

Distributed Multicell and Multiantenna Precoding: Characterization and Performance Evaluation

Emil Björnson
ACCESS Linnaeus Center
Royal Institute of Technology (KTH)
SE-100 44 Stockholm, Sweden
Email: emil.bjornson@ee.kth.se

Randa Zakhour, David Gesbert
Mobile Communications Department
EURECOM
06560 Sophia Antipolis, France
Email: {zakhour,gesbert}@eurecom.fr

Björn Ottersten
ACCESS Linnaeus Center
Royal Institute of Technology (KTH)
SE-100 44 Stockholm, Sweden
Email: bjorn.ottersten@ee.kth.se

Abstract—This paper considers downlink multiantenna communication with base stations that perform cooperative precoding in a *distributed* fashion. Most previous work in the area has assumed that transmitters have *common* knowledge of both data symbols of all users and full or partial channel state information (CSI). Herein, we assume that each base station only has *local* CSI, either instantaneous or statistical. For the case of instantaneous CSI, a parametrization of the beamforming vectors used to achieve the outer boundary of the achievable rate region is obtained for two multi-antenna transmitters and two single-antenna receivers. Distributed generalizations of classical beamforming approaches that satisfy this parametrization are provided, and it is shown how the distributed precoding design can be improved using the so-called virtual SINR framework [1]¹. Conceptually analog results for both the parametrization and the beamforming design are derived in the case of local statistical CSI. Heuristics on the distributed power allocation are provided in both cases, and the performance is illustrated numerically.

I. INTRODUCTION

The use of multiple antennas has the potential of greatly improving the performance of wireless systems. In single-cell scenarios, many algorithms have been proposed to enable simultaneous downlink transmission to multiple users [3]. These approaches exploit various amounts of channel state information (CSI) to improve the performance by optimizing the signal power and limiting the intra-cell interference. In multi-cell scenarios, the interference from adjacent cells has a significant and limiting impact on the performance—especially for users close to the cell boundary. The problem of handling inter-cell interference has recently attracted much attention [4].

Ideally, all base stations might share their CSI and data through backhaul links, which would enable coordinated precoding design that can manage the co-user interference as in the single-cell scenario. In practice, there are limitations in terms of delay and capacity on the backhaul and computational power at the transmitters, which makes it necessary to investigate schemes that reduce the signaling while still benefiting from a robust interference control. A practical

¹B. Ottersten is also with securityandtrust.lu, University of Luxembourg. This work is supported in part by the FP6 project Cooperative and Opportunistic Communications in Wireless Networks (COOPCOM), FP6-033533. It is also supported in part by the French project ORMAC.

¹A similar approach was proposed, though justified differently, in [2] where the authors define a metric similar to a virtual SINR which they term *signal to generating interference plus noise ratio* (SGINR).

iterative message passing procedure was proposed in [5] to exchange information between neighboring cells. An information theoretic approach was taken in [6] to determine the dependence of multi-cell rates on backhaul capacity.

The multiple-input single-output interference channel (MISO IC) represents the special case when only CSI is shared among the transmitters. Although each base station aims at maximizing the rate achieved by its own user, cooperation will greatly improve the performance [7]. The achievable rate region was characterized in [8] and the authors proposed a game-theoretical precoding design based on full CSI [7]. Precoding strategies that only exploits local CSI but still attains optimal rate points were proposed in [1], [9] based on a virtual signal-to-interference-and-noise ratio (SINR) framework.

Herein, we address the different problem of distributed multi-cell MIMO precoding where the cooperating base stations share knowledge of the data symbols but have local CSI only, thus much reducing the feedback load on the uplink and avoiding cell-to-cell CSI exchange. In this context, previous work includes [10] where the *layered virtual SINR framework* was introduced as a heuristic framework for distributed MIMO precoding design. In this paper, the novel contributions are:

- We characterize the outer boundary of the achievable rate region under linear precoding for the multicell MIMO channel for both instantaneous and statistical CSI.
- We provide theoretical optimality justifications for using the layered virtual SINR framework of [10] to design the precoding under local instantaneous CSI knowledge.
- We extend the layered virtual SINR approach to handle local statistical channel state information for cases when instantaneous fading information is not available.

Finally, we provide simulation results comparing the layered virtual SINR framework with generalizations of classical precoding approaches and with beamforming on the MISO IC.

Notation: The orthogonal projection matrix onto the column space of the matrix \mathbf{X} is $\Pi_{\mathbf{X}} = \mathbf{X}(\mathbf{X}^H \mathbf{X})^{-1} \mathbf{X}^H$; that onto its complement is $\Pi_{\mathbf{X}}^{\perp} = \mathbf{I} - \Pi_{\mathbf{X}}$, where \mathbf{I} is the identity matrix. $\mathcal{CN}(\bar{\mathbf{x}}, \mathbf{Q})$ is used to denote circularly symmetric complex Gaussian random vectors, where $\bar{\mathbf{x}}$ is the mean and \mathbf{Q} the covariance matrix. If $k \in \{1, 2\}$, then \bar{k} denotes the value not being k (e.g., $\bar{k} = \text{mod}(k, 2) + 1$ for $k = 1, 2$).

