On Carrier-Cooperation in Parallel Gaussian MIMO Relay Channels with Partial Decode-and-Forward

Christoph Hellings, and Wolfgang Utschick

48th Asilomar Conference on Signals, Systems and Computers
(ACSSC 2014)

Pacific Grove, CA, USA, 2nd - 5th November, 2014

©2014 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Technische Universität München
Fachgebiet Methoden der Signalverarbeitung
http://www.msv.ei.tum.de

On Carrier-Cooperation in Parallel Gaussian MIMO Relay Channels with Partial Decode-and-Forward

Christoph Hellings and Wolfgang Utschick
Associate Institute for Signal Processing, Technische Universität München, 80290 Munich, Germany
Telephone: +49 89 289-28516, Fax: +49 89 289-28504, Email: {hellings,utschick}@tum.de

***Abstract*—It is known that parallel relay channels are not separable, i.e., the capacity with joint processing of the subchannels can be higher than the sum of the individual capacities. The same holds for the data rates achievable using partial decode-and-forward in parallel Gaussian MIMO relay channels. However, in this paper, we show that it is sufficient to allow the relay to remap information from one subchannel to another between the decoding and the re-encoding. A carrier-cooperative transmission in the sense of spreading transmit symbols over several subchannels does not bring advantages in terms of achievable rate.**

I. Introduction

Since the capacity of relay networks is still an open problem except for special cases, many researchers have focused on deriving upper bounds such as the cut-set bound [1] and achievable schemes such as amplify-and-forward, compress-and-forward, and decode-and-forward [1], [2], [3, Ch. 9].

In the decode-and-forward protocol, the relay has to decode the complete message to then transmit it to the destination coherently with the source node. This strategy can be capacity-achieving if the channel between the source and the relay is strong [2], but otherwise, the source-relay link can become a bottleneck [2], [3, Section 9.2.1]. In this case, schemes where only a part of the message is decoded by the relay can be superior. Such a generalization of decode-and-forward is given by the partial decode-and-forward (PDF) scheme [2], [4]–[8], [3, Section 9.4.1].

In this paper, we consider the application of PDF to a set of parallel Gaussian multiple-input multiple-output (MIMO) relay channels. With the term *parallel* we refer to a setting with a single-relay node that supports the communication between a source and a destination over a set of parallel orthogonal resources (e.g., carriers). Note that is in accordance with the nomenclature of [9], [10], but differs from the nomenclature in [11], where parallel relay networks contain multiple relays.

In a communication system with such parallel orthogonal resources, data transmission can be performed separately on each resource or jointly across the resources. In the case of a single-hop transmission, the terms *joint coding* and *carrier-cooperative transmission* can be used interchangeably (see, e.g., [12], [13]). For multihop systems, we propose to make the following distinction.

- In a system with *separate coding*, messages are split into chunks that are then processed on a per-carrier basis. On the other hand, if a joint processing across carriers takes place at some encoding or decoding stage, we speak of *joint coding*.
- We say that *carrier-cooperative transmission* takes place in a time slot if the signals on the various carriers are statistically dependent in this time slot. This can be interpreted as spreading a transmit symbol over several carriers. If this is not the case, *carrier-noncooperative transmission* is performed.

The notion of a time slot (or block) is necessary since multihop transmission often relies on block coding schemes. In particular, the PDF rate can be achieved with a block-Markov coding scheme [3, Section 9.4.1]. The importance of above distinction will become clear in the interpretation that is given in Section V. Note that carrier-cooperative transmission requires joint coding in order to form a sensible transmit strategy, but the converse is not true.

It is known that separate coding is in general not capacity-achieving in parallel relay channels [10]. Imagine a two-carrier system where the source-relay channel is zero on the first carrier while the relay-destination channel is zero on the second carrier. Then, the relay can only be helpful if it can forward information on a carrier different from the one on which it has received the information.

Clearly, this observation also applies to the PDF scheme, i.e., joint coding can be necessary to achieve the optimal PDF rate. However, we show in this paper that carrier-cooperative transmission is not needed in the PDF scheme.

For parallel single-antenna relay channels, PDF was considered in [9], but mainly for the case where the relay and the source are not able to transmit in a coherent manner. This simplifies the implementation, but degrades the performance. For the case with coherent transmission, as assumed in our work, a complete solution of the PDF rate optimization was not obtained in [9] due to the nonconvexity of the problem.

In the recent work [14], it was proven that the optimal PDF rate in Gaussian MIMO relay channels can be achieved with Gaussian input signals. Moreover, it was shown in [8] that among all possible Gaussian inputs, circular symmetric ones achieve optimal performance. These results generalize to parallel MIMO relay channels since results obtained for single-carrier MIMO systems can be extended to the multicarrier case by introducing an equivalent single-carrier system with block-diagonal channel matrices (see [12], [13] and Section II). Therefore, we can assume circularly symmetric Gaussian input signals throughout the paper, and the optimal transmit strategy can be characterized by the joint covariance matrix of the source and relay inputs.

However, the existing literature on the optimization of PDF rates in Gaussian MIMO relay channels [5]–[7] has not yet overcome the difficulty of the nonconvexity of the arising optimization problems. Therefore, we cannot simply obtain the globally optimal covariance matrix and verify whether it corresponds to carrier-cooperative or carrier-noncooperative transmission.

Instead, we adopt the proof technique that was used in [8] to show optimality of circular symmetric transmit signals. To adapt this technique to multicarrier systems, we first propose a new parametrization of the involved covariance matrices in Section III. The proof of the main result is then provided in Section IV, and a discussion follows in Section V. It turns out that we need to show the optimality of carrier-noncooperative transmission in parallel MIMO broadcast channels with a certain type of shaping constraints as an ingredient for this proof. Therefore, this paper contains an excursus to parallel MIMO broadcast channels in Appendix A.

Notation: We use $\mathbf{0}$ for the zero matrix, $\mathbf{I}_N$ for the identity matrix of size N, $\bullet^{\mathrm{T}}$ for the transpose, $\bullet^{\mathrm{H}}$ for the conjugate transpose, and $\bullet^{\perp}$ for the orthogonal complement. The notation $\bullet^{\star}$ is used for optimizers and optimal values. The operators $\mathrm{I}(\bullet)$, $\mathrm{h}(\bullet)$, $\mathrm{E}[\bullet]$, and $\mathrm{tr}[\bullet]$ denote mutual information, differential entropy, expected value, and trace, respectively. We use $\boldsymbol{C}_{\boldsymbol{x}}$ for the covariance matrix of $\boldsymbol{x}$. The order relation $\succeq$ has to be understood in the sense of positive-semidefiniteness.

