On Optimal Gaussian Signaling in MIMO Relay Channels with Partial Decode-and-Forward

Christoph Hellings, Lennart Gerdes, Lorenz Weiland, and Wolfgang Utschick

IEEE Transactions on Signal Processing

vol. 62, no. 12, Jun. 2014

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

On Optimal Gaussian Signaling in MIMO Relay Channels with Partial Decode-and-Forward

Christoph Hellings, *Student Member, IEEE*, Lennart Gerdes, *Student Member, IEEE*, Lorenz Weiland, and Wolfgang Utschick, *Senior Member, IEEE*

***Abstract*—For Gaussian multiple-input multiple-output (MIMO) relay channels with partial decode-and-forward, the optimal type of input distribution is still an open question in general. Recent research has revealed that in some other scenarios with unknown optimal input distributions (e.g., interference channels), improper (i.e., noncircular) Gaussian distributions can outperform proper (circular) Gaussian distributions. In this paper, we show that this is not the case for partial decode-and-forward in the Gaussian MIMO relay channel with Gaussian transmit signals, i.e., we show that a proper Gaussian input distribution is the optimal one among all Gaussian distributions. In order to prove this property, an innovation covariance matrix is introduced, and a decomposition is performed by considering the optimization over this matrix as an outer problem. A key point for showing optimality of proper signals then is a reformulation that reveals that one of the subproblems is equivalent to a sum rate maximization in a two-user MIMO broadcast channel under a sum covariance constraint, for which the optimality of proper signals can be shown.**

***Index Terms*—Asymmetric complex signaling, composite real representation, Gaussian broadcast channel, Gaussian relay channel, multiple-input multiple-output (MIMO), partial decode-and-forward, proper and improper signals.**

I. Introduction

The idea of supporting wireless communication by means of relay nodes has attracted the interest of researchers both with an application oriented and with an information theoretic background. However, the capacity of relay networks is still an open problem except for special cases. In this work, we consider a network with a multiantenna transmitter, a multiantenna receiver, and one multiantenna relay node, where all channels are assumed to be additive circularly symmetric Gaussian noise channels, i.e., the setting under consideration is a so-called Gaussian multiple-input multiple-output (MIMO) relay channel (cf., e.g., [1]–[11]).

This setting with only one transmitter-receiver pair and only one relay is interesting both from a theoretical and a practical point of view. In information theory, it is a common approach to first understand fundamental properties in a minimal example before studying larger scenarios, and also when introducing relaying into practical wireless networks, it suggests itself to first start with simple settings before proceeding to more complicated scenarios like multiuser relay channels or settings with multiple relays. The study of such advanced settings is not considered in this paper and is left open for future research.

The assumption of additive Gaussian noise on all channels has a strong practical and theoretical justification as well. Practical systems designed based on this assumption perform quite well, which can be explained by the composite effect of many independent noise sources [12]. The assumption of circularly symmetric noise is due to the fact that Gaussian noise at the demodulator output of a bandpass system with real wide-sense stationary noise is circularly symmetric [13]. The justification from an information theoretic point of view is that Gaussian noise is the worst case noise in general wireless networks under very mild assumptions [12], and that the circularly symmetric Gaussian distribution maximizes the differential entropy [13], [14].

In addition to being an interesting setting in itself, the Gaussian MIMO relay channel can also be used as an elementary building block of larger multihop wireless networks [15]. To use it as such, a deep understanding of its properties is clearly necessary, which provides an additional motivation for studying the Gaussian MIMO relay channel.

A. The Gaussian MIMO Relay Channel with Partial Decode-and-Forward

An information theoretic model that is appropriate to analyze the Gaussian MIMO relay channel was first introduced in [16]. The application of this general model to the Gaussian MIMO relay channel was studied, e.g., in [1]–[11].

Since the optimal coding scheme is not known for relay channels, researchers have derived upper bounds such as the cut-set bound [17], which was studied for Gaussian (MIMO) relay channels in [1], [5], [7], [8], [17], [18], and achievable schemes (see also [19, Ch. 9]) such as amplify-and-forward [2], [4], [18], [20], [21], decode-and-forward [1], [5], [7], [8], [17], [18], compress-and-forward [6], [7], [17], [18], and combinations of these schemes [17].

While amplify-and-forward restricts the processing at the relay to consist of linear operations (or widely linear operations[1]), compress-and-forward relies on nonlinear quantization operations, but the message is not decoded by the relay. In

[1] Widely linear functions are linear functions of the real and imaginary parts of a complex number or, equivalently, linear functions of the complex number and its conjugate (e.g., [22]–[24]).

This work was supported by Deutsche Forschungsgemeinschaft (DFG) under grant Ut 36/11-1.

The authors are with the Associate Institute for Signal Processing, Technische Universität München, 80290 München, Germany, Telephone: +49 89 289-28516, Fax: +49 89 289-28504, e-mail: hellings@tum.de, gerdes@tum.de, lorenz.weiland@tum.de, utschick@tum.de.

the decode-and-forward protocol, the relay has to decode and re-encode the message in order to then transmit it to the destination jointly with the source node. This strategy can be capacity-achieving in special cases, in particular if the channel between the source and the relay is strong [18]. However, if this is not the case, the source-relay link can become a bottleneck [18], [19, 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 scheme was formulated in [17], where it was proposed to split the message transmitted by the source into a part that is transmitted using the decode-and-forward protocol and a second part that is not decoded by the relay, but relayed with compress-and-forward. As a special case of this general scheme, we can obtain the partial decode-and-forward scheme [3], [9], [10], [18], [25], [19, Section 9.4.1], where the second part is not relayed at all (i.e., it is compressed to a zero-message).

For Gaussian MIMO relay channels, partial decode-and-forward was previously considered in [3], [9]–[11]. Based on [3], where the achievable partial decode-and-forward rates were formulated for the Gaussian MIMO relay channel with jointly circularly symmetric Gaussian transmit signals, algorithms to find good suboptimal source and relay transmit covariance matrices were developed in [9], [10]. Due to the nonconvexity of the optimization problem, a method to find the globally optimal transmit covariance matrices under the assumption of circularly symmetric Gaussian transmit signals has not yet been proposed in the literature.

Moreover, the optimal type of probability distribution of the transmit signals is still unknown for partial decode-and-forward in Gaussian MIMO relay channels. This is in contrast to decode-and-forward, for which it can be shown that the circularly symmetric Gaussian distribution is the optimal input distribution [1].

In cases where the optimal partial decode-and-forward strategy becomes equivalent to decode-and-forward (or equivalent to point-to-point transmission without a relay), it can be concluded that the optimal input distribution of partial decode-and-forward is the same as for decode-and-forward (or direct transmission). These cases include physically (reversely) degraded relay channels [17], semideterministic relay channels [25], relay channels with orthogonal components [26], and stochastically (reversely) degraded relay channels [11], [21]. Special cases of the Gaussian MIMO relay channel can correspond to the cases considered in [26] or [11] (or [21], where the special case of [11] for single-antenna nodes was discussed). For these cases, we can directly conclude that the circularly symmetric Gaussian distribution is optimal for partial decode-and-forward. However, the optimal partial decode-and-forward input distribution remains unknown for general Gaussian MIMO relay channels.

B. Contributions

In this paper, we make a step towards finding the optimal type of input distribution for partial decode-and-forward in general Gaussian MIMO relay channels by providing an answer to the following question: under the assumption of Gaussian transmit signals, shall these signals be circularly symmetric (i.e., proper, see Section I-D) or not. As explained in detail in Section I-D, this question arises since many recent works have revealed that noncircular (i.e., improper) Gaussian distributions can outperform proper Gaussian distributions in various communication systems.

Following the commonly accepted approach in information theoretic studies of relay channels (e.g., [1], [3], [5]–[7], [17], [18]), our study is based on so-called achievable rates (*Shannon rates*, [27]), which give us the theoretical limits on the data rates that can be transmitted with vanishing error probability.

At least under the assumption of partial decode-and-forward with Gaussian input signals, we can show in Section IV that proper signals are indeed optimal from a Shannon rate perspective, i.e., they maximize the achievable rate. Note that this does not answer the question whether or not Gaussian signals are the optimal input signals for the partial decode-and-forward protocol. Nevertheless, the result obtained in this paper is a significant step towards a better understanding of partial decode-and-forward and towards finding the optimal input distribution.

Studying the special case of Gaussian inputs as a start when approaching difficult information theoretic problems involving Gaussian channels is a common practice (e.g., [3], [6], [9], [10], [21], [28], [29]). The assumption of Gaussian signals leads to tractable expressions and allows benefiting from the large variety of results on Gaussian signaling in the existing literature. In the outlook in Section VI, we shed some light on why it is not trivial to extend our considerations to the case where arbitrary transmit signals and not only Gaussian signals are allowed.

While the information theoretic relevance of the question treated in this paper comes from the pursuit of deeper insight into the problem of optimal relaying as explained above, there is also a strong motivation from a signal processing point of view. Since optimizing transmit strategies with proper signals requires far less optimization variables than optimizing improper signals (this becomes clear in Section IV), it is helpful to know that a restriction to proper signals can be introduced without a loss in performance.

Further contributions in the paper are a new parametrization of impropriety (see Section III) and a new decomposition of the optimization of the partial decode-and-forward rate based on a so-called innovation covariance matrix (see Section IV). We think that especially the latter might be a helpful tool for future research on partial decode-and-forward, e.g., for studying the abovementioned question whether Gaussian inputs are optimal as well as for algorithm design.

Finally, it turns out that as an ingredient for the proof of our main theorem, we also have to derive a new result concerning optimality of proper signals in MIMO broadcast channels with sum covariance constraints, which is presented as an excursus in Section V.

Before introducing the details of the system model and coding scheme in Section II, we want to introduce some notational conventions and devote the remainder of this section to a

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TSP.2014.2321737

brief introduction to proper and improper Gaussian signals. In addition, we provide a review of the communications-related literature on this subject, e.g., on related results for interference channels and broadcast channels.

C. Notation

In this paper, we use $\Re$ and $\Im$ for the real and imaginary parts, respectively, $\bullet^*$ for the complex conjugate, $\bullet^{\mathrm{T}}$ for the transpose, $\bullet^{\mathrm{H}}$ for the conjugate transpose, and $\bullet^{+}$ for the pseudoinverse. The notation $\bullet^{\star}$ is used for quantities that are optimizers or 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 $\mathbf{0}$ for the zero matrix or zero vector of appropriate size and $\mathbf{I}_N$ for the identity matrix of size N. Throughout the paper, we use $\boldsymbol{C}_{\boldsymbol{x}}$ for the covariance matrix of a real-valued or complex random vector $\boldsymbol{x}$, and $\tilde{\boldsymbol{C}}_{\boldsymbol{x}}$ for the pseudocovariance matrix (see Section I-D) of a complex random vector $\boldsymbol{x}$. Cross-covariance matrices and pseudo-cross-covariance matrices are denoted by $\boldsymbol{C}_{\boldsymbol{xy}}$ and $\tilde{\boldsymbol{C}}_{\boldsymbol{xy}}$, respectively. The sets $\mathbb{S}^N \subset \mathbb{R}^{N\times N}$ and $\mathbb{H}^N \subset \mathbb{C}^{N\times N}$ are the set of real-valued symmetric matrices and the set of complex Hermitian matrices, respectively. Orthogonal complements of linear subspaces are denoted by $\bullet^{\perp}$. The order relation $\succeq$ has to be understood in the sense of positive-semidefiniteness.

D. Proper and Improper Gaussian Signals

To exhaustively characterize the second-order statistical properties of a zero-mean complex random vector $\boldsymbol{x}$, we need the so-called pseudocovariance matrix $\tilde{\boldsymbol{C}}_{\boldsymbol{x}} = \mathrm{E}\left[\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}\right]$ in addition to the conventional covariance matrix $\boldsymbol{C}_{\boldsymbol{x}} = \mathrm{E}\left[\boldsymbol{x}\boldsymbol{x}^{\mathrm{H}}\right]$ (see [22]–[24] and references therein). Note that an alternative description of complex random vectors, which uses a composite real representation instead of the complex covariance matrix and pseudocovariance matrix, is introduced in Section I-E.

Only if the pseudocovariance vanishes, i.e., $\mathrm{E}\left[\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}\right] = \mathbf{0}$, the covariance matrix alone is sufficient to describe the second-order properties, in which case the random vector is called *proper*. Otherwise, it is called *improper*. In the case of a zero-mean Gaussian distribution, propriety is equivalent to circular symmetry of the probability density function.

In information theory, the proper Gaussian distribution plays an important role since it maximizes the differential entropy of a complex random vector for given covariance matrix [13], [14], i.e., it achieves equality in

$$\mathrm{h}(\boldsymbol{x}) \leq \log\det(\pi \mathrm{e} \boldsymbol{C}_{\boldsymbol{x}}). \tag{1}$$

In particular, this entails that the proper Gaussian distribution is the optimal input distribution of a point-to-point MIMO communication system with proper Gaussian noise [14].

The situation is different in interference-limited communication systems such as interference channels (see, e.g., [30]–[34]) or broadcast channels with (widely) linear transceivers (see, e.g., [33], [35], [36]). For such systems, the optimal input distribution is not known, but it can be shown that the proper Gaussian distribution is not always the optimal one. Advantages of using improper input distributions have also been shown for relay-assisted interference channels [37].

