

Joint Optimization of the Transmit Covariance and Relay Precoder in General Gaussian Amplify-and-Forward Relay Channels

Ramy H. Gohary, *Senior Member, IEEE*, and Halim Yanikomeroglu, *Senior Member, IEEE*

Abstract—The maximum data rate that can be achieved by the strictly causal full-duplex amplify-and-forward (AF) scheme in general Gaussian relay channels is achieved by Gaussian codebooks and can be cast as the solution of an optimization problem of the input transmit covariance and relay precoder. This problem possesses an intricate nonconvex structure and is hence difficult to solve. To circumvent this difficulty, the relay precoder is assumed to be given and then the Karush–Kuhn–Tucker conditions are used to obtain closed form expressions for the optimal input covariance corresponding to that precoder. These expressions are used to show that subdiagonal precoders suffice to attain the maximum achievable rate of the AF scheme at any source transmit power. In addition to significantly reducing the effort expended in searching for the optimal relay precoder, this observation enables us to find the optimal precoders at low and high source transmit powers. For asymptotically low transmit powers, the optimal relaying mechanism is shown to possess an interlacing structure, thereby resembling half-duplex operation. In contrast, for asymptotically high transmit powers, it is optimal for the relay to be silent. The asymptotic analysis enables us to develop an explicit formulation for a suboptimal precoder that, at intermediate source transmit powers, are shown numerically to outperform asymptotically optimal precoders.

Index Terms— Full-duplex AF relaying, KKT conditions, Schur basis, tridiagonal matrices.

I. INTRODUCTION

IN typical communication scenarios, the data rate that can be reliably communicated between a wireless source-destination pair can be significantly increased by using a communication-assisting node, known as a relay [1]. Relay operation modes can be classified into full-duplex, in which the relays can transmit and receive information on the same physical channel, and half-duplex, in which the relays transmit and receive information on orthogonal channels. To ensure effective collaboration between the source and the relay, both the transmis-

sion scheme and the relaying mechanism have to be designed jointly for the given channel conditions [2]. For instance, when the source transmitted signal, the signal observed by the relay and the signal observed by the receiver form a Markov chain, the relay channel is said to be physically degraded, and the optimal full-duplex relaying strategy is the decode-and-forward (DF) scheme [3]. On the other hand, when the signal received by the relay is a physically degraded version of the direct component of the signal observed by the receiver, the channel is said to be reversely degraded and the optimal strategy is for the relay to remain silent [3]. The maximum rate that can be achieved on a general relay channel was shown in [3] to be upper-bounded by the cut-set bound derived therein.

The capacity of a relatively broad class of half-duplex relay channels was obtained in [4]. A subset of this class was considered in [5], and therein, a relaying scheme that uses 1-dimensional piecewise linear mapping was proposed. Using the fact that, for half-duplex channels, the transmissions of the source and relay do not interfere, the scheme in [5] was shown to subsume and outperform linear relaying strategies.

In contrast with the half-duplex case, the capacity of general full-duplex relay channels that are neither degraded nor reversely degraded, remains an open problem, and relaying techniques other than DF are able to achieve higher data rates [6]. For instance, under certain channel conditions, the compress-and-forward (CF) relaying scheme developed in [3] was shown to yield higher rates than DF. Another relaying technique is the so-called amplify-and-forward (AF) scheme wherein the relay forwards to the receiver a linearly precoded version of its received signal without decoding it. Hence, this class of relays is typically referred to as being nonregenerative [2]. In comparison with DF and CF, AF relaying is computationally less demanding and is hence suitable for relaying terminals with low processing capabilities. In addition to its relative simplicity, the AF scheme was shown in [6] to outperform both DF and CF relaying schemes under certain channel conditions. Since in the full-duplex case considered herein the transmissions of the source and relay interfere at the destination, the advantages of the nonlinear scheme proposed in [5] for half-duplex systems do not necessarily carry over to this case.

In addition to scalar relay channels in which the source and the relay have one transmit antenna each, and the relay and the destination have one receive antenna each, the use of relays have been proposed for multiple-input multiple-output (MIMO) systems. For instance, bounds on the capacity of MIMO relay systems were studied in [7] and [8], in the presence of a direct

Manuscript received July 27, 2012; revised April 17, 2013; accepted April 30, 2013. Date of publication June 06, 2013; date of current version August 14, 2013. This work was supported in part by BlackBerry (formerly Research-In-Motion (RIM)) and in part by the Natural Sciences and Engineering Research Council of Canada (NSERC). This paper was presented at the 2012 IEEE International Symposium on Information Theory.

The authors are with the Department of Systems and Computer Engineering, Carleton University, Ottawa, ON K1S 5B6, Canada (e-mail: ramy.gohary@gmail.com; halim@sce.carleton.ca).

Communicated by M. Skoglund, 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.2013.2266656

MIMO link between the source and the destination. For the case in which the relay is restricted to operate in a nonregenerative mode and the channel used by the relay is orthogonal to that used by the source, optimal relay precoders are available; see e.g., [9] and [10]. For MIMO relay systems with multicarrier channels, relay designs based on the minimum mean square error criterion are obtained in [11]. Other criteria and approaches for designing relays are available in [2] and references therein.

Motivated by the potential gains and computational efficiency of AF relaying [6], herein we consider the joint design of the input transmit covariance and the relay precoder of an AF relay-assisted communication system with a direct link between the source and destination. Similar to [6], we assume that the relay operates in a strictly causal full-duplex mode wherein the transmissions of the relay are delayed by one symbol duration and the relay uses the same physical channel for transmission and reception, same time and frequency. A related design problem in which the relay precoder is restricted to have a Toeplitz structure was considered in [12]. By allowing the Toeplitz relay precoder to be noncausal, bounds on the rates that can be achieved by the AF relay system were provided in [13].

In contrast with [12] and [13], in this paper we consider general relay precoders that are not necessarily Toeplitz, but that are strictly causal. In other words, we consider the classic full-duplex AF relay channel considered in [14]. Despite being classic, the capacity of this channel is not known, and as mentioned in [6], the joint optimization of the input covariance and relay precoder for a block length greater than two is a “daunting” task. Indeed, it is an intricate nonconvex problem with potentially high dimensionality.

To address this difficulty, we propose a novel technique for performing the joint optimization of the input transmit covariance and the relay precoder. First, we note that when the relay precoder matrix is given, the design problem of the input covariance can be cast in a convex form for which the Karush–Kuhn–Tucker (KKT) conditions are necessary and sufficient for optimality. By analyzing these conditions, we obtain closed form expressions of the optimal input covariance corresponding to an arbitrary relay precoder. Substituting these expressions in the original formulation, it is shown that subdiagonal precoders suffice to achieve the maximum rate of the AF scheme. This observation reduces the number of optimization variables from being quadratic to linear in the block length. In addition to significantly reducing the effort expended in searching for the rate-optimal precoder, restricting attention to subdiagonal precoders enables us to identify precoders that are asymptotically optimal at low and high source transmit powers. At low source transmit powers, these precoders possess an interlacing structure, which, in a sense, contrasts the Toeplitz precoder structure proposed in [12] and [13]. The interlacing structure of relay precoders implies that the relay transmits every other time slot, which resembles, to some extent, half-duplex relaying. At high source transmit powers, it is shown that it is optimal for the relay precoder to be silent. Optimal precoders with interlacing subdiagonal structures are obtained using an efficient 1-dimensional search algorithm. This algorithm also generates precoder-covariance pairs, in which the precoders

are restricted to be rank-1; a restriction that is motivated by the asymptotic analysis. It is shown that, while the rates yielded by asymptotically optimal precoders (with interlacing structure) do not depend on the block length, the rates yielded by rank-1 precoders are block length dependent. Rank-1 precoders are shown to be more advantageous than their interlacing counterparts at intermediate source transmit powers. Another scenario in which increasing the block length was shown to be beneficial is the one in which the relay channel is binary symmetric with two orthogonal components [15]. Hence, unlike the case of Toeplitz precoders [12], increasing the block length is shown to be advantageous in both the scenario considered herein and the one considered in [15].

The paper is organized as follows. The system model and the design problem are described in Section II. Closed form expressions of the optimal input covariance corresponding to a given relay precoder are derived in Section III. In Section IV, the optimization of the relay precoder is considered. We begin in Section IV-A by showing that restricting the precoder to have a subdiagonal structure does not incur loss of optimality. Then in Section IV-B, we provide an alternate formulation of the relay design problem. This formulation is used in Section IV-C to draw insight into the optimal relay precoder structure at low transmit powers. In Section IV-D, we show that relay precoders with an interlacing subdiagonal structure are optimal for asymptotically low source transmit powers, and in Section IV-E we show that in the asymptotic case of high transmit powers, it is optimal for the relay to be silent. In Section IV-F, we provide an efficient algorithm for finding the optimal relay precoder with an interlacing subdiagonal structure and the corresponding input covariance. Precoders generated by this algorithm subsume asymptotically optimal ones and are later shown to outperform them at intermediate source transmit powers. In Section IV-G, the dependence of the AF achievable rate on the block length is studied in the asymptotic cases of low and high source transmit powers. In Section IV-H, we provide an explicit block length dependent formulation for the rate yielded by the optimal rank-1 precoder. This formulation is used to provide a tight lower bound on the rates achieved by full-duplex AF relaying. Section V provides numerical examples, and Section VI provides concluding remarks. For convenience, most of the notation in [6] is adopted herein and most of the proofs are relegated to the appendixes.

Notation: Standard notations are used throughout. Matrices are denoted by regular uppercase letters. For any integer m , the $m \times m$ identity matrix is denoted by I_m . The direct sum operation, and the trace and determinant operators are denoted by \oplus , $\text{Tr}(\cdot)$, and $\det(\cdot)$, respectively. The notation $\mathcal{O}(k^m)$ is used to imply that, for large k , the complexity, given by the number of multiplications, grows as k^m .

II. SYSTEM MODEL AND DESIGN PROBLEM

We consider the classic strictly causal full-duplex AF relaying scheme for the general Gaussian relay channel considered in [6] and [14] and shown in Fig. 1. In this scheme, the source organizes its transmissions in blocks of length k each. The signal vector transmitted by the source is denoted by $X \in \mathbb{R}^k$ and that

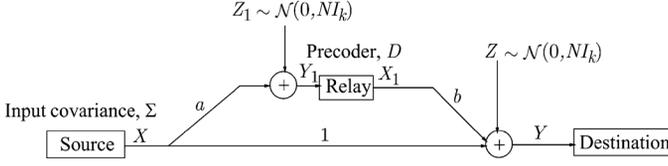


Fig. 1. AF full-duplex Gaussian relay channel [6].

received by the relay is denoted by $Y_1 \in \mathbb{R}^k$. The relay processes Y_1 and generates a signal $X_1 \in \mathbb{R}^k$. To allow for symbol synchronization and signal processing, the relay transmissions depend on its received signal in previous time slots, but not in the current one [6]. Hence, the first entry of X_1 is zero.

The signal vector received at the destination is denoted by $Y \in \mathbb{R}^k$. The channel gain between the source and the destination is normalized to unity, that between the source and the relay is denoted by a , and that between the relay and the destination is denoted by b . The relay received signal can be expressed as

$$Y_1 = aX + Z_1,$$

where $Z_1 \sim \mathcal{N}(0, NI_k)$ is the additive white Gaussian noise at the relay and N is the noise variance. The relay performs linear processing on Y_1 , which is equivalent to left multiplying Y_1 by a strictly lower triangular precoder matrix D . The relay output matrix is $X_1 = DY_1$, whence the received signal at the destination is given by

$$Y = X + bX_1 + Z = (I_k + abD)X + bDX_1 + Z, \quad (1)$$

where $Z \sim \mathcal{N}(0, NI_k)$ is the additive white Gaussian noise at the destination; Z_1 and Z are statistically independent. The source and relay transmit power budgets for a block of k channel uses are denoted by c_1 and c_2 , respectively.

The maximum rate that can be achieved by this AF scheme is given by [6] $R_{AF} = \sup_k R_{AF}^{(k)}$, where $R_{AF}^{(k)} = \frac{1}{k} \mathcal{C}_k$ is the maximum rate that can be achieved by this AF scheme with block length k , and \mathcal{C}_k is the supremum of the mutual information between the k -dimensional vectors X and Y over all input probability distributions satisfying the linear relaying constraint and the power budget constraints, i.e.,

$$\mathcal{C}_k = \sup_{X_1=DY_1, \mathbb{E}\{\text{Tr}(XX^T)\} \leq c_1, \mathbb{E}\{\text{Tr}(X_1X_1^T)\} \leq c_2} I(X; Y). \quad (2)$$

In Section IV-G, it will be shown that $\sup_k R_{AF}^{(k)} = \lim_{k \rightarrow \infty} R_{AF}^{(k)}$.

To facilitate the evaluation of R_{AF} , we will begin by considering the maximization of \mathcal{C}_k for any given k . We will later show that, for asymptotic source power cases, $\frac{1}{k} \mathcal{C}_k$ does not depend on k and hence, for these cases, the limit in the maximum AF rate expression is equal to $\frac{1}{2} \mathcal{C}_2$. Note that, since D is strictly lower triangular, the smallest nontrivial k is equal to 2 and corresponds to 1-dimensional relaying.

To evaluate \mathcal{C}_k for any given k , we note from (1) that, from a rate perspective, the full-duplex Gaussian AF relay channel is equivalent to a classic point-to-point vector Gaussian channel

with colored noise. Since the entropy of Y is maximized when its entries are Gaussian distributed, it follows from [16, Ch. 10] that the maximum rate of the channel in (1) is achieved when the entries of X are drawn from a Gaussian codebook; i.e., $X \sim \mathcal{N}(0, \Sigma)$, where $\Sigma = \mathbb{E}\{XX^T\} \in \mathbb{R}^{k \times k}$ is the input covariance matrix. Hence, restricting attention to the case in which X has a zero-mean Gaussian distribution with covariance Σ does not incur loss of optimality. In this case, the jointly rate-optimal input covariance, Σ , and relay precoder, D , that maximize the mutual information between X and Y can be expressed as the solution of the following optimization problem:

$$\max_{\Sigma, D} \log \frac{\det\left((I_k + abD)\Sigma(I_k + abD^T) + N(I_k + b^2DD^T)\right)}{\det\left(N(I_k + b^2DD^T)\right)}, \quad (3a)$$

$$\text{subject to } \Sigma \succeq 0, \quad \text{Tr}(\Sigma) \leq c_1, \quad (3b)$$

$$\text{Tr}(a^2D\Sigma D^T + NDD^T) \leq c_2, \quad (3c)$$

$$D_{ij} = 0, \quad j \geq i. \quad (3d)$$

The objective in this optimization represents twice the rate that can be reliably communicated over the AF relay channel used over k channel uses. The first constraint in (3b) follows from the fact that Σ is a covariance matrix and hence, must be positive semidefinite. The second constraint in (3b) and the constraint in (3c) ensure that the average source and relay transmit powers during blocks of k channel uses do not exceed prescribed power budgets, c_1 and c_2 , respectively. Finally, the constraint in (3d) ensures strict causality; i.e., that the relay transmissions depend only on its previously received signals.

Despite the intimate relevance of (3) to the maximum data rate that can be achieved by the classic AF relaying scheme, its intricate nonconvex structure conceals intrinsic features of the optimal (Σ, D) pair that solves it. To unravel some of these features, we begin by assuming that a strictly lower triangular relay precoder D is given. We then seek closed form expressions for the optimal covariance matrix Σ . We will later use these expressions to guide the optimization of the precoder matrix D .

III. OPTIMIZING THE TRANSMIT COVARIANCE MATRIX FOR A GIVEN LINEAR RELAY PRECODER

In this section, we will derive closed form expressions for the optimal Σ corresponding to an arbitrary relay precoder matrix D . First, we note that, with D fixed, the objective of the optimization problem in (3) is concave over the positive semidefinite cone given by the first constraint in (3b). The second constraint in (3b) and the constraint in (3c) are linear in Σ . Hence, it is immediate that, for a given D , the optimization problem in (3) is convex. Furthermore, it is easy to verify that for strictly positive c_1 and c_2 , the feasible set is not empty. Hence, Slater's condition holds and the (KKT) conditions are necessary and sufficient for optimality [17].

Let $L(\Sigma, \mu_1, \mu_2, \Phi)$ denote the Lagrangian function corresponding to (3), where μ_1 and μ_2 are the dual variables corresponding to the second constraint in (3b) and the constraint

in (3c), respectively, and Φ is the dual variable matrix corresponding to the first constraint in (3b). Using this notation, we have

$$L(\Sigma, \mu_1, \mu_2, \Phi) = -\log \det(H\Sigma H^T + I_k) + \mu_1(\text{Tr}(\Sigma) - c_1) + \mu_2(\text{Tr}(D\Sigma D^T) - c'_2) - \text{Tr}(\Phi\Sigma), \quad (4)$$

where for notational convenience, we have used the following definitions:

$$H \triangleq \frac{1}{\sqrt{N}}(I_k + b^2 DD^T)^{-1/2}(I_k + abD), \quad \text{and} \quad (5a)$$

$$c'_2 \triangleq \frac{1}{a^2}(c_2 - N \text{Tr}(DD^T)). \quad (5b)$$

Using matrix differentiation expressions from [18] and [19], the derivative of the objective in (3a) can be expressed as $-H^T(I_k + H\Sigma H^T)^{-1}H$. Invoking the matrix inversion lemma [20], this derivative can be cast as $-H^T H(I_k + \Sigma H^T H)^{-1}$, which is a more convenient form for subsequent analysis. Hence, the KKT conditions can be written as

$$\nabla_{\Sigma} L(\Sigma, \mu_1, \mu_2, \Phi) = -H^T H(I_k + \Sigma H^T H)^{-1} + \mu_1 I_k + \mu_2 D^T D - \Phi = 0, \quad (6a)$$

$$\Sigma \succeq 0, \quad \text{Tr}(\Sigma) \leq c_1, \quad \text{Tr}(D\Sigma D^T) \leq c'_2, \quad (6b)$$

$$\Phi \succeq 0, \quad \mu_1 \geq 0, \quad \mu_2 \geq 0, \quad (6c)$$

$$\text{Tr}(\Phi\Sigma) = 0, \quad \mu_1(\text{Tr}(\Sigma) - c_1) = 0, \quad (6d)$$

$$\mu_2(\text{Tr}(D\Sigma D^T) - c'_2) = 0. \quad (6e)$$

The conditions in (6b) ensure the feasibility of the primal problem, those in (6c) ensure the feasibility of the dual problem, and those in (6d) and (6e) represent complementarity slackness.

