \mathrm{tr}\left(\mathbf{E}_i^{\mathrm{H}}\mathbf{T}_{\mathrm{rZF}}^{(i)}(\mathbf{T}_{\mathrm{rZF}}^{(i)})^{\mathrm{H}}\mathbf{E}_i\right)\\ &= \sum_{i=1}^n \frac{P}{\Gamma_o^{(i)}}\mathrm{tr}\left(\frac{1}{M^2}\mathbf{E}_i\mathbf{E}_i^{\mathrm{H}}(\mathbf{C}^{(i)})^{-1}(\hat{\mathbf{H}}^{(i)})^{\mathrm{H}}\hat{\mathbf{H}}^{(i)}(\mathbf{C}^{(i)})^{-1}\right)\\ &\overset{(a)}{\asymp} \sum_{i=1}^n \frac{P}{\Gamma_o^{(i)}}\frac{\Gamma_o^{(i)}}{n}\\ &= P.\end{aligned}$$

(a) follows from Lemma 5 and $\Psi^{(i)} \asymp \Gamma_o^{(i)}$.

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

For the desired signal part at RX k, we can write

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

where we have defined

$$\mathbf{C}_{[k]}^{(j)}=\frac{\hat{\mathbf{H}}_{[k]}^{(j)}(\hat{\mathbf{H}}_{[k]}^{(j)})^{\mathrm{H}}}{M}+\alpha^{(j)}\mathbf{I}_M,\qquad\forall j$$

with

$$(\hat{\mathbf{H}}_{[k]}^{(j)})^{\mathrm{H}}=\left[\hat{\boldsymbol{h}}_1^{(j)}\quad\ldots\quad\hat{\boldsymbol{h}}_{k-1}^{(j)}\quad\hat{\boldsymbol{h}}_{k+1}^{(j)}\quad\ldots\quad\hat{\boldsymbol{h}}_K^{(j)}\right],\qquad\forall j,$$

which is the matrix $(\hat{\mathbf{H}}^{(j)})^{\mathrm{H}}$ with the kth column removed. Equality (a) follows then 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 and the fundamental Theorem 1. Note that $\delta^{(j)}$ can be calculated as illustrated in Theorem 1. It follows then directly that

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

C. Deterministic Equivalent for the Interference Term

According to the definition of $\mathbf{T}^{\mathrm{DCSI}}$ in (13), (11) and replace $\Psi^{(j)}$ by its deterministic equivalent.

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

To obtain a deterministic equivalent for the first summation term, apply Lemma 5. For the second summation term, we use the following relation

$$\begin{aligned}&(\mathbf{C}^{(j)})^{-1}-(\mathbf{C}_{[k]}^{(j)})^{-1}\\&=(\mathbf{C}^{(j)})^{-1}\left(\mathbf{C}_{[k]}^{(j)}-\mathbf{C}^{(j)}\right)(\mathbf{C}_{[k]}^{(j)})^{-1}\\&=-\frac{1}{M}(\mathbf{C}^{(j)})^{-1}\Big(c_{0,k}^{(j)}\boldsymbol{h}_k\boldsymbol{h}_k^{\mathrm{H}}+c_{1,k}^{(j)}\boldsymbol{\delta}_k^{(j)}(\boldsymbol{\delta}_k^{(j)})^{\mathrm{H}}\\&+c_{2,k}^{(j)}\boldsymbol{\delta}_k^{(j)}\boldsymbol{h}_k^{\mathrm{H}}+c_{2,k}^{(j)}\boldsymbol{h}_k(\boldsymbol{\delta}_k^{(j)})^{\mathrm{H}}\Big)(\mathbf{C}_{[k]}^{(j)})^{-1}\end{aligned}\qquad(30)$$

where

$$\begin{aligned}c_{0,k}^{(j)}&=1-(\sigma_k^{(j)})^2\\c_{1,k}^{(j)}&=(\sigma_k^{(j)})^2\\c_{2,k}^{(j)}&=\sigma_k^{(j)}\sqrt{1-(\sigma_k^{(j)})^2}\end{aligned}$$

