Statistical Beamforming on the Grassmann Manifold for the Two-User Broadcast Channel

Vasanthan Raghavan, *Senior Member, IEEE*, Stephen V. Hanly, *Member, IEEE*, and Venugopal V. Veeravalli, *Fellow, IEEE*

***Abstract*—A Rayleigh fading spatially correlated broadcast setting with $M = 2$ antennas at the transmitter and two users (each with a single antenna) is considered. It is assumed that the users have perfect channel information about their links, whereas the transmitter has only statistical information of each user's link (covariance matrix of the vector channel). A low-complexity linear beamforming strategy that allocates equal power and one spatial eigenmode to each user is employed at the transmitter. Beamforming vectors on the Grassmann manifold that depend only on statistical information are to be designed at the transmitter to maximize the ergodic sum-rate delivered to the two users. Toward this goal, the beamforming vectors are first fixed and a closed-form expression is obtained for the ergodic sum-rate in terms of the covariance matrices of the links. This expression is nonconvex in the beamforming vectors ensuring that the classical Lagrange multiplier technique is not applicable. Despite this difficulty, the optimal solution to this problem is shown to be the same as the solution to the maximization of an appropriately defined average signal-to-interference and noise ratio metric for each user. This solution is the dominant generalized eigenvector of a pair of positive-definite matrices where the first matrix is the covariance matrix of the forward link and the second is an appropriately designed "effective" interference covariance matrix. In this sense, our work is a generalization of optimal signalling along the dominant eigenmode of the transmit covariance matrix in the single-user case. Finally, the ergodic sum-rate for the general broadcast setting with M antennas at the transmitter and M-users (each with a single antenna) is obtained in terms of the covariance matrices of the links and the beamforming vectors.**

***Index Terms*—Adaptive signalling, broadcast channel, information rates, multiple-input single-output (MISO) systems, multiuser multiple-input multiple-output (MIMO), nonconvex optimization, precoding, spatial correlation.**

Manuscript received April 12, 2011; revised September 17, 2012; accepted April 28, 2013. Date of publication June 11, 2013; date of current version September 11, 2013. This work was supported in part by the ARC under Grant DP-0984862, in part by the NUS under Grant WBS #R-263-000-572-133, in part by the CSIRO Macquarie University Chair in Wireless Communications (established with funding provided by the Science and Industry Endowment Fund), and in part by the National Science Foundation under Grant CNS-0831670 at the University of Illinois. This paper was presented in part at the 2010 IEEE International Symposium on Information Theory.

V. Raghavan is with the Department of Mathematics, University of Southern California, Los Angeles, CA 90089 USA (e-mail: vasanthan_raghavan@ieee.org).

S. V. Hanly is with the Macquarie University, Sydney, N.S.W. 2109, Australia (e-mail: stephen.hanly@mq.edu.au).

V. V. Veeravalli is with the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: vvv@illinois.edu).

Communicated by A. M. Tulino, Associate Editor for Communications.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2013.2267721

I. Introduction

THE last 15 years of research in wireless communications has seen the emergence of multiantenna signalling as a viable option to realize high data-rates at practically acceptable reliability levels. While initial work on multiantenna design was primarily motivated by the single-user paradigm [1]–[6], more recent attention has been on the theory and practice of multiuser multiantenna communications [7]–[10]. The focus of this paper is on a broadcast setting that typically models a cellular downlink. We study the multiple-input single-output (MISO) broadcast problem where a central transmitter or base station with M antennas communicates with M users in the cell, each having a single antenna. Under the assumption of perfect channel state information (CSI) at both the transmitter and the user ends, significant progress has been made over the last few years on understanding optimal signalling that achieves the sum-capacity [11]–[16] as well as the capacity region [17] of the multiantenna Gaussian broadcast channel. The capacity-achieving *dirty-paper coding* (DPC) scheme [18] prenulls interference from simultaneous transmissions by other users to a specific user and hence results in a multiplexing gain of M.

Nevertheless, the high implementation complexity associated with DPC makes it less attractive in standardization efforts for practical systems. The consequent search for low-complexity signalling alternatives that are within a fixed power-offset[1] of the DPC scheme has resulted in an array of candidate linear (as well as nonlinear) precoding techniques [19]–[27]. In particular, a linear beamforming scheme that is developed as a generalization of the single-user beamforming scheme has attracted significant attention in the literature. Specifically, a scheme where the transmitter allocates one eigenmode to each user and shares the power budget equally among all the users is the focus of this study.

If perfect CSI is available at both ends, instantaneous nulls can be created in the interference subspace of each user (or interference can be zeroforced), and thus, this scheme remains order-optimal with respect to the DPC scheme [23]. However, the practical utility of the linear beamforming scheme is dependent on how gracefully its performance degrades with the quality of CSI at the transmitter. This is because while reasonably accurate CSI can be obtained at the user end via pilot-based

[1]Two schemes are within a fixed power-offset if the difference in power level necessary to achieve a fixed rate with the two schemes stays bounded independent of the rate.

TABLE I
STRUCTURE OF OPTIMAL BEAMFORMING VECTORS

Objective Function:	$\underset{\mathbf{w}_1,\mathbf{w}_2}{\operatorname{argmax}} E[R_i],\ i=1,2$	$\underset{\mathbf{w}_1,\mathbf{w}_2}{\operatorname{argmax}} E[R_1]+E[R_2]$
$\rho \to 0$	$\mathbf{w}_{i,\,\mathsf{opt}} = \mathbf{u}_1(\mathbf{\Sigma}_i)$ $\mathbf{w}_{j,\,\mathsf{opt}} =$ any vector on $\mathcal{G}(2,1)$, $j \neq i$ (See Prop. 4)	$\mathbf{w}_{1,\,\mathsf{opt}} = \mathbf{u}_1(\mathbf{\Sigma}_1)$ $\mathbf{w}_{2,\,\mathsf{opt}} = \mathbf{u}_1(\mathbf{\Sigma}_2)$ (See Prop. 3)
ρ intermediate	–	$\mathbf{w}_{1,\,\mathsf{opt}} = \mathbf{u}_1\left((\alpha^\star(\rho)\mathbf{\Sigma}_2+\mathbf{I})^{-1}\mathbf{\Sigma}_1\right)$ $\mathbf{w}_{2,\,\mathsf{opt}} = \mathbf{u}_1\left((\beta^\star(\rho)\mathbf{\Sigma}_1+\mathbf{I})^{-1}\mathbf{\Sigma}_2\right)$ $\{\alpha^\star(\rho),\beta^\star(\rho)\} \geq 0$, chosen appropriately (See Conjecture 1)
$\rho \to \infty$	$\mathbf{w}_{i,\,\mathsf{opt}} = \mathbf{u}_1(\mathbf{\Sigma}_i)$ $\mathbf{w}_{j,\,\mathsf{opt}} = \mathbf{u}_2(\mathbf{\Sigma}_i)$, $j \neq i$ (See Prop. 4)	$\mathbf{w}_{1,\,\mathsf{opt}} = \mathbf{u}_1\left(\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1\right)$ $\mathbf{w}_{2,\,\mathsf{opt}} = \mathbf{u}_1\left(\mathbf{\Sigma}_1^{-1}\mathbf{\Sigma}_2\right)$ (See Theorems 2 and 3)

training schemes, CSI at the transmitter requires either channel reciprocity or reverse link feedback, both of which place an overwhelming burden on the operating cost [10]. In the extreme (and pessimistic) setting of no CSI at the transmitter, the multiplexing gain reduces to 1 (that is, it is lost completely relative to the perfect CSI case).

In practice, the channel evolves fairly slowly on a statistical scale and it is possible to learn the spatial statistics[2] of the individual links at the transmitter with minimal cost. With only statistical information at the transmitter, the interference cannot be nulled out completely and a low-complexity decoder architecture that treats interference as noise is often preferred. Initial works assume an identity covariance matrix for all the users corresponding to an *independent and identically distributed* (i.i.d.) fading process in the spatial domain [7]–[10]. However, this model cannot be justified in practical systems that are often deployed in environments where the scattering is localized in certain spatial directions or where antennas are not spaced wide apart due to infrastructural constraints [28].

While signalling design for the single-user setting under a very general spatial correlation model is now well understood [1]–[6], the broadcast case where the channel statistics vary across users and different users experience different covariance matrices has not received much attention. In particular, Al-Naffouri *et al.* [29] study the problem where all the users share a common non-i.i.d. transmit covariance matrix and captures the impact of this common covariance matrix on the achievable rates. In [30], the authors show that second-order spatial statistics can be exploited to schedule users that enjoy better channel quality and hence improve the overall performance of an opportunistic beamforming scheme. In the same spirit, it is shown in [31] and [32] that second-order moments of the channel in combination with instantaneous norm (or weighted-norm) feedback is sufficient to extract almost all of the multiuser diversity gain in a broadcast setting. Spatial correlation is exploited to reduce the feedback overhead of a limited feedback codebook design in [33].

Summary of Main Contributions: With this background, the main focus of this paper is to fill some of the gaps in understanding the information-theoretic limits of broadcast channels with low-complexity signalling schemes (such as linear beamforming) under practical assumptions on CSI and decoder architecture. We study the simplest nontrivial version of this problem corresponding to the two-user ($M = 2$) case. We design optimal beamforming vectors on the Grassmann manifold[3] $\mathcal{G}(2,1)$ to maximize the ergodic sum-rate achievable with the linear beamforming scheme.

The first step in this goal is the computation of the ergodic sum-rate in closed form. For this, we develop insight into the structure of the density function of the weighted-norm of beamforming vectors isotropically distributed on $\mathcal{G}(2,1)$. Exploiting this knowledge, we derive an explicit expression for the ergodic sum-rate in terms of the covariance matrices of the users and the beamforming vectors. This expression can be rewritten in terms of a certain generalized "distance" measure between the beamforming vectors. As a result of this complicated nonlinear dependence, the sum-rate is nonconvex in the beamforming vectors thus precluding the use of the classical Lagrangian approach to convex optimization. Instead, a first-principles-based technique is developed where the beamforming vectors are decomposed along an appropriately chosen (in general, nonorthogonal) basis. Exploiting this decomposition structure, we obtain an upper bound for the ergodic sum-rate, which we show is tight for a specific choice of beamforming vectors (see Theorems 2 and 3). This optimal choice is the dominant generalized eigenvector[4] of a pair of covariance matrices, with one of them being the covariance matrix of the forward link and the other an appropriately designed "effective" interference covariance matrix. The generalized eigenvector structure is the solution to maximizing an appropriately defined average signal-to-interference and noise ratio (SINR) metric for each user and thus generalizes our intuition from the single-user case [1]–[6]. Table I summarizes the structure of the optimal beamforming vectors under

[2]With a Rayleigh (or a Ricean) fading model for the MISO channel, the complete statistical information of the link is captured by the covariance matrix (or the mean vector and the covariance matrix) of the vector channel.

[3]Informally, $\mathcal{G}(M,1)$ denotes the space of all M-dimensional unit-norm beamforming vectors modulo the phase of the first element of the vector. A more formal definition is provided in Section II (Definition 1).

[4]A generalized eigenvector generalizes the notion of an eigenvector to a pair of matrices. A more technical definition is provided in Section IV-B (Definition 2). The dominant eigenvector is the eigenvector corresponding to the dominant eigenvalue. Under the assumption that the eigenvector is unit-norm, it is unique on $\mathcal{G}(M,1)$.

different signal-to-noise ratio (SNR) assumptions, where $\mathbf{u}_1(\bullet)$ and $\mathbf{u}_2(\bullet)$ denote the dominant and nondominant eigenvectors of the matrices under consideration, respectively.

While a generalized eigenvector solution has been obtained in the perfect CSI case for the multiple-input multiple-output (MIMO) broadcast problem [26], [34] and the MIMO interference channel problem in the low-interference regime [35], its appearance in the statistical setting is a first. A closely related work of ours [36] reports the optimality of the generalized eigenvector solution for the statistical beamformer design in the MISO interference channel setting with two antennas. We also extend our intuition to the weighted ergodic sum-rate maximization problem [37] and conjecture on the structure of the optimal beamforming vectors. Numerical results justify our conjecture and the intuition behind it. Finally, closed-form expression for the ergodic sum-rate in terms of the covariance matrices of the links and the beamforming vectors are obtained in the general M-user case.

Organization: This paper is organized as follows. With Section II explaining the background of the problem, ergodic rate expressions in terms of the covariance matrices of the links and beamforming vectors are obtained in Section III for the $M = 2$ case. The nonconvex optimization problem of ergodic sum-rate maximization is the main focus of Section IV with the low- and the high-SNR extremes providing insight for the development in the intermediate-SNR regime. The focus then shifts to weighted ergodic sum-rate maximization in Section V. In addition, sum-rate expressions are generalized to the general M-user case. Numerical results are provided in Section VI and concluding remarks are provided in Section VII. Most of the proofs/details are relegated to the Appendixes.

Notation: We use uppercase and lowercase bold symbols for matrices and vectors, respectively. The notations $\mathbf{\Lambda}$ and $\mathbf{U}$ are usually reserved for eigenvalue and eigenvector matrices, whereas $\mathbf{I}$ is reserved for the identity matrix (of appropriate dimensionality). The ith diagonal element of $\mathbf{\Lambda}$ is denoted by $\mathbf{\Lambda}_i$, while the ith element of a vector $\mathbf{x}$ is denoted by $\mathbf{x}(i)$. At times, we also use $\lambda_1, \lambda_2, \ldots$ to denote the eigenvalues of a Hermitian matrix, and these eigenvalues are often arranged in decreasing order as $\lambda_1 \geq \lambda_2 \geq \cdots$. The Hermitian transpose and inverse operations of a matrix are denoted by $(\cdot)^H$ and $(\cdot)^{-1}$, while the trace operator is denoted by $\mathsf{Tr}(\cdot)$. The two-norm of a vector is denoted by the symbol $\|\cdot\|$. The operator $E[\cdot]$ stands for expectation, while the density function of a random variable is denoted by the symbol $p(\cdot)$. The symbols $\mathbb{C}$ and $\mathbb{R}^+$ stand for complex and positive real fields, respectively. $X \sim \mathcal{CN}(\mu, \sigma^2)$ indicates that X is a complex Gaussian random variable with mean μ and variance σ^2.

II. System Setup

We consider a broadcast setting that models a MISO cellular downlink with M antennas at the transmitter and M users, each with a single antenna. We denote the $M \times 1$ vector channel between the transmitter and user i as $\mathbf{h}_i, i = 1, \ldots, M$.

While different multiuser communication strategies can be considered [19]–[27], as motivated in Section I, the focus here is on a linear beamforming scheme where the information-bearing signal s_i meant for user i is beamformed from the transmitter with the $M \times 1$ unit-norm vector $\mathbf{w}_i$. We assume that s_i is unit energy and the transmitter divides its power budget of ρ equally across all the users. The received symbol y_i at user i is then written as

$$y_i = \sqrt{\frac{\rho}{M}} \cdot \mathbf{h}_i^H \left(\sum_{j=1}^{M} \mathbf{w}_j s_j \right) + n_i, \quad i = 1, \ldots, M \tag{1}$$

where ρ is the transmit power and n_i denotes the $\mathcal{CN}(0,1)$ complex Gaussian noise added at the receiver.

Initial works on the broadcast problem assume that $\mathbf{h}_i$ is ergodic and it evolves over time, frequency, and space in an i.i.d. fashion. While the above assumption can be justified in the time and frequency axes with a frame-based and multicarrier signalling approach (common in current-generation systems), it cannot be justified along the spatial axis. This is because the channel variation in the spatial domain cannot be i.i.d. unless the antennas at the transmitter end are spaced wide apart and the scattering environment connecting the transmitter with the users is sufficiently rich [28]. With this motivation, the main emphasis of this study is on understanding the impact of the users' spatial statistics on the performance of a linear beamforming scheme.

We assume a Rayleigh fading[5] (zero mean complex Gaussian) model for the channel, which implies that the complete spatial statistics are described by the second-order moments of $\{\mathbf{h}_i\}$. For the MISO model, the channel $\mathbf{h}_i$ of user i can be written as

$$\mathbf{h}_i = \mathbf{\Sigma}_i^{1/2} \mathbf{h}_{\mathsf{iid},i} \tag{2}$$

where $\mathbf{h}_{\mathsf{iid},i}$ is an $M \times 1$ vector with i.i.d. $\mathcal{CN}(0,1)$ entries and $\mathbf{\Sigma}_i \triangleq E\left[\mathbf{h}_i \mathbf{h}_i^H\right]$ is the transmit covariance matrix corresponding to user i. Note that (2) is the most general statistical model for $\mathbf{h}_i$ under the MISO assumption. With $\mathbf{\Sigma}_i = \mathbf{I}$ for all users, (2) reduces to the i.i.d. downlink model well studied in the literature [7]–[10].

Under the assumption of Gaussian inputs $\{s_i\}$, the instantaneous information-theoretic[6] rate, R_i, achievable by user i with the linear beamforming scheme using a mismatched[7] decoder [38] is given by

$$R_i = \log\left(1 + \frac{\frac{\rho}{M} \cdot |\mathbf{h}_i^H \mathbf{w}_i|^2}{1 + \frac{\rho}{M} \cdot \sum_{j \neq i} |\mathbf{h}_i^H \mathbf{w}_j|^2}\right) \tag{3}$$

$$= \underbrace{\log\left(1 + \frac{\rho}{M} \sum_{j=1}^{M} |\mathbf{h}_i^H \mathbf{w}_j|^2\right)}_{I_{i,1}} - \underbrace{\log\left(1 + \frac{\rho}{M} \sum_{j \neq i} |\mathbf{h}_i^H \mathbf{w}_j|^2\right)}_{I_{i,2}}. \tag{4}$$

With the spatial correlation model assumed in (2), we can write $I_{i,1}$ as

$$I_{i,1} = \log\left(1 + \frac{\rho}{M} \cdot \mathbf{h}_{\mathsf{iid},i}^H \mathbf{\Sigma}_i^{1/2} \left(\sum_{j=1}^{M} \mathbf{w}_j \mathbf{w}_j^H \right) \mathbf{\Sigma}_i^{1/2} \mathbf{h}_{\mathsf{iid},i}\right) \tag{5}$$

$$= \log\left(1 + \frac{\rho}{M} \cdot \mathbf{h}_{\mathsf{iid},i}^H \mathbf{V}_i \mathbf{\Lambda}_i \mathbf{V}_i^H \mathbf{h}_{\mathsf{iid},i}\right), \tag{6}$$

[5]While more general fading models such as Ricean or Nakagami-m models can be considered, this paper focuses on the Rayleigh model alone.

[6]All logarithms are to base e and all rate quantities are assumed to be in nats/s/Hz in this study, unless specified otherwise.

[7]Here, the decoding rule is different from the optimal decoding rule due to the presence of multiuser interference.

where we have used the following eigendecomposition in (6):

$$\mathbf{V}_i\mathbf{\Lambda}_i\mathbf{V}_i^H = \mathbf{\Sigma}_i^{1/2}\left(\sum_{j=1}^{M}\mathbf{w}_j\mathbf{w}_j^H\right)\mathbf{\Sigma}_i^{1/2} \tag{7}$$

$$\mathbf{\Lambda}_i = \mathsf{diag}\big([\mathbf{\Lambda}_{i,1},\ldots,\mathbf{\Lambda}_{i,M}]\big), \mathbf{\Lambda}_{i,1}\geq\cdots\geq\mathbf{\Lambda}_{i,M}\geq 0. \tag{8}$$

Similarly, we can write $I_{i,2}$ as

$$I_{i,2} = \log\left(1+\frac{\rho}{M}\cdot\mathbf{h}_{\mathsf{iid},i}^H\widetilde{\mathbf{V}}_i\widetilde{\mathbf{\Lambda}}_i\widetilde{\mathbf{V}}_i^H\mathbf{h}_{\mathsf{iid},i}\right) \tag{9}$$

$$\widetilde{\mathbf{V}}_i\widetilde{\mathbf{\Lambda}}_i\widetilde{\mathbf{V}}_i^H = \mathbf{\Sigma}_i^{1/2}\left(\sum_{j\neq i}\mathbf{w}_j\mathbf{w}_j^H\right)\mathbf{\Sigma}_i^{1/2} \tag{10}$$

$$\widetilde{\mathbf{\Lambda}}_i = \mathsf{diag}\big([\widetilde{\mathbf{\Lambda}}_{i,1},\ldots,\widetilde{\mathbf{\Lambda}}_{i,M}]\big), \widetilde{\mathbf{\Lambda}}_{i,1}\geq\cdots\geq\widetilde{\mathbf{\Lambda}}_{i,M}\geq 0. \tag{11}$$

The goal of this study is to maximize the throughput conveyed from the transmitter to the users by the choice of beamforming vectors. Specifically, the metric of interest is the ergodic sum-rate, $\mathcal{R}_{\mathsf{sum}}$, achievable with the linear beamforming scheme:

$$\mathcal{R}_{\mathsf{sum}} \triangleq \sum_{i=1}^{M} E\left[R_i\right]. \tag{12}$$

For this, note that the achievable rate in (3) is invariant to transformations of the form $\mathbf{w}_i \mapsto e^{j\theta}\mathbf{w}_i$ for any θ. Coupled with the unit-norm assumption for $\mathbf{w}_i$, the space over which optimization is performed is precisely defined as follows.

Definition 1 (Stiefel and Grassmann Manifold [39]): The unidimensional complex Stiefel manifold $\mathsf{St}(M,1)$ refers to the unit-radius complex sphere in M-dimensions and is defined as

$$\mathsf{St}(M,1) = \left\{\mathbf{x}\in\mathbb{C}^M : \|\mathbf{x}\| = 1\right\}. \tag{13}$$

The unidimensional complex Grassmann manifold $\mathcal{G}(M,1)$ consists of the set of 1-D subspaces of $\mathsf{St}(M,1)$. Here, a transformation of the form $\mathbf{x}\mapsto e^{j\theta}\mathbf{x}$ (for any θ) is treated as invariant by considering all vectors of the form $e^{j\theta}\mathbf{x}$ (for some θ) to belong to the 1-D subspace spanned by $\mathbf{x}$. ■

The optimization objective is then to understand the structure of the beamforming vectors, $\{\mathbf{w}_{i,\mathsf{opt}}\}$, that maximize $\mathcal{R}_{\mathsf{sum}}$:

$$\mathbf{w}_{i,\mathsf{opt}} = \underset{\mathbf{w}_i\in\mathcal{G}(M,1)}{\arg\max}\ \mathcal{R}_{\mathsf{sum}}, i = 1,\ldots,M. \tag{14}$$

In (14), the candidate beamforming vectors, $\{\mathbf{w}_i\}$, depend only on the long-term statistics of the channel, which (as noted before) in the MISO setting is the set of all transmit covariance matrices, $\{\mathbf{\Sigma}_i\}$.

Toward the goal of computing $\mathcal{R}_{\mathsf{sum}}$, we decompose $\mathbf{h}_{\mathsf{iid},i}$ into its magnitude and directional components as $\mathbf{h}_{\mathsf{iid},i} = \|\mathbf{h}_{\mathsf{iid},i}\|\cdot\widehat{\mathbf{h}}_{\mathsf{iid},i}$. It is well known [1] that $\|\mathbf{h}_{\mathsf{iid},i}\|^2$ can be written as

$$\|\mathbf{h}_{\mathsf{iid},i}\|^2 = \frac{1}{2}\sum_{j=1}^{2M}(c_j)^2 \tag{15}$$

where $(c_j)^2$ is a standard (real) chi-squared random variable and $\widehat{\mathbf{h}}_{\mathsf{iid},i}$ is a unit-norm vector that is isotropically distributed [39], [40] on $\mathcal{G}(M,1)$. Thus, we can rewrite $I_{i,1}$ and $I_{i,2}$ as

$$I_{i,1} = \log\left(1+\frac{\rho}{M}\cdot\|\mathbf{h}_{\mathsf{iid},i}\|^2\cdot\widehat{\mathbf{h}}_{\mathsf{iid},i}^H\mathbf{V}_i\mathbf{\Lambda}_i\mathbf{V}_i^H\widehat{\mathbf{h}}_{\mathsf{iid},i}\right), \tag{16}$$

$$I_{i,2} = \log\left(1+\frac{\rho}{M}\cdot\|\mathbf{h}_{\mathsf{iid},i}\|^2\cdot\widehat{\mathbf{h}}_{\mathsf{iid},i}^H\widetilde{\mathbf{V}}_i\widetilde{\mathbf{\Lambda}}_i\widetilde{\mathbf{V}}_i^H\widehat{\mathbf{h}}_{\mathsf{iid},i}\right). \tag{17}$$

Since the magnitude and directional information of an i.i.d. random vector are independent [40], $E\left[I_{i,1}\right]$ and $E\left[I_{i,2}\right]$ can be further written as in (18) and (19) at the bottom of the page, where we have also used the fact that a fixed[8] unitary transformation of an isotropically distributed vector on $\mathcal{G}(M,1)$ does not alter its distribution.