The matrix $(I_k + \Sigma H^T H)$ in (6a) is nonsingular because, for any $\Sigma \succeq 0$, $\det(I_k + \Sigma H^T H) = \det(I_k + H\Sigma H^T) > 0$. Furthermore, from the definition of the matrix H in (5), it can be seen that $H^T H \succ 0$. Hence, from (6a), we have

$$\det(\mu_1 I_k + \mu_2 D^T D - \Phi) = \frac{\det(H^T H)}{\det(I_k + \Sigma H^T H)} > 0, \quad (7)$$

where the strict inequality follows from the strict positive definiteness of $H^T H$. Using this observation and solving (6a) for Σ , we have

$$\Sigma = (\mu_1 I_k + \mu_2 D^T D - \Phi)^{-1} - (H^T H)^{-1}. \quad (8)$$

The matrix $\mu_1 I_k + \mu_2 D^T D - \Phi = (\Sigma + (H^T H)^{-1})^{-1}$ is positive definite (cf. (7)) and $\Phi \succeq 0$. Hence, we have $\mu_1 I_k + \mu_2 D^T D \succ 0$ but since $D^T D$ is rank deficient because D is strictly lower triangular, it follows that $\mu_1 > 0$, which, together with the second equality in (6d), yields

$$\text{Tr}(\Sigma) = c_1. \quad (9)$$

This implies that the rate-optimal input covariance exhausts all the available source transmit power.

From the first complementarity slackness condition in (6d), we have

$$\text{Tr}\left(\Phi((\mu_1 I_k + \mu_2 D^T D - \Phi)^{-1} - (H^T H)^{-1})\right) = 0. \quad (10)$$

Since Φ and Σ are positive semidefinite, it follows that either $\Phi = 0$ or its columns span the null-space of the columns of $\Sigma = (\mu_1 I_k + \mu_2 D^T D - \Phi)^{-1} - (H^T H)^{-1}$; that is, (10) is equivalent to

$$\left((\mu_1 I_k + \mu_2 D^T D - \Phi)^{-1} - (H^T H)^{-1}\right)\Phi = 0. \quad (11)$$

We now consider possible selections of μ_1 , μ_2 , and Φ that solve (11). First, we note from (8) that choosing μ_1 , μ_2 , and Φ such that $(\mu_1 I_k + \mu_2 D^T D - \Phi) = H^T H$ yields $\Sigma = 0$, which contradicts (9) and hence impossible. We are left with two possibilities: either $\Phi = 0$ or Φ lies in the null space of the matrix $\left((\mu_1 I_k + \mu_2 D^T D - \Phi)^{-1} - (H^T H)^{-1}\right)$. The first case can be considered as a special case of the second one, and hence for brevity, we will focus on the case of $\Phi \neq 0$ and $\mu_2 > 0$. The case of $\mu_2 = 0$ can be inferred from our analysis in a straightforward manner.

A. Analysis for the Case of $\phi \neq 0$ and $\mu_2 > 0$

Let $k_{\Phi} \triangleq \text{rank}(\Phi)$ and $k_{\Sigma} \triangleq \text{rank}(\Sigma)$, where $k_{\Phi} + k_{\Sigma} = k$; cf. (11). For (11) to hold, the vectors of Φ must span the null space of the matrix $\left((\mu_1 I_k + \mu_2 D^T D - \Phi)^{-1} - (H^T H)^{-1}\right)$; that is, if $U_{\Phi} \Lambda_{\Phi} U_{\Phi}^T$ and $U_{\Sigma} \Lambda_{\Sigma} U_{\Sigma}^T$ denote the eigendecompositions of Φ and Σ , respectively, then $U_{\Sigma} = U_{\Phi}^{\perp}$, where U_{Φ}^{\perp} denotes the orthonormal matrix that spans the null space of Φ . Using this notation, we write

$$\Sigma = (\mu_1 I_k + \mu_2 D^T D - \Phi)^{-1} - (H^T H)^{-1} = U_{\Phi}^{\perp} \Lambda_{\Sigma} (U_{\Phi}^{\perp})^T. \quad (12)$$

This equation can be manipulated to yield

$$\Lambda_{\Sigma}^{-1} - (U_{\Phi}^{\perp})^T (\mu_1 I_k + \mu_2 D^T D) (H^T H)^{-1} U_{\Phi}^{\perp} \Lambda_{\Sigma}^{-1} = (U_{\Phi}^{\perp})^T (\mu_1 I_k + \mu_2 D^T D) U_{\Phi}^{\perp}. \quad (13)$$

Using this expression, we have the following result.

Proposition 1: When $\Phi \neq 0$ and $\mu_2 > 0$, the optimal Σ is given by

$$\Sigma = \tilde{U}_1 \left(\nu_1 (I_{k_{\Sigma}} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)^{-1} - (\tilde{U}_1^T H^T H \tilde{U}_1)^{-1} \right) \tilde{U}_1^T, \quad (14)$$

where $\nu_1 = \frac{1}{\mu_1}$, $\nu_2 = \frac{\mu_2}{\mu_1}$ and \tilde{U}_1 denotes the matrix containing the Schur vectors corresponding to the smallest k_{Σ} eigenvalues of the (asymmetric) matrix

$$\tilde{B} = (H^T H)^{-1} (I_k + \nu_2 D^T D). \quad (15)$$

Proof: See Appendix A. ■

The optimal ν_2 can be obtained using the following algorithm, which is illustrated in the flow chart in Fig. 2. The main advantage of this algorithm is that it requires searching over one parameter only; viz., ν_2 , irrespective of k . For each value of

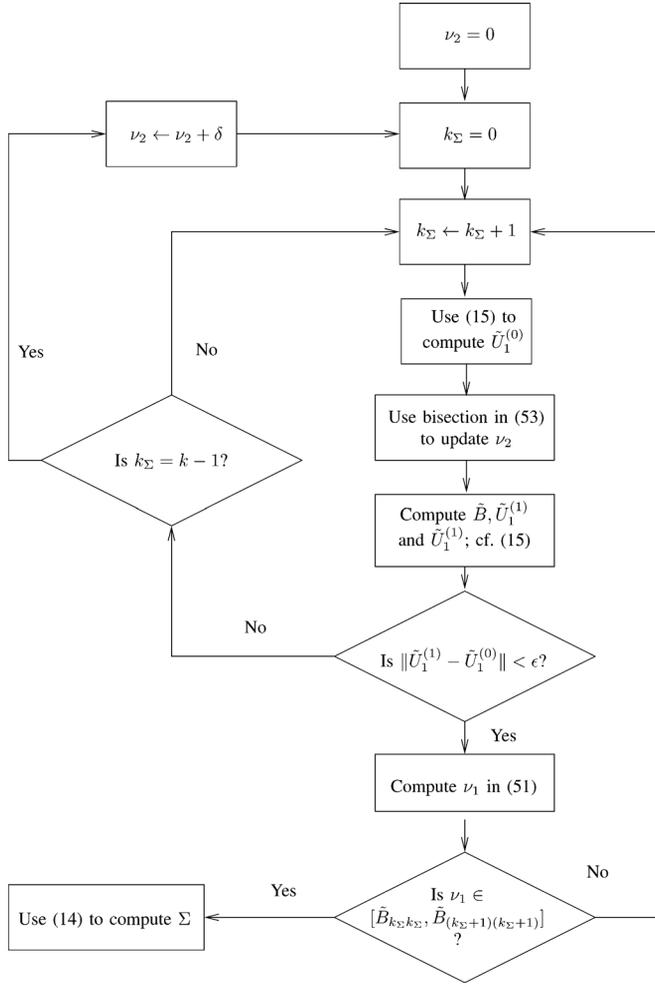


Fig. 2. Flow chart for finding optimal Σ when $\Phi \neq 0$ and $\mu_2 > 0$.

ν_2 , the complexity of obtaining the corresponding Σ is dominated by the computation of the Schur basis, which incurs a complexity of $O(k^3)$. This complexity is considerably less than that of generic interior point methods, which is $O(k^6)$; see e.g., [21].

Algorithm 1 (finding the optimal Σ when $\Phi \neq 0$, $\mu_2 > 0$):

- 1) Start from $\nu_2 = 0$ and $k_\Sigma = 1$.
- 2) Compute $\tilde{U}^{(0)} = [\tilde{U}_1^{(0)} \quad \tilde{U}_2^{(0)}] \in \mathbb{R}^{k \times k}$ to be the Schur decomposition that orders the eigenvalues of \tilde{B} in (15) in a nondecreasing order, where $\tilde{U}_1^{(0)}$ represents that first k_Σ columns of $\tilde{U}^{(0)}$.
- 3) Since $\Phi \neq 0$, the rank of Σ must satisfy $k_\Sigma \leq k - 1$.
 - a) With $\tilde{U}_1^{(0)}$ fixed, update ν_2 using a bisection search on the left-hand side of (53); cf., Lemma 6 in Appendix A.
 - b) Use updated ν_2 to obtain \tilde{B} and update the Schur basis. Denote the updated Schur basis by $\tilde{U}_1^{(1)}$.
 - c) If $\|\tilde{U}_1^{(0)} - \tilde{U}_1^{(1)}\| > \epsilon$, $k_\Sigma \leftarrow k_\Sigma + 1$ and repeat, else proceed to the next step.
 - d) Use (51) to determine ν_1 .
 - e) If the value of ν_1 computed from (51) lies between the k_Σ -th and $k_\Sigma + 1$ -th diagonal entries of \tilde{B} , determine Σ using (14). Otherwise, $k_\Sigma \leftarrow k_\Sigma + 1$, and repeat.

- 4) If no $k_\Sigma < k$ can be found, $\nu_2 \leftarrow \nu_2 + \delta$, where δ is a small step size and repeat. □

Remark 1:

- 1) When $\Phi = 0$, the optimal $k_\Sigma = k$ and $\tilde{U}_1 = I_k$. In that case, using Lemma 6 in Appendix A, Algorithm 1 can be shown to converge exponentially to the optimal ν_2 .
- 2) When $\nu_2 = \frac{\mu_2}{\mu_1} = 0$, Algorithm 1 can be readily modified to yield the optimal μ_1 and k_Σ in at most $k - 1$ steps, each with complexity $\mathcal{O}(k)$. In this case, \tilde{U}_1 contains the eigenvectors corresponding to the largest k_Σ eigenvalues of $H^T H$. □

In the next section, we will use the expressions obtained for the optimal Σ to gain insight into the structure of optimal relay precoders.

IV. OPTIMIZING THE RELAY PRECODER

In Section III, the KKT conditions were used to study the optimization of the input covariance matrix, Σ , when the precoder matrix D is given. The analysis of the KKT system was based on partitioning the feasible region of the dual variables into the following regions, which cover the entire set of admissible dual variables: 1) $\Phi = 0$, $\nu_1 > 0$, $\nu_2 \geq 0$, 2) $\Phi \neq 0$, $\nu_1 > 0$, $\nu_2 = 0$, and 3) $\Phi \neq 0$, $\nu_1 > 0$, $\nu_2 > 0$. The optimal Σ in each of these regions is parametrized by D , ν_1 , and ν_2 . Hence, finding a jointly optimal input covariance-relay precoder pair can be cast as an optimization problem in these variables. The cases in which $\nu_2 = 0$ and $\Phi = 0$ can be inferred from the case in which $\nu_2 > 0$ and $\Phi \neq 0$. Hence, in the forthcoming analysis we will focus on the latter case, and will allude to the other cases as necessary.

When $\nu_2 > 0$ and $\Phi \neq 0$, it was shown that the optimal input covariance matrix Σ corresponding to a given precoder D is given by (14). Substituting this expression for Σ in (3) yields that the joint design problem can be cast in the following form:

$$\max_{D, \nu_1, \nu_2} \log \det(\nu_1 \tilde{U}_1^T H^T H \tilde{U}_1 (I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)^{-1}), \quad (16a)$$

$$\text{subject to } \nu_1 (I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)^{-1} \succeq (\tilde{U}_1^T H^T H \tilde{U}_1)^{-1}, \quad (16b)$$

$$\nu_1 \text{Tr}((I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)^{-1}) - \text{Tr}((\tilde{U}_1^T H^T H \tilde{U}_1)^{-1}) = c_1, \quad (16c)$$

$$\text{Tr}\left(a^2 \tilde{U}_1^T D^T D \tilde{U}_1 (\nu_1 (I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)^{-1} - (\tilde{U}_1^T H^T H \tilde{U}_1)^{-1}) + N D^T D\right) = c_2, \quad (16d)$$

$$\nu_1 \geq 0, \quad \nu_2 \geq 0, \quad D_{ij} = 0, \quad j \geq i, \quad (16e)$$

where from (5),

$$H^T H = \frac{1}{N} (I_k + abD^T) (I_k + b^2 DD^T)^{-1} (I_k + abD). \quad (17)$$

Notice that when $\nu_2 = 0$, the equality in the constraint in (16d) is replaced with inequality " \leq ".

The problem in (16) is nonconvex in D and ν_2 and involves optimization over $\frac{k(k-1)}{2} + 1$ variables. In the next section, we

will show that the number of optimization variables can be reduced to k without loss of optimality. We will then use this simplification to obtain input covariance-relay precoder pairs, i.e., (Σ, D) pairs, that are asymptotically optimal at low and high source transmit powers.

A. Optimality of Subdiagonal Precoders

The algorithm developed in Section III for obtaining the optimal Σ corresponding to a given strictly lower triangular precoder $D \in \mathbb{R}^{k \times k}$ required the search over the dual variable ν_2 . Using this algorithm, the jointly optimal (Σ, D) pair can be determined by searching over the $k(k-1)/2$ degrees of freedom in D , in addition to ν_2 . The number of design degrees of freedom, and thereby the design complexity, can be significantly reduced using the following theorem.

Theorem 1: Relay precoders with the following subdiagonal structure are rate-optimal for Gaussian strictly causal full-duplex AF relaying

$$D = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ \pm\sigma_1 & 0 & \cdots & 0 & 0 \\ 0 & \pm\sigma_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \pm\sigma_{k-1} & 0 \end{bmatrix}, \quad (18)$$

where $\{\sigma_i\}_{i=1}^{k-1}$ are the singular values of D .

Proof: See Appendix D. ■

From this theorem and the fact that the dual variable ν_1 can be readily eliminated from (16), it can be seen that the optimal (Σ, D) pair can be determined by searching over ν_2 and the $k-1$ degrees of freedom of precoders of the form in (18). In other words, using Theorem 1, finding the jointly optimal (Σ, D) pair reduces the search from $k(k-1)/2 + 1$ variables to k variables. Although this reduction in the search space is quite significant, solving (16) for large k can be quite difficult. To circumvent this difficulty, in Sections IV-C and IV-D we will show how restricting D to be subdiagonal will enable us to obtain solutions of the optimization problem in (16) that are asymptotically optimal at low source transmit powers. In Section IV-E we will show that, for the other asymptote of high source transmit powers, setting $D = 0$ is optimal.

B. Equivalent Formulation for Designing Relay Precoders

In order to draw insight into the structure of the optimal precoder, we will develop an equivalent, but more convenient formulation than the one in (16).

We begin by making the following observation.

Remark 2: For any given ν_1 and ν_2 , either the problem (16) is infeasible or, at the optimal solution, the constraint in (16b) is redundant.

Proof: See Appendix E. ■

This remark implies that, for given $\nu_1, \nu_2 \geq 0$, if the optimization problem in (16) is feasible, the corresponding optimal D obtained with the constraint in (16b) eliminated is the same as that obtained if this constraint were present. Hence, for appropriate values of ν_1 and ν_2 , this constraint can be eliminated.

This observation will enable us to formulate an auxiliary optimization problem for the asymptotic case considered in the next section.

