The Multiplexing Gain of a Two-cell MIMO Channel with Unequal CSI

Paul de Kerret and David Gesbert
Mobile Communications Department, Eurecom
2229 route des Crêtes, 06560 Sophia Antipolis, France
{dekerret,gesbert}@eurecom.fr

***Abstract*—In this work[1], the joint precoding across two distant transmitters (TXs), sharing the knowledge of the data symbols to be transmitted, to two receivers (RXs), each equipped with one antenna, is discussed. We consider a distributed channel state information (CSI) configuration where each TX has its own local estimate of the channel and no communication is possible between the TXs. Based on the distributed CSI configuration, we introduce a concept of distributed MIMO precoding. We focus on the high signal-to-noise ratio (SNR) regime such that the two TXs aim at designing a precoding matrix to cancel the interference. Building on the study of the multiple antenna broadcast channel, we obtain the following key results: We derive the multiplexing gain (MG) as a function of the scaling in the SNR of the number of bits quantizing at each TX the channel to a given RX. Particularly, we show that the conventional Zero Forcing precoder is not MG maximizing, and we provide a precoding scheme optimal in terms of MG. Beyond the established MG optimality, simulations show that the proposed precoding schemes achieve better performances at intermediate SNR than known linear precoders.**

I. Introduction

One promising solution to answer the need for increased spectral efficiency in the future wireless networks consists in the joint transmission from several transmitter (TXs) to serve multiple receivers (RXs), so called Network MIMO [1], [2]. If all the TXs have access to the data symbols and to the global channel state information (CSI), the different TXs can then be seen as a unique virtual TX serving all the receivers (RXs). The precoding schemes of the multiple antenna broadcast channel (BC) can then be applied.

Yet, this requires the sharing of the data symbol and the CSI between the TXs, which represents a high requirement on the network infrastructure. Indeed, while in future wireless networks (e.g. LTE Advanced), it is considered to link the TXs with the Core Network via high capacity links to share the data symbols with the cooperating TXs, the sharing of the CSI is done through limited rate feedback channels and limited capacity signaling (so called X2) links between the TXs. Thus, an interesting information theoretic MIMO channel arises whereby multiple TXs may access the same data symbols, but have a limited CSI sharing capability. We define this channel as the *distributed CSI* (DCSI)-MIMO channel.

In the DCSI-MIMO channel, there may be inconsistencies between the different versions of CSI seen at the TXs due either to separate compression or separate feedback channels. Such inconsistencies can be detrimental to the channel capacity if they are not accounted for in the precoding design. This is the object of this work.

To put this in contrast, note that in the BC, the impact of finite rate feedback [3]–[6] and the derivation of robust solutions [7], [8] have been the focus of many works, which have been then extended to the MIMO network setting [9], [10]. However, these works only focus on the case of imperfect CSI yet *perfectly shared between the TXs* and do not consider the case when each TX has its own imperfect estimation of the multi-user channel, which will be our focus in this work. This setting was first studied in [11], and a tractable discrete optimization at finite SNR was derived. However, it does not lend itself to a more general performance analysis.

Our work can be seen as a generalization to the case of distributed CSI setting of the study by Jindal [3] of the multiple-antenna BC, in which the Multiplexing Gain (MG) is derived as a function of the number of feedback bits by each RX. We here consider only two TX-RX pairs, while the generalization to multiple TX-RX pairs is carried out in [12]. We consider only *Zero-Forcing* schemes which are known to achieve the maximal MG with perfect CSI in the MIMO BC.

Specifically, the main contributions are as follows. Let's first define the number of bits quantizing the estimate at TX j of the normalized channel $\tilde{\boldsymbol{h}}_i^{\mathrm{H}}$ from the two TXs to RX i as $\alpha_i^{(j)} \log_2(P)$ with $\alpha_i^{(j)} \in [0,1]$. Then, we show that:

- The MG achieved with conventional Zero Forcing at RX i is equal to $\min_{i,j\in\{1,2\}} \alpha_i^{(j)}$.
- The optimal MG at RX i is equal to $\max_{j\in\{1,2\}} \alpha_i^{(j)}$.
- We provide a precoding scheme achieving the maximal MG, as well as practical precoding schemes outperforming known linear precoding schemes at finite SNR for the DCSI-MIMO channel.

Notations: We denote by $\Pi_{\boldsymbol{u}}(\bullet)$ and $\Pi_{\boldsymbol{u}}^{\perp}(\bullet)$ the orthogonal projectors over the subspace spanned by $\boldsymbol{u}$ and over its orthogonal complement, respectively. $\bar{i}$ denotes the complementary indice of i, i.e., $\bar{i} = i \mod 2+1$.

II. System Model

We first present the classical multicell MIMO model before introducing our novel concepts of *distributed CSI* and *distributed precoding*.

