The Confidential MIMO Broadcast Capacity: A Simple Derivation

Anatoly Khina
EE—Systems Dept., TAU
Tel Aviv, Israel
Email: anatolyk@eng.tau.ac.il

Yuval Kochman
School of CSE, HUJI
Jerusalem, Israel
Email: yuvalko@cs.huji.ac.il

Ashish Khisti
ECE Dept., U. of Toronto
Toronto, Canada
Email: akhisti@comm.utoronto.ca

***Abstract*—We consider the problem of transmitting confidential messages over a two-user broadcast multiple-input multiple-output (MIMO) channel. Surprisingly, the capacity region of this setting under a covariance matrix constraint was shown by Liu et al. to be rectangular. That is, there is no tension, and both users can attain their respective MIMO wiretap capacities, *simultaneously*. In this work, we provide a new derivation of this result by proposing an alternative achievability scheme for the corner point of the capacity region. This derivation, in addition to being considerably shorter and simpler than the original, also provides a practical transmission scheme, in the sense that the codes used are scalar (single-antenna) ones. We use two main ingredients. The first is the explicit optimal input covariance matrix of Bustin et al. for the MIMO wiretap channel under a covariance matrix constraint, which we also re-derive in a simple manner. The second is a dirty-paper variant of a recently proposed optimal scheme for the MIMO wiretap channel, which uses scalar codes. The proposed treatment demonstrates the connection between the confidential broadcast problem and the MIMO wiretap one: the former almost reduces to the latter, except for the use of dirty-paper coding which is not mandatory in MIMO wiretap; the work sheds light on the reason for this difference.**

***Index Terms*—Confidential broadcast, wiretap channel, MIMO channel, dirty-paper coding, matrix decomposition**

I. Introduction

The confidential two-user broadcast (BC) channel is composed of a sender ("Alice") who wishes to convey different data to two users ("Bob" and "Charlie"), such that no information can be recovered by one user about the data intended for the other user. The Gaussian multiple-input multiple-output (MIMO) variant of this scenario, considered first in [1], is given by[1]

$$\boldsymbol{y}_B = \mathbf{H}_B \boldsymbol{x}_A + \boldsymbol{z}_B \tag{1a}$$

$$\boldsymbol{y}_C = \mathbf{H}_C \boldsymbol{x}_A + \boldsymbol{z}_C, \tag{1b}$$

where $\boldsymbol{y}_B$ and $\boldsymbol{y}_C$ are the received vector signals by Bob and Charlie, respectively, of lengths N_B and N_C; $\boldsymbol{x}_A$ is the transmitted vector signal by Alice of length N_A; $\boldsymbol{z}_B$ and $\boldsymbol{z}_C$ are Gaussian noise vectors, that are assumed, w.l.o.g., to be circularly-symmetric with zero mean and unit covariance matrix. The channel matrices $\mathbf{H}_B$ and $\mathbf{H}_C$ have the corresponding dimensions. The capacity region (under a constraint on the input) is the closure of the rates (R_B, R_C) such that reliable decoding and secrecy are guaranteed.[2]

The confidential BC channel can be seen as a generalization of the MIMO wiretap channel [3], [4], where no information is sent to Charlie ($R_C = 0$). Hence, it is usually referred to as an eavesdropper "Eve". Indeed, there is also a very close connection between the solutions to these two problems. For the case where the input is subject to an average *covariance constraint*

$$\mathbf{K}_A \triangleq E\left[\boldsymbol{x}\boldsymbol{x}^\dagger\right] \preceq \mathbf{K}, \tag{2}$$

