Enforcing Coordination in Network MIMO with Unequal CSIT

Paul de Kerret, Antonio Bazco, and David Gesbert
Communication Systems Department, EURECOM

***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 where 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. In that setting, a novel transmission scheme has been developed in [1] and surprisingly it was shown that in a so-called regime of weak-CSIT it was possible to achieve a centralized-outerbound obtained when all TXs exchange their local channel estimates. The main novelty of the scheme was the estimation of the interference at the TX having the most accurate CSIT, followed by the quantization and the retransmission of these interference terms using superposition coding. In this work, we develop a new precoding scheme, coined as *Hierarchical ZF* (HZF), which allows to efficiently exploit the CSI estimates at all the TXs instead of relying on the estimate at a single TX as in [1]. Combining HZF with the transmission scheme of [1], a novel transmission scheme significantly enlarging the so-called weak CSIT regime has been designed. Finally, HZF precoding leads to a stronger interference attenuation as state-of-the-art precoding from the literature and is expected to be useful in other wireless settings with unequal CSIT.**

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 [2]. 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 Broadcast Channel (BC) [3]. Going further, a long standing conjecture by Lapidoth, Shamai, and Wigger [4] has been recently settled in [5] 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however, *centralized* CSIT is typically assumed, i.e., precoding is done on the basis of a *single* imperfect/outdated multiuser channel estimate being common at *every* transmit antenna.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).

To account for TX dependent feedback limitations, a *distributed CSIT* model has been introduced in [6]. 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 [7]–[11].

In terms of DoF, it was shown in [6] that using a conventional ZF precoder 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. More recently, a transmission scheme improving the DoF was provided in [1]. Interestingly, it was shown that for some CSIT configurations forming a so-called *weak CSIT regime*, it was possible to achieve a *centralized-outerbound* obtained with full exchange of the channel estimates between all TXs.

In this work, we improve this result by developing a new precoding scheme called HZF which leads to an interference reduction stronger than state-of-the-art schemes from the literature. Combining HZF with the main ideas in [1] (interference estimation at the TX having the most accurate CSIT, quantization, and retransmission using superposition coding) has lead to a new transmission scheme extending significantly the so-called weak CSIT regime over which it is possible to achieve the centralized outerbound.

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. Let $\mathbf{A}$ be a matrix of size $n \times n$, we denote by $\mathbf{A}_{[i:j,k:\ell]}$ for $i,j,k,\ell \in \{1,\dots,n\}$ be the submatrix of $\mathbf{A}$ formed by selecting the rows $i,\dots,j$ and the columns $k,\dots,\ell$.

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 $\mathbf{H} \triangleq [\boldsymbol{h}_1, \ldots, \boldsymbol{h}_K]^{\mathrm{H}} \in \mathbb{C}^{K \times K}$. 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 multi-user channel estimate 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 $\hat{\mathbf{H}}^{(j)} \triangleq [\hat{\boldsymbol{h}}_1^{(j)}, \ldots, \hat{\boldsymbol{h}}_K^{(j)}]^{\mathrm{H}} \in \mathbb{C}^{K \times K}$ 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 $\hat{\mathbf{H}}^{(j)}$.

We model the CSI uncertainty at TX j 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. 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 [5], [12].

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

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

Without loss of generality, we assumed that the TXs are ordered such that

$$\alpha^{(1)} \geq \alpha^{(2)} \geq \ldots \geq \alpha^{(K)}. \tag{4}$$

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

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

Remark 1. This condition extends the condition provided in [5] which writes in our setting as

$$\max_{\mathbf{H} \in \mathbb{C}^{K \times K}} \left(p_{\mathbf{H}|\hat{\mathbf{H}}^{(j)}}(\mathbf{H}) \right) \doteq \sqrt{P^{\alpha^{(j)}}}, \quad \forall j \in \{1, \ldots, K\}. \tag{6}$$

□

Condition (6) is a mild technical assumption, which holds for the distributions usually considered, as for example with Gaussian random variables [1].