[1]This work has been performed in the framework of the European research project ARTIST4G, which is partly funded by the European Union under its FP7 ICT Objective 1.1 - The Network of the Future.

A. Multicell MIMO

We consider a joint downlink transmission from two TXs to two RXs using linear precoding and single user decoding. For ease of exposition, the TXs and the RXs are equipped with only one antenna, such that the received signal is written as

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{H}\boldsymbol{x} + \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{h}_1^{\mathrm{H}}\boldsymbol{x} \\ \boldsymbol{h}_2^{\mathrm{H}}\boldsymbol{x} \end{bmatrix} + \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} = \begin{bmatrix} \|\boldsymbol{h}_1^{\mathrm{H}}\|\tilde{\boldsymbol{h}}_1^{\mathrm{H}}\boldsymbol{x} \\ \|\boldsymbol{h}_2^{\mathrm{H}}\|\tilde{\boldsymbol{h}}_2^{\mathrm{H}}\boldsymbol{x} \end{bmatrix} + \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} \quad (1)$$

where y_i is the signal received at the i-th RX, $\boldsymbol{h}_i^{\mathrm{H}} \in \mathbb{C}^{1\times 2}$ is the channel from the TXs to the i-th RX, $\tilde{\boldsymbol{h}}_i^{\mathrm{H}} \triangleq \boldsymbol{h}_i^{\mathrm{H}}/\|\boldsymbol{h}_i^{\mathrm{H}}\|$ is the normalized channel, $\eta_i \sim \mathcal{CN}(0,1)$ is the noise at the i-th RX and is distributed as i.i.d. complex circularly symmetric Gaussian noise, and $\boldsymbol{x} \in \mathbb{C}^{2\times 1}$ is the transmitted signal from the TXs. The channel is block fading and the entries of the channel matrix $\mathbf{H}$ are distributed as i.i.d. complex circularly symmetric Gaussian with unit variance to model a Rayleigh fading channel. The transmitted signal $\boldsymbol{x}$ is obtained from the vector of transmit symbol $\boldsymbol{s} = [s_1, s_2]^{\mathrm{T}} \in \mathbb{C}^{2\times 1}$ (whose entries are assumed to be independent $\mathcal{CN}(0,1)$) as

$$\boldsymbol{x} = \mathbf{T}\boldsymbol{s} = \begin{bmatrix} \boldsymbol{t}_1 & \boldsymbol{t}_2 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} \quad (2)$$

where $\mathbf{T} \in \mathbb{C}^{2\times 2}$ and $\boldsymbol{t}_i \in \mathbb{C}^{2\times 1}$ is the beamforming vector used to transmit s_i. We consider a sum power constraint $\|\mathbf{T}\|_{\mathrm{F}}^2 = P$ and we also assume for simplicity and symmetry that $\boldsymbol{t}_i = \sqrt{P/2}\boldsymbol{u}_i$ with $\|\boldsymbol{u}_i\|_2^2 = 1$. Note that normalizing the individual columns does not alter the ability to zero-force the interference so that it does not affect the MG.

We also define the MG at RX i as

$$\mathrm{M}_{\mathrm{G}i} \triangleq \lim_{P\to\infty} \frac{R_i(P)}{\log_2(P)} \quad (3)$$

so that the total MG is $\mathrm{M}_{\mathrm{G}} \triangleq \mathrm{M}_{\mathrm{G}1} + \mathrm{M}_{\mathrm{G}2}$.

We will study the long-term average throughput over the fading distribution and also over the realizations of the Random Vector Quantization (RVQ) codebooks used for the CSI quantization (Cf. subsection II-C), such that the throughput for RX i reads as

$$R_i(P) \triangleq \mathrm{E}_{\mathbf{H},\mathcal{W}}\left[\log_2\left(1 + \frac{|\boldsymbol{h}_i^{\mathrm{H}}\boldsymbol{t}_i|^2}{1 + |\boldsymbol{h}_i^{\mathrm{H}}\boldsymbol{t}_{\bar{i}}|^2}\right)\right] \quad (4)$$

To achieve the maximal MG we aim at removing all the interference, i.e., at having

$$\mathcal{I}_1(\boldsymbol{t}_2) \triangleq |\boldsymbol{h}_1^{\mathrm{H}}\boldsymbol{t}_2|^2 = 0, \text{ and } \mathcal{I}_2(\boldsymbol{t}_1) \triangleq |\boldsymbol{h}_2^{\mathrm{H}}\boldsymbol{t}_1|^2 = 0. \quad (5)$$

From (5), it follows that the optimization of the two beamforming vectors $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$ can be uncoupled.