We now consider the objective in (16a) and the constraint in (16d). For the objective, substituting for Σ from (14) yields

$$\begin{aligned} \log \det(H\Sigma H^T + I_k) &= \log \det(H^T H \Sigma + I_k) \\ &= \log \det \left(\tilde{U}_1^T H^T H \tilde{U}_1 \left(\nu_1 (I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)^{-1} \right. \right. \\ &\quad \left. \left. - (\tilde{U}_1^T H^T H \tilde{U}_1)^{-1} \right) + I_{k_\Sigma} \right) \\ &= \log \det \left(\nu_1 (I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)^{-1} \tilde{U}_1^T H^T H \tilde{U}_1 \right). \end{aligned} \quad (19)$$

Since \tilde{U}_1 represents the first k_Σ vectors of the Schur basis of the matrix $\tilde{B} = (H^T H)^{-1}(I_k + \nu_2 D^T D)$, we have from the first statement of Lemma 5 in Appendix A that

$$\begin{aligned} (I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)^{-1} \tilde{U}_1^T H^T H \tilde{U}_1 \\ = \left(\tilde{U}_1^T (H^T H)^{-1} (I_k + \nu_2 D^T D) \tilde{U}_1 \right)^{-1}. \end{aligned} \quad (20)$$

Substituting from (20) in (19) yields

$$\begin{aligned} \log \det(H\Sigma H^T + I_k) &= k_\Sigma \log \nu_1 \\ &\quad - \log \det \left(\tilde{U}_1^T (H^T H)^{-1} (I_{k_\Sigma} + \nu_2 D^T D) \tilde{U}_1 \right). \end{aligned} \quad (21)$$

The matrix $\tilde{U}_1^T (H^T H)^{-1} (I_k + \nu_2 D^T D) \tilde{U}_1$ is upper triangular, and its diagonal entries are the smallest k_Σ eigenvalues of $\tilde{B} = (H^T H)^{-1} (I_k + \nu_2 D^T D)$.

Now, we consider the constraint in (16d), this constraint can be expressed as

$$\begin{aligned} a^2 \text{Tr} \left(\tilde{U}_1^T D^T D \tilde{U}_1 \left(\nu_1 (I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)^{-1} \right. \right. \\ \left. \left. - (\tilde{U}_1^T H^T H \tilde{U}_1)^{-1} \right) \right) + N \text{Tr}(D^T D) = c_2, \end{aligned} \quad (22)$$

which is equivalent to

$$\begin{aligned} \text{Tr} \left((I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1 - I_{k_\Sigma}) \left(\nu_1 (I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)^{-1} \right. \right. \\ \left. \left. - (\tilde{U}_1^T H^T H \tilde{U}_1)^{-1} \right) \right) = \frac{\nu_2}{a^2} (c_2 - N \text{Tr}(D^T D)). \end{aligned} \quad (23)$$

The left-hand side of this inequality can be expressed as follows:

$$\begin{aligned} \nu_1 k_\Sigma - \text{Tr} \left((\tilde{U}_1^T H^T H \tilde{U}_1)^{-1} (I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1) \right) \\ - \text{Tr} \left(\nu_1 (I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)^{-1} - (\tilde{U}_1^T H^T H \tilde{U}_1)^{-1} \right) \\ = \nu_1 k_\Sigma - c_1 - \text{Tr} \left(\tilde{U}_1^T (H^T H)^{-1} (I_k + \nu_2 D^T D) \tilde{U}_1 \right), \end{aligned} \quad (24)$$

where the second and last terms on the right-hand side of this equality follow from invoking (16a) and (20), respectively.

Using (24) to substitute for the left-hand side of (23) yields the constraint

$$\begin{aligned} \text{Tr} \left(\tilde{U}_1^T (H^T H)^{-1} (I_k + \nu_2 D^T D) \tilde{U}_1 \right) \\ = \nu_1 k_\Sigma - c_1 - \frac{\nu_2}{a^2} (c_2 - N \text{Tr}(D^T D)). \end{aligned} \quad (25)$$

$\nu_1 k_\Sigma - c_1$ is constant. Using Remark 2 and assuming that, for the given ν_1 , the problem in (26) is feasible, the constraint in (26b) can be dropped without loss of optimality.

Noting that for any $\nu_2 \geq 0$, the matrix $\tilde{U}_1^T \tilde{B} \tilde{U}_1$ is upper triangular, it can be seen, that when $c_1 \rightarrow 0$ and ν_1 is given, the optimization problem in (26) can be cast in the following form:

$$\min_{\tilde{B}} \prod_{i=1}^{k_\Sigma} \lambda_i(\tilde{B}), \quad (29a)$$

$$\text{subject to } \sum_{i=1}^{k_\Sigma} \lambda_i(\tilde{B}) = c, \quad (29b)$$

$$\lambda_{k_\Sigma}(\tilde{B}) \geq \dots \geq \lambda_1(\tilde{B}) \geq \gamma, \quad (29c)$$

where the scalars $c = \nu_1 k_\Sigma - c_1$ and γ are strictly positive because the eigenvalues of \tilde{B} are equal to the eigenvalues of $(I_k + \nu_2 D^T D)^{1/2} (H^T H)^{-1} (I_k + \nu_2 D^T D)^{1/2}$, which is strictly positive definite.

Necessary conditions that must be satisfied by the eigenvalues of $\tilde{B} = (H^T H)^{-1} (I_k + \nu_2 D^T D)$ that solves (29) are given in the following lemma.

Lemma 3: For any given $\gamma > 0$, the solution of the optimization problem in (29) must satisfy

$$\lambda_i(\tilde{B}) = \begin{cases} \gamma & i = 1, \dots, s, \text{ and} \\ \frac{c-s\gamma}{k_\Sigma-s} & i = s+1, \dots, k_\Sigma, \end{cases} \quad (30)$$

for some $s \in \{0, \dots, k_\Sigma\}$.

Proof: See Appendix F. \blacksquare

This lemma asserts that the smallest k_Σ eigenvalues of \tilde{B} must assume at most two distinct values. The smaller value has multiplicity $s \in \{0, \dots, k_\Sigma\}$ and the larger value has multiplicity $k_\Sigma - s$. We will later identify a class of precoder matrices D that yields \tilde{B} that has eigenvalues with the desired multiplicities. Before we do that, we will summarize the results obtained so far for the case of $c_1 \rightarrow 0$. In particular, using Lemma 3 and invoking (47) and (50) in Appendix A, we have the following proposition.

Proposition 2: As $c_1 \rightarrow 0$, the optimal ν_2 goes to 0, and the eigenvalues of the matrix $\tilde{B} = (H^T H)^{-1} (I_k + \nu_2 D^T D)$ that solves (29) and admits a jointly optimal input covariance-relay precoder pair must have the following structure:

- 1) $\lambda_k(\tilde{B}) \geq \dots \geq \lambda_{k_\Sigma+1}(\tilde{B}) > \lambda_{k_\Sigma}(\tilde{B}) \geq \dots \geq \lambda_1(\tilde{B})$, cf. (47) and (50) in Appendix A; and
- 2) $\{\lambda_i(\tilde{B})\}_{i=1}^{k_\Sigma}$ must assume at most two distinct values, cf. Lemma 3. \square

To identify precoder matrices D that yield \tilde{B} with the structure in this proposition, we note from Lemma 2, that for \tilde{B}^{-1} to have eigenvalues with multiplicity s and $k_\Sigma - s$, at least $\max\{s-1, k_\Sigma - s - 1\}$ of the subdiagonal entries of D in (18) must be zero. Since $\min_{s \in \{0, \dots, k_\Sigma\}} \max\{s-1, k_\Sigma - s - 1\} = \lceil k_\Sigma/2 \rceil - 1$, it follows that at least $\lceil k_\Sigma/2 \rceil - 1$ of the subdiagonal entries of D must be equal to zero, and it remains to determine these entries. For the ease of exposition, in the forthcoming analysis we will assume that $k = k_\Sigma + k_\Phi$ is even; the case of k odd can be considered similarly.

D. Asymptotic Optimality of D With Interlacing Subdiagonal Entries

In this section, we will use the conditions of Proposition 2 to determine the structure of asymptotically optimal precoders at low source transmit powers when $\Phi \neq 0$. The case of $\Phi = 0$ follows similarly and is omitted for brevity. The main result of this section is given in the following proposition.

Proposition 3: Relay precoders with interlacing subdiagonal entries are rate-optimal for strictly causal full-duplex AF relaying at asymptotically low source transmit powers, that is, as $c_1 \rightarrow 0$. In particular, such precoders have the form in (18), but with $\sigma_{2i-1} = \sigma_D$ and $\sigma_{2i} = 0$, $i = 1, \dots, \frac{k}{2}$.

Proof: See Appendix G. \blacksquare

Remark 3:

- 1) The analysis in Appendix G implies that the objective depends only on σ_D^2 . Hence, at asymptotically low source transmit powers, the sign of σ_D in Proposition 3 is immaterial.
- 2) The optimal σ_D in Proposition 3 corresponding to the case of $c_1 \rightarrow 0$ can be readily obtained from (83) by setting $\nu_2 = 0$ and finding the roots of the resulting polynomial. \square

In Appendix G, it is shown that when $c_1 \rightarrow 0$, the optimal precoder has rank $k_D = \frac{k}{2}$. However, for nonasymptotic source transmit powers, such a precoder may not be optimal. Restricting attention to the case of precoders with interlacing subdiagonal entries, it is shown that for nonasymptotic cases, the rank of the optimal precoder within that class is either $k_D = \frac{k}{2}$ or $k_D = 1$. In Section IV-F, we will present an efficient algorithm for finding the optimal precoder with interlacing subdiagonal entries in both cases.

E. Optimality of $D = 0$ for Finite c_2 and $c_1 \rightarrow \infty$

In this section, the optimization problem in (16) is studied when c_2 is finite and $c_1 \rightarrow \infty$, that is, when the relay transmit power is finite and the source transmit power is asymptotically high. We will show that, in that case, setting $D = 0$ solves the optimization problem in (16). Toward that end, we note that since $\text{rank}(\Sigma) = k_\Sigma$, the matrix $\nu_1 (I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)^{-1} - (\tilde{U}_1^T H^T H \tilde{U}_1)^{-1}$ in (16d) is full rank; see (14). Using this observation, we have, for any matrix $A \succeq 0$, $A \neq 0$, $\text{Tr}(A(\nu_1 (I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)^{-1} - (\tilde{U}_1^T H^T H \tilde{U}_1)^{-1})) \rightarrow \infty$ as $\text{Tr}(\nu_1 (I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)^{-1} - (\tilde{U}_1^T H^T H \tilde{U}_1)^{-1}) = c_1 \rightarrow \infty$.

Setting $A = \tilde{U}_1^T D^T D \tilde{U}_1$, it can be seen that, as $c_1 \rightarrow \infty$, the left-hand side of (16d) goes to infinity, which violates the constraint in (16d). Thus, when $c_1 \rightarrow \infty$, we must have $\tilde{U}_1^T D^T D \tilde{U}_1 \rightarrow 0$. Setting

$$\tilde{U}_1^T D^T D \tilde{U}_1 = 0, \quad (31)$$

it can be seen that the objective and the constraint of (16) are independent of ν_2 . Hence, setting $\nu_2 = 0$ does not incur loss of optimality and yields $\tilde{B}^{-1} = H^T H$. From Section III, we know that \tilde{U}_1 is the matrix containing the Schur vectors corresponding to the largest k_Σ eigenvalues of \tilde{B}^{-1} . Hence, when $\nu_2 = 0$, \tilde{U}_1 contains the eigenvectors of $H^T H$ corresponding to its largest k_Σ eigenvalues. Using (31) in (17) it can be seen that these are

the same as the eigenvectors corresponding to the smallest k_Σ eigenvalues of DD^T . Since the matrix D is square, the smallest k_Σ eigenvalues of DD^T are the same as the smallest k_Σ eigenvalues of $D^T D$. From (31), these eigenvalues are zero. It follows that $\tilde{U}_1^T H^T H \tilde{U}_1 = \frac{1}{N} \tilde{U}_1^T (I_k + b^2 DD^T)^{-1} \tilde{U}_1 = \frac{1}{N} I_{k_\Sigma}$. In other words, in this asymptotic case the optimization problem in (16) is trivial and is solved by setting $D = 0$.

F. Algorithm for Finding Optimal Precoders With an Interlacing Subdiagonal Structure

In this section, we develop an algorithm for finding the optimal D with interlacing subdiagonal structure. For the asymptotic case of $c_1 \rightarrow 0$, it was shown in Appendix G that such a D is optimal and to find it, the value of r in Appendix G is set to zero and σ_D must be determined. For nonasymptotic cases, finding the optimal D with interlacing subdiagonal structure requires determining both σ_D and r ; r equals either 0 or $\frac{k}{2} - 1$, cf. Appendix G. For $r = 0$, the proposed algorithm performs an exhaustive search over σ_D^2 , and for $\frac{k}{2} - 1$, it performs an exhaustive search over a (ν_1, σ_D^2) grid. The value of r that yields the higher objective is chosen. Note that (75) implies that the rank of the optimal D with interlacing structure is either $k_D = \frac{k}{2}$, corresponding to $r = 0$, or $k_D = 1$, corresponding to $r = \frac{k}{2} - 1$. For σ_D^2 , we note from (3d), that σ_D^2 must satisfy

$$\sigma_D^2 \leq \frac{c_2}{Nk_D}. \quad (32)$$

Using these observations, we have the following algorithm:

Algorithm 2:

- 1) Let N_s be the number of search steps, and define $\Delta\sigma_D^2 = \frac{2c_2}{NkN_s}$. For each $\sigma_D^2 = n_s \Delta\sigma_D^2$, $n_s = 0, \dots, N_s$, consider two possibilities: $k_\Sigma = k_D = \frac{k}{2}$, corresponding to $r = 0$, and $k_\Sigma = k - 1$, $k_D = 1$, corresponding to $r = \frac{k}{2} - 1$. For $k_\Sigma = k_D = \frac{k}{2}$ (corresponding to $r = 0$), perform the following steps:
 - a) Compute $\{\eta_i(c_y(\sigma_D))\}_{i=0}^2$. Solve (83) to obtain tentative values of ν_2 , the positive roots of (83).
 - b) Compute $\lambda_{\tilde{B}_1}^{-1}$ using (71) and $u_{\tilde{B}_1}^2$ using (72).
 - c) Compute ν_1 using (78). Notice that (78) ensures that, in accordance with Remark 2, the constraint in (80) is automatically satisfied.
 - d) Compute the objective in (76).
- For $k_\Sigma = k - 1$ and $k_D = 1$ (corresponding to $r = \frac{k}{2} - 1$), we note that, from (32) and (90), we must have $c_1 + (k - 2)(N - \nu_1) > 0$, which implies that $\nu_1 \leq N + \frac{c_1}{k-2}$. However, from (89), we have $\nu_1 \geq N$. Hence, to search for ν_1 , discretize the interval $\left[N, N + \frac{c_1}{k-2} \right]$. For each pair (σ_D, ν_1) , perform the following steps:
 - a) Compute $\{\eta_i(c_w(\sigma_D, \nu_1))\}_{i=0}^2$. Solve (83) to obtain ν_2 .
 - b) Compute $\lambda_{\tilde{B}_1}^{-1}$ using (71) and $u_{\tilde{B}_1}^2$ using (72).
 - c) If ν_1 satisfies (87), proceed to the next step. Otherwise, discard the current ν_1 and examine the next point in the (σ_D, ν_1) -grid.
 - d) Check if the pair (σ_D, ν_1) satisfies (89). If it does, compute the objective in (86). Otherwise, update the choice of ν_1 from the discretized interval.
- 2) Compare the values of the objectives corresponding to $r = 0$ and $r = \frac{k}{2} - 1$.

- 3) Select r , σ_D , ν_1 , and ν_2 that yield the smallest objective.
- 4) Construct the input covariance using (14). \square

Remark 4 (Asymptotically Optimal Precoders and Precoders Generated by Algorithm 2): Note that for asymptotically optimal precoders (as $c_1 \rightarrow 0$), the optimal r is equal to zero. However, for nonasymptotic cases, the optimal r for precoders with interlacing subdiagonals can be equal to $\frac{k}{2} - 1$. Hence, at asymptotically low transmit powers, the above algorithm yields optimal precoders, but at intermediate powers it yields the optimal rank-1 precoders, corresponding to $r = \frac{k}{2} - 1$. \square

Remark 4 (Algorithm for the Case of $\Phi = 0$): The previous analysis focused on the case of $\Phi \neq 0$. However, finding the globally optimal precoder at asymptotically low transmit powers requires considering the case of $\Phi = 0$. In that case, $\tilde{U}_1 = I_k$, and, similar to the case of $\Phi \neq 0$, the asymptotically optimal precoder possesses the interlacing structure. To consider that case, it is required to search over σ_D^2 as in the case of $r = 0$ in Algorithm 2. The optimal ν_1 , ν_2 and the corresponding objective for this case are provided in Appendix H. \square

G. Independence of the Block Length in Asymptotic Cases

The analysis in the previous sections assumed that the block length k is fixed. We will now provide results for the case in which k is allowed to vary. We begin by recalling that, in [6], \mathcal{C}_k in (2) was argued to form a superadditive sequence in k . Using this observation, the following remark was deduced in [6].

