Quantized Team Precoding: A Robust Approach for Network MIMO under General CSI Uncertainties

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

***Abstract*—The[1] obtaining of accurate channel state information (CSI) at the transmitter (TX) side is known to be one of the main limitations for joint precoding across distant multiple-antennas TXs. Indeed, the knowledge of the *multi-user* CSI usually requires an information sharing step between the TXs, which makes cooperation challenging in many wireless settings. To alleviate the problem, we study the transmission when the CSI is *imperfectly* shared between the TXs, a setting we refer as *Distributed* CSIT (D-CSIT). In such a case, the optimization of the joint precoder distributed across the distant TXs becomes a very different problem from the precoding usually encountered when all antennas receive the same channel estimate. Indeed, the precoder design then falls in the category of Team-Decision (TD) problems: Each TX does not know the information available (hence the precoder) at the other TXs. In this paper, we introduce a novel framework referred to as *Quantized Team Precoding* (QTP) which allows for robust distributed precoding in the presence of such CSI uncertainties. The method relies on the concept of consistency enforcement, which is obtained by quantizing the CSI space. The obtained precoders exhibit superior rate performance compared to precoders that classical robust designs from the literature.**

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 based on Cloud-RAN (C-RAN) [1], [2]. With perfect message and CSI sharing, the different TXs can be seen as a unique virtual multiple-antenna array serving all Receivers (RXs) in a multiple-antenna Broadcast Channel (BC) fashion, and well known precoding algorithms from the literature can be used [3], [4]. This joint precoding requires however the feedback of an accurate multi-user CSI to each TX in order to achieve the desired high performances [5], [6]. As a consequence, an important number of works have been focused on the feedback design (See [7] and references therein) and the design of robust precoders (See for example [8], [9]).

Yet, the large literature dealing with imperfect CSI on the TX side typically assumes *centralized* CSIT [2], i.e., that the precoding is based on the basis of a *single* imperfect channel estimate at every TX. In practice, this means assuming that the precoding is either done in a central node or that the channel estimates are *perfectly shared* between all TXs. Although meaningful in the single TX case with multiple-antennas, this assumption is likely to be breached in certain networks where flexible, heterogeneous and inexpensive backhaul links (or even flying relays) are to be used. In such cases, it is expected that the CSI exchange will introduce additional delay and quantization noise. It is thus practically relevant for joint precoding across distant TXs to study a model where each TX receives its *own* multi-user channel estimate, which we denote as the *Distributed CSI* (D-CSI) configuration [10].

We consider here a D-CSI setting where each TX designs its transmit coefficients solely on the basis of its own CSIT *without any additional communication with the others TXs*. Solving this optimization problem is particularly challenging (referred to in the literature as a Team Decision problem [11]) and only partial results are available. In [12], the number of Degrees-of-freedom (DoF) obtained with conventional zero-forcing precoding is derived and some Team precoding schemes improving the DoF are given. In [13], a robust precoding algorithm is designed for the case of two TXs having distributed CSIT. In [14], an algorithm is designed in the particular setting of so-called hierarchical D-CSIT.

In this work, we provide a novel precoding algorithm, called *Quantized Team Precoding (QTP)* precoding, which transmits in a robust manner with respect to the TXs having unequal channel estimates. This is the first robust precoding algorithm for Team Decision which is generic in the sense that it can be used in any antenna, fading, and CSI configurations. Our algorithm relies on the *quantization of the CSI space* which allows to enforce coordination in this Team Decision configuration.

II. System Model

A. Received Signal

We study Network MIMO transmission from K TXs to K RXs where the i-th TX is equipped with M_i antennas and transmits d_i streams to the i-th RX equipped with N_i antennas. The total number of RX antennas, the total number of TX antennas and the total number of streams are respectively given by

$$N_{\mathrm{tot}} \triangleq \sum_{i=1}^{K} N_i, \quad M_{\mathrm{tot}} \triangleq \sum_{i=1}^{K} M_i, \quad d_{\mathrm{tot}} \triangleq \sum_{i=1}^{K} d_i. \qquad (1)$$

We further assume that the RXs have perfect CSI as there is a common agreement that the information at the TXs is often a bottleneck. We consider that linear filtering is used at both the TXs and the RXs. The channel from the K TXs to the K RXs is represented by the multi-user channel matrix $\mathbf{H} \in$

[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).