Between these two extreme cases, there is the intermediate case of systems where only some interference is present. An example is a broadcast channel with interference precompensation using so-called dirty paper coding (DPC) [38], [39]. Interestingly, improper signaling is not beneficial in the case of dirty paper coding with a sum power constraint [38]. An intuitive explanation of this fact was given in [35]: with dirty paper coding, there is a user that does not see any interference, and for this user a proper transmit signal is optimal just like in a point-to-point link. This leads to a proper interference-plus-noise signal for the next user, and the same reasoning can be applied recursively to understand why proper transmit signals should be utilized for all users. In Section V, we provide a formal generalization of this result to the case of a sum covariance constraint instead of a sum power constraint. Moreover, in that section, it is also made clear why the same optimality of proper signals is true for multiple access channels with successive interference cancellation.

Against this background, the question arises to which of these cases the various coding schemes for the Gaussian MIMO relay channel belong. Not only for the cut-set upper bound, but also for the decode-and-forward strategy, proper Gaussian signals have been shown to be optimal in Gaussian (MIMO) relay channels [1], [17]. In the decode-and-forward scheme, both the relay and the destination decode the complete source signal so that no interfering signal remains. Therefore, it is not surprising that this scheme belongs to the interference-free case where proper signals are optimal.

In the partial decode-and-forward scheme, only a part of the signal transmitted by the source node is decoded by the relay. This is explained in detail together with the system model in Section II. The other part of the source signal causes interference at the relay and is treated as additional noise. Therefore, the question arises whether or not the presence of this interference leads to a situation in which impropriety can be helpful. This question is studied in this paper.

It has to be mentioned that the processing of improper signals by means of widely linear operations [40] has recently been considered in relay channels [41] and two-way MIMO relay channels [42] with amplify-and-forward for the case where the utilization of an improper transmit signal is a system assumption (e.g., due to BPSK, ASK, or GMSK modulation). Such studies are fundamentally different from the kind of study presented in this paper: while we ask whether or not improper input distributions can achieve a better performance than proper ones, the authors of [41] and [42] took improper signals as given and only asked how to adequately process these signals. To the best of the authors' knowledge, the question whether or not proper signals are optimal for amplify-and-forward has not yet been answered in the existing literature, and the same is true for compress-and-forward. However, since this paper focuses on the partial decode-and-forward scheme, we do not further elaborate on amplify-and-forward or compress-and-forward.

E. Composite Real Representation

As an alternative to describing the statistical properties of a complex random vector by means of the covariance matrix and the pseudocovariance matrix, a composite real representation can be used (e.g., [24]). To this end, we introduce the notations

$$\check{\boldsymbol{A}} = \begin{bmatrix} \Re(\boldsymbol{A}) & -\Im(\boldsymbol{A}) \\ \Im(\boldsymbol{A}) & \Re(\boldsymbol{A}) \end{bmatrix} \text{ and } \check{\boldsymbol{a}} = \begin{bmatrix} \Re(\boldsymbol{a}) \\ \Im(\boldsymbol{a}) \end{bmatrix} \tag{2}$$

for the real-valued counterparts of a complex matrix $\boldsymbol{A}$ and a complex vector $\boldsymbol{a}$. Note that $\check{\boldsymbol{A}}$ is defined in a way that [14, Lemma 1]

$$\check{\boldsymbol{b}} = \check{\boldsymbol{A}}\check{\boldsymbol{x}} \quad \Leftrightarrow \quad \boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}. \tag{3}$$

By definition, we have (e.g., [23, Section 2.2.3])

$$\mathrm{h}(\check{\boldsymbol{x}}) = \mathrm{h}(\boldsymbol{x}) \quad \text{and} \quad \mathrm{h}(\check{\boldsymbol{x}}|\check{\boldsymbol{y}}) = \mathrm{h}(\boldsymbol{x}|\boldsymbol{y}). \tag{4}$$

For a general (proper or improper) complex Gaussian random vector $\boldsymbol{x}$, we therefore have

$$\mathrm{h}(\boldsymbol{x}) = \mathrm{h}(\check{\boldsymbol{x}}) = \frac{1}{2}\log\det(2\pi \mathrm{e}\boldsymbol{C}_{\check{\boldsymbol{x}}}) \tag{5}$$

where

$$\boldsymbol{C}_{\check{\boldsymbol{x}}} = \begin{bmatrix} \boldsymbol{C}_{\Re\boldsymbol{x}} & \boldsymbol{C}_{\Re\boldsymbol{x}\Im\boldsymbol{x}} \\ \boldsymbol{C}_{\Re\boldsymbol{x}\Im\boldsymbol{x}}^{\mathrm{T}} & \boldsymbol{C}_{\Im\boldsymbol{x}} \end{bmatrix} \tag{6}$$

is the (real-valued) covariance matrix of the composite real vector $\check{\boldsymbol{x}}$. This matrix may not be confused with the composite real equivalent $\check{\boldsymbol{C}}_{\boldsymbol{x}}$ of the (complex) covariance matrix $\boldsymbol{C}_{\boldsymbol{x}}$.

The covariance matrix and the pseudocovariance matrix of a complex random vector $\boldsymbol{x}$ can be expressed as a function of $\boldsymbol{C}_{\check{\boldsymbol{x}}}$ by means of the equations [24]

$$\boldsymbol{C}_{\boldsymbol{x}} = \boldsymbol{C}_{\Re\boldsymbol{x}} + \boldsymbol{C}_{\Im\boldsymbol{x}} + \mathrm{j}(\boldsymbol{C}_{\Re\boldsymbol{x}\Im\boldsymbol{x}}^{\mathrm{T}} - \boldsymbol{C}_{\Re\boldsymbol{x}\Im\boldsymbol{x}}) \tag{7}$$

$$\tilde{\boldsymbol{C}}_{\boldsymbol{x}} = \boldsymbol{C}_{\Re\boldsymbol{x}} - \boldsymbol{C}_{\Im\boldsymbol{x}} + \mathrm{j}(\boldsymbol{C}_{\Re\boldsymbol{x}\Im\boldsymbol{x}}^{\mathrm{T}} + \boldsymbol{C}_{\Re\boldsymbol{x}\Im\boldsymbol{x}}). \tag{8}$$

In the recent literature, it has been argued that for many problems involving improper signals, an augmented complex representation $\underline{\boldsymbol{x}} = [\boldsymbol{x}^{\mathrm{T}}, \boldsymbol{x}^{\mathrm{H}}]^{\mathrm{T}}$ leads to more convenient mathematical expressions than the composite real representation (e.g., [22]–[24], [32], [33]). For instance, while we can directly see if a vector $\boldsymbol{x}$ is proper by verifying whether the off-diagonal blocks of the augmented covariance matrix $\boldsymbol{C}_{\underline{\boldsymbol{x}}}$ are zero (e.g., [24]), we have to check the conditions

$$\boldsymbol{C}_{\Re\boldsymbol{x}} = \boldsymbol{C}_{\Im\boldsymbol{x}} \quad \text{and} \quad \boldsymbol{C}_{\Re\boldsymbol{x}\Im\boldsymbol{x}}^{\mathrm{T}} = -\boldsymbol{C}_{\Re\boldsymbol{x}\Im\boldsymbol{x}} \tag{9}$$

in order to test for propriety based on the composite real formulation (e.g., [24]).

However, the composite real representation has an important advantage that is crucial for a proof in this paper: in the literature, there exist many results for real-valued systems, and in addition, many results that have been shown for systems with proper complex signals can easily be transferred to real-valued systems. By means of the composite real representation, these existing results can be used to gain new insights about systems with improper complex signals. Such an approach was, e.g., pursued in [35], [36] and for the special case of scalar complex random variables in [30], [31].

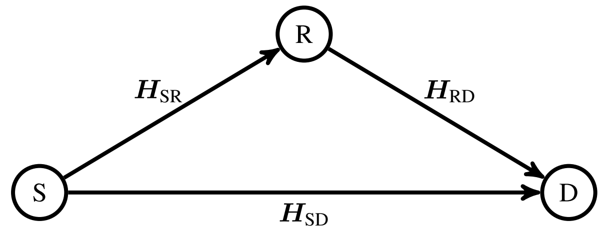

Fig. 1. Illustration of the Gaussian MIMO relay channel.

II. System Model and Coding Scheme

We consider data transmission from a multiantenna source node S to a multiantenna destination node D, where a multiantenna relay R is used to support the transmission. Under the assumption of frequency flat channels and additive circularly symmetric complex Gaussian noise at the relay and the destination, the transmission can be described by

$$\boldsymbol{y}_{\mathrm{R}} = \boldsymbol{H}_{\mathrm{SR}}\boldsymbol{x}_{\mathrm{S}} + \boldsymbol{\eta}_{\mathrm{R}} \tag{10}$$

$$\boldsymbol{y}_{\mathrm{D}} = \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{x}_{\mathrm{S}} + \boldsymbol{H}_{\mathrm{RD}}\boldsymbol{x}_{\mathrm{R}} + \boldsymbol{\eta}_{\mathrm{D}} \tag{11}$$

where $\boldsymbol{H}_{ij} \in \mathbb{C}^{N_j \times N_i}$ denotes the channel matrix between node i and node j with $i, j \in \{\mathrm{S}, \mathrm{R}, \mathrm{D}\}$, and N_i is the number of antennas at node i. The noise $\boldsymbol{\eta}_{\mathrm{R}} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{C}_{\boldsymbol{\eta}_{\mathrm{R}}})$ at the relay and the noise $\boldsymbol{\eta}_{\mathrm{D}} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{C}_{\boldsymbol{\eta}_{\mathrm{D}}})$ at the destination are assumed to be independent of each other and independent of the useful signals. Throughout the paper, we assume $\boldsymbol{C}_{\boldsymbol{\eta}_{\mathrm{R}}} = \mathbf{I}_{N_{\mathrm{R}}}$ and $\boldsymbol{C}_{\boldsymbol{\eta}_{\mathrm{D}}} = \mathbf{I}_{N_{\mathrm{D}}}$ without loss of generality. The system model is visualized in Fig. 1.

Throughout this work, we assume the transmit signals $\boldsymbol{x}_{\mathrm{S}}$ and $\boldsymbol{x}_{\mathrm{R}}$ at the source and at the relay to be jointly complex Gaussian, but not necessarily proper. Consequently, the receive signals $\boldsymbol{y}_{\mathrm{R}}$ and $\boldsymbol{y}_{\mathrm{D}}$ at the relay and the destination are general (proper or improper) complex Gaussian signals as well.

The relay is assumed to work in full-duplex mode with perfect self-interference cancellation, i.e., it is assumed to be able to transmit while receiving without disturbing its own reception. From an implementation point of view, this may be an overly optimistic assumption, and half-duplex operation, where self-interference cancellation is not necessary, might be preferable in practice. However, this paper is not meant to focus on aspects of practical implementations, but rather on fundamental properties of the optimal partial decode-and-forward solution. For the same reason, we restrict ourselves to the case of complete and perfect channel state information at all nodes. We conjecture that the results can be extended to the half-duplex case, and that they are also helpful to gain insights for the case of imperfect or incomplete channel state information. However, we leave a detailed study of these aspects open for future research.

In the partial decode-and-forward scheme, the transmit signal $\boldsymbol{x}_{\mathrm{S}}$ consists of a part $\boldsymbol{u}$ that is transmitted in cooperation with the relay and a part $\boldsymbol{v}$ which is received by the destination without the help of the relay, i.e., by making use of the direct link only. This implies that $\boldsymbol{u}$ has to be decoded by the relay. The transmit signal $\boldsymbol{x}_{\mathrm{S}}$ of the source node is created from the cooperation signal $\boldsymbol{u}$ and the noncooperative signal $\boldsymbol{v}$ by means of superposition coding (e.g., [25], [19, Section 9.4.1]), i.e., $\boldsymbol{x}_\mathrm{S} = \boldsymbol{u} + \boldsymbol{v}$ for continuous alphabets (e.g., [3]). While $\boldsymbol{v}$ is statistically independent of $\boldsymbol{u}$ and $\boldsymbol{x}_\mathrm{R}$, the cooperation signal $\boldsymbol{u}$ and the relay transmit signal $\boldsymbol{x}_\mathrm{R}$ are statistically dependent in general.

The achievable partial decode-and-forward rate is then given by

$$R = \min\{\underbrace{\mathrm{I}(\boldsymbol{x}_\mathrm{S}; \boldsymbol{y}_\mathrm{D}|\boldsymbol{u}\boldsymbol{x}_\mathrm{R}) + \mathrm{I}(\boldsymbol{u}; \boldsymbol{y}_\mathrm{R}|\boldsymbol{x}_\mathrm{R})}_{R_\mathrm{A}};\ \underbrace{\mathrm{I}(\boldsymbol{x}_\mathrm{S}\boldsymbol{x}_\mathrm{R}; \boldsymbol{y}_\mathrm{D})}_{R_\mathrm{B}}\}. \tag{12}$$

For a derivation of (12), the reader is referred to [25], [19, Section 9.4.1]. An intuitive interpretation of the involved mutual information expressions is given in Section IV. The optimization problem with power constraints reads as

$$\max_{p_{\boldsymbol{m}} \in \mathcal{M}} R \quad \text{s.t.} \quad \mathrm{E}\left[\boldsymbol{x}_\mathrm{S}^\mathrm{H}\boldsymbol{x}_\mathrm{S}\right] \leq P_\mathrm{S} \text{ and } \mathrm{E}\left[\boldsymbol{x}_\mathrm{R}^\mathrm{H}\boldsymbol{x}_\mathrm{R}\right] \leq P_\mathrm{R} \tag{13}$$

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. In (13), P_S and P_R denote the average transmit powers available at the source and at the relay, respectively.