Remark 6: The maximum rate achieved by full-duplex AF relaying is a monotonically increasing function of the block length, k . Furthermore,

$$\sup_k \frac{1}{k} \mathcal{C}_k = \lim_{k \rightarrow \infty} \frac{1}{k} \mathcal{C}_k.$$

Proof: With a slight abuse of notation, let X_j^k be the vector containing the symbols X_j, X_{j+1}, \dots, X_k transmitted by the source and let Y_j^k be the corresponding vector containing the symbols Y_j, Y_{j+1}, \dots, Y_k received by the destination. Invoking the definition of \mathcal{C}_k in (2), we have, for any $k, m > 0$,

$$\begin{aligned} \mathcal{C}_{k+m} &= \sup_{p(x_1, \dots, x_{k+m})} I(X_1^{k+m}; Y_1^{k+m}) \\ &\geq \sup_{p(x_1, \dots, x_k)} I(X_1^k; Y_1^k) \\ &\quad + \sup_{p(x_{k+1}, \dots, x_{k+m})} I(X_{k+1}^{k+m}; Y_{k+1}^{k+m}) \\ &= \mathcal{C}_k + \mathcal{C}_m, \end{aligned} \quad (33)$$

where the probability distributions in (33) and (34) satisfy the constraints in (2). To write the inequality in (34), we used the fact that the channel is memoryless and that factorizable probability distributions form a proper subset of general distributions. Dividing both sides by $k + m$ shows that encoding over a total signaling interval of $k + m$ channel uses yields a rate higher than the average rate yielded by separate encoding over two intervals of k and m channel uses for fractions $\frac{k}{k+m}$ and $\frac{m}{k+m}$ of the total signaling interval, respectively. \blacksquare

Motivated by this remark, we will now investigate the effect on increasing k on $R_{\text{AF}}^{(k)} = \frac{1}{k} \mathcal{C}_k$ in the two asymptotic cases of $c_1 \rightarrow \infty$ and $c_1 \rightarrow 0$. In the latter case, the optimal relay

precoders are rank- $\frac{k}{2}$ and are generated by Algorithm 2 with $r = 0$. The case of $r = \frac{k}{2} - 1$ pertaining to precoders generated by Algorithm 2 for intermediate source transmit powers will be investigated in the next section.

To begin with, we note that in the formulation in (3) the objective is to maximize twice the rate communicated over k channel uses. Hence, the per-channel-use rate is equal to the value of the objective in (3a) divided by $2k$. To study the effect of increasing k on the per-channel-use rate yielded by this formulation, let $c_1 = \tilde{c}_1 k$ and $c_2 = \tilde{c}_2 k$, where \tilde{c}_1 and \tilde{c}_2 are constants independent of k .

In Section IV-E we showed that, for any fixed $k < \infty$, when $c_1 \rightarrow \infty$, the optimal $D = 0$. In this case, it can be immediately seen that the covariance Σ that solves (3) is given by $\tilde{c}_1 J_k$ and $C_k = \frac{k}{2} \log(1 + \frac{\tilde{c}_1}{N})$. Hence, $R_{AF}^{(k)} = \frac{1}{k} C_k$ is independent of k and $R_{AF} = \lim_{k \rightarrow \infty} R_{AF}^{(k)} = \frac{1}{2} \log(1 + \frac{\tilde{c}_1}{N})$, that is, increasing k in this asymptotic case does not yield additional gain.

We will now show a similar result for the asymptotic case of $c_1 \rightarrow 0$, that is, we will show that for this case $\frac{1}{k} C_k$ does not depend on k . Hence, in this asymptotic case too, increasing k does not yield additional gain. To arrive at this result we recall that, when $c_1 \rightarrow 0$, the optimal $r = 0$ and hence, the optimal $k_\Sigma = \frac{k}{2}$. Substituting these values for r and k_Σ and using $k' = k$ and $c_i = \tilde{c}_i k$, $i = 1, 2$ in the objective in (76) and the constraints in (78) and (83) in Appendix G yields that $R_{AF}^{(k)}$ is given by

$$R_a = \max_{\nu_2, \sigma_D \geq 0} \frac{1}{4} \log \left(\frac{\nu_2}{a^2} \lambda_{\tilde{B}_1^{-1}} (2\tilde{c}_2 - N\sigma_D^2) + 2\lambda_{\tilde{B}_1^{-1}} \tilde{c}_1 + 1 \right), \quad (35a)$$

$$\text{subject to } \eta_0(c_y(\sigma_D))\nu_2^2 + \eta_1(c_y(\sigma_D))\nu_2 + \eta_2(c_y(\sigma_D)) = 0, \quad (35b)$$

where $c_y(\sigma_D) = \sigma_D^2(1 + a^2)^2 b^4 + 4a^2 b^2 - \frac{4a^4 b^2 \sigma_D^2 \tilde{c}_1}{2\tilde{c}_2 - N\sigma_D^2}$, and $\lambda_{\tilde{B}_1^{-1}}$ and $\{\eta_i(\cdot)\}_{i=0}^2$ are given in (71) and (84), respectively. Examining (35) reveals that for the asymptotic case of $c_1 \rightarrow 0$, $R_{AF}^{(k)}$ depends only on σ_D , but does not depend on k ; ν_2 can be readily eliminated from this formulation. Hence, we conclude that when $c_1 \rightarrow 0$, it suffices to set $k = 2$; k must be strictly greater than 1 for D to be strictly lower triangular.

Unfortunately, the intricate form of the Schur vectors of the general (asymmetric) tridiagonal matrices that appear in (16) has prohibited us from computing $R_{AF} = \lim_{k \rightarrow \infty} R_{AF}^{(k)}$ at intermediate source transmit powers. For these powers, the optimal D does not necessarily have the interlacing structure. To obtain a lower bound on the maximum AF rate (i.e., $\lim_{k \rightarrow \infty} R_{AF}^{(k)}$), in the next section we will consider the optimal relay precoders generated by Algorithm 2 with $r = \frac{k}{2} - 1$. Setting r to this value yields the optimal rank-1 relay precoder. The rate achieved by such a precoder and the corresponding optimal input covariance matrix is a (non-monotonic) function of k , and, at intermediate source transmit powers, this rate can be significantly higher than the rate achieved by the asymptotically optimal precoder; cf. Example 4 in Section V below.

Remark 7 (1-dimensional AF Relaying): The fact that R_a , the rate generated by asymptotically optimal precoders, is independent of k implies that it suffices for k to be set equal to 2. This corresponds to 1-dimensional relaying, in which the relay

scales its observed signal during one symbol interval and forwards it to the destination during the following symbol interval. Hence, asymptotically optimal relaying is in essence 1-dimensional relaying. \square

H. AF Achievable Rate

In the previous section it was shown that, when $c_1 \rightarrow \infty$, the optimal $D = 0$ and the relay channel reduces to the standard additive white Gaussian noise one. It was also shown that, for the case of $c_1 \rightarrow 0$, the maximum rate that can be achieved by strictly causal full-duplex AF relaying can be expressed as the solution of the, effectively 1-dimensional, optimization problem in (35). This problem corresponds to the case of $k_\Sigma = k_D = \frac{k}{2}$ and is solved by Algorithm 2 with $r = 0$. For intermediate source transmit powers, Algorithm 2 with $r = \frac{k}{2} - 1$ yields optimal rank-1 relay precoders, which are shown in Example 4 in Section V below to yield rates higher than those yielded by the asymptotically optimal precoders generated by the optimization in (35). Note that the latter optimization corresponds to setting $r = 0$, and as shown in Appendix G, setting r to lie in $\{1, \dots, \frac{k}{2} - 2\}$ yields strictly lower rates than setting $r = 0$ and $r = \frac{k}{2} - 1$.

We will now provide explicit expressions for the rate generated by Algorithm 2 with $r = \frac{k}{2} - 1$. Let $R_b^{(k)}$ denote this rate for finite k , and substitute for $k' = k$ and $r = \frac{k}{2} - 1$ in the objective in (86), the expression for $u_{\tilde{B}_1}^2$ in (90), the expressions for $c_z(\sigma_D, \nu_1)$ and $c_w(\sigma_D, \nu_1)$, and the constraint in (87). After simplification, the rate $R_b^{(k)}$ can be expressed as

$$R_b^{(k)} = \max_{\nu_1, \nu_2, \sigma_D \geq 0} \frac{1}{2} \log(N^{-1}\nu_1) + \frac{1}{2k} \log(\lambda_{\tilde{B}_1^{-1}} \nu_1^{-1} N^2), \quad (36a)$$

$$\text{subject to } \nu_1 - N - \tilde{c}_1 - \frac{\nu_2}{a^2} \tilde{c}_2 =$$

$$\frac{1}{k} \left(\nu_1 + \lambda_{\tilde{B}_1^{-1}}^{-1} - 2N - \frac{\nu_2}{a^2} N\sigma_D^2 \right), \quad (36b)$$

$$\eta_0(c_w(\sigma_D, \nu_1))\nu_2^2 + \eta_1(c_w(\sigma_D, \nu_1))\nu_2 + \eta_2(c_w(\sigma_D, \nu_1)) = 0, \quad (36c)$$

where $\lambda_{\tilde{B}_1^{-1}}$ is given in (71) and $\{\eta_i(\cdot)\}_{i=0}^2$ are given in (84), and

$$c_w(\sigma_D, \nu_1) = \sigma_D^2(1 + a^2)^2 b^4 + 4a^2 b^2 - 2\sigma_D^2 a^4 b^2 \frac{k(\tilde{c}_1 + N - \nu_1) - 2(N - \nu_1)}{k\tilde{c}_2 - N\sigma_D^2}.$$

For a given k , (36) constitutes a 2-dimensional optimization problem in ν_2 and σ_D ; ν_1 can be eliminated using (36b). Using the formulation in (36) and the one in Section IV-G yields the following result.

Proposition 4: For the full-duplex Gaussian relay channel in Fig. 1, the rate achieved by strictly causal AF relaying with block length k , $R_{AF}^{(k)}$, satisfies

$$R_{AF}^{(k)} \geq \max\{R_a, R_b^{(k)}\}, \quad (37)$$

where R_a is the rate of optimal 1-dimensional relaying given in (35), and $R_b^{(k)}$ is given in (36). \square

Notice that, as shown in Section IV-G, R_a in (37) does not depend on k . Furthermore, it represents the maximum AF rate at asymptotically low and high values of c_1 .

Remark 6 in the previous section shows that $R_{AF}^{(k)}$ is monotonically increasing in k . However, the intricacy of (36) conceals the effect of increasing k on $R_b^{(k)}$. Using (36), one can easily construct examples to show that, unlike $R_{AF}^{(k)}$, $R_b^{(k)}$ is not monotonically increasing in k . In fact, the formulation in (36) can be used to show that, for $a \geq 1$, $\lim_{k \rightarrow \infty} R_b^{(k)} = \frac{1}{2} \log\left(1 + \frac{\tilde{c}_1}{N}\right)$, which implies that, for sufficiently large k , rank-1 relaying is ineffective. However, in Section V, we will provide examples that show that, with $a \geq 1$, this relaying scheme can offer appreciable rate gain over 1-dimensional AF relaying when k is finite. Collecting these facts, we have the following result.

Corollary 1: The maximum rate that can be achieved by full-duplex AF satisfies

$$R_{AF} = \lim_{k \rightarrow \infty} R_{AF}^{(k)} \geq \max\{R_a, \sup_k R_b^k\}.$$

□

Analytical comparison between R_a , the rate generated by 1-dimensional AF relaying, and $R_b^{(k)}$, the rate corresponding to optimal rank-1 relay precoders seems formidable. Hence, in the next section, this comparison will be performed numerically. We will show that, for intermediate values of \tilde{c}_1 , the rate yielded by rank-1 relay precoders (corresponding to $r = \frac{k}{2} - 1$ in Algorithm 2) can be significantly higher than that yielded by its optimal 1-dimensional counterpart (corresponding to $r = 0$ in Algorithm 2).

V. NUMERICAL EXAMPLES

In this section, we consider the relay channel depicted in Fig. 1. We will provide four numerical examples. In all examples the noise variance N is set equal to 1. In the first example, the performance of the AF scheme in which the input covariance Σ is optimized for a given relay precoder D is compared with the case in which Σ is restricted to be a scaled identity. In the second example, the performance of AF when Σ and D are obtained by Algorithm 2 is compared with the performance of AF when the optimal Σ and D are obtained by exhaustive search over subdiagonal relay precoders. Note that in applying Algorithm 2 both cases of $\Phi = 0$ and $\Phi \neq 0$ must be considered; cf. Remark 4. In the third example, the rate achieved by AF when Σ and D are generated by Algorithm 2 is compared with the cut-set bound and the rate achieved by other relaying schemes including CF and the generalized DF scheme considered in [6]. Finally, in the last example we compare the rate yielded by asymptotically optimal 1-dimensional relaying, which corresponds to setting $r = 0$ in Algorithm 2 and the lower bound in Proposition 4, which is obtained by Algorithm 2 with either $r = 0$ corresponding to 1-dimensional relaying or $r = \frac{k}{2} - 1$ corresponding to multidimensional relaying; cf. Remarks 4 and 5 in Section IV-F and Remark 7 and Corollary 1 in Section IV-G.

Example 1: In this example, we consider the optimization of the input covariance Σ for a given relay precoder, D , when the

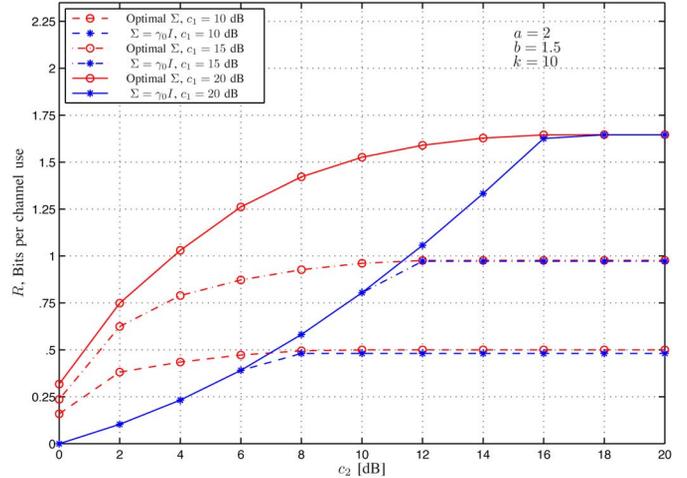


Fig. 3. Optimal input covariance, for $D = \frac{1}{\sqrt{\text{Tr}(\tilde{D}_{10}^T \tilde{D}_{10})}} \tilde{D}_{10}$.

block length, $k = 10$, the source-relay channel gain, $a = 2$, and the relay-destination channel gain, $b = 1.5$. The maximum rates that can be achieved with the optimal Σ for the given D are plotted in Fig. 3 for transmit powers, c_1 , of 10, 15, and 20 dB and a relay precoder $D = \frac{1}{\sqrt{\text{Tr}(\tilde{D}_{10}^T \tilde{D}_{10})}} \tilde{D}_{10}$, where the ij th entry of \tilde{D}_{10} is given by $[\tilde{D}_{10}]_{ij} = 1, i > j$, and $[\tilde{D}_{10}]_{ij} = 0, i \leq j$.

The optimal covariance for the given D is obtained using Algorithm 1. For improved efficiency, we begin by considering simplified versions of Algorithm 1 corresponding to the possibilities of $\Phi = 0$, and $\Phi \neq 0$ and $\nu_2 = 0$. If no solution of the KKT system is found for these cases, Algorithm 1 is run for the more general case of $\Phi \neq 0$ and $\nu_2 > 0$; cf. Remark 1 in Section III-A.

For comparison, the rate obtained by $\Sigma = \gamma_0 I$ is also shown, where γ_0 is chosen such that either the constraint in (3c) or the second constraint in (3b) is satisfied with equality; i.e., $\gamma_0 = \min\left\{\frac{c_1}{k}, \frac{c_2 - N \text{Tr}(D^T D)}{a^2 \text{Tr}(D^T D)}\right\}$.

From Fig. 3, it can be seen that in the low relay transmit power region, the relay power constraint is active and the maximum rate increases with the relay power budget, c_2 . However, the achievable rate ceases to increase after c_2 exceeds a certain threshold. In this region, the relay power constraint is not active and $\nu_2 = 0$. Roughly speaking, the ability of the relay to provide an increase in the achievable rate is constrained by both its power budget and the quality of the signal it receives. Hence, for fixed transmit power budget, c_1 , increasing c_2 beyond a certain threshold does not yield additional rate gain. It can be seen from the figure that, at low relay transmit powers, the optimal Σ yields achievable rates that are significantly greater than those achieved by scaled identity input covariance matrix. □

Example 2: In this example, we consider the joint optimization of the input covariance-relay precoder pair (Σ, D) when the block length, $k = 4$, the source-relay channel gain, $a = 1.25$, and the relay-destination channel gain, $b = 0.7$. The relay power budget c_2 is set equal to 10 dB and the transmitter power is varied between 0 and 20 dB.

In Fig. 4, we plot the rates yielded by the optimal input covariance-relay precoder pair. This pair was obtained using an ex-

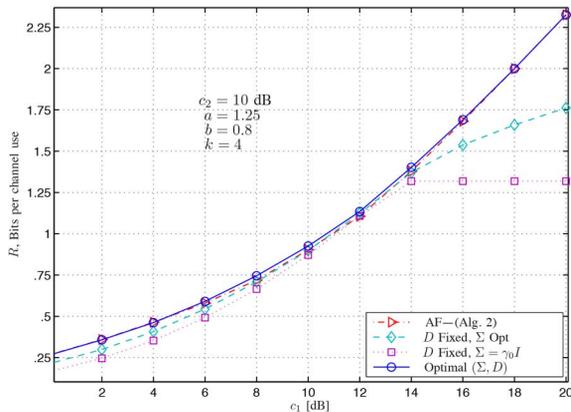


Fig. 4. Comparison between Σ and D obtained by exhaustive search and generated by Algorithm 2.

haustive search over subdiagonal precoders of the form in (18) and applying Algorithm 1 to obtain the corresponding optimal covariance matrix. In Fig. 4, we plot the rates corresponding to the asymptotically optimal relay precoders obtained using Algorithm 2. For comparison, in this figure we also plot the rates yielded by scaled identity input covariance and the optimal covariance corresponding to a Toeplitz subdiagonal precoder [12], [13], i.e., $D_{i+1,i} = 1/\sqrt{3}$, $i = 1, 2, 3$.