II. System Model and Coding Scheme

We consider data transmission from a source S to a destination D with the help of a relay R, where all nodes have multiple antennas. We collect the channel matrices $\boldsymbol{H}_{ij,c} \in \mathbb{C}^{N_j \times N_i}$ with $i, j \in \{\mathrm{S}, \mathrm{R}, \mathrm{D}\}$ in block-diagonal channel matrices

$$\grave{\boldsymbol{H}}_{ij} = \mathrm{blockdiag}(\boldsymbol{H}_{ij,1}, \ldots, \boldsymbol{H}_{ij,C}) \qquad (1)$$

where $c \in \{1, \ldots, C\}$ denotes the carrier index, and N_i is the number of antennas at node i. Data transmission in parallel Gaussian MIMO relay channels can then be described by

$$\boldsymbol{y}_{\mathrm{R}} = \grave{\boldsymbol{H}}_{\mathrm{SR}} \boldsymbol{x}_{\mathrm{S}} + \boldsymbol{\eta}_{\mathrm{R}} \qquad (2)$$

$$\boldsymbol{y}_{\mathrm{D}} = \grave{\boldsymbol{H}}_{\mathrm{SD}} \boldsymbol{x}_{\mathrm{S}} + \grave{\boldsymbol{H}}_{\mathrm{RD}} \boldsymbol{x}_{\mathrm{R}} + \boldsymbol{\eta}_{\mathrm{D}}. \qquad (3)$$

The system model is visualized in Fig. 1. We assume full-duplex transmission and perfect channel state information.

The noise $\boldsymbol{\eta}_{\mathrm{R}} = [\boldsymbol{\eta}_{\mathrm{R},1}^{\mathrm{T}}, \ldots, \boldsymbol{\eta}_{\mathrm{R},C}^{\mathrm{T}}]^{\mathrm{T}} \sim \mathcal{CN}(\mathbf{0}, \grave{\boldsymbol{C}}_{\boldsymbol{\eta}_{\mathrm{R}}})$ at the relay and the noise $\boldsymbol{\eta}_{\mathrm{D}} = [\boldsymbol{\eta}_{\mathrm{D},1}^{\mathrm{T}}, \ldots, \boldsymbol{\eta}_{\mathrm{D},C}^{\mathrm{T}}]^{\mathrm{T}} \sim \mathcal{CN}(\mathbf{0}, \grave{\boldsymbol{C}}_{\boldsymbol{\eta}_{\mathrm{D}}})$ at the destination are assumed to be independent of each other and independent of the useful signals. Moreover, the noise is assumed to be independent across carriers, i.e., $\grave{\boldsymbol{C}}_{\boldsymbol{\eta}_{\mathrm{R}}}$ and $\grave{\boldsymbol{C}}_{\boldsymbol{\eta}_{\mathrm{D}}}$ are block-diagonal. Throughout the paper, we assume $\grave{\boldsymbol{C}}_{\boldsymbol{\eta}_{\mathrm{R}}} = \mathbf{I}_{CN_{\mathrm{R}}}$ and $\grave{\boldsymbol{C}}_{\boldsymbol{\eta}_{\mathrm{D}}} = \mathbf{I}_{CN_{\mathrm{D}}}$. This is without loss of generality since other cases can be treated by introducing equivalent channels (which are still block-diagonal) after noise whitening.

It is optimal to use jointly circularly symmetric Gaussian transmit signals $\boldsymbol{x}_{\mathrm{S}}$ and $\boldsymbol{x}_{\mathrm{R}}$ at the source and at the relay (see Section I). Consequently, the receive signals $\boldsymbol{y}_{\mathrm{R}}$ and $\boldsymbol{y}_{\mathrm{D}}$ at the relay and at the destination are circularly symmetric Gaussian signals as well.

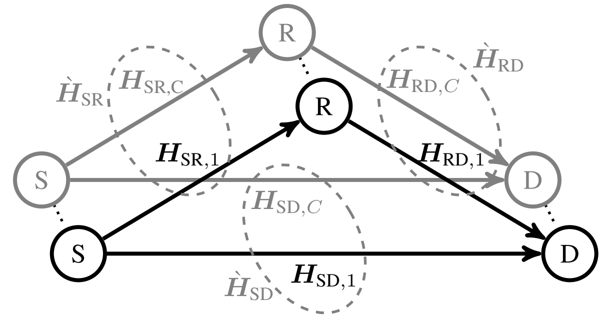

Fig. 1. Illustration of parallel Gaussian MIMO relay channels.

The source transmit signal $\boldsymbol{x}_{\mathrm{S}} = \boldsymbol{u} + \boldsymbol{v}$ is a superposition of a signal $\boldsymbol{u}$, which is decoded and forwarded by the relay, and a signal $\boldsymbol{v}$, which is transmitted without the help of the relay (e.g., [4], [5], [3, Section 9.4.1]). The decode-and-forward signal $\boldsymbol{u}$ is correlated with the relay transmit signal $\boldsymbol{x}_{\mathrm{R}}$, but $\boldsymbol{v}$ is independent of $\boldsymbol{u}$ and $\boldsymbol{x}_{\mathrm{R}}$.

The PDF rate is then given by [4], [3, Section 9.4.1]

$$R = \min\{\underbrace{\mathrm{I}(\boldsymbol{x}_{\mathrm{S}}; \boldsymbol{y}_{\mathrm{D}} | (\boldsymbol{u}, \boldsymbol{x}_{\mathrm{R}})) + \mathrm{I}(\boldsymbol{u}; \boldsymbol{y}_{\mathrm{R}} | \boldsymbol{x}_{\mathrm{R}})}_{R_{\mathrm{A}}};\ \underbrace{\mathrm{I}((\boldsymbol{x}_{\mathrm{S}}, \boldsymbol{x}_{\mathrm{R}}); \boldsymbol{y}_{\mathrm{D}})}_{R_{\mathrm{B}}}\}. \qquad (4)$$

The optimization of this rate with maximum transmit powers P_{S} and P_{R} at the source and the relay, respectively, reads as

$$\max_{p_{\boldsymbol{m}} \in \mathcal{M}} R \quad \text{s.t.} \quad \mathrm{E}\left[\boldsymbol{x}_{\mathrm{S}}^{\mathrm{H}} \boldsymbol{x}_{\mathrm{S}}\right] \leq P_{\mathrm{S}} \text{ and } \mathrm{E}\left[\boldsymbol{x}_{\mathrm{R}}^{\mathrm{H}} \boldsymbol{x}_{\mathrm{R}}\right] \leq P_{\mathrm{R}} \qquad (5)$$

where $\boldsymbol{m} = [\boldsymbol{u}^{\mathrm{T}}, \boldsymbol{x}_{\mathrm{S}}^{\mathrm{T}}, \boldsymbol{x}_{\mathrm{R}}^{\mathrm{T}}]^{\mathrm{T}}$, and $\mathcal{M}$ is the set of all valid probability distributions of $\boldsymbol{m}$ that have the property that $\boldsymbol{u}$–$(\boldsymbol{x}_{\mathrm{S}}, \boldsymbol{x}_{\mathrm{R}})$–$(\boldsymbol{y}_{\mathrm{R}}, \boldsymbol{y}_{\mathrm{D}})$ is a Markov chain.