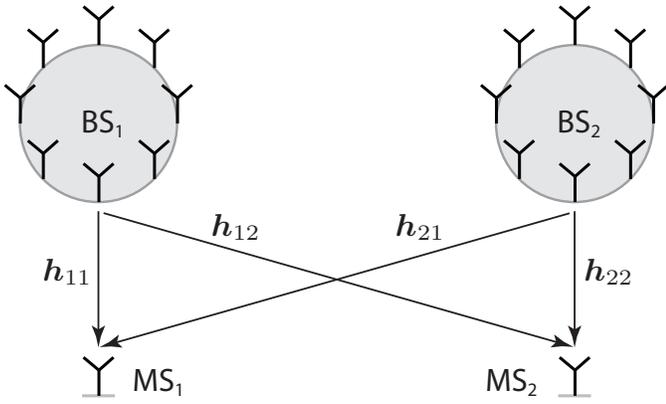


Fig. 1. The basic two-transmitter and two-receiver scenario with $N_t = 8$.

II. SYSTEM MODEL

The communication scenario herein consists of two transmitters (e.g., base stations in a cellular system) equipped with $N_t \geq 2$ antennas each and two single-antenna receivers (e.g., mobile terminals). The transmitters and receivers are denoted BS_j and MS_k , respectively, with $j, k \in \{1, 2\}$. This setup is illustrated in Fig. 1 for $N_t = 8$. The results of this paper can be generalized to $K > 2$ transmitters and receivers, but we limit ourselves to $K = 2$ in order to ease exposition.

Let $\mathbf{x}_j \in \mathbb{C}^{N_t}$ be the signal transmitted by BS_j and let the corresponding received signal at MS_k be denoted y_k , for $j, k \in \{1, 2\}$. The propagation channel to MS_k is assumed to be narrowband and flat-fading with the system model

$$y_k = \sum_{j=1}^2 \mathbf{h}_{jk}^H \mathbf{x}_j + n_k, \quad (1)$$

where $\mathbf{h}_{jk} \in \mathcal{CN}(0, \mathbf{Q}_{jk})$ is the channel between BS_j and MS_k and $n_k \in \mathcal{CN}(0, \sigma^2)$ is the additive noise. Throughout the paper, each receiver MS_k has full local CSI (i.e., perfect estimates of \mathbf{h}_{jk} for $j = 1, 2$). At the transmit side, we will distinguish between two different types of local CSI:

- *Local Instantaneous CSI*: BS_j knows the instantaneous value of \mathbf{h}_{jk} , for $k = 1, 2$, and the noise power σ^2 .
- *Local Statistical CSI*: BS_j knows the statistics of \mathbf{h}_{jk} (e.g., covariances), for $k = 1, 2$, and the noise power σ^2 .

Observe that in both cases, the base stations are only assumed to have CSI that can be obtained locally (either through feedback or reverse-link estimation). Hence, there is no exchange of CSI between the base stations.

A. Joint Multi-Transmitter Linear Precoding

Let $s_k \in \mathcal{CN}(0, 1)$ denote the data symbol intended for MS_k . Unlike the MISO IC [7], [8], we assume that the data symbols intended for both receivers are available at both transmitters. This will enable joint multi-transmitter precoding, although each transmitter only has local CSI. Proper transmission synchronization is however required to avoid inter-symbol interference. Under the assumption of linear precoding at each transmitter, the signal transmitted by BS_j is

$$\mathbf{x}_j = \sum_{k=1}^2 \sqrt{p_{jk}} \mathbf{w}_{jk} s_k, \quad (2)$$

where the beamformers \mathbf{w}_{jk} have unit norms ($\|\mathbf{w}_{jk}\| = 1$) and p_{jk} represents the power allocated for transmission to MS_k . The transmitter BS_j is subject to an average transmit power constraint of P_j ; that is, $\mathbb{E}\{\|\mathbf{x}_j\|^2\} = \sum_{k=1}^2 p_{jk} \leq P_j$. When the receivers treat the co-user transmissions as noise, the instantaneous SINR of MS_k is

$$\text{SINR}_k = \frac{|\sum_{j=1}^2 \sqrt{p_{jk}} \mathbf{h}_{jk}^H \mathbf{w}_{jk}|^2}{|\sum_{j=1}^2 \sqrt{p_{jk}} \mathbf{h}_{jk}^H \mathbf{w}_{jk}|^2 + \sigma^2}, \quad \text{for } k = 1, 2. \quad (3)$$

Under the given assumptions, the maximal achievable instantaneous transmission rate at MS_k is $R_k = \log_2(1 + \text{SINR}_k)$.

Observe that the main difference between the scenario at hand and the MISO broadcast channel (BC) is that in the latter case *all* transmit antennas are controlled by a central unit with CSI of *all* links and a power constraint for *all* antennas.

B. Spatial Channel Correlation

Next, we characterize the correlation matrices, \mathbf{Q}_{jk} , from a spatial point of view. This will be useful in the precoding design with statistical CSI. Depending on antenna distance and amount of scattering, the correlation of channels from transmit antennas to the receiver will vary; large antenna spacing and rich scattering correspond to low spatial correlation, and vice versa. High correlation translates into a large eigenvalue spread in \mathbf{Q}_{jk} , and low correlation to almost identical eigenvalues. The existence of spatial correlation has been shown experimentally in both outdoor [11] and indoor [12] scenarios.

