CSI feedback allocation in multicell MIMO channels

Paul de Kerret, David Gesbert

CSI Feedback Allocation in Multicell MIMO Channels

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], we consider the joint precoding across K transmitters (TXs), sharing the knowledge of the user's data symbols being transmitted to K single-antenna receivers (RXs). We consider a distributed channel state information (DCSI) configuration where each TX has its own local estimate of the overall multiuser MIMO channel. Our focus is on the optimization of the allocation of the CSI feedback subject to a constraint on the total amount of feedback. As a starting point, we consider the Wyner model where we derive a distance-based CSI allocation achieving close to the optimal performance using only a small percentage of the total feedback. The approach relies on the exploitation of the attenuation to restrict the cooperation at a local scale. Indeed, the CSI and the user's data symbols are then shared to only a finite number of neighbors such that our approach appears as an improved alternative to clustering.**

I. Introduction

Network or Multicell MIMO methods, whereby multiple interfering transmitters (TXs) share user messages and allow for joint precoding, are currently considered for next generation wireless networks [1]. With perfect message and channel state information (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. However, the allocation of the data symbols and the CSI to the cooperating TXs impose huge requirements on the architecture. The common solution is to use disjoint clusters to reduce the amount of data to be shared [2], [3]. Yet, clustering limits the performance of the system because of the interference created at the edge of the clusters.

In recent works, adaptive allocation of the CSI feedback bits in a multicell network has been studied [4]–[7], yet, always considering that the TXs jointly designing the beamformers were sharing the *same* channel estimate. On the opposite, we consider here the joint precoding from TXs having their own *local* channel estimates which models more realistically the cooperation of non-colocated TXs which are unlikely to receive the CSI with the same delay and the same quality. This setting, introduced in [8], [9] as the *distributed CSI* (DCSI)-MIMO channel, opens up completely new research problems as it can be seen as a team-decision problem [10].

Furthermore, it opens up the possibility to allocate each TX with exactly the CSI it really needs. Indeed, it is very intuitive that, in wireless networks, the precision with which a channel to a given RX should be known at a given TX depends on the distance between the TX and the RX. However, no result exists in the literature to describe this effect. Quantifying this well known intuition and translating it into an efficient allocation of the CSI is precisely the objective of this work.

We start by analyzing a very simplified channel model, referred to in the literature as the *Wyner* model, in which the TXs and the RXs are placed on a one-dimensional space (e.g. a line) and receive signals only from a few neighboring TXs. This simplistic model has the advantage of being more tractable while still offering valuable insights on more realistic channels. This model has been introduced in [11] and has been very successful since, particularly to model cooperation in wireless networks [12]–[14]. Furthermore, its simplicity has allowed the development of the approach which could then be extended to more general channel models [15].

In this work, we use an asymptotic analysis taking into account the geometry of the network to derive analytically a *distance-based* CSI allocation allowing for a large reduction of the CSI required at the TXs at the cost of reduced performance losses. Furthermore, the CSI allocation proposed reduces the sharing of the user's data symbol and the CSI at a local scale and consequently reveals itself as a realistic alternative to clustering, with improved performance and similar cost.

II. System Model

We consider the distributed CSI (DCSI)-MIMO channel, in which K transmitters (TXs) transmit *jointly* using linear precoding to K receivers (RXs) equipped with a single antenna and applying single user decoding. Each TX has the knowledge of the K symbols to transmit to the K RXs. Besides the knowledge of data symbol, each TX is supposed to acquire, through an unspecified feedback or sharing mechanism, its *own* estimate on the channel vectors to all the users.