From this figure, it can be seen that when the transmit power, c_1 , is low (less than 6 dB in this example) or high (greater than 16 dB), the rate achieved by the asymptotically optimal relay matrices coincides with that achieved by the optimal relay matrices. For high transmit powers, both exhaustive search and Algorithm 2 yield $D = 0$. For intermediate powers, the rate achieved by the precoder generated by Algorithm 2 is less than that achieved by the optimal precoder, but the gap between these rates appears to be generally small. Note that at intermediate powers, the precoders generated by Algorithm 2 are rank 1, i.e., not asymptotically optimal.

It is worth noting that, with the interlacing structure inherent in the asymptotically optimal relays, the relay transmits during one time slot and stays inactive during the next one, which resembles, to some extent, operating in a half-duplex mode. At intermediate powers, the rank-1 precoders generated by Algorithm 2 correspond to the relay being silent during all but one of the k channel uses, that is, the relay concentrates all its energy in one time slot.

Fig. 4 also shows that the rates achieved by the optimal (Σ, D) pair and the (Σ, D) pair generated by Algorithm 2 are significantly higher than those achieved by a fixed D and the corresponding optimal Σ . However, for example, when $c_1 \approx 12$ dB, the rate achieved by the Toeplitz precoder [12], [13] (i.e., with identical subdiagonal entries) is close to that achieved by the optimal precoder and slightly exceeds the rate achieved by the precoder generated by Algorithm 2. \square

Example 3: In the previous example, we showed that the precoders generated by Algorithm 2 yield rates that appear to be relatively close to the rates yielded by optimal precoders. Hence, to reduce the computational effort, in this example, we

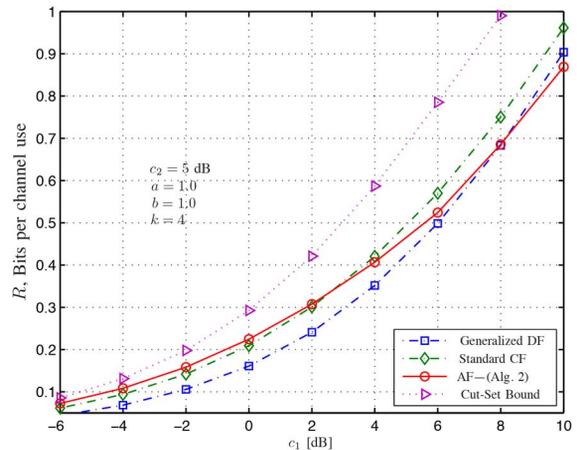


Fig. 5. Comparison between the rate achieved by AF with Σ and D generated by Algorithm 2, the cut-set bound, and the rate achieved by the generalized DF and CF schemes.

only consider the rate achieved by the AF scheme when Σ and D are generated by Algorithm 2. This rate is compared with the cut-set bound and the rate achieved by CF and the generalized DF scheme considered in [6]. Note that, as shown in Sections IV-C and IV-E, and Remarks 4 and 7, Algorithm 2 generates (Σ, D) pairs that are optimal at asymptotically low and high transmit powers, i.e., when $c_1 \rightarrow 0$ and $c_1 \rightarrow \infty$.

The rates achieved by the aforementioned schemes are plotted in Fig. 5 when $a = 1.0$, $b = 1.0$, $k = 4$, and $c_2 = 5$ dB. From this figure, it can be seen that for low transmit powers, AF performs better than both the generalized DF and the CF scheme. At these powers the optimal D assumes the interlacing structure described in Section IV-C, which corresponds to 1-dimensional relaying, and although we considered the case of $k = 4$, setting $k = 2$, at these powers, suffices for the considered AF scheme to attain its maximum rate; cf. Section IV-G.

At intermediate transmit powers, both the generalized DF and CF schemes yield higher rates than those achieved by the precoders generated by Algorithm 2, which, at these powers, correspond to either the asymptotically optimal 1-dimensional relaying, or multidimensional relaying with optimal rank-1 relay precoders. A comparison between these relaying schemes will be presented in Example 4 below for various block lengths.

Finally, numerical comparisons suggest that, in many cases, the gap between the better of the CF and the generalized DF, and the AF scheme with precoders generated by Algorithm 2 is relatively small (less than 10% for this example). This observation reinforces the case for using AF over either DF and CF for applications with stringent complexity and delay constraints, even at moderate source transmit powers. \square

Example 4: This example provides a comparison between asymptotically optimal (1-dimensional) relay precoders which correspond to setting $r = 0$ in Algorithm 2 and the lower bound in Proposition 4, which is obtained by Algorithm 2 with either $r = 0$, corresponding to 1-dimensional relaying, or $r = \frac{k}{2} - 1$, corresponding to multidimensional relaying with optimal rank-1 relay precoders. The rates achieved by these precoders are plotted in Fig. 6 when $a = 20.0$, $b = 4.0$, $c_2 = 5$ dB,

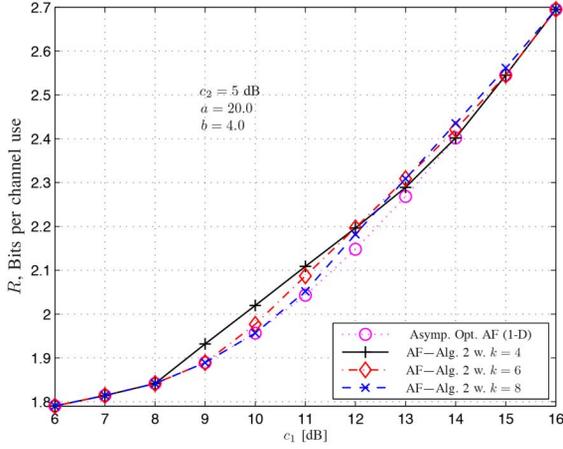


Fig. 6. Comparison between the rate achieved by 1-dimensional (asymptotically optimal AF) and multidimensional (rank-1) AF with block length $k = 4, 6, 8$.

and $k = 4, 6, 8$. The value of c_1 is varied between 6 and 16 dB. From this figure, we note that at intermediate source transmit powers, multidimensional relaying corresponding to $r = \frac{k}{2} - 1$ in Algorithm 2 can yield an appreciable gain over 1-dimensional relaying. For instance, at a data rate of 2 bits per channel use, the gain in this example is in excess of 0.8 dB. In addition, we note that, although $R_{AF}^{(k)}$ is monotonically increasing in k as shown in Remark 6, the multidimensional relaying rate, $R_b^{(k)}$, in (36) and subsequently the lower bound in Proposition 4 do not possess this property. For instance, setting $k = 4$ at $c_1 = 10$ dB yields a rate higher than that yielded by $k = 8$. However, at $c_1 = 14$ dB, setting $k = 8$ yields an appreciable gain over $k = 4$, which, at this source transmit power coincides with the rate of 1-dimensional relaying. \square

VI. CONCLUSION

The maximum data rate that can be achieved by the classic strictly causal full-duplex AF scheme in Gaussian relay channels is cast as the solution of a joint optimization problem of a positive semidefinite input covariance and a strictly lower triangular relay precoder. Although this problem is not convex, when the relay precoder is given, finding the corresponding optimal input covariance constitutes a convex problem for which the KKT conditions are sufficient and necessary for optimality. Using these conditions, closed form expressions for the optimal input covariance were derived, and these expressions were subsequently used to show that restricting the relay precoders to be subdiagonal does not incur loss of optimality. This result enabled us to reduce the number of optimization variables from quadratic to linear in the block length. Furthermore, the subdiagonal structure enabled us to identify a class of relay precoders that are asymptotically optimal at low and high source transmit powers. Precoders in this class transmit every other time slot in what resembles half-duplex operation. For intermediate transmit powers, we developed an efficient algorithm for generating optimal rank-1 precoders, which are shown, via numerical simulations, to achieve rates that are significantly higher than those achieved by asymptotically optimal (1-dimensional) precoders.

APPENDIX A PROOF OF PROPOSITION 1

To study the case of $\Phi \neq 0$ and $\mu_2 > 0$, we invoke (11) to cast (8) in forms that are more convenient to derive expressions for Σ and $\tilde{\Phi}$. In particular, to derive an explicit expression for Σ , we multiply both sides of (8) by $(\mu_1 I_k + \mu_2 D^T D - \Phi)$, and use the fact that $\Sigma \Phi = 0$ to write

$$(\mu_1 I_k + \mu_2 D^T D) \Sigma = I - (\mu_1 I_k + \mu_2 D^T D - \Phi)(H^T H)^{-1}. \quad (38)$$

Left multiplying (38) by Σ , and taking the transpose of both sides of the resulting equation yields

$$\Sigma(\mu_1 I_k + \mu_2 D^T D) \Sigma = \Sigma - (H^T H)^{-1}(\mu_1 I_k + \mu_2 D^T D) \Sigma. \quad (39)$$

For notational convenience, we use the fact that $\mu_1 > 0$ to define $\nu_1 \triangleq 1/\mu_1$ and $\nu_2 \triangleq \mu_2/\mu_1$. Furthermore, we define

$$\tilde{A} \triangleq I_k + \nu_2 D^T D, \quad \tilde{C} = (H^T H)^{-1}, \quad \text{and} \quad \tilde{B} \triangleq \tilde{C} \tilde{A}. \quad (40)$$

Denoting the (full) eigendecomposition of Σ by $U \begin{bmatrix} \Lambda_\Sigma & 0 \\ 0 & 0 \end{bmatrix} U^T$, the equality in (39) can be written as
$$U \begin{bmatrix} \Lambda_\Sigma & 0 \\ 0 & 0 \end{bmatrix} U^T \tilde{A} U \begin{bmatrix} \Lambda_\Sigma & 0 \\ 0 & 0 \end{bmatrix} U^T = \nu_1 U \begin{bmatrix} \Lambda_\Sigma & 0 \\ 0 & 0 \end{bmatrix} U^T - \tilde{B} U \begin{bmatrix} \Lambda_\Sigma & 0 \\ 0 & 0 \end{bmatrix} U^T,$$

where $U = [U_\Sigma \ U_\Phi]$ and $U_\Sigma = U_\Phi^\perp$. Left and right multiplication of both sides of this equation by U^T and U , respectively, and using A to denote $U^T \tilde{A} U$ and B to denote $U^T \tilde{B} U$ yields

$$\begin{bmatrix} \Lambda_\Sigma & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \Lambda_\Sigma & 0 \\ 0 & 0 \end{bmatrix} = \nu_1 \begin{bmatrix} \Lambda_\Sigma & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} \Lambda_\Sigma & 0 \\ 0 & 0 \end{bmatrix}, \quad (41)$$

where $\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ and $\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$ denote partitions of A and B with $A_{11}, B_{11} \in \mathbb{R}^{k_\Sigma \times k_\Sigma}$. Right multiplying (41) by $\begin{bmatrix} \Lambda_\Sigma^{-1} & 0 \\ 0 & 0 \end{bmatrix}$ and simplifying yield

$$\Lambda_\Sigma A_{11} = \nu_1 I - B_{11}, \quad \text{and} \quad B_{21} = 0. \quad (42)$$

We will use (42) to obtain the optimal Σ , but first, let us derive similar equations for the dual matrix, Φ .

Right multiplying (38) by Φ and using (6d) yields

$$\Phi = \frac{1}{\nu_1} \Phi \tilde{B} - \Phi (H^T H)^{-1} \Phi. \quad (43)$$

The full eigendecomposition of Φ can be expressed as $U \begin{bmatrix} 0 & 0 \\ 0 & \Lambda_\Phi \end{bmatrix} U^T$. Using this decomposition and left and right multiplication U^T and U , respectively, yields

$$\begin{bmatrix} 0 & 0 \\ 0 & \Lambda_\Phi \end{bmatrix} = \frac{1}{\nu_1} \begin{bmatrix} 0 & 0 \\ 0 & \Lambda_\Phi \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & \Lambda_\Phi \end{bmatrix} \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & \Lambda_\Phi \end{bmatrix}, \quad (44)$$

where $C = U^T \tilde{C} U = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$. Left multiplication of (44) by $\begin{bmatrix} 0 & 0 \\ 0 & \Lambda_\Phi^{-1} \end{bmatrix}$ and simplifying yields

$$C_{22} \Lambda_\Phi = \frac{1}{\nu_1} B_{22} - I, \quad \text{and} \quad B_{21} = 0. \quad (45)$$

We now consider (42) and (45). First, we note that A_{11} , C_{22} , Λ_Σ , and Λ_Φ are full rank and positive definite. Second, since, for any two square matrices W_1 and W_2 , the eigenvalues of $W_1 W_2$ are the same as the eigenvalues of $W_2 W_1$, it follows that k_Σ and $k_\Phi = k - k_\Sigma$ must be such that the k_Σ eigenvalues of B_{11} are less than ν_1 and the k_Φ eigenvalues of B_{22} are greater than ν_1 .

We now proceed to determine the orthogonal bases of Σ and Φ . From (42) and (45), it is seen that $B_{21} = U_\Phi^T \tilde{B} U_\Phi = 0$. To determine U_Φ and U_Σ , we invoke the Schur decomposition of an arbitrary (asymmetric) matrix W .

Lemma 4 (Schur decomposition [20]): Let $W \in \mathbb{R}^{n \times n}$ be an asymmetric matrix with real eigenvalues. For each such matrix there exists an orthonormal matrix Ψ such that the matrix $\Psi W \Psi^T$ is upper triangular. The diagonal entries of $\Psi W \Psi^T$ are the eigenvalues of W and their order can be arbitrarily chosen depending on Ψ . \square

We now apply Lemma 4 and the discussion preceding it to the matrix \tilde{B} . Let \tilde{U} be the Schur basis of \tilde{B} that arranges its eigenvalues in a nondecreasing order, i.e., $\tilde{U}^T \tilde{B} \tilde{U}$ is upper triangular with nondecreasing diagonal entries. From (42) and (45), we have that $B_{21} = 0$. However, these equations do not require B_{11} and B_{22} to be upper triangular. Hence, it can be seen that to satisfy these equations, we can choose $U = [U_\Sigma \quad U_\Phi]$ to be in the form

$$U = \tilde{U} \begin{bmatrix} Q_1 & 0 \\ 0 & Q_2 \end{bmatrix} = [\tilde{U}_1 \quad \tilde{U}_2] \begin{bmatrix} Q_1 & 0 \\ 0 & Q_2 \end{bmatrix}, \quad (46)$$

where $Q_1 \in \mathbb{R}^{k_\Sigma \times k_\Sigma}$ and $Q_2 \in \mathbb{R}^{k_\Phi \times k_\Phi}$ are orthonormal and $\tilde{U}_1 \in \mathbb{R}^{k \times k_\Sigma}$ and $\tilde{U}_2 \in \mathbb{R}^{k \times k_\Phi}$ are complementary partitions of \tilde{U} . Using (46) in the first equality in (45) yields

$$Q_2^T \tilde{U}_2^T \tilde{C} \tilde{U}_2 Q_2 \Lambda_\Phi = \frac{1}{\nu_1} Q_2^T \tilde{U}_2^T \tilde{B} \tilde{U}_2 Q_2 - I. \quad (47)$$

Left and right multiplication of (47) by Q_2 and Q_2^T , respectively, yields

$$Q_2 \Lambda_\Phi Q_2^T = \frac{1}{\nu_1} (\tilde{U}_2^T \tilde{C} \tilde{U}_2)^{-1} (\tilde{U}_2^T \tilde{B} \tilde{U}_2) - (\tilde{U}_2^T \tilde{C} \tilde{U}_2)^{-1}. \quad (48)$$

Using (48), the matrix Φ can be written as

$$\Phi = \tilde{U}_2 \left(\frac{1}{\nu_1} (\tilde{U}_2^T \tilde{C} \tilde{U}_2)^{-1} (\tilde{U}_2^T \tilde{B} \tilde{U}_2) - (\tilde{U}_2^T \tilde{C} \tilde{U}_2)^{-1} \right) \tilde{U}_2^T. \quad (49)$$

Since the matrix Φ is symmetric, the matrix $(\tilde{U}_2^T \tilde{C} \tilde{U}_2)^{-1} (\tilde{U}_2^T \tilde{B} \tilde{U}_2)$ must be symmetric, which is what we show next using the following lemma.

Lemma 5: Let $W_1, W_2 \in \mathbb{R}^{n \times n}$ be two nonsingular symmetric matrices. Let $\Psi = [\Psi_1 \quad \Psi_2]$ be a Schur orthonormal

basis of the matrix $W_1 W_2$, where Ψ_1 and Ψ_2 are arbitrary complementary partitions of this basis. Then,

$$(\Psi_1^T W_1 W_2 \Psi_1) (\Psi_1^T W_2 \Psi_1)^{-1} = (\Psi_1^T W_1^{-1} \Psi_1)^{-1}, \quad \text{and} \\ (\Psi_2^T W_1 \Psi_2)^{-1} (\Psi_2^T W_1 W_2 \Psi_2) = (\Psi_2^T W_2^{-1} \Psi_2)^{-1}.$$

Proof: See Appendix B. \blacksquare

Letting $W_1 = \tilde{C}$ and $W_2 = \tilde{A}$ implies that $W_1 W_2 = \tilde{B}$, which using Lemma 5 implies that Φ in (49) is symmetric.