Similar to [13], we partition the eigenvalue decomposition $\mathbf{Q}_{jk} = \mathbf{U}_{jk} \mathbf{\Lambda}_{jk} \mathbf{U}_{jk}^H$ of the correlation matrix \mathbf{Q}_{jk} in a certain way, based on the size of the different eigenvalues. Assume that the eigenvalues in the diagonal matrix $\mathbf{\Lambda}_{jk}$ are ordered decreasingly with the corresponding eigenvectors as columns of the unitary matrix \mathbf{U}_{jk} . Then, we partition \mathbf{U}_{jk} as

$$\mathbf{U}_{jk} = [\mathbf{U}_{jk}^{(D)} \quad \mathbf{U}_{jk}^{(0)}], \quad (4)$$

where $\mathbf{U}_{jk}^{(D)} \in \mathbb{C}^{N_t \times N_d}$ spans the subspace associated with the N_d largest and *dominating* eigenvalues. The signal and interference allocated to these eigendirections will clearly influence the SINR. Hence, data transmission should take place in $\mathbf{U}_{jk}^{(D)}$, while interference should be avoided. The remaining eigenvectors, in $\mathbf{U}_{jk}^{(0)} \in \mathbb{C}^{N_t \times N_t - N_d}$, span a subspace with an average gain close to *zero*, and hence the interference is negligible. The size of the parameter N_d depends strongly on the amount of spatial correlation, and can be just a fraction of N_t in an outdoor cellular scenario. In completely uncorrelated environments, the partitioning can be ignored since $N_d = N_t$.

III. CHARACTERIZATION OF THE PARETO BOUNDARY

In this section, we analyze the achievable rate region for the scenario at hand, which will provide a precoding structure that is used for practical precoding design in Section IV. Since the receivers are treating co-channel interference as noise (i.e., not attempting to decode and subtract the interference), the achievable rate region will in general be smaller than the capacity region. This assumption is however important to

achieve a simple receiver structure. In the case of instantaneous CSI, we define the achievable rate region as

$$\mathcal{R}_{\text{instant}} = \bigcup_{\substack{\mathbf{w}_{jk}, \|\mathbf{w}_{jk}\|=1 \\ p_{jk} \geq 0, \sum_k p_{jk} \leq P_j}} (R_1, R_2), \quad (5)$$

while in the case of statistical CSI we define the achievable expected rate region as

$$\mathcal{R}_{\text{statistic}} = \bigcup_{\substack{\mathbf{w}_{jk}, \|\mathbf{w}_{jk}\|=1 \\ p_{jk} \geq 0, \sum_k p_{jk} \leq P_j}} (\mathbb{E}\{R_1\}, \mathbb{E}\{R_2\}). \quad (6)$$

Observe that the regions, \mathcal{R} , are independent of the amount of CSI available at the transmitters; the assumption of local CSI does however affect which rate tuples within those regions can be reached in practice by selection of \mathbf{w}_{jk} and p_{jk} . The outer boundary of \mathcal{R} is called the Pareto boundary. The rate tuples on this boundary are Pareto optimal, which means that the rate achieved by MS_k cannot be increased without decreasing the rate of $\text{MS}_{\bar{k}}$. We have the following definition (for instantaneous CSI):

Definition 1. Consider all achievable rate tuples (R_1, R_2) . The Pareto boundary consists of all such tuples for which there is no *other* achievable tuple (S_1, S_2) with $S_1 \geq R_1$ and $S_2 \geq R_2$.

The corresponding definition in the case of statistical CSI is achieved by replacing all rates by their expectations.

Next, we will parameterize the Pareto boundary by showing that beamformers, \mathbf{w}_{jk} , that can be used to attain the boundary lie in a certain subspace defined using only local CSI. It will also be shown under which conditions full transmit power ($\sum_k p_{jk} = P_j$) should be used at both base stations. We begin with the case of instantaneous CSI and then we derive similar results in spatially correlated systems with statistical CSI.

A. Characterization with Instantaneous CSI

For the MISO IC and BC with instantaneous CSI, it is known that rate tuples on the Pareto boundary can only be achieved by beamformers that are linear combinations of MRT and ZF [8]. In Theorem 1 below, we show a similar result for the scenario at hand, for which the achievable rate region will in general be larger (since both base stations know all the data and can transmit simultaneously to both users).

Theorem 1. Each rate tuple (R_1, R_2) on the Pareto boundary can be achieved by beamforming vectors \mathbf{w}_{jk} that fulfill

$$\mathbf{w}_{jk} \in \text{span}(\mathbf{h}_{jk}, \mathbf{\Pi}_{\mathbf{h}_{j\bar{k}}}^\perp \mathbf{h}_{jk}) \quad \text{for all } j, k. \quad (7)$$

If $\frac{\mathbf{h}_{j1}}{\|\mathbf{h}_{j1}\|} \neq \frac{\mathbf{h}_{j2}}{\|\mathbf{h}_{j2}\|}$ for some j , then BS_j needs to use full power (i.e., $\sum_k p_{jk} = P_j$) and (7) becomes a necessary condition on \mathbf{w}_{jk} and $\mathbf{w}_{j\bar{k}}$ to attain rate tuples on the Pareto boundary.

Proof: The proof follows along the same lines as the proof of Corollary 1 in [8], but requires some additional considerations for the power allocation and beamforming design and is omitted due to the space limitation. ■

What the above theorem implies is that to attain rate tuples on the Pareto boundary, both base stations are required

to use full transmit power (except in a special case with zero probability) and use beamforming vectors that are linear combinations of the following straightforward extensions of MRT and ZF to (potentially) distributed precoding systems:

Definition 2 (Distributed MRT). $\mathbf{w}_{jk}^{(\text{MRT})} = \frac{\mathbf{h}_{jk}}{\|\mathbf{h}_{jk}\|}$ for all j, k .