It is important to note that $c_{0,k}^{(j)}c_{1,k}^{(j)}=(c_{2,k}^{(j)})^2, c_{0,k}^{(j)}+c_{1,k}^{(j)}=1$ as these relations will be used several times through the proof. Inserting (30) in $\mathcal{I}_k$ yields

$$\begin{aligned}\mathcal{I}_k&\asymp\frac{P}{M^2}\sum_{j=1}^{n}\sum_{j'=1}^{n}\frac{1}{\sqrt{\Gamma_o^{(j)}\Gamma_o^{(j')}}}\boldsymbol{h}_k^{\mathrm{H}}\mathbf{E}_j\mathbf{E}_j^{\mathrm{H}}(\mathbf{C}_{[k]}^{(j)})^{-1}(\hat{\mathbf{H}}_{[k]}^{(j)})^{\mathrm{H}}\\&\cdot\hat{\mathbf{H}}_{[k]}^{(j')}(\mathbf{C}^{(j')})^{-1}\mathbf{E}_{j'}\mathbf{E}_{j'}^{\mathrm{H}}\boldsymbol{h}_k\\&-\frac{P}{M^3}\sum_{j=1}^{n}\sum_{j'=1}^{n}\frac{1}{\sqrt{\Gamma_o^{(j)}\Gamma_o^{(j')}}}\boldsymbol{h}_k^{\mathrm{H}}\mathbf{E}_j\mathbf{E}_j^{\mathrm{H}}(\mathbf{C}^{(j)})^{-1}\left[\boldsymbol{h}_k c_{0,k}^{(j)}\boldsymbol{h}_k^{\mathrm{H}}\right]\\&\cdot(\mathbf{C}_{[k]}^{(j)})^{-1}(\hat{\mathbf{H}}_{[k]}^{(j)})^{\mathrm{H}}\hat{\mathbf{H}}_{[k]}^{(j')}(\mathbf{C}^{(j')})^{-1}\mathbf{E}_{j'}\mathbf{E}_{j'}^{\mathrm{H}}\boldsymbol{h}_k\\&-\frac{P}{M^3}\sum_{j=1}^{n}\sum_{j'=1}^{n}\frac{1}{\sqrt{\Gamma_o^{(j)}\Gamma_o^{(j')}}}\boldsymbol{h}_k^{\mathrm{H}}\mathbf{E}_j\mathbf{E}_j^{\mathrm{H}}(\mathbf{C}^{(j)})^{-1}\left[\boldsymbol{\delta}_k^{(j)}c_{1,k}^{(j)}(\boldsymbol{\delta}_k^{(j)})^{\mathrm{H}}\right]\\&\cdot(\mathbf{C}_{[k]}^{(j)})^{-1}(\hat{\mathbf{H}}_{[k]}^{(j)})^{\mathrm{H}}\hat{\mathbf{H}}_{[k]}^{(j')}(\mathbf{C}^{(j')})^{-1}\mathbf{E}_{j'}\mathbf{E}_{j'}^{\mathrm{H}}\boldsymbol{h}_k\\&-\frac{P}{M^3}\sum_{j=1}^{n}\sum_{j'=1}^{n}\frac{1}{\sqrt{\Gamma_o^{(j)}\Gamma_o^{(j')}}}\boldsymbol{h}_k^{\mathrm{H}}\mathbf{E}_j\mathbf{E}_j^{\mathrm{H}}(\mathbf{C}^{(j)})^{-1}\left[\boldsymbol{\delta}_k^{(j)}c_{2,k}^{(j)}\boldsymbol{h}_k^{\mathrm{H}}\right]\\&\cdot(\mathbf{C}_{[k]}^{(j)})^{-1}(\hat{\mathbf{H}}_{[k]}^{(j)})^{\mathrm{H}}\hat{\mathbf{H}}_{[k]}^{(j')}(\mathbf{C}^{(j')})^{-1}\mathbf{E}_{j'}\mathbf{E}_{j'}^{\mathrm{H}}\boldsymbol{h}_k\\&-\frac{P}{M^3}\sum_{j=1}^{n}\sum_{j'=1}^{n}\frac{1}{\sqrt{\Gamma_o^{(j)}\Gamma_o^{(j')}}}\boldsymbol{h}_k^{\mathrm{H}}\mathbf{E}_j\mathbf{E}_j^{\mathrm{H}}(\mathbf{C}^{(j)})^{-1}\left[\boldsymbol{h}_k c_{2,k}^{(j)}(\boldsymbol{\delta}_k^{(j)})^{\mathrm{H}}\right]\\&\cdot(\mathbf{C}_{[k]}^{(j)})^{-1}(\hat{\mathbf{H}}_{[k]}^{(j)})^{\mathrm{H}}\hat{\mathbf{H}}_{[k]}^{(j')}(\mathbf{C}^{(j')})^{-1}\mathbf{E}_{j'}\mathbf{E}_{j'}^{\mathrm{H}}\boldsymbol{h}_k\\&\triangleq A+B+C+D+E.\end{aligned}\qquad(31)$$