Substituting for Φ from (49) in (8) yields Σ . Alternatively, Σ can be obtained using (42). In particular, expressing A_{11} as $Q_1^T \tilde{U}_1^T \tilde{A} \tilde{U}_1 Q_1$ and left and right multiplying both sides of (42) by Q_1 and Q_1^T , respectively, yields

$$Q_1 \Lambda_\Sigma Q_1^T \tilde{U}_1^T \tilde{A} \tilde{U}_1 = \nu_1 I - \tilde{U}_1^T \tilde{B} \tilde{U}_1. \quad (50)$$

Right multiplying (50) by $\tilde{U}_1^T \tilde{A} \tilde{U}_1^{-1}$ and using Lemma 5 yields

$$Q_1 \Lambda_\Sigma Q_1^T = \nu_1 (\tilde{U}_1^T \tilde{A} \tilde{U}_1)^{-1} - (\tilde{U}_1^T \tilde{C}^{-1} \tilde{U}_1)^{-1},$$

which, after substituting for \tilde{A} and \tilde{B} , yields (14).

To find the optimal ν_2 , we note that since $\text{Tr}(\Sigma) = c_1$, we have

$$\nu_1 = \frac{c_1 + \text{Tr} \left((\tilde{U}_1^T H^T H \tilde{U}_1)^{-1} \right)}{\text{Tr} \left((I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)^{-1} \right)}. \quad (51)$$

Similarly, using the constraint $\text{Tr}(D^T D \Sigma) = c'_2$, we have

$$\nu_1 \text{Tr} \left(\tilde{U}_1^T D^T D \tilde{U}_1 (I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)^{-1} \right) \\ - \text{Tr} \left(\tilde{U}_1^T D^T D \tilde{U}_1 (\tilde{U}_1^T H^T H \tilde{U}_1)^{-1} \right) = c'_2. \quad (52)$$

Substituting for ν_1 from (51) in (52) and simplifying, yields

$$\frac{\text{Tr} \left(\tilde{U}_1^T D^T D \tilde{U}_1 (I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)^{-1} \right)}{\text{Tr} \left((I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)^{-1} \right)} \\ = \frac{c'_2 + \text{Tr} \left(\tilde{U}_1^T D^T D \tilde{U}_1 (\tilde{U}_1^T H^T H \tilde{U}_1)^{-1} \right)}{c_1 + \text{Tr} \left((\tilde{U}_1^T H^T H \tilde{U}_1)^{-1} \right)}. \quad (53)$$

The implementation of Algorithm 1 requires an efficient means for finding the optimal ν_2 for a given \tilde{U}_1 . Such a means can be obtained by using the following lemma.

Lemma 6: For any matrix $A = A^T \succeq 0$, the function $f(x) = \frac{\text{Tr}(A(I+xA)^{-1})}{\text{Tr}((I+xA)^{-1})}$ is monotonically decreasing in x . Furthermore, if the number of distinct eigenvalues of A is greater than one, then $f(x)$ is strongly monotonically decreasing.

Proof: See Appendix C. \blacksquare

Since $\tilde{U}_1^T D^T D \tilde{U}_1$ is positive semidefinite, Lemma 6 asserts that the left-hand side of (53) is monotonically decreasing in ν_2 . This observation ensures that, for any given \tilde{U}_1 , a standard bisection search will converge exponentially to the corresponding optimal ν_2 . However, we note that, in general, when $\Phi \neq 0$ the

Schur basis \tilde{U}_1 , depends on ν_2 which renders the search more difficult than the cases of $\Phi = 0$ and $\mu_2 = 0$ considered in Remark 1.

APPENDIX B
PROOF OF LEMMA 5

We will prove the first statement of the lemma. The proof of the second statement is similar, and is omitted for brevity.

Let $T \in \mathbb{R}^{k \times k}$ be any nonsingular matrix that is block partitioned as $T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}$, and let $S = T^{-1}$ be block partitioned in the same way, where $T_{11}, S_{11} \in \mathbb{R}^{k_{\Sigma} \times k_{\Sigma}}$, $T_{22}, S_{22} \in \mathbb{R}^{k_{\Phi} \times k_{\Phi}}$ and $k_{\Phi} = k - k_{\Sigma}$. Then, the north west block of S , i.e., S_{11} , is given by [20]

$$S_{11} = (T_{11} - T_{12}T_{22}^{-1}T_{21})^{-1}. \quad (54)$$

Let $\Psi = [\Psi_1 \ \Psi_2]$, where $\Psi_1 \in \mathbb{R}^{k \times k_{\Sigma}}$ and $\Psi_2 \in \mathbb{R}^{k \times k_{\Phi}}$. Let

$$T = (\Psi^T W_1 W_2 \Psi)^{-1} = \Psi^T (W_1 W_2)^{-1} \Psi,$$

from which we have

$$S_{11} = \Psi_1^T W_1 W_2 \Psi_1. \quad (55)$$

Now, from (54) we have

$$S_{11}^{-1} = \Psi_1^T (W_1 W_2)^{-1} \Psi_1 - \Psi_1^T (W_1 W_2)^{-1} \Psi_2 (\Psi_2^T (W_1 W_2)^{-1} \Psi_2)^{-1} \Psi_2^T (W_1 W_2)^{-1} \Psi_1.$$

Since Ψ is a Schur basis of $W_1 W_2$, we have $\Psi_2^T (W_1 W_2)^{-1} \Psi_1 = 0$, whence the second term in this equation vanishes yielding

$$S_{11}^{-1} = \Psi_1^T (W_1 W_2)^{-1} \Psi_1. \quad (56)$$

From (55) and (56), we have

$$(\Psi_1^T W_1 W_2 \Psi_1)^{-1} = \Psi_1^T (W_1 W_2)^{-1} \Psi_1.$$

Using this result, we have

$$(\Psi_1^T W_2 \Psi_1)(\Psi_1^T W_1 W_2 \Psi_1)^{-1} = \Psi_1^T W_2 \Psi_1 \Psi_1^T W_2^{-1} W_1^{-1} \Psi_1. \quad (57)$$

Now, the projector matrix $\Psi_1 \Psi_1^T$ can be expressed as $I - \Psi_2 \Psi_2^T$. Substituting this identity into (57) yields

$$\begin{aligned} & (\Psi_1^T W_2 \Psi_1)(\Psi_1^T W_1 W_2 \Psi_1)^{-1} \\ &= \Psi_1^T W_1^{-1} \Psi_1 - \Psi_1^T W_2 \Psi_2 \Psi_2^T W_2^{-1} W_1^{-1} \Psi_1. \end{aligned}$$

The second term of this equation vanishes because $\Psi_2^T (W_1 W_2)^{-1} \Psi_1 = 0$. Hence, we have

$$(\Psi_1^T W_1 \Psi_1)(\Psi_1^T W_1 W_2 \Psi_1)^{-1} = \Psi_1^T W_1^{-1} \Psi_1, \quad (58)$$

from which the first statement of the lemma directly follows.

APPENDIX C
PROOF OF LEMMA 6

To prove Lemma 6, we use $\{\lambda_{A_i}\}$ to denote the eigenvalues of A and write

$$\frac{\text{Tr}(A(I + xA)^{-1})}{\text{Tr}((I + xA)^{-1})} = \frac{\sum_{i=1}^k \frac{\lambda_{A_i}}{1+x\lambda_{A_i}}}{\sum_{i=1}^k \frac{1}{1+x\lambda_{A_i}}}.$$

Now,

$$\begin{aligned} \frac{d}{dx} \left(\frac{\sum_{i=1}^k \frac{\lambda_{A_i}}{1+x\lambda_{A_i}}}{\sum_{i=1}^k \frac{1}{1+x\lambda_{A_i}}} \right) &= \frac{\left(\sum_{i=1}^k \frac{1}{1+x\lambda_{A_i}} \right) \left(\sum_{j=1}^k \frac{-\lambda_{A_j}^2}{(1+x\lambda_{A_j})^2} \right)}{\left(\sum_{i=1}^k \frac{1}{1+x\lambda_{A_i}} \right)^2} \\ &+ \frac{\left(\sum_{i=1}^k \frac{\lambda_{A_i}}{1+x\lambda_{A_i}} \right) \left(\sum_{j=1}^k \frac{\lambda_{A_j}}{(1+x\lambda_{A_j})^2} \right)}{\left(\sum_{i=1}^k \frac{1}{1+x\lambda_{A_i}} \right)^2}. \end{aligned} \quad (59)$$

Let \mathcal{N} denote the numerator of (59). The first summation in \mathcal{N} can be expressed as

$$\sum_{i=1}^k \frac{1}{1+x\lambda_{A_i}} = \sum_{i=1}^k \left(1 - \frac{x\lambda_{A_i}}{1+x\lambda_{A_i}} \right) = k - x \sum_{i=1}^k \frac{\lambda_{A_i}}{1+x\lambda_{A_i}}. \quad (60)$$

Substituting from (60) into \mathcal{N} and simplifying, we have

$$\begin{aligned} \mathcal{N} &= \left(\sum_{i=1}^k \frac{\lambda_{A_i}}{1+x\lambda_{A_i}} \right) \left(\sum_{j=1}^k \frac{\lambda_{A_j} + x\lambda_{A_j}^2}{(1+x\lambda_{A_j})^2} \right) \\ &\quad - k \sum_{j=1}^k \frac{\lambda_{A_j}^2}{(1+x\lambda_{A_j})^2} \\ &= \sum_{i=1}^k \sum_{j=1}^k \left(\frac{\lambda_{A_i}}{1+x\lambda_{A_i}} \right) \left(\frac{\lambda_{A_j}}{1+x\lambda_{A_j}} \right) \\ &\quad - k \sum_{j=1}^k \frac{\lambda_{A_j}^2}{(1+x\lambda_{A_j})^2}. \end{aligned} \quad (61)$$

Let

$$y_i = \frac{\lambda_{A_i}}{1+x\lambda_{A_i}}, \quad i = 1, \dots, k. \quad (62)$$

Using this substitution in the right-hand side of (61), we have

$$\mathcal{N} = -k \sum_{j=1}^k y_j^2 + \sum_{i=1}^k \sum_{j=1}^k y_i y_j = -\frac{1}{2} \sum_{i=1}^k \sum_{j=1}^k (y_i - y_j)^2 \leq 0. \quad (63)$$

The last inequality holds strictly if $y_i \neq y_j$ for some $i, j \in \{1, \dots, k\}$. Hence, the derivative in (59) is monotonically decreasing as claimed in the lemma.

APPENDIX D
PROOF OF THEOREM 1

To prove this theorem, let the singular value decomposition of D be denoted by $D = U_D \Sigma_D V_D^T$. We will first show that the

optimization problem in (16) depends only on Σ_D and $Q_D = U_D^T V_D$, but not on U_D and V_D individually. Invoking the singular value decomposition of D in (17) yields

$$(I_k + \nu_2 D^T D)^{-1} H^T H = \frac{1}{N} V_D (I_k + \nu_2 \Sigma_D^2)^{-1} \\ \times (Q_D^T + ab \Sigma_D) (I_k + b^2 \Sigma_D^2)^{-1} (Q_D + ab \Sigma_D) V_D^T. \quad (64)$$

From this equation, it can be seen that, if $[\tilde{U}_1 \ \tilde{U}_2]$ denotes the Schur basis of the matrix $(H^T H)^{-1} (I_k + \nu_2 D^T D)$, then $[\tilde{U}_1 \ \tilde{U}_2]$ can be expressed as $V_D [\hat{U}_1 \ \hat{U}_2]$, where \hat{U}_1 and \hat{U}_2 depend only on Σ_D and $U_D^T V_D$, but not on U_D and V_D . We will now examine the objective and the constraints of (16).

For the objective, we note that, because \tilde{U}_1 is the Schur basis of the matrix $\tilde{B} = (H^T H)^{-1} (I_k + \nu_2 D^T D)$, the objective in (16a) is the product of the smallest k_Σ eigenvalues of \tilde{B} . Now, from (64) it can be seen that the eigenvalues of \tilde{B} depend only on Σ_D and $U_D^T V_D$, but not on U_D and V_D .

Using the singular value decomposition of D and the fact that $\tilde{U}_1 = V_D \hat{U}_1$, the constraint in (16b), can be expressed as

$$(\hat{U}_1^T (I_k + \nu_2 \Sigma_D^2) \hat{U}_1)^{-1} \succeq \frac{N}{\nu_1} (\hat{U}_1^T (Q_D^T + ab \Sigma_D) \\ \times (I_k + b^2 \Sigma_D^2)^{-1} (Q_D + ab \Sigma_D) \hat{U}_1)^{-1}.$$

Based on Remark 2, this constraint can be eliminated if (16) is feasible for the given ν_1 and ν_2 .

The constraint in (16c) can be expressed as

$$\nu_1 \text{Tr} \left((\hat{U}_1^T (I_k + \nu_2 \Sigma_D^2) \hat{U}_1)^{-1} \right) - c_1 \\ N \text{Tr} \left((\hat{U}_1^T (Q_D^T + ab \Sigma_D)^{-1} (I_k + b^2 \Sigma_D^2) (Q_D + ab \Sigma_D) \hat{U}_1)^{-1} \right),$$

and the constraint in (16d) can be expressed as

$$\text{Tr} \left(\hat{U}_1^T \Sigma_D^2 \hat{U}_1 \left(a^2 \nu_1 (I_k + \nu_2 \hat{U}_1^T \Sigma_D^2 \hat{U}_1)^{-1} - \right. \right. \\ \left. \left. a^2 N (\hat{U}_1^T (Q_D^T + ab \Sigma_D) (I_k + b^2 \Sigma_D^2)^{-1} (Q_D + ab \Sigma_D) \hat{U}_1)^{-1} \right. \right. \\ \left. \left. + NI \right) \right) = c_2.$$

Finally, we consider the constraint on D in (16e). We will show that this constraint is equivalent to the constraint that $\lambda_i(D) = 0$, $i = 1, \dots, k$. First, note that the constraint on D in (16e) implies that all the eigenvalues of D are zero; i.e., $\lambda_i(D) = 0$, $i = 1, \dots, k$. Since $\lambda_i(D) = \lambda_i(U_D \Sigma_D V_D^T) = \lambda_i(Q_D^T \Sigma_D) = 0$, $i = 1, \dots, k$, it can be seen that the optimization problem corresponding to (16) but with the constraint on D in (16e) replaced with the constraint that $\lambda_i(D) = 0$ is a relaxation of (16) and this relaxation depends only on Σ_D and Q_D but not on U_D and V_D . Suppose that solving this relaxation yields $\tilde{D}^* = \tilde{U}_D^* \tilde{\Sigma}_D^* \tilde{V}_D^{*T}$. If \tilde{D}^* is lower triangular, then \tilde{D}^* satisfies the constraints of (16). If \tilde{D}^* is not lower triangular, then there exists a Schur orthonormal matrix X such that $X \tilde{D}^* X^T = X \tilde{U}_D^* \tilde{\Sigma}_D^* \tilde{V}_D^{*T} X^T$ is lower triangular; cf., Lemma 4. However, since all the other constraints of (16) and the corresponding relaxation depend only on Σ_D and $Q_D = U_D^T V_D = (X U_D)^T (X V_D)$ but not on U_D and V_D , it follows that replacing \tilde{U}_D^* with $X \tilde{U}_D^*$ and \tilde{V}_D^* with

$X \tilde{V}_D^{*T}$ yields a lower triangular \tilde{D}^* that satisfies the constraints of (16). Since the optimization problem corresponding to (16) with the constraint on D in (16e) replaced with the constraint that $\lambda_i(D) = 0$ is a relaxation of (16), it follows that the value of the objective corresponding to \tilde{D}^* is a lower bound on the objective of the original optimization problem in (16). However, since \tilde{D}^* satisfies the constraints of (16), it can be seen that \tilde{D}^* solves (16). Hence, we have shown that the constraint on D in (16e) is equivalent to the constraint that $\lambda_i(D) = 0$, $i = 1, \dots, k$, and thus this constraint depends only on Σ_D and Q_D but not on U_D and V_D . Hence, we conclude that the optimization problem in (16) depends only on Σ_D and Q_D , but not on U_D and V_D .

To prove the statement of the theorem, it remains to determine $Q_D = U_D^T V_D$. Since the optimization problem in (16) depends only on Σ_D and Q_D , the matrix V_D (or U_D) can be arbitrarily chosen. In particular, suppose that the optimal matrix Q_D^* is known, then the optimization in (16) yields the same value for arbitrary V_D and $U_D = V_D Q_D^{*T}$. Using this observation and setting $V_D = I$, it can be seen that for $U_D \Sigma_D V_D^T$ to be lower triangular, U_D must be of the form

$$U_D = \begin{bmatrix} 0 & 0 & \cdots & 0 & \pm 1 \\ \pm 1 & 0 & \cdots & 0 & 0 \\ 0 & \pm 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \pm 1 & 0 \end{bmatrix}. \quad (65)$$

Using this U_D , $V_D = I$, and $\Sigma_D = \text{diag}(\sigma_1, \dots, \sigma_{k-1}, 0)$ yields D in (18).

APPENDIX E PROOF OF REMARK 2

The constraint in (16b) can be expressed as $\nu_1 (H^T H)^{1/2} (I_k + \nu_2 D^T D)^{-1} (H^T H)^{1/2} \succeq I$ and the objective can be expressed as $\log \det(\nu_1 (H^T H)^{1/2} (I_k + \nu_2 D^T D)^{-1} (H^T H)^{1/2})$. Now, for any two matrices W_1 and W_2 satisfying $W_1 \succeq W_2$, we have $\det(W_1) \geq \det(W_2)$ [20]. Hence, if (16) is feasible, removing the constraint in (16b) does not affect the optimal solution of (16). This is because, if (16) is feasible and, for a given D satisfying (16c) and (16d), the constraint in (16b) is violated, a greater value of the objective can be attained by selecting another D that satisfies (16b)–(16d). Conversely, if (16b) is removed and, at the optimal solution, (16b) is violated, then (16) must be infeasible.