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

$$\mathrm{DoF}^{\mathrm{DCSI}}(\boldsymbol{\alpha}) \triangleq \lim_{P \to \infty} \frac{\mathcal{C}(P)}{\log_2(P)}. \tag{7}$$

III. Main Results in the 3-user Case

In order to better convey the main intuition, we present in this work the 3-user case. The approach easily extends to an arbitrary number of user and the K-user case will be described in the extended journal version.

First, we extend the notion of *Weak CSIT regime* from [1].

Definition 1. *In the* 3*-user MISO BC with D-CSIT, we define the weak CSIT regime as comprising the CSIT configurations that satisfy*

$$\alpha^{(1)} \leq \frac{1}{4} + \frac{3}{4}\alpha^{(2)}. \tag{8}$$

We can now state one of our main results.

Theorem 1. *In the weak CSIT regime defined in Definition 1, the optimal sum-*DoF *of the* 3*-user MISO BC with D-CSIT is given by*

$$\mathrm{DoF}^{\mathrm{DCSI}}(\boldsymbol{\alpha}) = 1 + 2\alpha^{(1)}. \tag{9}$$

Comparing this result with the weak CSIT regime defined in [1], the contribution of this works resides in the addition of the term $\frac{3}{4}\alpha^{(2)}$. This additional term is a consequence of using the novel HZF precoding scheme as it will become clear in the following.

Going beyond the weak CSIT regime, we introduce the following definitions.

Definition 2. *In the* 3*-user MISO BC with D-CSIT, we define the heterogeneous CSIT regime as comprising the CSIT configurations that satisfy that*

$$\alpha^{(1)} \geq \min\left(2\alpha^{(2)}, \frac{1}{4} + \frac{3}{4}\alpha^{(2)}\right) \tag{10}$$

while the intermediate CSIT regime contains the CSIT configuration satisfying

$$\frac{1 + 3\alpha^{(2)}}{4} \leq \alpha^{(1)} \leq 2\alpha^{(2)}. \tag{11}$$

Extending the transmission scheme outside the weak-CSIT regime leads to the following achievable DoF.

Theorem 2. *In the* 3*-user MIMO BC with D-CSIT, it holds that*

$$\mathrm{DoF}^{\mathrm{DCSI}}(\boldsymbol{\alpha}) \geq \begin{cases} 1 + 2\alpha^{(1)} \text{ (Weak CSIT)} \\ \frac{3}{2}(1 + \alpha^{(2)}) \text{ (Intermediate CSIT)} \\ 1 + \alpha^{(1)} + \frac{3\alpha^{(1)}(1-\alpha^{(1)}) + \alpha^{(2)}\left(5\alpha^{(1)} - 3\alpha^{(2)} - 1\right)}{9\alpha^{(1)} - 8\alpha^{(2)}} \\ \qquad \text{(Heterogeneous CSIT)} \end{cases}$$

The different CSIT configurations are illustrated in Fig. 1. As it can be clearly seen, one of the main contribution of this work is the significant extension of the weak CSIT region.

This is practically interesting as it contains all the CSIT configurations over which it has been possible to achieve the optimal DoF. Furthermore, this weak CSIT regime does not depend only on $\alpha^{(1)}$ such that the results better adapt to more users. Note that the different CSIT regimes only depend on $\alpha^{(1)}$ and $\alpha^{(2)}$ but not on $\alpha^{(3)}$, in agreement with previous results [1], [6].