III. Rate Characterization: Two-User Case

We now restrict attention to the special case of two users $(M=2)$ and focus on computing the ergodic information-theoretic rates given in (18) and (19) in closed form. The following theorem computes the ergodic rates as a function of the covariance matrices of the two links ($\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$), and the choice of beamforming vectors ($\mathbf{w}_1$ and $\mathbf{w}_2$).

Theorem 1: The ergodic information-theoretic rate achievable at user i (where $i=1,2$) with linear beamforming in the two-user case is given as in (20) at the bottom of this page, where $h(\bullet)$ is a monotonically increasing function defined as

$$h(x) \triangleq \exp\left(\frac{1}{x}\right)E_1\left(\frac{1}{x}\right), \quad x\in(0,\infty) \tag{21}$$

[8]Note that the unitary transformation is independent of the channel realization when the beamforming vectors are dependent only on the long-term statistics of the channel.

$$E\left[I_{i,1}\right] = E_{\|\mathbf{h}_{\mathsf{iid},i}\|}\left[E_{\widehat{\mathbf{h}}_{\mathsf{iid},i}}\left[\log\left(1+\frac{\rho}{M}\cdot\|\mathbf{h}_{\mathsf{iid},i}\|^2\cdot\widehat{\mathbf{h}}_{\mathsf{iid},i}^H\mathbf{\Lambda}_i\widehat{\mathbf{h}}_{\mathsf{iid},i}\right)\right]\right] \tag{18}$$

$$E\left[I_{i,2}\right] = E_{\|\mathbf{h}_{\mathsf{iid},i}\|}\left[E_{\widehat{\mathbf{h}}_{\mathsf{iid},i}}\left[\log\left(1+\frac{\rho}{M}\cdot\|\mathbf{h}_{\mathsf{iid},i}\|^2\cdot\widehat{\mathbf{h}}_{\mathsf{iid},i}^H\widetilde{\mathbf{\Lambda}}_i\widehat{\mathbf{h}}_{\mathsf{iid},i}\right)\right]\right] \tag{19}$$

$$E\left[R_i\right] = E\left[I_{i,1}\right] - E\left[I_{i,2}\right] = \frac{\mathbf{\Lambda}_{i,1}h\left(\frac{\rho\mathbf{\Lambda}_{i,1}}{2}\right) - \mathbf{\Lambda}_{i,2}h\left(\frac{\rho\mathbf{\Lambda}_{i,2}}{2}\right)}{\mathbf{\Lambda}_{i,1}-\mathbf{\Lambda}_{i,2}} - h\left(\frac{\rho\widetilde{\mathbf{\Lambda}}_{i,1}}{2}\right) \tag{20}$$

with $E_1(x) = \int_x^\infty \frac{e^{-t}}{t} dt$ denoting the Exponential integral [41]. The corresponding eigenvalues [cf., (7) and (10)] can be written in terms of $\mathbf{\Sigma}_i$ and the beamforming vectors as follows:

$$\mathbf{\Lambda}_{i,1} = \frac{A_i + B_i + \sqrt{(A_i - B_i)^2 + 4C_i^2}}{2} \tag{22}$$

$$\mathbf{\Lambda}_{i,2} = \frac{A_i + B_i - \sqrt{(A_i - B_i)^2 + 4C_i^2}}{2} \tag{23}$$

$$\widetilde{\mathbf{\Lambda}}_{i,1} = B_i, \tag{24}$$

where $A_i = \mathbf{w}_i^H \mathbf{\Sigma}_i \mathbf{w}_i$, $B_i = \mathbf{w}_j^H \mathbf{\Sigma}_i \mathbf{w}_j$, and $C_i = |\mathbf{w}_i^H \mathbf{\Sigma}_i \mathbf{w}_j|$ with $j \neq i$ and $i, j \in \{1, 2\}$.

Proof: Since $E[R_i] = E[I_{i,1}] - E[I_{i,2}]$, we start by computing $E[I_{i,1}]$. From (18), we have

$$E[I_{i,1}] = E_{\mathbf{X}} \left[\int_{y=\mathbf{\Lambda}_{i,2}}^{\mathbf{\Lambda}_{i,1}} \log\left(1 + \frac{\rho}{2} \cdot xy\right) p_i(y) dy \right] \tag{25}$$

where $\mathbf{X}$ stands for the random variable $\mathbf{X} = \|\mathbf{h}_{\mathsf{iid},i}\|^2$, x is a realization of $\mathbf{X}$ and $p_i(y)$ denotes the density function of

$$\widehat{\mathbf{h}}_{\mathsf{iid},i}^H \mathbf{\Lambda}_i \widehat{\mathbf{h}}_{\mathsf{iid},i} = \sum_{j=1}^{2} \mathbf{\Lambda}_{i,j} \left| \widehat{\mathbf{h}}_{\mathsf{iid},i}(j) \right|^2, \tag{26}$$

evaluated at y with $\mathbf{\Lambda}_{i,2} \leq y \leq \mathbf{\Lambda}_{i,1}$. That is, a closed-form computation of $E[I_{i,1}]$ requires the density function of weighted-norm of vectors isotropically distributed on $\mathcal{G}(2, 1)$. In Lemma 1 of Appendix A, we show that

$$p_i(y) = \frac{1}{\mathbf{\Lambda}_{i,1} - \mathbf{\Lambda}_{i,2}}, \quad \mathbf{\Lambda}_{i,2} \leq y \leq \mathbf{\Lambda}_{i,1}. \tag{27}$$

Using this information along with the chi-squared structure of $\|\mathbf{h}_{\mathsf{iid},i}\|^2$ [see (15)], we have the expression in (28) at the bottom of the page. Integrating out the y variable, we have (29), again at the bottom of the page. Following a routine computation using the list of integral table formula [42, 4.337(2), p. 572], we have the expression for $E[I_{i,1}]$. Particularizing this expression to the case of $E[I_{i,2}]$ in (19) with $\widetilde{\mathbf{\Lambda}}_{i,2} = 0$ results in the rate expression as in (20). To complete the proposition, an elementary computation of the eigenvalues of the associated 2×2 matrices in (7) and (10) results in their characterization. ■

The increasing nature of $h(\bullet)$, defined in (21), is illustrated in Fig. 1. Toward the goal of obtaining physical intuition on the structure of the optimal beamforming vectors, it is of interest to obtain the limiting form of the ergodic rates in the low- and the high-SNR extremes.

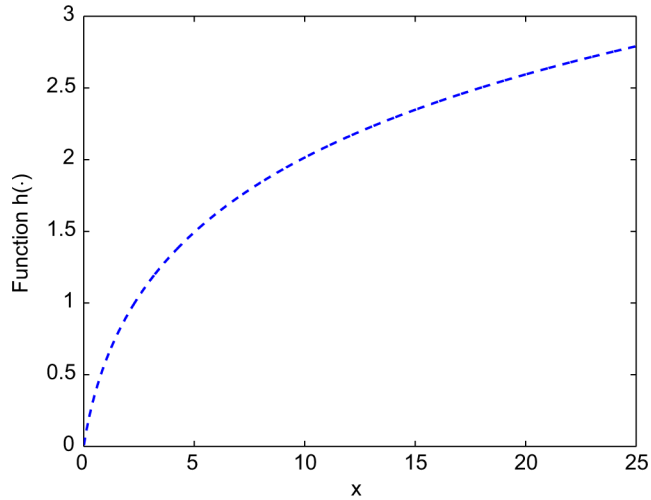

Fig. 1. Behavior of $h(x)$ for x satisfying $0 < x \leq 25$.

A. Low- SNR Extreme

Proposition 1: The ergodic rate $E[R_i]$ can be bounded as

$$1 - \rho \mathsf{C}_{\mathsf{low}} \leq \frac{E[R_i]}{\frac{\rho}{2}(\mathbf{\Lambda}_{i,1} + \mathbf{\Lambda}_{i,2} - B_i)} \leq 1 + \rho \mathsf{C}_{\mathsf{up}} \tag{30}$$

for some positive constants C_{up} and $\mathsf{C}_{\mathsf{low}}$ (not provided here for the sake of brevity) that depend only on the eigenvalues $\mathbf{\Lambda}_{i,1}, \mathbf{\Lambda}_{i,2}$, and $\widetilde{\mathbf{\Lambda}}_{i,1}$. Thus, as $\rho \to 0$, we have

$$\frac{E[R_i]}{\rho} \xrightarrow{\rho \to 0} \frac{1}{2}(\mathbf{\Lambda}_{i,1} + \mathbf{\Lambda}_{i,2} - B_i) \tag{31}$$

$$= \frac{A_i}{2} = \frac{\mathbf{w}_i^H \mathbf{\Sigma}_i \mathbf{w}_i}{2}. \tag{32}$$

Proof: We need the following bounds on the Exponential integral [41, 5.1.20, p. 229]:

$$\frac{x}{1 + 2x} \leq \frac{1}{2} \log(1 + 2x) \leq h(x) \leq \log(1 + x) \leq x \tag{33}$$

where the extremal inequalities are established by using the fact that

$$\frac{x}{x + 1} \leq \log(1 + x) \leq x. \tag{34}$$

Using these bounds, it is straightforward to see that the relationship in (30) holds. Note that both the upper and lower bounds converge to the same value as $\rho \to 0$, which results in the simplification in (32). ■

$$E[I_{i,1}] = \frac{1}{\mathbf{\Lambda}_{i,1} - \mathbf{\Lambda}_{i,2}} \cdot \int_{x=0}^{\infty} x e^{-x} \int_{y=\mathbf{\Lambda}_{i,2}}^{\mathbf{\Lambda}_{i,1}} \log\left(1 + \frac{\rho}{2} xy\right) dy dx \tag{28}$$

$$E[I_{i,1}] = \frac{2/\rho}{\mathbf{\Lambda}_{i,1} - \mathbf{\Lambda}_{i,2}} \left[\int_{x=0}^{\infty} \left[\left(1 + \frac{\rho}{2}\mathbf{\Lambda}_{i,1} x\right) \log\left(1 + \frac{\rho}{2}\mathbf{\Lambda}_{i,1} x\right) - \left(1 + \frac{\rho}{2}\mathbf{\Lambda}_{i,2} x\right) \log\left(1 + \frac{\rho}{2}\mathbf{\Lambda}_{i,2} x\right) \right] e^{-x} dx \right] - 1 \tag{29}$$

In the low-SNR extreme, the system is noise-limited, hence the linear scaling of $E[R_i]$ with ρ.

B. *High*-SNR *Extreme*

Proposition 2: As $\rho \to \infty$, we have

$$E[R_i] \to \frac{\mathbf{\Lambda}_{i,1}\log(\mathbf{\Lambda}_{i,1}) - \mathbf{\Lambda}_{i,2}\log(\mathbf{\Lambda}_{i,2})}{\mathbf{\Lambda}_{i,1} - \mathbf{\Lambda}_{i,2}} - \log(B_i) \quad (35)$$

$$= \frac{1}{2}\log\left(\frac{A_iB_i - C_i^2}{B_i^2}\right) + \frac{(A_i+B_i)/2}{\sqrt{(A_i-B_i)^2+4C_i^2}} \times \log\left(\frac{A_i+B_i+\sqrt{(A_i-B_i)^2+4C_i^2}}{A_i+B_i-\sqrt{(A_i-B_i)^2+4C_i^2}}\right) \quad (36)$$

with A_i, B_i, and C_i as in Theorem 1.

Proof: The following asymptotic expansion [41, 5.1.11, p. 229] of the Exponential integral is useful in obtaining the limiting form of $E[R_i]$ as $\rho \to \infty$:

$$E_1(x) = \log\left(\frac{1}{x}\right) + \sum_{k=1}^{\infty}\frac{(-1)^{k+1}x^k}{k\cdot k!} - \gamma \quad (37)$$

$$\xrightarrow{x\to 0} \log\left(\frac{1}{x}\right) + x - \gamma \quad (38)$$

where $\gamma \approx 0.577$ is the Euler–Mascheroni constant. Using the limiting value of $E_1(x)$ to approximate $E[R_i]$, we have the expression in (35). Expanding $\mathbf{\Lambda}_{i,1}$ and $\mathbf{\Lambda}_{i,2}$ in terms of A_i, B_i, and C_i, we have the expression in (36). ■

Unlike the low-SNR extreme, $E[R_i]$ is not a function of ρ here. The dominating impact of interference (due to the fixed nature of the beamforming vectors that are not adapted to the channel realizations) and the consequent boundedness of $E[R_i]$ in (35) as ρ increases should not be surprising. However, see the ergodic sum-rate comparison with a single-user scheme in Section VI-B.

IV. Sum-Rate Optimization: Two-User Case

We are now interested in understanding the structure of the optimal choice of beamforming vectors $(\mathbf{w}_{1,\mathsf{opt}}, \mathbf{w}_{2,\mathsf{opt}})$ that maximize $\mathcal{R}_{\mathsf{sum}}$ as a function of $\mathbf{\Sigma}_1$, $\mathbf{\Sigma}_2$ and ρ. This problem is difficult, in general. To obtain insight, we first consider the low- and the high-SNR extremes before studying the intermediate-SNR regime.

For simplicity, let us assume an eigendecomposition for $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ of the form

$$\mathbf{\Sigma}_1 = \mathbf{U}\mathsf{diag}([\lambda_1(\mathbf{\Sigma}_1), \lambda_2(\mathbf{\Sigma}_1)])\mathbf{U}^H, \quad (39)$$

$$\mathbf{\Sigma}_2 = \widetilde{\mathbf{U}}\mathsf{diag}([\lambda_1(\mathbf{\Sigma}_2), \lambda_2(\mathbf{\Sigma}_2)])\widetilde{\mathbf{U}}^H, \quad (40)$$

where $\mathbf{U} = [\mathbf{u}_1(\mathbf{\Sigma}_1), \mathbf{u}_2(\mathbf{\Sigma}_1)]$, $\widetilde{\mathbf{U}} = [\mathbf{u}_1(\mathbf{\Sigma}_2), \mathbf{u}_2(\mathbf{\Sigma}_2)]$, and $\lambda_1(\mathbf{\Sigma}_i) \geq \lambda_2(\mathbf{\Sigma}_i), i = 1, 2$. In particular, we assume that both $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ are positive definite, i.e., $\lambda_2(\mathbf{\Sigma}_i) > 0$.

A. *Low*-SNR *Extreme*

In the low-SNR regime, from Proposition 1, we see that the maximization of $E[R_i]$ involves optimizing over $\mathbf{w}_i$ alone. Thus, we have

$$\mathbf{w}_{i,\mathsf{opt}} = \arg\max_{\mathbf{w}_i}\mathcal{R}_{\mathsf{sum}} = \arg\max_{\mathbf{w}_i}\mathbf{w}_i^H\mathbf{\Sigma}_i\mathbf{w}_i = e^{j\nu_i}\mathbf{u}_1(\mathbf{\Sigma}_i) \quad (41)$$

for any choice of $\nu_i \in [0, 2\pi), i = 1, 2$.

Proposition 3: The ergodic sum-rate in the low-SNR extreme satisfies

$$\lim_{\rho\to 0}\frac{\mathcal{R}_{\mathsf{sum}}}{\rho} = \frac{1}{2}\cdot\Big[\lambda_1(\mathbf{\Sigma}_1) + \lambda_1(\mathbf{\Sigma}_2)\Big]. \quad (42)$$

■

In the low-SNR extreme, the optimal solution is such that the transmitter signals to a given user along the dominant statistical eigenmode of that user's channel and ignores the other user's channel completely. This is a solution motivated by the single-user viewpoint where the optimality of signalling along the dominant statistical eigenmode of the forward channel is well known [1]–[6]. This solution is not surprising since in the noise-limited regime, the broadcast channel is well approximated by separate single-user models connecting the transmitter to each receiver.

It is unreasonable to expect a low-complexity scheme such as linear beamforming to perform as well as the capacity-achieving DPC scheme at an arbitrary SNR. Nevertheless, it should not be surprising that in the low-SNR extreme, under the assumption that the DPC scheme is constrained to use equal power allocation, the low-SNR ergodic sum-rate performance of the linear beamforming scheme is the same as that of the DPC scheme. This "converse" result is presented in Appendix B.

B. *High*-SNR *Extreme*

Define $\mathbf{\Sigma}$ (and its corresponding eigendecomposition) as

$$\mathbf{\Sigma} \triangleq \mathbf{\Sigma}_2^{-\frac{1}{2}}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-\frac{1}{2}} = \mathbf{V}\mathsf{diag}([\eta_1, \eta_2])\mathbf{V}^H \quad (43)$$

where the unitary matrix $\mathbf{V} = [\mathbf{v}_1\mathbf{v}_2]$ and $\eta_1 \geq \eta_2$. Note that $\mathbf{\Sigma}$ is positive definite $(\eta_2 > 0)$ since both $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ are positive definite. The main result of this section is as follows.

Theorem 2: In the high-SNR extreme, the ergodic sum-rate is maximized by the following choice of beamforming vectors:

$$\mathbf{w}_{1,\mathsf{opt}} = e^{j\nu_1}\cdot\frac{\mathbf{\Sigma}_2^{-\frac{1}{2}}\mathbf{v}_1}{\|\mathbf{\Sigma}_2^{-\frac{1}{2}}\mathbf{v}_1\|}, \mathbf{w}_{2,\mathsf{opt}} = e^{j\nu_2}\cdot\frac{\mathbf{\Sigma}_2^{-\frac{1}{2}}\mathbf{v}_2}{\|\mathbf{\Sigma}_2^{-\frac{1}{2}}\mathbf{v}_2\|} \quad (44)$$

for any choice of $\nu_i \in [0, 2\pi), i = 1, 2$. The optimal ergodic sum-rate satisfies

$$\lim_{\rho\to\infty}\mathcal{R}_{\mathsf{sum}} = \frac{\kappa_1\log(\kappa_1)}{\kappa_1 - 1} + \frac{\log(\kappa_2)}{\kappa_2 - 1} \quad (45)$$

where

$$\kappa_1 \triangleq \frac{\eta_1\tau_2}{\eta_2\tau_1}, \quad \kappa_2 \triangleq \frac{\tau_2}{\tau_1}, \tag{46}$$

$$\tau_1 = \mathbf{v}_1^H\boldsymbol{\Sigma}_2^{-1}\mathbf{v}_1,\ \tau_2 = \mathbf{v}_2^H\boldsymbol{\Sigma}_2^{-1}\mathbf{v}_2,\ \tau_3 = \mathbf{v}_1^H\boldsymbol{\Sigma}_2^{-1}\mathbf{v}_2. \tag{47}$$

Proof: The first step in our proof is to rewrite the high-SNR rate expression in a form that permits further analysis. This is done in Appendixes C and D. With the definition of $\{\mathbf{v}_1, \mathbf{v}_2\}$ as in (43), since $\boldsymbol{\Sigma}_2$ is full-rank, we can decompose $\mathbf{w}_1$ and $\mathbf{w}_2$ as

$$\mathbf{w}_1 = \frac{\alpha\boldsymbol{\Sigma}_2^{-\frac{1}{2}}\mathbf{v}_1 + \beta\boldsymbol{\Sigma}_2^{-\frac{1}{2}}\mathbf{v}_2}{\|\alpha\boldsymbol{\Sigma}_2^{-\frac{1}{2}}\mathbf{v}_1 + \beta\boldsymbol{\Sigma}_2^{-\frac{1}{2}}\mathbf{v}_2\|} \tag{48}$$

$$\mathbf{w}_2 = \frac{\gamma\boldsymbol{\Sigma}_2^{-\frac{1}{2}}\mathbf{v}_1 + \delta\boldsymbol{\Sigma}_2^{-\frac{1}{2}}\mathbf{v}_2}{\|\gamma\boldsymbol{\Sigma}_2^{-\frac{1}{2}}\mathbf{v}_1 + \delta\boldsymbol{\Sigma}_2^{-\frac{1}{2}}\mathbf{v}_2\|} \tag{49}$$

for some choice of $\{\alpha, \beta, \gamma, \delta\}$ with $\alpha = |\alpha|e^{j\theta_\alpha}$ (similarly, for other quantities) satisfying $|\alpha|^2 + |\beta|^2 = |\gamma|^2 + |\delta|^2 = 1$. In Appendix E, we show that the ergodic sum-rate optimization over the 6-D parameter space $\{|\alpha|, |\gamma|, \theta_\alpha, \theta_\beta, \theta_\gamma, \theta_\delta\}$ results in the choice as in the statement of the theorem. ■

Many remarks are in order at this stage.

Remarks:

• Recall the definition of a generalized eigenvector:

Definition 2 (Generalized Eigenvector [43]): A generalized eigenvector $\mathbf{x}$ (with the corresponding generalized eigenvalue σ) of a pair of matrices $(\mathbf{A}, \mathbf{B})$ satisfies the relationship

$$\mathbf{A}\mathbf{x} = \sigma\mathbf{B}\mathbf{x}. \tag{50}$$

In the special case where $\mathbf{B}$ is invertible, a generalized eigenvector of the pair $(\mathbf{A}, \mathbf{B})$ is also an eigenvector of $\mathbf{B}^{-1}\mathbf{A}$. If $\mathbf{A}$ and $\mathbf{B}$ are also positive-definite, then all the generalized eigenvalues are positive. While a unit-norm generalized eigenvector (or an eigenvector) is not unique on $\mathsf{St}(M, 1)$, it is unique on $\mathcal{G}(M, 1)$. ■

We decompose $\{\mathbf{w}_1, \mathbf{w}_2\}$ in (48) and (49) along the basis[9] $\{\boldsymbol{\Sigma}_2^{-\frac{1}{2}}\mathbf{v}_1, \boldsymbol{\Sigma}_2^{-\frac{1}{2}}\mathbf{v}_2\}$ instead of the more routine basis $\{\mathbf{v}_1, \mathbf{v}_2\}$. The reason for this peculiar choice is as follows. It turns out that $\boldsymbol{\Sigma}_2^{-\frac{1}{2}}\mathbf{v}_1$ and $\boldsymbol{\Sigma}_2^{-\frac{1}{2}}\mathbf{v}_2$ are the dominant generalized eigenvectors (corresponding to the largest generalized eigenvalue) of the pairs $(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2)$ and $(\boldsymbol{\Sigma}_2, \boldsymbol{\Sigma}_1)$, respectively. For this claim, we use (43) to note that

$$\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2^{-\frac{1}{2}}\big(\boldsymbol{\Sigma}_2^{-\frac{1}{2}}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-\frac{1}{2}}\big)\boldsymbol{\Sigma}_2^{\frac{1}{2}} = \mathbf{M}\mathbf{D}\mathbf{M}^{-1} \tag{51}$$

$$\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_2 = \big(\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1\big)^{-1} = \mathbf{M}\mathbf{D}^{-1}\mathbf{M}^{-1} \tag{52}$$

where $\mathbf{M} = \boldsymbol{\Sigma}_2^{-\frac{1}{2}}\mathbf{V}$ and $\mathbf{D} = \mathsf{diag}\big([\eta_1\eta_2]\big)$. This means that we can write

$$\mathbf{w}_{1,\mathsf{opt}} = e^{j\nu_1} \cdot \mathbf{u}_1\big(\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1\big) \tag{53}$$

$$\mathbf{w}_{2,\mathsf{opt}} = e^{j\nu_2} \cdot \mathbf{u}_2\big(\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1\big) \tag{54}$$

[9]Note that $\boldsymbol{\Sigma}_2$ is a full-rank matrix and hence, the vectors $\boldsymbol{\Sigma}_2^{-\frac{1}{2}}\mathbf{v}_1$ and $\boldsymbol{\Sigma}_2^{-\frac{1}{2}}\mathbf{v}_2$ form a nonorthogonal basis (in general), whereas $\{\mathbf{v}_1, \mathbf{v}_2\}$ is orthonormal.

where $\mathbf{u}_1(\bullet)$ and $\mathbf{u}_2(\bullet)$ are the dominant and nondominant eigenvectors, respectively. Using the generalized eigenvector structure, it is easy to see that

$$\mathbf{u}_1\big(\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1\big) = \mathbf{u}_2\big(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_2\big) \tag{55}$$

$$\mathbf{u}_2\big(\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1\big) = \mathbf{u}_1\big(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_2\big) \tag{56}$$

and thus, we also have

$$\mathbf{w}_{1,\mathsf{opt}} = e^{j\nu_1} \cdot \mathbf{u}_1\big(\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1\big) \tag{57}$$

$$\mathbf{w}_{2,\mathsf{opt}} = e^{j\nu_2} \cdot \mathbf{u}_1\big(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_2\big). \tag{58}$$

• Given that the transmitter has only statistical information of the two links, a natural candidate for beamforming in the high-SNR extreme is the solution to the maximization of an appropriately defined average SINR metric for each user. Motivated by the fact [see (3)] that the instantaneous sum-rate for the ith user (R_i) is an increasing function of $|\mathbf{h}_i^H\mathbf{w}_i|^2$, whereas R_j (for $j \neq i$) is a decreasing function of $|\mathbf{h}_j^H\mathbf{w}_i|^2$, we define an "average" SINR metric as follows:

$$\mathsf{SINR}_i \triangleq \frac{E\left[|\mathbf{h}_i^H\mathbf{w}_i|^2\right]}{E\left[|\mathbf{h}_j^H\mathbf{w}_i|^2\right]} = \frac{\mathbf{w}_i^H\boldsymbol{\Sigma}_i\mathbf{w}_i}{\mathbf{w}_i^H\boldsymbol{\Sigma}_j\mathbf{w}_i}. \tag{59}$$

The optimization problem of interest is to maximize SINR_i which has the generalized eigenvector structure as solution [44]:

$$\underset{\mathbf{w}_i:\mathbf{w}_i^H\mathbf{w}_i=1}{\arg\max}\ \mathsf{SINR}_i = e^{j\nu_i}\mathbf{u}_1\big(\boldsymbol{\Sigma}_j^{-1}\boldsymbol{\Sigma}_i\big), \quad j \neq i,\ i, j \in \{1, 2\}. \tag{60}$$

It follows that if user i selfishly maximizes (its own) SINR_i metric, then the set of such beamforming vectors *also* maximize the ergodic sum-rate in the high-SNR regime. Note that this is just an observation and establishing the *equivalence* (if true) between the SINR maximization problem and the ergodic sum-rate maximization problem is a different task.