We proceed by calculating each of the 5 terms in (31) successively, using in particular Lemma 6:

$$\begin{aligned}A&=\frac{P}{M^2}\sum_{j=1}^{n}\sum_{j'=1}^{n}\frac{1}{\sqrt{\Gamma_o^{(j)}\Gamma_o^{(j')}}}\boldsymbol{h}_k^{\mathrm{H}}\mathbf{E}_j\mathbf{E}_j^{\mathrm{H}}(\mathbf{C}_{[k]}^{(j)})^{-1}(\hat{\mathbf{H}}_{[k]}^{(j)})^{\mathrm{H}}\\&\cdot\hat{\mathbf{H}}_{[k]}^{(j')}(\mathbf{C}^{(j')})^{-1}\mathbf{E}_{j'}\mathbf{E}_{j'}^{\mathrm{H}}\boldsymbol{h}_k\\&\asymp P\sum_{j=1}^{n}\sum_{j'=1}^{n}\frac{1}{\sqrt{\Gamma_o^{(j)}\Gamma_o^{(j')}}}\\&\cdot\Bigg[\frac{\mathrm{tr}\left(\mathbf{E}_{j'}\mathbf{E}_{j'}^{\mathrm{H}}\mathbf{E}_j\mathbf{E}_j^{\mathrm{H}}(\mathbf{C}_{[k]}^{(j)})^{-1}(\hat{\mathbf{H}}_{[k]}^{(j)})^{\mathrm{H}}\hat{\mathbf{H}}_{[k]}^{(j')}(\mathbf{C}_{[k]}^{(j')})^{-1}\right)}{M^2}\\&-c_{0,k}^{(j')}\frac{\mathrm{tr}\left(\mathbf{E}_j\mathbf{E}_j^{\mathrm{H}}(\mathbf{C}_{[k]}^{(j)})^{-1}(\hat{\mathbf{H}}_{[k]}^{(j)})^{\mathrm{H}}\hat{\mathbf{H}}_{[k]}^{(j')}(\mathbf{C}_{[k]}^{(j')})^{-1}\right)}{M^2}\\&\cdot\frac{\mathrm{tr}\left(\mathbf{E}_{j'}\mathbf{E}_{j'}^{\mathrm{H}}(\mathbf{C}_{[k]}^{(j')})^{-1}\right)}{M}\frac{1}{1+\frac{\mathrm{tr}\left((\mathbf{C}_{[k]}^{(j')})^{-1}\right)}{M}}\Bigg]\end{aligned}\qquad(32)$$

Since $\mathbf{E}_{j'}\mathbf{E}_{j'}^{\mathrm{H}}\mathbf{E}_j\mathbf{E}_j^{\mathrm{H}}=\mathbf{E}_j\mathbf{E}_j^{\mathrm{H}}\cdot\mathbb{1}_{j=j'}$, according to Lemma 5, it can be shown that

$$\frac{\mathrm{tr}\left(\mathbf{E}_j\mathbf{E}_j^{\mathrm{H}}(\mathbf{C}_{[k]}^{(j)})^{-1}(\hat{\mathbf{H}}_{[k]}^{(j)})^{\mathrm{H}}\hat{\mathbf{H}}_{[k]}^{(j')}(\mathbf{C}_{[k]}^{(j')})^{-1}\right)}{M^2}\asymp\frac{1}{n}\Gamma_{j,j'}^{o}\qquad(33)$$

Preprint submitted to 53rd Annual Allerton Conference on Communication, Control, and Computing. Received July 6, 2015.