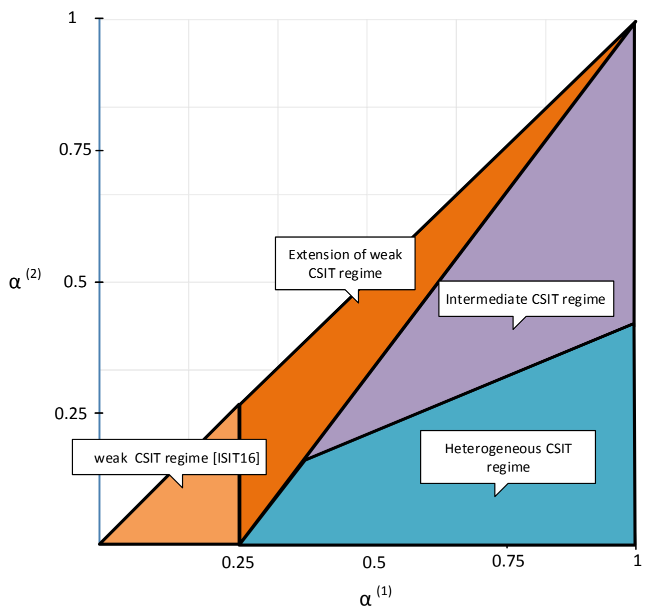

Fig. 1: Representation of the different CSIT configurations as a function of the CSIT scaling coefficients $\alpha^{(1)}$ and $\alpha^{(2)}$

In Fig. 2, the sum DoF is shown as a function of $\alpha^{(1)}$ for $\alpha^{(2)} = 2\alpha^{(1)}/3$ and $\alpha^{(3)} = 0$. The proposed algorithm is compared to the scheme in [1] and to conventional (regularized) ZF. The improvement compared to the approach in [1] is represented by the red triangle.

The novel transmission scheme extends the scheme of [1] and relies on the same ideas: estimate and quantize the interference (before their generation) at the TX having the most accurate estimate (i.e. TX 1), and then transmit these quantized interference terms using superposition coding to all the RXs. The main improvement with respect to [1] consists in using a novel precoding scheme, called *Hierarchical ZF* (HZF), which allows to reduce the amount of interference generated, thus reducing the amount of information to retransmit. The transmission scheme achieving the results in Theorem 2 then follows without major difficulty such that we focus in the rest of this work on the detailed presentation of HZF precoding. The description of the full scheme combining the approach of [1] and HZF will be given in the extended journal version.

IV. Hierarchical Zero-Forcing

HZF is a distributed precoding scheme in which each TX exploits its locally available CSIT to improve over the precoding of the TXs having less accurate CSIT. Therefore,

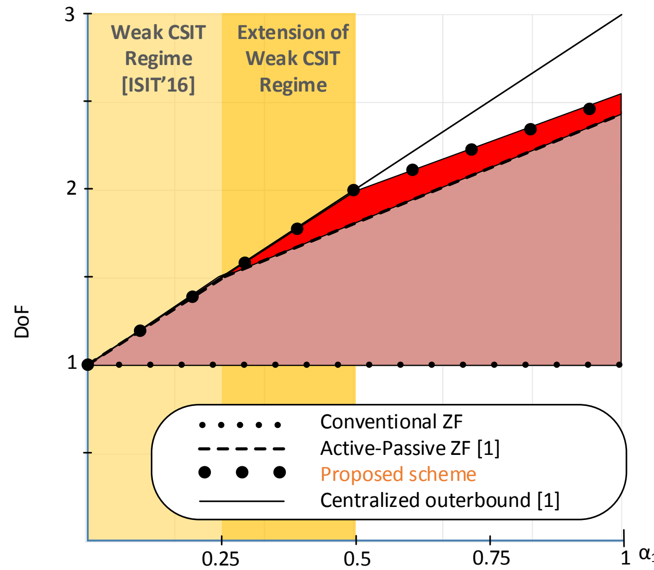

Fig. 2: Sum DoF as a function of $\alpha(1)$ for $\alpha^{(2)} = 2\alpha^{(1)}/3$ and $\alpha^{(3)} = 0$

the main interest of the HZF precoding lies in the fact that it satisfies the following Lemma.

Lemma 1. *In the K-user MISO BC with D-CSIT, if we denote by $\mathbf{t}_K^{\text{HZF}}$ the HZF beamformer towards user K with average power P, it then holds that*

$$|\boldsymbol{h}_j^{\text{H}} \boldsymbol{t}_K^{\text{HZF}}|^2 \dot{\leq} P^{1-\alpha^{(j)}}, \qquad j = 1, \dots, K-1. \tag{12}$$

The interference attenuation obtained using HZF is illustrated in the 3-user case in Fig. 3. Each of the $K-1$ best CSIT coefficients controls the interference attenuation at one user. The design of the HZF beamformer is given in the following while the proof of Lemma 1 is given in the Appendix.