III. Parametrization of Covariance Matrices

Let $\mathcal{D}_{K \times L}$ denote the projection of a matrix onto the set of block-diagonal matrices with block-size $K \times L$. We define the subspace of block-diagonal complex matrices

$$\mathbb{D}^{N, K \times L} = \{\grave{\boldsymbol{A}} \in \mathbb{C}^{NK \times NL} \mid \mathcal{D}_{K \times L}(\grave{\boldsymbol{A}}) = \grave{\boldsymbol{A}}\} \qquad (6)$$

and the subspace of complex matrices consisting of only off-diagonal blocks

$$\mathbb{O}^{N, K \times L} = \{\breve{\boldsymbol{A}} \in \mathbb{C}^{NK \times NL} \mid \mathcal{D}_{K \times L}(\breve{\boldsymbol{A}}) = \mathbf{0}\}. \qquad (7)$$

From this definition, it follows that $\mathbb{O}^{N, K \times L}$ is the orthogonal complement of $\mathbb{D}^{N, K \times L}$ in $\mathbb{C}^{NK \times NL}$.

Let $\mathbb{H}^{L} \subset \mathbb{C}^{L \times L}$ be the space of Hermitian matrices. It can be easily verified that the subspace of block-diagonal Hermitian matrices $\mathbb{D}^{N, M \times M} \cap \mathbb{H}^{NM}$ and the subspace of Hermitian matrices consisting of only off-diagonal blocks $\mathbb{O}^{N, M \times M} \cap \mathbb{H}^{NM}$ are orthogonal complements in $\mathbb{H}^{NM}$. Therefore, the covariance matrix $\boldsymbol{C}_{\boldsymbol{x}}$ of any random vector $\boldsymbol{x} \in \mathbb{C}^{NM}$ can be uniquely decomposed as

$$\boldsymbol{C}_{\boldsymbol{x}} = \grave{\boldsymbol{C}}_{\boldsymbol{x}} + \breve{\boldsymbol{C}}_{\boldsymbol{x}}, \quad \grave{\boldsymbol{C}}_{\boldsymbol{x}} \in \mathbb{D}^{N, M \times M}, \ \breve{\boldsymbol{C}}_{\boldsymbol{x}} \in \mathbb{O}^{N, M \times M}. \qquad (8)$$

If $\boldsymbol{x}$ represents a signal in a system with N carriers and M dimensions per carrier, the block-diagonal part $\grave{\boldsymbol{C}}_{\boldsymbol{x}}$ describes the power shaping within the carriers while the off-diagonal part $\breve{\boldsymbol{C}}_{\boldsymbol{x}}$ describes the correlation between the carriers.

A meaningful covariance matrix $\boldsymbol{C}_{\boldsymbol{x}}$ is obtained only for those pairs of $\grave{\boldsymbol{C}}_{\boldsymbol{x}} \in \mathbb{D}^{N,M\times M} \cap \mathbb{H}^{NM}$ and $\breve{\boldsymbol{C}}_{\boldsymbol{x}} \in \mathbb{O}^{N,M\times M} \cap \mathbb{H}^{NM}$ that fulfill $\grave{\boldsymbol{C}}_{\boldsymbol{x}} + \breve{\boldsymbol{C}}_{\boldsymbol{x}} \succeq \boldsymbol{0}$ (which implies $\grave{\boldsymbol{C}}_{\boldsymbol{x}} \succeq \boldsymbol{0}$). For convenience, we define

$$\breve{\mathcal{C}}(\grave{\boldsymbol{C}}) = \{\breve{\boldsymbol{C}} \in \mathbb{O}^{N,M\times M} \cap \mathbb{H}^{NM} \mid \grave{\boldsymbol{C}} + \breve{\boldsymbol{C}} \succeq \boldsymbol{0}\} \tag{9}$$

for $\grave{\boldsymbol{C}} \in \mathbb{D}^{N,M\times M} \cap \mathbb{H}^{NM}$.

As a consequence of [15, Appendix], we obtain the first of the following lemmas. The second one is easy to verify.

Lemma 1: For fixed $\grave{\boldsymbol{C}}_{\boldsymbol{x}}$, the entropy of a circularly symmetric Gaussian random vector $\boldsymbol{x}$ is maximized by $\breve{\boldsymbol{C}}_{\boldsymbol{x}} = \boldsymbol{0}$.

Lemma 2:

1) $\grave{\boldsymbol{A}}\grave{\boldsymbol{B}} \in \mathbb{D}^{N,K\times L}$ if $\grave{\boldsymbol{A}} \in \mathbb{D}^{N,K\times M}$ and $\grave{\boldsymbol{B}} \in \mathbb{D}^{N,M\times L}$,
2) $\grave{\boldsymbol{A}}\breve{\boldsymbol{B}} \in \mathbb{O}^{N,K\times L}$ if $\grave{\boldsymbol{A}} \in \mathbb{D}^{N,K\times M}$ and $\breve{\boldsymbol{B}} \in \mathbb{O}^{N,M\times L}$.

IV. Optimality of Carrier-Noncooperative Transmission

We now state and proof the main theorem of this paper. An interpretation is given afterwards.

Theorem 1: In parallel Gaussian MIMO relay channels with partial decode-and-forward, *carrier-noncooperative transmission* is optimal.

Proof of Theorem 1: Using a block-diagonal matrix $\grave{\boldsymbol{A}}$, we decompose $\boldsymbol{u}$ as

$$\boldsymbol{u} = \boldsymbol{q} + \grave{\boldsymbol{A}}\boldsymbol{x}_{\mathrm{R}} \quad \text{such that} \quad \mathcal{D}_{N_{\mathrm{S}}\times N_{\mathrm{R}}}(\mathrm{E}\left[\boldsymbol{q}\boldsymbol{x}_{\mathrm{R}}^{\mathrm{H}}\right]) = \boldsymbol{0} \tag{10}$$

i.e., the correlation between the components of $\boldsymbol{u}$ and $\boldsymbol{x}_{\mathrm{R}}$ on each carrier is completely covered by $\grave{\boldsymbol{A}}\boldsymbol{x}_{\mathrm{R}}$, but there could still be correlations between components of $\boldsymbol{q}$ and $\grave{\boldsymbol{A}}\boldsymbol{x}_{\mathrm{R}}$ that belong to different carriers. It will later be seen that using completely uncorrelated $\boldsymbol{q}$ and $\boldsymbol{x}_{\mathrm{R}}$ is optimal.