Nevertheless, in this sense, the solution to the broadcast problem mirrors and generalizes the single-user setting, where the optimality of signalling along the statistical eigenmodes of the channel is well understood [1]–[6]. Further, while optimal beamformer solutions in terms of the generalized eigenvectors are obtained in the perfect CSI case of the broadcast setting for the beamforming design problem [26], [34] and the interference channel problem [35], this solution in the statistical case is a first. A similar result is obtained in a related work of ours [36] on statistical beamforming vector design for the interference channel case. Since the generalized eigenvector solution has an intuitive explanation, it is of interest to obtain useful insights on the optimality of this solution in more general multiuser settings.

• The ergodic sum-rate in (45) is increasing in κ_1 and thus in $\frac{\eta_1}{\eta_2}$. We now observe that ill-conditioning of $\boldsymbol{\Sigma}_1$ is necessary and sufficient to ensure that $\frac{\eta_1}{\eta_2}$ is large. For this, we use standard eigenvalue inequalities for product of Hermitian matrices [44] to see that

$$\frac{\chi_1}{\chi_2} = \frac{\lambda_1(\boldsymbol{\Sigma}_1)\lambda_2(\boldsymbol{\Sigma}_2^{-1})}{\lambda_2(\boldsymbol{\Sigma}_1)\lambda_1(\boldsymbol{\Sigma}_2^{-1})} \leq \frac{\eta_1}{\eta_2} \leq \frac{\lambda_1(\boldsymbol{\Sigma}_1)\lambda_1(\boldsymbol{\Sigma}_2^{-1})}{\lambda_2(\boldsymbol{\Sigma}_1)\lambda_2(\boldsymbol{\Sigma}_2^{-1})} = \chi_1\chi_2 \tag{61}$$

where $\chi_i = \frac{\lambda_1(\mathbf{\Sigma}_i)}{\lambda_2(\mathbf{\Sigma}_i)}, i = 1, 2$. In other words, the more ill-conditioned $\mathbf{\Sigma}_1$ is, the larger the high-SNR statistical beamforming sum-rate asymptote is (and *vice versa*).

On the other hand, the ergodic sum-rate in (45) is not monotonic in $\frac{\tau_1}{\tau_2}$. Nevertheless, it can be seen that as a function of $\frac{\tau_1}{\tau_2}$, it has local maxima as $\frac{\tau_1}{\tau_2} \to 0$ and $\frac{\tau_1}{\tau_2} \to \infty$, and a minimum at $\frac{\tau_1}{\tau_2} = 1$. The more well-conditioned $\mathbf{\Sigma}_2$ is, the more closer $\frac{\tau_1}{\tau_2}$ is to 1, and hence, the high-SNR statistical beamforming sum-rate asymptote is minimized. If $\mathbf{\Sigma}_2$ is ill-conditioned, the value taken by $\frac{\tau_1}{\tau_2}$ depends on the angle between the dominant eigenvectors of $\mathbf{\Sigma}_2$ and $\mathbf{\Sigma}$. If the two eigenvectors are nearly parallel, $\frac{\tau_1}{\tau_2}$ is close to zero and if they are nearly perpendicular, $\frac{\tau_1}{\tau_2}$ is very large. In either case, the high-SNR statistical beamforming sum-rate asymptote is locally maximized.

The conclusion from the above analysis is that among all possible channels, the ergodic sum-rate is maximized (or minimized) when $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ are both ill-(or well-)conditioned. In other words, if both the users encounter poor scattering (that leads to an ill-conditioning of their respective covariance matrices), their fading is spatially localized. The transmitter can simultaneously excite these spatial localizations without causing a proportional increase in the interference level of the other user thus resulting in a higher ergodic sum-rate. On the other hand, rich scattering implies that fading is spatially isotropic for both the users. Any spatially localized excitation for one user will cause an isotropic interference level at the other user thus resulting in a smaller ergodic sum-rate.

• A special case that is of considerable interest is when $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ have the same set of orthonormal eigenvectors. This would be a suitable model for certain indoor scenarios where the antenna separation for the two users is the same [28]. Denoting (for simplicity) the set of common eigenvectors by $\mathbf{u}_1$ and $\mathbf{u}_2$, we can decompose $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ as

$$\mathbf{\Sigma}_1 = \big[\mathbf{u}_1, \mathbf{u}_2\big]\mathsf{diag}\big([\lambda_1, \lambda_2]\big)\big[\mathbf{u}_1, \mathbf{u}_2\big]^H, \tag{62}$$

$$\mathbf{\Sigma}_2 = \big[\mathbf{u}_1, \mathbf{u}_2\big]\mathsf{diag}\big([\mu_1, \mu_2]\big)\big[\mathbf{u}_1, \mathbf{u}_2\big]^H. \tag{63}$$

We reuse the notations χ_1 and χ_2 to denote

$$\chi_1 \triangleq \frac{\lambda_1}{\lambda_2} \text{ and } \chi_2 \triangleq \frac{\mu_1}{\mu_2}. \tag{64}$$

Without loss in generality, we can assume that $\chi_1 \geq 1$. Two scenarios[10] arise depending on the relationship between χ_1 and χ_2: i) $\chi_1 \geq \chi_2$, and ii) $\chi_1 < \chi_2$.

Theorem 3: In the high-SNR extreme, the ergodic sum-rate is maximized by the following choice of beamforming vectors:

$$\begin{array}{l} \mathbf{w}_{1,\mathsf{opt}} = e^{j\nu_1}\mathbf{u}_1, \mathbf{w}_{2,\mathsf{opt}} = e^{j\nu_2}\mathbf{u}_2 \quad \text{if } \chi_1 \geq \chi_2, \\ \mathbf{w}_{1,\mathsf{opt}} = e^{j\nu_2}\mathbf{u}_2, \mathbf{w}_{2,\mathsf{opt}} = e^{j\nu_1}\mathbf{u}_1 \quad \text{if } \chi_1 < \chi_2 \end{array} \tag{65}$$

for any choice of $\nu_i \in [0, 2\pi), i = 1, 2$. The optimal ergodic sum-rate satisfies

$$\lim_{\rho\to\infty} \mathcal{R}_{\mathsf{sum}} = \begin{cases} \frac{\chi_1 \cdot \log(\chi_1)}{\chi_1 - 1} + \frac{\log(\chi_2)}{\chi_2 - 1} & \text{if } \chi_1 \geq \chi_2 \\ \frac{\chi_2 \cdot \log(\chi_2)}{\chi_2 - 1} + \frac{\log(\chi_1)}{\chi_1 - 1} & \text{if } \chi_1 < \chi_2. \end{cases} \tag{66}$$

Proof: While Theorem 2 can be particularized to this special case easily, we pursue an alternate proof technique in Appendix F that exploits the comparative relationship between τ_1 and τ_2 (which is possible in the special case) and the fact that $\tau_3 = 0$. ■

A comparison of the proof techniques of Theorems 2 and 3 is presented in Appendix G.

• Some remarks on the optimization setup of this paper are necessary. The proofs of Theorems 2 and 3 require us to consider a 6-D optimization over the parameter space of $\{|\alpha|, |\gamma|, \theta_\alpha, \theta_\beta, \theta_\gamma, \theta_\delta\}$. As a result, any geometric interpretation of the optimization seems intractable. A naïve approach to the 6-D optimization problems in this paper is to adopt the method of matrix differentiation calculus. For this approach to work, we need to show that both the set over which optimization is done as well as the optimized function are convex, neither of which is true in our case. Specifically, neither $\mathcal{G}(2, 1)$ nor $\mathsf{St}(2, 1)$ are convex sets. It also turns out that the ergodic sum-rate is neither convex nor concave[11] over the set of beamforming vectors, even over a locally convex domain or an extended convex domain (like the interior of the sphere).

• The approach adopted in Appendixes E and F overcomes these difficulties, and it consists of two steps. In the first step, we produce an upper bound to the ergodic sum-rate that is independent of the optimization parameters. In the second step, we show that this upper bound can be realized by a specific choice of beamforming vectors, thereby confirming that choice's optimality. This approach seems to be the most natural (and first-principles-based) recourse to solving the nonconvex optimization problem at hand. An alternate approach to optimize the ergodic sum-rate is nonlinear/manifold optimization theory. But this approach is fraught with complicated Hessian calculations and technical difficulties such as distinguishing between local and global extrema.

C. *Intermediate*-SNR *Regime: Candidate Beamforming Vectors*

While physical intuition on the structure of the optimal ergodic sum-rate maximizing beamforming vectors has been obtained in the low- and the high-SNR extremes, the essentially intractable nature of the Exponential integral in the ergodic rate expressions of Theorem 1 means that such a possibility at an arbitrary SNR is difficult. Nevertheless, the single-user setup [45], [46] suggests that the optimal beamforming vectors (that determine the modes that are excited) and the power allocation across these modes can be continuously parameterized by a function of the SNR. Motivated by the single-user case, a desirable quality for a "good" beamforming vector structure ($\{\mathbf{w}_{i,\mathsf{cand}}(\rho), i = 1, 2\}$) at an arbitrary SNR of ρ is that the limiting behavior of such a structure in the low- and the high-SNR extremes should be the solutions of Proposition 3 and Theorem 2. That is,

$$\lim_{\rho\to 0} \mathbf{w}_{i,\mathsf{cand}}(\rho) = e^{j\nu_i}\mathbf{u}_1(\mathbf{\Sigma}_i), \tag{67}$$

$$\lim_{\rho\to\infty} \mathbf{w}_{i,\mathsf{cand}}(\rho) = e^{j\nu_i}\frac{\mathbf{\Sigma}_2^{-\frac{1}{2}}\mathbf{v}_i}{\|\mathbf{\Sigma}_2^{-\frac{1}{2}}\mathbf{v}_i\|}, i = 1, 2, \tag{68}$$

[10]These possibilities arise because even though the set of eigenvectors of $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ are the same, there is no specific reason to expect the dominant eigenvector of $\mathbf{\Sigma}_1$ to also be a dominant eigenvector of $\mathbf{\Sigma}_2$. Observe that the first case subsumes the setting where $\mu_1 = \mu_2 = \mu$ and $\mathbf{\Sigma}_2 = \mu\mathbf{I}$.

[11]It is possible that some function of the ergodic sum-rate may be convex. But we are not aware of any likely candidate that could work.

where the above limits are seen as manifold operations [39] on $\mathcal{G}(2,1)$.

A natural candidate that meets (67) and (68) is the following choice parameterized by $\alpha(\rho)$ and $\beta(\rho)$ satisfying $\{\alpha(\rho),\beta(\rho)\} \in [0,\infty)$ and $\nu_i \in [0,2\pi), i=1,2$:

$$\mathbf{w}_{1,\mathsf{cand}}(\rho) = e^{j\nu_1} \cdot \mathsf{Dom.eig.}\left(\left(\alpha(\rho)\mathbf{\Sigma}_2 + \mathbf{I}\right)^{-1}\mathbf{\Sigma}_1\right) \quad (69)$$

$$\mathbf{w}_{2,\mathsf{cand}}(\rho) = e^{j\nu_2} \cdot \mathsf{Dom.eig.}\left(\left(\beta(\rho)\mathbf{\Sigma}_1 + \mathbf{I}\right)^{-1}\mathbf{\Sigma}_2\right) \quad (70)$$

where the notation $\mathsf{Dom.eig}(\bullet)$ stands for the unit-norm dominant eigenvector operation. These vectors can be seen to be solutions to the following optimization problems:

$$\underset{\mathbf{w}_i:\mathbf{w}_i^H\mathbf{w}_i=1}{\arg\max}\ \mathsf{SINR}_i = \mathbf{w}_{i,\mathsf{cand}}(\rho), \quad i=1,2 \quad (71)$$

where

$$\mathsf{SINR}_1 = \frac{\mathbf{w}_1^H\mathbf{\Sigma}_1\mathbf{w}_1}{\mathbf{w}_1^H\mathbf{w}_1 + \alpha(\rho)\mathbf{w}_1^H\mathbf{\Sigma}_2\mathbf{w}_1} \quad (72)$$

$$\mathsf{SINR}_2 = \frac{\mathbf{w}_2^H\mathbf{\Sigma}_2\mathbf{w}_2}{\mathbf{w}_2^H\mathbf{w}_2 + \beta(\rho)\mathbf{w}_2^H\mathbf{\Sigma}_1\mathbf{w}_2}. \quad (73)$$

The choice in (69) and (70) is a low-dimensional mapping from $\mathcal{G}(2,1)\times\mathcal{G}(2,1)$ to $\mathbb{R}^+\times\mathbb{R}^+$, thus considerably simplifying the search space for candidate beamforming vectors. It must be noted that while the search space is simplified, the generalized eigenvector operation is a nonlinear mapping [43] in $\alpha(\rho)$ and $\beta(\rho)$.

While we are unable to prove the optimality structure of the proposed scheme in the intermediate-SNR regime, motivated by our numerical studies in Section IV-A, we make the following conjecture.

Conjecture 1: In the intermediate-SNR regime, the ergodic sum-rate is maximized by the following choice of beamforming vectors:

$$\mathbf{w}_{1,\mathsf{opt}} = e^{j\nu_1} \cdot \mathsf{Dom.eig.}\left(\left(\alpha^\star(\rho)\mathbf{\Sigma}_2 + \mathbf{I}\right)^{-1}\mathbf{\Sigma}_1\right) \quad (74)$$

$$\mathbf{w}_{2,\mathsf{opt}} = e^{j\nu_2} \cdot \mathsf{Dom.eig.}\left(\left(\beta^\star(\rho)\mathbf{\Sigma}_1 + \mathbf{I}\right)^{-1}\mathbf{\Sigma}_2\right) \quad (75)$$

for any choice of $\nu_i \in [0,2\pi), i=1,2$ and

$$\{\alpha^\star(\rho),\beta^\star(\rho)\} = \underset{\{\alpha(\rho),\beta(\rho)\}}{\arg\max}\ E[R_1] + E[R_2]$$
$$\text{s.t.}\ \mathbf{w}_1 = \mathbf{w}_{1,\mathsf{cand}}(\rho), \mathbf{w}_2 = \mathbf{w}_{2,\mathsf{cand}}(\rho). \quad (76)$$

■

V. Ergodic Sum-Rate: Generalizations

We studied the structure of ergodic sum-rate maximizing beamforming vectors in Section IV. In this section, we consider more general problems of this nature.

A. Maximizing $E[R_i]$

Consider a system where the quality-of-service (QoS) metric of one user significantly dominates that of the other user, e.g., one user is considerably more important to the network operator than the other. The relevant metric to optimize in this scenario is not the ergodic sum-rate, but the rate achievable by the more important user. In this setting, we have the following result.

Proposition 4: The optimal choice of the pair $(\mathbf{w}_{1,\mathsf{opt}}, \mathbf{w}_{2,\mathsf{opt}})$ that maximizes $E[R_i]$ is

i) Low-SNR Extreme:

$$\begin{cases} \mathbf{w}_{i,\mathsf{opt}} = e^{j\nu_1}\mathbf{u}_1(\mathbf{\Sigma}_i) \\ \mathbf{w}_{j,\mathsf{opt}} = \text{any vector on } \mathcal{G}(2,1),\ j\neq i. \end{cases} \quad (77)$$

ii) High-SNR Extreme:

$$\begin{cases} \mathbf{w}_{i,\mathsf{opt}} = e^{j\nu_1}\mathbf{u}_1(\mathbf{\Sigma}_i) \\ \mathbf{w}_{j,\mathsf{opt}} = e^{j\nu_2}\mathbf{u}_2(\mathbf{\Sigma}_i), \quad j\neq i \end{cases} \quad (78)$$

for any choice of $\nu_i \in [0,2\pi), i=1,2$.

Proof: See Appendix H. ■

With the above choice of beamforming vectors, $E[R_i]$ can be written as

$$E[R_i] \xrightarrow{\rho\to\infty} \frac{\chi_i\log(\chi_i)}{\chi_i - 1} \quad (79)$$

where $\chi_i = \frac{\lambda_1(\mathbf{\Sigma}_i)}{\lambda_2(\mathbf{\Sigma}_i)}$. From (79), it is to be noted that $E[R_i]$ increases as χ_i increases. That is, the more ill-conditioned $\mathbf{\Sigma}_i$ is, the larger the high-SNR statistical beamforming rate asymptote is (and *vice versa*). This should be intuitive as our goal is only to maximize $E[R_i]$ and the beamforming vectors in (78) achieve that goal.

B. Weighted Ergodic Sum-Rate Maximization

In a system where the QoS metrics of the two users are comparable (but not the same), the relevant metric to optimize is the weighted-sum of ergodic rates achievable by the two users [37]. Specifically, the objective function here is

$$\mathcal{R}_{\mathsf{weighted}} = \zeta_1 E[R_1] + \zeta_2 E[R_2] \quad (80)$$

for some choice of weights ζ_1 and ζ_2 satisfying (without loss in generality) $\{\zeta_1,\zeta_2\} \in [0,1]$. Note that $E[R_i]$ is a special case of this objective function with $\zeta_1=1,\zeta_2=0$ or $\zeta_1=0,\zeta_2=1$.

Maximizing $\mathcal{R}_{\mathsf{weighted}}$ to obtain a closed-form characterization of the optimal beamforming vectors seems hard in general. Motivated by the study for the sum-rate in the intermediate-SNR regime in Section IV-C, we now consider a set of candidate beamforming vectors that produce known optimal structures in special cases. For this, it is important to note that no choice of $\alpha(\rho)$ and $\beta(\rho)$ in (69) and (70) can produce the beamforming vectors in (78). A candidate set of beamforming vectors that not only produces the special (extreme) cases in the ergodic sum-rate setting, but also (78), is the following choice parameterized by four quantities, $\{\alpha(\rho),\beta(\rho),\gamma(\rho),\delta(\rho)\} \in [0,\infty)$:

$$\mathbf{w}_{1,\mathsf{weighted,cand}}(\rho) = e^{j\nu_1} \cdot \mathsf{Dom.eig.}\left(\left(\alpha(\rho)\mathbf{\Sigma}_2 + \mathbf{I}\right)^{-1}\left(\gamma(\rho)\mathbf{\Sigma}_1 + \mathbf{I}\right)\right) \quad (81)$$

$$\mathbf{w}_{2,\mathsf{weighted,cand}}(\rho) = e^{j\nu_2} \cdot \mathsf{Dom.eig.}\left(\left(\beta(\rho)\mathbf{\Sigma}_1 + \mathbf{I}\right)^{-1}\left(\delta(\rho)\mathbf{\Sigma}_2 + \mathbf{I}\right)\right) \quad (82)$$

where the notation $\mathsf{Dom.eig}(\bullet)$ stands for the usual unit-norm dominant eigenvector operation. As before, (81)–(82) corresponds to a low-dimensional map from $\{\mathcal{G}(2,1)\}^4$ to $\{[0,\infty)\}^4$ and thus a simplification in the search for a good beamformer structure.

While the structure of the optimal beamforming vectors is difficult to establish, motivated by the studies in Section VI-A, we pose the following conjecture.

Conjecture 2: In the intermediate-SNR regime, the weighted ergodic sum-rate, $\mathcal{R}_{\mathsf{weighted}}$, is maximized by the following choice of beamforming vectors:

$$\mathbf{w}_1 = e^{j\nu_1}\mathsf{Dom.eig.}\left(\left(\alpha^\star(\rho)\mathbf{\Sigma}_2+\mathbf{I}\right)^{-1}\left(\gamma^\star(\rho)\mathbf{\Sigma}_1+\mathbf{I}\right)\right) \tag{83}$$

$$\mathbf{w}_2 = e^{j\nu_2}\mathsf{Dom.eig.}\left(\left(\beta^\star(\rho)\mathbf{\Sigma}_1+\mathbf{I}\right)^{-1}\left(\delta^\star(\rho)\mathbf{\Sigma}_2+\mathbf{I}\right)\right) \tag{84}$$

for any choice of $\nu_i \in [0,2\pi), i = 1,2$ and where $\{\alpha^\star(\rho),\beta^\star(\rho),\gamma^\star(\rho),\delta^\star(\rho)\}$ satisfy

$$\begin{aligned}&\{\alpha^\star(\rho),\beta^\star(\rho),\gamma^\star(\rho),\delta^\star(\rho)\}\\ &= \underset{\{\alpha(\rho),\beta(\rho),\gamma(\rho),\delta(\rho)\}}{\arg\max}\ \zeta_1 E\left[R_1\right]+\zeta_2 E\left[R_2\right]\\ \text{s.t.}\quad &\mathbf{w}_1 = \mathbf{w}_{1,\mathsf{weighted,cand}}(\rho), \mathbf{w}_2 = \mathbf{w}_{2,\mathsf{weighted,cand}}(\rho).\end{aligned} \tag{85}$$

■

C. Rank-Deficient Case

Following up on Remark 3 in Section IV-B, we now consider the extreme case where both[12] $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ are rank-deficient in more detail.

Proposition 5: The ergodic information-theoretic rate achievable at user i is as in (86) at the bottom of the page, where $h(\bullet)$ is as in (21) and the eigendecomposition of $\mathbf{\Sigma}_i$ is

$$\mathbf{\Sigma}_i = \lambda_1\left(\mathbf{\Sigma}_i\right)\cdot\mathbf{u}_1\left(\mathbf{\Sigma}_i\right)\mathbf{u}_1\left(\mathbf{\Sigma}_i\right)^H, \quad i=1,2. \tag{87}$$

[12]The case when only one of the $\mathbf{\Sigma}_i$ is rank-deficient can be studied along analogous lines and no details are provided.

Proof: While Theorem 1 (as stated) is explicitly dependent on both $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ being of full rank and is hence not directly applicable in this extreme setting, much of the analysis follows through. The key to the proof is that all the results in Appendix A (Lemmas 1 and 2) also hold when some of the diagonal entries of $\mathbf{\Lambda}_i$ are zero. In fact, this fact is implicitly used to compute $E\left[I_{i,2}\right]$ in Theorem 1. ■

D. Three-User Case: $M = 3$

We now consider the task of generalizing Theorem 1 to the three-user ($M=3$) case.

Proposition 6: The ergodic information-theoretic rate achievable at user i (where $i = 1,2,3$) with linear beamforming in the three-user case

$$E\left[R_i\right] = E\left[I_{i,1}\right] - E\left[I_{i,2}\right] \tag{88}$$

is as in (89) at the bottom of the page, where $h(\bullet)$ is as in (21). The eigenvalue matrices $\mathbf{\Lambda}_i = \mathsf{diag}\left(\left[\mathbf{\Lambda}_{i,1},\mathbf{\Lambda}_{i,2},\mathbf{\Lambda}_{i,3}\right]\right)$ and $\widetilde{\mathbf{\Lambda}}_i = \mathsf{diag}\left(\left[\widetilde{\mathbf{\Lambda}}_{i,1},\widetilde{\mathbf{\Lambda}}_{i,2},\widetilde{\mathbf{\Lambda}}_{i,3}\right]\right)$ are defined as in (7) and (10), and can be obtained in terms of the beamforming vectors and the covariance matrices by solving the associated cubic equations.