The main idea of HZF consists in letting TX j reproduce the signal processing realized at TX k for $k > j$, i.e., at the TXs having a *less* accurate CSIT. This is made possible thanks to a particular Hierarchical Quantizer (HQ) introduced first in [14] in the different context of delayed CSIT, and recalled for completeness in Subsection IV-A. Applying this quantizer at each TX allows to make the CSIT configuration hierarchical.

Once the precoding at the TXs having less accurate CSIT is reproduced at a given TX, this TX then exploits its locally available CSIT to further reduce the interference at one user while preserving the interference reduction already realized by the TXs having less accurate CSIT. The precoder design achieving this goal is detailed in Subsection IV-B.

A. Hierarchical Quantizer

We start by recalling the original lemma given in [14, Lemma 1] before providing an intuitive example to illustrate its use in this work.

Lemma 2 ([14, Lemma 1])**.** *Let Y be a unit-variance zero-mean random variable of bounded density p_Y. Then, there*

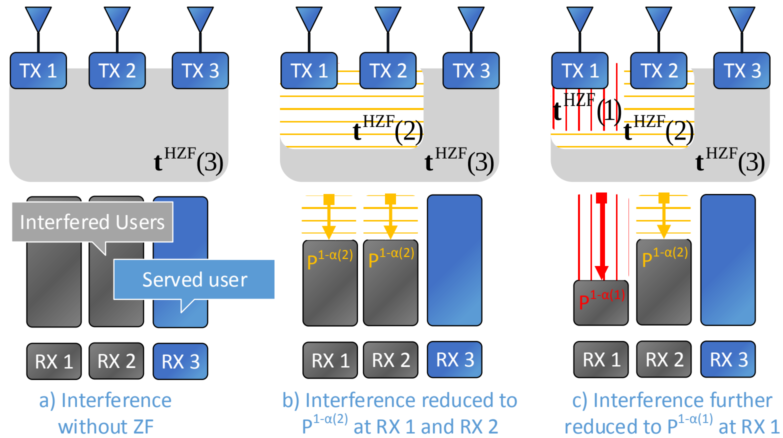

Fig. 3: Illustration of the interference reduction of the interference at each step of the HZF Algorithm.

exists a quantizer of rate $\beta \log_2(P)$ bits for any $0 < \beta \le 1$, denoted by Q_β, such that, for any unit-variance zero-mean random variable n, it holds that

$$\lim_{P\to\infty} \Pr\{Q_\beta(y+\sqrt{P^{-\beta}}n) = Q_\beta(y)\} = 1 \tag{13}$$

and at the same time

$$\mathbb{E}\left[|Q_\beta(y)-y|^2\right] \dot{\le} P^{-\beta}. \tag{14}$$

Let us now present a simple toy-example highlighting how this lemma can be helpful to deal with D-CSIT.

Example 1. *Let us consider two TXs, TX 1 and TX 2, with scaling coefficients $\alpha^{(1)}$ and $\alpha^{(2)}$, respectively, and $\alpha^{(1)} \ge \alpha^{(2)}$. Let us define*

$$\hat{\hat{\mathbf{H}}}^{(1)}_{\alpha^{(2)}} \triangleq Q_{\alpha^{(2)}}\left(\hat{\mathbf{H}}^{(1)}\right) \tag{15}$$

$$\hat{\hat{\mathbf{H}}}^{(2)}_{\alpha^{(2)}} \triangleq Q_{\alpha^{(2)}}\left(\hat{\mathbf{H}}^{(2)}\right). \tag{16}$$