Definition 3 (Distributed ZF). $\mathbf{w}_{jk}^{(\text{ZF})} = \frac{\mathbf{\Pi}_{\mathbf{h}_{j\bar{k}}}^\perp \mathbf{h}_{jk}}{\|\mathbf{\Pi}_{\mathbf{h}_{j\bar{k}}}^\perp \mathbf{h}_{jk}\|}$ for all j, k .

Note that these two approaches can be seen as the extremes in beamforming design, since they maximize the received signal power and minimize the co-user interference, respectively.

In terms of Definition 2 and 3, Theorem 1 implies that all beamforming vectors of importance can be expressed as

$$\mathbf{w}_{jk} = c_{jk} \mathbf{w}_{jk}^{(\text{MRT})} + d_{jk} \mathbf{w}_{jk}^{(\text{ZF})}, \quad \text{for all } j, k \quad (8)$$

for some complex-valued coefficients c_{jk}, d_{jk} (such that $\|\mathbf{w}_{jk}\| = 1$). This reduces the complexity of finding good beamforming vectors considerably (since now the vectors we are looking for each lie in a two-dimensional subspace), especially if the number of antennas N_t is large. In the special case when BS_j is not transmitting to MS_k (i.e., $p_{jk} = 0$), one can show that it is enough to have positive real-valued coefficients c_{jk}, d_{jk} , which further decreases the complexity. This is for example the case for the MISO IC [8].

An important observation for the following sections is that the two vectors in (8) are defined using only local CSI, which motivates the distributed precoding design in Section IV. Apart from selecting beamforming vectors, it is also necessary to perform optimal power allocation to attain the Pareto boundary. Some heuristics on how to perform efficient power allocation using only local CSI are given in Section IV.

B. Characterization with Statistical CSI

Next, we derive a parametrization of the Pareto boundary similar to that in Theorem 1, but for the case of statistical CSI. As shown in [14] for the MISO IC, a parametrization can be found when the correlation matrices are rank deficient. This is however rarely the case in practice, therefore we concentrate on spatially correlated channels and use the eigenvector partitioning in (4). By neglecting the statistical power in $\mathbf{U}_{jk}^{(0)}$, we achieve the following approximate parametrization.

Theorem 2. For each expected rate tuple $(\mathbb{E}\{R_1\}, \mathbb{E}\{R_2\})$ on the Pareto boundary, there is another achievable expected rate tuple $(\mathbb{E}\{\tilde{R}_1\}, \mathbb{E}\{\tilde{R}_2\})$ that fulfills $\mathbb{E}\{\tilde{R}_k\} \approx \mathbb{E}\{R_k\}$, for $k = 1, 2$, and that can be reached by

$$\mathbf{w}_{jk} \in \text{span}([\mathbf{U}_{jk}^{(D)} \ \mathbf{U}_{j\bar{k}}^{(D)}]) = \text{span}([\mathbf{U}_{jk}^{(D)} \ \mathbf{\Pi}_{\mathbf{U}_{j\bar{k}}^{(D)}}^\perp \mathbf{U}_{j\bar{k}}^{(D)}]), \quad (9)$$

for some power allocation p_{jk} , for all j, k . The equality holds with probability one (if $\text{span}(\mathbf{U}_{jk}^{(D)}) = \text{span}(\mathbf{\Pi}_{\mathbf{U}_{j\bar{k}}^{(D)}}^\perp \mathbf{U}_{j\bar{k}}^{(D)})$).

If $\text{span}([\mathbf{U}_{jk}^{(D)} \ \mathbf{U}_{j\bar{k}}^{(D)}]) \neq \mathbb{C}^{N_t}$ for some j , then BS_j can reach $(\mathbb{E}\{\tilde{R}_1\}, \mathbb{E}\{\tilde{R}_2\})$ using full power (i.e., $\sum_k p_{jk} = P_j$).

Proof: The theorem is proved in a similar way as Theorem 1, by ignoring the eigenvalues in $\mathbf{U}_{jk}^{(0)}$ for all j, k . ■

Observe that all approximation operators in Theorem 2 are due to neglecting power in the eigenspaces $\mathbf{U}_{jk}^{(0)}$, for all j, k . If the eigenvalues in these subspaces are identically zero, all approximation operators can be replaced by equalities.

The similarities between the parametrizations with instantaneous and statistical CSI are obvious when (7) and (9) are compared. The result with statistical CSI is however weaker, which is natural since each channel vector belongs (approximately) to a subspace of rank N_d while the channels with instantaneous CSI are known vectors (i.e., rank one). The common interpretation is that all interesting beamforming vectors are linear combinations of eigendirections of the local statistics that have non-negligible power in the links to either of the two users. In the special case of $N_d = 1$, the results of Theorem 2 are essentially the same as in Theorem 1.

IV. DISTRIBUTED PRECODING DESIGN WITH LOCAL CSI

In this section, we discuss beamforming design and power allocation based on local CSI. First, we briefly review the concepts of the layered virtual SINR framework proposed in [10] for instantaneous CSI. Then, based on the parametrizations of the Pareto boundary in Theorem 1, we justify the layered virtual SINR approach as a way to determine an optimized distributed precoding combination of MRT and ZF that satisfies the parametrization. Finally, we will extend the method to the case of local statistical CSI using Theorem 2.