The channel is represented by the channel matrix $\mathbf{H} \in \mathbb{C}^{K\times K}$ and the transmission is described mathematically as

$$\begin{bmatrix} y_1 \\ \vdots \\ y_K \end{bmatrix} = \mathbf{H}\boldsymbol{x} + \boldsymbol{\eta} = \begin{bmatrix} \boldsymbol{h}_1^{\mathrm{H}}\boldsymbol{x} \\ \vdots \\ \boldsymbol{h}_K^{\mathrm{H}}\boldsymbol{x} \end{bmatrix} + \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_K \end{bmatrix} \quad (1)$$

where y_i is the signal received at the i-th RX, $\boldsymbol{h}_i^{\mathrm{H}} \in \mathbb{C}^{1\times K}$ the channel to the i-th RX, $\boldsymbol{\eta} = [\eta_1, \ldots, \eta_K]^{\mathrm{T}} \in \mathbb{C}^{K\times 1}$ the zero

[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.

mean unit variance i.i.d. complex Gaussian ($\mathcal{CN}(0,1)$). $\boldsymbol{x} \in \mathbb{C}^{K\times 1}$ is the transmitted signal obtained from the symbol vector $\boldsymbol{s} = [s_1, \ldots, s_K]^{\mathrm{T}} \in \mathbb{C}^{K\times 1}$ (i.i.d. $\mathcal{CN}(0,1)$) as

$$\boldsymbol{x} = \mathbf{T}\boldsymbol{s} = \begin{bmatrix} \boldsymbol{t}_1 & \ldots & \boldsymbol{t}_K \end{bmatrix} \begin{bmatrix} s_1 \\ \vdots \\ s_K \end{bmatrix} \tag{2}$$

where $\mathbf{T} \in \mathbb{C}^{K\times K}$ is the precoding matrix and $\boldsymbol{t}_i \in \mathbb{C}^{K\times 1}$ is the beamforming vector used to transmit s_i to RX i. We consider a per-stream power constraint $\|\boldsymbol{t}_i\|^2 = P$.

Our focus is on the maximization of the sum rate averaged over the fading distribution, where the rate of user i reads as

$$R_i \triangleq \mathrm{E}_{\mathbf{H}}\left[\log_2\left(1 + \frac{|\boldsymbol{h}_i^{\mathrm{H}}\boldsymbol{t}_i|^2}{1+\sum_{\ell\neq i}|\boldsymbol{h}_i^{\mathrm{H}}\boldsymbol{t}_\ell|^2}\right)\right]. \tag{3}$$

A. Distributed CSI and Distributed Precoding

In the DCSI-MIMO channel, the j-th TX has its own individual estimate of its channel $\boldsymbol{h}_i^{\mathrm{H}}$ to RX i for all i, denoted by $\boldsymbol{h}_i^{(j)\mathrm{H}}$ and obtained from *Random Vector Quantization (RVQ)* [16] using $B_i^{(j)}$ bits. We focus on interference limited wireless networks working at high SNR so that we assume that Zero Forcing (ZF) precoders are used. The beamformers then read as

$$\forall i \in \{1,\ldots,K\},\ \boldsymbol{t}_i^{(j)} \triangleq \sqrt{P}\frac{\left(\mathbf{H}^{(j)}\right)^{-1}\boldsymbol{e}_i}{\|\left(\mathbf{H}^{(j)}\right)^{-1}\boldsymbol{e}_i\|}. \tag{4}$$

Although a given TX j may compute the whole precoding matrix $\mathbf{T}^{(j)}$, only the j-th row will be used in practice. Finally, the effective precoder is given by

$$\mathbf{T} \triangleq \begin{bmatrix} \boldsymbol{t}_1 & \ldots & \boldsymbol{t}_K \end{bmatrix} \triangleq \begin{bmatrix} T_{11}^{(1)} & T_{12}^{(1)} & \ldots & T_{1K}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ T_{K1}^{(K)} & T_{K2}^{(K)} & \ldots & T_{KK}^{(K)} \end{bmatrix}. \tag{5}$$

More detail on the distributed precoding can be found in [9].