Then, as result of the quantizer's properties mentioned above, it holds that

$$\lim_{P\to\infty} \Pr\left\{\hat{\hat{\mathbf{H}}}^{(1)}_{\alpha^{(2)}} = \hat{\hat{\mathbf{H}}}^{(2)}_{\alpha^{(2)}}\right\} = 1 \tag{17}$$

$$\mathbb{E}\left[\|\hat{\hat{\mathbf{H}}}^{(j)}_{\alpha^{(2)}} - \mathbf{H}\|^2_{\rm F}\right] \dot{\le} P^{-\alpha^{(2)}}, \qquad j = 1,2. \tag{18}$$

Thus, at high SNR, we can consider that TX 1 has access to $\hat{\hat{\mathbf{H}}}^{(1)}_{\alpha^{(1)}}$, and $\hat{\hat{\mathbf{H}}}^{(2)}_{\alpha^{(2)}}$, and TX 2 has access to $\hat{\hat{\mathbf{H}}}^{(2)}_{\alpha^{(2)}}$, with the same CSIT scaling coefficients. As the CSIT scaling of the estimate has remained the same, this approach effectively transforms the CSIT configuration into a hierarchical one, without degradation of the CSIT quality in terms of DoF. □

B. Hierarchical ZF Precoding at TX j

We now focus on the computation of the HZF precoder aimed at TX K, which we denote by $\boldsymbol{t}^{\rm HZF}$ (i.e., we omit the index K for the sake of clarity) at TX j using the locally available CSIT $\hat{\mathbf{H}}^{(j)}$ and the knowledge of the CSIT configuration (i.e., $\boldsymbol{\alpha}$). As a preliminary step, the precoder is decomposed as

$$\boldsymbol{t}^{\rm HZF} = \sum_{\ell=1}^{K} \begin{bmatrix} \boldsymbol{t}^{\rm HZF}(\ell) \\ \mathbf{0}_{K-\ell} \end{bmatrix} \tag{19}$$

where $\boldsymbol{t}^{\rm HZF}(\ell) \in \mathbb{C}^\ell$ is vector designed at the first (most accurate) ℓ TXs. As a consequence of the distributed precoding, this means that it is sufficient for TX j to compute the vectors $\boldsymbol{t}^{\rm HZF}(\ell)$ for $\ell \in \{j,\dots,K\}$ to obtain its precoding coefficient (i.e., the jth element of $\boldsymbol{t}^{\rm HZF}$). We now present the precoding algorithm used at TX j to compute these vectors.

First, the estimate $\hat{\mathbf{H}}^{(j)}$ is quantized using the Hierarchical quantizer described in Subsection IV-A such that TX j obtains

$$\hat{\hat{\mathbf{H}}}^{(\ell)} \triangleq Q_{\alpha^{(\ell)}}(\hat{\mathbf{H}}^{(j)}), \qquad \forall \ell \in \{K,\dots,j\}. \tag{20}$$

Remark 2. Note that the probability that TX j is able to accurately compute the estimate at TX ℓ for $\ell > j$, i.e., that

$$Q_{\alpha^{(\ell)}}(\hat{\mathbf{H}}^{(j)}) = Q_{\alpha^{(\ell)}}(\hat{\mathbf{H}}^{(\ell)}), \qquad \ell > j, \tag{21}$$

goes only to 1 as the SNR goes to infinity. However, as we study the DoF, and for the sake of clarity, we do the abuse of notation of simply writing $\hat{\hat{\mathbf{H}}}^{(\ell)}$. □

If TX j has the worst CSIT (i.e., $j = K$), then it computes

$$\boldsymbol{t}^{\rm HZF}(K) = \lambda^{\rm HZF} \hat{\hat{\mathbf{H}}}^{(j){\rm H}} \left(\hat{\hat{\mathbf{H}}}^{(j)} (\hat{\hat{\mathbf{H}}}^{(j)})^{\rm H} + \frac{1}{P}\mathbf{I}_K\right)^{-1} \tag{22}$$

with the power normalization $\lambda^{\rm HZF} \in \mathbb{R}^+$ being given by

$$\lambda^{\rm HZF} \triangleq \frac{\sqrt{P}}{\sqrt{\mathbb{E}\left[\left\|\sum_{\ell=1}^{K} \begin{bmatrix} \boldsymbol{t}^{\rm HZF}(\ell) \\ \mathbf{0}_{K-\ell} \end{bmatrix}\right\|^2\right]}}. \tag{23}$$

