Full-Duplex Bidirectional MIMO: Achievable Rates under Limited Dynamic Range

Brian P. Day,* Daniel W. Bliss,† Adam R. Margetts,† and Philip Schniter*
*Dept. of ECE, The Ohio State University, Columbus, OH 43210. Email: day.262@osu.edu, schniter@ece.osu.edu
†Advanced Sensor Techniques Group, MIT Lincoln Laboratory, Lexington, MA. Email: bliss@ll.mit.edu, margetts@ieee.org

***Abstract*—In this paper we consider the problem of full-duplex bidirectional communication between a pair of modems, each with multiple transmit and receive antennas. The principal difficulty in implementing such a system is that, due to the close proximity of each modem's transmit antennas to its receive antennas, each modem's outgoing signal can exceed the dynamic range of its input circuitry, making it difficult—if not impossible—to recover the desired incoming signal. To address these challenges, we consider systems that use pilot-aided channel estimates to perform transmit beamforming, receive beamforming, and interference cancellation. Modeling transmitter/receiver dynamic-range limitations explicitly, we derive tight upper and lower bounds on the achievable sum-rate, and propose a transmission scheme based on maximization of the lower bound, which requires us to (numerically) solve a nonconvex optimization problem. In addition, we derive an analytic approximation to the achievable sum-rate, and show, numerically, that it is quite accurate.** [1]

I. Introduction

Full-duplex bidirectional communication between two multiple-input multiple-output (MIMO) wireless modems has the potential to nearly double the system spectral efficiency [1]. By full-duplex, we mean that the two modems perform simultaneous transmission and reception (STAR) at the same carrier frequency. The fundamental difficulty with STAR is that, due to the close proximity of a given modem's transmit antennas to its receive antennas, the modem's outgoing signal can overwhelm its receiver circuitry, making it impossible to recover the incoming signal. To avoid this problem, existing practical systems tend to communicate in half-duplex mode (e.g., time-division duplex or frequency-division duplex). In this paper, we propose a realistic system model, including channel estimation errors and effects of limited dynamic range, and derive achievable-rate bounds for a proposed MIMO STAR protocol. It is shown that its spectral efficiency is uniformly better than optimized half-duplex and nearly double when operating within dynamic range constraints.

In this work, we assume that each of the two modems uses $N_{\mathsf{t}} \geq 1$ antennas for transmission and $N_{\mathsf{r}} \geq 1$ different antennas for reception (i.e., MIMO modems), and we assume a per-modem transmit power constraint. We then consider the problem of jointly optimizing the MIMO transmission and reception strategies in order to maximize the sum of the rates of reliable communication between the two modems (i.e., the sum-rate).

From our perspective, the primary challenges of MIMO-STAR are, in practice, due to the following:

1) high channel dynamic range (DR),
2) limited transmitter and receiver DR, and
3) imperfect channel state information (CSI).

Channel DR refers to the ratio of the (nominal) interference-channel gain to the (nominal) desired-channel gain, which may be as high as 100dB due to the relative separation between intra- and inter-modem antenna pairs. Limited transmitter and receiver-DR is a natural consequence of non-ideal amplifiers, oscillators, analog-to-digital converters (ADCs), and digital-to-analog converters (DACs). Imperfect CSI can result for several reasons, including channel time-variation, additive noise, and DR limitations.

Due to the practical importance of transmitter/receiver-DR and imperfect CSI, we model each artifact explicitly in this work. In particular, we model limited transmitter-DR by injecting, for each transmit antenna, an additive white Gaussian "transmitter noise" with variance κ times the energy of the intended transmit signal. Similarly, we model limited receiver-DR by injecting, for each receive antenna, an additive white Gaussian "receiver distortion" with variance β times the energy impinging on that receive antenna. Finally, we model CSI imperfections by assuming the use of pilot-aided least-squares (LS) channel estimation.

The problem that we consider, full-duplex bidirectional MIMO, is reminiscent yet fundamentally different than the well-studied two-user *MIMO interference channel* (ICh) problem [11], for which *interference alignment* [12] has recently been proposed. While in both problems the primary challenge is mitigation of other-user interference, in the MIMO-ICh problem, the other-user codewords are unknown, whereas in our problem, they are perfectly known because they are self-generated. In fact, in our problem, "other-user" interference manifests only through channel-estimation error and limited receiver-DR, both of which can become significant under very high channel-DR (e.g., 100 dB).

The contributions of this paper are as follows. For the full-duplex bidirectional MIMO communication problem: 1) an explicit model for transmitter/receiver-DR limitations is proposed; 2) pilot-aided least-squares MIMO-channel estimation, under DR limitations, is analyzed; 3) the residual

[1] This work was sponsored by the Defense Advanced Research Projects Agency under Air Force contract FA8721-05-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Government.

self-interference, resulting from DR limitations and channel-estimation error, is analyzed; 4) lower and upper bounds on the achievable sum-rate are derived; 5) a transmission scheme is proposed based on maximizing the sum-rate lower bound subject to a power constraint, requiring the solution of a nonconvex optimization problem; 6) an analytic approximation of the maximum achievable sum-rate is proposed; and, 7) the achievable sum-rate is numerically investigated as a function of signal-to-noise ratio, interference-to-noise ratio, transmitter/receiver dynamic range, and number of antennas.