We remark that the partial decode-and-forward scheme includes point-to-point transmission as the special case $\boldsymbol{u} \equiv \mathbf{0}$ and decode-and-forward as the special case $\boldsymbol{v} \equiv \mathbf{0}$. However, in the general case, where $\boldsymbol{u} \not\equiv \mathbf{0}$ and $\boldsymbol{v} \not\equiv \mathbf{0}$, we have to consider that the noncooperative signal $\boldsymbol{v}$ causes interference at the relay since it is not decoded by the relay. Due to this interference, it is not obvious whether the optimal input distribution is proper or improper, and it is not even clear whether Gaussian inputs are optimal (for a brief summary of special cases for which the optimal input distribution is known to be proper Gaussian, see Section I-A). Moreover, even if we restrict ourselves to jointly proper Gaussian signals as in the existing literature [3], [9], [10], the optimization of the transmit covariance matrices is a nonconvex problem for which only suboptimal solutions have been proposed [9], [10] for the general case.

III. Proposed Parametrization of Impropriety

As explained in Section I-E, the composite real representation enables us to transfer results that exist for real-valued systems or for systems with proper complex signals to the case of general (proper or improper) complex signals.

In order to benefit from this property, we have to overcome the disadvantages of the composite real-valued representation described in Section I-E. To this end, we propose to decompose the covariance matrix $\boldsymbol{C}_{\check{\boldsymbol{x}}}$ of the composite real representation (6) as

$$\boldsymbol{C}_{\check{\boldsymbol{x}}} = \underbrace{\frac{1}{2}\begin{bmatrix} \Re(\boldsymbol{C}_{\boldsymbol{x}}) & -\Im(\boldsymbol{C}_{\boldsymbol{x}}) \\ \Im(\boldsymbol{C}_{\boldsymbol{x}}) & \Re(\boldsymbol{C}_{\boldsymbol{x}}) \end{bmatrix}}_{\boldsymbol{P}_{\boldsymbol{x}}} + \underbrace{\frac{1}{2}\begin{bmatrix} \Re(\tilde{\boldsymbol{C}}_{\boldsymbol{x}}) & \Im(\tilde{\boldsymbol{C}}_{\boldsymbol{x}}) \\ \Im(\tilde{\boldsymbol{C}}_{\boldsymbol{x}}) & -\Re(\tilde{\boldsymbol{C}}_{\boldsymbol{x}}) \end{bmatrix}}_{\boldsymbol{N}_{\boldsymbol{x}}} \tag{14}$$

which is obtained by solving the system of equations (7) and (8) for $\boldsymbol{C}_{\Re\boldsymbol{x}}$, $\boldsymbol{C}_{\Im\boldsymbol{x}}$, and $\boldsymbol{C}_{\Re\boldsymbol{x}\Im\boldsymbol{x}}$.

The matrices $\boldsymbol{P}_{\boldsymbol{x}}$ and $\boldsymbol{N}_{\boldsymbol{x}}$ are both real-valued and symmetric with twice the size of the complex covariance matrix $\boldsymbol{C}_{\boldsymbol{x}}$. However, each of these matrices has its own characteristic block pattern. This leads to the following definition.

Definition 1: Let

$$\mathcal{P}^N = \left\{ \boldsymbol{P} \in \mathbb{S}^{2N} \;\middle|\; \boldsymbol{P} = \begin{bmatrix} \boldsymbol{A} & -\boldsymbol{B} \\ \boldsymbol{B} & \boldsymbol{A} \end{bmatrix},\ \boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{N\times N} \right\} \tag{15}$$

$$\mathcal{N}^N = \left\{ \boldsymbol{N} \in \mathbb{S}^{2N} \;\middle|\; \boldsymbol{N} = \begin{bmatrix} \boldsymbol{C} & \boldsymbol{D} \\ \boldsymbol{D} & -\boldsymbol{C} \end{bmatrix},\ \boldsymbol{C}, \boldsymbol{D} \in \mathbb{R}^{N\times N} \right\}. \tag{16}$$

Then, $\mathcal{P}^N$ is called the *power shaping space*, and $\mathcal{N}^N$ is called the *noncircularity space* or *impropriety space*.

It is easy to verify that $\mathcal{P}^N$ and $\mathcal{N}^N$ are linear subspaces of the space of real-valued symmetric matrices $\mathbb{S}^{2N}$. The names for $\mathcal{P}^N$ and $\mathcal{N}^N$ refer to the fact that $\boldsymbol{P}_{\boldsymbol{x}} \in \mathcal{P}^N$ completely determines the covariance matrix $\boldsymbol{C}_{\boldsymbol{x}}$ while $\boldsymbol{N}_{\boldsymbol{x}} \in \mathcal{N}^N$ determines the pseudocovariance matrix $\tilde{\boldsymbol{C}}_{\boldsymbol{x}}$ and, thus, the impropriety of $\boldsymbol{x} \in \mathbb{C}^N$.

Note that the set $\mathcal{P}^N$ is not restricted to the positive-semidefinite cone, and elements of $\mathcal{N}^N$ are indefinite by definition. A meaningful covariance matrix $\boldsymbol{C}_{\check{\boldsymbol{x}}}$ is obtained only for those $\boldsymbol{P}_{\boldsymbol{x}} \in \mathcal{P}^N$ and $\boldsymbol{N}_{\boldsymbol{x}} \in \mathcal{N}^N$ that fulfill $\boldsymbol{P}_{\boldsymbol{x}} + \boldsymbol{N}_{\boldsymbol{x}} \succeq \mathbf{0}$ (which implies $\boldsymbol{P}_{\boldsymbol{x}} \succeq \mathbf{0}$ due to the special structure of $\boldsymbol{P}_{\boldsymbol{x}}$ and $\boldsymbol{N}_{\boldsymbol{x}}$).

Due to the following lemma, any feasible composite real covariance matrix $\boldsymbol{C}_{\check{\boldsymbol{x}}}$ can be uniquely decomposed into a power shaping component $\boldsymbol{P}_{\boldsymbol{x}}$ and an impropriety component $\boldsymbol{N}_{\boldsymbol{x}}$.

Lemma 1: $\mathcal{N}^N$ is the orthogonal complement of $\mathcal{P}^N$ in $\mathbb{S}^{2N}$.

Proof: Orthogonality of $\mathcal{N}^N$ and $\mathcal{P}^N$ follows from $\mathrm{tr}[\boldsymbol{P}^\mathrm{T}\boldsymbol{N}] = \mathrm{tr}[\boldsymbol{P}\boldsymbol{N}] = 0$ for any choice of $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, and $\boldsymbol{D}$. Noting that $\boldsymbol{A}$, $\boldsymbol{C}$, and $\boldsymbol{D}$ must be symmetric while $\boldsymbol{B}$ must be skew-symmetric, we can count that $\mathcal{P}^N$ is N^2-dimensional, and $\mathcal{N}^N$ is (N^2+N)-dimensional. This adds up to $(2N^2+N)$, which is the dimensionality of $\mathbb{S}^{2N}$. ■

IV. Characterization of the Optimal Gaussian Signaling

In this section, we characterize the optimal Gaussian signaling in Gaussian MIMO relay channels with partial decode-and-forward by first stating and proving the main theorem of this paper and by giving an interpretation of the new insights afterwards.

Theorem 1: In a Gaussian MIMO relay channel with partial decode-and-forward, the zero-mean proper Gaussian distribution is the optimal one among all Gaussian input distributions.

The main difficulties in proving the theorem are

- the correlation between the cooperation signal $\boldsymbol{u}$ and the relay transmit signal $\boldsymbol{x}_\mathrm{R}$,
- the fact that the rate is given by a minimum of two rate expressions R_A and R_B [see (12)], and
- the interference that the noncooperative signal $\boldsymbol{v}$ causes at the relay.

To overcome the first problem, we split the cooperation signal $\boldsymbol{u}$ into a part $\boldsymbol{q}$ that has no linear relationship with $\boldsymbol{x}_\mathrm{R}$ and a part covering the linear relationship with $\boldsymbol{x}_{\mathrm{R}}$.[2] Then, we introduce the covariance matrix $\boldsymbol{C}_{\boldsymbol{v}+\boldsymbol{q}}$ of the sum $\boldsymbol{v}+\boldsymbol{q}$ and consider the separate maximization of R_{A} and R_{B} for a given $\boldsymbol{C}_{\boldsymbol{v}+\boldsymbol{q}}$. This separate maximization leads to an upper bound for the original problem, but in the end of the proof, it turns out that this bound is tight. Finally, the problem of interference at the relay is treated by identifying that for a fixed $\boldsymbol{C}_{\boldsymbol{v}+\boldsymbol{q}}$, the rate expression R_{A} can be reformulated in a way that it becomes mathematically equivalent to a dirty paper coding sum rate in a MIMO broadcast channel. As a last ingredient for the proof, it is then necessary to state and prove a new theorem for the MIMO broadcast channel, namely that proper Gaussian signaling is optimal for the sum rate maximization in a two-user MIMO broadcast channel not only under a sum power constraint, but also under a sum covariance constraint.

The detailed proof of the optimality of proper signaling for partial decode-and-forward in the MIMO relay channel with Gaussian transmit signals and proper Gaussian noise is provided below.

Proof of Theorem 1: Since a nonzero mean of the involved random vectors would not change the entropy, but require a part of the available transmit power, we can directly conclude that it is optimal to only use zero-mean random vectors.

We decompose $\boldsymbol{u}$ as

$$\boldsymbol{u}=\boldsymbol{q}+\boldsymbol{A}\boldsymbol{x}_{\mathrm{R}} \quad \text{such that} \quad \boldsymbol{C}_{\boldsymbol{q}\boldsymbol{x}_{\mathrm{R}}}=\mathrm{E}\left[\boldsymbol{q}\boldsymbol{x}_{\mathrm{R}}^{\mathrm{H}}\right]=\boldsymbol{0} \tag{17}$$

i.e., the linear relationship between $\boldsymbol{u}$ and $\boldsymbol{x}_{\mathrm{R}}$ is completely covered by $\boldsymbol{A}\boldsymbol{x}_{\mathrm{R}}$. For a better understanding of this decomposition, we have to consider that the partial decode-and-forward rate can be achieved by a block-Markov coding scheme [19, Section 9.4.1], and that the relay can only forward messages that it has received in earlier blocks for reasons of causality. With this in mind, we note the following. Both $\boldsymbol{q}$ and $\boldsymbol{A}\boldsymbol{x}_{\mathrm{R}}$ represent messages that are transmitted jointly with the relay. However, while $\boldsymbol{A}\boldsymbol{x}_{\mathrm{R}}$ represents the joint transmission that is currently taking place, $\boldsymbol{q}$ represents the message that is provided to the relay to allow cooperation in a future block. Since vanishing error probability can be assumed due to the formulation based on Shannon rates (cf. Section I-B), the source knows what the relay has received and decoded in the previous block and can conclude what the resulting relay transmit signal $\boldsymbol{x}_{\mathrm{R}}$ in the current block is. This explains why cooperative transmission of the source and the relay in the form of the signals $\boldsymbol{A}\boldsymbol{x}_{\mathrm{R}}$ and $\boldsymbol{x}_{\mathrm{R}}$ is possible even without a feedback link from the relay to the source. An illustration can be found in Fig. 2.

This correlation of the transmit signals of the source and the relay implies that the relay transmit signal and the matrix $\boldsymbol{A}$ have to be designed in a way that not only the power constraint at the relay is fulfilled, but also the power constraint at the source. This will be later seen in (19). Finding the value of $\boldsymbol{A}$

[2]Note that we do not use the term *correlation* since for improper random vectors, correlation may also be understood as the presence of a widely linear (but not necessarily linear) relationship [23, Section 2.2.1]. An example of vectors $\boldsymbol{q}$ and $\boldsymbol{x}_{\mathrm{R}}$ that are correlated in this sense without having a linear relationship is $\boldsymbol{C}_{\boldsymbol{q}\boldsymbol{x}_{\mathrm{R}}}=\boldsymbol{0}$ and $\tilde{\boldsymbol{C}}_{\boldsymbol{q}\boldsymbol{x}_{\mathrm{R}}}\neq\boldsymbol{0}$. This situation is explicitly allowed in our considerations.