Proof: The proof is tedious, but follows along the lines of Theorem 1. The first step is in characterizing $p_i(y)$, which is done in Lemma 2 of Appendix A. We can then generalize (28) using (148) as in (90)–(92) at the bottom of the next page. These integrals are cumbersome, but straightforward to compute using [42, 4.337(2), 4.337(5), p. 572]. The result is the expression in the statement of the proposition. ■

E. General M-User Case

As can be seen from Appendix A (Lemma 2), the expression for $p_i(y)$ becomes more complicated as M increases. Without a recourse to $p_i(y)$, closed-form expressions for the ergodic sum-rate of the linear beamforming scheme can be obtained using a recent advance in [31] and [32] that allows the computation of the density function of weighted-sum of standard central chi-squared terms (*generalized* chi-squared random variables).

$$E\left[R_i\right] = h\left(\frac{\rho}{2}\lambda_1\left(\mathbf{\Sigma}_i\right)\left(\left|\mathbf{u}_1\left(\mathbf{\Sigma}_i\right)^H\mathbf{w}_i\right|^2+\left|\mathbf{u}_1\left(\mathbf{\Sigma}_i\right)^H\mathbf{w}_j\right|^2\right)\right) - h\left(\frac{\rho}{2}\lambda_1\left(\mathbf{\Sigma}_i\right)\left|\mathbf{u}_1\left(\mathbf{\Sigma}_i\right)^H\mathbf{w}_j\right|^2\right), \quad j\neq i,\ i=1,2, \tag{86}$$

$$\begin{aligned}E[R_i] = &\frac{\mathbf{\Lambda}_{i,1}^2\cdot h\left(\frac{\rho\mathbf{\Lambda}_{i,1}}{3}\right)}{\left(\mathbf{\Lambda}_{i,1}-\mathbf{\Lambda}_{i,2}\right)\left(\mathbf{\Lambda}_{i,1}-\mathbf{\Lambda}_{i,3}\right)} - \frac{\mathbf{\Lambda}_{i,2}^2\cdot h\left(\frac{\rho\mathbf{\Lambda}_{i,2}}{3}\right)}{\left(\mathbf{\Lambda}_{i,1}-\mathbf{\Lambda}_{i,2}\right)\left(\mathbf{\Lambda}_{i,2}-\mathbf{\Lambda}_{i,3}\right)} + \frac{\mathbf{\Lambda}_{i,3}^2\cdot h\left(\frac{\rho\mathbf{\Lambda}_{i,3}}{3}\right)}{\left(\mathbf{\Lambda}_{i,1}-\mathbf{\Lambda}_{i,3}\right)\left(\mathbf{\Lambda}_{i,2}-\mathbf{\Lambda}_{i,3}\right)}\\ &- \frac{\widetilde{\mathbf{\Lambda}}_{i,1}^2\cdot h\left(\frac{\rho\widetilde{\mathbf{\Lambda}}_{i,1}}{3}\right)}{\left(\widetilde{\mathbf{\Lambda}}_{i,1}-\widetilde{\mathbf{\Lambda}}_{i,2}\right)\left(\widetilde{\mathbf{\Lambda}}_{i,1}-\widetilde{\mathbf{\Lambda}}_{i,3}\right)} + \frac{\widetilde{\mathbf{\Lambda}}_{i,2}^2\cdot h\left(\frac{\rho\widetilde{\mathbf{\Lambda}}_{i,2}}{3}\right)}{\left(\widetilde{\mathbf{\Lambda}}_{i,1}-\widetilde{\mathbf{\Lambda}}_{i,2}\right)\left(\widetilde{\mathbf{\Lambda}}_{i,2}-\widetilde{\mathbf{\Lambda}}_{i,3}\right)} - \frac{\widetilde{\mathbf{\Lambda}}_{i,3}^2\cdot h\left(\frac{\rho\widetilde{\mathbf{\Lambda}}_{i,3}}{3}\right)}{\left(\widetilde{\mathbf{\Lambda}}_{i,1}-\widetilde{\mathbf{\Lambda}}_{i,3}\right)\left(\widetilde{\mathbf{\Lambda}}_{i,2}-\widetilde{\mathbf{\Lambda}}_{i,3}\right)}\end{aligned} \tag{89}$$

For example, if $\mathbf{\Lambda}_i = \mathsf{diag}\left([\mathbf{\Lambda}_{i,1}, \ldots, \mathbf{\Lambda}_{i,M}]\right)$ and $\mathbf{\Lambda}_{i,j}, j = 1, \ldots, M$ are distinct,[13] we have

$$E\left[I_{i,1}\right] = \sum_{k=1}^{M} \prod_{j=1, j\neq k}^{M} \frac{\mathbf{\Lambda}_{i,k}}{\mathbf{\Lambda}_{i,k} - \mathbf{\Lambda}_{i,j}} \cdot h\left(\frac{\rho \mathbf{\Lambda}_{i,k}}{M}\right). \quad (93)$$

For $E\left[I_{i,2}\right]$, we replace $\left\{\mathbf{\Lambda}_{i,k}\right\}$ with $\left\{\widetilde{\mathbf{\Lambda}}_{i,k}\right\}$. It can be checked that these expressions match with the expressions in this paper for the $M = 2$ and $M = 3$ settings. Nevertheless, it is important to note that the formulas in (89) and (93) are in terms of the eigenvalue matrices $\{\mathbf{\Lambda}_i, \widetilde{\mathbf{\Lambda}}_i, i = 1, \ldots, M\}$, which become harder (and impossible for $M \geq 5$) to compute in closed form as a function of the beamforming vectors and the transmit covariance matrices as M increases. Tractable approximations to the ergodic sum-rate and beamforming vector optimization based on such approximations are necessary, which is the subject of ongoing work.

VI. Numerical Studies

A. Utility of Generalized Eigenvectors in the Intermediate-SNR Regime

We now measure the performance of the candidate beamforming vectors proposed in (69) and (70) via some numerical results. In the first study, we consider a well-conditioned system (note that $\mathsf{Tr}(\mathbf{\Sigma}_1) = \mathsf{Tr}(\mathbf{\Sigma}_2) = M = 2$) with condition numbers $\chi_1 = 2.3333$ and $\chi_2 = 3$:

$$\mathbf{\Sigma}_1 = \begin{bmatrix} 1.3042 & 0.0543 - 0.2540i \\ 0.0543 + 0.2540i & 0.6958 \end{bmatrix}, \quad (94)$$

$$\mathbf{\Sigma}_2 = \begin{bmatrix} 1.1161 & -0.2195 + 0.4340i \\ -0.2195 - 0.4340i & 0.8839 \end{bmatrix}. \quad (95)$$

Fig. 2(a) shows the ergodic sum-rate as a function of ρ for four schemes. For the first scheme, for every ρ, an optimal choice $\{\alpha^\star(\rho), \beta^\star(\rho)\}$ is obtained from the search space $\alpha(\rho) \times \beta(\rho) \in [0, \infty) \times [0, \infty)$ as given in (76). The performance of the beamforming vectors with $\alpha^\star(\rho)$ and $\beta^\star(\rho)$ for every ρ is plotted along with the performance of the candidate obtained via a numerical (Monte Carlo) search over $\mathcal{G}(2,1) \times \mathcal{G}(2,1)$. Supporting Conjecture 1, we observe that the performance with $\{\alpha^\star(\rho), \beta^\star(\rho)\}$ is indeed optimal.

Further, the performance of a set of beamforming vectors with $\alpha(\rho) = \beta(\rho) = 0$ and $\alpha(\rho) = 100, \beta(\rho) = 15$ [fixed for all ρ in (69) and (70)] are also plotted. Observe that these two sets approximate the low- and the high-SNR solutions of Proposition 3 and Theorem 2, respectively.

In the second study, we consider an ill-conditioned system with condition numbers $\chi_1 = 28.4009$ and $\chi_2 = 38.2970$:

$$\mathbf{\Sigma}_1 = \begin{bmatrix} 1.7745 & -0.5178 + 0.0247i \\ -0.5178 - 0.0247i & 0.2255 \end{bmatrix}, \quad (96)$$

$$\mathbf{\Sigma}_2 = \begin{bmatrix} 1.2522 & -0.8739 - 0.2711i \\ -0.8739 + 0.2711i & 0.7478 \end{bmatrix}. \quad (97)$$

Fig. 2(b) plots the performance of the proposed scheme, the low- and the high-SNR solutions in addition to the candidate obtained via a numerical search over $\mathcal{G}(2,1) \times \mathcal{G}(2,1)$. As before, the low- and the high-SNR solutions are optimal in their respective extremes, while the candidate $\{\alpha^\star(\rho), \beta^\star(\rho)\}$ is optimal across all ρ.

Note that in Fig. 2(a), the high-SNR solution essentially coincides with the numerical search for all ρ, whereas at the low-SNR extreme, the performance of the low-SNR solution is as expected. This suggests that the high-SNR solution can be uniformly good across all ρ in some scenarios. In contrast, in Fig. 2(b), we see that there exists an SNR-regime where both the low- and the high-SNR solutions are suboptimal. We now explain why the high-SNR solution performs as well as $\{\alpha^\star(\rho), \beta^\star(\rho)\}$ for all ρ in the former setting. For this, we need to understand the behavior of the angle between the proposed set of beamforming vectors in (69) and (70) and the low-SNR solution as a function of α and β. In Fig. 3(a), we plot $\cos\left(\mathsf{Angle}_1(\alpha(\rho))\right)$ as a function of $\alpha(\rho)$ and $\cos\left(\mathsf{Angle}_2(\beta(\rho))\right)$ as a function of $\beta(\rho)$ where

$$\cos\left(\mathsf{Angle}_1(\alpha(\rho))\right) = \left|\left(\mathsf{Dom.eig.}\left(\left(\alpha(\rho)\mathbf{\Sigma}_2 + \mathbf{I}\right)^{-1}\mathbf{\Sigma}_1\right)\right)^H \mathsf{Dom.eig.}\left(\mathbf{\Sigma}_1\right)\right| \quad (98)$$

$$\cos\left(\mathsf{Angle}_2(\beta(\rho))\right) = \left|\left(\mathsf{Dom.eig.}\left(\left(\beta(\rho)\mathbf{\Sigma}_1 + \mathbf{I}\right)^{-1}\mathbf{\Sigma}_2\right)\right)^H \mathsf{Dom.eig.}\left(\mathbf{\Sigma}_2\right)\right|. \quad (99)$$

[13] More complicated expressions can be obtained in case $\{\mathbf{\Lambda}_{i,j}\}$ are not distinct. These expressions can be derived in a straightforward manner using the results in [31].

$$E\left[I_{i,1}\right] = \frac{\mathsf{I}_1}{\left(\mathbf{\Lambda}_{i,1} - \mathbf{\Lambda}_{i,2}\right)\left(\mathbf{\Lambda}_{i,1} - \mathbf{\Lambda}_{i,3}\right)} + \frac{\mathsf{I}_2}{\left(\mathbf{\Lambda}_{i,1} - \mathbf{\Lambda}_{i,3}\right)\left(\mathbf{\Lambda}_{i,2} - \mathbf{\Lambda}_{i,3}\right)} \quad (90)$$

$$\mathsf{I}_1 = \int_{x=0}^{\infty} x^2 e^{-x} \int_{y=0}^{\mathbf{\Lambda}_{i,1} - \mathbf{\Lambda}_{i,2}} y \log\left(1 + \frac{\rho}{3}\mathbf{\Lambda}_{i,1}x - \frac{\rho}{3}xy\right) dydx \quad (91)$$

$$\mathsf{I}_2 = \int_{x=0}^{\infty} x^2 e^{-x} \int_{y=0}^{\mathbf{\Lambda}_{i,2} - \mathbf{\Lambda}_{i,3}} y \log\left(1 + \frac{\rho}{3}\mathbf{\Lambda}_{i,3}x + \frac{\rho}{3}xy\right) dydx \quad (92)$$

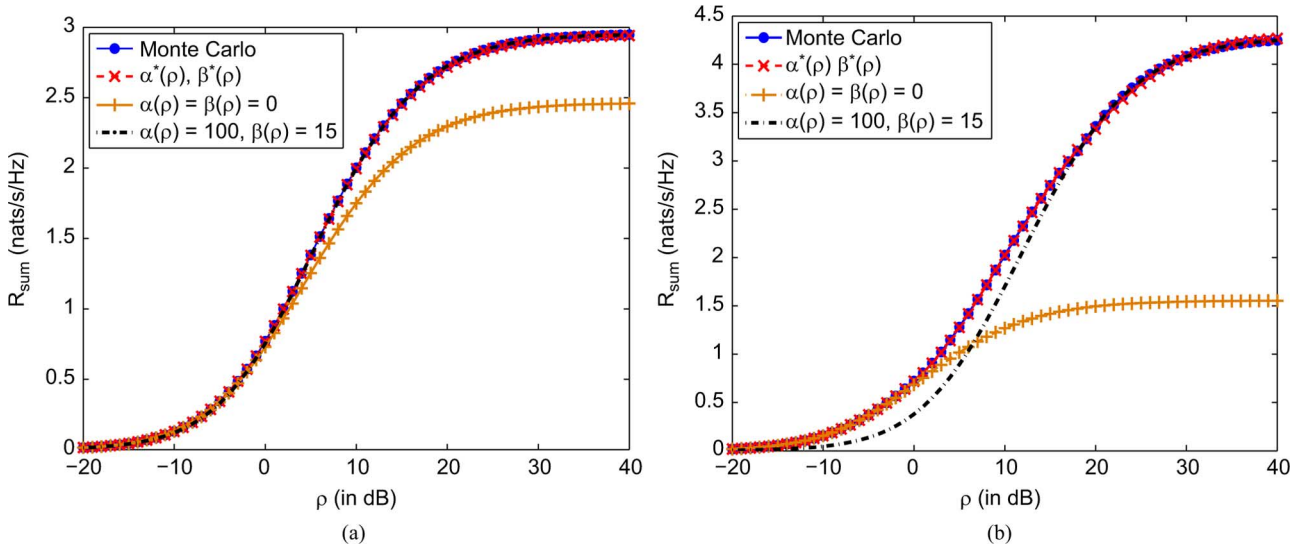

Fig. 2. Performance of the proposed scheme with $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$: (a) as in (94) and (95), (b) as in (96) and (97).

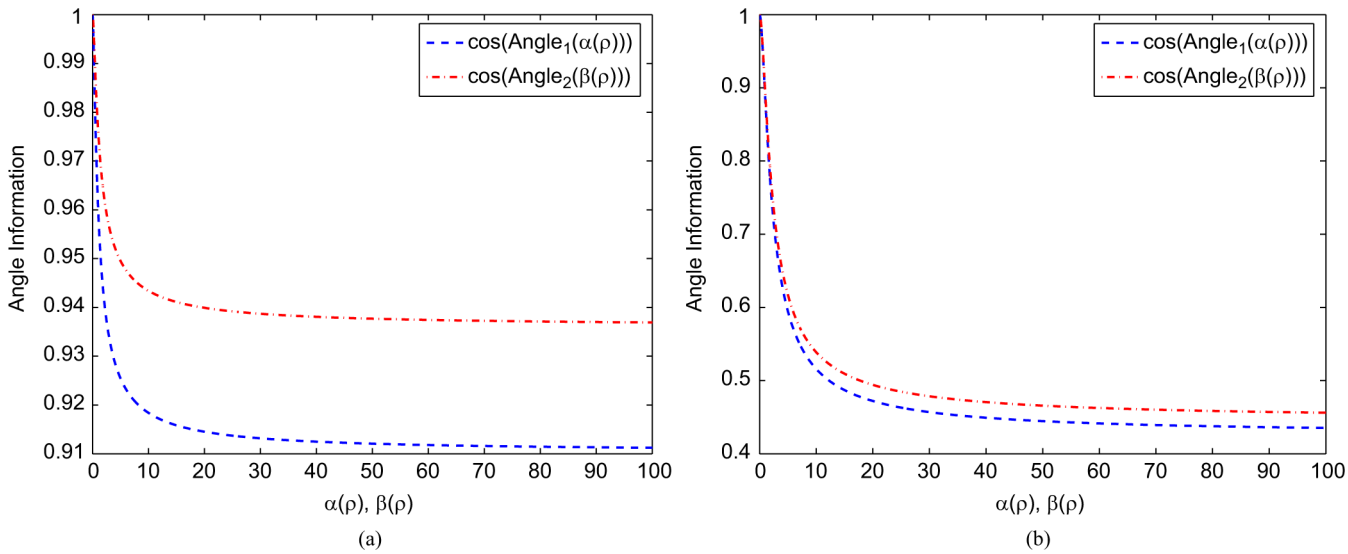

Fig. 3. Angle_1 and Angle_2, defined in (98) and (99), as a function of α and β for the setting in: (a) (94) and (95), (b) (96) and (97).

From Fig. 3(a), we note that the chordal distance[14] between the low- and the high-SNR solutions is small (on the order of 0.05). Also, observe that there is a quick convergence of (69) and (70) as α (or β) increases to the high-SNR solution and hence the high-SNR solution is a good approximation to the choice $\{\alpha^\star(\rho), \beta^\star(\rho)\}$ over a large SNR range. On the other hand, from Fig. 3(b), we see that the chordal distance between the low- and the high-SNR solutions is large (on the order of 0.90), which translates to the suboptimality gap in Fig. 2(b).

We now study the performance of the proposed beamforming vectors in (81) and (82) for the weighted ergodic sum-rate of the system with $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ as in (96) and (97). Two sets of weights are considered: i) $\zeta_1 = 1, \zeta_2 = 0.5$ and ii) $\zeta_1 = 0.2, \zeta_2 = 0.8$. Fig. 4 plots the performance of two schemes. The first scheme corresponds to a Monte Carlo search over $\{\mathcal{G}(2,1)\}^2$, whereas the second scheme corresponds to the use of an optimal choice $\{\alpha^\star(\rho), \beta^\star(\rho), \gamma^\star(\rho), \delta^\star(\rho)\}$ (for every ρ) as in (85). As can be seen from Fig. 4, the proposed scheme in (81) and (82) performs as well as the Monte Carlo search for both sets of weights. Numerical studies suggest that similar performance is seen across all possible $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$, and all possible weights ζ_1 and ζ_2.

[14]In short, the chordal distance is the square-root of the difference of 1 and the square of the quantity computed in (98) [or (99)]. See (171) in Appendix C for more details.

B. Utility of Statistical Beamforming in Practice

Interference limitation ensures that the degrees-of-freedom (DoFs) realized by the multiuser statistical beamforming scheme is zero unlike the case with a scheme where the DoFs are time-shared across the two users (single-user statistical beamforming). Thus, the multiuser scheme is suboptimal relative to the single-user scheme eventually, as ρ increases. So it could be asked: why is statistical multiuser beamforming of interest in practice? The most important answer to this question is that our work on statistical beamforming is but a first step on a path to utilizing limited channel state feedback. This program of research is interested in designing precoders when the number of feedback bits (B) per channel state is small, and the work in this paper considers the extreme case of $B = 0$. However, the numerical results that we now present will show

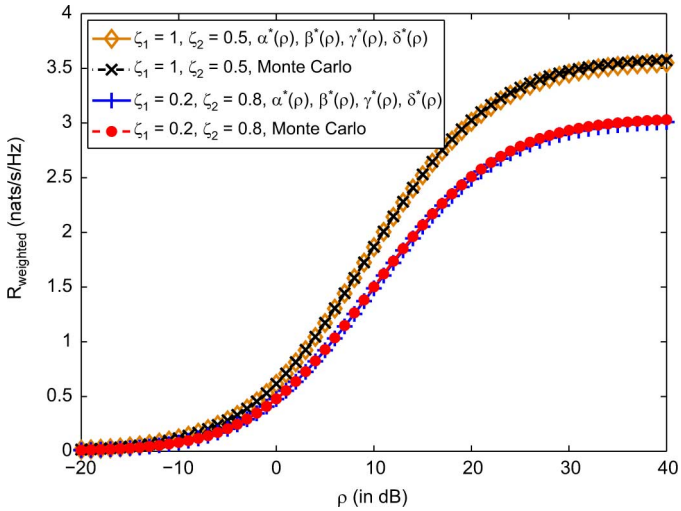

Fig. 4. Weighted ergodic sum-rate of the proposed scheme with $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ as in (96).

that even when $B = 0$, there can still be a benefit in multiuser beamforming over single-user techniques.

For this comparison, we start with the single-user scheme operating over $M = 2$ time-slots (with $M = 2$ antennas) where time-slot 1 and 2 are meant exclusively for users 1 and 2, respectively. The ergodic sum-rate achievable with this scheme is given by

$$\mathcal{R}_{\mathsf{sum}}\Big|_{\mathsf{su}} = \sum_{i=1}^{M=2} E\left[\log\left(1+\mathbf{h}_i^H\mathbf{Q}_i\mathbf{h}_i\right)\right] \tag{100}$$

where $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are statistically optimally chosen (see [1]–[6] for details) input covariance matrices with $\mathsf{Tr}(\mathbf{Q}_1) = \rho = \mathsf{Tr}(\mathbf{Q}_2)$. The special case of uniform power allocation is realized by $\mathbf{Q}_i = \frac{\rho}{2}\mathbf{I}$. For the multiuser scheme, the ergodic sum-rate is given by

$$\mathcal{R}_{\mathsf{sum}}\Big|_{\mathsf{mu}} = 2\sum_{i=1}^{M=2} E\left[\log\left(1+\frac{\frac{\rho}{2}|\mathbf{h}_i^H\mathbf{w}_i|^2}{1+\frac{\rho}{2}|\mathbf{h}_i^H\mathbf{w}_j|^2}\right)\right],\ j\neq i \tag{101}$$

where we use the high-SNR optimizing beamforming structure from Theorem 2 for $\mathbf{w}_1$ and $\mathbf{w}_2$. Note that a prelog factor of 2 comes from using both time-slots for simultaneous transmission to both users.

In Fig. 5(a) and (b), we illustrate the ergodic rate (for user 1) with the single-user (with both the optimal and uniform power allocation strategies) and multiuser schemes as a function of ρ for the following two scenarios:

$$\mathsf{I}\quad :\mathbf{\Sigma}_1 = \begin{bmatrix} 0.5942 & 0.4471-0.5601i \\ 0.4471+0.5601i & 1.4058 \end{bmatrix} \tag{102}$$

$$\mathbf{\Sigma}_2 = \begin{bmatrix} 1.2886 & -0.1173-0.2180i \\ -0.1173+0.2180i & 0.7114 \end{bmatrix} \tag{103}$$

$$\mathsf{II}:\quad \mathbf{\Sigma}_1 = \begin{bmatrix} 0.1193 & 0.3613+0.0990i \\ 0.3613-0.0990i & 1.8807 \end{bmatrix} \tag{104}$$

$$\mathbf{\Sigma}_2 = \begin{bmatrix} 1.7436 & 0.2319-0.1855i \\ 0.2319+0.1855i & 0.2564 \end{bmatrix}. \tag{105}$$

Note that the scenarios considered above correspond to the following condition numbers:

$$\mathsf{I}: \chi_1 = \frac{1.8236}{0.1764} = 10.3366 \approx 10.14\ \mathsf{dB},$$
$$\chi_2 = \frac{1.3802}{0.6198} = 2.2270 \approx 3.48\ \mathsf{dB}, \tag{106}$$
$$\mathsf{II}: \chi_1 = \frac{1.9571}{0.0429} = 45.5808 \approx 16.59\ \mathsf{dB},$$
$$\chi_2 = \frac{1.8007}{0.1993} = 9.0354 \approx 9.56\ \mathsf{dB}. \tag{107}$$

While there has been little work on the typical distribution(s) of condition number of channels encountered in the Third Generation Partnership Project for Long Term Evolution (3GPP-LTE) or LTE-Advanced standards, initial indications suggest that the examples considered here are representative of real systems [47, Fig. 3]. Clearly, the multiuser scheme is better than the single-user scheme for all $\rho < 18$ dB in Scenario-I [see Fig. 5(a)], whereas it is better for all $\rho < 25$ dB (and even beyond this SNR) in Scenario-II [see Fig. 5(b)].