Let $\mathcal{X} = (\grave{\boldsymbol{C}}_{\boldsymbol{v}}, \grave{\boldsymbol{C}}_{\boldsymbol{q}}, \grave{\boldsymbol{C}}_{\boldsymbol{x}_{\mathrm{R}}}, \grave{\boldsymbol{A}}, \breve{\boldsymbol{C}}_{\boldsymbol{v}}, \breve{\boldsymbol{C}}_{\boldsymbol{q}}, \breve{\boldsymbol{C}}_{\boldsymbol{x}_{\mathrm{R}}}, \breve{\boldsymbol{C}}_{\boldsymbol{q}\boldsymbol{x}_{\mathrm{R}}})$, and let

$$\mathbb{X} = \left\{ \mathcal{X} \;\middle|\; \begin{array}{l} \grave{\boldsymbol{C}}_{\boldsymbol{v}} \succeq \boldsymbol{0}, \grave{\boldsymbol{C}}_{\boldsymbol{q}} \succeq \boldsymbol{0}, \grave{\boldsymbol{C}}_{\boldsymbol{x}_{\mathrm{R}}} \succeq \boldsymbol{0}, \\ \breve{\boldsymbol{C}}_{\boldsymbol{v}} \in \breve{\mathcal{C}}(\grave{\boldsymbol{C}}_{\boldsymbol{v}}), \breve{\boldsymbol{C}}_{\boldsymbol{\rho}} \in \breve{\mathcal{C}}(\grave{\boldsymbol{C}}_{\boldsymbol{\rho}}) \end{array} \right\} \tag{11}$$

where we have used the abbreviation $\boldsymbol{\rho} = [\boldsymbol{q}^{\mathrm{T}} \;\; \boldsymbol{x}_{\mathrm{R}}^{\mathrm{T}}]^{\mathrm{T}}$ and the parametrization of covariance matrices introduced in Section III. The maximization (5) can then be written as

$$\begin{aligned} \max_{\mathcal{X}\in\mathbb{X}} \quad & \min\{R_{\mathrm{A}}(\mathcal{X});\; R_{\mathrm{B}}(\mathcal{X})\} \\ \text{s.t.} \quad & \mathrm{tr}[\grave{\boldsymbol{C}}_{\boldsymbol{v}} + \grave{\boldsymbol{C}}_{\boldsymbol{q}} + \grave{\boldsymbol{A}}\grave{\boldsymbol{C}}_{\boldsymbol{x}_{\mathrm{R}}}\grave{\boldsymbol{A}}^{\mathrm{H}}] \leq P_{\mathrm{S}} \\ & \mathrm{tr}[\grave{\boldsymbol{C}}_{\boldsymbol{x}_{\mathrm{R}}}] \leq P_{\mathrm{R}} \end{aligned} \tag{12}$$

with R_{A} and R_{B} from (4). To formulate the constraints, we have made use of Lemma 2 and of the fact that only the block-diagonal part is relevant for the trace.

Following the lines of [8], we introduce an auxiliary variable $\grave{\boldsymbol{C}}_{\boldsymbol{v}+\boldsymbol{q}}$ and apply the max-min-inequality [16, Section 5.4.1] to obtain the following upper bound to the optimal value:

$$\begin{aligned} \max_{\grave{\boldsymbol{C}}_{\boldsymbol{v}+\boldsymbol{q}} \succeq \boldsymbol{0}} \quad & \min\{R_{\mathrm{A}}^{\star}(\grave{\boldsymbol{C}}_{\boldsymbol{v}+\boldsymbol{q}});\; R_{\mathrm{B}}^{\star}(\grave{\boldsymbol{C}}_{\boldsymbol{v}+\boldsymbol{q}})\} \\ \text{s.t.} \quad & \mathrm{tr}[\grave{\boldsymbol{C}}_{\boldsymbol{v}+\boldsymbol{q}}] \leq P_{\mathrm{S}} \end{aligned} \tag{13}$$

where $\grave{\boldsymbol{C}}_{\boldsymbol{v}+\boldsymbol{q}}$ is block-diagonal, and

$$\begin{aligned} R_i^{\star}(\grave{\boldsymbol{C}}_{\boldsymbol{v}+\boldsymbol{q}}) = \max_{\mathcal{X}\in\mathbb{X}} \quad & R_i(\mathcal{X}) \\ \text{s.t.} \quad & \grave{\boldsymbol{C}}_{\boldsymbol{v}} + \grave{\boldsymbol{C}}_{\boldsymbol{q}} = \grave{\boldsymbol{C}}_{\boldsymbol{v}+\boldsymbol{q}} \\ & \mathrm{tr}[\grave{\boldsymbol{A}}\grave{\boldsymbol{C}}_{\boldsymbol{x}_{\mathrm{R}}}\grave{\boldsymbol{A}}^{\mathrm{H}}] \leq P_{\mathrm{S}} - \mathrm{tr}[\grave{\boldsymbol{C}}_{\boldsymbol{v}+\boldsymbol{q}}] \\ & \mathrm{tr}[\grave{\boldsymbol{C}}_{\boldsymbol{x}_{\mathrm{R}}}] \leq P_{\mathrm{R}} \end{aligned} \tag{14}$$

for $i \in \{\mathrm{A}, \mathrm{B}\}$. We now show that there exists an optimizer $\mathcal{X}^{\star}(\grave{\boldsymbol{C}}_{\boldsymbol{v}+\boldsymbol{q}})$ that maximizes R_{A} and R_{B} simultaneously for any given $\grave{\boldsymbol{C}}_{\boldsymbol{v}+\boldsymbol{q}}$, which implies that the upper bound is tight.