II. System Model

Our bidirectional communication problem involves two modems ("A" and "B"), and thus two communicating transmitter-receiver pairs ($i \in \{1,2\}$). We assume, without loss of generality, that modem A houses transmitter $i=1$ and receiver $i=2$, whereas modem B houses transmitter $i=2$ and receiver $i=1$. In the sequel, we use $t \in \mathbb{Z}^+$ to denote the channel-use index, $\boldsymbol{s}_i(t) \in \mathbb{C}^{N_\mathsf{t}}$ to denote the noisy signal radiated by the antenna array of transmitter i, and $\boldsymbol{u}_j(t) \in \mathbb{C}^{N_\mathsf{r}}$ to denote the undistorted signal collected by the antenna array of receiver j, where N_t is the number of transmit antennas and N_r is the number of receive antennas.

We assume that the signal radiated by transmitter j and collected by receiver i propagates through an additive white Gaussian noise (AWGN) corrupted Raleigh-fading MIMO channel $\boldsymbol{H}_{ij} \in \mathbb{C}^{N_\mathsf{r}\times N_\mathsf{t}}$. By "Rayleigh fading," we mean that $\mathrm{vec}(\boldsymbol{H}_{ij}) \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_{N_\mathsf{r}N_\mathsf{t}})$. The time-$t$ radiated signals $\{\boldsymbol{s}_j(t)\}_{j=1}^2$ are then related to each received signals $\boldsymbol{u}_i(t)$ via

$$\boldsymbol{u}_1(t) = \sqrt{\rho}\boldsymbol{H}_{11}\boldsymbol{s}_1(t) + \sqrt{\eta}\boldsymbol{H}_{12}\boldsymbol{s}_2(t) + \boldsymbol{n}_1(t) \quad (1)$$

$$\boldsymbol{u}_2(t) = \sqrt{\rho}\boldsymbol{H}_{22}\boldsymbol{s}_2(t) + \sqrt{\eta}\boldsymbol{H}_{21}\boldsymbol{s}_1(t) + \boldsymbol{n}_2(t). \quad (2)$$

In (1)-(2), $\boldsymbol{n}_i(t) \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_{N_\mathsf{r}})$ denotes AWGN, $\rho > 0$ denotes the signal-to-noise ratio (SNR), and $\eta > 0$ denotes the interference-to-noise ratio (INR). The size of η will depend on, e.g., antenna separation and analog-domain suppression.

We assume that the signaling epoch $\mathcal{T}$ is partitioned into a training period $\mathcal{T}_{\mathsf{train}}$ and a subsequent data communication period $\mathcal{T}_{\mathsf{data}}$. For reasons that will become clear in the sequel, the training period is itself partitioned into two equal-length portions (i.e., $\mathcal{T}_{\mathsf{train}}[1]$ and $\mathcal{T}_{\mathsf{train}}[2]$), as is the data period (i.e., $\mathcal{T}_{\mathsf{data}}[1]$ and $\mathcal{T}_{\mathsf{data}}[2]$). Within each of these four sub-periods, we assume that the transmitted signals are zero-mean and wide-sense stationary.

We model the effect of limited transmitter dynamic range (DR) by injecting, per transmit antenna, an independent zero-mean Gaussian "transmitter noise" whose variance is κ times the energy of the *intended* transmit signal at that antenna. In particular, say that $\boldsymbol{x}_j(t) \in \mathbb{C}^{N_\mathsf{t}}$ denotes the the j^{th} transmitter's intended time-t transmit signal, and say $\boldsymbol{Q}_j \triangleq \mathrm{Cov}\{\boldsymbol{x}_j(t)\}$ over the relevant time period (e.g., $t \in \mathcal{T}_{\mathsf{data}}[1]$). We then write the time-t noisy radiated signal as

$$\boldsymbol{s}_j(t) = \boldsymbol{x}_j(t) + \boldsymbol{c}_j(t) \text{ s.t.} \begin{cases} \boldsymbol{c}_j(t) \sim \mathcal{CN}(\boldsymbol{0}, \kappa\,\mathrm{diag}(\boldsymbol{Q}_j)) \\ \boldsymbol{c}_j(t) \perp\!\!\!\perp \boldsymbol{x}_j(t) \\ \boldsymbol{c}_j(t) \perp\!\!\!\perp \boldsymbol{c}_j(t')\big|_{t'\neq t} \end{cases}, \quad (3)$$

where $\boldsymbol{c}_j(t) \in \mathbb{C}^{N_\mathsf{t}}$ denotes the transmitter noise. Typically, $\kappa \ll 1$. The model (3) closely approximates the combined effects of additive power-amp noise, non-linearities in the DAC and power-amp, and oscillator phase noise (e.g., [14]).

We model the effect of limited receiver-DR by injecting, per receive antenna, an independent zero-mean Gaussian "receiver distortion" whose variance is β times the energy collected by that antenna. In particular, say that $\boldsymbol{u}_i(t) \in \mathbb{C}^{N_\mathsf{r}}$ denotes the i^{th} receiver's undistorted time-t received vector, and say $\boldsymbol{\Phi}_i \triangleq \mathrm{Cov}\{\boldsymbol{u}_i(t)\}$ over the relevant time period (e.g., $t \in \mathcal{T}_{\mathsf{data}}[1]$). We then write the distorted post-ADC received signal as

$$\boldsymbol{y}_i(t) = \boldsymbol{u}_i(t) + \boldsymbol{e}_i(t) \text{ s.t.} \begin{cases} \boldsymbol{e}_i(t) \sim \mathcal{CN}(\boldsymbol{0}, \beta\,\mathrm{diag}(\boldsymbol{\Phi}_i)) \\ \boldsymbol{e}_i(t) \perp\!\!\!\perp \boldsymbol{u}_i(t) \\ \boldsymbol{e}_i(t) \perp\!\!\!\perp \boldsymbol{e}_i(t')\big|_{t'\neq t} \end{cases}, \quad (4)$$

where $\boldsymbol{e}_i(t) \in \mathbb{C}^{N_\mathsf{r}}$ is additive distortion. Typically, $\beta \ll 1$. The model (4) closely approximates the combined effects of additive gain-control noise, non-linearities in the ADC and gain-control, and oscillator phase noise (e.g., [15]).