To understand this aspect further, we note from Proposition 3 and [5] that

$$\lim_{\rho\to 0}\frac{\mathcal{R}_{\mathsf{sum}}\big|_{\mathsf{su}}}{\rho} = \lim_{\rho\to 0}\frac{\mathcal{R}_{\mathsf{sum}}\big|_{\mathsf{mu}}}{\rho} = \lambda_1(\mathbf{\Sigma}_1)+\lambda_1(\mathbf{\Sigma}_2). \tag{108}$$

Thus, we can define ρ_{cutoff} as the smallest SNR above which the single-user scheme dominates the multiuser scheme (similar ideas have been explored in [45]).

We now consider a family of channels for users 1 and 2 with fixed unitary matrices

$$\mathbf{U} = \begin{bmatrix} -0.0975+0.4312i & -0.8906-0.1069i \\ 0.4765-0.7600i & -0.3967-0.1951i \end{bmatrix} \tag{109}$$

$$\widetilde{\mathbf{U}} = \begin{bmatrix} -0.7955+0.4980i & 0.2849-0.1949i \\ -0.2376-0.2505i & -0.6731-0.6541i \end{bmatrix}, \tag{110}$$

where $\mathbf{U}$ and $\widetilde{\mathbf{U}}$ are as in (39) and (40). We study the behavior of ρ_{cutoff} for this family as a function of χ_1 and χ_2 in Fig. 5(c). With $\lambda_1(\mathbf{\Sigma}_2) \geq 1.1$ or equivalently, $\chi_2 = \frac{\lambda_1(\mathbf{\Sigma}_2)}{\lambda_2(\mathbf{\Sigma}_2)} = \frac{\lambda_1(\mathbf{\Sigma}_2)}{M-\lambda_1(\mathbf{\Sigma}_2)} \geq 1.2222$, it can be seen that the multiuser scheme performs better than the single-user scheme for $\rho < 5$ dB and $\rho < 10$ dB provided $\chi_1 > 3$ and $\chi_1 > 7$, respectively. In general, it can be seen that ρ_{cutoff} increases as either χ_1 or χ_2 increases, thus suggesting a large SNR range where the multiuser scheme is useful in practice.

In Fig. 5(d), we plot the ergodic sum-rate ratio, defined as,

$$\mathsf{Ergodic\ Sum\ Rate\ Ratio} \triangleq \frac{\mathcal{R}_{\mathsf{sum}}\big|_{\mathsf{mu}}}{\mathcal{R}_{\mathsf{sum}}\big|_{\mathsf{su}}} \tag{111}$$

as a function of χ_1 with $\lambda_1(\mathbf{\Sigma}_2) = 1.3$ (or equivalently, $\chi_2 = \frac{\lambda_1(\mathbf{\Sigma}_2)}{\lambda_2(\mathbf{\Sigma}_2)} = 1.8571$) for different practically observed values of SNR in the range $[-5, 5]$ dB. As can be seen from this figure, the improvement of the multiuser scheme relative to the single-user scheme increases as either ρ decreases or χ_1 increases. The multiuser scheme results in an at least 20% improvement in performance provided that $\chi_1 > 6$. With ρ smaller than 0 dB and $\chi_1 > 10$, the multiuser scheme results in an improvement of

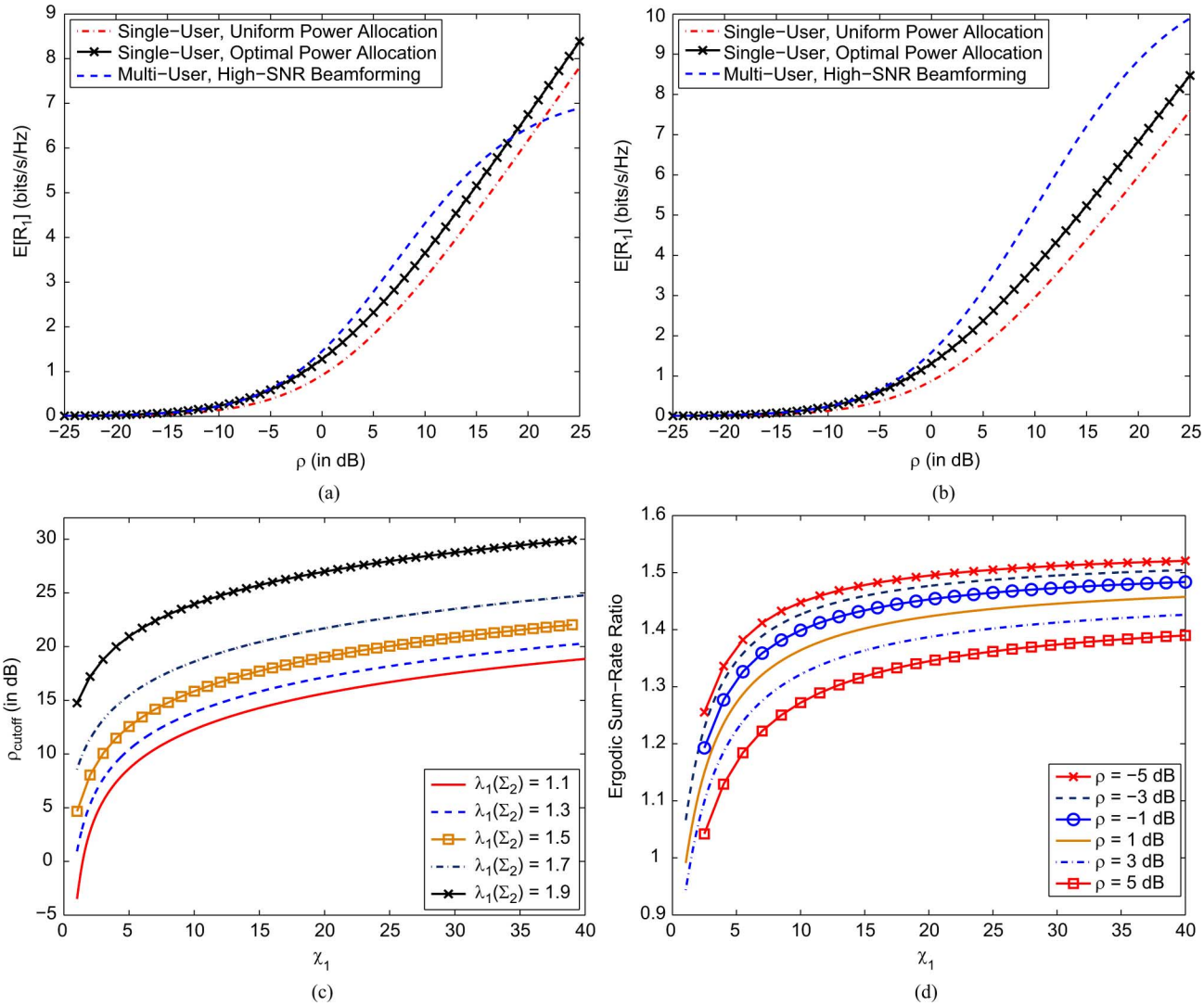

Fig. 5. (a) and (b) Ergodic rate of user 1 as a function of ρ for single- and multiuser schemes in Scenarios-I and II. (c) and (d) ρ_{cutoff} and ergodic sum-rate ratio for a family of channels (for users 1 and 2) as a function of χ_1 and χ_2.

over 40%. These four numerical studies illustrate that while the single-user scheme eventually dominates as $\rho \to \infty$, the proposed statistical beamforming scheme is better over a significant range of SNR for many practical scenarios.

While the primary attention of this paper is restricted to the $M = 2$ case, we now demonstrate that the SNR range, $[0, \rho_{\text{cutoff}}]$, over which statistical multiuser beamforming is better than single-user statistical beamforming, does not become insignificant as M increases. For this, we upper bound the ergodic sum-rate of the single-user scheme as

$$\mathcal{R}_{\text{sum}}\Big|_{\text{su}} = \sum_{i=1}^{M} E\left[\log\left(1 + \text{Tr}\left(\mathbf{Q}_i \mathbf{h}_i \mathbf{h}_i^H\right)\right)\right] \quad (112)$$

$$\leq \sum_{i=1}^{M} \log\left(1 + E\left[\text{Tr}\left(\mathbf{Q}_i \mathbf{h}_i \mathbf{h}_i^H\right)\right]\right) \quad (113)$$

$$= \sum_{i=1}^{M} \log\left(1 + \text{Tr}\left(\mathbf{Q}_i \mathbf{\Sigma}_i\right)\right) \quad (114)$$

$$\leq \sum_{i=1}^{M} \log\left(1 + \rho \lambda_1(\mathbf{\Sigma}_i)\right) \quad (115)$$

where the second step follows from Jensen's inequality and the final step from beamforming all the power along the dominant eigenmode.

Using the fact that $|\mathbf{h}_i^H \mathbf{w}_j|^2 \leq \|\mathbf{h}_i\|^2 \cdot \|\mathbf{w}_j\|^2 = \|\mathbf{h}_i\|^2$, we can upper bound the interference in the multiuser scheme as

$$1 + \frac{\rho}{M} \sum_{j \neq i} |\mathbf{h}_i^H \mathbf{w}_j|^2 \leq 1 + \frac{\rho}{M} \cdot (M-1) \cdot \|\mathbf{h}_i\|^2 \quad (116)$$

$$= \frac{\rho}{M} \cdot \|\mathbf{h}_i\|^2 \cdot (M - 1 + C) \quad (117)$$

where $C = \frac{M}{\rho \|\mathbf{h}_i\|^2}$. Thus, the sum-rate is lower bounded as

$$\mathcal{R}_{\text{sum}}\Big|_{\text{mu}} \geq M \sum_{i=1}^{M} E\left[\log\left(1 + \frac{\frac{\rho}{M} \cdot |\mathbf{h}_i^H \mathbf{w}_i|^2}{\frac{\rho}{M} \|\mathbf{h}_i\|^2 (M - 1 + C)}\right)\right] \quad (118)$$

$$= M \sum_{i=1}^{M} E\left[\log\left(1 + \frac{\cos^2(\theta_i)}{M - 1 + C}\right)\right] \quad (119)$$

where θ_i is the angle between $\widehat{\mathbf{h}}_i = \frac{\mathbf{h}_i}{\|\mathbf{h}_i\|}$ and $\mathbf{w}_i$. We use the fact that $\log(1+x) \to x$ as $x \to 0$ to bound $\mathcal{R}_{\mathsf{sum}}\Big|_{\mathsf{mu}}$ as

$$\mathcal{R}_{\mathsf{sum}}\Big|_{\mathsf{mu}} \geq M \sum_{i=1}^{M} E\left[\frac{\cos^2(\theta_i)}{M-1+C}\right] \tag{120}$$

$$\xrightarrow{M\to\infty} \frac{M}{M-1+1/\rho} \cdot \sum_{i=1}^{M} E\left[\cos^2(\theta_i)\right] \tag{121}$$

where $\mathsf{Tr}(\mathbf{\Sigma}_i) = M$ for all the users implies that

$$C \xrightarrow{M\to\infty} \frac{M}{\rho \mathsf{Tr}(\mathbf{\Sigma}_i)} = \frac{1}{\rho}. \tag{122}$$

A lower bound to ρ_{cutoff} is obtained by $\widetilde{\rho}$ that satisfies

$$\sum_{i=1}^{M} \log\left(1+\widetilde{\rho}\lambda_1(\mathbf{\Sigma}_i)\right) = \frac{M}{M-1+1/\widetilde{\rho}} \cdot \sum_{i=1}^{M} E\left[\cos^2(\theta_i)\right]. \tag{123}$$

While this requires an involved computation, a simpler bound for ρ_{cutoff} is

$$\rho_{\mathsf{cutoff}} \geq \widetilde{\rho} \geq \widehat{\rho} \text{ where}$$

$$\log\left(\widehat{\rho}\right) = \frac{1}{M}\sum_{i=1}^{M}\left[E\left[\cos^2(\theta_i)\right] - \log\left(\lambda_1\left(\mathbf{\Sigma}_i\right)\right)\right] \tag{124}$$

$$= \mathcal{O}(1) \text{ as } M \to \infty \tag{125}$$

under the assumption that

$$\sup_M \max_i \lambda_1(\mathbf{\Sigma}_i) = \mathcal{O}(1). \tag{126}$$

It is important to emphasize that given the complexities of the considered optimization, power control for the users has been ignored in this study. On the one hand, equal power allocation is popular in current-generation cellular standards where low-complexity schemes are preferred. For example, in the 3GPP-LTE standard, in contrast to slow power-control for the uplink, no power-control mechanism is specified for the downlink and hence, equal power allocation is implicitly assumed (see [48], [49, p. 119], and [50, Ch. 4.5.3] for details). On the other hand, the performance of statistical beamforming could seriously deteriorate without power control, especially in the high-SINR extreme. However, as seen above, a single-user scheme could be more useful in this regime than the multiuser scheme. Nevertheless, it is of interest in understanding the possibilities with power control and this will be the subject of future work.

VII. Concluding Remarks

This paper considered the design of statistical beamforming vectors in a MISO broadcast setting to maximize the ergodic sum-rate. The approach pursued here for the simplest nontrivial problem with two users is as follows: first, the beamforming vectors are fixed and ergodic rate expressions are computed in closed form in terms of the covariance matrices of the links and the beamforming vectors. The optimization of this nonconvex function results in a generalized eigenvector structure for the optimal beamforming vectors, the solution to maximizing an appropriately defined SINR metric for each user. This structure generalizes the single-user setup where the dominant eigenmodes of the transmit covariance matrix of the links are excited. The main results of this paper are presented in Table I for different SNR (ρ) assumptions where we use $\mathbf{u}_1(\bullet)$ and $\mathbf{u}_2(\bullet)$ to denote the dominant and subdominant eigenvectors of the matrix under consideration.

Possible extensions of this work include unifying the special case of Theorem 3 with the general case of Theorem 2, and proving Conjectures 1 and 2. Developing intuition in the three-user case as well as tractable approximations in the general M-user ($M > 2$) case critically depend on exploiting the functional structure of the ergodic sum-rate expression. The generalized eigenvector solution has been seen in other multiuser scenarios as well, for example, the interference channel problem with two antennas [36]. Generalizing the theme developed in the broadcast setting to the interference channel setting, the Rayleigh case to the Ricean case, and the perfect CSI case to the case where only statistical information is available are all important tasks.

Appendix

A) Density Function of Weighted-Norm of Isotropically Distributed Unit-Norm Vectors: Toward computing $E[I_{i,1}]$, we generalize the technique expounded in [51] where the surface area (that is required) to be computed is treated as a differential element of a corresponding solid volume (at a specific radius value), and the volume of the necessary solid object is calculated using tools from higher dimensional integration (geometry). In this direction, we have

$$p_i(x)dx \triangleq \mathsf{P}\left(\widehat{\mathbf{h}}_{\mathsf{iid},i}^H \mathbf{\Lambda}_i \widehat{\mathbf{h}}_{\mathsf{iid},i} \in [x, x+dx]\right) \tag{127}$$

$$p_i(x) = \frac{\partial}{\partial x}\mathsf{P}\left(\widehat{\mathbf{h}}_{\mathsf{iid},i}^H \mathbf{\Lambda}_i \widehat{\mathbf{h}}_{\mathsf{iid},i} \leq x\right) \tag{128}$$

with

$$\mathsf{P}\left(\widehat{\mathbf{h}}_{\mathsf{iid},i}^H \mathbf{\Lambda}_i \widehat{\mathbf{h}}_{\mathsf{iid},i} \leq x\right) = 1 - \frac{\mathsf{Area}\,(x,1)}{\mathsf{Area}\,(1)} \tag{129}$$

where

$$\mathsf{Area}\,(x,y) \triangleq \mathsf{Area}\left(\widehat{\mathbf{h}}_{\mathsf{iid},i}^H \mathbf{\Lambda}_i \widehat{\mathbf{h}}_{\mathsf{iid},i} \geq x, \|\widehat{\mathbf{h}}_{\mathsf{iid},i}\|^2 = y\right) \tag{130}$$

$$\mathsf{Area}\,(y) \triangleq \mathsf{Area}\left(\|\widehat{\mathbf{h}}_{\mathsf{iid},i}\|^2 = y\right) \tag{131}$$

denote the area of a (unit radius) spherical cap carved out by the ellipsoid

$$\left\{\widehat{\mathbf{h}}_{\mathsf{iid},i} : \widehat{\mathbf{h}}_{\mathsf{iid},i}^H \mathbf{\Lambda}_i \widehat{\mathbf{h}}_{\mathsf{iid},i} = x\right\} \tag{132}$$

and the area of a (unit radius) complex sphere, respectively. The volume of the objects desired in the computation of $p_i(x)$ are

$$\mathsf{Vol}\left(x, r^2\right) \triangleq \mathsf{Vol}\left(\widehat{\mathbf{h}}_{\mathsf{iid},i}^H \mathbf{\Lambda}_i \widehat{\mathbf{h}}_{\mathsf{iid},i} \geq x, \|\widehat{\mathbf{h}}_{\mathsf{iid},i}\|^2 \leq r^2\right) \quad (133)$$

$$= \int_{y=0}^{r^2} \mathsf{Area}\,(x, y)\, dy, \text{ and} \quad (134)$$

$$\mathsf{Vol}(r^2) \triangleq \mathsf{Vol}\left(\|\widehat{\mathbf{h}}_{\mathsf{iid},i}\|^2 \leq r^2\right) = \int_{x=0}^{r^2} \mathsf{Area}(x) dx. \quad (135)$$

Thus, we have

$$\mathsf{Area}\,(x, 1) = \frac{\partial}{\partial r^2} \mathsf{Vol}\left(x, r^2\right)\Big|_{r=1}, \quad (136)$$

$$\mathsf{Area}\,(x) = \frac{\partial}{\partial r^2} \mathsf{Vol}(r^2)\Big|_{r=1}, \text{ and hence,} \quad (137)$$

$$p_i(x) = -\frac{\frac{\partial^2}{\partial x r^2} \mathsf{Vol}\left(x, r^2\right)\Big|_{r=1}}{\frac{\partial}{\partial r^2} \mathsf{Vol}\left(r^2\right)\Big|_{r=1}}. \quad (138)$$

It is important to realize that computing $\mathsf{Vol}\left(x, r^2\right)$ is *non-trivial* even in the simplest case of $M = 2$. This is because every additional dimension to the complex ellipsoid corresponds to addition of two real dimensions, thus rendering a geometric visualization impossible. For example, with $M = 2$, we have the intersection of two 4-D real objects. Nevertheless, the following lemma captures the complete structure of $p_i(x)$ when $M = 2$.

Lemma 1: If $M = 2$, the random variable $\widehat{\mathbf{h}}_{\mathsf{iid},i}^H \mathbf{\Lambda}_i \widehat{\mathbf{h}}_{\mathsf{iid},i}$ is uniformly distributed in the interval $[\mathbf{\Lambda}_{i,2}, \mathbf{\Lambda}_{i,1}]$.

Proof: First, note that it follows from [51, Lemma 2] that

$$\mathsf{Vol}(r^2) = \frac{\pi^M r^{2M}}{M!}. \quad (139)$$

For computing $\mathsf{Vol}\left(x, r^2\right)$, we follow the same variable transformation as in [51]. We set $\widehat{\mathbf{h}}_{\mathsf{iid},i}(k) = r_k \exp(j\phi_k)$ for $k = 1, 2$. The ellipsoid is contained completely in the sphere of radius r if r is such that $r \geq \sqrt{\frac{x}{\mathbf{\Lambda}_{i,2}}}$, whereas the sphere is contained completely in the ellipsoid if $r \leq \sqrt{\frac{x}{\mathbf{\Lambda}_{i,1}}}$. In the intermediate regime for r, a nontrivial intersection between the two objects is observed and one can compute the volume by performing a 2-D integration as follows:

$$\mathsf{Vol}\left(x, r^2\right) = \int\int_{\mathcal{A}} r_1 r_2 \phi_1 \phi_2 dr_1 dr_2 d\phi_1 d\phi_2 \quad (140)$$

$$= (2\pi)^2 \cdot \int\int_{\mathcal{B}} r_1 dr_1 r_2 dr_2 \quad (141)$$

$$= (2\pi)^2 \cdot \int_0^{U_2} r_2 dr_2 \int_{L_1}^{U_1} r_1 dr_1 \quad (142)$$

where

$$\mathcal{A} = \left\{r_1, r_2 : r_1^2 \mathbf{\Lambda}_{i,1} + r_2^2 \mathbf{\Lambda}_{i,2} \geq x,\ r_1^2 + r_2^2 \leq r^2\right\} \text{ and } \left\{\phi_1, \phi_2 : [0, 2\pi)\right\} \quad (143)$$

$$\mathcal{B} = \left\{r_1, r_2 : r_1^2 \mathbf{\Lambda}_{i,1} + r_2^2 \mathbf{\Lambda}_{i,2} \geq x,\ r_1^2 + r_2^2 \leq r^2\right\} \quad (144)$$

$$L_1 = \sqrt{\frac{x - r_2^2 \mathbf{\Lambda}_{i,2}}{\mathbf{\Lambda}_{i,1}}},\ U_1 = \sqrt{r^2 - r_2^2},\ U_2 = \frac{r^2 \mathbf{\Lambda}_{i,1} - x}{\mathbf{\Lambda}_{i,1} - \mathbf{\Lambda}_{i,2}}. \quad (145)$$

Trivial computation establishes the following:

$$\mathsf{Vol}\left(x, r^2\right) = \begin{cases} 0, & r \leq \sqrt{\frac{x}{\mathbf{\Lambda}_{i,1}}} \\ \frac{\pi^2}{2} \cdot \frac{\left(r^2 \mathbf{\Lambda}_{i,1} - x\right)^2}{\mathbf{\Lambda}_{i,1} \cdot (\mathbf{\Lambda}_{i,1} - \mathbf{\Lambda}_{i,2})}, & \sqrt{\frac{x}{\mathbf{\Lambda}_{i,1}}} \leq r \leq \sqrt{\frac{x}{\mathbf{\Lambda}_{i,2}}} \\ \frac{\pi^2}{2} \cdot \left(r^4 - \frac{x^2}{\mathbf{\Lambda}_{i,1} \mathbf{\Lambda}_{i,2}}\right), & r \geq \sqrt{\frac{x}{\mathbf{\Lambda}_{i,2}}}. \end{cases} \quad (146)$$

Another trivial computation using (138) results in

$$p_i(x) = \frac{1}{\mathbf{\Lambda}_{i,1} - \mathbf{\Lambda}_{i,2}}. \quad (147)$$

That is, $\widehat{\mathbf{h}}_{\mathsf{iid},i}^H \mathbf{\Lambda}_i \widehat{\mathbf{h}}_{\mathsf{iid},i}$ is uniformly distributed in its range. ■

The structure of $p_i(x)$ gets more complicated as M increases. We now provide its structure in the $M = 3$ and $M = 4$ cases without proof.

Lemma 2: With $M = 3$, the density function $p_i(x)$ is of the form:

$$p_i(x) = \begin{cases} 0, & x \leq \mathbf{\Lambda}_{i,3} \\ \frac{2(x - \mathbf{\Lambda}_{i,3})}{(\mathbf{\Lambda}_{i,1} - \mathbf{\Lambda}_{i,3})(\mathbf{\Lambda}_{i,2} - \mathbf{\Lambda}_{i,3})}, & \mathbf{\Lambda}_{i,3} \leq x \leq \mathbf{\Lambda}_{i,2} \\ \frac{2(\mathbf{\Lambda}_{i,1} - x)}{(\mathbf{\Lambda}_{i,1} - \mathbf{\Lambda}_{i,2})(\mathbf{\Lambda}_{i,1} - \mathbf{\Lambda}_{i,3})}, & \mathbf{\Lambda}_{i,2} \leq x \leq \mathbf{\Lambda}_{i,1} \\ 0, & x \geq \mathbf{\Lambda}_{i,1}. \end{cases} \quad (148)$$

With $M = 4$, the density function $p_i(x)$ takes the form:

$$p_i(x) = \begin{cases} 0, & x \leq \mathbf{\Lambda}_{i,4} \\ \frac{3(x - \mathbf{\Lambda}_{i,4})^2 / (\mathbf{\Lambda}_{i,3} - \mathbf{\Lambda}_{i,4})}{(\mathbf{\Lambda}_{i,1} - \mathbf{\Lambda}_{i,4})(\mathbf{\Lambda}_{i,2} - \mathbf{\Lambda}_{i,4})}, & \mathbf{\Lambda}_{i,4} \leq x \leq \mathbf{\Lambda}_{i,3} \\ \frac{3 \cdot \mathcal{L}_0}{(\mathbf{\Lambda}_{i,1} - \mathbf{\Lambda}_{i,3})(\mathbf{\Lambda}_{i,2} - \mathbf{\Lambda}_{i,4})}, & \mathbf{\Lambda}_{i,3} \leq x \leq \mathbf{\Lambda}_{i,2} \\ \frac{3(\mathbf{\Lambda}_{i,1} - x)^2 / (\mathbf{\Lambda}_{i,1} - \mathbf{\Lambda}_{i,4})}{(\mathbf{\Lambda}_{i,1} - \mathbf{\Lambda}_{i,2})(\mathbf{\Lambda}_{i,1} - \mathbf{\Lambda}_{i,3})}, & \mathbf{\Lambda}_{i,2} \leq x \leq \mathbf{\Lambda}_{i,1} \\ 0, & x \geq \mathbf{\Lambda}_{i,1} \end{cases} \quad (149)$$

where

$$\mathcal{L}_0 = \frac{(x - \mathbf{\Lambda}_{i,3})(\mathbf{\Lambda}_{i,2} - x)}{\mathbf{\Lambda}_{i,2} - \mathbf{\Lambda}_{i,3}} + \frac{(x - \mathbf{\Lambda}_{i,4})(\mathbf{\Lambda}_{i,1} - x)}{\mathbf{\Lambda}_{i,1} - \mathbf{\Lambda}_{i,4}}. \quad (150)$$

■

Fig. 6 illustrates the trends of the cumulative distribution function (CDF) by plotting the fit between the theoretical expressions in Lemmas 1 and 2, and the CDF estimated by Monte Carlo methods. The cases considered are: a) $\mathbf{\Lambda}_i = \mathsf{diag}([2\,1])$ for $M = 2$, b) $\mathbf{\Lambda}_i = \mathsf{diag}([3\,2\,1])$ for $M = 3$, and c) $\mathbf{\Lambda}_i = \mathsf{diag}([4\,3\,2\,1])$ for $M = 4$. The figure shows the excellent match between theory and Monte Carlo estimates.