B. Optimized CSI Allocation

Depending on the CSI accuracy at a TX, the coefficients implemented will be very different or not from the coefficients obtained with perfect CSI. Considering a realistic constraint on the total number of CSI bits transmitted via the multiuser feedback channel, which we denote by $B_{\max}$, the ideal approach would be to find the allocation of the feedback bits maximizing the average sum rate:

$$\underset{\{B_i^{(j)}\}}{\text{maximize}} \sum_{i=1}^{K} R_i, \text{ s.t. } \sum_{j=1}^{K}\sum_{i=1}^{K} B_i^{(j)} \leq B_{\max}. \tag{6}$$

However, this problem is very hard to solve because it is a discrete optimization where in each step we have to derive the sum rate with the partial CSI allocation.

C. Asymptotic Approach

Thus, we consider this problem asymptotically in the SNR and we aim in fact at finding the minimal CSI allocation achieving the full Multiplexing Gain (MG) when the interference attenuation increases in terms of P. We will then show that this leads to an efficient solution for the optimization problem (6). We start by recalling an asymptotic result for the DCSI-MIMO channel without pathloss.

Theorem 1. *[9] In the DCSI-MIMO Rayleigh channel with K single antennas RXs and K single antennas TXs, it is necessary and sufficient with RVQ in order to achieve the maximal Multiplexing Gain to quantize each normalized channel vector $\tilde{\boldsymbol{h}}_i$ with a number of bits in the order of $O((K-1)\log(P))$. Furthermore, the maximal Multiplexing Gain is achieved if and only if $E[\|\boldsymbol{t}_i - \boldsymbol{t}_i^{PCSI}\|^2] = O(1/(P))$.*

In the Wyner model, the factor K is replaced by the number of non-zero coefficient, i.e., 3. We study the mean square error between the beamformer implemented and the beamformer with perfect CSI because this has been recognized in [9] as being well adapted to the DCSI-MIMO channel. Indeed, each TX implements one beamforming coefficient and the consistency between the TXs has a critical impact on the performance.

III. Wyner Model

We consider a finite linear version of the Wyner model where K TXs equipped with only one antenna are uniformly distributed along a line. A RX receives interference coming from its two direct neighboring TXs with an interference attenuation factor equal to $\mu \in (0,1)$, while the short term fading is assumed to be Rayleigh distributed. Thus, the multiuser channel matrix $\mathbf{H} \in \mathbb{C}^{K\times K}$ is a tridiagonal matrix defined as

$$\mathbf{H} \triangleq \begin{bmatrix} d_1 & a_1\mu & 0 & \ldots & \ldots & \ldots & 0 \\ \mu b_2 & d_2 & \mu a_2 & 0 & \ldots & \ldots & 0 \\ \ldots & \ddots & \ddots & \ddots & 0 & \ldots & 0 \\ \ldots & 0 & \mu b_i & d_i & \mu a_i & 0 & \ldots \\ 0 & \ldots & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & \mu b_{K-1} & d_{K-1} & \mu a_{K-1} \\ 0 & \ldots & \ldots & \ldots & 0 & \mu b_K & d_K \end{bmatrix}$$

where b_i, d_i, and a_i are distributed as standard i.i.d. complex Gaussian ($\mathcal{CN}(0,1)$). We denote the estimates of b_i, d_i, and a_i at TX j by $b_i^{(j)}$, $d_i^{(j)}$, and $a_i^{(j)}$, respectively, while the estimate errors are then $\Delta b_i^{(j)}$, $\Delta d_i^{(j)}$, and $\Delta a_i^{(j)}$, respectively.

A. Inverse of a Tridiagonal Matrix

The Wyner model allows to obtain a closed form for the channel inverse [17], which we recall in the following and will be useful to quantify the effect of the limited CSI.