$\mathbb{C}^{N_{\text{tot}}\times M_{\text{tot}}}$ where $\mathbf{H}_{i,k}\in\mathbb{C}^{N_i\times M_k}$ denotes the channel matrix from TX k to RX i.

The transmission is then described as

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

where $\boldsymbol{y}_i\in\mathbb{C}^{N_i\times 1}$ is the signal received at the i-th RX, $\mathbf{H}_i\in\mathbb{C}^{N_i\times M_{\text{tot}}}$ the channel from all TXs to the i-th RX, and $\boldsymbol{\eta}\triangleq[\boldsymbol{\eta}_1,\ldots,\boldsymbol{\eta}_K]^{\mathrm{T}}\in\mathbb{C}^{N_{\text{tot}}\times 1}$ the normalized Gaussian noise with its elements i.i.d. as $\mathcal{CN}(0,1)$.

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

$$\mathbf{x}=\mathbf{T}\boldsymbol{s}=\begin{bmatrix}\mathbf{T}_1,\ldots,\mathbf{T}_K\end{bmatrix}\begin{bmatrix}\boldsymbol{s}_1\\ \vdots\\ \boldsymbol{s}_K\end{bmatrix}=\sum_{k=1}^{K}\mathbf{T}_k\boldsymbol{s}_k \tag{3}$$

with $\mathbf{T}_j\in\mathbb{C}^{M_{\text{tot}}\times d_j}$ being the joint precoder serving user j and $\mathbf{T}\in\mathbb{C}^{M_{\text{tot}}\times d_{\text{tot}}}$ being the multi-user precoder. We also introduce the matrix $\mathbf{W}_j\in\mathbb{C}^{M_j\times d_{\text{tot}}}$ to denote the precoding coefficients at TX j such that the signal transmitted by TX j, denoted by $\mathbf{x}_j\in\mathbb{C}^{M_j\times 1}$, is given by

$$\mathbf{x}_j=\mathbf{W}_j\boldsymbol{s}. \tag{4}$$

The multi-user precoder $\mathbf{T}$ can then alternatively be written as

$$\mathbf{T}=\begin{bmatrix}\mathbf{W}_1\\ \vdots\\ \mathbf{W}_K\end{bmatrix}. \tag{5}$$

Finally, the received signal at RX k is filtered through by $\mathbf{G}_k^{\mathrm{H}}\in\mathbb{C}^{d_k\times N_k}$. The Mean Square Error (MSE) matrix at RX k for given precoders and RX filters, denoted by $\mathbf{M}_k\in\mathbb{C}^{d_k\times d_k}$, is then

$$\begin{aligned}&\mathbf{M}_k\\ &\triangleq\mathbb{E}_{\boldsymbol{d}_k}[(\boldsymbol{d}_k-\mathbf{G}_k^{\mathrm{H}}\boldsymbol{y}_k)(\boldsymbol{d}_k-\mathbf{G}_k^{\mathrm{H}}\boldsymbol{y}_k)^{\mathrm{H}}]\\ &=\mathbf{I}_{N_k}+\mathbf{G}_k^{\mathrm{H}}\mathbf{G}_k+\mathbf{G}_k^{\mathrm{H}}\mathbf{H}_k\mathbf{T}\mathbf{T}^{\mathrm{H}}\mathbf{H}_k^{\mathrm{H}}\mathbf{G}_k-\mathbf{G}_k^{\mathrm{H}}\mathbf{H}_k\mathbf{T}_k-\mathbf{T}_k^{\mathrm{H}}\mathbf{H}_k^{\mathrm{H}}\mathbf{G}_k.\end{aligned} \tag{6}$$

Under classical Gaussian signaling, the rate of user k can be written as

$$R_k\triangleq\log_2\left|\mathbf{M}_k^{-1}\right|,\qquad \forall k\in\{1,\ldots,K\}. \tag{7}$$

Finally, we define the average sum rate $\mathbb{E}\left[R\right]$ as

$$\mathbb{E}\left[R\right]\triangleq\sum_{k=1}^{K}\mathbb{E}\left[R_k\right]. \tag{8}$$

B. Distributed CSIT

In the D-CSIT setting, each TX receives its own copy of noisy CSI based on which it designs its transmission parameters without any additional communication to the other TXs.

