Network MIMO: Transmitters with no CSI Can Still be Very Useful

Paul de Kerret and David Gesbert
Communication Systems Department, EURECOM, Sophia-Antipolis, France

***Abstract*—In this paper[1] we consider the Network MIMO channel under the so-called Distributed Channel State Information at the Transmitters (D-CSIT) configuration. In this setting, the precoder is designed in a distributed manner at each Transmitter (TX) on the basis of local versions of Channel State information (CSI) of various quality. Although the use of simple Zero-Forcing (ZF) was recently shown to reach the optimal DoF for a Broadcast Channel (BC) under noisy, yet centralized, CSI at the TX (CSIT), it can turn very inefficient when faced with D-CSIT: The number of Degrees-of-Freedom (DoF) achieved is then limited by the worst CSI accuracy across TXs. To circumvent this effect, we develop a new robust transmission scheme improving the DoF. A surprising result is uncovered by which, in the regime of so-called weak CSIT, the proposed scheme is shown to be DoF-optimal and to achieve a centralized outerbound consisting in the DoF of a genie-aided centralized setting in which the CSIT versions of all TXs are available everywhere. Building upon the insight obtained in the weak CSIT regime, we develop a general D-CSIT robust scheme for the 3-user case which improves over the DoF obtained by conventional robust approaches for any arbitrary CSIT configuration.**

I. Introduction

Multiple-antennas at the TX can be exploited to serve multiple users at the same time, thus offering a strong DoF improvement over time-division schemes [1]. This DoF improvement is however critically dependent on the accuracy of the CSIT. Indeed, the absence of CSIT is known to lead to the complete loss of the DoF improvement in the case of an isotropic BC [2]. Going further, a long standing conjecture by Lapidoth, Shamai, and Wigger [3] has been recently settled in [4] by showing that a scaling of the CSIT error in $P^{-\alpha}$ for $\alpha \in [0,1]$ leads to a DoF of $1+(K-1)\alpha$ in the K antennas BC.

In the above literature, however, *centralized* CSIT is typically assumed, i.e., precoding is done on the basis of a *single* imperfect/outdated multiuser channel estimate being common at *every* transmit antenna. Although meaningful in the case of a BC with a single TX, this assumption can be challenged when the joint precoding is carried out across distant TXs linked by heterogeneous and imperfect backhaul links, as in the Network MIMO context. In this case, it is expected that the CSI exchange introduces further delay and quantization noise such that it becomes necessary to study the impact of *TX dependent* CSI noise.

[1]D. Gesbert and P. de Kerret are supported by the European Research Council under the European Union's Horizon 2020 research and innovation program (Agreement no. 670896).

In order to account for TX dependent feedback limitations, a *distributed CSIT* model is introduced in [5]. In this model, TX j receives its own multi-user imperfect estimate $\hat{\mathbf{H}}^{(j)}$ on the basis of which it designs its transmit coefficients, without additional communications with the other TXs. The finite-SNR performance of regularized ZF under D-CSIT has been computed in the large system limit in [6] while heuristic robust precoding schemes have been provided in [7], [8] for practical cellular networks. In [9], [10], Interference Alignment with distributed CSIT is studied and methods to reduce the required CSIT gaining at each TX are provided.

In terms of DoF, it was shown in [5] that using a conventional ZF precoder (regularized or not) in the Network MIMO setting with distributed CSIT leads to a severe DoF degradation caused by the lack of a consistent CSI shared by the cooperating TXs. A two-user specific scheme called "Active-Passive ZF (AP-ZF)" was proposed to lift the DoF.

In this work, we study the DoF for a general K-user Network MIMO channel with D-CSIT. More precisely, our main results read as follows.

- In a certain weak CSIT regime, which will be defined rigorously below, we provide a transmission scheme achieving the optimal DoF. Surprisingly, the DoF obtained is the same as in a genie-aided setting in which all TXs share the knowledge of the most accurate CSI estimate.
- In the arbitrary CSIT regime, the above D-CSIT robust scheme is extended, helping lift the DoF substantially above what is achieved by conventional ZF precoding.

Notations: We will use $\doteq$ to denote exponential equality, i.e., we write $f(P) \doteq P^x$ to denote $\lim_{P\to\infty} \frac{\log f(P)}{\log P} = x$. The exponential inequalities $\dot{\leq}$ and $\dot{\geq}$ are defined in the same way.