Figure 1 summarizes our model.

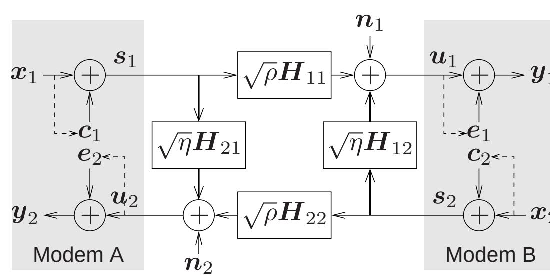

Fig. 1. A model of bidirectional MIMO communication under limited transmitter/receiver-DR. The dashed lines denote statistical dependence. In labeled quantities, the time index t has been suppressed for brevity.

III. Analysis of Achievable Sum-Rate

A. Pilot-Aided Channel Estimation

In this section, we describe the pilot-aided channel estimation procedure that is used to learn the channel matrices $\{\boldsymbol{H}_{ij}\}$. In our protocol, the training interval consists of two sub-periods, $\mathcal{T}_{\mathsf{train}}[1]$ and $\mathcal{T}_{\mathsf{train}}[2]$, each of duration TN_t channel uses (for some $T \in \mathbb{Z}^+$). For all $t \in \mathcal{T}_{\mathsf{train}}[1]$, we assume that transmitter $j=1$ transmits a known pilot signal and $j=2$ remains silent, while, for all $t \in \mathcal{T}_{\mathsf{train}}[2]$, $j=2$ transmits and $j=1$ remains silent. As we shall see, it suffices to choose the pilot sequence $\boldsymbol{X}_j = [\boldsymbol{x}_j(1), \dots, \boldsymbol{x}_j(TN_\mathsf{t})] \in \mathbb{C}^{N_\mathsf{t}\times TN_\mathsf{t}}$ arbitrarily so long as it satisfies $\frac{1}{2T}\boldsymbol{X}_j\boldsymbol{X}_j^\mathsf{H} = \boldsymbol{I}_{N_\mathsf{t}}$, where the scaling is chosen to satisfy a per-period power constraint of the form $\mathrm{tr}(\boldsymbol{Q}_j) = 2$, consistent with the data power constraints that will be described in the sequel.

Our limited transmitter/receiver-DR model implies that the (distorted) space-time pilot signal observed by receiver i is

$$\boldsymbol{Y}_i = \sqrt{\alpha_{ij}}\boldsymbol{H}_{ij}(\boldsymbol{X}_j + \boldsymbol{C}_j) + \boldsymbol{N}_i + \boldsymbol{E}_i, \quad (5)$$

where, for notational convenience, we define

$$\alpha_{ij} \triangleq \begin{cases} \rho & \text{if } i = j \\ \eta & \text{if } i \neq j. \end{cases} \quad (6)$$

In (5), $\boldsymbol{C}_j, \boldsymbol{E}_i$ and $\boldsymbol{N}_i$ are $N_\mathsf{t} \times TN_\mathsf{t}$ matrices of transmitter noise, receiver distortion, and AWGN, respectively. At the conclusion of training, we assume that the i^{th} receiver estimates the channels $\{\boldsymbol{H}_{ij}\}_{j=1}^2$ via least-squares (LS), yielding

$$\sqrt{\alpha_{ij}}\hat{\boldsymbol{H}}_{ij} \triangleq \tfrac{1}{2T}\boldsymbol{Y}_i\boldsymbol{X}_j^\mathsf{H}, \tag{7}$$

and communicates them to the other modem.In the sequel, it will be useful to decompose the channel estimate into the true channel plus some estimation error. It can be shown that such a decomposition takes the form of

$$\sqrt{\alpha_{ij}}\hat{\boldsymbol{H}}_{ij} = \sqrt{\alpha_{ij}}\boldsymbol{H}_{ij} + \boldsymbol{D}_{ij}^{\frac{1}{2}}\tilde{\boldsymbol{H}}_{ij}, \tag{8}$$

where the entries of $\tilde{\boldsymbol{H}}_{ij}$ are i.i.d $\mathcal{CN}(0,1)$, and where

$$\begin{aligned}\boldsymbol{D}_{ij} = \tfrac{1}{2T}\big[&(1+\beta)\boldsymbol{I} + \alpha_{ij}\tfrac{2\kappa}{N_\mathsf{t}}\boldsymbol{H}_{ij}\boldsymbol{H}_{ij}^\mathsf{H} \\ &+ \alpha_{ij}\tfrac{2\beta}{N_\mathsf{t}}(1+\kappa)\operatorname{diag}(\boldsymbol{H}_{ij}\boldsymbol{H}_{ij}^\mathsf{H})\big]\end{aligned} \tag{9}$$

characterizes the spatial covariance of the estimation error.