B. Distributed CSI

We assume a limited CSI setting where finite quality channel estimates are obtained from quantizing the true channel vectors. The *distributed* CSI is defined here in the sense that each TX has a different estimate of the normalized channel $\tilde{\boldsymbol{h}}_i$ from all TXs to RX i. Moreover, the estimates for $\tilde{\boldsymbol{h}}_1$ and $\tilde{\boldsymbol{h}}_2$ are also a priori of statistically different qualities. We denote by $\tilde{\boldsymbol{h}}_i^{(j)}$ the estimate of the normalized channel vector $\tilde{\boldsymbol{h}}_i$ acquired at TX j. Furthermore, the number of quantizing bits for $\tilde{\boldsymbol{h}}_i^{(j)}$ is given by $B_i^{(j)}$.

In the context of MIMO BC, it is shown in [3] that the number of quantization bits should scale indefinitely with the SNR in order to achieve a positive MG with ZF. It also holds in a distributed CSI configuration so that we focus on the *scaling in the logarithm of the SNR* of the number of quantization bits

$$\alpha_i^{(j)} \triangleq \lim_{P\to\infty} \frac{B_i^{(j)}}{\log_2(P)}. \quad (6)$$

Since $\alpha_i^{(j)} = 1, \forall i,j \in \{1,2\}$ is shown later in Theorem 1 to be sufficient to achieve the maximal MG, we will always consider that $\alpha_i^{(j)} \in [0,1]$.

C. Random Vector Quantization

We consider the performances averaged over codebooks used to quantize the channels randomly generated. This follows a result in [3] stating that in the case of two antennas at the TX, no codebook can achieve in the single TX case a better MG than the MG achieved with RVQ.

However, in the MIMO BC, a codeword $\boldsymbol{c}$ is selected to quantize $\boldsymbol{h}$ if it maximizes the inner product $|\boldsymbol{h}^{\mathrm{H}}\boldsymbol{c}|$ over the codebook. Any other codeword of the form $\boldsymbol{c}e^{j\phi}$ where ϕ is any real number achieves the same performances and can be selected indifferently. This is problematic in a distributed setting since we are now interested in $\|\tilde{\boldsymbol{h}}_i^{(1)} - \tilde{\boldsymbol{h}}_i^{(2)}\|$ and even if the codewords at TX1 and TX2 are $e^{j\phi_1}\tilde{\boldsymbol{h}}_i$ and $e^{j\phi_2}\tilde{\boldsymbol{h}}_i$ respectively, i.e., exactly in the direction of $\tilde{\boldsymbol{h}}_i$, the two estimates differ greatly in norm.

Our solution is for each codeword and each channel estimate to choose $e^{j\phi}$ as the complex conjugate of the first vector element divided by its absolute value, thus making the first vector element real valued. Because of this choise, the quantization scheme is not any longer in the Grassmann manifold and we have to consider the isomorphisme between $\mathbb{C}$ and $\mathbb{R}^2$. Thus, for the quantization, each complex vector is considered as a vector of $\mathbb{R}^4$ made of the stacked real and imaginary parts. Moreover, since the first coefficient is real valued only, we have to consider in fact $\mathbb{R}^3$ only. A vector $\boldsymbol{u} \in \mathbb{C}^2$ with is first coefficient real valued is represented in $\mathbb{R}^3$ as $\boldsymbol{u}_{\mathbb{R}^3}$ and is defined as

$$\boldsymbol{u}_{\mathbb{R}^3} \triangleq \begin{bmatrix} \mathrm{Re}(u_1) \\ \mathrm{Re}(u_2) \\ \mathrm{Im}(u_2) \end{bmatrix} \quad (7)$$

Thus, we define the angles between $\boldsymbol{u}_{\mathbb{R}^3}$ and $\boldsymbol{v}_{\mathbb{R}^3}$ in $\mathbb{R}^3$ as

$$\angle(\boldsymbol{u}_{\mathbb{R}^3}, \boldsymbol{v}_{\mathbb{R}^3}) = \arccos\left(\frac{|\boldsymbol{u}_{\mathbb{R}^3}^{\mathrm{T}}\boldsymbol{v}_{\mathbb{R}^3}|}{\|\boldsymbol{u}_{\mathbb{R}^3}\|\|\boldsymbol{v}_{\mathbb{R}^3}\|}\right). \quad (8)$$

Finally, the estimate $\tilde{\boldsymbol{h}}_i^{(j)}$ is chosen as the element of the random codebook $\mathcal{C}$ which maximizes the cosinus of the angle between the codeword and the true channel in $\mathbb{R}^3$:

$$\tilde{\boldsymbol{h}}_{i\mathbb{R}^3}^{(j)} = \operatorname*{argmax}_{\boldsymbol{c}_{\mathbb{R}^3}\in\mathcal{C}_{\mathbb{R}^3}} \cos(\angle(\boldsymbol{c}_{\mathbb{R}^3}, \tilde{\boldsymbol{h}}_{i\mathbb{R}^3})) = |\boldsymbol{c}_{\mathbb{R}^3}^{\mathrm{T}}\tilde{\boldsymbol{h}}_{i\mathbb{R}^3}|. \quad (9)$$