II. System Model

A. Transmission Model

We study a communication system where K TXs jointly serve K Receivers (RXs) over a Network (Broadcast) MIMO channel. We consider that each TX is equipped with a single-antenna. Each RX is also equipped with a single antenna and we further assume that the RXs have perfect CSI so as to focus on the impact of the imperfect CSI on the TX side.

The signal received at RX i is written as

$$y_i = \boldsymbol{h}_i^{\mathrm{H}} \mathbf{x} + z_i \tag{1}$$

where $\boldsymbol{h}_i^{\mathrm{H}} \in \mathbb{C}^{1\times K}$ is the channel to user i, $\mathbf{x} \in \mathbb{C}^K$ is the transmitted multi-user signal, and $z_i \in \mathbb{C}$ is the additive noise at RX i, being independent of the channel and the transmitted signal, and distributed as $\mathcal{N}_{\mathbb{C}}(0,1)$. We further define the channel matrix $\mathbf{H} \triangleq [\boldsymbol{h}_1, \dots, \boldsymbol{h}_K]^{\mathrm{H}} \in \mathbb{C}^{K\times K}$. The channel is assumed to be drawn from a continuous ergodic distribution such that all the channel matrices and all their sub-matrices are full rank with probability one.

B. Distributed CSIT Model

The D-CSIT setting differs from the conventional centralized one in that each TX receives a possibly different (global) CSI on the basis of which it designs its own transmission parameters without any additional communication to the other TXs. Specifically, TX j receives the imperfect multi-user channel estimate $\hat{\mathbf{H}}^{(j)} = [\hat{\boldsymbol{h}}_1^{(j)}, \dots, \hat{\boldsymbol{h}}_K^{(j)}]^{\mathrm{H}} \in \mathbb{C}^{K\times K}$ where $(\hat{\boldsymbol{h}}_i^{(j)})^{\mathrm{H}}$ refers to the estimate of the channel from all TXs to user i, at TX j. TX j then designs its transmit coefficients solely as a function of $\hat{\mathbf{H}}^{(j)}$ and the statistics of the channel.

We model the CSI uncertainty at the TXs by

$$\hat{\mathbf{H}}^{(j)} = \mathbf{H} + \sqrt{P^{-\alpha^{(j)}}}\boldsymbol{\Delta}^{(j)} \tag{2}$$

where $\boldsymbol{\Delta}^{(j)}$ is a random variable with zero mean and bounded covariance matrix verifying (2). The scalar $\alpha^{(j)}$ is called the *CSIT scaling coefficient* at TX j. It takes its value in $[0,1]$ where $\alpha^{(j)} = 0$ corresponds to a CSIT being essentially useless in terms of DoF while $\alpha^{(j)} = 1$ corresponds to a CSIT being essentially perfect in terms of DoF [4], [11].

The distributed CSIT quality is represented through the multi-user CSIT scaling vector $\boldsymbol{\alpha} \in \mathbb{R}^K$ defined as

$$\boldsymbol{\alpha} \triangleq \begin{bmatrix} \alpha^{(1)} \\ \vdots \\ \alpha^{(K)} \end{bmatrix}. \tag{3}$$

For ease of exposition, we consider the simple configuration where the channel realizations and the channel estimates are drawn in an i.i.d manner. Furthermore, we consider that for a given transmission power P, the conditional probability density functions verify that

$$\max_{\mathbf{H}\in\mathbb{C}^{K\times K}} \left(p_{\mathbf{H}|\hat{\mathbf{H}}^{(1)},\dots,\hat{\mathbf{H}}^{(K)}}(\mathbf{H})\right) \doteq \sqrt{P^{\alpha^{\max}}}. \tag{4}$$

Remark 1. This condition extends the condition provided in [4] which writes in our setting as

$$\max_{\mathbf{H}\in\mathbb{C}^{K\times K}} \left(p_{\mathbf{H}|\hat{\mathbf{H}}^{(j)}}(\mathbf{H})\right) \doteq \sqrt{P^{\alpha^{(j)}}}, \quad \forall j \in \{1, \dots, K\}. \tag{5}$$

□

Condition (4) is a mild technical assumption, which holds for the distributions usually considered.