B. Partial Self-Interference Cancellation

Recall that the data communication period is partitioned into two sub-periods, $\mathcal{T}_{\mathsf{data}}[1]$ and $\mathcal{T}_{\mathsf{data}}[2]$, and that—within each—the transmitted signals are wide-sense stationary. The (instantaneous, distorted) signal at receiver $i=1$ and any time $t \in \mathcal{T}_{\mathsf{data}}[l]$ then takes the form

$$\begin{aligned}\boldsymbol{y}_1[l] &= \sqrt{\rho}\boldsymbol{H}_{11}(\boldsymbol{x}_1[l] + \boldsymbol{c}_1[l]) + \boldsymbol{n}_1[l] + \boldsymbol{e}_1[l] \\ &\quad + \sqrt{\eta}\boldsymbol{H}_{12}(\boldsymbol{x}_2[l] + \boldsymbol{c}_2[l]) &(10)\\ &= (\sqrt{\rho}\hat{\boldsymbol{H}}_{11} - \boldsymbol{D}_{11}^{\frac{1}{2}}\tilde{\boldsymbol{H}}_{11})(\boldsymbol{x}_1[l] + \boldsymbol{c}_1[l]) + \boldsymbol{n}_1[l] + \boldsymbol{e}_1[l] \\ &\quad + (\sqrt{\eta}\hat{\boldsymbol{H}}_{12} - \boldsymbol{D}_{12}^{\frac{1}{2}}\tilde{\boldsymbol{H}}_{12})(\boldsymbol{x}_2[l] + \boldsymbol{c}_2[l]). &(11)\end{aligned}$$

Defining the aggregate interference term

$$\begin{aligned}\boldsymbol{v}_1[l] \triangleq\ &\sqrt{\rho}\hat{\boldsymbol{H}}_{11}\boldsymbol{c}_1[l] - \boldsymbol{D}_{11}^{\frac{1}{2}}\tilde{\boldsymbol{H}}_{11}(\boldsymbol{x}_1[l] + \boldsymbol{c}_1[l]) + \boldsymbol{n}_1[l] &(12)\\ &+ \sqrt{\eta}\hat{\boldsymbol{H}}_{12}\boldsymbol{c}_2[l] - \boldsymbol{D}_{12}^{\frac{1}{2}}\tilde{\boldsymbol{H}}_{12}(\boldsymbol{x}_2[l] + \boldsymbol{c}_2[l]) + \boldsymbol{e}_1[l],\end{aligned}$$

we can write $\boldsymbol{y}_1[l] = \sqrt{\rho}\hat{\boldsymbol{H}}_{11}\boldsymbol{x}_1[l] + \sqrt{\eta}\hat{\boldsymbol{H}}_{12}\boldsymbol{x}_2[l] + \boldsymbol{v}_1[l]$, where the self-interference term $\sqrt{\eta}\hat{\boldsymbol{H}}_{12}\boldsymbol{x}_2[l]$ is known and thus can be canceled. The interference-canceled signal $\boldsymbol{z}_1[l] \triangleq \boldsymbol{y}_1[l] - \sqrt{\eta}\hat{\boldsymbol{H}}_{12}\boldsymbol{x}_2[l]$ can then be written as

$$\boldsymbol{z}_1[l] = \sqrt{\rho}\hat{\boldsymbol{H}}_{11}\boldsymbol{x}_1[l] + \boldsymbol{v}_1[l]. \tag{13}$$

Equation (13) shows that, in effect, the information signal $\boldsymbol{x}_1[l]$ propagates through a known channel $\sqrt{\rho}\hat{\boldsymbol{H}}_{11}$ corrupted by an aggregate (possibly non-Gaussian) noise $\boldsymbol{v}_1[l]$, whose $(\hat{\boldsymbol{H}}_{11}, \hat{\boldsymbol{H}}_{12})$-conditional covariance we denote as $\hat{\boldsymbol{\Sigma}}_1[l] \triangleq \operatorname{Cov}\{\boldsymbol{v}_1[l] \,|\, \hat{\boldsymbol{H}}_{11}, \hat{\boldsymbol{H}}_{12}\}$. It can be shown that

$$\begin{aligned}\hat{\boldsymbol{\Sigma}}_1[l] \approx\ &\boldsymbol{I} + \kappa\rho\hat{\boldsymbol{H}}_{11}\operatorname{diag}(\boldsymbol{Q}_1[l])\hat{\boldsymbol{H}}_{11}^\mathsf{H} + \hat{\boldsymbol{D}}_{11}\operatorname{tr}(\boldsymbol{Q}_1[l]) \\ &+ \kappa\eta\hat{\boldsymbol{H}}_{12}\operatorname{diag}(\boldsymbol{Q}_2[l])\hat{\boldsymbol{H}}_{12}^\mathsf{H} + \hat{\boldsymbol{D}}_{12}\operatorname{tr}(\boldsymbol{Q}_2[l]) \\ &+ \beta\rho\operatorname{diag}(\hat{\boldsymbol{H}}_{11}\boldsymbol{Q}_1[l]\hat{\boldsymbol{H}}_{11}^\mathsf{H}) \\ &+ \beta\eta\operatorname{diag}(\hat{\boldsymbol{H}}_{12}\boldsymbol{Q}_2[l]\hat{\boldsymbol{H}}_{12}^\mathsf{H}),\end{aligned} \tag{14}$$

where $\hat{\boldsymbol{D}}_{ij} \triangleq \mathrm{E}\{\boldsymbol{D}_{ij} \,|\, \hat{\boldsymbol{H}}_{ij}\}$ obeys

$$\hat{\boldsymbol{D}}_{ij} \approx \tfrac{1}{2T}\big[\boldsymbol{I} + \alpha_{ij}\tfrac{2\kappa}{N_\mathsf{t}}\hat{\boldsymbol{H}}_{ij}\hat{\boldsymbol{H}}_{ij}^\mathsf{H} + \alpha_{ij}\tfrac{2\beta}{N_\mathsf{t}}\operatorname{diag}(\hat{\boldsymbol{H}}_{ij}\hat{\boldsymbol{H}}_{ij}^\mathsf{H})\big] \tag{15}$$

and where the approximations in (14)-(15) follow from $\kappa \ll 1$ and $\beta \ll 1$. A similar analysis applies to $\hat{\boldsymbol{\Sigma}}_2[l]$. We note, for later use, that the channel estimation error terms $\hat{\boldsymbol{D}}_{ij}$ can be made arbitrarily small through appropriate choice of T.