Let us first consider R_{B}, which can be written as

$$R_{\mathrm{B}} = \underbrace{\mathrm{h}(\boldsymbol{y}_{\mathrm{D}})}_{\leq \mathrm{h}(\boldsymbol{y}_{\mathrm{D,sep}})} - \underbrace{\mathrm{h}(\boldsymbol{\eta}_{\mathrm{D}})}_{\text{const.}} \tag{15}$$

where the inequality is due to Lemma 1 if $\boldsymbol{y}_{\mathrm{D,sep}}$ is a circularly symmetric Gaussian vector with

$$\begin{aligned} \grave{\boldsymbol{C}}_{\boldsymbol{y}_{\mathrm{D,sep}}} = \grave{\boldsymbol{C}}_{\boldsymbol{y}_{\mathrm{D}}} = {} & \grave{\boldsymbol{H}}_{\mathrm{SD}}(\grave{\boldsymbol{C}}_{\boldsymbol{v}} + \grave{\boldsymbol{C}}_{\boldsymbol{q}})\grave{\boldsymbol{H}}_{\mathrm{SD}}^{\mathrm{H}} + \\ & (\grave{\boldsymbol{H}}_{\mathrm{SD}}\grave{\boldsymbol{A}} + \grave{\boldsymbol{H}}_{\mathrm{RD}})\grave{\boldsymbol{C}}_{\boldsymbol{x}_{\mathrm{R}}}(\grave{\boldsymbol{H}}_{\mathrm{SD}}\grave{\boldsymbol{A}} + \grave{\boldsymbol{H}}_{\mathrm{RD}})^{\mathrm{H}} + \grave{\boldsymbol{C}}_{\boldsymbol{\eta}_{\mathrm{D}}} \end{aligned} \tag{16}$$

and $\breve{\boldsymbol{C}}_{\boldsymbol{y}_{\mathrm{D,sep}}} = \boldsymbol{0}$. Since $\grave{\boldsymbol{H}}_{\mathrm{SD}}$, $\grave{\boldsymbol{H}}_{\mathrm{RD}}$, $\grave{\boldsymbol{A}}$, and $\grave{\boldsymbol{C}}_{\boldsymbol{\eta}_{\mathrm{D}}}$ are block-diagonal, equality in (15) can be achieved by setting $\breve{\boldsymbol{C}}_{\boldsymbol{v}}$, $\breve{\boldsymbol{C}}_{\boldsymbol{q}}$, $\breve{\boldsymbol{C}}_{\boldsymbol{x}_{\mathrm{R}}}$, and $\breve{\boldsymbol{C}}_{\boldsymbol{q}\boldsymbol{x}_{\mathrm{R}}}$ to zero for any fixed choice of the matrices $\grave{\boldsymbol{C}}_{\boldsymbol{v}}$, $\grave{\boldsymbol{C}}_{\boldsymbol{q}}$, and $\grave{\boldsymbol{C}}_{\boldsymbol{x}_{\mathrm{R}}}$ (due to Lemma 2). Moreover, R_{B} does not depend on $\grave{\boldsymbol{C}}_{\boldsymbol{v}}$ and $\grave{\boldsymbol{C}}_{\boldsymbol{q}}$, but only on $\grave{\boldsymbol{C}}_{\boldsymbol{v}} + \grave{\boldsymbol{C}}_{\boldsymbol{q}} = \grave{\boldsymbol{C}}_{\boldsymbol{v}+\boldsymbol{q}}$. Thus, R_{B} is maximized by an optimizer $\mathcal{X}_{\mathrm{B}}^{\star}$ with the structure

$$\mathcal{X}_{\mathrm{B}}^{\star} = (*, *, \grave{\boldsymbol{C}}_{\boldsymbol{x}_{\mathrm{R}}}^{\star}, \grave{\boldsymbol{A}}^{\star}, \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0}) \tag{17}$$

where $*$ denotes "don't care." Note that due to the fact that $\grave{\boldsymbol{C}}_{\boldsymbol{q}\boldsymbol{x}_{\mathrm{R}}} = \boldsymbol{0}$ by assumption and $\breve{\boldsymbol{C}}_{\boldsymbol{q}\boldsymbol{x}_{\mathrm{R}}} = \boldsymbol{0}$ in the optimum, R_{B} is maximized by uncorrelated vectors $\boldsymbol{q}$ and $\boldsymbol{x}_{\mathrm{R}}$.

The rate R_{A} can be written as

$$\begin{aligned} R_{\mathrm{A}} = {} & \mathrm{h}(\grave{\boldsymbol{H}}_{\mathrm{SD}}\boldsymbol{v} + \boldsymbol{\eta}_{\mathrm{D}}) - \mathrm{h}(\boldsymbol{\eta}_{\mathrm{D}}) \\ & + \underbrace{\mathrm{h}(\grave{\boldsymbol{H}}_{\mathrm{SR}}(\boldsymbol{v}+\boldsymbol{q}) + \boldsymbol{\eta}_{\mathrm{R}} | \boldsymbol{x}_{\mathrm{R}})}_{\leq \mathrm{h}(\grave{\boldsymbol{H}}_{\mathrm{SR}}(\boldsymbol{v}+\boldsymbol{q}) + \boldsymbol{\eta}_{\mathrm{R}})} - \mathrm{h}(\grave{\boldsymbol{H}}_{\mathrm{SR}}\boldsymbol{v} + \boldsymbol{\eta}_{\mathrm{R}}). \end{aligned} \tag{18}$$

Since conditioning reduces uncertainty unless in the case of statistical independence [17, Section 8.6], (18) is maximized by independent $\boldsymbol{q}$ and $\boldsymbol{x}_{\mathrm{R}}$, i.e., by $\breve{\boldsymbol{C}}_{\boldsymbol{q}\boldsymbol{x}_{\mathrm{R}}} = \boldsymbol{0}$. Thus, the conditioning on $\boldsymbol{x}_{\mathrm{R}}$ can be dropped, and the probability distribution of $\boldsymbol{x}_{\mathrm{R}}$ does not play a role for the optimal R_{A}.

We rewrite the optimization as

$$\max_{\boldsymbol{C}_{\boldsymbol{v}} \succeq \boldsymbol{0}, \boldsymbol{C}_{\boldsymbol{q}} \succeq \boldsymbol{0}} \; R_{\mathrm{A}}(\boldsymbol{C}_{\boldsymbol{v}}, \boldsymbol{C}_{\boldsymbol{q}}) \quad \text{s.t.} \quad \grave{\boldsymbol{C}}_{\boldsymbol{v}} + \grave{\boldsymbol{C}}_{\boldsymbol{q}} \preceq \grave{\boldsymbol{C}}_{\boldsymbol{v}+\boldsymbol{q}}. \tag{19}$$