Example 1. *We show in the extended version [12] that condition* (4) *is satisfied in the usual case where the noise realizations $\boldsymbol{\Delta}^{(j)} \in \mathbb{C}^{K\times K}$ are i.i.d. according to $\mathcal{N}_{\mathbb{C}}(\mathbf{0}_K, \mathbf{I}_K)$ and all the CSI noise error terms $\boldsymbol{\Delta}^{(j)}$ are independent of each other.* ■

C. Degrees-of-Freedom Analysis

Let us denote by $\mathcal{C}(P)$ the sum capacity [13] of the MISO BC with distributed CSIT considered. The optimal sum DoF is then denoted by

$$\mathrm{DoF}^{\mathrm{DCSI}}(\boldsymbol{\alpha}) \triangleq \lim_{P\to\infty} \frac{\mathcal{C}(P)}{\log_2(P)}. \tag{6}$$

III. Main Results

As one of the key observations made in this paper, we found that the DoF behavior in a Network MIMO channel with D-CSIT quite depends on the CSI quality regime. To this end, notions of "weak CSIT" regime and "arbitrary CSIT" regimes are introduced to characterize the interval in which the CSIT scaling coefficients $\{\alpha^{(j)}\}_{j=1}^K$ are allowed to take their values.

The weak CSIT regime is defined as follows:

Definition 1. *In the K-user Network MIMO with distributed CSIT and $K \geq 2$, we define the weak CSIT regime as comprising all the CSIT configurations satisfying that*

$$\max_{j\in\{1,\dots,K\}} \alpha^{(j)} \leq \frac{1}{1 + K(K-2)}. \tag{7}$$

Hence in the two-user case, the condition reduces to $\alpha^{(j)} \leq 1, \forall j \in \{1,2\}$, such that the notion of weak CSIT coincides with the arbitrary regime in that case, which will be shown to be coherent with the DoF result obtained in [5]. In the three-user case, CSIT is said to be weak if $\alpha^{(j)} \leq 1/4, \forall j \in \{1,2,3\}$, and so forth.

A. Weak CSIT Regime

Theorem 1. *In the K-user Network MIMO with distributed CSIT, the optimal sum DoF in the weak CSIT regime is*

$$\mathrm{DoF}^{\mathrm{DCSI}}(\boldsymbol{\alpha}) = 1 + (K-1) \max_{j\in\{1,\dots,K\}} \alpha^{(j)}. \tag{8}$$

Proof. The achievability follows from the description of the scheme in Section V while the outerbound follows from the intuitive fact that CSIT discrepancies between the TXs cannot improve the DoF. The rigorous proof is given in the extended version [12]. □

Quite surprisingly, it can be seen that the DoF expression only depends on the best CSIT accuracy among all TXs. This indicates that even TXs with no CSI are useful in lifting the DoF in a decentralized setting. This result is in sharp contrast with the performance obtained using a conventional ZF approach, for which the DoF is shown in [5] to be limited by the worst accuracy across the TXs (more precisely to be equal to $1 + (K-1)\min_j \alpha^{(j)}$ when considering successive decoding). Note that this is despite the fact that the ZF approach was recently shown to be DoF optimal for the BC under centralized CSIT setting [4].

B. Arbitrary CSIT Regime with $K=3$

Developing a transmission scheme optimally exploiting the available CSIT *for any CSIT configuration* is very challenging due to the large number of parameters. Consequently, we focus first on the 3-user case while finding the optimal DoF for any number of users and any CSIT configuration is an ongoing research.

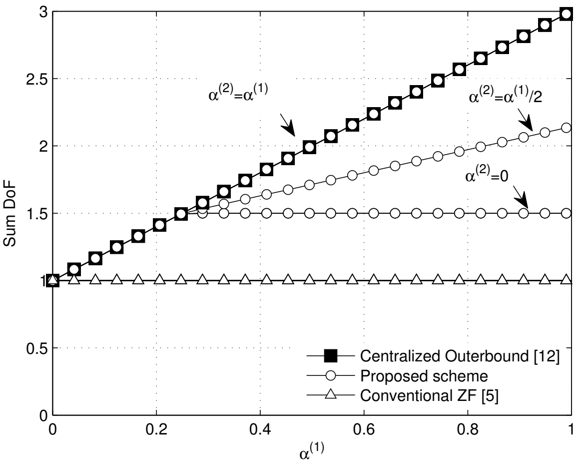

Fig. 1: Sum DoF as a function of $\alpha^{(1)}$. The DoF achieved is presented for exemplary values of $\alpha^{(2)}$, while $\alpha^{(3)}=0$.