We start by introducing two $K+1$ dimensional vectors $\boldsymbol{\beta} \triangleq [\beta_1, \ldots, \beta_{K+1}]$ and $\boldsymbol{\alpha} \triangleq [\alpha_0, \ldots, \alpha_K]$ which will be

used in the matrix inverse and are defined as

$$\alpha_i \triangleq \begin{cases} 1 & \text{, if i=0} \\ d_1 & \text{, if i=1} \\ d_i\alpha_{i-1} - \mu^2 b_i a_{i-1}\alpha_{i-2} & \text{, if i=2,...,K} \end{cases}$$

and

$$\beta_i \triangleq \begin{cases} 1 & \text{, if i=K+1} \\ d_K & \text{, if i=K} \\ d_i\beta_{i+1} - \mu^2 b_{i+1} a_i \beta_{i+2} & \text{, if i=1,...,K-1} \end{cases}.$$

The channel inverse $\mathbf{H}^{-1}$ then reads as follows. First, the diagonal elements for $i = 2, \ldots, n-1$ are

$$\{\mathbf{H}^{-1}\}_{ii} = \left(d_i - \frac{\mu^2 b_i a_{i-1}\alpha_{i-2}}{\alpha_{i-1}} - \frac{\mu^2 b_{i+1} a_i \beta_{i+2}}{\beta_{i+1}} \right)^{-1}$$

while the two extremal diagonal elements read as

$$\{\mathbf{H}^{-1}\}_{11} = \left(d_1 - \frac{\mu^2 b_2 a_1 \beta_3}{\beta_2} \right)^{-1}$$
$$\{\mathbf{H}^{-1}\}_{nn} = \left(d_n - \frac{\mu^2 b_n a_{n-1}\alpha_{n-2}}{\alpha_{n-1}} \right)^{-1}.$$

Last, the off-diagonal elements are given by

$$\{\mathbf{H}^{-1}\}_{ij} = \begin{cases} (-\mu)^{j-i} (\prod_{k=1}^{j-i} a_{j-k}) \frac{\alpha_{i-1}}{\alpha_{j-1}} \{\mathbf{H}^{-1}\}_{jj} & \text{if } i < j \\ (-\mu)^{i-j} (\prod_{k=1}^{i-j} b_{j+k}) \frac{\beta_{i+1}}{\beta_{j+1}} \{\mathbf{H}^{-1}\}_{jj} & \text{if } i > j. \end{cases}$$

B. CSI Allocation in the Wyner Model

We can observe in the expressions given in Subsection III-A for the inverse of a tridiagonal channel that the amplitude of the coefficients in the inverse decreases exponentially as the elements get away from the diagonal.

To obtain analytical results at high SNR, we consider that the interference attenuation factor μ can be written as a fixed fraction of P, i.e., $P^{-\zeta}$ for a certain $\zeta \in (0,1)$. Thus, the interference attenuation factor decreases as the SNR increases. Note that this is only an artifact to model the impact of the attenuation at high SNR, as the impact of the pathloss is otherwise neglected. We confirm by simulations that the performance remain good as μ increases.

Theorem 2. *Considering that $\mu = P^{-\zeta}$ for some $\zeta \in (0,1)$, it is sufficient in order to achieve the maximal MG to quantize the channel vectors $\boldsymbol{h}_i^H$ at TX j with a number of bits*

$$B_i^{(j)} = \max(2\log_2(P\mu^{2|i-j|}), 0). \tag{7}$$

Proof: A sketch of the proof is given in the following and more detail is given in [15].

We start by inserting the CSI error estimates in the off-diagonal elements of the inverse given in Subsection III-A. We consider the coefficient with $j < i$ as the case $j > i$ follows by symmetry. In a first step we consider the coefficient at TX j corresponding to the transmission of stream i.

$$\begin{aligned} \frac{\{\mathbf{H}^{(j)-1}\}_{ji}}{\{\mathbf{H}^{(j)-1}\}_{ii}} &= (-\mu)^{i-j} \frac{\hat{\alpha}_{j-1}^{(j)} + \Delta\alpha_{j-1}^{(j)}}{\hat{\alpha}_{i-1}^{(j)} + \Delta\alpha_{i-1}^{(j)}} \prod_{k=1}^{i-j} (\hat{a}_{i-k}^{(j)} + \Delta a_{i-k}^{(i)}) \\ &\approx (-\mu)^{j-i} \frac{\hat{\alpha}_{j-1}^{(j)} + \Delta\alpha_{j-1}^{(j)}}{\hat{\alpha}_{i-1}^{(j)} + \Delta\alpha_{i-1}^{(j)}} \left(\prod_{k=1}^{i-j} \hat{a}_{i-k}^{(j)} \right) \left(\sum_{\ell=1}^{i-j} \frac{\Delta a_{j+\ell}^{(j)}}{\hat{a}_{j+\ell}^{(j)}} \right) \end{aligned} \tag{8}$$