From Theorem 1 and 2 we know that both base stations should use full transmit power P_j (almost surely). A remaining issue is how to split the power between the two users. To simplify the notation, we introduce the splitting parameters $0 \leq \alpha_j \leq 1$, for which the power splitting at BS $_j$ becomes

$$p_{j1} = P_j \alpha_j \quad \text{and} \quad p_{j2} = P_j (1 - \alpha_j). \quad (10)$$

Observe that $\alpha_1 = 1, \alpha_2 = 0$ corresponds to the standard MISO IC, while $\alpha_1 = \alpha_2 = 0$ or $= 1$ means that only one of the receivers gets all the attention. Some heuristics on how to determine good α_j will be given in this section.

A. Precoding with Local Instantaneous CSI

The purpose of distributed precoding design is to find beamforming vectors and power splitting that achieve performance close to the Pareto boundary, but is only based on local instantaneous CSI. As shown by Theorem 1, the two extremes in the precoding design with local instantaneous CSI are distributed MRT and ZF (see Definition 2 and 3), and these satisfy the requirement of only depending on local CSI. More formally, Theorem 1 states that rate tuples on the Pareto boundary can only be achieved by beamforming vectors that are linear combinations of distributed MRT and ZF. In order to derive an efficient linear combination we propose to use the framework of virtual SINR, which has been shown to attain the Pareto boundary of MISO interference channels [1]. Virtual SINR maximizing beamformers are obtained (see [10]) as

$$\mathbf{w}_{jk}^{(\text{VSINR})} = \arg \max_{\|\mathbf{w}\|^2=1} \frac{|\mathbf{h}_{jk}^H \mathbf{w}|^2}{\frac{\sigma^2}{p_{jk}} + |\mathbf{h}_{j\bar{k}}^H \mathbf{w}|^2}. \quad (11)$$

This Rayleigh quotient is solved by straightforward eigenvalue techniques and the solution is a linear combination of the MRT and ZF vectors in (8) with real-valued positive coefficients [1]. This means that signals arriving at a given user from different base stations will do so constructively. Moreover, by its very definition (cf. (11)), maximizing a virtual SINR effectively balances between the useful signal power at a target user and the interference power generated at others; assuming a judicious assignment of the powers p_{jk} , for $j, k \in \{1, 2\}$, this should lead to good performance, under the given local CSI.

Finally, regarding the power splitting between the users, an intuitive heuristic is to allocate more power to the one with the stronger instantaneous channel gain [10]. This can be formalized as

$$\alpha_j^{\text{instant}} = \frac{\|\mathbf{h}_{j1}\|^2}{\sum_{k=1}^2 \|\mathbf{h}_{jk}\|^2}. \quad (12)$$

B. Precoding with Local Statistical CSI

As seen by comparing Theorem 1 and 2, the Pareto optimal precoding design with statistical CSI resembles that of instantaneous CSI. Next, we will show the corresponding relationships for the distributed precoding design. In analogy with the distributed MRT and ZF approaches with instantaneous CSI in Definition 2 and 3, we propose extensions assuming local statistical CSI. The straightforward generalization of MRT is to use the dominating eigenvector of \mathbf{Q}_{jk} as beamformer. We will denote the normalized eigenvector of the largest eigenvalue of \mathbf{Q}_{jk} by $\mathbf{u}_{jk}^{(D)}$. The generalization of ZF is to maximize the average received signal power under the condition that the beamformer lies in the negligible eigen-subspace $\mathbf{U}_{j\bar{k}}^{(0)}$ of the co-user. We summarize this as:

Definition 4 (Distributed Generalized MRT). In this approach, $\mathbf{w}_{jk}^{(\text{G-MRT})} = \mathbf{u}_{jk}^{(D)}$ for all j, k .

Definition 5 (Distributed Generalized ZF). In this approach, $\mathbf{w}_{jk}^{(\text{G-ZF})} = \mathbf{U}_{j\bar{k}}^{(0)} \mathbf{v}_{jk}$, where \mathbf{v}_{jk} is the normalized dominating eigenvector of $(\mathbf{U}_{j\bar{k}}^{(0)})^H \mathbf{Q}_{jk} \mathbf{U}_{j\bar{k}}^{(0)}$, for all j, k .

As in the case of instantaneous CSI, the virtual SINR framework can be applied to balance the signal and interference powers. We propose the following novel extension where the Rayleigh quotient represents maximization of a lower bound on the SINR, assuming the channels of both users are uncorrelated (and using that $\mathbb{E}\{|\mathbf{h}_{jk}^H \mathbf{w}|^2\} = \mathbf{w}^H \mathbf{Q}_{jk} \mathbf{w}$):

$$\mathbf{w}_{jk}^{(\text{LVSINR})} = \arg \max_{\|\mathbf{w}\|^2=1} \frac{\mathbf{w}^H \mathbf{Q}_{jk} \mathbf{w}}{\frac{\sigma^2}{p_{jk}} + \mathbf{w}^H \mathbf{Q}_{j\bar{k}} \mathbf{w}}. \quad (13)$$

Unlike the instantaneous local CSI case, the resulting beamformers cannot be written out as linear combinations of the generalized MRT and ZF vectors, unless only the most dominant eigenvectors in each correlation matrix are considered. However, this formulation does result in vectors that satisfy the condition for (approximative) Pareto optimality specified in Theorem 2. Note also that unlike the instantaneous CSI case, we cannot guarantee the coherent arrival of useful signals at a given user (unless it feeds back a phase correction to its

'main' base station). On average, an increase in signal power will however increase the rate (see Corollary 1 below).