APPENDIX F PROOF OF LEMMA 3

For notational convenience, let $\{z_i\}_{i=1}^k$ denote the eigenvalues of \tilde{B} , where $z_k \geq \dots \geq z_1 \geq \gamma$. The Lagrangian corresponding to (29) is given by

$$L(\{z_i\}_{i=1}^k, \mu, \{\alpha_i\}_{i=1}^k) = \prod_{i=1}^{k_\Sigma} z_i + \mu \left(\sum_{i=1}^{k_\Sigma} z_i - c \right) + \sum_{i=1}^{k_\Sigma} \alpha_i (\gamma - z_i).$$

Since the problem in (29) is not convex, the KKT conditions are only necessary for optimality. Using the Lagrangian expression, these conditions can be written as

$$\prod_{i \neq j} z_i = \alpha_j - \mu, \quad z_i \geq \gamma, \quad \sum_{i=1}^k z_i = c, \quad (66)$$

$$\mu \text{ free}, \quad \alpha_i \geq 0, \quad \alpha_i(z_i - \gamma) = 0, \quad i = 1, \dots, k. \quad (67)$$

Since $z_k \geq \dots \geq z_1 \geq \gamma$, we can assume that $z_k \geq \dots \geq z_{s+1} > \gamma$ and $z_s = \dots = z_1 = \gamma$, where $s \in \{0, \dots, k\}$. Now, the last condition in (67) yields $\alpha_{s+1} = \dots = \alpha_k = 0$. Using this in (66) yields

$$\mu = - \prod_{i=1, i \neq s+1}^k z_i = - \prod_{i=1, i \neq s+2}^k z_i = \dots = - \prod_{i=1, i \neq k}^k z_i. \quad (68)$$

From (68), we have $z_{s+1} = \dots = z_k = \beta$ and $z_1 = \dots = z_s = \gamma$, where β is strictly greater than γ .

APPENDIX G PROOF OF PROPOSITION 3

The goal of analysis in this appendix is two fold: one, to prove Proposition 3, and two, to facilitate the design of an efficient algorithm for finding optimal relay precoders with interlacing subdiagonal structures. To achieve this goal with minimal redundancy, we will keep the analysis general for any $\nu_2 \geq 0$, even though we have shown in Section IV-C that $\nu_2 \rightarrow 0$ as $c_1 \rightarrow 0$. This generalization will enable us to explore the use of subdiagonal precoders with interlacing structure in cases other than the one of $c_1 \rightarrow 0$. Such precoders are shown to provide close-to-optimal performance in Section V.

From Proposition 2 and Lemma 2 we know that $\lceil \frac{k_\Sigma}{2} \rceil - 1$ of the subdiagonal entries of D must be zero in order for the eigenvalues of the matrix \tilde{B} to have the required multiplicities. The objective and constraints in (29) are functions only of the minimum k_Σ eigenvalues of \tilde{B} , which must have the multiplicities in Lemma 3. Hence, only the part of D that yields \tilde{B} with eigenvalues with the desired multiplicities contribute directly to the objective in (29a). Now, from Lemma 2 and the structure of \tilde{B} in (27), it can be seen that if some consecutive subdiagonal entries of D are nonzero, the resulting \tilde{B} will have a corresponding diagonal block with distinct eigenvalues. Hence, for D to yield \tilde{B} with eigenvalues with the required multiplicities, some of its nonzero subdiagonal entries must be identical and each must be followed by a zero. We will make these observations precise in the following analysis.

Suppose that D has the subdiagonal structure in (18) and let us assume that the k_Σ eigenvalues of \tilde{B} in (29a) correspond to the first $k' - 1$ subdiagonal entries of D , where $k' \leq k$, and for ease of exposition, k' is assumed even. We will later show that the optimal $k' = k$.

To yield \tilde{B} with the desired eigenvalue multiplicities in Lemma 3, the considered $k' - 1$ entries must be interlacing.

In particular, suppose that $\sigma_{2i} = 0$ for $i = 1, \dots, \frac{k'}{2} - 1$, and let $\sigma_{2i-1} = 0$ for $i = 1, \dots, r$, and $\sigma_{2i-1} = \sigma_D > 0$ for $i = r + 1, \dots, \frac{k'}{2}$, $r = 0, \dots, \frac{k'}{2} - 1$.

With this structure, the matrix $D^T D$ is diagonal with the first $2r + 1$ diagonal entries equal to zero, the following $k' - 2r - 1$ entries interlacing with values σ_D^2 and 0, respectively, and the remaining $k - k'$ entries are nonzero and do not contribute to the objective in (29a); cf., (27). We will study the optimization problem in (26) when D is restricted to have this structure.

The structure of D yields a matrix \tilde{B}^{-1} of the form

$$\tilde{B}^{-1} = \frac{1}{N} I_{2r} \oplus \left(\oplus_{i=r+1}^{\frac{k'}{2}} G \right) \oplus \hat{G}, \quad (69)$$

where

$$G = \frac{1}{N(1 + b^2 \sigma_D^2)} \begin{bmatrix} \frac{1 + (1 + a^2)b^2 \sigma_D^2}{1 + \nu_2 \sigma_D^2} & \frac{ab \sigma_D}{1 + \nu_2 \sigma_D^2} \\ ab \sigma_D & 1 \end{bmatrix}, \quad (70)$$

and \hat{G} is the diagonal block of \tilde{B}^{-1} corresponding to the part of D with consecutive nonzero entries. The eigenvalues of \hat{G} are distinct and therefore do not satisfy the conditions of Lemma 3 and do not contribute to the objective in (29a).

The construction in (69) yields a \tilde{B}^{-1} with an upper left $k' \times k'$ corner with the following distinct three eigenvalues: $\lambda_{\tilde{B}_0^{-1}} = \frac{1}{N}$ with multiplicity $2r$, and

$$\lambda_{\tilde{B}_1^{-1}} = \frac{1}{2N(1 + b^2 \sigma_D^2)} \left(1 + \frac{1 + (1 + a^2)b^2 \sigma_D^2}{1 + \nu_2 \sigma_D^2} + \sqrt{\left(1 - \frac{1 + (1 + a^2)b^2 \sigma_D^2}{1 + \nu_2 \sigma_D^2} \right)^2 + 4 \frac{a^2 b^2 \sigma_D^2}{1 + \nu_2 \sigma_D^2}} \right), \quad (71a)$$

$$\lambda_{\tilde{B}_2^{-1}} = \frac{1}{2N(1 + b^2 \sigma_D^2)} \left(1 + \frac{1 + (1 + a^2)b^2 \sigma_D^2}{1 + \nu_2 \sigma_D^2} - \sqrt{\left(1 - \frac{1 + (1 + a^2)b^2 \sigma_D^2}{1 + \nu_2 \sigma_D^2} \right)^2 + 4 \frac{a^2 b^2 \sigma_D^2}{1 + \nu_2 \sigma_D^2}} \right), \quad (71b)$$

each with multiplicity $\frac{k'}{2} - r$.

The matrix \tilde{U}_1 in (14) is given by the Schur vectors that correspond to the smallest k_Σ eigenvalues of \tilde{B} . These are the same as the Schur vectors that correspond to the largest k_Σ eigenvalues of \tilde{B}^{-1} . These eigenvalues are given by $\lambda_{\tilde{B}_1^{-1}}$, and possibly $\lambda_{\tilde{B}_0^{-1}}$, depending on the optimal values of k_Σ and r .

The Schur vectors corresponding to $\lambda_{\tilde{B}_0^{-1}}$ correspond to the zero diagonal entries of $D^T D$, and without loss of generality can be written as $\tilde{U}_{\tilde{B}_0} = [I_{2r} \ 0_{2r \times (k-2r)}]^T$.

The Schur vectors corresponding to these eigenvalues and can be expressed as $\tilde{U}_{\tilde{B}_1} = [0_{\frac{k'}{2}-r \times 2r} \ I_{\frac{k'}{2}-r} \otimes [u_{\tilde{B}_1} \ u_{\tilde{B}_2}]]^T$, where $u_{\tilde{B}_i}$, $i = 1, 2$ are defined below. See (72) and (73), shown at the bottom of the next page.

To study the optimization problem in (26) when D is subdiagonal and its first $k' - 1$ subdiagonal entries are restricted to have the interlacing structure, we consider matrices of the form $\tilde{U}_{\tilde{B}_i}^T D^T D \tilde{U}_{\tilde{B}_i}$ for $i = 0, 1$. First, we note that, because

$\tilde{U}_{\tilde{B}_0}$ corresponds to the zero diagonal entries of $D^T D$, we have $\tilde{U}_{\tilde{B}_0}^T D = 0_{2r \times k}$. Second, we have

$$\tilde{U}_{\tilde{B}_1}^T D^T D \tilde{U}_{\tilde{B}_1} = u_{\tilde{B}_1}^2 \sigma_D^2 I_{\frac{k'}{2}-r}. \quad (74)$$

Before examining the details of the optimization problem in (26), we note that the constraint in (26e) is satisfied by the subdiagonal construction of D and that there are two possibilities for the matrix Σ to satisfy the conditions of Proposition 2 for any given D with the interlacing diagonal entries: either $k_\Sigma = \frac{k'}{2} - r$ or $k_\Sigma = \frac{k'}{2} + r$, $r = 0, \dots, \frac{k'}{2} - 1$. We will consider these cases separately. For both cases

$$\text{rank}(D) = k_D = \frac{k'}{2} - r + k - k' = k - \frac{k'}{2} - r. \quad (75)$$

1) *Case of $k_\Sigma = \frac{k'}{2} - r$:* We will show that in this case the optimal $k' = k$ and the optimal $r = 0$. To do that, we begin by noting that for any given $r \in \{0, \dots, \frac{k'}{2} - 1\}$, the eigenvectors of Σ will be the Schur vectors corresponding to the smallest eigenvalue of \tilde{B} , which is given by $\lambda_{\tilde{B}_1}^{-1}$ and has multiplicity $\frac{k'}{2} - r$; that is, $\tilde{U}_1 = \tilde{U}_{\tilde{B}_1}$. We will now examine the objective and the constraints of (26).

a) *Objective in (26c):* The objective is given by

$$\nu_1^{-k_\Sigma} \det(\tilde{U}_1^T \tilde{B} \tilde{U}_1) = (\lambda_{\tilde{B}_1}^{-1} \nu_1)^{-k_\Sigma}. \quad (76)$$

b) *Constraint in (26c):* The left-hand side of this constraint can be expressed as

$$\begin{aligned} & \text{Tr}(\nu_1(I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)^{-1}) - (\tilde{U}_1^T H^T H \tilde{U}_1)^{-1} \\ &= \text{Tr}((\nu_1 I_{k_\Sigma} - (\tilde{U}_1^T H^T H \tilde{U}_1)^{-1} (I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)) \\ & \quad \times (I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)^{-1}) \\ &= \text{Tr}((\nu_1 I_{k_\Sigma} - \tilde{U}_1^T (H^T H)^{-1} (I_k + \nu_2 D^T D) \tilde{U}_1) \\ & \quad \times (I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)^{-1}) \end{aligned} \quad (77)$$

where in (77), we have used Lemma 5 and the fact that \tilde{U}_1 contains k_Σ vectors of the Schur basis of \tilde{B} to replace $(\tilde{U}_1^T H^T H \tilde{U}_1)^{-1} (I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)$ with $\tilde{U}_1^T (H^T H)^{-1} (I_k + \nu_2 D^T D) \tilde{U}_1$.

From (74), we have

$$(I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)^{-1} = \frac{1}{1 + \nu_2 u_{\tilde{B}_1}^2 \sigma_D^2} I_{\frac{k'}{2}-r}.$$

Substituting in (77) yields that the constraint in (26c) is equivalent to

$$\frac{(\nu_1 - \lambda_{\tilde{B}_1}^{-1}) k_\Sigma}{1 + \nu_2 u_{\tilde{B}_1}^2 \sigma_D^2} = c_1. \quad (78)$$

Note that here we used the fact that, because ν_1 is greater than zero, the constraint in (77) is satisfied with equality at the optimal solution.

c) *Constraint in (26d):* Noting that the diagonal entries of $\tilde{U}_1^T (H^T H)^{-1} (I_k + \nu_2 D^T D) \tilde{U}_1$ are the smallest k_Σ eigenvalues of \tilde{B} , this constraint can be expressed as

$$k_\Sigma \lambda_{\tilde{B}_1}^{-1} = \nu_1 k_\Sigma - c_1 - \frac{\nu_2}{a^2} \left(c_2 - N \left(\frac{k'}{2} - r \right) \sigma_D^2 - N \sum_{i=k'}^{k-1} \sigma_i^2 \right), \quad (79)$$

where $\{\sigma_i\}_{i=k'}^{k-1}$ are the consecutive nonzero $k - k'$ subdiagonal entries of D . Note that, from (79), it can be readily seen that the optimal k' must be equal to k . This is because the $k - k'$ entries $\{\sigma_i\}_{i=k'}^{k-1}$ do not appear except in this constraint and their effect is to reduce the fraction of c_2 available for the relay to transmit its signal. We will later make this observation more formal. In writing (79), we assumed that ν_2 is greater than zero and hence, the constraint in (26d) is satisfied with equality. However, when $\nu_2 = 0$, this constraint can be satisfied with strict inequality.

d) *Constraint in (26b):* This constraint is equivalent to the constraint that

$$\begin{aligned} \nu_1 I_{k_\Sigma} & \succeq (I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)^{1/2} (\tilde{U}_1^T H^T H \tilde{U}_1)^{-1} \\ & \quad \times (I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)^{1/2}, \end{aligned}$$

which is equivalent to the constraint that the maximum eigenvalue of the matrix $(\tilde{U}_1^T H^T H \tilde{U}_1)^{-1} (I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)$ is less than ν_1 . Using Lemma 5, and the fact that the eigenvalues of this matrix are equal and are given by $\lambda_{\tilde{B}_1}^{-1}$, the constraint in (26b) can be cast as

$$\nu_1 \geq \lambda_{\tilde{B}_1}^{-1}. \quad (80)$$

e) *Deriving an auxiliary constraint from (78) and (79):* From (78) and (79), it can be deduced that

$$u_{\tilde{B}_1}^2 = \frac{\left(c_2 - N \left(\frac{k'}{2} - r \right) \sigma_D^2 - N \sum_{i=k'}^{k-1} \sigma_i^2 \right)}{a^2 c_1 \sigma_D^2}. \quad (81)$$

$$u_{\tilde{B}_1} = \frac{ab\sigma_D}{\sqrt{a^2 b^2 \sigma_D^2 + (N \lambda_{\tilde{B}_1}^{-1} (1 + b^2 \sigma_D^2) (1 + \nu_2 \sigma_D^2) - 1 - (1 + a^2) b^2 \sigma_D^2)^2}}, \quad (72)$$

$$u_{\tilde{B}_2} = \frac{2N \lambda_{\tilde{B}_1}^{-1} (1 + b^2 \sigma_D^2) (1 + \nu_2 \sigma_D^2) - 1 - (1 + a^2) b^2 \sigma_D^2}{\sqrt{a^2 b^2 \sigma_D^2 + (N \lambda_{\tilde{B}_1}^{-1} (1 + b^2 \sigma_D^2) (1 + \nu_2 \sigma_D^2) - 1 - (1 + a^2) b^2 \sigma_D^2)^2}}. \quad (73)$$

Let $c_x(\sigma_D) \triangleq \frac{2a^4 b^2 \sigma_D^2 c_1}{\left(c_2 - N\left(\frac{k'}{2} - r\right)\sigma_D^2 - N\sum_{i=k'}^{k-1}\sigma_i^2\right)}$. Substituting for $u_{\tilde{B}_1}^2$ from (72), the constraint in(81) can be cast as

$$\frac{c_x(\sigma_D) - 2a^2 b^2 (2 + \nu_2 \sigma_D^2) - (\nu_2 - (1 + a^2)b^2)^2 \sigma_D^2}{\nu_2 - (1 + a^2)b^2} = \sqrt{(\nu_2 - (1 + a^2)b^2)^2 \sigma_D^4 + 4a^2 b^2 \sigma_D^2 (1 + \nu_2 \sigma_D^2)}.$$

Manipulating this constraint, and using the notation

$$c_y(\sigma_D) \triangleq \sigma_D^2 (1 + a^2)^2 b^4 + 4a^2 b^2 - c_x(\sigma_D) \quad (82)$$

yields the following quadratic polynomial in ν_2 :

$$\eta_0(c_y(\sigma_D))\nu_2^2 + \eta_1(c_y(\sigma_D))\nu_2 + \eta_2(c_y(\sigma_D)) = 0, \quad (83)$$

where

$$\begin{aligned} \eta_0(x) &\triangleq 2\sigma_D^2(x - b^4\sigma_D^2 - 2a^2b^4\sigma_D^2 - 2a^2b^2 + a^4b^4\sigma_D^2), \\ \eta_1(x) &\triangleq 4\sigma_D^2b^2\left(b^2(1 + a^2)(2a^2 + (a^2 + 1)b^2\sigma_D^2) - x\right), \\ \eta_2(x) &\triangleq x^2 - b^6\sigma_D^2(1 + a^2)^2((1 + a^2)^2b^2\sigma_D^2 + 4a^2). \end{aligned} \quad (84)$$

Now, we note that using the inequality in (80) in (76) shows that, when $k_\Sigma = \frac{k'}{2} - r$, the index of the objective is greater than 1. Hence, for any feasible ν_1 and σ_D (yielding $\lambda_{\tilde{B}_1}^{-1}$), the objective is maximized when k_Σ is maximum, that is, when $k' = k$ and $r = 0$.