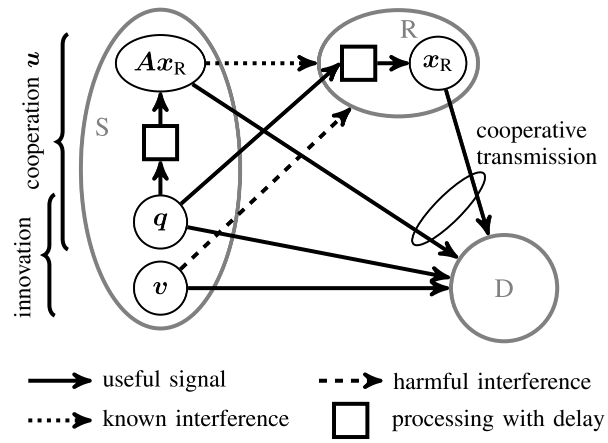

Fig. 2. Illustration of the proposed decomposition of the source signal.

and the distribution of $\boldsymbol{x}_{\mathrm{R}}$ that maximize the achievable rate while fulfilling these two constraints is one of the goals of the optimization.

Note that even under the assumption of Gaussian signals, $\boldsymbol{C}_{\boldsymbol{q}\boldsymbol{x}_{\mathrm{R}}}=\boldsymbol{0}$ does not imply statistical independence of $\boldsymbol{q}$ and $\boldsymbol{x}_{\mathrm{R}}$ since the pseudo-cross-covariance matrix $\tilde{\boldsymbol{C}}_{\boldsymbol{q}\boldsymbol{x}_{\mathrm{R}}}$ might be nonzero [23, Section 2.2.1], i.e., there might still be a widely linear relationship between $\boldsymbol{q}$ and $\boldsymbol{x}_{\mathrm{R}}$. However, we will later show that independence of $\boldsymbol{q}$ and $\boldsymbol{x}_{\mathrm{R}}$ is indeed optimal.

Our aim is to maximize the partial decode-and-forward rate R subject to power constraints at the source and at the relay. Under the assumption of zero-mean complex Gaussian transmit signals, the mutual information expressions in R can be written as functions of $\boldsymbol{A}$ and of the (joint) covariance and pseudocovariance matrices, which are, therefore, the optimization variables.

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

$$\mathbb{X}=\left\{\mathcal{X}\;\middle|\;\begin{array}{c}\boldsymbol{C}_{\boldsymbol{v}}\succeq\boldsymbol{0},\boldsymbol{C}_{\boldsymbol{q}}\succeq\boldsymbol{0},\boldsymbol{C}_{\boldsymbol{x}_{\mathrm{R}}}\succeq\boldsymbol{0},\\ \tilde{\boldsymbol{C}}_{\boldsymbol{v}}\in\tilde{\mathcal{C}}(\boldsymbol{C}_{\boldsymbol{v}}),\tilde{\boldsymbol{C}}_{\boldsymbol{\rho}}\in\tilde{\mathcal{C}}(\boldsymbol{C}_{\boldsymbol{\rho}})\end{array}\right\} \tag{18}$$

where we have used the abbreviation $\boldsymbol{\rho}=[\boldsymbol{q}^{\mathrm{T}}\ \boldsymbol{x}_{\mathrm{R}}^{\mathrm{T}}]^{\mathrm{T}}$. The set $\tilde{\mathcal{C}}(\boldsymbol{C})$ is the set of valid pseudocovariance matrices for given covariance matrix $\boldsymbol{C}$, i.e., the set of all complex symmetric matrices $\tilde{\boldsymbol{C}}$ for which $\boldsymbol{C}-\tilde{\boldsymbol{C}}(\boldsymbol{C}^{*})^{+}\tilde{\boldsymbol{C}}^{*}\succeq\boldsymbol{0}$ and for which the null space of $\boldsymbol{C}$ is contained in the null space of $\tilde{\boldsymbol{C}}^{*}$ [23, Section 2.2.2].

Using the tuple of optimization variables $\mathcal{X}$ and the definitions of R_{A} and R_{B} in (12), the optimization (13) can be written as

$$\begin{aligned}\max_{\mathcal{X}\in\mathbb{X}}\quad &\min\{R_{\mathrm{A}}(\mathcal{X});\ R_{\mathrm{B}}(\mathcal{X})\}\\ \text{s.t.}\quad &\mathrm{tr}[\boldsymbol{C}_{\boldsymbol{v}}+\boldsymbol{C}_{\boldsymbol{q}}+\boldsymbol{A}\boldsymbol{C}_{\boldsymbol{x}_{\mathrm{R}}}\boldsymbol{A}^{\mathrm{H}}]\leq P_{\mathrm{S}}\\ &\mathrm{tr}[\boldsymbol{C}_{\boldsymbol{x}_{\mathrm{R}}}]\leq P_{\mathrm{R}}.\end{aligned} \tag{19}$$

Due to the max-min-inequality [43, Section 5.4.1], an upper bound to the optimal value is given by

$$\begin{aligned}\max_{\boldsymbol{C}_{\boldsymbol{v}+\boldsymbol{q}}\succeq\boldsymbol{0}}\quad &\min\{R_{\mathrm{A}}^{\star}(\boldsymbol{C}_{\boldsymbol{v}+\boldsymbol{q}});\ R_{\mathrm{B}}^{\star}(\boldsymbol{C}_{\boldsymbol{v}+\boldsymbol{q}})\}\\ \text{s.t.}\quad &\mathrm{tr}[\boldsymbol{C}_{\boldsymbol{v}+\boldsymbol{q}}]\leq P_{\mathrm{S}}\end{aligned} \tag{20}$$

where $R_i^\star(\boldsymbol{C}_{\boldsymbol{v}+\boldsymbol{q}})$, $i \in \{\mathrm{A}, \mathrm{B}\}$ is the optimal value of

$$\begin{aligned} &\max_{\mathcal{X} \in \mathbb{X}} \; R_i(\mathcal{X}) \\ &\text{s.t.} \quad \boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{C}_{\boldsymbol{q}} = \boldsymbol{C}_{\boldsymbol{v}+\boldsymbol{q}} \\ &\qquad\; \operatorname{tr}[\boldsymbol{A}\boldsymbol{C}_{\boldsymbol{x}_\mathrm{R}}\boldsymbol{A}^\mathrm{H}] \leq P_\mathrm{S} - \operatorname{tr}[\boldsymbol{C}_{\boldsymbol{v}+\boldsymbol{q}}] \\ &\qquad\; \operatorname{tr}[\boldsymbol{C}_{\boldsymbol{x}_\mathrm{R}}] \leq P_\mathrm{R}. \end{aligned} \tag{21}$$

This upper bound is tight if there exists an optimizer $\mathcal{X}^\star(\boldsymbol{C}_{\boldsymbol{v}+\boldsymbol{q}})$ that maximizes R_A and R_B simultaneously for given $\boldsymbol{C}_{\boldsymbol{v}+\boldsymbol{q}}$. In the following, we show that such an optimizer exists for any possible $\boldsymbol{C}_{\boldsymbol{v}+\boldsymbol{q}}$. Therefore, the following results are not restricted to some fixed $\boldsymbol{C}_{\boldsymbol{v}+\boldsymbol{q}}$, but still hold if the optimal $\boldsymbol{C}_{\boldsymbol{v}+\boldsymbol{q}}$ is searched in the outer maximization of (20).

Let us first consider the solution of $R_\mathrm{B}^\star(\boldsymbol{C}_{\boldsymbol{v}+\boldsymbol{q}})$. In terms of differential entropies, R_B can be written as

$$R_\mathrm{B} = \underbrace{\mathrm{h}(\boldsymbol{y}_\mathrm{D})}_{\leq \mathrm{h}(\boldsymbol{y}_{\mathrm{D,proper}})} - \underbrace{\mathrm{h}(\boldsymbol{\eta}_\mathrm{D})}_{\text{const.}} \tag{22}$$

where $\boldsymbol{y}_{\mathrm{D,proper}}$ is a proper Gaussian signal with the same covariance matrix as $\boldsymbol{y}_\mathrm{D}$, i.e.,

$$\begin{aligned} \boldsymbol{C}_{\boldsymbol{y}_{\mathrm{D,proper}}} = \boldsymbol{C}_{\boldsymbol{y}_\mathrm{D}} = \boldsymbol{H}_\mathrm{SD}\boldsymbol{C}_{\boldsymbol{v}+\boldsymbol{q}}\boldsymbol{H}_\mathrm{SD}^\mathrm{H} + \\ (\boldsymbol{H}_\mathrm{SD}\boldsymbol{A} + \boldsymbol{H}_\mathrm{RD})\boldsymbol{C}_{\boldsymbol{x}_\mathrm{R}}(\boldsymbol{H}_\mathrm{SD}\boldsymbol{A} + \boldsymbol{H}_\mathrm{RD})^\mathrm{H} + \boldsymbol{C}_{\boldsymbol{\eta}_\mathrm{D}}. \end{aligned} \tag{23}$$

Since the noise $\boldsymbol{\eta}_\mathrm{D}$ is assumed to be proper, equality in (22) can be achieved by setting all pseudocovariance matrices to zero for any fixed choice of the covariance matrices (due to [14, Lemmas 3 and 4]). Also note that for fixed $\boldsymbol{C}_{\boldsymbol{v}+\boldsymbol{q}}$, the receive covariance matrix $\boldsymbol{C}_{\boldsymbol{y}_\mathrm{D}}$ does not depend on $\boldsymbol{C}_{\boldsymbol{v}}$ and $\boldsymbol{C}_{\boldsymbol{q}}$. Therefore, there is an optimizer $\mathcal{X}_\mathrm{B}^\star$ that maximizes R_B and has the structure

$$\mathcal{X}_\mathrm{B}^\star = (*, *, \boldsymbol{C}_{\boldsymbol{x}_\mathrm{R}}^\star, \boldsymbol{A}^\star, \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0}) \tag{24}$$

where $*$ denotes "don't care." Note that due to the fact that $\boldsymbol{C}_{\boldsymbol{q}\boldsymbol{x}_\mathrm{R}} = \boldsymbol{0}$ by assumption and $\tilde{\boldsymbol{C}}_{\boldsymbol{\rho}} = \boldsymbol{0}$ in the optimum, R_B is maximized by statistically independent random vectors $\boldsymbol{q}$ and $\boldsymbol{x}_\mathrm{R}$.

The rate R_A can be written as

$$\begin{aligned} R_\mathrm{A} &= \mathrm{h}(\boldsymbol{H}_\mathrm{SD}\boldsymbol{v} + \boldsymbol{\eta}_\mathrm{D}) - \mathrm{h}(\boldsymbol{\eta}_\mathrm{D}) \\ &\quad + \underbrace{\mathrm{h}(\boldsymbol{H}_\mathrm{SR}(\boldsymbol{v}+\boldsymbol{q}) + \boldsymbol{\eta}_\mathrm{R} | \boldsymbol{x}_\mathrm{R})}_{\leq \mathrm{h}(\boldsymbol{H}_\mathrm{SR}(\boldsymbol{v}+\boldsymbol{q}) + \boldsymbol{\eta}_\mathrm{R})} - \mathrm{h}(\boldsymbol{H}_\mathrm{SR}\boldsymbol{v} + \boldsymbol{\eta}_\mathrm{R}). \end{aligned} \tag{25}$$

Since conditioning reduces uncertainty unless in the case of statistical independence [44, Section 8.6], equality holds in (25) if and only if $\boldsymbol{q}$ and $\boldsymbol{x}_\mathrm{R}$ are independent, i.e., $\tilde{\boldsymbol{C}}_{\boldsymbol{q}\boldsymbol{x}_\mathrm{R}} = \boldsymbol{0}$. We can therefore assume this independence and drop the conditioning on $\boldsymbol{x}_\mathrm{R}$. As a result, the probability distribution of $\boldsymbol{x}_\mathrm{R}$ does not play a role for the optimal R_A.

Based on the composite real representations of all complex random vectors, we can use (5) to express R_A as

$$\begin{aligned} R_\mathrm{A}(\boldsymbol{C}_{\check{\boldsymbol{v}}}, \boldsymbol{C}_{\check{\boldsymbol{q}}}) &= \frac{1}{2}\log\frac{\det\left(\frac{1}{2}\mathbf{I}_{2N_\mathrm{D}} + \check{\boldsymbol{H}}_\mathrm{SD}\boldsymbol{C}_{\check{\boldsymbol{v}}}\check{\boldsymbol{H}}_\mathrm{SD}^\mathrm{T}\right)}{\det\left(\frac{1}{2}\mathbf{I}_{2N_\mathrm{D}}\right)} \\ &+ \frac{1}{2}\log\frac{\det\left(\frac{1}{2}\mathbf{I}_{2N_\mathrm{R}} + \check{\boldsymbol{H}}_\mathrm{SR}\boldsymbol{C}_{\check{\boldsymbol{v}}}\check{\boldsymbol{H}}_\mathrm{SR}^\mathrm{T} + \check{\boldsymbol{H}}_\mathrm{SR}\boldsymbol{C}_{\check{\boldsymbol{q}}}\check{\boldsymbol{H}}_\mathrm{SR}^\mathrm{T}\right)}{\det\left(\frac{1}{2}\mathbf{I}_{2N_\mathrm{R}} + \check{\boldsymbol{H}}_\mathrm{SR}\boldsymbol{C}_{\check{\boldsymbol{v}}}\check{\boldsymbol{H}}_\mathrm{SR}^\mathrm{T}\right)} \end{aligned} \tag{26}$$

where we have used $\boldsymbol{C}_{\check{\boldsymbol{\eta}}_\mathrm{R}} = \boldsymbol{P}_{\boldsymbol{\eta}_\mathrm{R}} = \frac{1}{2}\mathbf{I}_{2N_\mathrm{R}}$ and $\boldsymbol{C}_{\check{\boldsymbol{\eta}}_\mathrm{D}} = \boldsymbol{P}_{\boldsymbol{\eta}_\mathrm{D}} = \frac{1}{2}\mathbf{I}_{2N_\mathrm{D}}$.