C. Bounds on Achievable Sum-Rate

Equation (13) succinctly characterizes the effective communication channel, under limited transmitter/receiver-DR and pilot-aided LS MIMO-channel estimation, for transmitter/receiver pair $i=1$ during data communication period $l \in \{1,2\}$; an equivalent model can be stated for the pair $i=2$. Due to the channel estimation error components in (12), the aggregate noise $\boldsymbol{v}_1[l]$ is generally non-Gaussian, which complicates the analysis of the channel (13). It is known, however, that among all distributions on $\boldsymbol{v}_1[l]$ with fixed covariance $\hat{\boldsymbol{\Sigma}}_1[l]$, the Gaussian one is worst from a mutual-information perspective [16]. Thus, we can lower-bound the sum mutual information $I(\boldsymbol{\mathcal{Q}})$, written as a function of the transmit covariance matrices $\boldsymbol{\mathcal{Q}} \triangleq \big(\boldsymbol{Q}_1[1], \boldsymbol{Q}_1[2], \boldsymbol{Q}_2[1], \boldsymbol{Q}_2[2]\big)$, as $I(\boldsymbol{\mathcal{Q}}) \geq \underline{I}(\boldsymbol{\mathcal{Q}})$, where [17]

$$\begin{aligned}\underline{I}(\boldsymbol{\mathcal{Q}}) &= \sum_{i=1}^{2}\frac{1}{2}\sum_{l=1}^{2}\log\det\big(\boldsymbol{I} + \rho\hat{\boldsymbol{H}}_{ii}\boldsymbol{Q}_i[l]\hat{\boldsymbol{H}}_{ii}^\mathsf{H}\hat{\boldsymbol{\Sigma}}_i^{-1}[l]\big) &(16)\\ &= \frac{1}{2}\sum_{i=1}^{2}\sum_{l=1}^{2}\log\det\big(\rho\hat{\boldsymbol{H}}_{ii}\boldsymbol{Q}_i[l]\hat{\boldsymbol{H}}_{ii}^\mathsf{H} + \hat{\boldsymbol{\Sigma}}_i[l]\big) - \log\det(\hat{\boldsymbol{\Sigma}}_i[l]).\end{aligned}$$

Furthermore, standard communication theoretic arguments imply that it is possible to achieve a sum-rate equal to $\underline{I}(\boldsymbol{\mathcal{Q}})$ in (16) by using independent Gaussian codebooks at each transmitter and maximum-likelihood detection at each receiver [17]. Taking "log" in (16) to be base-2, the units of sum-rate are bits-per-channel-use (bpcu).

A straightforward upper bound $\overline{I}(\boldsymbol{\mathcal{Q}})$ on achievable sum-rate then follows from the perfect-CSI case (i.e., $\hat{\boldsymbol{D}}_{ij} = \boldsymbol{0}$), where $\boldsymbol{v}_1[l]$ becomes Gaussian under our distortion model. Moreover, the lower bound $\underline{I}(\boldsymbol{\mathcal{Q}})$ converges to the upper bound $\overline{I}(\boldsymbol{\mathcal{Q}})$ as the training length $T \to \infty$.

D. Transmit Covariance Optimization

We would now like to find the transmit covariance matrices $\boldsymbol{\mathcal{Q}} \triangleq \big(\boldsymbol{Q}_1[1], \boldsymbol{Q}_1[2], \boldsymbol{Q}_2[1], \boldsymbol{Q}_2[2]\big)$ that maximize the sum-rate lower bound $\underline{I}(\boldsymbol{\mathcal{Q}})$ in (16) subject to the per-user power constraint (17b). This yields the optimization problem

$$\max_{\boldsymbol{Q}_1[1],\boldsymbol{Q}_1[2],\boldsymbol{Q}_2[1],\boldsymbol{Q}_2[2]} \quad \underline{I}\big(\boldsymbol{Q}_1[1], \boldsymbol{Q}_1[2], \boldsymbol{Q}_2[1], \boldsymbol{Q}_2[2]\big) \tag{17a}$$

$$\text{s.t.} \quad \frac{1}{2}\sum_{l=1}^{2}\operatorname{tr}\big(\boldsymbol{Q}_i[l]\big) \leq 1, \quad i=1,2 \tag{17b}$$

$$\boldsymbol{Q}_i[l] \geq 0, \quad \forall i,l \in \{1,2\}, \tag{17c}$$

where the inequality (17c) constrains each $\boldsymbol{Q}_i[l]$ to be positive semi-definite. We solve[2] this non-convex optimization problem via Gradient Projection (GP), taking inspiration from [18].

[2]In general, (17) is a non-convex optimization problem, and so finding the global maximum can be difficult. Although GP is not guaranteed to find the global maximum, our experience with different initializations suggests that, in our problem, GP is indeed finding the global maximum.

E. Sum-Rate Approximation

The complicated nature of the optimization problem (17) motivates us to approximate its solution, i.e., the transmit-covariance optimized sum-rate $I_* \triangleq \max_{\mathcal{Q}\in\mathbb{Q}} \underline{I}(\mathcal{Q})$, where $\mathbb{Q}$ represents the constraint set implied by (17b)-(17c). Here, we focus on the case of $T \to \infty$, where channel estimation error is driven to zero so that $\underline{I}(\mathcal{Q}) = I(\mathcal{Q}) = \overline{I}(\mathcal{Q})$.