with

$$\begin{aligned} R_{\mathrm{A}}(\boldsymbol{C}_{\boldsymbol{v}}, \boldsymbol{C}_{\boldsymbol{q}}) = {} & \log \frac{\det\left(\mathbf{I}_{CN_{\mathrm{D}}} + \grave{\boldsymbol{H}}_{\mathrm{SD}}\boldsymbol{C}_{\boldsymbol{v}}\grave{\boldsymbol{H}}_{\mathrm{SD}}^{\mathrm{H}}\right)}{\det\left(\mathbf{I}_{CN_{\mathrm{D}}}\right)} \\ & + \log \frac{\det\left(\mathbf{I}_{CN_{\mathrm{R}}} + \grave{\boldsymbol{H}}_{\mathrm{SR}}\boldsymbol{C}_{\boldsymbol{v}}\grave{\boldsymbol{H}}_{\mathrm{SR}}^{\mathrm{H}} + \grave{\boldsymbol{H}}_{\mathrm{SR}}\boldsymbol{C}_{\boldsymbol{q}}\grave{\boldsymbol{H}}_{\mathrm{SR}}^{\mathrm{H}}\right)}{\det\left(\mathbf{I}_{CN_{\mathrm{R}}} + \grave{\boldsymbol{H}}_{\mathrm{SR}}\boldsymbol{C}_{\boldsymbol{v}}\grave{\boldsymbol{H}}_{\mathrm{SR}}^{\mathrm{H}}\right)} \end{aligned} \tag{20}$$

and we note that this is mathematically equivalent to a sum rate maximization in a two-user MIMO broadcast channel with dirty paper coding (cf. (24) and, e.g., [18]). A similar equivalence to a MIMO broadcast channel was exploited in [8], but for the case of a single carrier. Here, we have the case of parallel MIMO broadcast channels (e.g., [13], [15]) due to the block-diagonal channel matrices $\grave{\boldsymbol{H}}_{\text{SD}}$ and $\grave{\boldsymbol{H}}_{\text{SR}}$. The constraint affects only the block-diagonal parts $\grave{\boldsymbol{C}}_{\boldsymbol{v}}$ and $\grave{\boldsymbol{C}}_{\boldsymbol{q}}$, which is equivalent to a set of per-carrier shaping constraints $\grave{\boldsymbol{C}}_{\boldsymbol{v},c} + \grave{\boldsymbol{C}}_{\boldsymbol{q},c} \preceq \grave{\boldsymbol{C}}_{\boldsymbol{v}+\boldsymbol{q},c}$, $c \in \{1,\ldots,C\}$. The relaxation to an inequality constraint does not change the optimum since it can be shown that the constraint is active (R_{A} is increasing in $\grave{\boldsymbol{C}}_{\boldsymbol{v}+\boldsymbol{q}}$, the proof follows the lines of [8]).

In Theorem 2 in Appendix A, we show that carrier-noncooperative transmission is optimal for this kind of sum rate maximization in parallel MIMO broadcast channels. Due to the mathematical equivalence, we can conclude that there is an optimal solution of (19) with $\breve{\boldsymbol{C}}_{\boldsymbol{v}} = \breve{\boldsymbol{C}}_{\boldsymbol{q}} = \mathbf{0}$. Thus, there is an optimizer $\mathcal{X}_{\text{A}}^\star$ that maximizes R_{A} and has the structure

$$\mathcal{X}_{\text{A}}^\star = (\grave{\boldsymbol{C}}_{\boldsymbol{v}}^\star, \grave{\boldsymbol{C}}_{\boldsymbol{q}}^\star, *, *, \mathbf{0}, \mathbf{0}, *, \mathbf{0}). \tag{21}$$

It is easy to see that there exists an $\mathcal{X}^\star$ that is compatible with both structures $\mathcal{X}_{\text{A}}^\star$ and $\mathcal{X}_{\text{B}}^\star$. Therefore, the upper bound (13) is tight and can be achieved with block-diagonal covariance matrices $\boldsymbol{C}_{\boldsymbol{v}} = \grave{\boldsymbol{C}}_{\boldsymbol{v}}$, $\boldsymbol{C}_{\boldsymbol{q}} = \grave{\boldsymbol{C}}_{\boldsymbol{q}}$, and $\boldsymbol{C}_{\boldsymbol{x}_{\text{R}}} = \grave{\boldsymbol{C}}_{\boldsymbol{x}_{\text{R}}}$. Since this reasoning holds for any feasible $\grave{\boldsymbol{C}}_{\boldsymbol{v}+\boldsymbol{q}}$, it also holds for the optimal $\grave{\boldsymbol{C}}_{\boldsymbol{v}+\boldsymbol{q}}^\star$, which proves that carrier-noncooperative transmission is optimal for PDF in Gaussian MIMO relay channels. ∎

V. Discussion

We have shown that the PDF rate in parallel Gaussian MIMO relay channels is maximized by carrier-noncooperative transmission, but this does not mean that the setting is separable in the sense of equality between the achievable PDF rate and the sum of the rates that can be achieved with PDF individually on each carrier. For optimal PDF, the relay must have the possibility to forward a signal on a carrier different from the one on which it has been received.

For the optimal strategy, Theorem 1 states that all three source signals $\boldsymbol{v}$, $\boldsymbol{q}$, and $\grave{\boldsymbol{A}}\boldsymbol{x}_{\text{R}}$ as well as the relay signal $\boldsymbol{x}_{\text{R}}$ consist of per-carrier signals that are not correlated across carriers. Thus, R_{A} and R_{B} can be achieved by summing up the respective per-carrier expressions over all carriers, but this is not the same as assuming separate coding:

$$R = \min\left\{\sum\nolimits_{c=1}^{C} R_{\text{A},c}\ ;\ \sum\nolimits_{c=1}^{C} R_{\text{B},c}\right\} \tag{22}$$

$$\geq \sum\nolimits_{c=1}^{C} \min\left\{R_{\text{A},c}\ ;\ R_{\text{B},c}\right\} = R_{\text{separate}}. \tag{23}$$

The minimum operation in the PDF rate represents the fact that the rate of information leaving the source (towards relay and destination) has to be balanced with the rate of information arriving at the destination (from the source and the relay). Without joint coding, we have the stricter condition that this balance has to hold on each carrier individually.

For further interpretation, we recall that the PDF rate can be achieved by a block-Markov coding scheme [3, Section 9.4.1]. Apparently, $\grave{\boldsymbol{A}}\boldsymbol{x}_{\text{R}}$ represents the coherent transmission that is

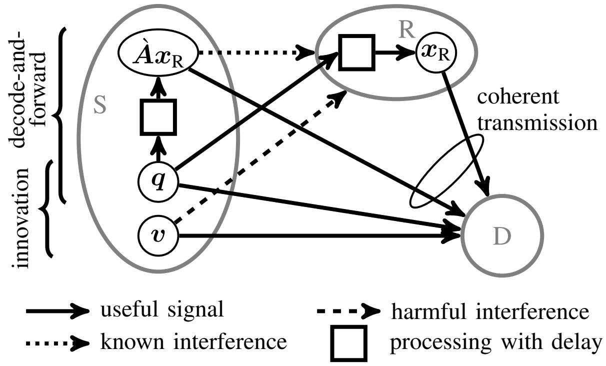

Fig. 2. Decomposition of the source signal (adapted from [8]).

currently taking place while $\boldsymbol{q}$ represents the message that is provided to the relay to allow coherent transmission in a future block [8]. Thus, $\boldsymbol{q}$, which is not correlated with $\boldsymbol{x}_{\text{R}}$, can be interpreted as part of the innovation that is introduced into the system by the source. An illustration can be found in Fig. 2. In the block-Markov scheme, the important distinction is whether components of $\boldsymbol{x}_{\text{R}}$ and $\grave{\boldsymbol{A}}\boldsymbol{x}_{\text{R}}$ on some carrier may depend on components of $\boldsymbol{q}$ on other carriers *in earlier blocks*. Such a dependence complies with carrier-noncooperative transmission in each block, but requires joint coding across carriers.