Specifically, TX j receives the channel estimate $\hat{\mathbf{H}}^{(j)}\in\mathbb{C}^{N_{\text{tot}}\times M_{\text{tot}}}$ and designs its transmit coefficient $\mathbf{x}_j\in\mathbb{C}^{M_j\times 1}$ as a function of $\hat{\mathbf{H}}^{(j)}$, *without any form of information exchange with the other TXs*. Note that this model is in fact very general as it allows for any joint distribution $p_{\mathbf{H},\hat{\mathbf{H}}^{(1)},\ldots,\hat{\mathbf{H}}^{(K)}}$ and thus any level of correlation between the channel estimates at the TXs. Interestingly, our model encompasses various scenarios of interest as particular cases: Partial CSIT, perfect CSIT, and the classical noisy centralized CSIT.

C. Distributed Precoding

Let us assume a cooperation (team) regime where TX j aims at maximizing the sum rate for each channel realization[2]. The key challenge lies in that TX j does not know the transmit coefficients used at the other TXs. In the Bayesian approach, TX j then needs to account for all possible precoder design decisions at other TXs and average over their distribution.

In principle this problem of finding a good precoding policy can be carried out off line via a *functional optimization*. The precoding function of TX j is denoted by

$$\boldsymbol{w}_j:\mathbb{C}^{N_{\text{tot}}\times M_{\text{tot}}}\rightarrow\mathbb{C}^{M_j\times d_{\text{tot}}} \tag{9}$$

such that the transmit signal $\mathbf{x}_j$ at TX j for a given channel realization $\hat{\mathbf{H}}^{(j)}$ is equal to

$$\mathbf{x}_j=\boldsymbol{w}_j(\hat{\mathbf{H}}^{(j)})\boldsymbol{s}. \tag{10}$$

Upon concatenation of all TX's precoding decisions, the global precoder used for the transmission for a given channel realization is equal to

$$\mathbf{T}\triangleq\begin{bmatrix}\boldsymbol{w}_1(\hat{\mathbf{H}}^{(1)})\\ \boldsymbol{w}_2(\hat{\mathbf{H}}^{(2)})\\ \vdots\\ \boldsymbol{w}_K(\hat{\mathbf{H}}^{(K)})\end{bmatrix}. \tag{11}$$

D. Unquantized Team Precoding

With D-CSIT, the TD problem of joint precoding can be written as the following optimization problem:

$$(\boldsymbol{w}_1^\star,\ldots,\boldsymbol{w}_K^\star)=\underset{(\boldsymbol{w}_1,\ldots,\boldsymbol{w}_K)\in\mathcal{W}}{\operatorname{argmax}}\ \mathbb{E}[\mathrm{R}(\boldsymbol{w}_1(\hat{\mathbf{H}}^{(1)}),\ldots,\boldsymbol{w}_K(\hat{\mathbf{H}}^{(K)}))] \tag{12}$$

where $\mathcal{W}$ is defined as

$$\begin{aligned}\mathcal{W}\triangleq\{(\boldsymbol{w}_1,\ldots,\boldsymbol{w}_K)|\boldsymbol{w}_j:\mathbb{C}^{N_{\text{tot}}\times M_{\text{tot}}}\rightarrow\mathbb{C}^{M_j\times d_{\text{tot}}},\\ \forall\mathbf{X}\in\mathbb{C}^{N_{\text{tot}}\times M_{\text{tot}}},\|\boldsymbol{w}_j(\mathbf{X})\|^2\leq P_j,\forall j\}.\end{aligned} \tag{13}$$

Finding the solution of a TD problem such as (12) is often out of reach. Inspired from the game-theoretic literature, it is however possible to define *best-response* strategies, which in our setting are written as follows.