Liu et al. [5] established the capacity region by showing that it is rectangular. Namely, it is given by

$$R_B \le C(\mathbf{H}_B, \mathbf{H}_C, \mathbf{K}) \tag{3a}$$

$$R_C \le C(\mathbf{H}_C, \mathbf{H}_B, \mathbf{K}), \tag{3b}$$

where $C(\mathbf{H}_B, \mathbf{H}_C, \mathbf{K})$ is the capacity of the MIMO wiretap channel when Charlie acts the part of an eavesdropper, under a covariance input constraint. The converse is immediate, as both users achieve their maximal possible secrecy rates simultaneously; it is the direct part that is quite striking. The MIMO wiretap capacity under a covariance constraint was, in turn, shown by Liu and Shamai [6] to be achieved by a Gaussian input; the solution is given as a maximization over covariance matrices satisfying the constraint (2):

$$C(\mathbf{H}_B, \mathbf{H}_C, \mathbf{K}) = \max_{\mathbf{K}_A \preceq \mathbf{K}} I_S(\mathbf{H}_B, \mathbf{H}_C, \mathbf{K}_A) \tag{4}$$

where

$$I_S(\mathbf{H}_B, \mathbf{H}_C, \mathbf{K}) \triangleq I(\mathbf{H}_B, \mathbf{K}) - I(\mathbf{H}_C, \mathbf{K}),$$

and the Gaussian vector mutual information (MI) is

$$I(\mathbf{H}, \mathbf{K}) = \log\det\left\{\mathbf{I} + \mathbf{H}\mathbf{K}\mathbf{H}^\dagger\right\}. \tag{5}$$

Later, Bustin et al. [7] provided an explicit solution to this maximization problem. In order to obtain these results, all of the works [5]–[7] used heavy machinery such as channel enhancement and vector extensions of the I-MMSE relation.

In this work, we show that some of these results can be derived in a simpler manner, once we know that the solution to the MIMO wiretap problem is Gaussian (that is, given (4)).

[1]The Gaussian single-input single-output (SISO) scenario reduces to messages for the stronger user only (the Gaussian SISO wiretap channel [2]), as the BC channel is degraded.

[2]Throughout the paper, we are only interested in weak secrecy.

As an added value, we use constructive proofs which provide schemes that are practical, in the sense that they employ scalar (SISO) codes for the MIMO secrecy problems at hand.

In Section II we re-interpret the explicit solution of [7] in terms of the generalized singular value decomposition (GSVD) [8], [9], and then derive it from (4) using only linear algebra, without any information-theoretic considerations.

Then, we note that this solution seems related to the confidential BC channel. Namely, some "directions" are useful for Bob, while others would be useful for Charlie if we inverted the roles. However, we need yet another ingredient. To that end, we present in Section III a dirty-paper coding (DPC) variant of a recently proposed successive interference cancellation (SIC) scalar-codes scheme [10] for the MIMO wiretap channel.

Finally in Section IV we use the above to construct a DPC scalar-codes scheme for the confidential BC channel. This scheme is optimal, thus analyzing its performance provides an alternative achievability proof for the MIMO confidential BC capacity [5].

II. MIMO Wiretap and Confidential Broadcast Capacitities

In this section we re-derive the result of Bustin et al. [7] in terms of the GSVD [8], [9].

To that end, construct the augmented matrices $\mathbf{G}_B = G(\mathbf{H}_B, \mathbf{K})$ and $\mathbf{G}_C = G(\mathbf{H}_C, \mathbf{K})$, where[3]

$$G(\mathbf{H}, \mathbf{K}) \triangleq \begin{pmatrix} \mathbf{H}\mathbf{K}^{1/2} \\ \mathbf{I} \end{pmatrix} . \tag{6}$$

Recall that $\mathbf{K}$ is the constraining covariance matrix (2). Now choose some unitary matrix $\mathbf{V}_A$ and apply the QR decompositions:

$$\mathbf{G}_B \mathbf{V}_A = \mathbf{U}_B \mathbf{T}_B \, , \tag{7a}$$

$$\mathbf{G}_C \mathbf{V}_A = \mathbf{U}_C \mathbf{T}_C \, , \tag{7b}$$

where $\mathbf{U}_B$ and $\mathbf{U}_C$ are unitary, and $\mathbf{T}_B$ and $\mathbf{T}_C$ are (generalized) upper-triangular of dimensions $(N_B + N_A) \times N_A$ and $(N_C + N_A) \times N_A$, respectively, viz. $T_{B;ij} = T_{C;ij} = 0$ for $i > j$. We have obtained a family of joint unitary decompositions, depending on the choice of $\mathbf{V}_A$. Let $\{b_i\}$ and $\{c_i\}$ denote the diagonal values of $\mathbf{T}_B$ and $\mathbf{T}_C$, respectively. Then, the Gaussian MI (5) satisfies:

$$I(\mathbf{H}_B, \mathbf{K}) = \log\det\left\{\mathbf{G}_B^\dagger \mathbf{G}_B\right\} = \sum \log b_i^2 \tag{8}$$

and similarly for Charlie. Thus, for any $\mathbf{V}_A$,

$$I_S(\mathbf{H}_B, \mathbf{H}_C, \mathbf{K}) = \sum_{i=1}^{N_A} \log \frac{b_i^2}{c_i^2} \, . \tag{9}$$

A special choice of $\mathbf{V}_A$ gives the GSVD,[4] where the generalized singular values (GSVs) are given by $\mu_i\left(\mathbf{H}_B, \mathbf{H}_C, \mathbf{K}\right) \triangleq b_i/c_i$. Without loss of generality, we assume that the GSV vector is non-increasing. In terms of the GSVs, we can re-write (4) as:

$$C(\mathbf{H}_B, \mathbf{H}_C, \mathbf{K}) = \max_{\mathbf{K}_A \preceq \mathbf{K}} \sum_{i=1}^{N_A} \log \mu_i^2 \left(\mathbf{H}_B, \mathbf{H}_C, \mathbf{K}_A\right) .$$

Indeed, in these terms the capacity expression of [7] can be re-stated as follows.

Theorem 1: The secrecy capacity under a covariance matrix constraint $\mathbf{K}$ is equal to

$$C(\mathbf{H}_B, \mathbf{H}_C, \mathbf{K}) = \sum_{i=1}^{N_A} \left[\log \mu_i^2 \left(\mathbf{H}_B, \mathbf{H}_C, \mathbf{K}\right)\right]^+ ,$$

where $[x]^+ \triangleq \max\{0, x\}$.

The key to our proof of this result is the following lemma.

Lemma 1: Let $\mathbf{K}$ and $\mathbf{K}_A$ be two matrices satisfying $\mathbf{0} \preceq \mathbf{K}_A \preceq \mathbf{K}$. Then for all $i = 1, \ldots, N_A$,

$$\left|\log \mu_i^2(\mathbf{H}_b, \mathbf{H}_C, \mathbf{K})\right| \geq \left|\log \mu_i^2(\mathbf{H}_b, \mathbf{H}_C, \mathbf{K}_A)\right| .$$

That is, as we "decrease" the input covariance, the GSVs move towards $\mu_i = 1$. The proof, which appears in Appendix B, uses standard matrix calculus to show that the differential of the i-th GSV, $d\mu_i$, w.r.t. a change in the covariance matrix $d\mathbf{K}_A$, is given by

$$d\mu_i = \left(\mu_i^2 - 1\right) \cdot \gamma_i(d\mathbf{K}_A) \, ,$$

where $\gamma_i(d\mathbf{K}_A) \geq 0$ for $d\mathbf{K}_A \succeq \mathbf{0}$.

By Lemma 1, clearly Theorem 1 gives an upper bound on the capacity. To see that it is achievable, consider the matrix:

$$\mathbf{K}_B = \mathbf{K}^{1/2} \mathbf{V}_A \mathbf{I}_B \mathbf{V}_A^\dagger \mathbf{K}^{1/2\dagger} \tag{10}$$

where $\mathbf{V}_A$ is the right unitary matrix of the triangular form of the GSVD of $\mathbf{G}_B$ and $\mathbf{G}_C$ (7), $\mathbf{I}_B$ is a diagonal matrix whose diagonal values corresponding to GSVs that are greater than 1 — equal to 1, and the others are 0. Trivially, $\mathbf{K}_B \preceq \mathbf{K}$. The choice of $\mathbf{K}_B$ effectively truncates the GSVs of $\mathbf{K}$:

$$\log \mu_i^2 \left(\mathbf{H}_B, \mathbf{H}_C, \mathbf{K}_B\right) = \left[\log \mu_i^2 \left(\mathbf{H}_B, \mathbf{H}_C, \mathbf{K}\right)\right]^+ .$$

This is formally proved in Appendix A.