The optimization of R_A can thus be written as

$$\begin{aligned} &\max_{\boldsymbol{C}_{\check{\boldsymbol{v}}} \succeq \boldsymbol{0}, \boldsymbol{C}_{\check{\boldsymbol{q}}} \succeq \boldsymbol{0}} \; R_\mathrm{A}(\boldsymbol{C}_{\check{\boldsymbol{v}}}, \boldsymbol{C}_{\check{\boldsymbol{q}}}) \\ &\qquad \text{s.t.} \quad \mathbb{P}_{\mathcal{P}^{N_\mathrm{S}}}(\boldsymbol{C}_{\check{\boldsymbol{v}}} + \boldsymbol{C}_{\check{\boldsymbol{q}}}) = \boldsymbol{P}_{\boldsymbol{v}+\boldsymbol{q}} \end{aligned} \tag{27}$$

where $\mathbb{P}_{\mathcal{P}^{N_\mathrm{S}}}$ denotes projection onto the power shaping space $\mathcal{P}^{N_\mathrm{S}}$, and $\boldsymbol{P}_{\boldsymbol{v}+\boldsymbol{q}}$ is the constant real-valued power shaping matrix corresponding to the constant complex covariance matrix $\boldsymbol{C}_{\boldsymbol{v}+\boldsymbol{q}}$. Due to the projection, the constraint on $\boldsymbol{C}_{\check{\boldsymbol{v}}} + \boldsymbol{C}_{\check{\boldsymbol{q}}}$ only concerns the power shaping component $\boldsymbol{P}_{\boldsymbol{v}} + \boldsymbol{P}_{\boldsymbol{q}}$ and not the impropriety component. Therefore, it is equivalent to the constraint on $\boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{C}_{\boldsymbol{q}}$ in (21).

Since $\mathcal{P}^{N_\mathrm{S}}$ is the orthogonal complement of $\mathcal{N}^{N_\mathrm{S}}$, the optimal real-valued covariance matrices $\boldsymbol{C}_{\check{\boldsymbol{v}}}^\star, \boldsymbol{C}_{\check{\boldsymbol{q}}}^\star$ can be uniquely decomposed into power shaping matrices $\boldsymbol{P}_{\boldsymbol{v}}^\star, \boldsymbol{P}_{\boldsymbol{q}}^\star$ and impropriety matrices $\boldsymbol{N}_{\boldsymbol{v}}^\star, \boldsymbol{N}_{\boldsymbol{q}}^\star$, which, in turn, uniquely determine the complex covariance matrices $\boldsymbol{C}_{\boldsymbol{v}}^\star, \boldsymbol{C}_{\boldsymbol{q}}^\star$ and pseudocovariance matrices $\tilde{\boldsymbol{C}}_{\boldsymbol{v}}^\star, \tilde{\boldsymbol{C}}_{\boldsymbol{q}}^\star$, respectively.

By comparing (26) with [39, Eq. (43)], we note that the equation for R_A is the same mathematical expression as for the sum rate in a real-valued Gaussian MIMO broadcast channel with dirty paper coding.[3] Since the channel matrices have the special block structure (2) while no particular structure is assumed for the real-valued covariance matrices, the optimization in (27) can also be understood as the maximization of the sum rate in a complex broadcast channel with dirty paper coding and general complex (possibly improper) Gaussian signals subject to a constraint $\boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{C}_{\boldsymbol{q}} = \boldsymbol{C}_{\boldsymbol{v}+\boldsymbol{q}}$ on the sum transmit covariance matrix. This observation is the key point of our proof since it can be shown that for such a setting, proper complex signals (i.e., vanishing pseudocovariance matrices) are optimal. However, this fact is stated and proven separately as Theorem 2 in Section V since the proof is nontrivial (see the discussion at the beginning of the next section).

Due to the mathematical equivalence, we can conclude that also for (27), there must exist an optimal solution that corresponds to proper complex signals in the original complex system. Summing up, there is an optimizer $\mathcal{X}_\mathrm{A}^\star$ that maximizes R_A and has the structure

$$\mathcal{X}_\mathrm{A}^\star = (\boldsymbol{C}_{\boldsymbol{v}}^\star, \boldsymbol{C}_{\boldsymbol{q}}^\star, *, *, \boldsymbol{0}, \boldsymbol{0}, *, \boldsymbol{0}). \tag{28}$$

If we let $\mathbb{X}_\mathrm{A}$ denote the set of all $\mathcal{X}_\mathrm{A}^\star$ with this structure, and let $\mathbb{X}_\mathrm{B}$ denote the set of all $\mathcal{X}_\mathrm{B}^\star$ with the structure given in (24), it is easy to verify that $\mathbb{X}_\mathrm{A} \cap \mathbb{X}_\mathrm{B} \neq \varnothing$. Therefore,

$$\mathcal{X}^\star = (\boldsymbol{C}_{\boldsymbol{v}}^\star, \boldsymbol{C}_{\boldsymbol{q}}^\star, \boldsymbol{C}_{\boldsymbol{x}_\mathrm{R}}^\star, \boldsymbol{A}^\star, \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0}) \in \mathbb{X}_\mathrm{A} \cap \mathbb{X}_\mathrm{B} \tag{29}$$

maximizes R_A and R_B simultaneously.

This shows that the upper bound in (20) is tight and that it can be achieved with vanishing pseudocovariance matrices $\tilde{\boldsymbol{C}}_{\boldsymbol{v}} = \tilde{\boldsymbol{C}}_{\boldsymbol{q}} = \boldsymbol{0}$, $\tilde{\boldsymbol{C}}_{\boldsymbol{x}_\mathrm{R}} = \boldsymbol{0}$, and $\tilde{\boldsymbol{C}}_{\boldsymbol{q}\boldsymbol{x}_\mathrm{R}} = \boldsymbol{0}$, i.e., with a transmit strategy consisting of proper signals. Since this reasoning holds for any feasible $\boldsymbol{C}_{\boldsymbol{v}+\boldsymbol{q}}$, it also holds for the optimal $\boldsymbol{C}_{\boldsymbol{v}+\boldsymbol{q}}^\star$, which proves that propriety is optimal for partial decode-and-forward with Gaussian signals. ■

[3] Note that the encoding order does not play any role for sum rate maximization (see Theorem 2).

Since both $\boldsymbol{v}$ and $\boldsymbol{q}$ are statistically independent of the relay transmit signal $\boldsymbol{x}_{\mathrm{R}}$ in the optimal Gaussian signaling, the sum $\boldsymbol{v}+\boldsymbol{q}$ can be interpreted as the innovation the source introduces into the system, and $\boldsymbol{C}_{\boldsymbol{v}+\boldsymbol{q}}$ can be called innovation covariance matrix.

In the above proof, we have seen that for a fixed innovation covariance matrix $\boldsymbol{C}_{\boldsymbol{v}+\boldsymbol{q}}$, the optimization of the partial decode-and-forward rate can be solved by separately solving the optimizations of R_{A} and R_{B}. However, it is not clear to which value the matrix $\boldsymbol{C}_{\boldsymbol{v}+\boldsymbol{q}}$ should be fixed in order to obtain a good (or even optimal) overall solution. Therefore, it is not obvious whether this decomposition is useful to derive new optimization algorithms. Nevertheless, such an approach should be investigated in future research.

On the other hand, introducing the innovation signal $\boldsymbol{v}+\boldsymbol{q}$ is very insightful from a theoretical point of view since it reveals that we have the following behavior when using proper Gaussian signaling.

Apart from the dependence on the power shaping of the innovation, the first mutual information expression R_{A} then depends only on how this power shaping is distributed between $\boldsymbol{q}$ and $\boldsymbol{v}$, but not on the statistical properties of the non-innovative signal $\boldsymbol{A}\boldsymbol{x}_{\mathrm{R}}$ and the relay signal $\boldsymbol{x}_{\mathrm{R}}$ [see (28)]. This is similar as in the cut-set bound (cf. [17]), where we also have the situation that the mutual information expression corresponding to the cut at the source depends on the innovation only. The only difference is that the mutual information expression used in the cut-set bound is equivalent to joint reception by the relay and the destination, whereas the corresponding term in the partial decode-and-forward rate results from the restriction to a particular coding scheme, which takes into account the distributed reception.

The second mutual information expression R_{B}, by contrast, only depends on the power shaping of the innovation and on the properties of $\boldsymbol{A}\boldsymbol{x}_{\mathrm{R}}$ and $\boldsymbol{x}_{\mathrm{R}}$, but not on how the innovation power is distributed between $\boldsymbol{q}$ and $\boldsymbol{v}$ [see (24)].

For an intuitive understanding of this observation, we have to recall that in the block-Markov coding scheme discussed earlier in this section, $\boldsymbol{q}$ is provided to the relay to allow future cooperation while the joint transmission that is currently taking place is represented by $\boldsymbol{A}\boldsymbol{x}_{\mathrm{R}}$. This means that in the current block, $\boldsymbol{q}$ is transmitted without the help of the relay—just like the noncooperative signal $\boldsymbol{v}$. Consequently, for R_{B}, which represents the reception of all currently transmitted signals at the destination, there is no difference between $\boldsymbol{q}$ and $\boldsymbol{v}$, and the distribution of the power between these two signals does not play a role.

We think that the insights obtained by fixing the innovation covariance matrix $\boldsymbol{C}_{\boldsymbol{v}+\boldsymbol{q}}$ might be helpful not only for the proof in this paper, but also for future research on partial decode-and-forward.

V. Excursus: Gaussian MIMO Broadcast Channel with Sum Covariance Constraint

As stated before, (27) is equivalent to a sum rate maximization in a two-user MIMO broadcast channel. As in [36], we distinguish three different cases of Gaussian MIMO broadcast channels. In a real-valued broadcast channel, the channel matrices $\boldsymbol{H}_k$ are real-valued, and the transmit signal vectors and noise vectors follow real-valued Gaussian distributions. The second case we consider is a complex broadcast channel with complex channel matrices $\boldsymbol{H}_k$, proper complex noise and proper complex transmit signals. By allowing the transmit signals to be general complex (proper or improper) Gaussian vectors even though the noise is proper, we obtain a broadcast channel with improper signaling as the third case.

Even though it is accepted as common knowledge (e.g., [30], [32], [33], [36]) and intuitively understandable (see [35] and Section I-D) that proper Gaussian signals are capacity achieving in MIMO broadcast channels with proper Gaussian noise, the existing literature on MIMO broadcast channels does (to the best of the authors' knowledge) not include any results about optimality of proper signals under a sum covariance constraint.

In [38], it was shown that dirty paper coding with proper complex Gaussian signals is capacity-achieving for a MIMO broadcast channel with a *sum power constraint*. Thus, under a sum power constraint, we could directly conclude from [38] that $\tilde{\boldsymbol{C}}_{\boldsymbol{v}} = \tilde{\boldsymbol{C}}_{\boldsymbol{q}} = \boldsymbol{0}$ in the optimum of (27). However, due to the *sum covariance constraint* in (27), this conclusion is not possible.

In [39], a real-valued broadcast channel with a sum covariance constraint was considered, but the results do not help, either, for the following two reasons. Firstly, since [39] does not assume any special structure of the channel matrices, the results from [39] do not allow conclusions about the structure of the optimal covariance matrices $\boldsymbol{C}^{\star}_{\check{\boldsymbol{v}}}, \boldsymbol{C}^{\star}_{\check{\boldsymbol{q}}}$ that we obtain for channels with the block structure (2). In particular, we cannot conclude whether or not the impropriety matrices $\boldsymbol{N}^{\star}_{\boldsymbol{v}}, \boldsymbol{N}^{\star}_{\boldsymbol{q}}$ that correspond to the optimal solution are zero matrices, i.e., whether or not the optimal pseudocovariance matrices $\tilde{\boldsymbol{C}}^{\star}_{\boldsymbol{v}}, \tilde{\boldsymbol{C}}^{\star}_{\boldsymbol{q}}$ are zero. Secondly, the constraint on the real-valued sum covariance matrix in [39] is not the same as a constraint on the power shaping component of the real-valued sum covariance (or, equivalently, the complex sum covariance) in our case.

Since optimality of proper signals under a sum covariance constraint cannot be concluded from the abovementioned references, we state this optimality in the following theorem and devote the remainder of this section to proving this statement.

Theorem 2: In a complex two-user MIMO broadcast channel with proper Gaussian noise, the optimal sum rate under a sum covariance constraint $\boldsymbol{C}_{\boldsymbol{x}_1} + \boldsymbol{C}_{\boldsymbol{x}_2} = \boldsymbol{C}$ is achievable with proper Gaussian transmit signals by using dirty paper coding with arbitrary encoding order.

Following the derivation in [38], optimality of proper signals under a sum power constraint can be shown by the help of a dual uplink, i.e., a multiple access channel. Since the classical uplink-downlink duality from [38] only holds under a sum power constraint and not under a sum covariance constraint, we make use of the uplink-downlink minimax duality with linear conic constraints from [45], [46]. This duality has been derived under the assumption of proper signals in both uplink and downlink. However, to study the case of general complex signals, we can apply the duality to the composite real representation. To do so, we first introduce some notations, equations, and lemmas.