Next, we will derive explicit expressions for the expected rates for a given power allocation and set of beamforming vectors. These results are important both for computational reasons and to enable analysis of the impact of the precoding design on the average performance. Recall from (3) that the SINR at MS_k can be expressed as

$$\text{SINR}_k = \frac{|a_{kk}|^2}{|a_{\bar{k}k}|^2 + \sigma^2}, \quad (14)$$

where signal and interference coefficients are distributed as

$$a_{lk} \in \mathcal{CN}(0, \sum_{j=1}^2 p_{jl} \mathbf{w}_{jl}^H \mathbf{Q}_{jk} \mathbf{w}_{jl}) \quad \text{for } l = k, \bar{k}. \quad (15)$$

As a result of the precoding design, the signal and interference coefficients will in general be correlated with the factor

$$\varrho_k = \mathbb{E}\{a_{kk}^H a_{\bar{k}k}\} = \sum_{j=1}^2 \sqrt{p_{jk} p_{j\bar{k}}} \mathbf{w}_{jk}^H \mathbf{Q}_{j\bar{k}} \mathbf{w}_{j\bar{k}}. \quad (16)$$

Intuitively, the expected rate should decrease with ϱ_k , since the probability of realizations with strong signal and weak interference reduces. The following theorem gives an explicit expression of the expected rate and its corollary clarify its dependence on signal coefficients a_{kk} and correlation ϱ_k .

Theorem 3. Consider the rate $R = \log_2(1 + \frac{|a_1|^2}{|a_2|^2 + \sigma^2})$, where $a_l \in \mathcal{CN}(0, \lambda_l)$ with $\lambda_l > 0$ for $l = 1, 2$. The coefficients a_l are correlated with $\varrho = \mathbb{E}\{a_1^H a_2\}$. Then, the expected rate is

$$\mathbb{E}\{R\} = \frac{\mu_1 E_1(\frac{\sigma^2}{\mu_1}) e^{\sigma^2/\mu_1} - \mu_2 E_1(\frac{\sigma^2}{\mu_2}) e^{\sigma^2/\mu_2}}{(\mu_1 - \mu_2) \log(2)} - \frac{E_1(\frac{\sigma^2}{\lambda_2}) e^{\sigma^2/\lambda_2}}{\log(2)}, \quad (17)$$

where $\mu_l = \frac{\lambda_1 + \lambda_2 \pm \sqrt{(\lambda_1 - \lambda_2)^2 + 4|\varrho|^2}}{2}$ has different signs in front of the square root for $l = 1$ and $l = 2$. The exponential integral is denoted $E_1(z) = \int_z^\infty \frac{e^{-u}}{u} du$. In the special case of $\lambda_1 = \lambda_2$ and $\varrho = 0$, we instead have

$$\mathbb{E}\{R\} = \frac{1 - \frac{\sigma^2}{\lambda_1} E_1(\frac{\sigma^2}{\lambda_1}) e^{\sigma^2/\lambda_1}}{\log(2)}. \quad (18)$$

Proof: Observe that $|a_1|^2 + |a_2|^2$ has the same distribution as $|z_1|^2 + |z_2|^2$, with the whitened variables $z_1 \in \mathcal{CN}(0, \mu_1)$ and $z_2 \in \mathcal{CN}(0, \mu_2)$, and μ_1, μ_2 being the eigenvalues of the correlation matrix $\mathbf{Q} = \begin{bmatrix} \lambda_1 & \varrho \\ \varrho^* & \lambda_2 \end{bmatrix}$. Then, the expected rate is $\mathbb{E}\{R\} = \mathbb{E}\{\log_2(|z_1|^2 + |z_2|^2 + \sigma^2)\} - \mathbb{E}\{\log_2(|b|^2 + \sigma^2)\}$ and each term can be calculated using [15, Theorem 2]. ■

Corollary 1. The expected rate in (17) is increasing with decreasing magnitude $|\varrho|$ of the correlation factor. It is also an increasing function in the signal variance λ_1 .

Proof: The first part follows since the first term in (17) is Schur-concave with respect to $[\mu_1 \ \mu_2]^T$ and since increasing $|\varrho|$ will increase the difference between μ_1 and μ_2 . The second part follows since both μ_1 and μ_2 increases with λ_1 . ■

There are two implications of the corollary. Firstly, each beamformer \mathbf{w}_{jk} can be replaced with $\mathbf{w}_{jk} e^{i\phi_{jk}}$ without