Inserting (33) in (32) and using the fundamental Theorem 1 yields

$$A \asymp P\sum_{j=1}^{n}\sum_{j'=1}^{n}\left(\frac{1}{n}\mathbb{1}_{j=j'} - c_{0,k}^{(j')}\frac{\Gamma^o_{j,j'}}{n\sqrt{\Gamma_o^{(j)}\Gamma_o^{(j')}}}\frac{\delta^{(j')}}{n}\frac{1}{1+\delta^{(j')}}\right)$$
$$= P\sum_{j=1}^{n}\left[\frac{1}{n} - \frac{1}{n^2}\frac{c_{0,k}^{(j)}\delta^{(j)}}{1+\delta^{(j)}}\right]$$
$$- P\sum_{j=1}^{n}\sum_{j'=1,j'\neq j}^{n}\frac{\Gamma^o_{j,j'}}{\sqrt{\Gamma_o^{(j)}\Gamma_o^{(j')}}}\frac{c_{0,k}^{(j')}\delta^{(j')}}{n^2\left(1+\delta^{(j')}\right)} \quad (34)$$

Similar to the derivation of the term A in (31), applying Lemma 6 and Lemma 5 to derive the B,C,D,E term and adding all the terms together result in the following expression for the interference term.

$$I_k \asymp P - P\sum_{j=1}^{n}\sum_{j'=1}^{n}\frac{\Gamma^o_{j,j'}}{\sqrt{\Gamma_o^{(j)}\Gamma_o^{(j')}}}\left[\frac{2c_{0,k}^{(j)}}{n^2}\frac{\delta^{(j)}}{1+\delta^{(j)}} - \frac{\left((\rho_k^{(j,j')})^2 c_{2,k}^{(j)}c_{2,k}^{(j')} + c_{0,k}^{(j)}c_{0,k}^{(j')}\right)\delta^{(j)}\delta^{(j')}}{n^2\left(1+\delta^{(j)}\right)\left(1+\delta^{(j')}\right)}\right]$$

V. Simulation Results

A. Numerical Verification of Theorem 2

We now verify using Monte-Carlo simulations the accuracy of the asymptotic expression derived in Theorem 2. We consider a network MIMO system consisting of $n = 5$ TXs with a sum power constraint given by $P = 10$ dB and we assume that $\sigma_k^{(j)}, \forall j = 1,\ldots,n,\ k = 1,\ldots,K$ is uniformly distributed between $(0,1)$. Assume $\beta = 1$, $\rho_k^{(j,j')}, j \neq j', \forall j,j' = 1,\ldots,n$, $\forall k = 1,\ldots,K$ is also uniformly distributed between $(0,1)$ and the $n\times n$ error correlation matrix for which the (j,j') entry is $\rho_k^{(j,j')}$ is symmetric positive semi-definite. The regularization parameter $\alpha^{(j)}, \forall j = 1,\ldots,n$ is chosen uniformly distributed between $(0,1)$.

In Fig. 1, we plot the log scaled gap of average rate per user between the simulation and the deterministic equivalent as a function of the number of users for a square setting where $M = nM_{\text{TX}} = K$ (i.e., $\beta = 1$). It reveals that Monte-Carlo simulation converges to the deterministic equivalent as the system becomes large.