Our approximation is built around the special case that each $\boldsymbol{H}_{ij}$ is diagonal, although not necessarily square, with $N_{\mathsf{min}} \triangleq \min\{N_{\mathsf{t}}, N_{\mathsf{r}}\}$ identical diagonal entries equal to $\sqrt{N_{\mathsf{t}}N_{\mathsf{r}}/N_{\mathsf{min}}}$. (The latter value is chosen so that $\mathrm{E}\{\mathrm{tr}(\boldsymbol{H}_{ij}\boldsymbol{H}_{ij}^{\mathsf{H}})\} = N_{\mathsf{t}}N_{\mathsf{r}}$ as assumed in Section II.) In this case, the mutual information expression (16) becomes (for $j \neq i$)

$$I(\mathcal{Q}) \approx \frac{1}{2}\sum_{i,l} \log\det\Big(\boldsymbol{I} + \rho\tfrac{N_{\mathsf{t}}N_{\mathsf{r}}}{N_{\mathsf{min}}}\boldsymbol{Q}_i[l]\big(\boldsymbol{I} + (\kappa+\beta)\tfrac{N_{\mathsf{t}}N_{\mathsf{r}}}{N_{\mathsf{min}}} \times \big[\rho\,\mathrm{diag}(\boldsymbol{Q}_i[l]) + \eta\,\mathrm{diag}(\boldsymbol{Q}_j[l])\big]\big)^{-1}\Big). \quad (18)$$

When $\eta \ll \rho$, the η-dependent term in (18) can be ignored, after which it is straightforward to show that, under the constraints (17b)-(17c), the optimal covariances are the "full duplex" $\mathcal{Q}_{\mathsf{FD}} \triangleq (\frac{1}{N_{\mathsf{t}}}\boldsymbol{I}, \frac{1}{N_{\mathsf{t}}}\boldsymbol{I}, \frac{1}{N_{\mathsf{t}}}\boldsymbol{I}, \frac{1}{N_{\mathsf{t}}}\boldsymbol{I})$, for which (18) gives

$$I(\mathcal{Q}_{\mathsf{FD}}) \approx 2N_{\mathsf{min}}\log\left(1 + \frac{\rho}{\frac{N_{\mathsf{min}}}{N_{\mathsf{r}}} + (\kappa+\beta)(\rho+\eta)}\right). \quad (19)$$

When $\eta \gg \rho$, the η-dependent term in (18) dominates unless $\boldsymbol{Q}_j[l] = \boldsymbol{0}$. In this case, the optimal covariances are the "half duplex" ones $\mathcal{Q}_{\mathsf{HD}} \triangleq (\frac{2}{N_{\mathsf{t}}}\boldsymbol{I}, \boldsymbol{0}, \boldsymbol{0}, \frac{2}{N_{\mathsf{t}}}\boldsymbol{I})$, for which (18) gives

$$I(\mathcal{Q}_{\mathsf{HD}}) \approx N_{\mathsf{min}}\log\left(1 + \frac{\rho}{\frac{N_{\mathsf{min}}}{2N_{\mathsf{r}}} + (\kappa+\beta)\rho}\right). \quad (20)$$

Finally, for any given pair (η, ρ), we approximate the optimized sum-rate as follows: $I_* \approx \max\{I(\mathcal{Q}_{\mathsf{FD}}), I(\mathcal{Q}_{\mathsf{HD}})\}$. From (19)-(20), it is straightforward to show that the boundary between full- and half-duplex occurs at $\eta = \big(\sqrt{\xi^2 + 2\rho\xi/(\kappa+\beta)} - (\xi - 2\rho)\big)/2$ for $\xi \triangleq \frac{N_{\mathsf{min}}}{N_{\mathsf{r}}(\kappa+\beta)} + 2\rho$.

We now make some additional observations about (19)-(20). First, suppose that $\eta \ll \rho$, in which case $\mathcal{Q}_{\mathsf{FD}}$ is appropriate. From (19), we see that $I(\mathcal{Q}_{\mathsf{FD}})$ will not significantly benefit from further increase in SNR ρ when $(\kappa+\beta)\rho > (\kappa+\beta)\eta + \frac{N_{\mathsf{min}}}{N_{\mathsf{r}}}$, i.e., when $\rho > \eta + \frac{N_{\mathsf{min}}}{N_{\mathsf{r}}(\kappa+\beta)}$. Since $\eta \ll \rho$, this ρ-saturation occurs when $\rho > \frac{N_{\mathsf{min}}}{N_{\mathsf{r}}(\kappa+\beta)}$. Next, suppose that $\eta \gg \rho$, in which case $\mathcal{Q}_{\mathsf{HD}}$ is appropriate. Here, (20) shows that $I(\mathcal{Q}_{\mathsf{HD}})$ will not significantly benefit from SNRs above $\rho = \frac{N_{\mathsf{min}}}{2N_{\mathsf{r}}(\kappa+\beta)}$. Thus, in both the $\eta \ll \rho$ and $\eta \gg \rho$ cases, we can interpret $\rho \approx \frac{N_{\mathsf{min}}}{N_{\mathsf{r}}(\kappa+\beta)}$ as the transition between SNR-limited and distortion-limited regimes. (See Fig. 2.)

Figure 2 shows a contour plot of the proposed optimized-sum-rate approximation as a function of INR η and SNR ρ. We shall see in Section IV that our approximation of the covariance-optimized sum-rate is surprisingly close, on average, to that found by solving (17) using gradient projection.