Remark 1: The covariance matrix $\mathbf{K}_B$ (10) is called $\mathbf{K}_x^*$ in [7], where it is given in terms of the diagonal form of the GSVD (see also Appendix A):[5]

$$\mathbf{K}_B = \mathbf{K}^{1/2} \mathbf{Y} \begin{bmatrix} \left(\mathbf{Y}_B^\dagger \mathbf{Y}_B\right)^{-1} & \mathbf{0}_{L_B \times L_C} \\ \mathbf{0}_{L_C \times L_B} & \mathbf{0}_{L_C \times L_C} \end{bmatrix} \mathbf{Y}^\dagger \mathbf{K}^{1/2} \, , \tag{12}$$

where $\mathbf{Y} = \mathbf{X}^{-\dagger}$ and $\mathbf{X}$ is the right invertible matrix of the diagonal form of the GSVD, L_B and L_C denote the GSVs that are greater and smaller than 1, respectively, $\mathbf{Y}_B$ is the submatrix composed of the first L_B columns of $\mathbf{Y}$, and $\mathbf{0}_{m \times n}$ denotes the all-zero matrix of dimensions $m \times n$. Using the triangular form of the GSVD, simplifies the presentation and

[3] $\mathbf{K}^{1/2}$ is any matrix $\mathbf{B}$ satisfying $\mathbf{B}\mathbf{B}^\dagger = \mathbf{K}$.

[4] Here we use the triangular form of the GSVD; see Appendix A for further details.

[5] In [7] a specific choice of $\mathbf{K}^{1/2}$ was used: the matrix $\mathbf{B}$ that satisfies $\mathbf{B}\mathbf{B} = \mathbf{K}$.

construction of $\mathbf{K}_B$, as is evident from comparing (10) with (12)

Remark 2: One may wonder why, of all possible choices of $\mathbf{V}_A$, the capacity is given in terms of the GSVD. An intuitive reason is as follows. Among all achievable diagonal ratios, the GSV series is the "least balanced" possible in a multiplicative majorization sense [11]. In particular, for any $\mathbf{V}_A$,

$$\sum_{i=1}^{N_A}\left[\log \mu_i^2\right]^+ \geq \sum_{i=1}^{N_A}\left[\log \frac{b_i^2}{c_i^2}\right]^+ .$$

Remark 3: Denote the capacity of the MIMO wiretap channel under a power constraint P by $C(\mathbf{H}_B, \mathbf{H}_C, P)$. Since by, e.g., [12, Lemma 1],

$$C(\mathbf{H}_B, \mathbf{H}_C, P) = \max_{\mathbf{K}:\text{trace}\{\mathbf{K}\}\leq P} C(\mathbf{H}_B, \mathbf{H}_C, \mathbf{K}),$$

it follows that

$$C(\mathbf{H}_B, \mathbf{H}_C, P) = \max_{\mathbf{K}:\text{trace}\{\mathbf{K}\}\leq P} \sum_{i=1}^{N_A}\left[\log \mu_i^2\left(\mathbf{H}_B, \mathbf{H}_C, \mathbf{K}\right)\right]^+ .$$

For the optimal $\mathbf{K}$, all the GSVs are greater or equal to 1. To the contrary, assume that some are strictly smaller than 1; then, we can use a matrix $\mathbf{K}_A$ with the appropriate directions "nullified". But since $\text{trace}\{\mathbf{K}_A\} < \text{trace}\{\mathbf{K}\} \leq P$, we can then use amplification to improve the rate.

Now we note that, if we were interested in confidential communication with Charlie rather than with Bob, we would get the same solution with the roles of $\mathbf{H}_B$ and $\mathbf{H}_C$ reversed. But then, this means inversion of the GSVs:

$$\log \mu_i(\mathbf{H}_C, \mathbf{H}_B, \mathbf{K}) = -\log \mu_i(\mathbf{H}_B, \mathbf{H}_C, \mathbf{K}).$$

Thus, we can write the rectangular capacity-region of the confidential BC channel (3) as follows.