*B) Low-*SNR *"Converse":* From [8], the sum-rate achieved by the DPC scheme with a given channel realization pair $\{\mathbf{h}_1, \mathbf{h}_2\}$ is, as in (151) and (152) at the bottom

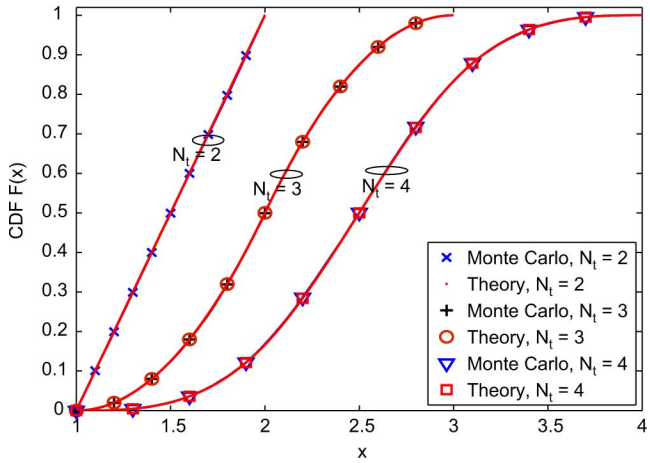

Fig. 6. CDF of weighted-norm of isotropically distributed beamforming vectors.

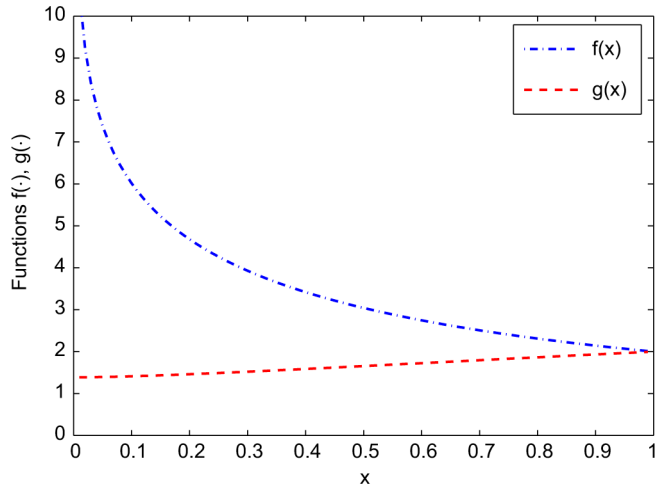

Fig. 7. Behavior of $f(x)$ and $g(x)$.

of the page, corresponding to a power allocation $\mathcal{P}$ of ρ_1 to user 1 and ρ_2 to user 2 with $\rho_1 + \rho_2 = \rho$ and $\Pi = [\pi_1, \pi_2]$ is a user permutation/ordering. If we constrain the DPC scheme to have a fixed power allocation $\rho_1 = \rho_2 = \frac{\rho}{2}$, the sum-rate achievable reduces to

$$\mathcal{R}_{\mathsf{sum,dpc}}\Big|_{\text{eq. pow.}} = \max_{\Pi} \left\{ \log\left(1 + \frac{\rho}{2}|\mathbf{h}_{\pi_1}^H \mathbf{w}_{\pi_1}|^2 + \frac{\rho}{2}|\mathbf{h}_{\pi_2}^H \mathbf{w}_{\pi_2}|^2\right) \right\} \tag{153}$$

$$= \log\left(1 + \frac{\rho}{2}|\mathbf{h}_1^H \mathbf{w}_1|^2 + \frac{\rho}{2}|\mathbf{h}_2^H \mathbf{w}_2|^2\right). \tag{154}$$

It is straightforward to check that the ergodic sum-rate of this scheme satisfies

$$\lim_{\rho\to 0} \frac{E\left[\mathcal{R}_{\mathsf{sum,dpc}}|_{\text{eq. pow.}}\right]}{\rho} = \frac{1}{2} \cdot \left[\mathbf{E}\left[|\mathbf{h}_1^H \mathbf{w}_1|^2 + |\mathbf{h}_2^H \mathbf{w}_2|^2\right]\right] \tag{155}$$

$$= \frac{1}{2} \cdot \left[\mathbf{w}_1^H \mathbf{\Sigma}_1 \mathbf{w}_1 + \mathbf{w}_2^H \mathbf{\Sigma}_2 \mathbf{w}_2\right] \tag{156}$$

$$\leq \frac{1}{2} \cdot \left[\lambda_1(\mathbf{\Sigma}_1) + \lambda_1(\mathbf{\Sigma}_2)\right]. \tag{157}$$

The upper bound is the same as that of the linear beamforming scheme in the low-SNR extreme and is achieved with the same structure. ■

C) Rewriting the Rate Expression in the High-SNR Extreme: A straightforward exercise shows that (36) can be rewritten as in (158) as follows:

$$E\left[R_i\right] \xrightarrow{\rho\to\infty} \frac{1}{2} g\left(d_{\mathbf{\Sigma}_i}(\mathbf{w}_1, \mathbf{w}_2)\right) + \log\left(1 + \frac{A_i}{B_i}\right) - \log(2) \tag{158}$$

where $g(\bullet)$ is a function defined as

$$g(z) \triangleq f(z) + 2\log(z), \tag{159}$$

$$f(z) = \frac{1}{\sqrt{1-z^2}} \log\left(\frac{1+\sqrt{1-z^2}}{1-\sqrt{1-z^2}}\right), \; 0 < z < 1. \tag{160}$$

In (158), $d_{\mathbf{\Sigma}_i}(\mathbf{w}_1, \mathbf{w}_2)$ is defined as

$$d_{\mathbf{\Sigma}_i}(\mathbf{w}_1, \mathbf{w}_2) \triangleq \sqrt{\frac{4\left(A_i B_i - C_i^2\right)}{\left(A_i + B_i\right)^2}} \tag{161}$$

with A_i, B_i, and C_i as in Theorem 1. As illustrated in Fig. 7, $f(\bullet)$ is monotonically decreasing as a function of its argument and $g(\bullet)$ is increasing with

$$2\log(2) = \lim_{z\to 0} g(z) \leq g(z) \leq \lim_{z\to 1} g(z) = 2 \tag{162}$$

$$\infty = \lim_{z\to 0} f(z) \geq f(z) \geq \lim_{z\to 1} f(z) = 2. \tag{163}$$

Formal proofs of these facts are provided in Appendix D next. Some properties of $d_{\mathbf{\Sigma}_i}(\mathbf{w}_1, \mathbf{w}_2)$ are now established.

$$\mathcal{R}_{\mathsf{sum,dpc}} = \max_{\mathcal{P},\Pi} \left\{ \log\left(1 + \frac{\rho_{\pi_1}|\mathbf{h}_{\pi_1}^H \mathbf{w}_{\pi_1}|^2}{1 + \rho_{\pi_2}|\mathbf{h}_{\pi_2}^H \mathbf{w}_{\pi_2}|^2}\right) + \log\left(1 + \rho_{\pi_2}|\mathbf{h}_{\pi_2}^H \mathbf{w}_{\pi_2}|^2\right) \right\} \tag{151}$$

$$= \max_{\mathcal{P},\Pi} \left\{ \log\left(1 + \rho_{\pi_1}|\mathbf{h}_{\pi_1}^H \mathbf{w}_{\pi_1}|^2 + \rho_{\pi_2}|\mathbf{h}_{\pi_2}^H \mathbf{w}_{\pi_2}|^2\right) \right\} \tag{152}$$

• The quantity $d_{\mathbf{\Sigma}_i}(\mathbf{w}_1, \mathbf{w}_2)$ is a generalized "distance" semimetric[15] between $\mathbf{w}_1$ and $\mathbf{w}_2$ satisfying

$$0 \leq d_{\mathbf{\Sigma}_i}(\mathbf{w}_1, \mathbf{w}_2) \leq 1. \tag{164}$$

To establish the lower bound in (164), note that an application of the Cauchy–Schwarz inequality implies that $C_i^2 \leq A_i B_i$. Equality in the lower bound in (164) is achieved if and only if $\mathbf{\Sigma}_i^{1/2}\mathbf{w}_1 = \zeta \mathbf{\Sigma}_i^{1/2}\mathbf{w}_2$ for some $\zeta \in \mathbb{C}$. Since $\mathbf{\Sigma}_i$ is positive definite, this is possible only when $\mathbf{w}_1 = \zeta \mathbf{w}_2$. Since both $\mathbf{w}_1$ and $\mathbf{w}_2$ are unit-norm, this is possible only with $|\zeta| = 1$. In other words, equality in the lower bound only occurs for $\mathbf{w}_1 = \mathbf{w}_2$ on $\mathcal{G}(2,1)$. The fact that $d_{\mathbf{\Sigma}_i}(\mathbf{w}_1, \mathbf{w}_2) \leq 1$ is obvious. Symmetry of the distance metric in $\mathbf{w}_1$ and $\mathbf{w}_2$ is obvious.

• The triangle inequality does not hold in general. One counterexample is as follows:

$$\mathbf{\Sigma}_i = \mathsf{diag}\big([20, 1]\big), \quad \mathbf{w}_1 = \left[\frac{1}{\sqrt{3}}, \sqrt{\frac{2}{3}}\right], \tag{165}$$

$$\mathbf{w}_2 = \left[\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right], \quad \mathbf{w}_3 = \left[-\sqrt{\frac{3.3}{7}}, \sqrt{\frac{3.7}{7}}\right]. \tag{166}$$

This choice results in

$$d_{\mathbf{\Sigma}_i}(\mathbf{w}_1, \mathbf{w}_3) \approx 0.2536, \tag{167}$$

$$d_{\mathbf{\Sigma}_i}(\mathbf{w}_1, \mathbf{w}_2) + d_{\mathbf{\Sigma}_i}(\mathbf{w}_2, \mathbf{w}_3) \approx 0.2534. \tag{168}$$

Many such counterexamples can be listed out via a routine numerical search. Numerical studies also suggest that the triangle inequality holds for almost all choices of $\{\mathbf{w}_i\}$ provided that $\chi_i = \frac{\lambda_1(\mathbf{\Sigma}_i)}{\lambda_2(\mathbf{\Sigma}_i)}$ is not too large (unlike the example in (165) and (166)).

• The upper bound in (164) is achieved only if $A_i = B_i$ and $C_i = 0$. By decomposing $\mathbf{w}_1$ and $\mathbf{w}_2$ along the orthonormal set of basis vectors $\{\mathbf{u}_1(\mathbf{\Sigma}_i), \mathbf{u}_2(\mathbf{\Sigma}_i)\}$, it can be checked that $A_i = B_i$ and $C_i = 0$ is possible only if

$$\mathbf{w}_1 = e^{j\nu_1} \cdot \left[\mathbf{u}_1(\mathbf{\Sigma}_i) \cdot \sqrt{\frac{1}{\chi_i + 1}} + e^{j\nu_2} \cdot \mathbf{u}_2(\mathbf{\Sigma}_i) \cdot \sqrt{\frac{\chi_i}{\chi_i + 1}}\right] \tag{169}$$

$$\mathbf{w}_2 = e^{j(\nu_1+\nu_3)} \cdot \left[\mathbf{u}_1(\mathbf{\Sigma}_i) \cdot \sqrt{\frac{1}{\chi_i + 1}} - e^{j\nu_2} \cdot \mathbf{u}_2(\mathbf{\Sigma}_i) \cdot \sqrt{\frac{\chi_i}{\chi_i + 1}}\right] \tag{170}$$

for any choice of $\nu_j \in [0, 2\pi), j = 1, 2, 3$.

• The semimetric reduces to the standard *chordal* distance metric [39] on $\mathcal{G}(2,1)$

$$d_{\mathbf{\Sigma}_i}(\mathbf{w}_1, \mathbf{w}_2) = \sqrt{1 - |\mathbf{w}_1^H \mathbf{w}_2|^2} \tag{171}$$

if $\mathbf{\Sigma}_i = \lambda \mathbf{I}$ for some $\lambda > 0$.

[15]A semimetric satisfies all the properties necessary for a distance metric except the triangle inequality.

D) Monotonicity of $f(\bullet)$ and $g(\bullet)$: We first claim that

$$2 \leq \frac{1}{\sqrt{1-z^2}} \cdot \log\left(\frac{1+\sqrt{1-z^2}}{1-\sqrt{1-z^2}}\right) \leq \frac{2}{z^2}, \; 0 < z < 1, \tag{172}$$

which is equivalent to

$$\exp\left(2 \cdot \sqrt{1-z^2}\right) \leq \frac{1+\sqrt{1-z^2}}{1-\sqrt{1-z^2}} \leq \exp\left(\frac{2\sqrt{1-z^2}}{z^2}\right), \quad 0 < z < 1. \tag{173}$$

For this, we start with the exponential series expansion of $\exp\left(2 \cdot \sqrt{1-z^2}\right)$ that results in

$$\exp\left(2 \cdot \sqrt{1-z^2}\right) = 1 + \sum_{k=1}^{\infty} \frac{2^k \cdot \left(1-z^2\right)^{\frac{k}{2}}}{\Gamma(k+1)} \tag{174}$$

$$\leq 1 + 2\sum_{k=1}^{\infty} \left(1-z^2\right)^{\frac{k}{2}} = 1 + \frac{2\sqrt{1-z^2}}{1-\sqrt{1-z^2}} \tag{175}$$

where $\Gamma(\cdot)$ is the Gamma function, the second inequality follows from the fact that $\frac{2^{k-1}}{\Gamma(k+1)} \leq 1$ for all $k \geq 1$, and the last equality from the sum of an infinite geometric series. For the other side of (172), note that

$$\exp\left(\frac{2\sqrt{1-z^2}}{z^2}\right) \geq 1 + \frac{2\sqrt{1-z^2}}{z^2} + \frac{2(1-z^2)}{z^4} \tag{176}$$

$$\geq 1 + \frac{2\sqrt{1-z^2}}{z^2}\left(1+\sqrt{1-z^2}\right) \tag{177}$$

$$= 1 + \frac{2\sqrt{1-z^2}}{1-\sqrt{1-z^2}} \tag{178}$$

where the first inequality follows by truncating the terms of the asymptotic expansion and the second follows by using the fact that $z^2 < 1$. The proof is complete by noting that

$$\frac{\partial f(z)}{\partial z} = \frac{-z}{1-z^2}\left[\frac{2}{z^2} - \frac{1}{\sqrt{1-z^2}} \log\left(\frac{1+\sqrt{1-z^2}}{1-\sqrt{1-z^2}}\right)\right] < 0 \tag{179}$$

$$\frac{\partial g(z)}{\partial z} = \frac{z}{1-z^2}\left[\frac{1}{\sqrt{1-z^2}} \log\left(\frac{1+\sqrt{1-z^2}}{1-\sqrt{1-z^2}}\right) - 2\right] > 0. \tag{180}$$

■

E) Completing the Proof of Theorem 2: With the description of $\mathbf{w}_1$ and $\mathbf{w}_2$ as in (48) and (49), elementary algebra shows that

$$A_1 = \mathbf{w}_1^H \mathbf{\Sigma}_1 \mathbf{w}_1 = \frac{|\alpha|^2\eta_1 + |\beta|^2\eta_2}{X^2} \tag{181}$$

$$B_1 = \mathbf{w}_2^H \mathbf{\Sigma}_1 \mathbf{w}_2 = \frac{|\gamma|^2\eta_1 + |\delta|^2\eta_2}{Y^2} \tag{182}$$

$$C_1 = |\mathbf{w}_1^H \mathbf{\Sigma}_1 \mathbf{w}_2| = \frac{\left|\alpha^\star\gamma\eta_1 + \beta^\star\delta\eta_2\right|}{XY} \tag{183}$$

$$A_2 = \mathbf{w}_2^H \mathbf{\Sigma}_2 \mathbf{w}_2 = \frac{1}{Y^2} \tag{184}$$

$$B_2 = \mathbf{w}_1^H \mathbf{\Sigma}_2 \mathbf{w}_1 = \frac{1}{X^2} \tag{185}$$

$$C_2 = |\mathbf{w}_1^H \mathbf{\Sigma}_2 \mathbf{w}_2| = \frac{\left|\alpha^\star \gamma + \beta^\star \delta\right|}{XY}, \text{ where} \tag{186}$$

$$X = \|\alpha \mathbf{\Sigma}_2^{-\frac{1}{2}} \mathbf{v}_1 + \beta \mathbf{\Sigma}_2^{-\frac{1}{2}} \mathbf{v}_2\|, \tag{187}$$

$$Y = \|\gamma \mathbf{\Sigma}_2^{-\frac{1}{2}} \mathbf{v}_1 + \delta \mathbf{\Sigma}_2^{-\frac{1}{2}} \mathbf{v}_2\|. \tag{188}$$

As in the statement of the theorem, let $\tau_i,\ i = 1, 2, 3$ denote

$$\tau_1 = \mathbf{v}_1^H \mathbf{\Sigma}_2^{-1} \mathbf{v}_1,\ \tau_2 = \mathbf{v}_2^H \mathbf{\Sigma}_2^{-1} \mathbf{v}_2,\ \tau_3 = \mathbf{v}_1^H \mathbf{\Sigma}_2^{-1} \mathbf{v}_2. \tag{189}$$

We can rewrite X^2 and Y^2 in terms of $\{\tau_i\}$ as

$$X^2 = |\alpha|^2 \tau_1 + |\beta|^2 \tau_2 + 2|\alpha||\beta||\tau_3| \cos(\theta_1) \tag{190}$$

$$Y^2 = |\gamma|^2 \tau_1 + |\delta|^2 \tau_2 + 2|\gamma||\delta||\tau_3| \cos(\theta_2) \tag{191}$$

where $\theta_1 = \mathsf{arg}(\tau_3) + \theta_\beta - \theta_\alpha$ and $\theta_2 = \mathsf{arg}(\tau_3) + \theta_\delta - \theta_\gamma$. Now note that if $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a pair of eigenvectors for $\mathbf{\Sigma}$, then so is the pair $\left\{e^{j\nu_1}\mathbf{v}_1, e^{j\nu_2}\mathbf{v}_2\right\}$ for any choice of ν_1 and ν_2. In other words, the choice of $\{\mathbf{v}_1, \mathbf{v}_2\}$ is unique *only* on $\mathcal{G}(2,1)$, and not on $\mathsf{St}(2,1)$. Hence, $\mathsf{arg}(\tau_3)$ can be chosen arbitrarily and *independently* in determining the values of X^2 and Y^2. With the specific choice that $\mathsf{arg}(\mathbf{v}_1^H \mathbf{\Sigma}_2^{-1} \mathbf{v}_2) = \frac{\pi}{2} + \theta_\alpha - \theta_\beta$ in (190) and $\mathsf{arg}(\mathbf{v}_1^H \mathbf{\Sigma}_2^{-1} \mathbf{v}_2) = \frac{\pi}{2} + \theta_\gamma - \theta_\delta$ in (191), we have

$$X^2 = |\alpha|^2 \tau_1 + |\beta|^2 \tau_2 \tag{192}$$

$$Y^2 = |\gamma|^2 \tau_1 + |\delta|^2 \tau_2. \tag{193}$$

Thus, the high-SNR expression for the ergodic sum-rate can be simplified as in (194) at the bottom of the page. Using (160) to rewrite the above equation in terms of $f(\bullet)$, we have after simplification the (195) also at the bottom of the page. We now claim that the following two inequalities hold:

$$\frac{XY}{X^2 + Y^2} \geq \frac{\sqrt{\tau_1 \tau_2}}{\tau_1 + \tau_2} \tag{196}$$

$$\frac{XY \cdot \left(|\gamma|^2 \eta_1 + |\delta|^2 \eta_2\right)}{\left(|\alpha|^2 \eta_1 + |\beta|^2 \eta_2\right) Y^2 + \left(|\gamma|^2 \eta_1 + |\delta|^2 \eta_2\right) X^2} \geq \frac{\sqrt{\tau_1 \tau_2} \cdot \eta_2}{\tau_1 \eta_2 + \tau_2 \eta_1}. \tag{197}$$

The proof of (196) and (197) will be tackled later.