Next, we consider the case of $k_\Sigma = \frac{k'}{2} + r$. We will show that for this case also setting $k' = k$ and $r = 0$ is optimal as $c_1 \rightarrow 0$.

2) *Case of $k_\Sigma = \frac{k'}{2} + r$:* We will show that for the asymptotic case of $c_1 \rightarrow 0$, the optimal number of interlacing subdiagonal entries is $k' = k$ and the optimal $r = 0$. For nonasymptotic cases, the optimal number of interlacing subdiagonal entries is also $k' = k$, but in those cases the optimal r is either $r = 0$ or $r = \frac{k}{2} - 1$. We begin by noting that, for any given $r \in \{1, \dots, \frac{k}{2} - 1\}$, when $k_\Sigma = \frac{k'}{2} + r$, the eigenvectors of Σ are the Schur vectors corresponding to the two smallest distinct eigenvalues of \tilde{B} : $\lambda_{\tilde{B}_0}^{-1}$, which has multiplicity $2r$ and $\lambda_{\tilde{B}_1}^{-1}$, which has multiplicity $\frac{k'}{2} - r$. The matrix \tilde{U}_1 in this case is given by $\tilde{U}_1 = [\tilde{U}_{\tilde{B}_0} \quad \tilde{U}_{\tilde{B}_1}]$, and subsequently, it can be readily seen that

$$\tilde{U}_1^T D^T D \tilde{U}_1 = 0_{2r \times 2r} \oplus u_{\tilde{B}_1}^2 \sigma_D^2 I_{\frac{k'}{2} - r}. \quad (85)$$

We will now examine the objective and the constraints of (26) for this case.

a) *Objective in (26a):* The objective is given by

$$\begin{aligned} \nu_1^{-k_\Sigma} \det(\tilde{U}_1^T \tilde{B} \tilde{U}_1) &= \nu_1^{-k_\Sigma} \lambda_{\tilde{B}_1}^{-\left(\frac{k'}{2} - r\right)} N^{2r} \\ &= (\nu_1 \lambda_{\tilde{B}_1}^{-1})^{-\frac{k'}{2}} (\nu_1^{-1} \lambda_{\tilde{B}_1}^{-1} N^2)^r. \end{aligned} \quad (86)$$

b) *Constraint in (26c):* From (85), we have

$$(I_{k_\Sigma} + \nu_2 \tilde{U}_1^T D^T D \tilde{U}_1)^{-1} = I_{2r} \oplus \frac{1}{1 + \nu_2 u_{\tilde{B}_1}^2 \sigma_D^2} I_{\frac{k'}{2} - r}.$$

Using this in (77), the constraint in (26c) can be expressed in the following form:

$$\frac{(\nu_1 - \lambda_{\tilde{B}_1}^{-1})\left(\frac{k'}{2} - r\right)}{1 + \nu_2 u_{\tilde{B}_1}^2 \sigma_D^2} + 2r(\nu_1 - N) = c_1. \quad (87)$$

c) *Constraint in (26d):* The first $2r$ diagonal entries of $\tilde{U}_1^T (H^T H)^{-1} (I_k + \nu_2 D^T D) \tilde{U}_1$ are equal to N and the next $\frac{k'}{2} - r$ entries are equal to $\lambda_{\tilde{B}_1}^{-1}$. Hence, the constraint in (26d) can be written as

$$\begin{aligned} 2rN + \left(\frac{k'}{2} - r\right) \lambda_{\tilde{B}_1}^{-1} &= \nu_1 k_\Sigma - c_1 \\ &- \frac{\nu_2}{a^2} \left(c_2 - N\left(\frac{k'}{2} - r\right)\sigma_D^2 - N\sum_{i=k'}^{k-1}\sigma_i^2\right). \end{aligned} \quad (88)$$

Similar to the discussion following (79), it can be seen that the optimal k' must be equal to k , because the $k - k'$ entries $\{\sigma_i\}_{k'}^{k-1}$ do not appear except in this constraint and their effect is to reduce c_2 .

d) *Constraint in (26b):* This constraint is equivalent to the constraint that

$$\nu_1 \geq \max\left\{\lambda_{\tilde{B}_1}^{-1}, \lambda_{\tilde{B}_0}^{-1}\right\} = \max\left\{\lambda_{\tilde{B}_1}^{-1}, N\right\}. \quad (89)$$

e) *Deriving an auxiliary constraint from (87) and (88):* Analogous to the case of $k_\Sigma = \frac{k'}{2} - r$, we now derive a simplified constraint from (87) and(88) for the current case of $k_\Sigma = \frac{k'}{2} + r$. Combining (87) and(88) yields

$$u_{\tilde{B}_1}^2 = \frac{c_2 - N\left(\frac{k'}{2} - r\right)\sigma_D^2 - N\sum_{i=k'}^{k-1}\sigma_i^2}{(c_1 + 2r(N - \nu_1))a^2\sigma_D^2}. \quad (90)$$

Let $c_z(\sigma_D, \nu_1) \triangleq \frac{2a^4 b^2 (c_1 + 2r(N - \nu_1))\sigma_D^2}{c_2 - N\left(\frac{k'}{2} - r\right)\sigma_D^2 - N\sum_{i=k'}^{k-1}\sigma_i^2}$, and $c_w(\sigma_D, \nu_1) \triangleq \sigma_D^2 (1 + a^2)^2 b^4 + 4a^2 b^2 - c_z(\sigma_D, \nu_1)$. Using these definitions and simplifying yields a quadratic polynomial in ν_2 similar to the one in (83), but with coefficients $\eta_i(c_w(\sigma_D, \nu_1))$, $i = 0, 1, 2$, where $\{\eta_i\}_{i=0}^2$ are defined in (84).

Now, we note that the inequality in (89) implies that the term $\nu_1 \lambda_{\tilde{B}_1}^{-1}$ in (86) is greater than 1. Hence, the optimal k' in this case is always equal to k . However, the term $\nu_1^{-1} \lambda_{\tilde{B}_1}^{-1} N^2$ may be greater or less than 1. If, for a feasible ν_1 and σ_D (yielding $\lambda_{\tilde{B}_1}^{-1}$), this term is greater than 1, the optimal r is equal to 0. On the other hand, if this term is less than 1, the optimal $r = \frac{k'}{2} - 1 = \frac{k}{2} - 1$, which, from (75), implies that $\text{rank}(D) = k_D = 1$.

To complete the proof of Proposition 3 we will now show that, similar to the previous case of $k_\Sigma = \frac{k}{2} - r$, when $k_\Sigma = \frac{k}{2} + r$ the asymptotically optimal D as $c_1 \rightarrow 0$ must have $r = 0$. To do so, we note that since ν_1 satisfies the inequality in (89), each term on the left-hand side of (87) must be nonnegative. This implies that, as $c_1 \rightarrow 0$, both terms must tend to 0. Since $r < \frac{k}{2}$, we must have $\nu_1 \searrow \lambda_{\tilde{B}_1}^{-1}$ and at the same time either $r = 0$ or $\nu_1 \searrow N$. The latter possibility can be dismissed by using (88) to show that if $\nu_1 \searrow \lambda_{\tilde{B}_1}^{-1}$ and $\nu_1 \searrow N$, $c_1 \nearrow 0$, that is, in this case c_1 approaches 0 from the left, which contradicts the fact that $c_1 > 0$. Hence, we conclude that as $c_1 \rightarrow 0$, r must be equal to 0, which completes the proof of the proposition.

APPENDIX H

OPTIMAL INTERLACING PRECODER: CASE OF $\Phi = 0$

When $\Phi = 0$, the optimal precoder at asymptotically low transmit power possesses an interlacing subdiagonal structure analogous to the one in the case of $\Phi \neq 0$. Since a complete search for the optimal precoder requires considering both cases, we herein provide expressions for this case. The derivation of these expressions resembles that in Appendix G and is omitted for brevity.

When $\Phi = 0$, it can be seen from (12) that $\tilde{U}_1 = I_k$. In that case, it can be verified that

- the objective in (26a) is given by

$$\begin{aligned} \nu_1^{-k} \det(\tilde{B}) &= (\nu_1^2 \lambda_{\tilde{B}_1^{-1}} \lambda_{\tilde{B}_2^{-1}})^{-k/2} \\ &= \left(\frac{N}{\nu_1} \sqrt{(1 + b^2 \sigma_D^2)(1 + \nu_2 \sigma_D^2)} \right)^k, \end{aligned} \quad (91)$$

- the constraint in (26b) is equivalent to

$$\nu_1 \lambda_{\tilde{B}_2^{-1}} \geq 1, \quad (92)$$

where $\lambda_{\tilde{B}_2^{-1}}$ is given in (71),

- the constraint in (26c) can be shown to be equivalent to

$$\nu_1 = \frac{1 + \nu_2 \sigma_D^2}{2 + \nu_2 \sigma_D^2} \left(N(2 + (1 + a^2)b^2 \sigma_D^2) + \frac{2c_1}{k} \right), \quad \text{and} \quad (93)$$

- the constraint in (26d) is given by

$$\begin{aligned} \frac{Nk}{2} (2 + \nu_2 \sigma_D^2 + (1 + a^2)b^2 \sigma_D^2) \\ = \nu_1 k - c_1 - \frac{\nu_2}{a^2} \left(c_2 - \frac{Nk}{2} \sigma_D^2 \right). \end{aligned} \quad (94)$$

Using (93) and (94) to solve for ν_2 , we obtain

$$\nu_2 = \max \left\{ \frac{a^2(1 + a^2)b^2 Nk \sigma_D^4 + 2\sigma_D^2(c_1 a^2 + Nk) - 4c_2}{\sigma_D^2(Nk \sigma_D^2(a^2 - 1) + 2c_2)}, 0 \right\}. \quad (95)$$

Using (91), (92), (93), and (95), the $\Phi = 0$ counterpart of Algorithm 2 can be readily derived.

ACKNOWLEDGMENT

The authors would like to thank C. (Sekhar) Bontu, C. Royer, and J. Womack of BlackBerry (formerly Research-In-Motion (RIM)) for their comments and support. The authors would also like to express their gratitude to the Associate Editor and the anonymous Reviewers for their constructive remarks and suggestions.

REFERENCES

- [1] R. Pabst, B. H. Walke, D. C. Schultz, P. Herold, H. Yanikomeroglu, S. Mukherjee, H. Viswanath, M. Lott, W. Zirwas, M. Dohler, H. Aghvami, D. D. Falconer, and G. P. Fettweis, "Relay-based deployment concepts for wireless and mobile broadband radio," *IEEE Commun. Mag.*, vol. 42, no. 9, pp. 80–89, Sep. 2004.
- [2] G. Kramer, I. Maric, and R. D. Yates, "Cooperative communications," *Found. Trend. Commun. Inf. Theory*, vol. 1, no. 3–4, pp. 271–425, 2006.
- [3] T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 5, pp. 572–584, Sep. 1979.
- [4] A. A. El Gamal and S. Zahedi, "Capacity of a class of relay channels with orthogonal components," *IEEE Trans. Inf. Theory*, vol. 51, no. 5, pp. 1815–1817, May 2005.
- [5] M. N. Khormuji and M. Skoglund, "On instantaneous relaying," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3378–3394, Jul. 2010.
- [6] A. A. El Gamal, M. Mohseni, and S. Zahedi, "Bounds on capacity and minimum energy-per-bit for AWGN relay channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1545–1561, Apr. 2006.
- [7] B. Wang, J. Zhang, and A. Host-Madsen, "On the capacity of MIMO relay channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 29–43, Jan. 2005.
- [8] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.
- [9] X. Tang and Y. Hua, "Optimal design of non-regenerative MIMO wireless relays," *IEEE Trans. Wireless Commun.*, vol. 6, no. 4, pp. 1398–1407, Apr. 2007.
- [10] O. Muñoz-Medina, J. Vidal, and A. Agustín, "Linear transceiver design in nonregenerative relays with channel state information," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2593–2604, Jun. 2007.
- [11] Y. Rong, "Linear non-regenerative multicarrier MIMO relay communications based on MMSE criterion," *IEEE Trans. Commun.*, vol. 58, no. 7, pp. 1918–1923, Jul. 2010.
- [12] C. Kim, Y. Sung, and Y. H. Lee, "A joint time-invariant filtering approach to the linear Gaussian relay problem," *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 4360–4375, Aug. 2012.
- [13] Y. Sung and C. Kim, "The capacity for the linear time-invariant Gaussian relay channel," Sep. 2011 [Online]. Available: <http://arxiv.org/abs/1109.5426>
- [14] A. El Gamal, N. Hassanpour, and J. Mammen, "Relay networks with delays," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3413–3431, Oct. 2007.
- [15] M. N. Khormuji and M. Skoglund, "Capacity bounds and mapping design for binary symmetric relay channels," *Entropy*, vol. 14, pp. 2589–2610, Dec. 2012.
- [16] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley, 1991.
- [17] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [18] D. P. Palomar and S. Verdú, "Gradient of mutual information in linear vector Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 141–154, Jan. 2006.
- [19] A. Graham, *Kronecker Products and Matrix Calculus: With Applications*. New York, NY, USA: Ellis Horwood Ltd., 1981.
- [20] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1999.
- [21] M. J. Wainwright and M. I. Jordan, "Log-determinant relaxation for approximate inference in discrete Markov random fields," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2099–2109, Jun. 2006.
- [22] P. Henrici, "Bounds for eigenvalues of certain tridiagonal matrices," *J. Soc. Ind. Appl. Math.*, vol. 11, pp. 281–290, Jun. 1963.
- [23] J. Cheng and T. Berger, "On minimal eigenvalues of a class of tridiagonal matrices," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5024–5031, Nov. 2009.
- [24] B. N. Parlett, *The Symmetric Eigenvalue Problem.*, ser. Prentice-Hall Series Computational Mathematics. Toronto, ON, Canada: Prentice-Hall, 1980.
- [25] G. H. Golub and C. F. van Loan, *Matrix Computations*, First ed. Baltimore, MD, USA: The Johns Hopkins Univ. Press, 1983.

Ramy H. Gohary (S'02–M'06–SM'13) received the B.Sc. (Hons.) degree from Assiut University, Egypt in 1996, the M.Sc. degree from Cairo University, Egypt, in 2000, and the Ph.D. degree from McMaster University, ON, Canada in 2006, all in electronics and communications engineering. He received the Natural Sciences and Engineering Research Council visiting fellowship award in 2007.

Dr. Gohary was a visiting scientist with the Terrestrial Wireless Systems Branch, Communications Research Centre, Canada. He is currently the project manager of the Carleton-BlackBerry (formerly Research In Motion (RIM)) research project.

His research interests include analysis and design of MIMO wireless communication systems, applications of optimization and geometry in signal processing and communications, information theoretic aspects of multiuser communication systems, and applications of iterative detection and decoding techniques in multiple antenna and multiuser systems.

Halim Yanikomeroglu (S'96–M'98–SM'12) was born in Giresun, Turkey, in 1968. He received the B.Sc. degree in Electrical and Electronics Engineering from the Middle East Technical University, Ankara, Turkey, in 1990, and the M.A.Sc. degree in Electrical Engineering (now ECE) and the Ph.D. degree in Electrical and Computer Engineering from the University of Toronto, Canada, in 1992 and 1998, respectively. He was with the R&D Group of Marconi Kominikasyon A.S., Ankara, Turkey, during 1993–1994.

Since 1998 Dr. Yanikomeroglu has been with the Department of Systems and Computer Engineering at Carleton University, Ottawa, where he is now a Full Professor. His research interests cover many aspects of wireless technologies with a special emphasis on cellular networks. Dr. Yanikomeroglu coauthored more than 50 journal and 160 conference papers, and has given more than 20 tutorials in leading international conferences on wireless technologies. In recent years, Dr. Yanikomeroglu's research has been funded by Huawei, RIM, Samsung, Communications Research Centre of Canada (CRC), Telus, and Nortel. This collaborative research resulted in about 15 patent applications.

Dr. Yanikomeroglu is a member of the Steering Committee of the IEEE Wireless Communications and Networking Conference (WCNC), and has been involved in the organization of this conference over the years, including serving

as the Technical Program Co-Chair of WCNC 2004 and the Technical Program Chair of WCNC 2008; he will also serve as the Technical Program Co-Chair of WCNC 2014 to be held in Istanbul. Dr. Yanikomeroglu was the General Co-Chair of the IEEE Vehicular Technology Conference Fall 2010 held in Ottawa. Dr. Yanikomeroglu has served in the editorial boards of the IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and *IEEE Communications Surveys & Tutorials*. He was the Chair of the IEEE's Technical Committee on Personal Communications (now called, Wireless Technical Committee). He is a Distinguished Lecturer for the IEEE Vehicular Technology Society.

Dr. Yanikomeroglu is a recipient of the Carleton University Faculty Graduate Mentoring Award in 2010, the Carleton University Graduate Students Association Excellence Award in Graduate Teaching in 2010, and the Carleton University Research Achievement Award in 2009. Dr. Yanikomeroglu spent the 2011–2012 academic year at TOBB University of Economics and Technology, Ankara, Turkey, as a Visiting Professor. He is a registered Professional Engineer in the province of Ontario, Canada.