Theorem 2. *In the* 3*-user Network MIMO with distributed CSIT and* $\alpha^{(1)} \geq \alpha^{(2)} \geq \alpha^{(3)}$*, it holds that*

$$\mathrm{DoF}^{\mathrm{DCSI}}(\boldsymbol{\alpha}) \geq \begin{cases} 1+2\alpha^{(1)} & \text{if } \alpha^{(1)} \leq \frac{1}{4} \\ 3\frac{2\alpha^{(1)}-\alpha^{(2)}+2\alpha^{(1)}\alpha^{(2)}}{4\alpha^{(1)}-\alpha^{(2)}} & \text{if } \alpha^{(1)} \geq \frac{1}{4}. \end{cases} \quad (9)$$

Proof. The detailed description of the scheme can be found in the extended version [12]. □

The DoF achieved with the proposed scheme is illustrated in Fig. 1. For $\alpha^{(1)} \leq \frac{1}{4}$, the transmission occurs in the weak CSIT regime, and the proposed scheme coincides with the transmission scheme for the weak CSIT, as described in Section V. Note that the achieved DoF does not depend on $\alpha^{(3)}$. This is coherent with the results in [5] and follows from the use of Active-Passive (AP-) ZF.

The proposed D-CSI robust transmission schemes rely on several ingredients which are (i) AP-ZF precoding, (ii) interference quantization, and (iii) superposition coding. As a necessary preliminary, we start by extending the AP-ZF precoding scheme introduced in [5] to more than one passive TX.

IV. Preliminary: Active-Passive Zero Forcing

Let us consider a setting in which K single-antenna TXs aim to send $K-n$ symbols to serve one user (e.g., a user having $K-n$ antennas) while zero-forcing interference to n other single-antenna users, where $n<K$. The channel from the K TXs to the interfered users is denoted by $\mathbf{H} \in \mathbb{C}^{n\times K}$. We divide the TXs between so-called *active TXs* and *passive TXs* and we consider without loss of generality that the first n TXs are the active TXs while the remaining $K-n$ TXs are the passive ones.

We define the *active channel* as the channel coefficients from the active TXs, denoted by $\mathbf{H}_{\mathrm{A}} \in \mathbb{C}^{n\times n}$, and the *passive channel* as the channel coefficients from the passive TXs, denoted by $\mathbf{H}_{\mathrm{P}} \in \mathbb{C}^{n\times (K-n)}$, such that

$$\mathbf{H} = \begin{bmatrix}\mathbf{H}_{\mathrm{A}} & \mathbf{H}_{\mathrm{P}}\end{bmatrix}. \quad (10)$$

Turning to the CSIT configuration, we assume that an estimate $\hat{\mathbf{H}}^{(j)} \in \mathbb{C}^{n\times K}$ is available at TX j, for $j\in\{1,\dots,K\}$. We define the estimated active channel $\hat{\mathbf{H}}_{\mathrm{A}}^{(j)} \in \mathbb{C}^{n\times n}$ and the estimated passive channel $\hat{\mathbf{H}}_{\mathrm{P}}^{(j)} \in \mathbb{C}^{n\times (K-n)}$ similarly such that

$$\hat{\mathbf{H}}^{(j)} = \begin{bmatrix}\hat{\mathbf{H}}_{\mathrm{A}}^{(j)} & \hat{\mathbf{H}}_{\mathrm{P}}^{(j)}\end{bmatrix}. \quad (11)$$

Turning to the computation of the precoder at TX j, the AP-ZF precoder computed at TX j is then denoted by $\mathbf{T}^{\mathrm{APZF}(j)} \in \mathbb{C}^{K\times K-n}$. The part of the precoder which should be implemented at the active TXs is denoted by $\lambda^{\mathrm{APZF}}\mathbf{T}^{\mathrm{A}(j)} \in \mathbb{C}^{n\times K-n}$, where λ^{APZF} is used to satisfy an average sum power constraint (see its exact value in (15)), and is called the *active precoder* (computed at TX j) while the part of the precoder which should be implemented at the passive TXs is denoted by $\lambda^{\mathrm{APZF}}\mathbf{T}^{\mathrm{P}} \in \mathbb{C}^{(K-n)\times(K-n)}$ and is called the *passive precoder*. It then holds that

$$\mathbf{T}^{\mathrm{APZF}(j)} = \lambda^{\mathrm{APZF}} \begin{bmatrix}\mathbf{T}^{\mathrm{A}(j)} \\ \mathbf{T}^{\mathrm{P}}\end{bmatrix}. \quad (12)$$