where we have done a first order approximation in the error terms as the CSI allocation will be derived so as to fulfill the scaling of the mean square error in $1/P$ as given in Theorem 1. To proceed, we approximate the first ratio in (8) by keeping only the first order coefficients in μ which gives:

$$\begin{aligned} \frac{\hat{\alpha}_{j-1}^{(j)} + \Delta\alpha_{j-1}^{(j)}}{\hat{\alpha}_{i-1}^{(j)} + \Delta\alpha_{i-1}^{(j)}} &\approx \prod_{k=1}^{i-j} (\hat{d}_{i-k}^{(j)} + \Delta d_{i-k}^{(j)})^{-1} \\ &\approx \left(\prod_{k=1}^{i-j} \hat{d}_{i-k}^{(j)} \right)^{-1} \left(\sum_{\ell=1}^{i-j} \frac{\Delta d_{j+\ell}^{(j)}}{\hat{d}_{j+\ell}^{(j)}} \right)^{-1}. \end{aligned} \tag{9}$$

To obtain an approximate closed form for the channel inverse, we also need to approximate the diagonal terms in the denominator of the Left Hand Side in (8) as

$$\{\mathbf{H}^{-1}\}_{ii} = \left(d_i - \frac{\mu^2 b_{i+1} a_i \beta_{i+2}}{\beta_{i+1}} - \frac{\mu^2 b_i a_{i-1}\alpha_{i-2}}{\alpha_{i-1}} \right) \approx \frac{1}{d_i}. \tag{10}$$

The sum of all the error terms arising in (8) has to tend to zero as $O(1/P)$ due to the fact that the channel inverse elements given in (8) are further normalized so as to fulfilled the *per-stream* power constraint of P.

Inserting (9) and (10) in (8) we conclude by inspection that this condition is fulfilled on the first order when $a_\ell^{(j)}$ and $d_\ell^{(j)}$ for $\ell \in \{i-1, i-2, \ldots, j\}$ as well as $d_i^{(j)}$ are obtained from the quantization of the channel vector $\boldsymbol{h}_\ell$ for $\ell \in \{i-1, i-2, \ldots, j\}$ with $\max(2\log_2(P\mu^{2(i-j)}), 0)$ bits.

These conditions correspond only to the strongest requirement for that stream at that TX, but the other channel parameters need obviously to be also known, although with a lower accuracy. These additional requirements are obtained by proceeding similarly but with higher order approximation in (9). Furthermore, for a given TX, this has to be done for all the streams to be transmitted.

Yet, only the strongest requirements are of interest and going through all the conditions, we obtain that the strongest requirements are fulfilled by allocating to TX j $\max(2\log_2(P\mu^{2(i-j)}), 0)$ bits for the vector quantization of $\boldsymbol{h}_i^{(j)}$ from which $\hat{d}_i^{(j)}$ stems, and $\max(2\log_2(P\mu^{2(i-j+1)}), 0)$ bits for the vector quantization of $\boldsymbol{h}_i^{(j)}$ from which $\hat{a}_i^{(j)}$ and $\hat{b}_i^{(j)}$ come. Taking the first, more constraining, condition concludes the proof. ■

Intuitively, the CSI allocation is based on the fundamental property that a given TX contributes a large power only to the streams transmitted to neighboring RXs. When contributing with only a small amount of power, a less accurate CSI is

then needed to maintain a constant interference level created by the transmission of that stream.