D. Distributed Precoding

In the distributed CSI setting, each TX has a different estimate of the channel, which it uses to compute the precoding matrix. We denote the precoder computed at TX j as

$$\mathbf{T}^{(j)} \triangleq \begin{bmatrix} \boldsymbol{t}_1^{(j)} & \boldsymbol{t}_2^{(j)} \end{bmatrix} \triangleq \begin{bmatrix} T_{11}^{(j)} & T_{12}^{(j)} \\ T_{21}^{(j)} & T_{22}^{(j)} \end{bmatrix}. \tag{10}$$

Note that although a given TX j may compute the whole precoding matrix $\mathbf{T}^{(j)}$, only the j-th row will be used in practice since the other row corresponds to the coefficients being implemented at the other TX. Practically, it means that

$$\mathbf{T} = \begin{bmatrix} T_{11}^{(1)} & T_{12}^{(1)} \\ T_{21}^{(2)} & T_{22}^{(2)} \end{bmatrix}. \tag{11}$$

III. Main Theorems on the Multiplexing Gain

In the multiple antenna BC with perfect CSI, ZF achieves the maximal MG and can be conjectured to be also optimal with imperfect CSI. The central question of this paper is whether this result still holds in the DCSI-MIMO channel, and what are otherwise the MG optimal precoding strategies.

A. Conventional Zero Forcing

The conventional ZF precoder applied distributively consists in transmitting symbol i with the beamformer $\boldsymbol{t}_i^{\mathrm{ZF}} \triangleq [t_{1i}^{\mathrm{ZF}(1)}, t_{2i}^{\mathrm{ZF}(2)}]^{\mathrm{T}}$, with its elements defined as

$$\boldsymbol{t}_i^{\mathrm{ZF}(j)} \triangleq \begin{bmatrix} t_{1i}^{\mathrm{ZF}(j)} \\ t_{2i}^{\mathrm{ZF}(j)} \end{bmatrix} \triangleq \sqrt{\frac{P}{2}} \frac{\Pi^{\perp}_{\tilde{\boldsymbol{h}}_{\bar{i}}^{(j)}} \left(\tilde{\boldsymbol{h}}_i^{(j)}\right)}{\|\Pi^{\perp}_{\tilde{\boldsymbol{h}}_{\bar{i}}^{(j)}} \left(\tilde{\boldsymbol{h}}_i^{(j)}\right)\|}, \quad j \in \{1,2\}. \tag{12}$$

Intuitively, this means that each TX applies ZF with its own CSI implicitely assuming that the other TX is sharing it.

Theorem 1. *Conventional ZF achieves the following MG:*

$${M_G}^{\mathrm{ZF}} = 2 \min_{i,j \in \{1,2\}} \alpha_i^{(j)}. \tag{13}$$

Proof: A detailed proof is given in [12]. ■

Corollary 1. *If the CSI scaling is identical across the RXs and the TXs, i.e.,*

$$\forall i,j,\ell,k \in \{1,2\}, \alpha_i^{(\ell)} = \alpha_k^{(j)} \tag{14}$$

then ZF is optimal. Moreover, the differences in the CSI realizations between the TXs lead then to no loss in MG.

Proof: The result follows from the comparison between the MG in Theorem 1 and in a multiple antenna BC. ■

B. Robust Zero Forcing

To reduce the harmful effect of the imperfect CSI, robust precoding schemes have been derived in the literature either as statistical robust ZF precoder or precoder optimizing the worst case performances [7]. However, the robust versions improve the rate offset but do not have any impact on the MG.

C. Limited Zero Forcing

Comparing the MG in Theorem 1 and in a multiple antenna BC [3], it appears that in the case of imperfectly shared CSI, the MG is limited by the worst quality of the CSI across the channels to the RXs and across the TXs, which is a very pessimistic result. This leads to investigate schemes which are more adapted to this CSI setting, and we now propose a modification of the ZF scheme which improves the MG when the channels $\tilde{\boldsymbol{h}}_1$ and $\tilde{\boldsymbol{h}}_2$ are of different qualities. We call it *limited ZF* (ℓZF). The beamformer used to transmit symbol i is then $\boldsymbol{t}_i^{\ell\mathrm{ZF}} \triangleq [t_{1i}^{\ell\mathrm{ZF}(1)}, t_{2i}^{\ell\mathrm{ZF}(2)}]^{\mathrm{T}}$, with its elements defined as

$$\boldsymbol{t}_i^{\ell\mathrm{ZF}(j)} \triangleq \begin{bmatrix} t_{1i}^{\ell\mathrm{ZF}(j)} \\ t_{2i}^{\ell\mathrm{ZF}(j)} \end{bmatrix} \triangleq \sqrt{\frac{P}{2}} \frac{\Pi^{\perp}_{\tilde{\boldsymbol{h}}_{\bar{i}}^{(j)}} (\boldsymbol{c}_i)}{\|\Pi^{\perp}_{\tilde{\boldsymbol{h}}_{\bar{i}}^{(j)}} (\boldsymbol{c}_i)\|} \tag{15}$$

where $\boldsymbol{c}_i$ is a vector chosen beforehand and known at the TXs. Due to the isotropy of the channel, the choice of $\boldsymbol{c}_i$ is arbitrary and does not influence the performances.