Theorem 2: The capacity region of the confidential MIMO BC channel under an input covariance constraint $\mathbf{K}$ is given by all rates (R_B, R_C) satisfying:

$$\begin{aligned} R_B &\leq \sum_{i=1}^{N_A}\left[\log \mu_i^2\left(\mathbf{H}_B, \mathbf{H}_C, \mathbf{K}\right)\right]^+ \\ R_C &\leq \sum_{i=1}^{N_A}\left[-\log \mu_i^2\left(\mathbf{H}_B, \mathbf{H}_C, \mathbf{K}\right)\right]^+ . \end{aligned}$$

The converse part of this result is trivial by Theorem 1. For the direct part, it is tempting to think that since different GSVs are nullified for Bob and for Charlie, Alice can achieve their optimal rates simultaneously by communicating over orthogonal "subspaces". However, since the matrices $\mathbf{T}_B$ and $\mathbf{T}_C$ are not diagonal, these "subspaces" are not orthogonal, and some more care is needed. In the rest of this paper we develop a DPC scheme for the wiretap channel that leads to an optimal transmission scheme for the confidential BC channel. Thus, this derivation provides a proof for the direct part of Theorem 2, which is an alternative to the proof in [5].

Remark 4: Similarly to the MIMO wiretap channel, the capacity region under a power constraint P is just the union of all (rectangular) regions under a covariance constraint with small enough trace.

III. DPC-based Scalar Schemes for MIMO

In this section we present DPC-based schemes for the Gaussian MIMO channel (without secrecy) and the MIMO wiretap channel. These schemes, which build upon the matrix decomposition (7), allow to approach the optimal rate for any input covariance matrix, using scalar dirty-paper codes. SIC counterparts of these schemes were previously presented in [10].

A. Without Secrecy Constraints

We now briefly review the connection between matrix decompositions and scalar transmission schemes, without secrecy requirements. For a more thorough account, the reader is referred to [11], [13]. Consider the channel (1a). Recalling (6), construct the augmented matrix $\mathbf{G}_B = G(\mathbf{H}_B, \mathbf{K})$. For some unitary matrix $\mathbf{V}_A$,[6] decompose $\mathbf{G}_B$ as in (7a).

We start by describing a scheme that utilizes successive interference cancellation (SIC) to approach capacity using scalar codes. We then discuss a similar scheme that pre-cancels the interferences at the transmitter by means of DPC.

Let $\tilde{\boldsymbol{x}}$ be a vector of standard Gaussian variables, and set

$$\boldsymbol{x} = \mathbf{K}^{1/2}\mathbf{V}_A\tilde{\boldsymbol{x}}\,. \tag{13}$$

Denote by $\tilde{\mathbf{U}}_B$ the sub-matrix consisting of the upper-left $N_B \times N_A$ block of $\mathbf{U}_B$, define $\tilde{\mathbf{T}} = \tilde{\mathbf{U}}_B^\dagger\mathbf{K}^{1/2}\mathbf{V}_A$, and let

$$\tilde{\boldsymbol{y}}_B = \tilde{\mathbf{U}}_B^\dagger\boldsymbol{y}_B = \tilde{\mathbf{U}}_B^\dagger\mathbf{K}^{1/2}\mathbf{V}_A\tilde{\boldsymbol{x}} + \tilde{\mathbf{U}}_B^\dagger\boldsymbol{z}_B = \tilde{\mathbf{T}}\tilde{\boldsymbol{x}} + \tilde{\boldsymbol{z}}_B\,. \tag{14}$$

Since $\tilde{\mathbf{U}}_B$ is not unitary, the statistics of $\tilde{\boldsymbol{z}} \triangleq \tilde{\mathbf{U}}_B^\dagger\boldsymbol{z}$ differ from those of $\boldsymbol{z}$, and its covariance matrix is given by $\mathbf{K}_{\tilde{\boldsymbol{z}}} \triangleq \tilde{\mathbf{U}}_B\tilde{\mathbf{U}}_B^\dagger$. Now, for $i = 1, \ldots, N_A$, define

$$\begin{aligned} y'_{B;i} &= \tilde{y}_{B;i} - \sum_{\ell=i+1}^{N_A}\tilde{T}_{i,\ell}\tilde{x}_\ell \\ &= \tilde{T}_{i,i}\tilde{x}_i + \sum_{\ell=1}^{i-1}\tilde{T}_{i,\ell}\tilde{x}_\ell + \tilde{z}_i \triangleq \tilde{T}_{i,i}\tilde{x}_i + z_i^{\text{eff}}\,. \end{aligned} \tag{15}$$