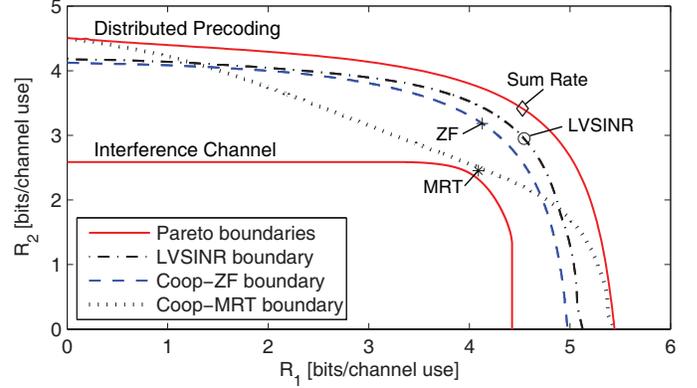


Fig. 2. Achievable rate regions with different precoding and power allocation approaches for a random realization with $\mathbf{Q}_{11} = \mathbf{Q}_{22} = \mathbf{I}$, $\mathbf{Q}_{12} = \mathbf{Q}_{21} = 0.5\mathbf{I}$, and 5 dB. The sum rate point and the points achieved by MRT, ZF, and LVSINR with the proposed heuristic power splitting are given for comparison.

affecting the signal and interference coefficients in (15). Thus, the phases, ϕ_{jk} , should be used to minimize the correlation in (16) and thereby improve the performance. Secondly, increasing the average signal power improves performance, even if the signal and interference powers are clearly correlated.

Finally, we want to derive a heuristic power splitting and we design it based on generalized ZF to avoid the problem that an increase in both signal power and correlation factor affects the performance in opposite ways (see Corollary 1). Since we cannot guarantee the signals from multiple transmitters to interact constructively using only statistical CSI, an intuitive heuristic is to allocate all power to the user that benefits the most from the increased average gain:

$$\alpha_j^{\text{statistic}} = \begin{cases} 1, & \mathbf{w}_{j1}^H \mathbf{Q}_{j1} \mathbf{w}_{j1} > \mathbf{w}_{j2}^H \mathbf{Q}_{j2} \mathbf{w}_{j2}, \\ 0, & \text{otherwise,} \end{cases} \quad (19)$$

with beamformers \mathbf{w}_{jk} designed as in Definition 5. With this power splitting, the transmitters will either serve one user each or they will concentrate on a single user. The main difference from the MISO IC in [14] is the distributed selection of which receiver that should be served by which transmitter.

V. NUMERICAL EXAMPLES

The performance of the proposed precoding and power splitting schemes are evaluated numerically. For instantaneous local CSI, we illustrate the performance in an uncorrelated system with $N_t = 3$, $\mathbf{Q}_{11} = \mathbf{Q}_{22} = \mathbf{I}$, $\mathbf{Q}_{12} = \mathbf{Q}_{21} = \beta\mathbf{I}$, where β is the average cross link power that will be varied.

In Fig. 2, the Pareto boundary of the distributed precoding scenario is given for a randomly chosen realization with $\beta = 0.5$ and an SNR, P_j/σ^2 , of 5 dB. As a comparison, we give the Pareto boundary of the MISO IC [8] and the outer boundaries of achievable rate regions for the precoding approaches in Section IV (distributed MRT, distributed ZF, and layered VSINR). The rate tuples achieved with the heuristic power splitting in (12) and the sum rate maximizing point are given as references. For the specific realization there is a clear performance gain of allowing distributed precoding compared to forcing each transmitter to only communicate with a single

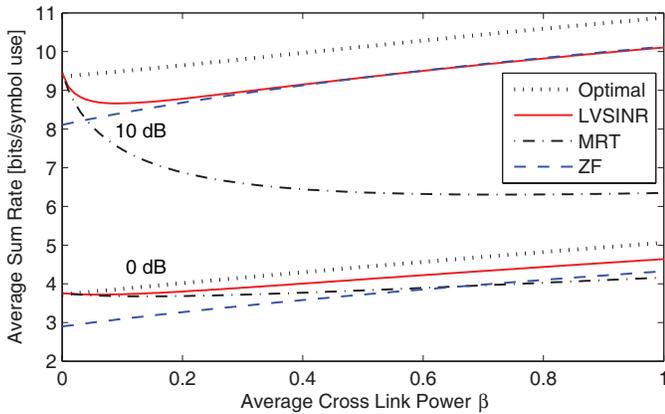


Fig. 3. Performance of different precoding designs with instantaneous CSI, as function of the average cross link power: $\mathbf{Q}_{11} = \mathbf{Q}_{22} = \mathbf{I}$, $\mathbf{Q}_{12} = \mathbf{Q}_{21} = \beta\mathbf{I}$.

receiver. As expected, MRT is useful when the rate of only one of the users should be maximized; on the other hand, ZF and layered VSINR are quite close to the optimal sum rate point. The heuristic power splitting provides performance close to the boundary of each achievable rate region.

In Fig. 3, the average sum rate over different realizations is given with optimal precoding and power allocation, and for distributed MRT, distributed ZF, and layered VSINR with heuristic power splitting. The performance is shown for varying cross link power β and at an SNR of 0 or 10 dB. From Fig. 3, we observe that MRT works well at low SNR and/or low cross link power, while ZF is better at high SNR and cross link power. For the considered scenario, the layered VSINR is the clearly better choice since it provides better performance at low SNR and combines the benefits of MRT and ZF at high SNR. Of course, due to the distributed nature of all these schemes there is always some loss with respect to the performance with sum rate maximizing beamformers (except when the cross-link is very weak and MRT is optimal).