C. Discussion of the Derived CSI Allocation

The increase of the feedback allocation to a given TX when the total number of cooperating TXs increases, i.e., its scaling in K, is a very good indicator of the feasibility of the CSI allocation. In the full CSI allocation given in Theorem 1, the number of feedback bits allocated per TX increases linearly with the number of cooperating TXs since each TX needs to know all the channel vectors. This is problematic because adding one TX increases the feedback required at *all* the TXs of the network and we discuss in the following the improvement brough by the proposed CSI allocation.

Proposition 1. *For finite P and μ, the CSI allocation given in Theorem 2 leads to a number of CSI bits allocated per TX which does not increase with the number of cooperating TXs.*

Proof: The proof follows easily from the CSI allocation given in Theorem 2 and a detailed proof is given in [15]. ■

Intuitively, the boundedness of the number of feedback bits allocated to each TX follows from the exponential decay of the off-diagonal coefficients. This reduces the cooperation to a *local scale* since the power allocated to the transmission of datas to users asymptotically far away is asymptotically small.

With the sharing of the CSI, the second main requirement for joint precoding comes from the sharing of the user's data symbols to all the TXs. This condition cannot be fulfilled in realistic networks when the size of the network increases. This represents an hindrance for the practical use of large cooperation areas and we now discuss whether full sharing of the data symbols is really necessary.

Corollary 1. *Considering that $\mu = P^{-\zeta}$ for a fixed $\zeta \in (0,1)$, the CSI allocation proposed in Theorem 2 requires only the sharing of symbol s_i to TX j if $2\log_2(P\mu^{2|i-j|}) > 0$. Thus, for finite P and μ, each symbol is shared only to a finite number of cooperating TXs which does not increase with the total number of cooperating TXs.*

Proof: The proof is similar to the proof of Proposition 1 and is available in [15]. ■

The distance-based CSI allocation extends to the routing of the user's data symbols. Combining the two properties, the proposed solution offer a practical alternative to clustering.

D. Heuristic Finite SNR CSI Allocation

The CSI allocation has been discussed only asymptotically and we aim now at using this insight to design a feedback scheme performing well in realistic conditions. It is proved in [18] that it is necessary and sufficient for achieving an optimally scaling sum rate to use for each normalized channel a number of bits in the order of $O((K-1)\log(K^2P))$ in a Rayleigh fading channel with distributed CSI. The adaptation of this formula to the Wyner channel is out of the scope of this paper, but we conjecture it to be $O(2\log(KP))$ which can be shown to be at least a lower bound for the scaling [15].

Achieving good performance requires the proposed CSI allocation to be close to $O(2\log(KP))$ when μ increases. Thus, we propose a heuristic solution which preserves the property of local cooperation described above, yet converges to the desired CSI allocation when μ tends to one. It consists in finding iteratively a K_{SNR} verifying that $\mu^{2K_{\text{SNR}}} < 1/(K_{\text{SNR}}P)$ and then use the CSI allocation from Theorem 2 with $\min(K_{\text{SNR}}, K)P$ instead of only P. The number K_{SNR} can be understood as being the number of TXs which are inside the cooperation neighborhood and for which we apply the CSI allocation without pathloss.

E. Extension to More General Channel Models

The analysis has been carried out in the Wyner model but has been extended to *exponentially decaying* channel matrices and *polynomially decaying* channel matrices. These matrices model much more realistically practical channels as they have only nonzero elements and solely a condition on the rate of decrease of the outer-diagonal elements. Furthermore, this can be shown to model well the wireless transmission from a linear array (e.g., TXs along a road). Interestingly, the properties of the CSI allocation derived (scaling in K, user's data allocation) are not proper to the Wyner model and are preserved in the other models considered. The extension can be found in [15].

IV. Simulations

We simulate the average sum rate in the Wyner model presented at the beginning of Section III for a network made of 25 TXs. For comparison, we show the performance achieved when *perfect CSI* is available at all TXs. To evaluate the efficiency of the *distance-based* CSI allocation proposed in Theorem 2 using the heuristic described in Subsection III-D, we also compare it to two alternative CSI allocations using the same total number of feedback bits but allocated differently: the *uniform* CSI allocation and the conventional *clustering* solution with disjoint clusters made of 5 TXs.