Using (196) and (197) in conjunction with the decreasing nature of $f(\bullet)$, we have (198) and (199) at the bottom of the page. Note that $\{\theta_\bullet\}$ enter the above optimization only via the term $|\beta\gamma - \alpha\delta|$ and

$$|\beta\gamma - \alpha\delta| \leq |\alpha| \sqrt{1 - |\gamma|^2} + |\gamma| \sqrt{1 - |\alpha|^2} \tag{200}$$

with equality achieved if and only if $\theta_\alpha + \theta_\delta - \theta_\beta - \theta_\gamma = \pi$ (modulo 2π). Parameterizing $|\alpha|$ and $|\gamma|$ as $|\alpha| = \sin(\theta)$ and $|\gamma| = \sin(\phi)$ for some $\{\theta, \phi\} \in [0, \pi/2]$, we have

$$|\beta\gamma - \alpha\delta| \leq \sin(\theta)\cos(\phi) + \cos(\theta)\sin(\phi) = \sin(\theta + \phi) \leq 1 \tag{201}$$

since $0 \leq \theta + \phi \leq \pi$. Further, $\eta_1 \geq \eta_2$ implies that $|\gamma|^2 \eta_1 + |\delta|^2 \eta_2 \geq \eta_2$ and hence, we have

$$\frac{|\beta\gamma - \alpha\delta|}{|\gamma|^2 \eta_1 + |\delta|^2 \eta_2} \leq \frac{1}{\eta_2}. \tag{202}$$

$$2E\left[R_1\right] + 2E\left[R_2\right] + 4\log(2) = g\left(\frac{2\sqrt{\eta_1\eta_2} XY \cdot |\beta\gamma - \alpha\delta|}{\left(|\alpha|^2\eta_1 + |\beta|^2\eta_2\right) \cdot Y^2 + \left(|\gamma|^2\eta_1 + |\delta|^2\eta_2\right) \cdot X^2}\right) + 2\log\left(1 + \frac{X^2}{Y^2}\right) + g\left(\frac{2XY \cdot |\beta\gamma - \alpha\delta|}{X^2 + Y^2}\right) + 2\log\left(1 + \frac{Y^2}{X^2} \cdot \frac{|\alpha|^2\eta_1 + |\beta|^2\eta_2}{|\gamma|^2\eta_1 + |\delta|^2\eta_2}\right) \tag{194}$$

$$2E\left[R_1\right] + 2E\left[R_2\right] - \log(\eta_1\eta_2) = f\left(\frac{2\sqrt{\eta_1\eta_2} XY \cdot |\beta\gamma - \alpha\delta|}{\left(|\alpha|^2\eta_1 + |\beta|^2\eta_2\right) Y^2 + \left(|\gamma|^2\eta_1 + |\delta|^2\eta_2\right) X^2}\right) + f\left(\frac{2XY \cdot |\beta\gamma - \alpha\delta|}{X^2 + Y^2}\right) + 2\log\left(\frac{|\beta\gamma - \alpha\delta|^2}{|\gamma|^2\eta_1 + |\delta|^2\eta_2}\right) \tag{195}$$

$$2E\left[R_1\right] + 2E\left[R_2\right] - \log(\eta_1\eta_2) \leq f\left(\frac{2\sqrt{\eta_1\eta_2\tau_1\tau_2} \cdot \eta_2}{\tau_1\eta_2 + \tau_2\eta_1} \cdot \frac{|\beta\gamma - \alpha\delta|}{\left(|\gamma|^2\eta_1 + |\delta|^2\eta_2\right)}\right) + f\left(\frac{2\sqrt{\tau_1\tau_2}}{\tau_1 + \tau_2} \cdot |\beta\gamma - \alpha\delta|\right) + 2\log\left(\frac{|\beta\gamma - \alpha\delta|^2}{|\gamma|^2\eta_1 + |\delta|^2\eta_2}\right) \tag{198}$$

$$= g\left(\frac{2\sqrt{\eta_1\eta_2\tau_1\tau_2} \cdot \eta_2}{\tau_1\eta_2 + \tau_2\eta_1} \cdot \frac{|\beta\gamma - \alpha\delta|}{\left(|\gamma|^2\eta_1 + |\delta|^2\eta_2\right)}\right) + g\left(\frac{2\sqrt{\tau_1\tau_2}}{\tau_1 + \tau_2} \cdot |\beta\gamma - \alpha\delta|\right) + 2\log\left(\frac{\left(\tau_1\eta_2 + \tau_2\eta_1\right)\left(\tau_1 + \tau_2\right)}{4\tau_1\tau_2\eta_2\sqrt{\eta_1\eta_2}}\right) \tag{199}$$

Using the fact that $g(\bullet)$ is an increasing function, we get an upper bound for the sum-rate as

$$2E[R_1] + 2E[R_2] \leq f\left(\frac{2\sqrt{\eta_1\eta_2\tau_1\tau_2}}{\eta_1\tau_2+\eta_2\tau_1}\right) + f\left(\frac{2\sqrt{\tau_1\tau_2}}{\tau_1+\tau_2}\right) + \log\left(\frac{\eta_1}{\eta_2}\right). \tag{203}$$

This upper bound is achievable with the choice of $|\alpha| = 1$ and $|\gamma| = 0$ in (48) and (49), which is equivalent to (44). Substituting this choice in the sum-rate expression yields (204) and (205) at the bottom of the page. To simplify (205), we define κ_1 and κ_2 as

$$\kappa_1 = \frac{\eta_1\tau_2}{\eta_2\tau_1} \text{ and } \kappa_2 = \frac{\tau_2}{\tau_1}. \tag{206}$$

The fact that $\eta_1 \geq \eta_2$ implies that $\kappa_1 \geq \kappa_2$. Thus, there are three possibilities: 1) $\kappa_1 \geq \kappa_2 \geq 1$, 2) $\kappa_1 \geq 1 \geq \kappa_2$, and 3) $1 \geq \kappa_1 \geq \kappa_2$. It is straightforward but tedious to check that in all of the three cases, (205) reduces to (45) as in the statement of the theorem. The proof will be complete if the inequalities (196) and (197) can be established.

6) Proof of (196): For the first inequality, note that

$$\frac{X^2+Y^2}{XY} = \frac{X}{Y} + \frac{Y}{X} = t + \frac{1}{t} \triangleq q(t) \tag{207}$$

can be written as a symmetric function $q(t)$ in t where $t = \frac{X}{Y}$. Further, noting that $q(t)$ is decreasing in t for $t \leq 1$ and is increasing in t for $t \geq 1$, the maximum of $\frac{X^2+Y^2}{XY}$ is achieved either when $\frac{X}{Y}$ achieves its largest or smallest value. The inequality in (196) follows since

$$\sqrt{\frac{\min(\tau_1,\tau_2)}{\max(\tau_1,\tau_2)}} \leq \frac{X}{Y} \leq \sqrt{\frac{\max(\tau_1,\tau_2)}{\min(\tau_1,\tau_2)}}. \tag{208}$$

7) Proof of (197): The proof of (197) is more involved. For this, we define $\mathcal{L}_1$ as in (209) at the bottom of the page. By taking derivative of this expression with respect to $|\alpha|^2$, note that the first term in (209) is increasing in $|\alpha|^2$ for any fixed choice of $|\gamma|$ if and only if $\frac{\tau_1}{\tau_2} \geq 1$. Similarly, for any fixed choice of $|\gamma|$, the second term in (209) is increasing in $|\alpha|^2$ if and only if $\frac{\eta_1}{\eta_2} \geq \frac{\tau_1}{\tau_2}$. Thus, the condition

$$1 \leq \frac{\tau_1}{\tau_2} \leq \frac{\eta_1}{\eta_2} \tag{210}$$

is necessary and sufficient to ensure that for any choice of $|\gamma|$, $\mathcal{L}_1$ is maximized by the choice $|\alpha| = 1$. On analogous lines, taking the derivative with respect to $|\gamma|^2$, it can be seen that for any fixed choice of $|\alpha|$, both the terms in (209) are decreasing in $|\gamma|^2$ if and only if the same condition in (210) holds. In other words, under (210), $\mathcal{L}_1$ is maximized by $|\alpha| = 1$ and $|\gamma| = 0$. At this stage, two other possibilities need to be considered: 1) $\frac{\tau_1}{\tau_2} \leq 1 \leq \frac{\eta_1}{\eta_2}$, and 2) $1 \leq \frac{\eta_1}{\eta_2} \leq \frac{\tau_1}{\tau_2}$. In either case, we will show that

$$\left(\frac{|\alpha|^2(\eta_1-\eta_2)+\eta_2}{|\gamma|^2(\eta_1-\eta_2)+\eta_2} - \frac{\eta_1}{\eta_2}\right)\cdot\frac{Y}{X} + \left(\frac{X}{Y} - \sqrt{\frac{\tau_1}{\tau_2}}\right)\cdot\left(1 - \frac{\eta_1}{\eta_2}\sqrt{\frac{\tau_2}{\tau_1}}\frac{Y}{X}\right) \leq 0, \tag{211}$$

which is equivalent to (197), or the statement that $\mathcal{L}_1$ is maximized by $|\alpha| = 1$ and $|\gamma| = 0$. For this, note that in either case, we have

$$\frac{|\alpha|^2(\eta_1-\eta_2)+\eta_2}{|\gamma|^2(\eta_1-\eta_2)+\eta_2} \leq \frac{\eta_1}{\eta_2}. \tag{212}$$

In the first case, we also have

$$\sqrt{\frac{\tau_1}{\tau_2}} \leq \frac{X}{Y} \leq \sqrt{\frac{\tau_2}{\tau_1}} \leq \frac{\eta_1}{\eta_2}\sqrt{\frac{\tau_2}{\tau_1}}, \tag{213}$$

where the last step in (213) follows from $\frac{\eta_1}{\eta_2} \geq 1$. Combining (212) and (213), we note that (211) is immediate when $\frac{\tau_1}{\tau_2} \leq 1 \leq \frac{\eta_1}{\eta_2}$. In the second case, however, (213) is replaced with

$$\sqrt{\frac{\tau_2}{\tau_1}} \leq \frac{X}{Y} \leq \sqrt{\frac{\tau_1}{\tau_2}}. \tag{214}$$

It can be seen that if $|\alpha|$ and $|\gamma|$ are such that

$$\frac{X}{Y} = D\cdot\frac{\eta_1}{\eta_2}\sqrt{\frac{\tau_2}{\tau_1}}, \tag{215}$$

$$E[R_1] + E[R_2] \xrightarrow{\rho\to\infty} \frac{1}{2}f\left(\frac{2\sqrt{\eta_1\eta_2\tau_1\tau_2}}{\eta_1\tau_2+\eta_2\tau_1}\right) + \frac{1}{2}f\left(\frac{2\sqrt{\tau_1\tau_2}}{\tau_1+\tau_2}\right) + \frac{1}{2}\log\left(\frac{\eta_1}{\eta_2}\right) \tag{204}$$

$$= \frac{1}{2}\cdot\frac{\eta_1\tau_2+\eta_2\tau_1}{|\eta_1\tau_2-\eta_2\tau_1|}\cdot\log\left(\frac{\eta_1\tau_2+\eta_2\tau_1+|\eta_1\tau_2-\eta_2\tau_1|}{\eta_1\tau_2+\eta_2\tau_1-|\eta_1\tau_2-\eta_2\tau_1|}\right) + \frac{1}{2}\cdot\frac{\tau_1+\tau_2}{|\tau_1-\tau_2|}\cdot\log\left(\frac{\tau_1+\tau_2+|\tau_1-\tau_2|}{\tau_1+\tau_2-|\tau_1-\tau_2|}\right) + \frac{1}{2}\log\left(\frac{\eta_1}{\eta_2}\right) \tag{205}$$

$$\mathcal{L}_1 \triangleq \frac{\left(|\alpha|^2\eta_1+|\beta|^2\eta_2\right)Y}{\left(|\gamma|^2\eta_1+|\delta|^2\eta_2\right)X} + \frac{\left(|\gamma|^2\eta_1+|\delta|^2\eta_2\right)X}{\left(|\gamma|^2\eta_1+|\delta|^2\eta_2\right)Y} = \sqrt{\frac{|\alpha|^2(\tau_1-\tau_2)+\tau_2}{|\gamma|^2(\tau_1-\tau_2)+\tau_2}}\left[1 + \frac{|\alpha|^2(\eta_1-\eta_2)+\eta_2}{|\gamma|^2(\eta_1-\eta_2)+\eta_2}\cdot\frac{|\gamma|^2(\tau_1-\tau_2)+\tau_2}{|\alpha|^2(\tau_1-\tau_2)+\tau_2}\right] \tag{209}$$

for some choice of D satisfying $1 \leq D \leq \frac{\tau_1}{\tau_2} \cdot \frac{\eta_2}{\eta_1}$, (211) holds immediately. Thus, we only need to show that (211) holds when $|\alpha|$ and $|\gamma|$ are such that

$$\frac{X}{Y} = D \cdot \frac{\eta_1}{\eta_2}\sqrt{\frac{\tau_2}{\tau_1}}, \tag{216}$$

for some choice of D satisfying $\frac{\eta_2}{\eta_1} \leq D \leq 1$. After some elementary algebra, our task is to show that

$$\frac{|\alpha|^2(\eta_1-\eta_2)+\eta_2}{|\gamma|^2(\eta_1-\eta_2)+\eta_2} \leq \frac{\eta_1}{\eta_2}\left[1-(1-D)\left[1-\frac{D\eta_1\tau_2}{\eta_2\tau_1}\right]\right] \tag{217}$$

$$\frac{X}{Y} = D \cdot \frac{\eta_1}{\eta_2}\sqrt{\frac{\tau_2}{\tau_1}}. \tag{218}$$

By bounding the denominator of (217) as $\eta_2 \leq |\gamma|^2(\eta_1 - \eta_2) + \eta_2 \leq \eta_1$, it can be seen that (217) holds if the quadratic inequality in D as in (219) at the bottom of the page is true. For this, we note that the left-hand side represents a convex parabola in D with maximum achieved at either $D = \frac{\eta_2}{\eta_1}$ or $D = 1$. Substituting $D = \frac{\eta_2}{\eta_1}$ and $D = 1$ and simplifying, we see the equations in (220) and (221) at the bottom of the page to be true. Since the maximum of the parabola in the domain $\frac{\eta_2}{\eta_1} \leq D \leq 1$ is below 0, (211) holds. Thus, we are done with the aspect of showing that $|\alpha| = 1, |\gamma| = 0$ is sum-rate optimal. ■

H) Proof of Theorem 3: Following the logic of Appendix E, we decompose $\mathbf{w}_1$ and $\mathbf{w}_2$ along $\{\mathbf{u}_1, \mathbf{u}_2\}$ since they form an orthonormal basis:

$$\mathbf{w}_1 = \alpha\mathbf{u}_1 + \beta\mathbf{u}_2 \tag{222}$$

$$\mathbf{w}_2 = \gamma\mathbf{u}_1 + \delta\mathbf{u}_2 \tag{223}$$

for some choice of $\{\alpha, \beta, \gamma, \delta\}$ with $\alpha = |\alpha|e^{j\theta_\alpha}$ (similarly, for other quantities) satisfying $|\alpha|^2 + |\beta|^2 = |\gamma|^2 + |\delta|^2 = 1$. We now study the ergodic sum-rate optimization over the 6-D parameter space. With the description of $\mathbf{w}_1$ and $\mathbf{w}_2$ as in (222) and (223), elementary algebra shows that

$$A_1 = \mathbf{w}_1^H\mathbf{\Sigma}_1\mathbf{w}_1 = |\alpha|^2\lambda_1 + |\beta|^2\lambda_2 \tag{224}$$

$$B_1 = \mathbf{w}_2^H\mathbf{\Sigma}_1\mathbf{w}_2 = |\gamma|^2\lambda_1 + |\delta|^2\lambda_2 \tag{225}$$

$$C_1 = |\mathbf{w}_1^H\mathbf{\Sigma}_1\mathbf{w}_2| = |\alpha^\star\gamma\lambda_1 + \beta^\star\delta\lambda_2| \tag{226}$$

$$A_2 = \mathbf{w}_2^H\mathbf{\Sigma}_2\mathbf{w}_2 = |\gamma|^2\mu_1 + |\delta|^2\mu_2 \tag{227}$$

$$B_2 = \mathbf{w}_1^H\mathbf{\Sigma}_2\mathbf{w}_1 = |\alpha|^2\mu_1 + |\beta|^2\mu_2 \tag{228}$$

$$C_2 = |\mathbf{w}_1^H\mathbf{\Sigma}_2\mathbf{w}_2| = |\alpha^\star\gamma\mu_1 + \beta^\star\delta\mu_2| \tag{229}$$

and hence,

$$d_{\mathbf{\Sigma}_1}(\mathbf{w}_1, \mathbf{w}_2)^2 = \frac{4\lambda_1\lambda_2 \cdot |\beta\gamma - \alpha\delta|^2}{\left[(|\alpha|^2 + |\gamma|^2)\lambda_1 + (|\beta|^2 + |\delta|^2)\lambda_2\right]^2} \tag{230}$$

$$d_{\mathbf{\Sigma}_2}(\mathbf{w}_1, \mathbf{w}_2)^2 = \frac{4\mu_1\mu_2 \cdot |\beta\gamma - \alpha\delta|^2}{\left[(|\alpha|^2 + |\gamma|^2)\mu_1 + (|\beta|^2 + |\delta|^2)\mu_2\right]^2}. \tag{231}$$

The high-SNR expression of the ergodic sum-rate can be written as in (232) at the bottom of the page. Note that $\{\theta_\bullet\}$ enter the above optimization only via the term $|\beta\gamma - \alpha\delta|$ and as in (200), we have

$$|\beta\gamma - \alpha\delta| \leq |\beta||\gamma| + |\alpha||\delta| \leq 1. \tag{233}$$

Given that $\chi_1 = \frac{\lambda_1}{\lambda_2} \geq 1$, three possibilities arise depending on the relationship between 1, χ_1 and $\chi_2 = \frac{\mu_1}{\mu_2}$: 1) $\chi_1 > 1 \geq \chi_2$, 2) $\chi_1 > \chi_2 > 1$, and 3) $\chi_2 \geq \chi_1 > 1$.

$$D^2 \cdot \frac{\eta_1^2\tau_2}{\eta_2^2\tau_1} \cdot \left(2 - \frac{\tau_1\eta_2 - \tau_2\eta_1}{\eta_1(\tau_1 - \tau_2)}\right) - D \cdot \frac{\eta_1}{\eta_2} \cdot \left(1 + \frac{\eta_1\tau_2}{\eta_2\tau_1}\right) + \frac{\tau_1\eta_2 - \tau_2\eta_1}{\eta_2(\tau_1 - \tau_2)} \leq 0 \tag{219}$$

$$\text{LHS of (219)}\Big|_{D=\frac{\eta_2}{\eta_1}} = -\frac{\tau_2}{\eta_1\eta_2\tau_1} \cdot \frac{\eta_1 - \eta_2}{\tau_1 - \tau_2} \cdot (\eta_1(\tau_1 - \tau_2) + \tau_1(\eta_1 - \eta_2)) \leq 0 \tag{220}$$

$$\text{LHS of (219)}\Big|_{D=1} = -\frac{(\eta_1 - \eta_2) \cdot (\eta_2\tau_1 - \eta_1\tau_2)}{\eta_2^2(\tau_1 - \tau_2)} \leq 0 \tag{221}$$

$$2E\left[R_1\right] + 2E\left[R_2\right] + 4\log(2) = g\left(\frac{\sqrt{4\lambda_1\lambda_2} \cdot |\beta\gamma - \alpha\delta|}{(|\alpha|^2 + |\gamma|^2)\lambda_1 + (|\beta|^2 + |\delta|^2)\lambda_2}\right) + g\left(\frac{\sqrt{4\mu_1\mu_2} \cdot |\beta\gamma - \alpha\delta|}{(|\alpha|^2 + |\gamma|^2)\mu_1 + (|\beta|^2 + |\delta|^2)\mu_2}\right) + 2\log\left(1 + \frac{|\alpha|^2\lambda_1 + |\beta|^2\lambda_2}{|\gamma|^2\lambda_1 + |\delta|^2\lambda_2}\right) + 2\log\left(1 + \frac{|\gamma|^2\mu_1 + |\delta|^2\mu_2}{|\alpha|^2\mu_1 + |\beta|^2\mu_2}\right) \tag{232}$$

Case i): In the first case where $\chi_2 \leq 1$, we use the fact that $g(\bullet)$ is an increasing function to bound the sum-rate as in (234) at the bottom of the page. Observing that

$$\begin{aligned}&(\sin^2(\theta)+\sin^2(\phi))(\lambda_1-\lambda_2)+2\lambda_2\\&\quad\leq \sin^2(\phi)(\lambda_1-\lambda_2)+\lambda_1+\lambda_2 \qquad (235)\\&2\mu_2-(\sin^2(\theta)+\sin^2(\phi))(\mu_2-\mu_1)\\&\quad\leq 2\mu_2-\sin^2(\theta)(\mu_2-\mu_1) \qquad (236)\end{aligned}$$

and $f(\bullet)$ is a decreasing function, we have (237) and (238) at the bottom of the page.

It is straightforward to note that the right-hand side of (238) is decreasing in $\sin^2(\phi)$ and increasing in $\sin^2(\theta)$. If in addition, the condition that $\theta+\phi=\pi/2$ is satisfied, then an upper bound on $E[R_1]+E[R_2]$ can be maximized. This results in the choice $\theta=\pi/2$ and $\phi=0$ (that is, $|\alpha|=1$ and $|\gamma|=0$) and with this choice, we have

$$\begin{aligned}&2E[R_1]+2E[R_2]\\&\leq f\left(\frac{\sqrt{4\lambda_1\lambda_2}}{\lambda_1+\lambda_2}\right)+f\left(\frac{\sqrt{4\mu_1\mu_2}}{\mu_1+\mu_2}\right)+\log\left(\frac{\lambda_1\mu_2}{\lambda_2\mu_1}\right).\end{aligned} \qquad (239)$$

It is also straightforward to check that the choice of $\mathbf{w}_1$ and $\mathbf{w}_2$ as in the statement of the theorem meets this upper bound and is thus optimal.

Case ii): In the second case where $\chi_1 > \chi_2 > 1$, we start as in Case i) and after optimization over $\{\theta_\bullet\}$, we can bound the sum-rate as in (240) and (241) at the bottom of the page. The joint dependence between θ and ϕ in the right-hand side of (241) precludes the possibility of breaking down the double variable optimization of (241) into a pair of single variable optimizations. That is, the technique from *Case i)* fails here and this case needs to be studied differently.

The proof in this case follows in three steps. In the first step, when $\chi_1 > \chi_2$, we show that $\mathcal{L}_2$ (defined as in (242) and (243) at the bottom of the next page) is maximized by $\theta=\pi/2$ and $\phi=0$. For this, we set $\sin^2(\theta)+\sin^2(\phi)$ to take a specific value of α. Since the numerator is only a function of α, maximizing $\mathcal{L}_2$ is equivalent to minimizing the denominator of (243). There are two possible cases depending on whether $\alpha \leq 1$ or $1 < \alpha \leq 2$. In the former case, it can be seen that the denominator is minimized by $\phi=0$ and $\theta=\sin^{-1}(\sqrt{\alpha})$, whereas in the latter

$$\begin{aligned}2E[R_1]+2E[R_2]+4\log(2) \leq\ & f\left(\frac{\sqrt{4\lambda_1\lambda_2}\cdot\sin(\theta+\phi)}{(\sin^2(\theta)+\sin^2(\phi))(\lambda_1-\lambda_2)+2\lambda_2}\right)+f\left(\frac{\sqrt{4\mu_1\mu_2}\cdot\sin(\theta+\phi)}{2\mu_2-(\sin^2(\theta)+\sin^2(\phi))(\mu_2-\mu_1)}\right)\\&+2\log\left(\frac{\sqrt{4\lambda_1\lambda_2}\cdot\sin(\theta+\phi)}{\sin^2(\phi)(\lambda_1-\lambda_2)+\lambda_2}\right)+2\log\left(\frac{\sqrt{4\mu_1\mu_2}\cdot\sin(\theta+\phi)}{\mu_2-\sin^2(\theta)(\mu_2-\mu_1)}\right)\end{aligned} \qquad (234)$$

$$\begin{aligned}2E[R_1]+2E[R_2]+4\log(2) \leq\ & f\left(\frac{\sqrt{4\lambda_1\lambda_2}\cdot\sin(\theta+\phi)}{\sin^2(\phi)(\lambda_1-\lambda_2)+\lambda_1+\lambda_2}\right)+f\left(\frac{\sqrt{4\mu_1\mu_2}\cdot\sin(\theta+\phi)}{2\mu_2-\sin^2(\theta)(\mu_2-\mu_1)}\right)\\&+2\log\left(\frac{\sqrt{4\lambda_1\lambda_2}\cdot\sin(\theta+\phi)}{\sin^2(\phi)(\lambda_1-\lambda_2)+\lambda_2}\right)+2\log\left(\frac{\sqrt{4\mu_1\mu_2}\cdot\sin(\theta+\phi)}{\mu_2-\sin^2(\theta)(\mu_2-\mu_1)}\right) \qquad (237)\\=\ & g\left(\frac{\sqrt{4\lambda_1\lambda_2}\cdot\sin(\theta+\phi)}{\sin^2(\phi)(\lambda_1-\lambda_2)+\lambda_1+\lambda_2}\right)+g\left(\frac{\sqrt{4\mu_1\mu_2}\cdot\sin(\theta+\phi)}{2\mu_2-\sin^2(\theta)(\mu_2-\mu_1)}\right)\\&+2\log\left(1+\frac{\lambda_1}{\sin^2(\phi)(\lambda_1-\lambda_2)+\lambda_2}\right)+2\log\left(1+\frac{\mu_2}{\mu_2-\sin^2(\theta)(\mu_2-\mu_1)}\right)\end{aligned} \qquad (238)$$

$$\begin{aligned}2E[R_1]+2E[R_2]+4\log(2) \leq\ & f\left(\frac{2\sqrt{\chi_1}\cdot\sin(\theta+\phi)}{(\chi_1-1)(\sin^2(\theta)+\sin^2(\phi))+2}\right)+f\left(\frac{2\sqrt{\chi_2}\cdot\sin(\theta+\phi)}{(\chi_2-1)(\sin^2(\theta)+\sin^2(\phi))+2}\right)\\&+2\log\left(\frac{2\sqrt{\chi_1}\cdot\sin(\theta+\phi)}{(\chi_1-1)\sin^2(\phi)+1}\right)+2\log\left(\frac{2\sqrt{\chi_2}\cdot\sin(\theta+\phi)}{(\chi_2-1)\sin^2(\theta)+1}\right) \qquad (240)\\=\ & g\left(\frac{2\sqrt{\chi_1}\cdot\sin(\theta+\phi)}{(\chi_1-1)(\sin^2(\theta)+\sin^2(\phi))+2}\right)+g\left(\frac{2\sqrt{\chi_2}\cdot\sin(\theta+\phi)}{(\chi_2-1)(\sin^2(\theta)+\sin^2(\phi))+2}\right)\\&+2\log\left(\frac{(\chi_1-1)(\sin^2(\theta)+\sin^2(\phi))+2}{(\chi_1-1)\sin^2(\phi)+1}\right)+2\log\left(\frac{(\chi_2-1)(\sin^2(\theta)+\sin^2(\phi))+2}{(\chi_2-1)\sin^2(\theta)+1}\right)\end{aligned} \qquad (241)$$

case, it is minimized by $\phi = \sin^{-1}(\sqrt{\alpha - 1})$ and $\theta = \pi/2$. Substituting these values, it can be seen that

$$\mathcal{L}_2 \leq \begin{cases} \frac{(\chi_1-1)(\chi_2-1)\alpha^2+2\alpha(\chi_1+\chi_2-2)+4}{1+(\chi_2-1)\alpha} & \text{if } \alpha \leq 1 \\ \frac{(\chi_1-1)(\chi_2-1)\alpha^2+2\alpha(\chi_1+\chi_2-2)+4}{\chi_2\cdot[1+(\chi_1-1)(\alpha-1)]} & \text{if } 1 < \alpha \leq 2. \end{cases} \tag{244}$$

A straightforward derivative calculation (using the critical fact that $\chi_1 > \chi_2$) shows that while the right-hand side of (244) is increasing for $\alpha \leq 1$, it is decreasing for $1 < \alpha \leq 2$.