Finally, expected sum rates are shown in Fig. 4 in the case of local statistical CSI. We consider base stations located in the opposite corners of a unit square with two uniformly distributed users within the square. Uniform circular arrays with $N_t = 5$ and half a wavelength separation are used. The channels are spatially correlated with varying angular spread (as seen from a base station). The power decay is $1/r^4$, where r is the distance, and the SNR is either 0 or 10 dB (for $r = 1$). The behavior in Fig. 4 is similar to that for instantaneous CSI, which means that generalized MRT is good at low SNR and generalized ZF at high SNR. The layered VSINR approach provides a clearly better performance at both SNRs. The overall performance decreases with angular spread (low angular spread means high correlation). The upper bound (UB) represents optimal beamforming and power allocation.

VI. CONCLUSION

In this paper, distributed precoding was considered in a multicell environment where the data is shared among the base stations, but only local CSI is available. The beamforming vectors used to attain the outer boundary of the achievable rate region were characterized to lie in certain subspaces

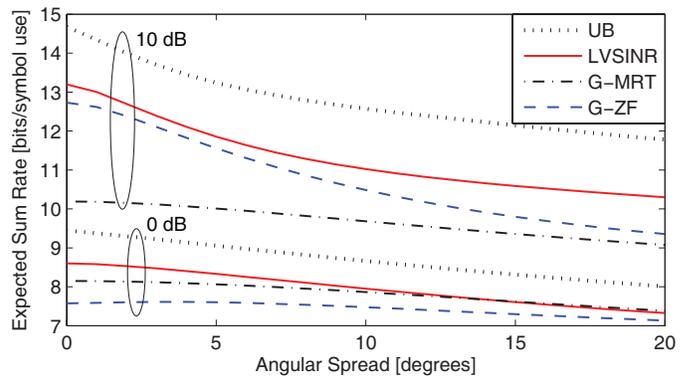


Fig. 4. Expected performance with different precoding designs and statistical CSI, as function of the angular spread. The users are uniformly distributed in a square with the base stations in opposite corners.

based on generalizations of MRT and ZF, assuming either instantaneous or statistical CSI. The recently proposed layered virtual SINR precoding design, and its statistical extension introduced herein, was shown to belong to these characterized subspaces and to provide good performance. Finally, heuristics on the power allocation was given and evaluated numerically.

REFERENCES

- [1] R. Zakhour and D. Gesbert, "Coordination on the MISO interference channel using the virtual SINR framework," in *Proc. International ITG Workshop on Smart Antennas*, 2009.
- [2] B. O. Lee, H. W. Je, I. Sohn, O.-S. Shin, and K. B. Lee, "Interference-aware decentralized precoding for multicell MIMO TDD systems," in *Proc. IEEE GLOBECOM'08*, 2008.
- [3] D. Gesbert, M. Kountouris, R. Heath, C.-B. Chae, and T. Sälzer, "Shifting the MIMO paradigm," *IEEE Signal Process. Mag.*, vol. 24, pp. 36–46, 2007.
- [4] M. Karakayali, G. Foschini, and R. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *IEEE Wireless Commun. Mag.*, vol. 13, pp. 56–61, 2006.
- [5] B. L. Ng, J. Evans, S. Hanly, and D. Aktas, "Distributed downlink beamforming with cooperative base stations," *IEEE Trans. Inf. Theory*, vol. 54, pp. 5491–5499, 2008.
- [6] O. Simeone, O. Somekh, H. V. Poor, and S. Shamai, "Distributed MIMO in multi-cell wireless systems via finite-capacity links," in *Proc. IEEE ISCCSP'08*, 2008.
- [7] E. Larsson and E. Jorswieck, "Competition versus cooperation on the MISO interference channel," *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 1059–1069, 2008.
- [8] E. Jorswieck, E. Larsson, and D. Danev, "Complete characterization of the Pareto boundary for the MISO interference channel," *IEEE Trans. Signal Process.*, vol. 56, pp. 5292–5296, 2008.
- [9] R. Zakhour, Z. Ho, and D. Gesbert, "Distributed beamforming coordination in multicell MIMO channels," in *Proc. IEEE VTC'09-Spring*.
- [10] R. Zakhour and D. Gesbert, "Distributed multicell-MIMO precoding using the layered virtual SINR framework," in *IEEE Trans. Wireless Commun.*, submitted.
- [11] D. Chizhik, J. Ling, P. Wolniansky, R. Valenzuela, N. Costa, and K. Huber, "Multiple-input-multiple-output measurements and modeling in Manhattan," *IEEE J. Sel. Areas Commun.*, vol. 21, pp. 321–331, 2003.
- [12] K. Yu, M. Bengtsson, B. Ottersten, D. McNamara, P. Karlsson, and M. Beach, "Modeling of wideband MIMO radio channels based on NLOS indoor measurements," *IEEE Trans. Veh. Technol.*, vol. 53, pp. 655–665, 2004.
- [13] D. Hammarwall, M. Bengtsson, and B. Ottersten, "Utilizing the spatial information provided by channel norm feedback in SDMA systems," *IEEE Trans. Signal Process.*, vol. 56, pp. 3278–3293, 2008.
- [14] J. Lindblom, E. Larsson, and E. Jorswieck, "Parameterization of the MISO interference channel with transmit beamforming and partial channel state information," in *Proc. Asilomar'08*, 2008.
- [15] E. Björnson and B. Ottersten, "Post-user-selection quantization and estimation of correlated Frobenius and spectral channel norms," in *Proc. IEEE PIMRC'08*, 2008.