B. Effect of the D-CSI

We now discuss the cost of the distributiveness of the CSI. We consider a network consisting of $n = 3$ TXs and $K = 30$ RXs with a sum power constraint given by $P = 10$ dB and we assume that at TX1, $\sigma_k^{(1)} = \sigma^{(1)}, \forall k = 1,\ldots,K$ is uniformly distributed between $(0,0.1)$, at TX 2, $\sigma_k^{(2)} = \sigma^{(2)}, \forall k = 1,\ldots,K$ is uniformly distributed between $(0.3,0.4)$, and at TX 3, $\sigma_k^{(3)} = \sigma^{(3)}, \forall k = 1,\ldots,K$ is uniformly distributed between $(0.6,0.7)$. This corresponds to an asymmetric CSI setting that TX1 has relatively good CSI, TX2 has moderate CSI and TX3 has relatively bad CSI. We choose $\beta = 1$, $\rho_k^{(j,j')} = \rho, j \neq j', \forall j,j' = 1,\ldots,n, \forall k = 1,\ldots,K$ and we let ρ vary from 0 to 1. This indicates that the CSI model will vary from fully distributed CSI to centralized CSI. We compare the performance obtained using D-CSI regularized ZF with the optimal regularization coefficients (with legend DCSI robust RZF) obtained numerically in Section II-E and the naive regularized ZF in Section II-D (with legend naive RZF) where the CSI inconsistencies between the TXs are not considered. The case when all TXs use a common regularize parameter which is optimized numerically (with legend DCSI RZF with equal regularize parameter) is also considered in simulation.

Fig. 2 reveals that the average rate per user increases as the correlation among the estimation error at each TX increases, i.e., the more centralized the better user rate we can achieve. It also shows that the proposed D-CSI robust Regularized ZF outperforms the naive Regularized ZF algorithm.

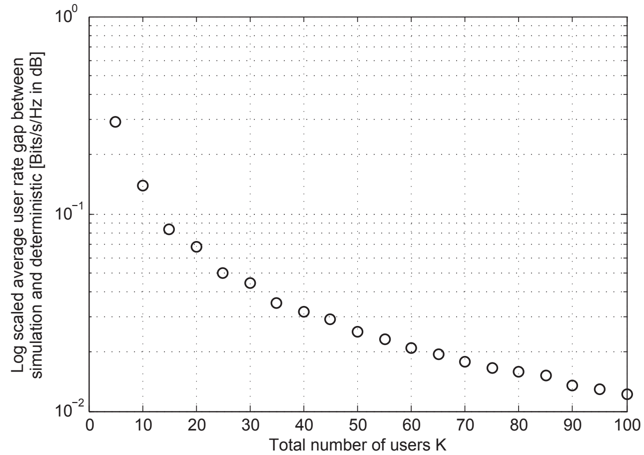

Fig. 1: Log scaled gap of average rate per user between simulation and deterministic as a function of the number of users K.

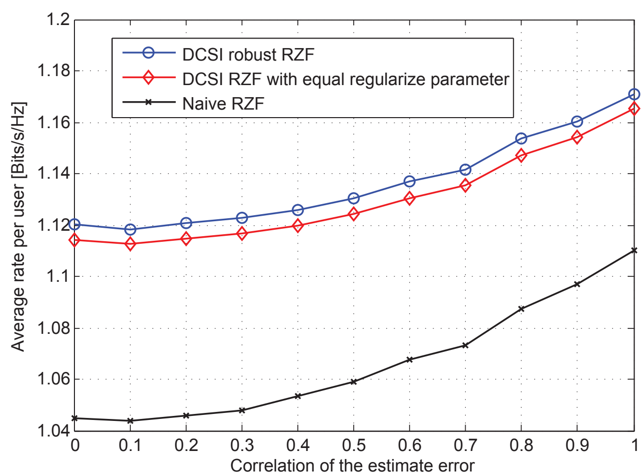

Fig. 2: Average rate per user as a function of estimation error correlation ρ.