Appendix A
Excursus: Parallel MIMO Broadcast Channels

In this appendix, we extend a result that is known for parallel MIMO broadcast channels with a sum power constraint (see [15]) to a certain class of shaping constraints.

Theorem 2: If the noise is independent across carriers, the optimal sum rate in parallel MIMO broadcast channels with shaping constraints that affect only the per-carrier covariance matrices is achieved with carrier-noncooperative transmission.

For the proof, we make use of the uplink-downlink minimax duality with linear conic constraints from [19], [20]. We use $\boldsymbol{x}_k$ and $\boldsymbol{\xi}_k$, $k \in \{1,\ldots,K\}$ for the input signals in the downlink and in the dual uplink, respectively. Moreover, we use $\boldsymbol{\eta}_k$ for the downlink noise, and $\boldsymbol{\eta}$ for the uplink noise. The number of downlink transmit antennas is denoted by M, and the number of antennas at the kth downlink receiver is N_k. The data rate of user k can be expressed as (e.g., [18])

$$r_k = \log \frac{\det\left(\grave{\boldsymbol{C}}_{\boldsymbol{\eta}_k} + \grave{\boldsymbol{H}}_k\left(\sum_{j\in\mathcal{I}_k\cup\{k\}} \boldsymbol{C}_{\boldsymbol{x}_j}\right)\grave{\boldsymbol{H}}_k^{\text{H}}\right)}{\det\left(\grave{\boldsymbol{C}}_{\boldsymbol{\eta}_k} + \grave{\boldsymbol{H}}_k\left(\sum_{j\in\mathcal{I}_k} \boldsymbol{C}_{\boldsymbol{x}_j}\right)\grave{\boldsymbol{H}}_k^{\text{H}}\right)} \tag{24}$$

where $\mathcal{I}_k$ is the set of users causing interference to user k, i.e., the set of users encoded after user k. In the dual uplink,

$$r_k^{\text{UL}} = \log \frac{\det\left(\grave{\boldsymbol{C}}_{\boldsymbol{\eta}} + \sum_{j\in\mathcal{I}_k^{\text{UL}}\cup\{k\}} \grave{\boldsymbol{H}}_j^{\text{H}}\boldsymbol{C}_{\boldsymbol{\xi}_j}\grave{\boldsymbol{H}}_j\right)}{\det\left(\grave{\boldsymbol{C}}_{\boldsymbol{\eta}} + \sum_{j\in\mathcal{I}_k^{\text{UL}}} \grave{\boldsymbol{H}}_j^{\text{H}}\boldsymbol{C}_{\boldsymbol{\xi}_j}\grave{\boldsymbol{H}}_j\right)}. \tag{25}$$

Just like in [18], the decoding order in the uplink is the reverse downlink encoding order, i.e., $\mathcal{I}_k^{\text{UL}} = \{1,\ldots,K\} \setminus (\mathcal{I}_k \cup \{k\})$.

It will be seen later [due to symmetry in (30)] that this order can be chosen arbitrarily.

The minimax duality with linear conic constraints and worst-case noise optimization from [19], [20] reads as follows.

Lemma 3: The downlink minimax problem

$$\min_{\substack{(\boldsymbol{C}_{\boldsymbol{\eta}_k}\succeq \mathbf{0})_{\forall k}:\ (\boldsymbol{C}_{\boldsymbol{\eta}_k})_{\forall k}\in\mathcal{Y}^\perp \\ \sum_{k=1}^K \operatorname{tr}[\boldsymbol{C}_{\boldsymbol{\eta}_k}]=\sum_{k=1}^K CN_k}} \ \max_{\substack{(\boldsymbol{C}_{\boldsymbol{x}_k}\succeq \mathbf{0})_{\forall k},\boldsymbol{Z}\in\mathcal{Z} \\ \sum_{k=1}^K \boldsymbol{C}_{\boldsymbol{x}_k}\preceq \boldsymbol{C}+\boldsymbol{Z}}} \sum\nolimits_{k=1}^K r_k \tag{26}$$

and the uplink minimax problem

$$\min_{\substack{\boldsymbol{C}_{\boldsymbol{\eta}}\succeq \mathbf{0},\boldsymbol{C}_{\boldsymbol{\eta}}\in\mathcal{Z}^\perp \\ \operatorname{tr}[\boldsymbol{C}\boldsymbol{C}_{\boldsymbol{\eta}}]=\sum_{k=1}^K CN_k}} \ \max_{\substack{(\boldsymbol{C}_{\boldsymbol{\xi}_k}\succeq \mathbf{0})_{\forall k},(\boldsymbol{Y}_k)_{\forall k}\in\mathcal{Y} \\ \boldsymbol{C}_{\boldsymbol{\xi}_k}\preceq \mathbf{I}_{CN_k}+\boldsymbol{Y}_k\ \forall k}} \sum\nolimits_{k=1}^K r_k^{\mathrm{UL}} \tag{27}$$

have the same optimal value [19], [20], where $\mathcal{Z}\subseteq\mathbb{H}^M$ and $\mathcal{Y}\subseteq\bigotimes_{k=1}^K\mathbb{H}^{N_k}$ are linear subspaces.
The subspaces $\mathcal{Z}$ and $\mathcal{Y}$ can be used to model various constraints on the transmit covariance matrices (cf. [20]) while their orthogonal complements $\mathcal{Y}^\perp$ and $\mathcal{Z}^\perp$ determine constraints for the worst-case noise optimizations.

Proof of Theorem 2: To model that only the diagonal blocks of the transmit covariance matrices (i.e., the per-carrier covariance matrices) are affected by the shaping constraint, let $\mathcal{Z}=\mathbb{O}^{C,M\times M}\cap\mathbb{H}^{CM}$, which allows adding arbitrary off-diagonal blocks (as long as $\boldsymbol{C}_{\boldsymbol{x}_k}\succeq\mathbf{0}$). This translates to $\mathcal{Z}^\perp=\mathbb{D}^{C,M\times M}\cap\mathbb{H}^{CM}$ in the uplink optimization (27). Thus, the noise in the uplink is constrained to have a block-diagonal covariance matrix, i.e., to be independent across carriers.