Throughout the remainder of this section, we use $\boldsymbol{x}_k$ and $\boldsymbol{\xi}_k$, $k \in \{1,2\}$ for the transmit signals in the downlink and in the uplink, respectively, $\boldsymbol{\eta}_k$ 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 for the first two cases (real-valued and proper complex) can be expressed as

$$\begin{aligned} R_{\text{DL}} = \mu \log \frac{\det\left(\boldsymbol{C}_{\boldsymbol{\eta}_2} + \boldsymbol{H}_2 \boldsymbol{C}_{\boldsymbol{x}_2} \boldsymbol{H}_2^{\text{H}}\right)}{\det\left(\boldsymbol{C}_{\boldsymbol{\eta}_2}\right)} \\ + \mu \log \frac{\det\left(\boldsymbol{C}_{\boldsymbol{\eta}_1} + \boldsymbol{H}_1 \boldsymbol{C}_{\boldsymbol{x}_2} \boldsymbol{H}_1^{\text{H}} + \boldsymbol{H}_1 \boldsymbol{C}_{\boldsymbol{x}_1} \boldsymbol{H}_1^{\text{H}}\right)}{\det\left(\boldsymbol{C}_{\boldsymbol{\eta}_1} + \boldsymbol{H}_1 \boldsymbol{C}_{\boldsymbol{x}_2} \boldsymbol{H}_1^{\text{H}}\right)} \end{aligned} \quad (30)$$

where the pre-log factor is $\mu = 1$ in the complex case and $\mu = \frac{1}{2}$ in the real-valued case. Instead of explicitly stating a rate equation for improper signaling, we can treat the third case by studying its composite real representation as a real-valued broadcast channel.

Lemma 2: In a real-valued or proper complex two-user MIMO broadcast channel, the sum rate maximization with a sum covariance constraint $\boldsymbol{C}_{\boldsymbol{x}_1} + \boldsymbol{C}_{\boldsymbol{x}_2} \preceq \boldsymbol{C}$ has an optimizer for which the sum covariance constraint is active.

Proof: Let $R_{\text{DL}}(\boldsymbol{C}_{\boldsymbol{x}_1}, \boldsymbol{C}_{\boldsymbol{x}_2})$ denote the sum rate (30) that is achieved with the transmit covariances $\boldsymbol{C}_{\boldsymbol{x}_1}$ and $\boldsymbol{C}_{\boldsymbol{x}_2}$. Since $\boldsymbol{H}\boldsymbol{X}\boldsymbol{H}^{\text{H}} \succeq \boldsymbol{H}\boldsymbol{X}'\boldsymbol{H}^{\text{H}}$ for $\boldsymbol{X} \succeq \boldsymbol{X}'$ and $\det(\boldsymbol{X})$ is nondecreasing in $\boldsymbol{X} \succeq \mathbf{0}$ [47, Section 7.7], it is easy to verify that $R_{\text{DL}}(\boldsymbol{C}_{\text{sum}} - \boldsymbol{C}_{\boldsymbol{x}_2}, \boldsymbol{C}_{\boldsymbol{x}_2})$ is nondecreasing in $\boldsymbol{C}_{\text{sum}} = \boldsymbol{C}_{\boldsymbol{x}_1} + \boldsymbol{C}_{\boldsymbol{x}_2}$. Moreover, the constraint on the covariance matrices can be written as $\mathbf{0} \preceq \boldsymbol{C}_{\boldsymbol{x}_2} \preceq \boldsymbol{C}_{\text{sum}} \preceq \boldsymbol{C}$, where $\boldsymbol{C}_{\boldsymbol{x}_2} \preceq \boldsymbol{C}_{\text{sum}}$ is relaxed if $\boldsymbol{C}_{\text{sum}}$ is increased. ■

The following lemma states the minimax duality with linear conic constraints from [45], [46]. A short interpretation is given below.

Lemma 3: For the real-valued case and for the proper complex case, the two-user downlink minimax problem

$$\min_{\substack{\boldsymbol{C}_{\boldsymbol{\eta}_1} \succeq \mathbf{0}, \boldsymbol{C}_{\boldsymbol{\eta}_2} \succeq \mathbf{0} : (\boldsymbol{C}_{\boldsymbol{\eta}_1}, \boldsymbol{C}_{\boldsymbol{\eta}_2}) \in \mathcal{Y}^{\perp} \\ \operatorname{tr}[\boldsymbol{B}_1 \boldsymbol{C}_{\boldsymbol{\eta}_1}] + \operatorname{tr}[\boldsymbol{B}_2 \boldsymbol{C}_{\boldsymbol{\eta}_2}] = \sigma^2}} \; \max_{\substack{\boldsymbol{C}_{\boldsymbol{x}_1} \succeq \mathbf{0}, \boldsymbol{C}_{\boldsymbol{x}_2} \succeq \mathbf{0}, \boldsymbol{Z} \in \mathcal{Z} \\ \boldsymbol{C}_{\boldsymbol{x}_1} + \boldsymbol{C}_{\boldsymbol{x}_2} \preceq \boldsymbol{C} + \boldsymbol{Z}}} R_{\text{DL}} \quad (31)$$

and the two-user uplink minimax problem

$$\min_{\substack{\boldsymbol{C}_{\boldsymbol{\eta}} \succeq \mathbf{0} : \boldsymbol{C}_{\boldsymbol{\eta}} \in \mathcal{Z}^{\perp} \\ \operatorname{tr}[\boldsymbol{C} \boldsymbol{C}_{\boldsymbol{\eta}}] = \sigma^2}} \; \max_{\substack{\boldsymbol{C}_{\boldsymbol{\xi}_1} \succeq \mathbf{0}, \boldsymbol{C}_{\boldsymbol{\xi}_2} \succeq \mathbf{0}, (\boldsymbol{Y}_1, \boldsymbol{Y}_2) \in \mathcal{Y} \\ \boldsymbol{C}_{\boldsymbol{\xi}_k} \preceq \boldsymbol{B}_k + \boldsymbol{Y}_k \; \forall k}} R_{\text{UL}} \quad (32)$$

have the same optimal value, where $\mathcal{Z} \subseteq \mathbb{H}^M$ and $\mathcal{Y} \subseteq \mathbb{H}^{N_1} \times \mathbb{H}^{N_2}$ are linear subspaces ($\mathbb{S}$ instead of $\mathbb{H}$ in the real-valued case).

Proof: See [45] for the proper complex case. The real-valued case can easily be proven by repeating all steps of the proof in [45] for the real-valued setting. ■

The maximizations in the minimax problems (31) and (32) correspond to the design of the optimal transmit strategies while the minimizations can be interpreted as finding the worst-case noise properties [45], [46]. By appropriately choosing the linear subspaces $\mathcal{Y}$ and $\mathcal{Z}$, various constraints on the transmit covariance matrices such as sum power constraints or shaping constraints can be modeled [46]. Accordingly, the subspaces $\mathcal{Y}^{\perp}$ and $\mathcal{Z}^{\perp}$ restrict the noise covariance matrices so that various kinds of noise can be allowed in the worst-case noise optimization and even fixed noise covariances can be modeled (see below and [46]).

Lemma 4: In a MIMO multiple access channel with proper Gaussian noise, an optimum of the sum rate maximization under constraints on the covariance matrices $\boldsymbol{C}_{\boldsymbol{x}_1}$ and $\boldsymbol{C}_{\boldsymbol{x}_2}$ is achieved with proper transmit signals.

Proof: The sum rate in the multiple access channel is given by (e.g., [44, Section 15.3])

$$R_{\text{UL}} = \underbrace{\mathrm{h}(\boldsymbol{H}_1^{\text{H}} \boldsymbol{\xi}_1 + \boldsymbol{H}_2^{\text{H}} \boldsymbol{\xi}_2 + \boldsymbol{\eta})}_{= \mathrm{h}(\boldsymbol{y}) \leq \mathrm{h}(\boldsymbol{y}_{\text{proper}})} - \underbrace{\mathrm{h}(\boldsymbol{\eta})}_{\text{const.}} \quad (33)$$

where $\boldsymbol{y} = \boldsymbol{H}_1^{\text{H}} \boldsymbol{\xi}_1 + \boldsymbol{H}_2^{\text{H}} \boldsymbol{\xi}_2 + \boldsymbol{\eta}$, and $\boldsymbol{y}_{\text{proper}}$ is a proper Gaussian signal with the same covariance matrix as $\boldsymbol{y}$. The differential entropy $\mathrm{h}(\boldsymbol{y})$ is bounded from above by $\mathrm{h}(\boldsymbol{y}_{\text{proper}})$, and $\mathrm{h}(\boldsymbol{y}_{\text{proper}})$ only depends on the covariance matrices of $\boldsymbol{\xi}_1$, $\boldsymbol{\xi}_2$ and $\boldsymbol{\eta}$. Given any choice of these covariances, we can achieve equality $\mathrm{h}(\boldsymbol{y}) = \mathrm{h}(\boldsymbol{y}_{\text{proper}})$ by setting the pseudocovariance matrices $\tilde{\boldsymbol{C}}_{\boldsymbol{\xi}_1}$ and $\tilde{\boldsymbol{C}}_{\boldsymbol{\xi}_2}$ to zero since $\tilde{\boldsymbol{C}}_{\boldsymbol{\eta}} = \mathbf{0}$ (due to [14, Lemmas 3 and 4]). ■

Based on these lemmas, we can now prove the optimality of dirty paper coding with proper Gaussian signals under a sum covariance constraint in the two-user MIMO broadcast channel with proper Gaussian noise.

Proof of Theorem 2: Due to the equivalence with a composite real broadcast channel, we can directly conclude from [39] that dirty paper coding with Gaussian signals is optimal, but we have to show the optimality of proper signals. We can assume identity matrices as noise covariance matrices since any other case could be treated by introducing equivalent channels after noise whitening.

We express the sum rate maximization for the broadcast channel with general complex signals in the equivalent composite real broadcast channel. Similar as in the proper complex case in [46], the sum rate maximization with fixed noise covariance matrices $\boldsymbol{C}_{\check{\boldsymbol{\eta}}_1} = \frac{1}{2}\mathbf{I}_{2N_1}$ and $\boldsymbol{C}_{\check{\boldsymbol{\eta}}_2} = \frac{1}{2}\mathbf{I}_{2N_2}$ can be rewritten as a minimax problem with worst-case noise. To this end, we define a feasible set for the noise covariance matrices that is compatible with the structure of (31), but only contains one feasible element. This formulation is given in (34) in Table I. Since the constraint on the complex sum covariance matrix translates to a constraint on the power shaping component of the real-valued sum covariance matrix, we can add a matrix $\boldsymbol{Z}$ out of the impropriety space $\mathcal{N}^M$ to the constant power shaping matrix $\boldsymbol{P} = \frac{1}{2}\check{\boldsymbol{C}} \in \mathcal{P}^M$ on the right hand side of the sum covariance constraint.

Due to Lemma 3, the optimal minimax sum rate $R_{\text{UL}}^{\star}$ in the dual uplink given by (35) in Table I is equal to the optimal minimax rate $R_{\text{DL}}^{\star}$ in the downlink (34), i.e., we have $R_{\text{UL}}^{\star} = R_{\text{DL}}^{\star}$.