Theorem 2. *The MG achieved with limited ZF is*

$${M_G}^{\ell\mathrm{ZF}} = \min_{j \in \{1,2\}} \alpha_1^{(j)} + \min_{j \in \{1,2\}} \alpha_2^{(j)} \tag{16}$$

and is the maximal MG if $\forall i \in \{1,2\}, \alpha_i^{(1)} = \alpha_i^{(2)}$.

Proof: A detailed proof is given in [12]. ■

D. Cooperative Zero Forcing

We now propose a scheme called *cooperative* Zero Forcing (cZF) which consists in the precoder whose beamformer to transmit symbol i is given by

$$\boldsymbol{t}_i^{\mathrm{cZF}} \triangleq \sqrt{\frac{P(1+\rho_i^{(2)})}{2\log_2(P)}} \boldsymbol{u}_i^{\mathrm{cZF}} \triangleq \sqrt{\frac{P}{2\log_2(P)}} \begin{bmatrix} 1 \\ -\frac{\tilde{h}_{\bar{i}1}^{(2)}}{\tilde{h}_{\bar{i}2}^{(2)}} \end{bmatrix} \tag{17}$$

where $\tilde{\boldsymbol{h}}_{\bar{i}}^{(2)\mathrm{H}} \triangleq [\tilde{h}_{\bar{i}1}^{(2)}, \tilde{h}_{\bar{i}2}^{(2)}]$, $\rho_i^{(2)} \triangleq |\tilde{h}_{\bar{i}1}^{(2)}|^2 / |\tilde{h}_{\bar{i}2}^{(2)}|^2$, $\|\boldsymbol{u}_i^{\mathrm{cZF}}\| = 1$, and we have assumed wlog that $\alpha_{\bar{i}}^{(2)} \geq \alpha_{\bar{i}}^{(1)}$.

Theorem 3. *Cooperative ZF achieves the maximal MG of*

$${M_G}^{\mathrm{cZF}} = \max_{j \in [1,2]} \alpha_1^{(j)} + \max_{j \in [1,2]} \alpha_2^{(j)}. \tag{18}$$

Proof: Due to the symmetry between the two RXs, we consider only the MG at RX 1, and we consider that the beamformers $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$ are given by (17). We assume wlog that $\alpha_1^{(2)} \geq \alpha_1^{(1)}$, i.e., TX 2 has the best CSI over $\tilde{\boldsymbol{h}}_1$. Using cooperative ZF, the MG at RX 1 reads as

$$\mathrm{M}_{\mathrm{G}1} = \lim_{P \to \infty} \frac{\mathrm{E}_{\mathbf{H},\mathcal{W}} \left[\log_2\left(1 + \frac{\|\boldsymbol{h}_1\|^2 \|\boldsymbol{t}_1\|^2 |\tilde{\boldsymbol{h}}_1^{\mathrm{H}} \boldsymbol{u}_1|^2}{\sigma_1^2 + \mathcal{I}_1(\boldsymbol{t}_2)}\right)\right]}{\log_2(P)}$$

$$\mathrm{M}_{\mathrm{G}1} = \lim_{P \to \infty} \frac{\mathrm{E}_{\mathbf{H},\mathcal{W}} \left[\log_2\left(\frac{(\rho_2^{(2)}+1)P}{\log_2(P)}\right) - \log_2(\mathcal{I}_1(\boldsymbol{t}_2))\right]}{\log_2(P)}$$

$$\mathrm{M}_{\mathrm{G}1} = 1 - \lim_{P \to \infty} \frac{\mathrm{E}_{\mathbf{H},\mathcal{W}} [\log_2(\mathcal{I}_1(\boldsymbol{t}_2))]}{\log_2(P)}. \tag{19}$$

We now consider the interference term $\mathcal{I}_1(\boldsymbol{t}_2)$:

$$\mathcal{I}_1(\boldsymbol{t}_2) = |\boldsymbol{h}_1^{\mathrm{H}}\boldsymbol{t}_2|^2 = \frac{P}{2\log_2(P)}\left|\boldsymbol{h}_1^{\mathrm{H}}\begin{bmatrix}1\\ -\frac{\tilde{h}_{11}^{(2)}}{\tilde{h}_{12}^{(2)}}\end{bmatrix}\right|^2. \tag{20}$$