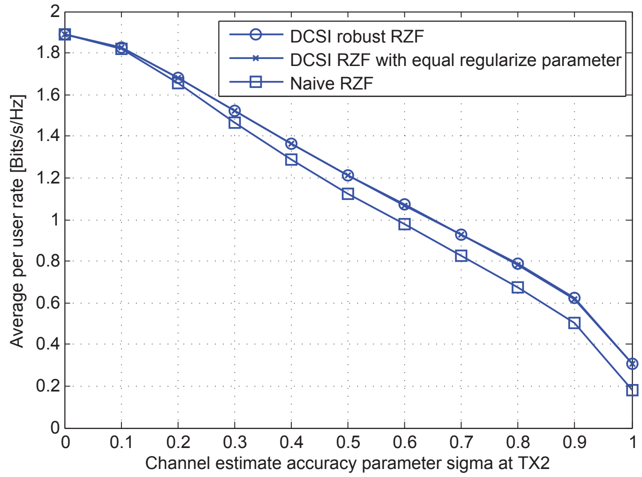

Fig. 3: Average rate per user as a function of CSI estimate accuracy parameter σ at TX2, with TX1 having perfect CSI.

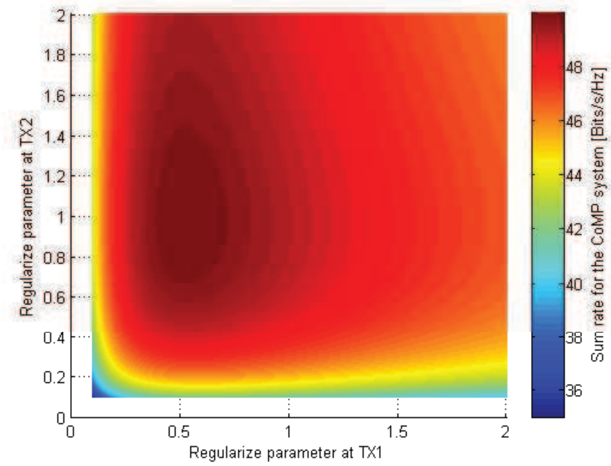

Fig. 4: Sum rate as a function of regularization parameter at TX1 and TX2, TX1 has good CSI estimate and TX2 has bad CSI estimate.

Fig. 3 exhibits the average user rate as the function of the accuracy parameter σ. In this simulation, we consider $n = 2$ TXs and $K = 30$ RXs with a sum power constraint given by $P = 10$ dB. Assume $\beta = 1$, $\rho_k^{(j,j')}, j \neq j', \forall j, j' = 1, 2,\ \forall k = 1, \ldots, K$ is also uniformly distributed between $(0, 1)$ and the $n \times n$ error correlation matrix for which the (j, j') entry is $\rho_k^{(j,j')}$ is symmetric positive semi-definite. At TX 1, $\sigma_k^{(1)} = 0, \forall k = 1, \ldots, K$, which indicates that CSI is perfect at TX 1. At TX 2, $\sigma_k^{(2)} = \sigma, \forall k = 1, \ldots, K$ with σ varies from 0 to 1, which indicates that CSI at TX 2 varies from perfect CSI to no CSI case.

We can conclude from Fig. 3 that when the CSI is more accurate, the better average user rate we can achieve. When both TXs has symmetrically good CSI, the optimization for the regularized parameter will not enhance much of the system performance. While the CSI at the two TXs becomes more and more asymmetric, the gap between the proposed D-CSI robust Regularized ZF and the naive Regularized ZF becomes non-negligible.

C. Optimal Regularization Coefficient

We now analysis how the choice of the regularization parameter will interact with the user rate. We consider a network consisting of $n = 2$ TXs and $K = 50$ RXs with a sum power constraint given by $P = 10$ dB. Assume $\beta = 1$, $\rho_k^{(j,j')}, j \neq j', \forall j, j' = 1, 2,\ \forall k = 1, \ldots, K$ is also uniformly distributed between $(0, 1)$ and the $n \times n$ error correlation matrix for which the (j, j') entry is $\rho_k^{(j,j')}$ is symmetric positive semi-definite. If $(\sigma_k^{(j)})^2, \forall j = 1, 2,\ k = 1, \ldots, K$ is uniformly distributed between $(0, 0.1)$, we call it good CSI. If $(\sigma_k^{(j)})^2, \forall j = 1, 2,\ k = 1, \ldots, K$ is uniformly distributed between $(0.6, 0.7)$, we call it bad CSI.