Otherwise, for every $n \in \{K-1,\dots,j\}$, it computes

$$\boldsymbol{t}^{\rm HZF}(n) = -\hat{\hat{\mathbf{H}}}^{(n){\rm H}}_{[1:n,1:n]} \left(\hat{\hat{\mathbf{H}}}^{(n)}_{[1:n,1:n]} \hat{\hat{\mathbf{H}}}^{(n){\rm H}}_{[1:n,1:n]} + \frac{1}{P}\mathbf{I}_n\right)^{-1} \cdot \hat{\hat{\mathbf{H}}}^{(n)}_{[1:n,1:K]} \left(\sum_{\ell=n+1}^{K} \begin{bmatrix} \boldsymbol{t}^{\rm HZF}(\ell) \\ \mathbf{0}_{K-\ell} \end{bmatrix}\right). \tag{24}$$

It is then shown in the Appendix how this design allows to satisfy the interference attenuation claimed in Lemma 1 and illustrated in Fig. 3.

V. Conclusion

Considering the MISO BC with distributed CSIT, we have developed a new precoding scheme, called Hierarchical ZF, which allows to efficiently exploit the CSI estimates of different qualities at all TXs thanks to an increased coordination between the TXs. Combining HZF with the new approach of interference estimation, quantization, and retransmission using superposition coding at the TX having the most accurate TX, has allowed to significantly enlarge the achievable DoF region. Interestingly, the novel scheme also extends the so-called weak CSIT regime over which it is possible to achieve the centralized-outerbound, thus providing the optimal sum-DoF expression. Whether this weak CSIT regime can be further extended or finding the optimal DoF outside this CSIT regime are very interesting directions of research. Finally, the novel HZF precoding is expected to be a useful tool in other wireless settings with unequal CSIT.

VI. Appendix

A. Proof of Lemma 1

For the sake of clarity, we present the proof in the 3-user case. The proof for an arbitrary number of user will be given in the extended journal version but does not present any additional difficulties. Furthermore, as already mentioned in Remark 2, we always consider that TX j has been able to obtain accurately $\hat{\hat{\mathbf{H}}}^{(\ell)}$ for $\ell > j$ as the probability that this holds true tends to one as P tends to infinity.

a) Interference at RX 2: Let us focus first on RX 2.

$$\boldsymbol{h}_2^{\mathrm{H}}\boldsymbol{t}^{\mathrm{HZF}} = \boldsymbol{h}_2^{\mathrm{H}} \begin{bmatrix} t^{\mathrm{HZF}}(1) \\ 0 \\ 0 \end{bmatrix} + \boldsymbol{h}_2^{\mathrm{H}} \left(\begin{bmatrix} \boldsymbol{t}^{\mathrm{HZF}}(2) \\ 0 \end{bmatrix} + \boldsymbol{t}^{\mathrm{HZF}}(3) \right). \quad (25)$$

By design of $\boldsymbol{t}^{\mathrm{HZF}}(2)$, the second term of (25) satisfies that

$$\boldsymbol{h}_2^{\mathrm{H}} \left(\begin{bmatrix} \boldsymbol{t}^{\mathrm{HZF}}(2) \\ 0 \end{bmatrix} + \boldsymbol{t}^{\mathrm{HZF}}(3) \right) \doteq \boldsymbol{\delta}_2^{(2)\mathrm{H}} \left(\begin{bmatrix} \boldsymbol{t}^{\mathrm{HZF}}(2) \\ 0 \end{bmatrix} + \boldsymbol{t}^{\mathrm{HZF}}(3) \right) \quad (26)$$

due to the fact that

$$\hat{\hat{\boldsymbol{h}}}_2^{(2)\mathrm{H}} \left(\begin{bmatrix} \boldsymbol{t}^{\mathrm{HZF}}(2) \\ 0 \end{bmatrix} + \boldsymbol{t}^{\mathrm{HZF}}(3) \right) \xrightarrow[P\to\infty]{} 0. \quad (27)$$