By construction, $\boldsymbol{t}_2$ is orthogonal to $\boldsymbol{h}_1^{(2)}$, so that

$$\mathcal{I}_1(\boldsymbol{t}_2) = \frac{P(1+\rho_2^{(2)})}{2\log_2(P)}\left|\left(\Pi^{\perp}_{\boldsymbol{h}_1^{(2)}}(\boldsymbol{h}_1) + \Pi_{\boldsymbol{h}_1^{(2)}}(\boldsymbol{h}_1)\right)^{\mathrm{H}}\boldsymbol{u}_2\right|^2$$

$$\mathcal{I}_1(\boldsymbol{t}_2) = \frac{P(1+\rho_2^{(2)})}{2\log_2(P)}\|\boldsymbol{h}_1\|^2\left|\Pi^{\perp}_{\boldsymbol{h}_1^{(2)}}(\boldsymbol{h}_1)^{\mathrm{H}}\boldsymbol{u}_2\right|^2. \tag{21}$$

We now define $Z \triangleq \left|\Pi^{\perp}_{\boldsymbol{h}_1^{(2)}}(\boldsymbol{h}_1)^{\mathrm{H}}\boldsymbol{u}_2\right|^2$, the following result is proven in the appendix of [12] using results on the Grassmanian manifolds from [13].

Lemma 1. *It exists two constants such that the expectation of the logarithm of the quantization error is bounded as*

$$\frac{B_1^{(2)}+C_1}{(n-1)} \leq E_{\mathcal{C},\tilde{\boldsymbol{h}}_1}\left[-\log_2(Z)\right] \leq \frac{B_1^{(2)}+C_2}{(n-1)}. \tag{22}$$

Since the difference between the bounds in Lemma 1 is only a constant not scaling in P, clearly the lower and the upper bound for the MG will be the same. Indeed, inserting (21) in the MG expression (19) and using (22), we can write

$$\mathrm{M_{G1}} = \lim_{P\to\infty}\frac{-E_{\mathbf{H},\mathcal{W}}\left[\log_2(|\Pi^{\perp}_{\boldsymbol{h}_1^{(2)}}(\boldsymbol{h}_1)^{\mathrm{H}}\boldsymbol{u}_2|^2)\right]}{\log_2(P)}$$

$$\mathrm{M_{G1}} = \lim_{P\to\infty}\frac{B_1^{(2)}}{\log_2(P)} = \alpha_1^{(2)} \tag{23}$$

which is the best scaling among the TXs. ■

The intuition behind the result is that the only way to achieve the MG from the best CSI accuracy is if TX 2 (which has the best knowledge of $\tilde{\boldsymbol{h}}_1$) can adapt to the transmission at TX 1 to reduce the interference. This is possible only if TX 2 knows the coefficient used to transmit at TX 1.

The last point to discuss is the choice of the coefficient used to transmit at TX 1. Actually, the beamformer can be multiplied arbitrarily by any unit norm complex number without impacting the rate achieved, so that only the power used at TX 1 needs to be decided. In (17), the power used is set to $P/(2\log_2(P))$, which follows the fact that the channel h_{22} might have a very small amplitude, in which case it would be necessary for TX 2 to transmit with a very large power to cancel the interference. To ensure that the interference are canceled for all channel realizations while respecting the power constraint, it is necessary to have the ratio between the power used at TX 1 and the total power tending to zero. The factor $\log_2(P)$ is used because it fulfills this property while not reducing the MG due to the partial power consumption.

E. Power Control for Cooperative Precoding

We have seen that cooperative ZF could achieve a much better MG than ZF. However, this comes at the cost of using only a small share of the available power, which is clearly inefficient and leads to bad performances at finite SNR. To improve the performances, the TX with the worst accuracy needs to adapt its power consumption to the channel realizations. In the following, we propose two possible solutions.

- Firstly, TX 1 can use its local CSI to normalize the beamformer which is then given by

$$\boldsymbol{t}_i^{\mathrm{cZF}} = \sqrt{\frac{P}{2}}\begin{bmatrix}\frac{1}{\sqrt{1+\rho_i^{(1)}}}\\ -\frac{\frac{\tilde{h}_{\bar{i}1}^{(2)}}{\tilde{h}_{\bar{i}2}^{(2)}}}{\sqrt{1+\rho_i^{(2)}}}\end{bmatrix} \tag{24}$$

  with $\rho_i^{(j)} \triangleq |\tilde{h}_{\bar{i}1}^{(j)}|^2/|\tilde{h}_{\bar{i}2}^{(j)}|^2$. This beamformer is not MG maximizing because the local CSI is used at TX 1 so that TX 2 cannot adapt to it to cancel the interference, and the beamformer is not orthogonal to $\tilde{\boldsymbol{h}}_{\bar{i}}^{(2)}$. Yet, this solution achieves good performance at intermediate SNR.
- Another possibility is to assume that TX 1 receives the scalar $\rho_i^{(2)}$ (or ρ_i) and use it to control its power. This means that either the RX or TX 2 needs to feedback this scalar. It requires an additionnal feedback, but only a few bits are necessary, because it is only used to improve the power efficiency and does not impact the MG. Thus, the feedback of this scalar does not change the scaling of the CSI in terms of the SNR nor the performances, and appears thus as an interesting practical solution.