In this scalar channel from $\tilde{x}_i$ to $y'_{B;i}$, we see other $\tilde{x}_\ell$ as "interference", $\tilde{z}_i$ — as "noise", and their sum z_i^{eff} — as "effective noise". The resulting signal-to-interference-and-noise ratio (SINR) is given by:

$$S_i \triangleq \frac{(\tilde{T}_{i,i})^2}{K_{\boldsymbol{z}^{\text{eff}};i,i}} \triangleq \frac{(\tilde{T}_{i,i})^2}{K_{\tilde{\boldsymbol{z}};i,i} + \sum_{\ell=1}^{i-1}(\tilde{T}_{i,\ell})^2}\,,$$

where $K_{\tilde{\boldsymbol{z}};i,j}$ denotes the (i,j) entry of $\mathbf{K}_{\tilde{\boldsymbol{z}}}$. The following key result achieves the mutual information [13, Lemma III.3], [14][7]

$$\begin{aligned} I\left(\tilde{x}_i; \boldsymbol{y}_B \middle| \tilde{x}_{i+1}^{N_A}\right) &= I\left(\tilde{x}_i; y'_{B;i}\right) \\ &= \log(1+S_i) = \log(b_i^2)\,. \end{aligned} \tag{16}$$

[6] See [10], [11], [13] for interesting choices of $\mathbf{V}_A$.

[7] Note that, even though $\tilde{\boldsymbol{z}}$ has dependent components, the entries of the effective noise $\boldsymbol{z}^{\text{eff}}$, are independent.

On account of (8), the sum of these rates amounts to $I(\mathbf{H}_B, \mathbf{K})$, which equals the channel capacity for the optimal $\mathbf{K}$.

This analysis leads immediately to an optimal SIC-based scheme, since the decoder can perform iteratively the interference cancellation (15). Indeed, such a scheme, which can be found in, e.g., [10], is a variant of the renowned V-BLAST/GDFE scheme [14]–[16]. A different approach is that of pre-canceling the interferences at the transmitter using DPC. Such pre-cancellation incurs no loss in performance compared to the interference-free channel [17]. This results in the following scheme.

Scheme 1 (MIMO point-to-point via DPC):

Offline: Construct N_A good dirty-paper codebooks as follows. Codebook i ($1 \le i \le N_A$) is constructed for a channel with AWGN of power 1, SNR $S_i = b_i^2 - 1$ and interference[8]

$$\sum_{\ell=i+1}^{N_A} T_{i,\ell}\tilde{x}_\ell$$

that is available as side information at the transmitter.

Alice: At each time instance:

- Generates $\tilde{x}_i$ from last ($i = N_A$) to first ($i = 1$), where $\tilde{x}_i$ is generated according to the message to be conveyed and the interference signals $\{\tilde{x}_\ell | \ell = i+1, \ldots, N_A\}$.
- Forms $\tilde{\boldsymbol{x}}$ with entries $\{\tilde{x}_i\}$.
- Transmits $\boldsymbol{x}$ according to (13):

$$\boldsymbol{x} = \mathbf{K}^{1/2}\mathbf{V}_A\tilde{\boldsymbol{x}}$$

.

Bob:

- At each time instance forms $\tilde{\boldsymbol{y}}_B$ according to (14):

$$\tilde{\boldsymbol{y}}_B = \tilde{\mathbf{U}}_B^\dagger \boldsymbol{y}_B \,.$$

- Decodes the codebooks using dirty-paper decoders, where $\tilde{x}_i$ is decoded from $\tilde{y}_{B;i}$.

By using good dirty-paper codes, capacity is achieved; see, e.g., [13].