Figure. 4 exploits the case when TX1 has rather good CSI estimate for all the user channels, while TX2 has rather bad CSI estimates for all user channels. The heat map indicates the system sum rate when different regularization parameters are chosen at TX1 and TX2. Dark red represents for higher sum rate and dark blue represents lower sum rate. Fig. 4 shows that it's not always optimal to let all TXs choose the same regularize parameter. In this asymmetric CSI setting, TX1 with good CSI intends to choose a small regularize parameter and the precoder resembles more a ZF precoder while TX2 with bad CSI intends to choose a large regularize parameter and the precoder is more like a matched filter.

VI. Conclusion

We have studied regularized ZF joint precoding in a distributed CSI configuration. We extend the conventional distributed CSI scenario by allowing the CSI errors at the different TXs to be correlated. This novelty offers new perspectives for modeling the CSI in partially centralized setting. In addition, we extend the analysis of regularized ZF by allowing each TX to choose its own regularization coefficient. Using RMT tools, an analytical expression is derived to approximate the average rate per user in the large system limit. This analytical expression is then used to optimize the regularization coefficients at the different TXs in order to reduce the impact of the distributed CSI configuration. Interestingly, it is shown that it is *not* always optimal to use the same regularization coefficient at all the TXs. The CSI model is extremely general and can be adapted to many wireless settings.

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 $\mathbf{Q} = \left(\mathbf{H}^{\mathrm{H}}\mathbf{H} + \alpha \mathbf{I}_M\right)^{-1}$ and $\mathbf{Q}_k = \left(\mathbf{H}_k^{\mathrm{H}}\mathbf{H}_k + \alpha \mathbf{I}_M\right)^{-1}$*

Preprint submitted to 53rd Annual Allerton Conference on Communication, Control, and Computing. Received July 6, 2015.

with $\alpha > 0$ respectively. It then holds that

$$\mathbf{Q} = \mathbf{Q}_k - \frac{1}{M}\frac{\mathbf{Q}_k \boldsymbol{h}_k \boldsymbol{h}_k^{\mathrm{H}} \mathbf{Q}_k}{1+\frac{1}{M}\boldsymbol{h}_k^{\mathrm{H}}\mathbf{Q}_k \boldsymbol{h}_k} \tag{35}$$

and

$$\boldsymbol{h}_k^{\mathrm{H}}\mathbf{Q} = \frac{\boldsymbol{h}_k^{\mathrm{H}}\mathbf{Q}_k}{1+\frac{1}{M}\boldsymbol{h}_k^{\mathrm{H}}\mathbf{Q}_k \boldsymbol{h}_k}. \tag{36}$$

Lemma 2 ([10], [11]). *Let $(\mathbf{A}_N)_{N\geq 1}, \mathbf{A}_N \in \mathbb{C}^{N\times N}$ be a sequence of matrices such that $\limsup \|\mathbf{A}_N\| < \infty$, and $(\mathbf{x}_N)_{N\geq 1}, \mathbf{x}_N \in \mathbb{C}^{N\times 1}$ 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,*

$$\frac{1}{N}\mathbf{x}_N^{\mathrm{H}}\mathbf{A}_N\mathbf{x}_N - \frac{1}{N}\operatorname{tr}(\mathbf{A}_N) \xrightarrow[N\to\infty]{a.s.} 0. \tag{37}$$

Lemma 3 ([10], [11]). *Let $(\mathbf{A}_N)_{N\geq 1}, \mathbf{A}_N \in \mathbb{C}^{N\times N}$ 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,*

$$\frac{1}{N}\mathbf{x}_N^{\mathrm{H}}\mathbf{A}_N\mathbf{y}_N \xrightarrow[N\to\infty]{a.s.} 0. \tag{38}$$

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*

$$\operatorname{tr}(\mathbf{A}(\mathbf{Q}-\mathbf{Q}_k)) \leq \|\mathbf{A}\|_2. \tag{39}$$