IV. Simulations

We consider two models for the imperfect channel CSI, a statistical model and RVQ. In the statistical model, the quantization error is modeled by adding a Gaussian i.i.d. quantization noise to the channel with the covariance matrix at TX j equal to $\mathrm{diag}([1/P^{\alpha_1^{(j)}}, 1/P^{\alpha_2^{(j)}}])$. When considering given finite number of feedback bits, we compute $\alpha_i^{(j)} = B_i^{(j)}/\log_2(P)$, so that $\mathrm{diag}([1/P^{\alpha_1^{(j)}}, 1/P^{\alpha_2^{(j)}}]) = \mathrm{diag}([1/2^{B_1^{(j)}}, 1/2^{B_2^{(j)}}])$. For RVQ, we consider a number of quantizing bits either numerically given or obtained from the CSI scaling as $q_i^{(j)} = \lfloor\alpha_i^{(j)}\log_2(P)\rfloor$. In the statistical model, we average over 10000 realization and for RVQ we average over 100 codebooks and 1000 channel realizations. In the simulations. we consider the following precoders: ZF with perfect CSI, conventional ZF [cf. (12)], limited ZF [cf. (15)], and cooperative ZF [cf. (17)] with heuristic power control and with 4-bits power control.

In Fig. 1, we consider the statistical model with the CSI scaling $[\alpha_1^{(1)}, \alpha_1^{(2)}] = [1, 0.5]$ and $[\alpha_2^{(1)}, \alpha_2^{(2)}] = [0, 0.7]$. To emphasize the MG (i.e., the slope of the curve in the figure), we let the SNR grow large. As expected theoretically, conventional ZF saturates at high SNR, while limited ZF has a positive slope and cooperative ZF performs close to perfect ZF with a slope only slightly smaller than the optimal one.

In Fig. 2 and Fig. 3 we plot the sum rate achieved with the CSI feedback $[B_1^{(1)}, B_1^{(2)}] = [6, 3]$ and $[B_2^{(1)}, B_2^{(2)}] = [3, 6]$ for

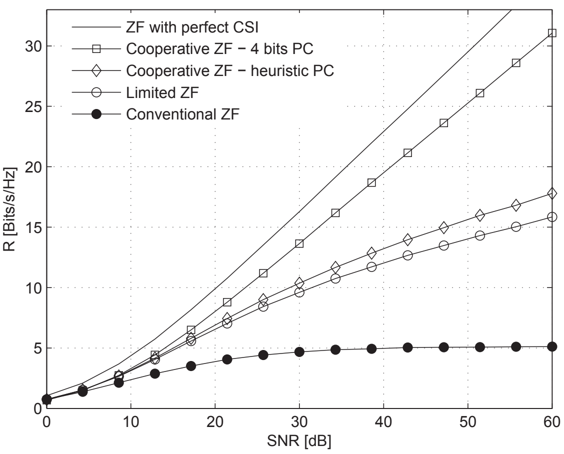

Fig. 1. Sum rate in terms of the SNR with a statistical modeling of the error from RVQ using $[\alpha_1^{(1)}, \alpha_1^{(2)}] = [1, 0.5]$ and $[\alpha_2^{(1)}, \alpha_2^{(2)}] = [0, 0.7]$.

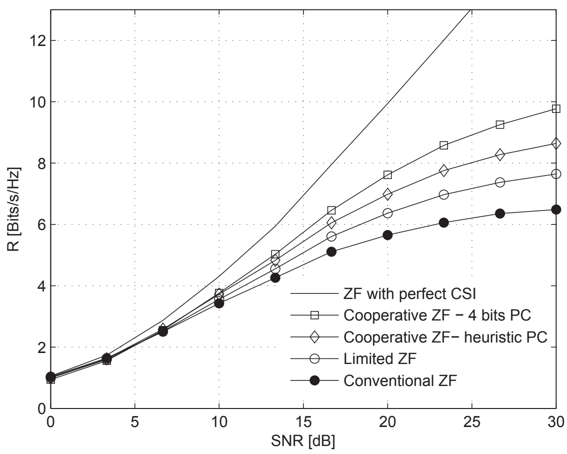

Fig. 3. Sum rate in terms of the SNR with RVQ using $[B_1^{(1)}, B_1^{(2)}] = [6, 3]$ and $[B_2^{(1)}, B_2^{(2)}] = [3, 6]$.