Since the uplink channels $\check{\boldsymbol{H}}_k^{\text{T}}$ have the block structure (2), we can return to a complex formulation. The key point now is to note that we have the constraint $\boldsymbol{C}_{\check{\boldsymbol{\eta}}} \in \mathcal{P}^M$ in the optimization for the worst-case noise. Therefore, even without solving the minimization in (35), we know that $\tilde{\boldsymbol{C}}_{\boldsymbol{\eta}} = \mathbf{0}$, i.e., the uplink noise $\boldsymbol{\eta}$ is proper. Moreover, it is easy to verify that also the constraints on the real-valued covariance

TABLE I
STEPS OF THE PROOF OF THEOREM 2.

	composite real ($\hat{=}$ general complex)	proper complex
downlink	$\min\limits_{\substack{\boldsymbol{C}_{\check{\boldsymbol{\eta}}_1}\succeq\mathbf{0},\boldsymbol{C}_{\check{\boldsymbol{\eta}}_2}\succeq\mathbf{0}:(\boldsymbol{C}_{\check{\boldsymbol{\eta}}_1},\boldsymbol{C}_{\check{\boldsymbol{\eta}}_2})\in\check{\mathcal{Y}}^{\perp}\\ \mathrm{tr}[\boldsymbol{C}_{\check{\boldsymbol{\eta}}_1}]+\mathrm{tr}[\boldsymbol{C}_{\check{\boldsymbol{\eta}}_2}]=N_1+N_2}} \max\limits_{\substack{\boldsymbol{C}_{\check{\boldsymbol{x}}_1}\succeq\mathbf{0},\boldsymbol{C}_{\check{\boldsymbol{x}}_2}\succeq\mathbf{0},\boldsymbol{Z}\in\check{\mathcal{Z}}\\ \boldsymbol{C}_{\check{\boldsymbol{x}}_1}+\boldsymbol{C}_{\check{\boldsymbol{x}}_2}\preceq\boldsymbol{P}+\boldsymbol{Z}}} R_{\mathrm{DL}} \quad (34)$ active due to Lemma 2 $\check{\mathcal{Z}}=\mathcal{N}^M$ $\check{\mathcal{Y}}^{\perp}=\{(\boldsymbol{C}_{\check{\boldsymbol{\eta}}_1},\boldsymbol{C}_{\check{\boldsymbol{\eta}}_2})\in\mathbb{S}^{2N_1}\times\mathbb{S}^{2N_2} \mid \boldsymbol{C}_{\check{\boldsymbol{\eta}}_k}=\alpha\mathbf{I}_{2N_k}\ \forall k,\ \alpha\in\mathbb{R}\}$ $R_{\mathrm{DL}}=\frac{1}{2}\log\frac{\det\left(\boldsymbol{C}_{\check{\boldsymbol{\eta}}_2}+\check{\boldsymbol{H}}_2\boldsymbol{C}_{\check{\boldsymbol{x}}_2}\check{\boldsymbol{H}}_2^{\mathrm{T}}\right)}{\det\left(\boldsymbol{C}_{\check{\boldsymbol{\eta}}_2}\right)} +\frac{1}{2}\log\frac{\det\left(\boldsymbol{C}_{\check{\boldsymbol{\eta}}_1}+\check{\boldsymbol{H}}_1\boldsymbol{C}_{\check{\boldsymbol{x}}_2}\check{\boldsymbol{H}}_1^{\mathrm{T}}+\check{\boldsymbol{H}}_1\boldsymbol{C}_{\check{\boldsymbol{x}}_1}\check{\boldsymbol{H}}_1^{\mathrm{T}}\right)}{\det\left(\boldsymbol{C}_{\check{\boldsymbol{\eta}}_1}+\check{\boldsymbol{H}}_1\boldsymbol{C}_{\check{\boldsymbol{x}}_2}\check{\boldsymbol{H}}_1^{\mathrm{T}}\right)}$ Lemma 3 $\Downarrow$	$\min\limits_{\substack{\boldsymbol{C}_{\boldsymbol{\eta}_1}\succeq\mathbf{0},\boldsymbol{C}_{\boldsymbol{\eta}_2}\succeq\mathbf{0}:(\boldsymbol{C}_{\boldsymbol{\eta}_1},\boldsymbol{C}_{\boldsymbol{\eta}_2})\in\mathcal{Y}^{\perp}\\ \mathrm{tr}[\boldsymbol{C}_{\boldsymbol{\eta}_1}]+\mathrm{tr}[\boldsymbol{C}_{\boldsymbol{\eta}_2}]=N_1+N_2}} \max\limits_{\substack{\boldsymbol{C}_{\boldsymbol{x}_1}\succeq\mathbf{0},\boldsymbol{C}_{\boldsymbol{x}_2}\succeq\mathbf{0}\\ \boldsymbol{C}_{\boldsymbol{x}_1}+\boldsymbol{C}_{\boldsymbol{x}_2}\preceq\boldsymbol{C}}} R_{\mathrm{DL,proper}} \quad (37)$ active due to Lemma 2 $\mathcal{Z}=(\mathbb{H}^M)^{\perp}=\{\mathbf{0}\}$ $\mathcal{Y}^{\perp}=\{(\boldsymbol{C}_{\boldsymbol{\eta}_1},\boldsymbol{C}_{\boldsymbol{\eta}_2})\in\mathbb{H}^{N_1}\times\mathbb{H}^{N_2} \mid \boldsymbol{C}_{\boldsymbol{\eta}_k}=\alpha\mathbf{I}_{N_k}\ \forall k,\ \alpha\in\mathbb{R}\}$ $R_{\mathrm{DL,proper}}=\log\frac{\det\left(\boldsymbol{C}_{\boldsymbol{\eta}_2}+\boldsymbol{H}_2\boldsymbol{C}_{\boldsymbol{x}_2}\boldsymbol{H}_2^{\mathrm{H}}\right)}{\det\left(\boldsymbol{C}_{\boldsymbol{\eta}_2}\right)} +\log\frac{\det\left(\boldsymbol{C}_{\boldsymbol{\eta}_1}+\boldsymbol{H}_1\boldsymbol{C}_{\boldsymbol{x}_2}\boldsymbol{H}_1^{\mathrm{H}}+\boldsymbol{H}_1\boldsymbol{C}_{\boldsymbol{x}_1}\boldsymbol{H}_1^{\mathrm{H}}\right)}{\det\left(\boldsymbol{C}_{\boldsymbol{\eta}_1}+\boldsymbol{H}_1\boldsymbol{C}_{\boldsymbol{x}_2}\boldsymbol{H}_1^{\mathrm{H}}\right)}$
uplink	$\min\limits_{\substack{\boldsymbol{C}_{\check{\boldsymbol{\eta}}}\succeq\mathbf{0}:\boldsymbol{C}_{\check{\boldsymbol{\eta}}}\in\check{\mathcal{Z}}^{\perp}\\ \mathrm{tr}[\boldsymbol{P}\boldsymbol{C}_{\check{\boldsymbol{\eta}}}]=N_1+N_2}} \max\limits_{\substack{\boldsymbol{C}_{\check{\boldsymbol{\xi}}_1}\succeq\mathbf{0},\boldsymbol{C}_{\check{\boldsymbol{\xi}}_2}\succeq\mathbf{0},(\boldsymbol{Y}_1,\boldsymbol{Y}_2)\in\check{\mathcal{Y}}\\ \boldsymbol{C}_{\check{\boldsymbol{\xi}}_k}\preceq\mathbf{I}_{2N_k}+\boldsymbol{Y}_k\ \forall k}} R_{\mathrm{UL}} \quad (35)$ $\check{\mathcal{Z}}^{\perp}=(\mathcal{N}^M)^{\perp}=\mathcal{P}^M$ due to Lemma 1 $\check{\mathcal{Y}}=\left\{(\boldsymbol{Y}_1,\boldsymbol{Y}_2)\in\mathbb{S}^{2N_1}\times\mathbb{S}^{2N_2} \,\middle\vert\, \mathrm{tr}[\boldsymbol{Y}_1]+\mathrm{tr}[\boldsymbol{Y}_2]=0\right\}$ $R_{\mathrm{UL}}=\frac{1}{2}\log\frac{\det\left(\boldsymbol{C}_{\check{\boldsymbol{\eta}}}+\check{\boldsymbol{H}}_1^{\mathrm{T}}\boldsymbol{C}_{\check{\boldsymbol{\xi}}_1}\check{\boldsymbol{H}}_1+\check{\boldsymbol{H}}_2^{\mathrm{T}}\boldsymbol{C}_{\check{\boldsymbol{\xi}}_2}\check{\boldsymbol{H}}_2\right)}{\det\left(\boldsymbol{C}_{\check{\boldsymbol{\eta}}}\right)}$ $\xRightarrow{\text{Lemma 4}}$	Lemma 3 $\Uparrow$ $\min\limits_{\substack{\boldsymbol{C}_{\boldsymbol{\eta}}\succeq\mathbf{0}\\ \mathrm{tr}[\boldsymbol{C}\boldsymbol{C}_{\boldsymbol{\eta}}]=N_1+N_2}} \max\limits_{\substack{\boldsymbol{C}_{\boldsymbol{\xi}_1}\succeq\mathbf{0},\boldsymbol{C}_{\boldsymbol{\xi}_2}\succeq\mathbf{0},(\boldsymbol{Y}_1,\boldsymbol{Y}_2)\in\mathcal{Y}\\ \boldsymbol{C}_{\boldsymbol{\xi}_k}\preceq\mathbf{I}_{N_k}+\boldsymbol{Y}_k\ \forall k}} R_{\mathrm{UL,proper}} \quad (36)$ $\mathcal{Z}^{\perp}=\mathbb{H}^M$ (any $\boldsymbol{C}_{\boldsymbol{\eta}}$ corresponds to a $\boldsymbol{P}_{\boldsymbol{\eta}}\in\check{\mathcal{Z}}^{\perp}=\mathcal{P}^M$) $\mathcal{Y}=\left\{(\boldsymbol{Y}_1,\boldsymbol{Y}_2)\in\mathbb{H}^{N_1}\times\mathbb{H}^{N_2} \,\middle\vert\, \mathrm{tr}[\boldsymbol{Y}_1]+\mathrm{tr}[\boldsymbol{Y}_2]=0\right\}$ $R_{\mathrm{UL,proper}}=\log\frac{\det\left(\boldsymbol{C}_{\boldsymbol{\eta}}+\boldsymbol{H}_1^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{\xi}_1}\boldsymbol{H}_1+\boldsymbol{H}_2^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{\xi}_2}\boldsymbol{H}_2\right)}{\det\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right)}$

matrices $\boldsymbol{C}_{\check{\boldsymbol{\xi}}_1}$ and $\boldsymbol{C}_{\check{\boldsymbol{\xi}}_2}$ translate only to constraints on the complex covariance matrices $\boldsymbol{C}_{\boldsymbol{\xi}_1}$ and $\boldsymbol{C}_{\boldsymbol{\xi}_2}$ and do not affect the pseudocovariance matrices $\tilde{\boldsymbol{C}}_{\boldsymbol{\xi}_1}$ and $\tilde{\boldsymbol{C}}_{\boldsymbol{\xi}_2}$.[4] Consequently, due to Lemma 4, proper signals achieve the optimum in the complex representation of the uplink, i.e., we have $R^{\star}_{\mathrm{UL,proper}}=R^{\star}_{\mathrm{UL}}=R^{\star}_{\mathrm{DL}}$, where $R^{\star}_{\mathrm{UL,proper}}$ is the optimizer of (36) in Table I.[5]

Again due to the uplink-downlink duality (Lemma 3), the optimizer $R^{\star}_{\mathrm{DL,proper}}$ of (37) in Table I fulfills $R^{\star}_{\mathrm{DL,proper}}=R^{\star}_{\mathrm{UL,proper}}$. The chosen encoding order is without loss of generality: as usual in uplink-downlink duality (cf., e.g., [38]), the encoding order in the downlink is the reverse uplink decoding order, and from the symmetry in (33), we can conclude that this order can be chosen arbitrarily.

The constraints in the worst-case noise optimization correspond again to a feasible set consisting of only one element, i.e., to fixed noise covariance matrices $\boldsymbol{C}_{\boldsymbol{\eta}_1}=\mathbf{I}_{N_1}$ and $\boldsymbol{C}_{\boldsymbol{\eta}_2}=\mathbf{I}_{N_2}$. This is in accordance with the original complex broadcast channel.

Due to Lemma 2, the sum covariance constraint of the proper complex downlink minimax problem is active in the optimum. This shows that we can indeed find a proper complex downlink strategy with $\boldsymbol{C}_{\boldsymbol{x}_1}+\boldsymbol{C}_{\boldsymbol{x}_2}=\boldsymbol{C}$ which achieves $R^{\star}_{\mathrm{DL,proper}}=R^{\star}_{\mathrm{UL,proper}}=R^{\star}_{\mathrm{UL}}=R^{\star}_{\mathrm{DL}}$. ■

This concludes our excursus on the MIMO broadcast channel, which was necessary as an ingredient for the proof of the main theorem of this paper in Section IV.

VI. SUMMARY AND OUTLOOK

We have shown that the zero-mean proper Gaussian distribution is the optimal one among all Gaussian distributions for the problem of optimizing the partial decode-and-forward rate in a Gaussian MIMO relay channel. A key point in the proof was to decompose the optimization in a way that we could exploit the mathematical equivalence between one of the arising expressions and the optimization of the sum rate of a Gaussian MIMO broadcast channel under a sum covariance constraint.

Since for such broadcast channels it is not only known that proper signals are optimal under the assumption of Gaussian signals, but also that the Gaussian distribution is the optimal input distribution, the question arises whether the same proof technique can be applied to show optimality of Gaussian signals in Gaussian MIMO relay channels with partial decode-and-forward. Unfortunately, this is not the case since the

[4] In fact, the constraints are equivalent to a sum power constraint (similar as in the proper complex case in [46]): since $\boldsymbol{Y}_k$ can be optimized and only has to satisfy a trace constraint, the constraint on $\boldsymbol{C}_{\check{\boldsymbol{\xi}}_k}$ is equivalent to a trace constraint $\mathrm{tr}[\boldsymbol{C}_{\check{\boldsymbol{\xi}}_k}]\leq 2N_k+\mathrm{tr}[\boldsymbol{Y}_k]$. As the trace constraint on $\boldsymbol{Y}_k$ only affects the sum $\boldsymbol{Y}_1+\boldsymbol{Y}_2$, the individual constraints on $\mathrm{tr}[\boldsymbol{C}_{\check{\boldsymbol{\xi}}_k}]$ become equivalent to a sum power constraint $\mathrm{tr}[\boldsymbol{C}_{\check{\boldsymbol{\xi}}_1}+\boldsymbol{C}_{\check{\boldsymbol{\xi}}_2}]\leq 2N_1+2N_2$.