The precoder $\mathbf{T}^{\mathrm{P}}$ is arbitrarily chosen as any full rank matrix known to all TXs while the precoder $\mathbf{T}^{\mathrm{A}(j)}$ is computed as

$$\mathbf{T}^{\mathrm{A}(j)} = -\left((\hat{\mathbf{H}}_{\mathrm{A}}^{(j)})^{\mathrm{H}}\hat{\mathbf{H}}_{\mathrm{A}}^{(j)} + \frac{1}{P}\mathbf{I}_n\right)^{-1}(\hat{\mathbf{H}}_{\mathrm{A}}^{(j)})^{\mathrm{H}}\hat{\mathbf{H}}_{\mathrm{P}}^{(j)}\mathbf{T}^{\mathrm{P}} \quad (13)$$

where P is the total available average transmit power. Note that the precoder $\mathbf{T}^{\mathrm{P}}$ is a CSI independent precoder which is commonly agreed upon by all TXs beforehand.

Remark 2. The interpretation of (13) is as follows: The passive TXs transmit with constant coefficients and the active TXs adapt their transmission strategy to eliminate the interferences generated. To gain more intuition, we refer the reader to the easier case with a single passive TX in [5] □

The effective AP-ZF precoder is implemented in a distributed manner and is denoted by $\mathbf{T}^{\mathrm{APZF}} \in \mathbb{C}^{K\times K-n}$. It is a composite version of the precoders computed at each TX and is hence given by

$$\mathbf{T}^{\mathrm{APZF}} \triangleq \lambda^{\mathrm{APZF}} \begin{bmatrix}\boldsymbol{e}_1^{\mathrm{H}}\mathbf{T}^{\mathrm{A}(1)} \\ \vdots \\ \boldsymbol{e}_n^{\mathrm{H}}\mathbf{T}^{\mathrm{A}(n)} \\ \mathbf{T}_{\mathrm{P}}\end{bmatrix}. \quad (14)$$

where $\boldsymbol{e}_i \in \mathbb{C}^n$ for $i \in \{1,\ldots,n\}$ is the ith row of the identity $\mathbf{I}_n$ and where the normalization coefficient λ^{APZF} is chosen as

$$\lambda^{\mathrm{APZF}} \triangleq \frac{\sqrt{P}}{\mathbb{E}\left[\left\|\begin{bmatrix} -\left(\mathbf{H}_{\mathrm{A}}^{\mathrm{H}}\mathbf{H}_{\mathrm{A}} + \frac{1}{P}\mathbf{I}_n\right)^{-1}\mathbf{H}_{\mathrm{A}}^{\mathrm{H}}\mathbf{H}_{\mathrm{P}}\mathbf{T}^{\mathrm{P}} \\ \mathbf{T}^{\mathrm{P}} \end{bmatrix}\right\|_{\mathrm{F}}\right]}. \tag{15}$$

This normalization constant λ^{APZF} requires only statistical CSI and can hence be applied at every TX. It ensures that an average sum power constraint is fulfilled, i.e., that

$$\mathbb{E}\left[\|\mathbf{T}^{\mathrm{APZF}}\|_{\mathrm{F}}^2\right] = P. \tag{16}$$

The key following properties can be easily shown from the precoder design and their proofs are relegated to the extended version [12].

Lemma 1. *With perfect channel knowledge at all (active) TXs, the AP-ZF precoder with n active TXs and $K-n$ passive TXs satisfies*

$$\mathbf{H}\mathbf{T}^{\mathrm{APZF}\star} \xrightarrow[P\to\infty]{} \mathbf{0}_{n\times(K-n)} \tag{17}$$

where $\mathbf{T}^{\mathrm{APZF}\star}$ denotes the AP-ZF precoder based on perfect CSIT and is given as

$$\mathbf{T}^{\mathrm{APZF}\star} \triangleq \lambda^{\mathrm{APZF}} \begin{bmatrix} \mathbf{T}^{\mathrm{A}\star} \\ \mathbf{T}^{\mathrm{P}} \end{bmatrix}. \tag{18}$$

Lemma 2. *The AP-ZF precoder with n active TXs and $K-n$ passive TXs is of rank $K-n$.*