Definition 1. *A* best-response *precoding function $(\boldsymbol{w}_1^{BR},\ldots,\boldsymbol{w}_K^{BR})$ for the team decision problem* (12) *is a precoding function satisfying*

$$\boldsymbol{w}_j^{BR}=\underset{\boldsymbol{w}_j\in\mathcal{W}_j}{\operatorname{argmax}}\,\mathbb{E}[\mathrm{R}(\boldsymbol{w}_j,\bar{\boldsymbol{w}}_j^{BR})],\qquad\forall j\in\{1,\ldots,K\} \tag{14}$$

[2]Hence, there is no conflict between TX. Consequently, we depart from classical game theoretic approaches to rate maximizing coordination.

where we have used the short-hand notation $(\boldsymbol{w}_j, \bar{\boldsymbol{w}}_j^{BR})$ to replace $(\boldsymbol{w}_1^{BR}, \ldots, \boldsymbol{w}_{j-1}^{BR}, \boldsymbol{w}_j, \boldsymbol{w}_{j+1}^{BR}, \ldots, \boldsymbol{w}_K^{BR})$ and $\mathcal{W}_j$ is the space of possible values of $\boldsymbol{w}_j$ given $\bar{\boldsymbol{w}}_j^{BR}$.

A best-response precoding function corresponds to a TD precoding function where each TX transmits optimally *given the precoding functions of the other TXs*. Note that solving for $\boldsymbol{w}_j$ in (14) remains difficult for a continuous CSI observation space as it requires optimizing over an infinite dimensional space. Consequently, we reduce the dimensionality of the optimized space by quantizing the CSI space.

III. Quantized Team Precoding

A. Dimensionality Reduction by Discretization

The main difficulty of the optimization problem (14) comes from the optimization of each TX precoding function being carried out over a *functional space*, which is hence of infinite dimension. This makes it especially difficult to obtain any iterative solutions. To circumvent this difficulty, we reduce the dimensionality of the search space by projecting the channel state space over a finite codebook, thus reducing the search space from a space of infinite dimension to a space of finite dimensionality.

For ease of presentation, we will use the same codebook at every TX but the extension to codebooks of different sizes presents no difficulty. Specifically, let us denote the codebook used at each TX by $\mathcal{Y}$, and assume that it contains n instances of the multi-user channel state $\mathbf{H}$, i.e.,

$$\mathcal{Y} \triangleq \{\mathbf{H}_i | \mathbf{H}_i \in \mathbb{C}^{N^{\text{tot}} \times M^{\text{tot}}}, i = 1, \ldots, n\}. \tag{15}$$

The optimization of the codebook design , although an interesting research topic, is out of the scope of this work. We use hence a random codebook obtained from n random Monte-Carlo realizations according to $p_{\mathbf{H}}$ as it is known to be an efficient solution for large codebooks [5], [15]. We then denote by $\mathcal{Q}$ the quantizer which maps an estimate $\hat{\mathbf{H}}^{(j)}$ received at TX j to an index corresponding to an element in $\mathcal{Y}$. We chose here the Grassmannian quantization [15] such that:

$$\mathcal{Q}(\hat{\mathbf{H}}) \triangleq \operatorname*{argmax}_{i \in \{1,\ldots,n\}} \left| \frac{\text{vect}(\mathbf{H}_i)^{\text{H}}}{\|\text{vect}(\mathbf{H}_i)\|} \cdot \frac{\text{vect}(\hat{\mathbf{H}})}{\|\text{vect}(\hat{\mathbf{H}})\|} \right|. \tag{16}$$

The Grassmannian quantization corresponds to the usual choice in the MIMO literature and achieves the optimal scaling in the high precision regime [5], [15].

Following this quantization step at each TX, the TD optimization problem (12) is reduced to

$$(\boldsymbol{w}_1^{\star}, \ldots, \boldsymbol{w}_K^{\star}) = \operatorname*{argmax}_{(\boldsymbol{w}_1,\ldots,\boldsymbol{w}_K) \in \mathcal{W}^n} \mathbb{E}[\text{R}(\boldsymbol{w}_1(\mathcal{Q}(\hat{\mathbf{H}}^{(1)})), \ldots, \boldsymbol{w}_K(\mathcal{Q}(\hat{\mathbf{H}}^{(K)})))] \tag{17}$$

where we have defined $\mathcal{W}^n$ as

$$\mathcal{W}^n \triangleq \{(\boldsymbol{w}_1^n, \ldots, \boldsymbol{w}_K^n) | \boldsymbol{w}_j^n : \boldsymbol{w}_j^n : \{1, \ldots, n\} \rightarrow \mathbb{C}^{M_j \times d^{\text{tot}}}, \forall i \in \{1, \ldots, n\}, \|\boldsymbol{w}_j(i)\|^2 \leq P_j, \forall j\}. \tag{18}$$