Thus, it holds that

$$\left| \boldsymbol{h}_2^{\mathrm{H}} \left(\begin{bmatrix} \boldsymbol{t}^{\mathrm{HZF}}(2) \\ 0 \end{bmatrix} + \boldsymbol{t}^{\mathrm{HZF}}(3) \right) \right|^2 \dot{\leq} P^{1-\alpha^{(2)}}. \quad (28)$$

Turning to the first term of (25), the precoding coefficient $t^{\mathrm{HZF}}(1)$ is obtained from (24) and can be written as

$$t^{\mathrm{HZF}}(1) = -\hat{\hat{\mathbf{H}}}_{[1,1]}^{(1)\mathrm{H}} \left(\hat{\hat{\mathbf{H}}}_{[1,1]}^{(1)} \hat{\hat{\mathbf{H}}}_{[1,1]}^{(1)\mathrm{H}} + \frac{1}{P} \right)^{-1} \cdot \hat{\hat{\boldsymbol{h}}}_1^{(1)} \left(\begin{bmatrix} \boldsymbol{t}^{\mathrm{HZF}}(2) \\ 0 \end{bmatrix} + \boldsymbol{t}^{\mathrm{HZF}}(3) \right) \quad (29)$$

Applying the same calculation as for (28), we obtain

$$\left| \boldsymbol{h}_1^{\mathrm{H}} \left(\begin{bmatrix} \boldsymbol{t}^{\mathrm{HZF}}(2) \\ 0 \end{bmatrix} + \boldsymbol{t}^{\mathrm{HZF}}(3) \right) \right|^2 \dot{\leq} P^{1-\alpha^{(2)}}. \quad (30)$$

Inserting (30) in (29) then gives

$$|t^{\mathrm{HZF}}(1)|^2 \dot{\leq} P^{1-\alpha^{(2)}}. \quad (31)$$

Inserting (31) and (28) in (25) yields the result for RX 2.

b) Interference at RX 1: Similar to (27), it also holds by construction of $t^{\mathrm{HZF}}(1)$ that

$$\hat{\hat{\boldsymbol{h}}}_1^{(1)\mathrm{H}} \left(\begin{bmatrix} t^{\mathrm{HZF}}(1) \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} \boldsymbol{t}^{\mathrm{HZF}}(2) \\ 0 \end{bmatrix} + \boldsymbol{t}^{\mathrm{HZF}}(3) \right) \xrightarrow[P\to\infty]{} 0 \quad (32)$$

from which it follows directly that

$$\left| \boldsymbol{h}_1^{\mathrm{H}} \begin{bmatrix} t^{\mathrm{HZF}}(1) \\ 0 \\ 0 \end{bmatrix} + \boldsymbol{h}_2^{\mathrm{H}} \left(\begin{bmatrix} \boldsymbol{t}^{\mathrm{HZF}}(2) \\ 0 \end{bmatrix} + \boldsymbol{t}^{\mathrm{HZF}}(3) \right) \right|^2 \dot{\leq} P^{1-\alpha^{(1)}}, \quad (33)$$

which concludes the proof.

References

[1] P. de Kerret and D. Gesbert, "Network MIMO: Transmitters with no CSI Can Still be Very Useful," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, 2016.

[2] 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.

[3] 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.

[4] 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

[5] A. G. Davoodi and S. A. Jafar, "Aligned image sets under channel uncertainty: Settling conjectures on the collapse of Degrees of Freedom under finite precision CSIT," *IEEE Trans. Inf. Theo.*, vol. 62, no. 10, pp. 5603–5618, Oct. 2016.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

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

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

[14] P. de Kerret, D. Gesbert, J. Zhang, and P. Elia, "Optimally Bridging the Gap from Delayed to Perfect CSIT in the K-user MISO BC," in *Proc. IEEE Information Theory Workshop (ITW)*, 2016.