Remark 5: The dirty-paper codes that are used can be generated together using random binning; in this case Costa's auxiliaries $\{u_i\}$ as well as all the interferences are Gaussian, and hence Costa's results follow through. Alternatively, one can generate the codebooks one-by-one and rely on the extended analysis for non Gaussian noise and interference of [18]. Furthermore, structured versions can be applied as well, which are valid for arbitrary interference sequences; see [19].

B. MIMO Wiretap

In this section we describe an optimal scheme for the MIMO wiretap channel using scalar dirty-paper wiretap codes. We note that a SIC-based counterpart of the scheme was presented in [10]. The scheme is optimal for any covariance matrix $\mathbf{K}$. Without loss of generality, we assume that $\mathbf{K}$ is such that $\log \mu_i(\mathbf{H}_B, \mathbf{H}_C, \mathbf{K}) \ge 0$ for all $i = 1, \ldots, N_A$; otherwise we can replace $\mathbf{K}$ by $\mathbf{K}_B$ (10).

[8]Note that $\tilde{T}_{i;\ell} = T_{i;\ell}$ for $\ell > i$.

Scheme 2 (MIMO wiretap via DPC):

Offline:

- Apply the QR decomposition to $\mathbf{G}_B\mathbf{V}_A$ and to $\mathbf{G}_C\mathbf{V}_A$, where $\mathbf{G}_B \triangleq G(\mathbf{H}_B, \mathbf{K})$ and $\mathbf{G}_C \triangleq G(\mathbf{H}_C, \mathbf{K})$:

$$\mathbf{G}_k\mathbf{V}_A = \mathbf{U}_k\mathbf{T}_k \,, \quad k \in \{B, C\} \,,$$

  $\{b_i\}$ and $\{c_i\}$ are the diagonal values of $\mathbf{T}_B$ and $\mathbf{T}_C$, respectively, and $\tilde{\mathbf{U}}_k$ is the upper-left $N_k \times N_A$ sub-matrix of $\mathbf{U}_k$.
- Construct good scalar wiretap codes as follows. Codebook i ($1 \le i \le N_A$) is of unit power with entries denoted by $\tilde{x}_i$ (with the time index omitted to simplify notation). It is constructed for an AWGN channel to Bob of SNR $b_i^2 - 1$ and interference

$$\sum_{\ell=i+1}^{N_A} T_{B;i,\ell}\tilde{x}_\ell \,, \tag{17}$$

  and for an AWGN channel to Charlie of SNR $c_i^2 - 1$ and interference

$$\sum_{\ell=i+1}^{N_A} T_{C;i,\ell}\tilde{x}_\ell \,. \tag{18}$$

Alice: At each time instance:

- Generates $\tilde{x}_i$ from last to first, where $\tilde{x}_i$ is generated according to the message to be conveyed and the interference signals $\{\tilde{x}_\ell | \ell = i+1, \ldots, N_A\}$.
- Forms $\tilde{\boldsymbol{x}}$ with entries $\{\tilde{x}_i\}$.
- Transmits $\boldsymbol{x}$ according to (13): $\boldsymbol{x} = \mathbf{K}^{1/2}\mathbf{V}_A\tilde{\boldsymbol{x}}$.

Bob:

- At each time instance forms $\tilde{\boldsymbol{y}}_B$ according to (14).
- Decodes the codebooks using dirty-paper decoders, where $\tilde{x}_i$ is decoded from $\tilde{y}_{B;i}$.

The following theorem proves the optimality of this scheme when using good scalar dirty-paper codes.

Theorem 3: Let $\epsilon > 0$, however small, and define $\xi = N_A\epsilon$. Then, for any covariance $\mathbf{K}$ and any unitary $\mathbf{V}_A$, there exist scalar codebooks of secrecy rates $R_i = \log(b_i^2/c_i^2) - \epsilon$, such that Scheme 2 achieves the secrecy rate $C(\mathbf{H}_B, \mathbf{H}_C, \mathbf{K}) - \xi$.

Proof: The proof follows by a standard extension of the proof of Theorem 3 of [10] (which is a specialization of Theorem 2 of [10] for the Gaussian MIMO setting) to the dirty-paper case.