B. Best-Response Optimization

Coming back to the best-response optimization problem, we consider the best-response optimization at TX j with the precoding functions at the other TXs being fixed equal to $\bar{\boldsymbol{w}}_j^{\text{BR}}$. The quantized best-response optimization is mathematically written as, $\forall i \in \{1, \ldots, n\}$,

$$\boldsymbol{w}_j^{\text{BR}}(i) = \operatorname*{argmax}_{\mathbf{W}_j \in \mathbb{C}^{M_j \times d_{\text{tot}}}, \|\mathbf{W}_j\|^2 \leq P_j} \mathbb{E}\left[\text{R}(\mathbf{W}_j, \bar{\boldsymbol{w}}_j^{\text{BR}}) | \mathcal{Q}(\hat{\mathbf{H}}^{(j)}) = \mathbf{H}_i\right]. \tag{19}$$

C. Stochastic Optimization using Monte-Carlo Sampling

Since the precoding functions at the other TXs are fixed to $\boldsymbol{w}_j^{\text{BR}}$, the expectation in (19) is numerically tractable. It consists simply in finding the complex matrix $\mathbf{W}_j \in \mathbb{C}^{M_j \times d^{\text{tot}}}$ which maximizes the objective in (19). This optimization problem falls in the category of so-called stochastic optimization problems for which a vast literature is available (see for example [16] and references therein). For instance, Monte-Carlo approximations with n^{MC} trials can be used to approximate the expectation operator. The best-response at TX j is then approximated as

$$\boldsymbol{w}_j^{\text{BR}}(i) = \operatorname*{argmax}_{\mathbf{W}_j \in \mathbb{C}^{M_j \times d_{\text{tot}}}, \|\mathbf{W}_j\|^2 \leq P_j} \frac{1}{n^{\text{MC}}} \sum_{\ell=1}^{n^{\text{MC}}} \text{R}(\mathbf{W}_j, \bar{\boldsymbol{w}}_j^{\text{BR}}(\mathcal{Q}(\bar{\mathbf{H}}^j(\ell)))). \tag{20}$$

with the samples $(\mathbf{H}, \hat{\mathbf{H}}^{(1)}, \ldots, \hat{\mathbf{H}}^{(j-1)}, \hat{\mathbf{H}}^{(j+1)}, \ldots, \hat{\mathbf{H}}^{(K)})$ being drawn in an i.i.d. manner from the conditional joint distribution $p_{\mathbf{H}, \hat{\mathbf{H}}^{(1)}, \ldots, \hat{\mathbf{H}}^{(K)} | \hat{\mathbf{H}}^{(j)} = \mathbf{H}_i}$.

At this step, the initial TD optimization problem has been replaced by a conventional deterministic optimization problem. However, depending on the structure of the objective, this optimization can still be difficult. In particular, the sum rate objective is non-convex and even its deterministic counterpart is a challenging optimization problem [4]. Therefore, the last step of our precoding algorithm consists in using a convex approximation for the optimization. Indeed, a convex function has the advantage that the sum of convex function (or the expectation) of a convex function remains a convex function. Therefore, if the initial objective is convex, optimization problem (20) will also be a convex problem. It will then be possible to use a convex solver to efficiently obtain a solution. In fact, we will show that it is then even possible to design a specific algorithm converging to a local optimum.

D. Update of the Precoder using Convex Approximation

Our convex approximation relies on the algorithm given in [4] and then successfully used in many other settings [14], [17]–[19]. Following this approach, the instantaneous sum rate defined in (7) is rewritten as

$$R = \max_{\mathbf{G}_k, \mathbf{\Omega}_k} \sum_{k=1}^{K} \log_2 |\mathbf{\Omega}_k| - \text{tr}\left(\mathbf{\Omega}_k \mathbf{M}_k\right) + d_k \tag{21}$$

with the matrix $\mathbf{M}_k$ being defined in (6) and $\mathbf{\Omega}_k \in \mathbb{C}^{d_k \times d_k}$ being a weighting matrix left to be optimized. The optimization problem (21) is convex in each of the optimization variable. Setting the derivative to zero, it is easily shown that their optimal values verify [4], [14]