[5] Based on (2), we would obtain $\frac{1}{2}\boldsymbol{C}$ instead of $\boldsymbol{C}$ in the trace constraint and $\frac{1}{2}\boldsymbol{C}_{\boldsymbol{\xi}_k}$ instead of $\boldsymbol{C}_{\boldsymbol{\xi}_k}$ in the covariance constraints. However, since jointly scaling the signal covariance matrices and the noise covariance does not change the data rate, these two factors can be canceled out against each other.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TSP.2014.2321737

equivalence that we have observed is only an equivalence of the rate expressions obtained *after* applying the assumption of Gaussian signals. However, as explained in the following, the underlying mutual information expressions are not equivalent.

Below (25), we have shown that $\boldsymbol{q}$ and $\boldsymbol{x}_{\mathrm{R}}$ have to be independent. Thus, R_{A} can be written as

$$R_{\mathrm{A}} = \mathrm{I}(\boldsymbol{v}; \boldsymbol{H}_{\mathrm{SD}}(\boldsymbol{v}+\boldsymbol{q})+\boldsymbol{\eta}_{\mathrm{D}} \,|\, \boldsymbol{q})+\mathrm{I}(\boldsymbol{q}; \boldsymbol{H}_{\mathrm{SR}}(\boldsymbol{v}+\boldsymbol{q})+\boldsymbol{\eta}_{\mathrm{R}}). \quad (38)$$

This can be interpreted as a broadcast channel with a restriction to superposition coding and with a genie that provides the receiver D with knowledge of the interfering signal $\boldsymbol{q}$.

In the conventional MIMO broadcast channel model, we do not have the restriction to superposition coding and we do not have the abovementioned genie. Dirty paper coding is a special case of the Gelfand-Pinsker coding scheme. Using the Gelfand-Pinsker scheme in a broadcast channel without the assumption of Gaussian signaling, we have [19, Section 7.4]

$$\begin{aligned} R_{\mathrm{BC}} = &\, \mathrm{I}(\boldsymbol{s}_1; \boldsymbol{H}_1 \boldsymbol{x}(\boldsymbol{s}_1, \boldsymbol{s}_2) + \boldsymbol{\eta}_1) - \mathrm{I}(\boldsymbol{s}_1; \boldsymbol{s}_2) \\ &+ \mathrm{I}(\boldsymbol{s}_2; \boldsymbol{H}_2 \boldsymbol{x}(\boldsymbol{s}_1, \boldsymbol{s}_2) + \boldsymbol{\eta}_2) \end{aligned} \quad (39)$$

where the transmit signal $\boldsymbol{x}(\boldsymbol{s}_1, \boldsymbol{s}_2)$ is a function of the two input signals $\boldsymbol{s}_1, \boldsymbol{s}_2$, and $\boldsymbol{s}_2$ is considered as a disturbance that is known to the transmitter when encoding the signals for user 1, but not to the receiver.

Obviously, the expressions in (38) and (39) do not match. However, in the case of Gaussian signals, there exists a choice of $\boldsymbol{x}(\bullet, \bullet)$ that leads to the same rate as if no interference was present at user 1 [19, Section 6.7], i.e., as if there was a genie. This explains why we have obtained a mathematical equivalence of the rate expressions for the two scenarios under the assumption of Gaussian signals in Section IV.

We know that Gaussian signals are optimal for R_{BC}, i.e., using non-Gaussian signals reduces R_{BC}. On the other hand, equality between R_{A} and R_{BC} no longer holds without the assumption of Gaussian signals. Therefore, it might be possible to increase R_{A} by using non-Gaussian signals despite the fact that R_{BC} decreases. We conclude that it is not obvious whether R_{A} can be increased by allowing non-Gaussian input distributions. We also remark that the proof of optimality of proper signals was performed under the assumption of Gaussian signals and does not allow conclusions on whether or not improper non-Gaussian signaling can outperform proper non-Gaussian signaling. These questions are left open for future research.

References

[1] B. Wang, J. Zhang, and A. Høst-Madsen, "On the capacity of MIMO relay channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 29–43, Jan. 2005.

[2] Y. Fan and J. Thompson, "MIMO configurations for relay channels: Theory and practice," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1774–1786, May 2007.

[3] 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.

[4] Y. Rong and F. Gao, "Optimal beamforming for non-regenerative MIMO relays with direct link," *IEEE Commun. Lett.*, vol. 13, no. 12, pp. 926–928, 2009.

[5] S. Simoens, O. Muñoz-Medina, J. Vidal, and A. del Coso, "On the Gaussian MIMO relay channel with full channel state information," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3588–3599, Sep. 2009.

[6] ——, "Compress-and-forward cooperative MIMO relaying with full channel state information," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 781–791, Feb. 2010.

[7] C. T. K. Ng and G. J. Foschini, "Transmit signal and bandwidth optimization in multiple-antenna relay channels," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 2987–2992, Jul. 2011.

[8] L. Gerdes and W. Utschick, "Optimized capacity bounds for the MIMO relay channel," in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP) 2011*, May 2011, pp. 3336–3339.

[9] 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.

[10] 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.

[11] L. Gerdes, L. Weiland, M. Riemensberger, and W. Utschick, "Optimal partial decode-and-forward rates for stochastically degraded Gaussian relay channels," presented at the 48th Annual Conf. Inf. Sci. Syst. (CISS), Princeton, NJ, USA., Mar. 2014.

[12] I. Shomorony and A. S. Avestimehr, "Is Gaussian noise the worst-case additive noise in wireless networks?" in *Proc. Int. Symp. Inf. Theory (ISIT) 2012*, Jul. 2012, pp. 214–218.

[13] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1293–1302, Jul. 1993.

[14] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov./Dec. 1999.

[15] L. Gerdes, M. Riemensberger, and W. Utschick, "Bounds on the capacity regions of half-duplex Gaussian MIMO relay channels," *EURASIP J. Advances Signal Process.*, vol. 2013, no. 43, pp. 1–17, Mar. 2013.

[16] E. C. van der Meulen, "Three-terminal communication channels," *Advances in Applied Probability*, vol. 3, no. 1, pp. 120–154, 1971.

[17] 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.

[18] 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.

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

[20] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[21] A. El Gamal, M. Mohseni, and S. Zahedi, "Bounds on capacity and minimum energy-per-bit for AWGN relay channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1545–1561, Apr. 2006.

[22] D. P. Mandic and S. L. Goh, *Complex valued nonlinear adaptive filters: noncircularity, widely linear and neural models*. West Sussex, UK: John Wiley & Sons, 2009.

[23] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals*. Cambridge, UK: Cambridge University Press, 2010.

[24] T. Adalı, P. J. Schreier, and L. L. Scharf, "Complex-valued signal processing: The proper way to deal with impropriety," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5101–5125, Nov. 2011.

[25] 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.

[26] A. El Gamal and S. Zahedi, "Capacity of a class of relay channels with orthogonal components," *IEEE Trans. Inf. Theory*, vol. 51, no. 5, pp. 1815–1817, May 2005.

[27] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical J.*, vol. 27, pp. 379–423 and 623–656, Jul./Oct. 1948.

[28] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 49–60, Jan. 1981.

[29] I. Sason, "On achievable rate regions for the Gaussian interference channel," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1345–1356, Jun. 2004.

[30] V. R. Cadambe, S. A. Jafar, and C. Wang, "Interference alignment with asymmetric complex signaling—settling the Høst-Madsen–Nosratinia conjecture," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4552–4565, Sep. 2010.

[31] Z. K. M. Ho and E. Jorswieck, "Improper Gaussian signaling on the two-user SISO interference channel," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3194–3203, Sep. 2012.

[32] Y. Zeng, C. M. Yetis, E. Gunawan, Y. L. Guan, and R. Zhang, "Transmit optimization with improper Gaussian signaling for interference channels," *IEEE Trans. Signal Process.*, vol. 61, no. 11, pp. 2899–2913, Jun. 2013.

[33] Y. Zeng, R. Zhang, E. Gunawan, and Y. L. Guan, "Optimized transmission with improper Gaussian signaling in the K-user MISO interference channel," Oct. 2013, unpublished, preprint available at http://arxiv.org/abs/1303.2766.

[34] C. Lameiro and I. Santamaría, "Degrees-of-freedom for the 4-user SISO interference channel with improper signaling," in *Proc. Int. Conf. Commun. (ICC) 2013*, Jun. 2013, pp. 1646–1650.

[35] C. Hellings and W. Utschick, "Performance gains due to improper signals in MIMO broadcast channels with widely linear transceivers," in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP) 2013*, May 2013, pp. 4379–4383.

[36] C. Hellings, M. Joham, and W. Utschick, "QoS feasibility in MIMO broadcast channels with widely linear transceivers," *IEEE Signal Process. Lett.*, vol. 20, no. 11, pp. 1134–1137, Nov. 2013.

[37] Z. Ho and E. Jorswieck, "Secure degrees of freedom on widely linear instantaneous relay-assisted interference channel," in *Proc. IEEE 14th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2013, pp. 684–688.

[38] 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.

[39] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.

[40] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Trans. Signal Process.*, vol. 43, no. 8, pp. 2030–2033, Aug. 1995.

[41] Z. Ho, A. Zappone, and E. Jorswieck, "Optimal non-regenerative relay processing with improper signals," in *Proc. 10th Int. Symp. Wireless Commun. Syst. (ISWCS)*, Aug. 2013, pp. 567–571.

[42] J. Zhang and M. Haardt, "Widely linear signal processing for two-way relaying with MIMO amplify and forward relays," in *Proc. 10th Int. Symp. Wireless Commun. Syst. (ISWCS)*, Aug. 2013, pp. 264–268.

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

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

[45] A. Dotzler, M. Riemensberger, and W. Utschick, "Uplink-downlink minimax duality with linear conic constraints," 2012, to be published.

[46] 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.

[47] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, UK: Cambridge University Press, 1985.

Christoph Hellings (S'10) received the B.Sc. degree and the Dipl.-Ing. degree in electrical engineering (both with high distinction) from the Technische Universität München (TUM) in 2008 and 2010, respectively.

He is currently working towards the doctoral degree at the Associate Institute for Signal Processing, TUM. His research interests include energy efficient communications, improper signaling, and the optimization of linear precoding strategies in multicarrier systems.

Mr. Hellings was awarded a scholarship of the Max Weber Program of the Bavarian state, and for his Diploma thesis, he received an award of the German Association for Electrical, Electronic & Information Technologies (VDE). In 2011, he received an outstanding teaching assistant award from the student representatives of the Department of Electrical Engineering and Information Technology, TUM.

Lennart Gerdes (S'10) received the B.Sc. degree and the Dipl.-Ing. degree (with high distinction) in electrical engineering in 2008 and 2009, respectively, both from the Technische Universität München (TUM), Germany.

He is currently pursuing his doctoral degree at the Associate Institute for Signal Processing, TUM. His research interests include resource allocation in wireless communication networks, especially MIMO relay networks, and network information theory.

Mr. Gerdes was a co-recipient of a Student Best Paper Award at the 14th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2013). In 2013, he received an outstanding teaching assistant award from the student representatives of the Department of Electrical Engineering and Information Technology, TUM.

Lorenz Weiland received the B.Sc. degree in electrical engineering and economics in 2009 and the B.Sc. degree in electrical engineering in 2010 both from Technische Universität Darmstadt. In 2012, he received the M.Sc. degree in electrical engineering (with high distinction) from Technische Universität München (TUM).

He is currently working towards the doctoral degree at the Associate Institute for Signal Processing, TUM. His research interests include signal processing in multiantenna systems, especially for MIMO relaying and in MIMO radar systems.

Mr. Weiland received a Student Best Paper Award at the 14th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2013).

Wolfgang Utschick (SM'06) completed several industrial education programs before he received the diploma ('93) and doctoral degrees ('98) in electrical engineering, both with honors, from Technische Universität München (TUM). In 2002 Dr. Utschick was appointed Professor at TUM where he is director of the Fachgebiet Methoden der Signalverarbeitung (Associate Institute for Signal Processing). Dr. Utschick teaches courses on Signal Processing, Stochastic Processes, and Optimization Theory in the field of Wireless Communications and Signal Processing applications. In 2005, he has been guest professor at the University of Edinburgh and since 2011 he is serving as a regular guest professor at the Singapur Institute of Technology, where he teaches undergraduate and graduate courses. He holds some 20 patents in the field of multiantenna signal processing and has authored and co-authored more than 300 technical articles in international journals and conference proceedings. He has edited several books and is founder and editor of the Springer book series Foundations in Signal Processing, Communications and Networking. Dr. Utschick has been Principal Investigator in multiple research projects funded by the German Research Fund (DFG). He currently serves as a coordinator and spokesman of the German DFG focus program Communications over Interference limited Networks (COIN) which is devoted to topics as cooperative communications, crosslayer design, ad-hoc wireless networks, etc. He is a member of the VDE and senior member of the IEEE where he recently finished his service as an associate editor for IEEE Transactions on Signal Processing after his third term. Already before, from '06 to '09, he was instrumental as an associate editor for IEEE Transactions on Circuits and Systems 1 and 2. He is elected member of the IEEE Signal Processing Society Technical Committee on Signal Processing for Communications and Networking since 2010 during his second term. Since 2011 he is in charge of study affairs (Dean for Academic Affairs) and member of the steering committee of the Faculty of Electrical Engineering and Information Technology. In 2014 he has been elected Chair of Deans for Academic Affairs at TUM and member of the TUM Extended Board of Management.