B. New Lemmas

Lemma 5. *Let $\boldsymbol{h}'_k = \sqrt{1-(\sigma'_k)^2}\boldsymbol{h}_k + \sigma'_k\boldsymbol{\delta}'_k$ and $\boldsymbol{h}''_k = \sqrt{1-(\sigma''_k)^2}\boldsymbol{h}_k + \sigma''_k\boldsymbol{\delta}''_k$. $\sigma'_k, \sigma''_k \in [0,1]$ with $\boldsymbol{h}_k$ independent with $\boldsymbol{\delta}'_k, \boldsymbol{\delta}''_k$. $\boldsymbol{\delta}'_k, \boldsymbol{\delta}''_k$ have their elements i.i.d. $\mathcal{N}_{\mathbb{C}}(0,1)$, $\mathbb{E}\left[\boldsymbol{\delta}'_k(\boldsymbol{\delta}''_k)^H\right] = \rho_k^2\mathbf{I}_M$. $\boldsymbol{h}_k^H, {\boldsymbol{h}'_k}^H, {\boldsymbol{h}''_k}^H, \forall k = 1\ldots K$ are the row vectors for $\mathbf{H}, \mathbf{H}', \mathbf{H}''$ respectively. Let $\mathbf{Q}' = \left(\frac{\mathbf{H}'^{\mathrm{H}}\mathbf{H}'}{M} + \alpha'\mathbf{I}_M\right)^{-1}$ and $\mathbf{Q}'' = \left(\frac{\mathbf{H}''^{\mathrm{H}}\mathbf{H}''}{M} + \alpha''\mathbf{I}_M\right)^{-1}$ with $\alpha', \alpha'' > 0$. Let $\mathbf{A} \in \mathbb{C}^{M\times M}$ be of uniformly bounded spectral norm with respect to M. Then,*

$$\frac{1}{M^2}\operatorname{tr}\left(\mathbf{A}\mathbf{Q}'\mathbf{H}'^{\mathrm{H}}\mathbf{H}''\mathbf{Q}''\right) - \frac{\operatorname{tr}(\mathbf{A})}{M}Y_0 \xrightarrow{a.s.} 0 \tag{40}$$

where Y_0 is defined as

$$Y_o = \frac{\delta'\delta''\frac{1}{M}\sum_{\ell=1}^K \eta_\ell}{(1+\delta')(1+\delta'') - \delta'\delta''\frac{1}{M}\sum_{\ell=1}^K \eta_\ell^2}$$
$$\eta_\ell = \sqrt{(1-(\sigma'_\ell)^2)(1-(\sigma''_\ell)^2)} + \sigma'_\ell\sigma''_\ell\rho_\ell^2.$$

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 $\mathbf{x}, \mathbf{x}', \mathbf{y}$ have i.i.d. complex entries of zero mean, variance $1/M$ and finite 8th order moment. $\mathbf{x}, \mathbf{y}$ and $\mathbf{x}', \mathbf{y}$ are mutually independent as well as independent of $\mathbf{L}, \mathbf{R}, \bar{\mathbf{A}}$. $\mathbf{x}, \mathbf{x}'$ satisfies $\mathbb{E}\left[\mathbf{x}'\mathbf{x}^H\right] = \rho^2\mathbf{I}_M$. Then we have:*

$$\mathbf{x}^{\mathrm{H}}\mathbf{L}\mathbf{A}^{-1}\mathbf{R}\mathbf{x} \asymp u_{\mathrm{LR}} - \frac{c_0 u_{\mathrm{L}} u_{\mathrm{R}}}{1+u}$$
$$\mathbf{x}^{\mathrm{H}}\mathbf{L}\mathbf{A}^{-1}\mathbf{R}\mathbf{y} \asymp -\frac{c_2 u_{\mathrm{L}} u_{\mathrm{R}}}{1+u}$$
$$\mathbf{x}^{\mathrm{H}}\mathbf{L}\mathbf{A}'^{-1}\mathbf{R}\mathbf{y} \asymp -\rho^2\frac{c_2 u_{\mathrm{L}} u_{\mathrm{R}}}{1+u}$$

with

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

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

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

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] P. de Kerret, D. Gesbert, and U. Salim, "Regularized Zero Forcing in Cooperative Broadcast Channels under Distributed CSIT: A Large system analysis," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, 2015.

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

[14] 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.

[15] 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.

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