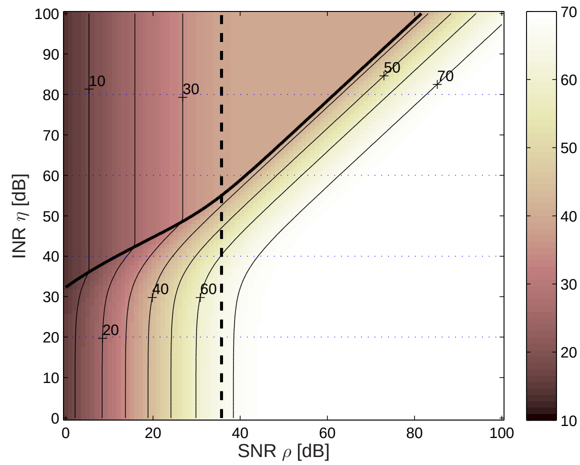

Fig. 2. Contour plot of the optimized-sum-rate approximation I_* versus SNR ρ and INR η, for $N_{\mathsf{t}} = 3$, $N_{\mathsf{r}} = 4$, and $\beta = \kappa = -40$dB. The dark curve shows the boundary $\eta = (\sqrt{\xi^2 + 2\rho\xi/(\kappa+\beta)} - (\xi - 2\rho))/2$ between full- and half-duplex regimes, and the vertical dashed line shows the boundary $\rho = \frac{N_{\mathsf{min}}}{N_{\mathsf{r}}(\kappa+\beta)}$ between SNR-limited and distortion-limited regimes.

IV. Numerical Results and Conclusions

We now study the average behavior of the GP-optimized sum-rate $\max_{\mathcal{Q}} \underline{I}(\mathcal{Q})$ as a function of SNR ρ, INR η, dynamic range κ^{-1} and β^{-1}, and number of antennas N_{t} and N_{r}. We also investigate the role of interference cancellation, the role of two distinct data sub-periods, and the relation to optimized half-duplex (OHD) signaling. In doing so, we find close agreement with the optimized-sum-rate approximation proposed in Section III-E and illustrated in Fig. 2. Throughout, we used $T = 50$ training duration. All results were averaged over 1000 realizations, unless specified otherwise.

Below, we denote the full scheme proposed in Section III by "TCO-2-IC," which indicates the use of interference cancellation (IC) and transmit covariance optimization (TCO) performed individually over the 2 data sub-periods (i.e., $\mathcal{T}_{\mathsf{data}}[1]$ and $\mathcal{T}_{\mathsf{data}}[2]$). To test the impact of IC and of two data sub-periods, we also implemented the proposed scheme but without IC, which we refer to as "TCO-2," as well as the proposed scheme with only one data sub-period (i.e., $\boldsymbol{Q}_i[1] = \boldsymbol{Q}_i[2]\ \forall i$), which we refer to as "TCO-1-IC." To optimize half-duplex, we used GP to maximize $\underline{I}(\mathcal{Q})$ under the power constraint (17b) and the additional half-duplex constraint $\boldsymbol{Q}_1[2] = \boldsymbol{0} = \boldsymbol{Q}_2[1]$.

In Fig. 3, we examine sum-rate performance versus INR η for the TCO-2-IC, TCO-1-IC, TCO-2, and OHD schemes, using several different dynamic range parameters $\beta = \kappa$. For OHD, we see that sum-rate is invariant to INR η, as expected. For the proposed TCO-2-IC, we observe "full duplex" performance for low-to-mid values of η and a transition to OHD performance at high values of η, just as predicted by the approximation in Section III-E. In fact, the sum-rates in Fig. 3 are nearly identical to the approximate values in Fig. 2. To see the importance of two distinct data-communication periods, we study the TCO-1-IC trace, where we observe TCO-2-IC-like performance at low-to-mid values of η, but performance that drops below OHD at high η. Essentially, TCO-1-IC forces full-duplex signaling at high INR η, where half-duplex signaling is optimal, while TCO-2-IC facilitates the possibility of half-duplex signaling through the use of two distinct data-communication sub-periods, similar to the interference-channel scheme [20]. Finally, from the TCO-2 trace, we conclude that partial interference cancellation is essential for all but extreme values of INR η.

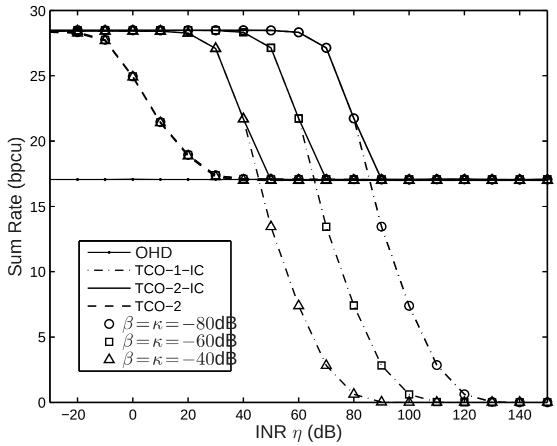

Fig. 3. Achievable sum-rate lower bound $\underline{I}(\boldsymbol{\mathcal{Q}})$ for TCO-2-IC, TCO-2, TCO-1-IC, and OHD versus INR η. Here, $N_t = 3$, $N_r = 4$, $\rho = 15$dB, and $T = 50$. OHD is plotted for $\beta = \kappa = -60$dB, but was observed to give nearly identical results for all three values of $\beta = \kappa$.