Lemma 3. *If $\hat{\mathbf{H}}^{(j)} \triangleq \mathbf{H} + \sqrt{P^{-\alpha^{(j)}}}\boldsymbol{\Delta}^{(j)}$ for $\alpha^{(j)} \in [0,1]$ with $\boldsymbol{\Delta}^{(j)}$ a random variable with zero mean and bounded covariance matrix, it then holds that*

$$\left\|\mathbf{H}\frac{\mathbf{T}^{\mathrm{APZF}}}{\|\mathbf{T}^{\mathrm{APZF}}\|_{\mathrm{F}}}\right\|_{\mathrm{F}}^2 \dot{\leq} P^{-\min_{j\in\{1,\ldots,n\}}\alpha^{(j)}}. \tag{19}$$

The interpretation behind this result is that the interference attenuation of AP-ZF precoding is only limited by the CSIT accuracy at the active TXs, and does not depend on the CSIT accuracy at the passive TXs.

V. Weak CSIT Regime: Achievable Scheme

We now consider the weak CSIT regime and we describe the transmission scheme achieving the DoF expression given in Theorem 1. We then verify that this transmission scheme achieves the claimed DoF. Without loss of generality, we assume that the TX with the best CSIT accuracy is TX 1, i.e., that $\alpha^{(1)} = \max_{j\in\{1,\ldots,K\}}\alpha^{(j)}$.

A. Encoding

The proposed transmission scheme consists in a single channel use during which the transmitted signal $\mathbf{x} \in \mathbb{C}^K$ is

$$\mathbf{x} = \sum_{i=1}^{K}\mathbf{T}_i^{\mathrm{APZF}}\boldsymbol{s}_i + \begin{bmatrix} 1 \\ \mathbf{0}_{K-1}\end{bmatrix} s_0 \tag{20}$$

where

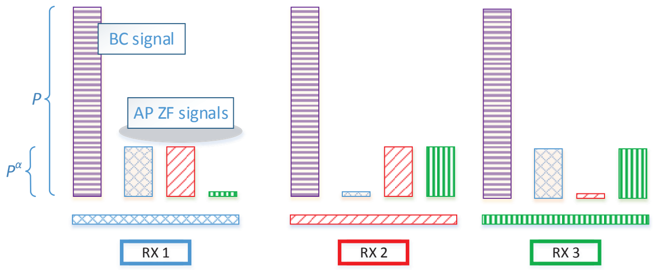

Fig. 2: Illustration of the transmission scheme for the weak CSIT regime in the case of $K=3$ users.

- $\boldsymbol{s}_i \in \mathbb{C}^{K-1}$ is a vector containing data symbols having each a rate of $\alpha^{(1)}\log_2(P)$ bits and the power $\mathbb{E}[|\{\boldsymbol{s}_i\}_\ell|^2] = P^{\alpha^{(1)}}/(K(K-1)), \forall \ell \in 1,\ldots,K-1$. $\mathbf{T}_i^{\mathrm{APZF}}$ is the AP-ZF precoder described in Section IV, with the interference being zero-forced at a single user, user $\overline{i+1}$ where $\overline{i+1} = i \bmod [K]+1$, and with TX 1 being the only active TX.
- $s_0 \in \mathbb{C}$ is a data symbol of rate $(1-\alpha^{(1)})\log_2(P)$ bits and is broadcast with the power $P - P^{\alpha^{(1)}}$ from TX 1.

Upon omitting the signals received at the noise floor, the received signal at user i is

$$y_i = \underbrace{\boldsymbol{h}_i^{\mathrm{H}}\begin{bmatrix} 1 \\ \mathbf{0}_{K-1}\end{bmatrix} s_0}_{\sim P} + \underbrace{\boldsymbol{h}_i^{\mathrm{H}}\mathbf{T}_i^{\mathrm{APZF}}\boldsymbol{s}_i}_{\sim P^{\alpha^{(1)}}} + \underbrace{\boldsymbol{h}_i^{\mathrm{H}}\sum_{k=1,k\neq i,k\neq\overline{i-1}}^{K}\mathbf{T}_k^{\mathrm{APZF}}\boldsymbol{s}_k}_{\sim(K-2)P^{\alpha^{(1)}}} \tag{21}$$

where we have written under the bracket the power scaling to help the reader. The term $\boldsymbol{h}_i^{\mathrm{H}}\mathbf{T}_{\overline{i-1}}^{\mathrm{APZF}}\boldsymbol{s}_{\overline{i-1}}$ has been removed as it was scaling in P^0 due to the attenuation by $P^{-\alpha^{(1)}}$ through AP-ZF with TX 1 being the only active TX, as shown in Lemma 3. The scaling of the received signals are illustrated in Fig. 2.