$$\mathbf{G}_k = \left(\mathbf{H}_k \mathbf{T}\mathbf{T}^{\mathrm{H}}\mathbf{H}_k^{\mathrm{H}} + \mathbf{I}_{N_k}\right)^{-1} \mathbf{H}_k \mathbf{T}_k, \qquad \forall k, \tag{22}$$

and

$$\mathbf{\Omega}_k = \mathbf{M}_k^{-1}, \qquad \forall k. \tag{23}$$

We also define for ease of use the matrices

$$\mathbf{G} \triangleq \operatorname{diag}\left(\mathbf{G}_1, \dots, \mathbf{G}_K\right), \quad \mathbf{\Omega} \triangleq \operatorname{diag}\left(\mathbf{\Omega}_1, \dots, \mathbf{\Omega}_K\right). \tag{24}$$

It can easily be verified that inserting the optimal values obtained in (22) and in (23) inside (21) gives the sum rate R. This reformulation of the objective has the advantage that a local optimal is easily obtained by iteratively updating the optimization variables.

Indeed, coming back to the optimization of $\boldsymbol{w}_j^{\mathrm{BR}}(i)$ in (20), let us consider the RX filters and the weighting matrices for each Monte-Carlo realization to be fixed, and let us denote by $\mathbf{G}(\ell)$ and $\mathbf{\Omega}(\ell)$ their values for the ℓ-th Monte-Carlo realization. The optimization problem can then be seen to be convex in the precoding coefficients $\mathbf{W}_j$ at TX j.

As a preliminary step for writting the optimal update for $\boldsymbol{w}_j^{\mathrm{BR}}(i)$, we need to introduce some notations for the ℓ-th Monte-Carlo realization: We denote by $\mathbf{H}(\ell)$ the multi-user channel and by $\hat{\mathbf{H}}^{(j)}(\ell)$ the estimate at TX j, by $\mathbf{H}_k(\ell)$ the channel from all TXs to RX k, by $\mathbf{K}_j(\ell)$ the channel from TX j to all the RXs and by $\bar{\mathbf{K}}_j(\ell)$ the channel from all the TXs at the exception of TX j to all the RXs, and finally by $\bar{\boldsymbol{w}}_j(\boldsymbol{i}_\ell)$ the precoding decision at all TXs except TX j. Using these notations, the optimal precoder is given by (25) at the top of next page, where λ_j is the Lagrangian multiplier associated with the power constraint at TX j and is obtained by bisection. Updating iteratively the RX filters $\mathbf{G}(\ell)$ with (22) and the weighting matrices $\mathbf{\Omega}(\ell)$ with (23) for each Monte-Carlo realization, and the precoder at each TX according to (25), the algorithm converges to a local maximum since the objective is increased at each step. For clarity, all the steps of the algorithm are put together in Algorithm 1 in the case of $K = 2$ TX/RX pairs.

Algorithm 1 DT Precoding for $K = 2$ TXs

- Input: $\mathcal{Y}$,$\boldsymbol{w}_1^0 \in (\mathbb{C}^{M_1 \times d_1})^n$ and $\boldsymbol{w}_2^0 \in (\mathbb{C}^{M_2 \times d_2})^n$
- Initialization: Set $\boldsymbol{w}_1^{\mathrm{BR}} = \boldsymbol{w}_1^0$, and $\boldsymbol{w}_2^{\mathrm{BR}} = \boldsymbol{w}_2^0$
- Until convergence:
  - At TX 1 for each $i \in \{1, \dots, n\}$
    - $\forall \ell \in \{1, \dots, n^{\mathrm{MC}}\}$, set

      $$\mathbf{T}(\ell) = \begin{bmatrix} \boldsymbol{w}_1^{\mathrm{BR}}(\mathcal{Q}(\hat{\mathbf{H}}^{(1)}(\ell))) \\ \boldsymbol{w}_2^{\mathrm{BR}}(\mathcal{Q}(\hat{\mathbf{H}}^{(2)}(\ell))) \end{bmatrix}$$
    - $\forall \ell \in \{1, \dots, n^{\mathrm{MC}}\}$, update $\mathbf{G}(\ell)$ such that $\forall k$