In Fig. 1, we plot the average rate per user as a function of the normalized transmit power for $\mu = 0.5$. The distance-based CSI allocation outperforms the clustering solution which fails to achieve a positive MG as well as the uniform CSI allocation which becomes efficient only at asymptotically high SNR.

In Fig. 2, we plot the percentage of the sum rate with perfect CSI which is achieved using the different CSI allocations, in terms of the interference attenuation factor μ. Additionally, we plot in Fig. 3 the matching number of bits used by the distance-based CSI allocation normalized over the "full" CSI allocation $2\log_2(KP)$. The proposed solution achieves most of the performance attained with perfect CSI while using only a small percentage of the total feedback required.

V. Conclusion

In this work, we have studied the allocation of the CSI bits to non-colocated TXs cooperating to transmit jointly to their associated RXs. For the Wyner model, we have derived

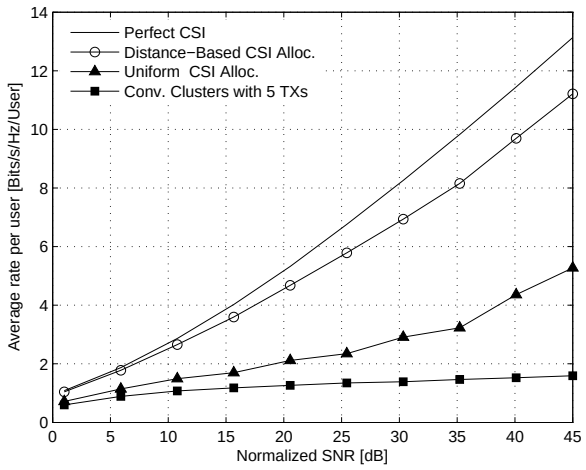

Fig. 1. Average rate per user in terms of the normalized transmit power per TX P for $\mu = 0.5$.

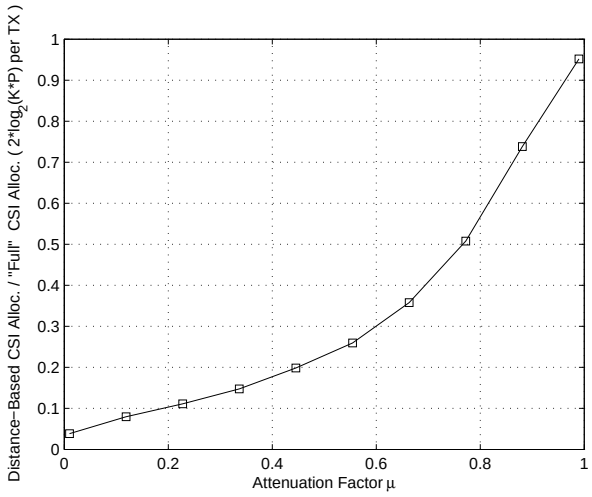

Fig. 3. Percentage of the total feedback with full CSI allocation (i.e., $K(2\log_2(KP))$ bits) for $P = 20$ dB.

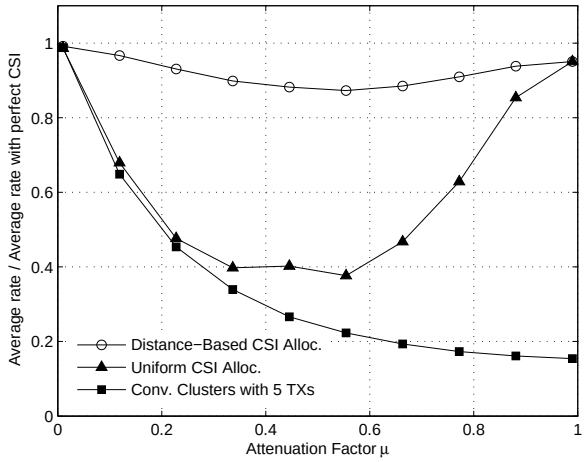

Fig. 2. Percentage of the sum rate achieved with perfect CSI in terms of the interference attenuation factor μ for $P = 20$ dB.