B. Interference Quantization

In fact, the broadcast data symbol s_0 does not contain only "fresh" new data symbol destinated to one user, but also quantized side information, which once decoded at the users, will be needed to decode the other data symbols. More precisely, being the most informed user, TX 1 uses its locally available CSI to compute the terms $(\hat{\boldsymbol{h}}_i^{(1)})^{\mathrm{H}}\mathbf{T}_k^{\mathrm{APZF}}\boldsymbol{s}_k$ for $k \neq i, k \neq \overline{i-1}$ which it then quantizes using $\alpha^{(1)}\log_2(P)$ bits for each term. Following well known results from rate-distortion theory [14], the distortion noise then remains at the noise floor as the SNR increases.

It can be seen from (21) that there are in total $K(K-2)$ interference terms such that the total number of bits to transmit is $K(K-2)\alpha^{(1)}\log_2(P)$. In the weak interference regime, it holds that $K(K-2)\alpha^{(1)} \leq 1-\alpha^{(1)}$. Hence, all these bits can be transmitted through the broadcast data symbol s_0 of data rate $(1-\alpha^{(1)})\log_2(P)$ bits. If the previous inequality is strict, the quantized bits transmitted in s_0 are completed with information bits destined to any particular user.

C. Decoding and DoF Analysis

It remains to verify that this scheme leads to the claimed DoF. Let us consider without loss of generality the decoding at user 1 as the decoding at the other users will follow with a circular permutation of the user's indices.

Using successive decoding, the data symbol s_0 is decoded first, followed by $\boldsymbol{s}_1$. The data symbol s_0 of rate of $(1-\alpha^{(1)})\log_2(P)$ bits can be decoded with vanishing probability of error as its SINR can be seen in (21) to scale in $P^{1-\alpha^{(1)}}$.

Upon decoding s_0, the quantized interference $(\hat{\boldsymbol{h}}_1^{(1)})^{\mathrm{H}}\mathbf{T}_k^{\mathrm{APZF}}\boldsymbol{s}_k,\ k \in \{2,\ldots,K-1\}$ are computed (up to quantization noise) and exploited by user 1 for interference cancellation. Yet, it remains to evaluate the impact of the imperfect estimation at TX 1:

$$(\hat{\boldsymbol{h}}_1^{(1)})^{\mathrm{H}}\mathbf{T}_k^{\mathrm{APZF}}\boldsymbol{s}_k = \boldsymbol{h}_1^{\mathrm{H}}\mathbf{T}_k^{\mathrm{APZF}}\boldsymbol{s}_k + \underbrace{P^{-\alpha^{(1)}}(\boldsymbol{\delta}_1^{(1)})^{\mathrm{H}}\mathbf{T}_k^{\mathrm{APZF}}\boldsymbol{s}_k}_{\sim P^0} \tag{22}$$

It follows from (22) that subtracting the estimated interference from the received signals can be done perfectly in terms of DoF.

Remark 3. Note that this is possible only because the transmit power is lower than $P^{\alpha^{(1)}}$. It this is not the case, TX 1 is not able to estimate the interference generated up to the noise floor. □

After having subtracted the quantized interference terms, the remaining signal at user 1 is then (up to the noise floor)

$$y_1 = \boldsymbol{h}_1^{\mathrm{H}}\mathbf{T}_1^{\mathrm{APZF}}\boldsymbol{s}_1. \tag{23}$$

The quantized interference terms $(\hat{\boldsymbol{h}}_i^{(1)})^{\mathrm{H}}\mathbf{T}_1^{\mathrm{APZF}}\boldsymbol{s}_1, i = 3,\ldots,K$, which have been obtained through s_0, are then used to form a virtual received vector $\boldsymbol{y}_1^{\mathrm{v}} \in \mathbb{C}^{K-1}$ equal to

$$\boldsymbol{y}_1^{\mathrm{v}} \triangleq \begin{bmatrix} \boldsymbol{h}_1^{\mathrm{H}} \\ (\hat{\boldsymbol{h}}_3^{(1)})^{\mathrm{H}} \\ \vdots \\ (\hat{\boldsymbol{h}}_K^{(1)})^{\mathrm{H}} \end{bmatrix} \mathbf{T}_1^{\mathrm{APZF}}\boldsymbol{s}_1. \tag{24}$$

Each component of $\boldsymbol{y}_1^{\mathrm{v}}$ has a SINR scaling in $P^{\alpha^{(1)}}$ and the AP-ZF precoder is of rank $K-1$ (See Lemma 2) such that user 1 can decode its desired $K-1$ data symbols, each with the rate of $\alpha^{(1)}\log_2(P)$ bits.