In the downlink, we assume identity matrices as noise covariance matrices without loss of generality. As in [20], the sum rate maximization with fixed noise covariance matrices $\boldsymbol{C}_{\boldsymbol{\eta}_k}=\mathbf{I}_{CN_k}$ can be rewritten as a minimax problem by defining a feasible set that contains only one element:

$$\mathcal{Y}^\perp = \tag{28}$$
$$\left\{(\boldsymbol{C}_{\boldsymbol{\eta}_k})_{\forall k}\in\bigotimes_{k=1}^K\mathbb{H}^{CN_k}\ \middle|\ \boldsymbol{C}_{\boldsymbol{\eta}_k}=\alpha\mathbf{I}_{CN_k}\ \forall k,\ \alpha\in\mathbb{R}\right\}$$

The orthogonal complement

$$\mathcal{Y}=\left\{(\boldsymbol{Y}_k)_{\forall k}\in\bigotimes_{k=1}^K\mathbb{H}^{CN_k}\ \middle|\ \sum\nolimits_{k=1}^K\operatorname{tr}[\boldsymbol{Y}_k]=0\right\} \tag{29}$$

leads to shaping constraints in (27) that affect only the diagonal blocks of the uplink transmit covariance matrices.[1]

The uplink sum rate equals (e.g., [17, Section 15.3])

$$R_{\mathrm{UL}}=\underbrace{\mathrm{h}\left(\sum\nolimits_{k=1}^K\grave{\boldsymbol{H}}_k^{\mathrm{H}}\boldsymbol{\xi}_k+\boldsymbol{\eta}\right)}_{=\mathrm{h}(\boldsymbol{y})\leq\mathrm{h}(\boldsymbol{y}_{\mathrm{sep}})}-\underbrace{\mathrm{h}\left(\boldsymbol{\eta}\right)}_{\text{const.}} \tag{30}$$

where $\boldsymbol{y}=\sum_{k=1}^K\grave{\boldsymbol{H}}_k^{\mathrm{H}}\boldsymbol{\xi}_k+\boldsymbol{\eta}$, and $\boldsymbol{y}_{\mathrm{sep}}$ is a circularly symmetric Gaussian signal with $\grave{\boldsymbol{C}}_{\boldsymbol{y}_{\mathrm{sep}}}=\grave{\boldsymbol{C}}_{\boldsymbol{y}}$ and $\breve{\boldsymbol{C}}_{\boldsymbol{y}_{\mathrm{sep}}}=\mathbf{0}$. The inequality is due to Lemma 1. To comply with the constraints, we only have to consider the block-diagonal components of the transmit covariance matrices. For any fixed choice of them, equality $\mathrm{h}(\boldsymbol{y})=\mathrm{h}(\boldsymbol{y}_{\mathrm{sep}})$ can be achieved by setting the off-diagonal components $\breve{\boldsymbol{C}}_{\boldsymbol{x}_k}$ to zero since the off-diagonal component $\breve{\boldsymbol{C}}_{\boldsymbol{\eta}}$ of the noise covariance is zero.

Thus, the optimal uplink transmit covariance matrices are block-diagonal. By transforming this optimal solution to the downlink as described in [19], we obtain block-diagonal covariance matrices in the downlink, i.e., carrier-noncooperative transmission is optimal in the downlink. ■

[1] In fact, it can be shown that constraints of this form are equivalent to a sum power constraint [20].

References

[1] T. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.

[2] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.

[3] G. Kramer, "Topics in multi-user information theory," *Foundations and Trends® Commun. and Inf. Theory*, vol. 4, no. 4–5, pp. 265–444, 2007.

[4] A. El Gamal and M. Aref, "The capacity of the semideterministic relay channel," *IEEE Trans. Inf. Theory*, vol. 28, no. 3, pp. 536–536, May 1982.

[5] C. K. Lo, S. Vishwanath, and R. W. Heath, "Rate bounds for MIMO relay channels," *J. Commun. Netw.*, vol. 10, no. 2, pp. 194–203, Jun. 2008.

[6] L. Gerdes, L. Weiland, and W. Utschick, "A zero-forcing partial decode-and-forward scheme for the Gaussian MIMO relay channel," in *Proc. Int. Conf. Commun. (ICC) 2013*, Jun. 2013, pp. 1942–1947.

[7] L. Weiland, L. Gerdes, and W. Utschick, "Partial decode-and-forward rates for the Gaussian MIMO relay channel: Inner approximation of non-convex rate constraints," in *Proc. IEEE 14th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2013, pp. 540–544.

[8] C. Hellings, L. Gerdes, L. Weiland, and W. Utschick, "On optimal Gaussian signaling in MIMO relay channels with partial decode-and-forward," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3153–3164, Jun. 2014.

[9] Y. Liang, V. Veeravalli, and H. Poor, "Resource allocation for wireless fading relay channels: Max-min solution," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3432–3453, Oct. 2007.

[10] Y. Liang and G. Kramer, "Rate regions for relay broadcast channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3517–3535, Oct. 2007.

[11] B. E. Schein, "Distributed coordination in network information theory," Ph.D. Dissertation, Massachusetts Institute of Technology, Sep. 2001.

[12] D. P. Palomar, M. A. Lagunas, and J. M. Cioffi, "Optimum linear joint transmit-receive processing for MIMO channels with QoS constraints," *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1179–1197, May 2004.

[13] C. Hellings and W. Utschick, "On the inseparability of parallel MIMO broadcast channels with linear transceivers," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 6273–6278, Dec. 2011.

[14] L. Gerdes, C. Hellings, L. Weiland, and W. Utschick, "The optimal input distribution for partial decode-and-forward in the MIMO relay channel," submitted to IEEE Trans. Inf. Theory. Preprint available at http://arxiv.org/abs/1409.8624.

[15] P. Tejera, W. Utschick, J. Nossek, and G. Bauch, "Rate balancing in multiuser MIMO OFDM systems," *IEEE Trans. Commun.*, vol. 57, no. 5, pp. 1370–1380, May 2009.

[16] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2009, 7th printing with corrections.

[17] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ: Wiley-Interscience, 2006.

[18] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.

[19] A. Dotzler, M. Riemensberger, and W. Utschick, "Minimax duality for MIMO interference networks," 2014, to be published. Preprint available at http://www.msv.ei.tum.de/ando/MiniMaxDuality.pdf.

[20] A. Dotzler, M. Riemensberger, W. Utschick, and G. Dietl, "Interference robustness for cellular MIMO networks," in *Proc. IEEE 13th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2012, pp. 229–233.