a distance-based CSI allocation achieving most of the performance attained with perfect CSI while requiring only a small share of the total feedback. Particularly, the solution proposed requires only the sharing of the CSI and the user's data symbol to a local neighborhood around each TX. Thus, the solution proposed has the same requirements as a conventional clustering solution and appears as a realistic alternative. Our approach has been presented for the simplistic Wyner model but its extension to more realistic channel models has been carried in [15]. As an alternative to clustering, the optimization of the CSI allocation at finite SNR and in other settings represents an interesting direction of research.

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 Journal on Selected Areas in Communications*, vol. 28, no. 9, pp. 1380–1408, Dec. 2010.

[2] J. Zhang, R. Chen, J. Andrews, A. Ghosh, and R. Heath, "Networked MIMO with Clustered Linear Precoding," *IEEE Trans. on Wireless Communications*, vol. 8, no. 4, pp. 1910 –1921, April 2009.

[3] A. Papadogiannis, D. Gesbert, and E. Hardouin, "A Dynamic Clustering Approach in Wireless Networks with Multi-Cell Cooperative Processing," in *Proc. ICC 2008*.

[4] J. Zhang and J. Andrews, "Adaptive Spatial Intercell Interference Cancellation in Multicell Wireless Networks," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 9, pp. 1455 –1468, Dec. 2010.

[5] W. W. L. Ho, T. Q. S. Quek, S. Sun, and R. W. Heath, "Decentralized Precoding for Multicell MIMO Downlink," *IEEE Trans. on Wireless Communications*, vol. 10, no. 6, pp. 1798 –1809, June 2011.

[6] R. Bhagavatula and R. Heath, "Adaptive Limited Feedback for Sum-Rate Maximizing Beamforming in Cooperative Multicell Systems," *IEEE Trans. on Signal Processing*, vol. 59, no. 2, pp. 800–811, Feb. 2011.

[7] ——, "Adaptive Bit Partitioning for Multicell Intercell Interference Nulling with Delayed Limited Feedback," *IEEE Trans. on Signal Processing*, vol. PP, no. 99, p. 1, 2011.

[8] P. de Kerret and D. Gesbert, "The Multiplexing Gain of a Two-cell MIMO Channel with Unequal CSI," in *Proc. ISIT 2011*.

[9] ——, "The Multiplexing Gain of the Network MIMO Channel With Distributed CSI," 2011, http://arxiv.org/submit/257686/view. Submitted to Trans. Information Theory Aug. 2011.

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

[11] A. Wyner, "Shannon-Theoretic Approach to a Gaussian Cellular Multiple-Access Channel," *IEEE Trans. on Information Theory*, vol. 40, no. 6, pp. 1713 –1727, nov 1994.

[12] R. Bhagavatula, R. Heath, and B. Rao, "Limited Feedback with Joint CSI Quantization for Multicell Cooperative Generalized Eigenvector Beamforming," in *Proc. ICASSP*, 2010.

[13] N. Levy, S. Shamai, M. Wigger, and A. Lapidoth, "A Cognitive Network with Clustered Decoding," in *Proc. ISIT 2009*.

[14] S. Shamai (Shitz) and M. Wigger, "Rate-Limited Transmitter-Cooperation in Wyners Asymmetric Interference Network," in *Proc. ISIT 2011*.

[15] P. de Kerret and D. Gesbert, "Distance-Based CSI Allocation in Multicell MIMO Channels," 2012, in preparation.

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

[17] M. El-Mikkawy and A. Karawia, "Inversion of general tridiagonal matrices," *Elsevier Applied Mathematics Letters*, Nov. 2005.

[18] P. de Kerret and D. Gesbert, "Feedback Design in MIMO Broadcast Channels with Distributed CSI," 2012, in preparation. Available upon request.