In one channel use, $(K-1)\alpha^{(1)}\log_2(P)$ bits are transmitted to each user and $(1-\alpha^{(1)}-K(K-2)\alpha^{(1)})\log_2(P)$ bits are transmitted to any particular user through the data symbol s_0. Adding all the terms provides the desired DoF expression.

Remark 4. Interestingly, the above scheme builds on the principle of interference quantization and retransmission, which has already been exploited for the different context of precoding with delayed CSIT (see e.g. [14], [15]). In contrast, the distributed nature of the CSIT is exploited here such that the interference terms are estimated and transmitted from the TX having the most accurate CSIT, *during the same channel use in which the interference terms were generated.* □

VI. Conclusion

We have described a new D-CSIT robust transmission schemes improving over the DoF achieved by conventional precoding approaches when faced with distributed CSIT and being even DoF-optimal over a certain "weak CSIT regime". We have then uncovered the surprising result that in that case, the optimal DoF was the same as in a genie-aided CSIT configuration where the most accurate channel estimate is made available at all the TXs. The new scheme relies on several new methods such as the estimation of the interference and their retransmission from a single TX, and the AP-ZF precoding with multiple passive TXs, which could certainly prove useful to improve the DoF in other wireless configurations.

VII. Acknowledgment

The authors are grateful to Inaki Estella, Petros Elia, Dirk Slock, and Shlomo Shamai for interesting and helpful discussions.

References

[1] G. Caire and S. Shamai (Shitz), "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, 2003.

[2] S. A. Jafar and A. J. Goldsmith, "Isotropic fading vector broadcast channels: The scalar upper bound and loss in degrees of freedom," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 848–857, Mar. 2005.

[3] M. W. Amos Lapidoth, Shlomo Shamai, "On the capacity of fading MIMO broadcast channels with imperfect transmitter side-information," 2006. [Online]. Available: http://arxiv.org/pdf/cs/0605079.pdf

[4] A. G. Davoodi and S. A. Jafar, "Aligned image sets under channel uncertainty: Settling a conjecture by Lapidoth, Shamai and Wigger on the collapse of degrees of freedom under finite precision CSIT," 2014. [Online]. Available: http://arxiv.org/abs/1403.1541

[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] Q. Li, P. de Kerret, D. Gesbert, and N. Gresset, "Robust Regularized ZF in Decentralized Broadcast Channel with Correlated CSI Noise," in *Proc. Allerton Conference on Communication, Control, and Computing (Allerton)*, 2015.

[7] P. de Kerret, R. Fritzsche, D. Gesbert, and U. Salim, "Robust precoding for network MIMO with hierarchical CSIT," in *Proc. IEEE International Symposium on Wireless Communication Systems (ISWCS)*, 2014.

[8] R. Fritzsche and G. Fettweis, "Distributed robust sum rate maximization in cooperative cellular networks," in *Proc. IEEE Workshop on Cooperative and Cognitive Mobile Networks (CoCoNet)*, 2013.

[9] X. Rao, L. Ruan, and V. K. N. Lau, "CSI feedback reduction for MIMO interference alignment," *IEEE Trans. Signal Process.*, vol. 61, no. 18, pp. 4428–4437, Sept. 2013.

[10] P. de Kerret and D. Gesbert, "Interference alignment with incomplete CSIT sharing," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2563–2573, May. 2014.

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

[12] P. de Kerret and D. Gesbert, "Network MIMO: Transmitters with no CSI Can Still be Very Useful," 2016, https://arxiv.org/abs/1601.07399.

[13] T. Cover and A. Thomas, *Elements of information theory.* Wiley-Interscience, Jul. 2006.

[14] S. Yang, M. Kobayashi, D. Gesbert, and X. Yi, "Degrees of freedom of time correlated MISO broadcast channel with delayed CSIT," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 315–328, Jan. 2013.

[15] T. Gou and S. Jafar, "Optimal use of current and outdated channel state information: Degrees of freedom of the MISO BC with mixed CSIT," *IEEE Communications Letters*, vol. 16, no. 7, pp. 1084–1087, Jul. 2012.