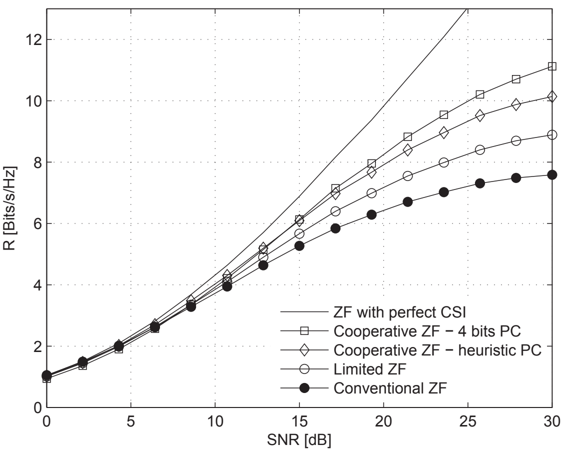

Fig. 2. Sum rate in terms of the SNR with a statistical modeling of the error obtained using $[B_1^{(1)}, B_1^{(2)}] = [6, 3]$ and $[B_2^{(1)}, B_2^{(2)}] = [3, 6]$.

the statistical modeling and RVQ, respectively. Firstly, we can observe the good match between the two models used. From the theoretical analysis the MG is null for all the precoding schemes for a finite number of feedback bits, which can be observed by the saturation of the sum rate as the SNR grows. Yet, the saturation occurs at higher SNR for limited ZF compared to conventional ZF, and at even higher SNR for cooperative ZF, which leads to an improvement of the sum rate even at intermediate SNR.

V. Conclusion

In this work, the multiplexing gain in a two-cell broadcast channel where the TXs have different estimates of the multi-user channel has been studied. We have shown that usual Zero Forcing precoding applied without taking into account the differences in CSI quality achieves far from the maximal MG. We have also derived the value of the maximal MG in that distributed CSI configuration and provided a MG maximizing precoding scheme. Moreover, we have shown by simulations that the new precoding approach outperforms known linear precoding schemes at intermediate SNR. We have considered only two TXs and two RXs with a single antenna to keep the notations simple, but the extension to multiple-antenna TXs or RXs appears to be tractable while the analysis in the case of K TX-RX pairs with a single antenna is done in [12].

References

[1] M. Karakayali, G. Foschini, and R. Valenzuela, “Network Coordination for Spectrally Efficient Communications in Cellular Systems,” *IEEE Wireless Communications*, vol. 13, no. 4, pp. 56–61, Aug. 2006.

[2] O. Somekh, O. Simeone, Y. Bar-Ness, and A. Haimovich, “Distributed Multi-Cell Zero-Forcing Beamforming in Cellular Downlink Channels,” in *Proc. GLOBECOM*, 2006.

[3] N. Jindal, “MIMO Broadcast Channels With Finite-Rate Feedback,” *IEEE Trans. on Inf. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.

[4] D. Samardzija and N. Mandayam, “Unquantized and Uncoded Channel State Information Feedback in Multiple-antenna Multiuser Systems,” *IEEE Trans. on Comm.*, vol. 54, no. 7, pp. 1335–1345, july 2006.

[5] P. Ding, D. Love, and M. Zoltowski, “Multiple Antenna Broadcast Channels With Shape Feedback and Limited Feedback,” *IEEE Trans. on Sign. Process.*, vol. 55, no. 7, pp. 3417–3428, July 2007.

[6] T. Yoo, N. Jindal, and A. Goldsmith, “Multi-Antenna Downlink Channels with Limited Feedback and User Selection,” *IEEE Journal on Selected Areas in Comm.*, vol. 25, no. 7, pp. 1478–1491, Sept. 2007.

[7] M. Shenouda and T. Davidson, “Robust Linear Precoding for Uncertain Miso Broadcast Channels,” in *Proc. ICASSP*, 2006.

[8] N. Vucic, H. Boche, and S. Shi, “Robust Transceiver Optimization in Downlink Multiuser MIMO Systems,” *IEEE Trans. on Sign. Process.*, vol. 57, no. 9, pp. 3576–3587, Sept. 2009.

[9] M. Kobayashi, M. Debbah, and J.-C. Belfiore, “Outage Efficient Strategies for Network MIMO with Partial CSIT,” in *Proc. ISIT*, 2009.

[10] H. Huh, A. Tulinoy, and G. Caire, “Network MIMO large-system analysis and the impact of CSIT Estimation,” in *Proc. CISS*, 2010.

[11] R. Zakhour and D. Gesbert, “Team Decision for the Cooperative MIMO Channel with Imperfect CSIT Sharing,” in *Proc. ITA*, 2010.

[12] P. de Kerret and D. Gesbert, “Zero Forcing in a Multicell MIMO Channels with Unequal CSI,” 2011, http://arxiv.org/submit/257686/view.

[13] W. Dai, Y. Liu, and B. Rider, “Quantization Bounds on Grassmann Manifolds and Applications to MIMO Communications,” *IEEE Trans. on Inf. Theory*, vol. 54, no. 3, pp. 1108–1123, March 2008.