In Fig. 4, we examine sum-rate of the proposed TCO-IC-2 and OHD versus SNR ρ, using the dynamic range parameters $\beta = \kappa = -40$dB and various fixed values of INR η. All the behaviors in Fig. 4 are almost exactly as predicted by the sum-rate approximation described in Section III-E and illustrated in Fig. 2. In particular, we see OHD's sum-rate increase with SNR ρ up to the distortion-limited regime, i.e., $\rho \gtrsim \frac{N_{min}}{N_r(\kappa+\beta)} \approx 36$dB. For TCO-IC-2, we see sum-rate increase with ρ when $\rho \in [0, 36]$ (i.e., the SNR-limited regime), saturate when $\rho \in [36, \eta]$ (i.e., distortion-limited high-INR regime), increase again around $\rho \approx \eta$ (i.e., the transition to the low-INR regime), and then saturate when $\rho \gg \eta$ (i.e., the distortion-limited low-INR regime). In fact, the sum-rates in Fig. 4 are nearly identical to the approximations in Fig. 2.

Finally, in Fig. 5, we explore the sum-rate of TCO-2-IC and OHD versus the number of antennas, N_t and N_r. There,we see that sum-rate increases with both N_r and N_t, as expected. More interesting is the sum-rate behavior when the total number of antennas is fixed, e.g., at $N_t + N_r = 7$, as illustrated by the triangles in Fig. 5. The fact that the configuration $(N_t, N_r) = (3, 4)$ outperforms $(N_t, N_r) = (4, 3)$ is predicted by the approximations (19)-(20): given fixed N_{min} (here, $N_{min} = 3$), one should strive to maximize N_r.

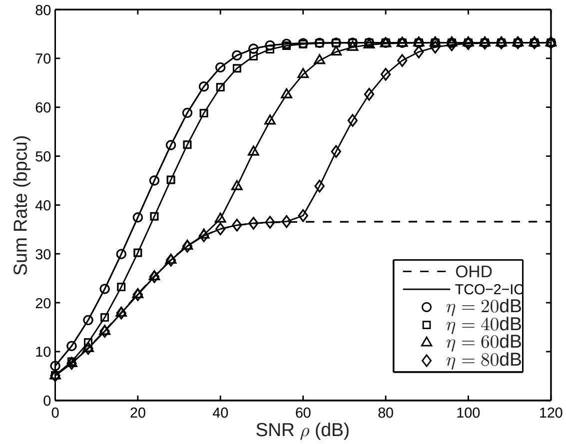

Fig. 4. Achievable sum-rate lower bound $\underline{I}(\boldsymbol{\mathcal{Q}})$ for TCO-2-IC and OHD versus SNR ρ. Here, $N_t = 3$, $N_r = 4$, $\beta = \kappa = -40$dB, and $T = 50$.

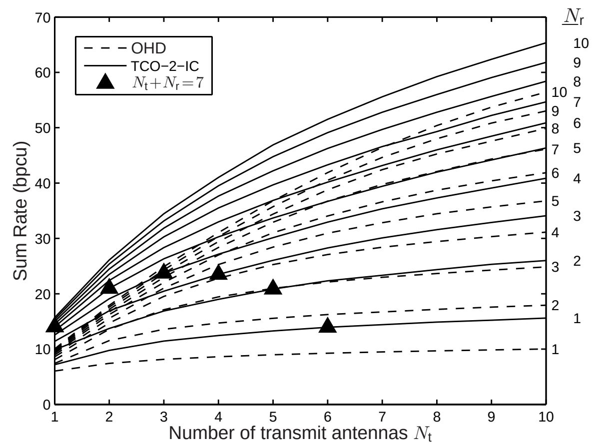

Fig. 5. Achievable sum-rate lower bound $\underline{I}(\boldsymbol{\mathcal{Q}})$ for TCO-2-IC and OHD versus number of antennas N_t with various N_r. Here, $\rho = 15$dB, $\eta = 60$dB, $\beta = \kappa = -60$dB, and $T = 50$.

References

[1] D. W. Bliss, P. A. Parker, and A. R. Margetts, "Simultaneous transmission and reception for improved wireless network performance," in *Proc. IEEE Workshop Statist. Signal Process.*, (Madison, WI), pp. 478–482, Aug. 2007.

[2] X. Shang, B. Chen, G. Kramer, and H. V. Poor, "Capacity regions and sum-rate capacities of vector Gaussian interference channels," *IEEE Trans. Inform. Theory*, vol. 56, pp. 5030–5044, Oct. 2010.

[3] S. A. Jafar and M. J. Fakhereddin, "Degrees of freedom for the MIMO interference channel," *IEEE Trans. Inform. Theory*, vol. 53, pp. 2637–2642, July 2007.

[4] H. Suzuki, T. V. A. Tran, I. B. Collings, G. Daniels, and M. Hedley, "Transmitter noise effect on the performance of a MIMO-OFDM hardware implementation achieving improved coverage," *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 867–876, Aug. 2008.

[5] W. Namgoong, "Modeling and analysis of nonlinearities and mismatches in AC-coupled direct-conversion receiver," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 163–173, Jan. 2005.

[6] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links," *IEEE Trans. Inform. Theory*, vol. 49, pp. 951–963, Apr. 2003.

[7] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. New York: Cambridge University Press, 2005.

[8] S. Ye and R. S. Blum, "Optimized signaling for MIMO interference systems with feedback," *IEEE Trans. Signal Process.*, vol. 51, pp. 2839–2848, Nov. 2003.

[9] Y. Rong and Y. Hua, "Optimal power schedule for distributed MIMO links," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 2896–2900, Aug. 2008.