      $$\mathbf{G}_k(\ell) = \left(\mathbf{H}_k(\ell)\mathbf{T}(\ell)\mathbf{T}^{\mathrm{H}}(\ell)\mathbf{H}_k^{\mathrm{H}}(\ell) + \mathbf{I}_{N_k}\right)^{-1} \cdot \mathbf{H}_k(\ell)\mathbf{T}_k(\ell),$$
    - $\forall \ell \in \{1, \dots, n^{\mathrm{MC}}\}$, update $\mathbf{\Omega}(\ell)$ such that $\forall k$

      $$\begin{aligned} \mathbf{\Omega}_k^{-1}(\ell) &= \mathbf{I}_{N_k} + \mathbf{G}_k^{\mathrm{H}}(\ell)\mathbf{G}_k(\ell) \\ &+ \mathbf{G}_k^{\mathrm{H}}(\ell)\mathbf{H}_k(\ell)\mathbf{T}(\ell)\mathbf{T}^{\mathrm{H}}(\ell)\mathbf{H}_k^{\mathrm{H}}(\ell)\mathbf{G}_k(\ell) \\ &- \mathbf{G}_k^{\mathrm{H}}(\ell)\mathbf{H}_k(\ell)\mathbf{T}_k(\ell) - \mathbf{T}_k^{\mathrm{H}}(\ell)\mathbf{H}_k^{\mathrm{H}}(\ell)\mathbf{G}_k(\ell) \end{aligned}$$
    - Update $\boldsymbol{w}_1^{\mathrm{BR}}(i)$ according to (25).
  - Proceed symmetrically at TX 2 to obtain $\boldsymbol{w}_2^{\mathrm{BR}}$

IV. Simulations Results

In this section, we evaluate through Monte-Carlo simulations the efficiency of the proposed algorithm. We consider the most simple configuration with $K = 2$ TXs and all the nodes having a single-antenna. We use also a simple imperfect CSIT model where the estimate $\hat{\mathbf{H}}^{(j)}$ at TX j is given by

$$\hat{\mathbf{H}}^{(j)} = \sqrt{1 - \sigma_j^2}\mathbf{H} + \sigma_j \mathbf{\Delta}_j \tag{26}$$

with $\mathbf{\Delta}_j \in \mathbb{R}^{2\times 2}$ having its elements distributed as $\mathcal{N}_{\mathbb{C}}(0,1)$ while $\mathbf{H}$ is a Rayleigh fading channel and has its elements distributed as $\mathcal{N}_{\mathbb{C}}(0,1)$. To evaluate the efficiency of our novel DT precoding scheme, we compare its performance with the upper bound obtained in the case where both TXs have access to the perfect instantaneous CSI and use the sum-rate maximization algorithm from [4]. We also compare our precoding scheme to the conventional distributed precoding approach where each TX designs its precoder using the robust sum-rate maximization algorithm from [19], i.e., each TX uses only the estimate locally available and does not take into account the distributed aspect of the CSI configuration. We call hence this scheme *conventional robust* precoding.

We use Algorithm 1 with a codebook of $n = 10000$ elements and $n^{\mathrm{MC}} = 100$ Monte-Carlo realizations. We also use as initialization of the algorithm the precoder using conventional robust precoding.

In Fig. 1, we show the ergodic rate achieved for the CSIT qualities $\sigma_1^2 = 0.5$ and $\sigma_2^2 = 0.1$. It can be seen that our approach outperforms the conventional robust precoding at any SNR value. DT precoding performs well at low to medium SNR and, in contrast to conventional robust precoding, is able to achieve a positive DoF by serving only one user at high SNR. DT precoding suffers at high SNR from a degradation of the performance due to the quantization noise. Reducing this quantization noise requires to use larger codebooks or to modify the algorithm so as to reduce the impact of this quantization noise.