In other words,

$$\mathcal{L}_2 \leq \frac{(1+\chi_1)\cdot(1+\chi_2)}{\chi_2}, \tag{245}$$

and this upper bound is achieved with $\theta = \pi/2, \phi = 0$.

In the second step, if $\theta + \phi > \pi/2$, we have

$$\frac{\sin(\theta+\phi)}{(\chi_i - 1)\left(\sin^2(\theta)+\sin^2(\phi)\right)+2} \leq \frac{\sin(\theta+\phi)}{\chi_i+1} \leq \frac{1}{\chi_i+1}, \quad i = 1, 2 \tag{246}$$

since $\sin(\theta) > \sin(\pi/2 - \phi) = \cos(\phi)$. Therefore, we have the equations in (247) and (248) at the bottom of the page with (248) following from the monotonicity of $g(\bullet)$. Using Step 1, this results in

$$\begin{aligned} &2E\left[R_1\right] + 2E\left[R_2\right] \\ &\quad \leq f\left(\frac{2\sqrt{\chi_1}}{\chi_1+1}\right) + f\left(\frac{2\sqrt{\chi_2}}{\chi_2+1}\right) + \log\left(\frac{\chi_1}{\chi_2}\right). \end{aligned} \tag{249}$$

Note that (246) fails if $\theta + \phi = \nu \leq \pi/2$. Thus, in the third step, we consider this possibility. Here, a straightforward manipulation shows that

$$\begin{aligned} &\sin^2(\theta) + \sin^2(\nu - \theta) \\ &\quad = \sin^2(\nu) - 2\sin(\theta)\sin(\nu-\theta)\cos(\nu) \leq \sin^2(\nu), \end{aligned} \tag{250}$$

where the last inequality follows because $\nu \leq \pi/2$. Hence, we have the equation in (251) at the bottom of the page. It can be easily seen that $\mathcal{L}_3$ is maximized by $\theta = \pi/2$ and $\phi = 0$. The upper bound is the same in both the cases $\theta + \phi \leq \pi/2$ and $\theta + \phi > \pi/2$. And this upper bound is met by the choice $\theta = \pi/2$ and $\phi = 0$ (that is, $|\alpha| = 1$ and $|\gamma| = 0$) and is hence optimal.

Case iii): Since the expression for the sum-rate is symmetric in χ_1 and χ_2, an argument analogous to Case ii) completes the theorem in the case $\chi_1 \leq \chi_2$.

The optimal sum-rate in all the three cases is given by the unified expression

$$\begin{aligned} &2E\left[R_1\right] + 2E\left[R_2\right] \\ &\xrightarrow{\rho\to\infty} f\left(\frac{2\sqrt{\chi_1}}{\chi_1+1}\right) + f\left(\frac{2\sqrt{\chi_2}}{\chi_2+1}\right) + \Big|\log(\chi_1) - \log(\chi_2)\Big| \end{aligned} \tag{252}$$

where $f(\bullet)$ is as defined in (160). This expression can be simplified as in the statement of the theorem. ■

I) Comparison of Proof Techniques of Theorems 2 and 3: We first show that Theorem 2 reduces to Theorem 3 under the

$$\mathcal{L}_2 \triangleq \frac{(\chi_1-1)(\sin^2(\theta)+\sin^2(\phi))+2}{(\chi_1-1)\sin^2(\phi)+1} \cdot \frac{(\chi_2-1)(\sin^2(\theta)+\sin^2(\phi))+2}{(\chi_2-1)\sin^2(\theta)+1} \tag{242}$$

$$= \frac{(\chi_1-1)(\chi_2-1)\left[\sin^2(\theta)+\sin^2(\phi)\right]^2 + 2\left(\chi_1+\chi_2-2\right)\left[\sin^2(\theta)+\sin^2(\phi)\right] + 4}{\left[(\chi_1-1)\sin^2(\phi)+1\right]\left[(\chi_2-1)\sin^2(\theta)+1\right]} \tag{243}$$

$$2E\left[R_1\right] + 2E\left[R_2\right] + 4\log(2) \leq g\left(\frac{2\sqrt{\chi_1}\cdot\sin(\theta+\phi)}{(\chi_1-1)(\sin^2(\theta)+\sin^2(\phi))+2}\right) + g\left(\frac{2\sqrt{\chi_2}\cdot\sin(\theta+\phi)}{(\chi_2-1)(\sin^2(\theta)+\sin^2(\phi))+2}\right) + 2\log(\mathcal{L}_2) \tag{247}$$

$$\leq g\left(\frac{2\sqrt{\chi_1}}{\chi_1+1}\right) + g\left(\frac{2\sqrt{\chi_2}}{\chi_2+1}\right) + 2\log(\mathcal{L}_2) \tag{248}$$

$$2E\left[R_1\right] + 2E\left[R_2\right] + 4\log(2) \leq g\left(\frac{2\sqrt{\chi_1}\sin(\nu)}{(\chi_1-1)\sin^2(\nu)+2}\right) + g\left(\frac{2\sqrt{\chi_2}\sin(\nu)}{(\chi_2-1)\sin^2(\nu)+2}\right) + 2\log\left(\frac{\left[(\chi_1-1)\sin^2(\nu)+2\right]\cdot\left[(\chi_2-1)\sin^2(\nu)+2\right]}{\left[(\chi_1-1)\sin^2(\nu-\theta)+1\right]\cdot\left[(\chi_2-1)\sin^2(\theta)+1\right]}\right) \triangleq \mathcal{L}_3 \tag{251}$$

assumption that the eigenvectors of $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ coincide. For this, we set

$$\alpha' = \frac{\alpha\sqrt{\tau_1}}{X},\ \beta' = \frac{\beta\sqrt{\tau_2}}{X},\ \gamma' = \frac{\gamma\sqrt{\tau_1}}{Y},\ \delta' = \frac{\delta\sqrt{\tau_2}}{Y} \quad (253)$$

where X and Y are as in (190) and (191), respectively. Note that the above transformation is a bijection from the space $\left\{\alpha,\beta,\gamma,\delta : |\alpha|^2+|\beta|^2 = 1 = |\gamma|^2+|\delta|^2\right\}$ to the space $\left\{\alpha',\beta',\gamma',\delta' : |\alpha'|^2+|\beta'|^2 = 1 = |\gamma'|^2+|\delta'|^2\right\}$. With this transformation, it can be checked that (194) reduces to (232). It can also be seen that the sum-rate expression in (45) reduces to that in (66) in both cases.

The technique pursued in the general case diverges from that in Appendix F in two ways.

Difference 1: It can be easily checked that $\tau_3 = 0$ if and only if the set of eigenvectors of $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ coincide. In general, $\tau_3 \neq 0$ and $\mathsf{arg}(\tau_3)$ could affect the sum-rate optimization. The first step in Appendix E is to show that this is not the case and $\mathsf{arg}(\tau_3)$ plays no role in the optimization.

This is done by exploiting the fact that the sum-rate optimization [see (14)] is a problem over $\mathcal{G}(2,1)$ and not over $\mathsf{St}(2,1)$.

Difference 2: The second complication is that there is a definitive (and easily classifiable) comparative relationship between $\tau_1 = \mathbf{v}_1^H\mathbf{\Sigma}_2^{-1}\mathbf{v}_1$ and $\tau_2 = \mathbf{v}_2^H\mathbf{\Sigma}_2^{-1}\mathbf{v}_2$ in the special case. This comparative relationship does not generalize to the setting where the eigenvectors of $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ are different.

Specifically, under Case i) of the discussion in Theorem 3, $\tau_1 > \tau_2$ if and only if $\chi_2 < 1$ whereas under Case ii), $\tau_1 > \tau_2$ if and only if $\chi_2 > 1$. On the other hand, in the general case, all the three possibilities: 1) $\tau_1 > \tau_2$, 2) $\tau_1 < \tau_2$, and 3) $\tau_1 = \tau_2$ can occur for appropriate choices of $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$. We now illustrate this with a numerical example. Let $\mathbf{\Sigma}_2$ be fixed such that

$$\mathbf{\Sigma}_2 = \begin{bmatrix} 1 & -0.6897 \\ -0.6897 & 1 \end{bmatrix}. \quad (254)$$

With the choice

$$\mathbf{\Sigma}_1 = \begin{bmatrix} 1 & 0.8 \\ 0.8 & 1 \end{bmatrix}, \quad (255)$$

it can be seen that $\eta_1 = 5.8, \eta_2 = 0.1184, \tau_1 = 3.2222$ and $\tau_2 = 0.5918$, whereas if

$$\mathbf{\Sigma}_1 = \begin{bmatrix} 1 & -0.8 \\ -0.8 & 1 \end{bmatrix}, \quad (256)$$

it can be seen that $\eta_1 = 1.0653, \eta_2 = 0.6444, \tau_1 = 0.5918$ and $\tau_2 = 3.2222$. It can be seen that $\eta_1 = 1.4603, \eta_2 = 0.5397$ and $\tau_1 = \tau_2 = 1.90725$ if $\mathbf{\Sigma}_1$ satisfies

$$\mathbf{\Sigma}_1 = \begin{bmatrix} \frac{2}{3} & -0.34485 \\ -0.34485 & \frac{1}{3} \end{bmatrix}. \quad (257)$$

These differences imply that, in general, there exists no bijective transformation [as in (253)] to transform the objective function from the form in (194) to that in (232). Despite these issues, it would be of interest to pursue a theme that could unify the general case with the special case. ■

J) Proof of Proposition 4: The proof in the low-SNR extreme is obvious. In the high-SNR extreme, we first note that the optimization problem over the choice of a pair $(\mathbf{w}_1, \mathbf{w}_2)$ that results in a corresponding choice of (A_i, B_i, C_i) can be recast in the form of a two-parameter optimization problem over (M_i, N_i) with $M_i = \frac{A_i}{B_i}$ and $N_i = \frac{C_i}{B_i}$ under the constraint that

$$0 \leq N_i^2 \leq M_i \leq \chi_i = \frac{\lambda_1(\mathbf{\Sigma}_i)}{\lambda_2(\mathbf{\Sigma}_i)}. \quad (258)$$

For this, observe that for any given choice of $(\mathbf{w}_1, \mathbf{w}_2)$, the resultant (A_i, B_i, C_i) has to satisfy $C_i^2 \leq A_iB_i$ (Cauchy–Schwarz inequality) and $\frac{A_i}{B_i} \leq \chi_i$ (Ritz–Raleigh ratio) [44].

Thus, we have

$$\max_{\mathbf{w}_1,\mathbf{w}_2} \lim_{\rho\to\infty} E[R_i] \leq \max_{0\leq N_i^2\leq M_i\leq\chi_i} \lim_{\rho\to\infty} E[R_i]. \quad (259)$$

Since the high-SNR expression for $E[R_i]$ satisfies

$$2E[R_i] + 2\log(2) = g\left(\frac{2\sqrt{M_i - N_i^2}}{M_i+1}\right) + 2\log(1+M_i), \quad (260)$$

and $g(\bullet)$ is an increasing function, optimization over N_i which affects only the first term on the right-hand side implies that the optimal choice of N_i is zero. Plugging this choice and using the structure of $g(\bullet)$ and $f(\bullet)$ from (159) and (160), respectively, we have

$$E[R_i] \leq \frac{M_i}{|M_i-1|}\cdot|\log(M_i)|. \quad (261)$$

The right-hand side of (261) is increasing in M_i since the derivative function satisfies

$$\frac{d\frac{M_i\cdot|\log(M_i)|}{|M_i-1|}}{dM_i} = \begin{cases} \frac{M_i-1-\log(M_i)}{(M_i-1)^2} & \text{if } M_i > 1 \\ \frac{1}{2} & \text{if } M_i = 1 \\ \frac{\log\left(\frac{1}{M_i}\right)-(1-M_i)}{(1-M_i)^2} & \text{if } M_i < 1, \end{cases} \quad (262)$$

where all the three pieces are positive and hence, the derivative function is smooth. The goal of maximizing M_i under the constraints of $N_i = 0$ and unit-normedness of $\mathbf{w}_1$ and $\mathbf{w}_2$ is met by the choice as in the statement of the proposition. This upper bound to $E[R_i]$ is also met by the same choice of beamforming vectors and this choice is thus optimal.

Recall from (158) (the high-SNR expression) that $E[R_i]$ is the sum of two terms. The increasing nature of $g(\bullet)$ means that the first term of (158) is maximized when $d_{\mathbf{\Sigma}_i}(\mathbf{w}_1, \mathbf{w}_2)$ is maximized. That is, by the choice $\{\mathbf{w}_1, \mathbf{w}_2\}$ as in (169) and (170). On the other hand, the second term as well as $E[R_i]$ (which is the sum of the two terms) are maximized by the choice in (78). With this choice of beamforming vectors, $d_{\mathbf{\Sigma}_i}(\cdot,\cdot)$ can be written as

$$d_{\mathbf{\Sigma}_i}(\mathbf{w}_{i,\mathsf{opt}}, \mathbf{w}_{j,\mathsf{opt}}) = \frac{2\sqrt{\chi_i}}{\chi_i+1}. \quad (263)$$

Note that $d_{\mathbf{\Sigma}_i}(\cdot,\cdot)$ decreases as χ_i increases. ■

REFERENCES

[1] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.

[2] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2632–2639, Sep. 2001.

[3] A. J. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 684–702, Jun. 2003.

[4] D. Gesbert, H. Bölcskei, D. A. Gore, and A. J. Paulraj, "Outdoor MIMO wireless channels: Models and performance prediction," *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 1926–1934, Dec. 2002.

[5] V. V. Veeravalli, Y. Liang, and A. M. Sayeed, "Correlated MIMO Rayleigh fading channels: Capacity, optimal signaling and asymptotics," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2058–2072, Jun. 2005.

[6] A. M. Tulino, A. Lozano, and S. Verdú, "Impact of antenna correlation on the capacity of multiantenna channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2491–2509, Jul. 2005.

[7] Q. H. Spencer, C. B. Peel, A. L. Swindlehurst, and M. Haardt, "An introduction to the multi-user MIMO downlink," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 60–67, Oct. 2004.

[8] D. Gesbert, M. Kountouris, R. W. Heath, Jr, C.-B. Chae, and T. Salzer, "Shifting the MIMO paradigm: From single user to multiuser communications," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 36–46, Oct. 2007.

[9] B. Hassibi and M. Sharif, "Fundamental limits in MIMO broadcast channels," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1333–1344, Sep. 2007.

[10] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2845–2866, Jun. 2010.

[11] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.

[12] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector Gaussian broadcast channel and downlink-uplink duality," *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 1912–1921, Aug. 2003.

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

[14] N. Jindal, S. Vishwanath, and A. J. Goldsmith, "On the duality of Gaussian multiple-access and broadcast channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 5, pp. 768–783, May 2004.

[15] W. Yu and J. M. Cioffi, "Sum capacity of Gaussian vector broadcast channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1875–1892, Sep. 2004.

[16] M. Schubert and H. Boche, "Solution of multiuser downlink beamforming problem with individual SINR constraint," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 18–28, Jan. 2004.

[17] 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.

[18] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 3, pp. 439–441, May 1983.

[19] R. F. H. Fischer, C. Windpassinger, A. Lampe, and J. B. Huber, "MIMO precoding for decentralized receivers," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun.-Jul. 2002, pp. 496–496.

[20] M. Joham, W. Utschick, and J. Nossek, "Linear transmit processing in MIMO communications systems," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2700–2712, Aug. 2005.

[21] F. Boccardi, F. Tosato, and G. Caire, "Precoding schemes for the MIMO-GBC," in *Proc. Int. Zurich Semin. Commun.*, Feb. 2006, pp. 10–13.

[22] B. M. Hochwald, C. B. Peel, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication—Part II: Perturbation," *IEEE Trans. Commun.*, vol. 53, no. 3, pp. 537–544, Mar. 2005.

[23] T. Yoo and A. J. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 528–541, Mar. 2006.

[24] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multi-user MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, Feb. 2004.

[25] A. Wiesel, Y. C. Eldar, and S. Shamai, "Zero forcing precoding and generalized inverses," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4409–4418, Sep. 2008.

[26] C.-B. Chae, D. Mazzarese, N. Jindal, and R. W. Heath, Jr, "Coordinated beamforming with limited feedback in the MIMO broadcast channel," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1505–1515, Oct. 2008.

[27] C.-B. Chae, S. Shim, and R. W. Heath, Jr, "Block diagonalized vector perturbation for multi-user MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4051–4057, Nov. 2008.

[28] V. Raghavan, J. H. Kotecha, and A. M. Sayeed, "Why does the Kronecker model result in misleading capacity estimates?," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4843–4864, Oct. 2010.

[29] T. Y. Al-Naffouri, M. Sharif, and B. Hassibi, "How much does transmit correlation affect the sum-rate scaling of MIMO Gaussian broadcast channels?," *IEEE Trans. Commun.*, vol. 57, no. 2, pp. 562–572, Feb. 2009.

[30] M. Kountouris, D. Gesbert, and L. Pittman, "Transmit correlation-aided opportunistic beamforming and scheduling," in *Proc. Eur. Signal Process. Conf.*, Sep. 2006, pp. 2409–2414.

[31] D. Hammarwall, M. Bengtsson, and B. E. Ottersten, "Acquiring partial CSI for spatially selective transmission by instantaneous channel norm feedback," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1188–1204, Mar. 2008.

[32] D. Hammarwall, M. Bengtsson, and B. E. Ottersten, "Utilizing the spatial information provided by channel norm feedback in SDMA systems," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3278–3293, Jul. 2008.

[33] B. Clerckx, G. Kim, and S. Kim, "Correlated fading in broadcast MIMO channels: curse or blessing?," in *Proc. IEEE Global Telecommun. Conf.*, Dec. 2008, pp. 3830–3834.

[34] A. Wiesel, Y. C. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2006.

[35] V. S. Annapureddy and V. V. Veeravalli, "Sum capacity of MIMO interference channels in the low interference regime," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2565–2581, May 2011.

[36] V. Raghavan and S. V. Hanly, "Statistical beamformer design for the two-antenna interference channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2010, pp. 2278–2282.

[37] M. Kobayashi and G. Caire, "An iterative water-filling algorithm for maximum weighted sum-rate of Gaussian MIMO-BC," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1640–1646, Aug. 2006.

[38] N. Merhav, G. Kaplan, A. Lapidoth, and S. Shamai, "On information rates for mismatched decoders," *IEEE Trans. Inf. Theory*, vol. 40, no. 6, pp. 1953–1967, Jun. 1994.

[39] A. Edelman, T. Arias, and S. Smith, "The geometry of algorithms with orthogonality constraints," *SIAM J. Matrix Anal. Appl.*, vol. 20, no. 2, pp. 303–353, Apr. 1999.

[40] B. Hassibi and T. L. Marzetta, "Multiple-antennas and isotropically random unitary inputs: The received signal density in closed form," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1473–1484, Jun. 2002.

[41] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, 10th ed. Gaithersburg, MD, USA: National Bureau of Standards, 1972.

[42] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 4th ed. New York, NY, USA: Academic, 1965.

[43] G. W. Stewart, *Matrix Algorithms: Eigensystems*. Philadelphia, PA, USA: SIAM, 2001.

[44] C. R. Rao and M. B. Rao, *Matrix Algebra and its Applications to Statistics and Econometrics*. Singapore: World Scientific, 1998.

[45] V. Raghavan, V. V. Veeravalli, and R. W. Heath, Jr, "Reduced rank signaling in spatially correlated MIMO channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2007, pp. 1081–1085.

[46] V. Raghavan, A. M. Sayeed, and V. V. Veeravalli, "Semiunitary precoding for spatially correlated MIMO channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 3, pp. 1284–1298, Mar. 2011.

[47] Agilent Application Note MIMO Performance and Condition Number in LTE Test Agilent Technologies, Santa Clara, CA, USA, Tech. Rep. 5990-4759EN, 2009 [Online]. Available: http://cp.literature.agilent.com/litweb/pdf/5990-4759EN.pdf

[48] *Technical specification Group Radio Access Network, E-UTRA, Physical layer procedures (Rel. 8)*, 3GPP TS 36.213 V8.8.0, Sep. 2009.

[49] H. Holma and A. Toskala, *LTE for UMTS: OFDMA and SC-FDMA Based Radio Access*. New York, NY, USA: Wiley, 2009.

[50] A. Ghosh, J. Zhang, J. G. Andrews, and R. Muhamed, *Fundamentals of LTE*. Englewood Cliffs, NJ, USA: Prentice-Hall, 2010.

[51] K. K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, "On beamforming with finite rate feedback in multiple antenna systems," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2562–2579, Oct. 2003.

Vasanthan Raghavan (S'01–M'06–SM'11) received the B.Tech degree in Electrical Engineering from the Indian Institute of Technology at Madras, India in 2001, the M.S. and the Ph.D. degrees in Electrical and Computer Engineering in 2004 and 2006 respectively, and the M.A. degree in Mathematics in 2005, all from the University of Wisconsin-Madison, Madison, WI. He is currently at the University of Southern California, Los Angeles, CA. His research interests span multiantenna communication techniques, quickest changepoint detection, quantitative criminology, and random matrix theory.

Stephen V. Hanly (M'98) received his B.Sc. (Hons) and M.Sc. from the University of Western Australia, and the Ph.D. degree in Mathematics in 1994 from Cambridge University, UK. From 1993 to 1995, he was a post-doctoral member of technical staff at AT&T Bell Laboratories. From 1996 to 2009 he was on the research and teaching staff at the University of Melbourne, and between 2010 and 2011 he was an Associate Professor at the National University of Singapore. He is presently a Professor in the Department of Engineering at Macquarie University, Sydney, Australia, where he holds the CSIRO-Macquarie University Chair in Wireless Communications. He was an Associate Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2005–2009, a guest editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, Special issue on "Cooperative Communications in MIMO Cellular Networks" in 2010, and he is presently a guest editor for EURASIP *Journal of Wireless Communications and Networking*, Special issue on "Recent Advances in Optimization Techniques in Wireless Communication Networks." In 2005, he was the technical co-chair for the IEEE International Symposium on Information Theory held in Adelaide, Australia. He was a co-recipient of the best paper award at the Infocom 1998 conference, and the 2001 Joint IEEE Communications Society and Information Theory Society best paper award, both for his work with David Tse. His research interests are in the areas of information theory, signal processing, and wireless networking.

Venugopal V. Veeravalli (S'86–M'92–SM'98–F'06) received the Ph.D. degree in 1992 from the University of Illinois at Urbana-Champaign, the M.S. degree in 1987 from Carnegie-Mellon University, Pittsburgh, PA, and the B.Tech degree in 1985 from the Indian Institute of Technology, Bombay, (Silver Medal Honors), all in Electrical Engineering. He joined the University of Illinois at Urbana-Champaign in 2000, where he is currently a Professor in the department of Electrical and Computer Engineering, and a Research Professor in the Coordinated Science Laboratory. He served as a program director for communications research at the U.S. National Science Foundation in Arlington, VA from 2003 to 2005. He has previously held academic positions at Harvard University, Rice University, and Cornell University.

His research interests include distributed sensor systems and networks, wireless communications, detection and estimation theory, and information theory. He is a Fellow of the IEEE and was on the Board of Governors of the IEEE Information Theory Society from 2004 to 2007. He was an Associate Editor for Detection and Estimation for the IEEE TRANSACTIONS ON INFORMATION THEORY from 2000 to 2003, and an associate editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 1999 to 2000. Among the awards he has received for research and teaching are the IEEE Browder J. Thompson Best Paper Award, the National Science Foundation CAREER Award, and the Presidential Early Career Award for Scientists and Engineers (PECASE).