V. Discussion and Future Works

The proposed Quantized Team Precoding scheme allows for a versatile robust precoding scheme for distributed CSI which can be applied in any antenna, fading, and CSI configuration. The precoder design relies on a functional approximation where the precoding function is approximated by discretization of the channel space estimate.

$$
\boldsymbol{w}_j^{\mathrm{BR}}(i)=\left(\frac{1}{n^{\mathrm{MC}}}\sum_{\ell=1}^{\mathrm{MC}}\mathbf{K}_j^{\mathrm{H}}(\ell)\mathbf{G}(\ell)\mathbf{\Omega}(\ell)\mathbf{G}^{\mathrm{H}}(\ell)\mathbf{K}_j(\ell)+\lambda_j\mathbf{I}_{M_j}\right)^{-1}\left(\frac{1}{n^{\mathrm{MC}}}\sum_{\ell=1}^{\mathrm{MC}}\mathbf{K}_j^{\mathrm{H}}(\ell)\mathbf{G}(\ell)\mathbf{\Omega}(\ell)\left(\mathbf{I}_{d_{\mathrm{tot}}}-\mathbf{G}^{\mathrm{H}}(\ell)\bar{\mathbf{K}}_j(\ell)\bar{\boldsymbol{w}}_j^{\mathrm{BR}}(\boldsymbol{i}_\ell)\right)\right)
\tag{25}
$$

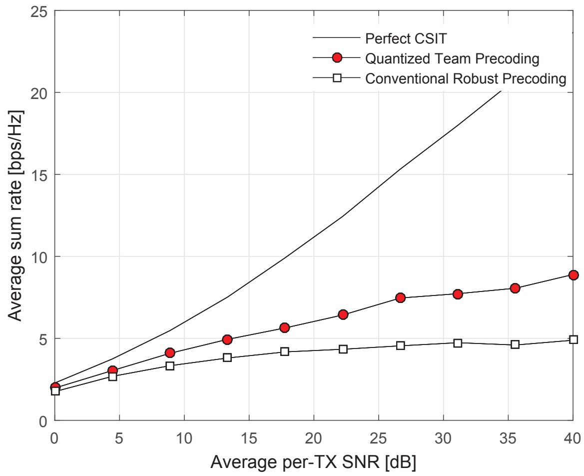

Fig. 1: Average sum rate as a function of the per-TX power constraint.

As the complexity of the algorithm increases very quickly with the dimensionality of the channel estimate space, we give a preliminary study based on Monte-Carlo sampling. Optimizing the sampling along the optimization is a very promising direction of research, also investigated in relation to other optimization and learning problems. Finally, codebook design is not considered here and could be investigated in the future as an interesting parameter to trade-off complexity and accuracy.

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] B. Dai and W. Yu, "Backhaul-aware multicell beamforming for downlink cloud radio access network," in *Proc. IEEE International Conference on Communications (ICC)*, 2015.

[3] 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.

[4] S. S. Christensen, R. Agarwal, E. Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. on Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.

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

[6] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2845–2866, Jun. 2010.

[7] D. J. Love, R. W. Heath, V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1341–1365, Oct. 2008.

[8] N. Vucic, H. Boche, and S. Shi, "Robust transceiver optimization in downlink multiuser MIMO systems," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3576–3587, Sep. 2009.

[9] M. B. Shenouda and T. N. Davidson, "Convex conic formulations of robust downlink precoder designs with quality of service Constraints," *IEEE Journal of Sel. Topics in Sign. Process.*, vol. 1, no. 4, pp. 714–724, 2007.

[10] P. de Kerret, "Transmitter cooperation with distributed feedback in wireless networks," Ph.D. dissertation, TELECOM ParisTech, 2013. [Online]. Available: http://tel.archives-ouvertes.fr/tel-00952820

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

[12] 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.

[13] R. Zakhour and D. Gesbert, "Team decision for the cooperative MIMO channel with imperfect CSIT sharing," in *Proc. Information Theory and Applications Workshop (ITA)*, 2010.

[14] 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.

[15] W. Dai, Y. Liu, and B. Rider, "Quantization bounds on Grassmann manifolds and applications to MIMO communications," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1108–1123, Mar. 2008.

[16] P. Kall and S. W. Wallace, *Stochastic Programming*. John Wiley & Sons, Chichester, 1994.

[17] F. Negro, I. Ghauri, and D. T. Slock, "Deterministic annealing design and analysis of the Noisy MIMO Interference Channel," in *Proc. Information Theory and Applications Workshop (ITA)*, 2011.

[18] M. Razaviyayn, M. S. Boroujeni, and Z.-Q. Luo, "A stochastic weighted MMSE approach to sum rate maximization for a MIMO interference channel," in *Proc. IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2013.

[19] R. Fritzsche and G. P. Fettweis, "Robust sum rate maximization in the multi-cell MU-MIMO downlink," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, 2013.