Codebook construction: For each $k = 1, \ldots, N_A$, generate a codebook $\mathcal{C}$ of $2^{n(R_k + \tilde{R}_k)}$ sub-codebooks. Each such sub-codebook is assigned a unique index-pair (m_k, f_k), where $m_k \in \left\{1, 2, \ldots, 2^{nR_k}\right\}$ and $f_k \in \left\{1, 2, \ldots, 2^{n\tilde{R}_k}\right\}$, and contains $2^{n[\tilde{R}_k^{\mathrm{GP}} - (R_k + \tilde{R}_k)]}$ codewords. Each codeword is generated independently in an i.i.d. manner w.r.t. $p(\mathsf{u}_k)$ which is Gaussian with parameters dictated by

$$\mathsf{u}_k = \tilde{T}_{B;k,k}\tilde{\mathsf{x}}_k + \alpha_k \sum_{\ell=k+1}^{N_A} \tilde{T}_{B;k,\ell}\tilde{\mathsf{x}}_\ell \,,$$

$$\alpha_k \triangleq \frac{b_k^2-1}{b_k^2}\ ,$$

where $\{\tilde{\mathsf{x}}_k|k=1,\ldots,N_A\}$ are unit power i.i.d. Gaussian random variables.

The rates are chosen as follows.

$$\begin{aligned}
R_k &\triangleq \left[I\left(\mathsf{u}_k;\mathbf{y}_B\right)-I\left(\mathsf{u}_k;\mathsf{u}_{k+1}^{N_A}\right)\right]-I\left(\mathsf{u}_k;\mathbf{y}_E\middle|\mathsf{u}_{k+1}^{N_A}\right)-\epsilon\\
&= I\left(\mathsf{u}_k;\mathbf{y}_B\right)-I\left(\mathsf{u}_k;\mathbf{y}_E,\mathsf{u}_{k+1}^{N_A}\right)-\epsilon\\
&= I\left(\tilde{\mathsf{x}}_k;\mathbf{y}_B\middle|\tilde{\mathsf{x}}_{k+1}^{N_A}\right)-I\left(\tilde{\mathsf{x}}_k;\mathbf{y}_E\middle|\tilde{\mathsf{x}}_{k+1}^{N_A}\right)-\epsilon\,,\\
\tilde{R}_k &\triangleq I\left(\mathsf{u}_k;\mathbf{y}_E\middle|\mathsf{u}_{k+1}^{N_A}\right)-\epsilon\\
&= I\left(\tilde{\mathsf{x}}_k;\mathbf{y}_E\middle|\tilde{\mathsf{x}}_{k+1}^{N_A}\right)-\epsilon\,,\\
\tilde{R}_k^{\mathrm{GP}} &\triangleq I\left(\mathsf{u}_k;\mathbf{y}_B\right)-\epsilon,
\end{aligned}$$

for unit power i.i.d. Gaussian random variables $\{\tilde{\mathsf{x}}_k|k=1,\ldots,N_A\}$. The transitions above from u_k to $\tilde{x}_k$ are justified since the interference (transmitter side-information) in sub-channel k is composed of messages $\{x_\ell|\ell=1,\ldots,N_A\}$. Note that by (16), $R_k=\log(b_k^2/c_k^2)-\epsilon$, thus by (9) the sum of these rates approaches the desired secrecy rate $I_S(\mathbf{H}_B,\mathbf{H}_C,\mathbf{K})-\xi$.

Encoding (Alice): Encoding is carried in a successive manner, from last ($k=N_A$) to first ($k=1$). Within codebook k, the index of the sub-codebook to be used is determined by the secret message m_k and a fictitious message f_k drawn uniformly over their respective ranges. The codeword u_k, within sub-codebook (m_k,f_k) that is selected, is the one that is jointly typical with the side information $\sum_{\ell=k+1}^{N_A}\tilde{T}_{B;k,\ell}\tilde{x}_\ell$. If no such codeword u_k exists, then the first codeword is selected.

Decoding (Bob): Bob recovers (m_k,f_k) using standard dirty-paper decoding (as discussed in Section III-A) and discards f_k. The error probability can be made arbitrary small by taking large enough n.

Secrecy analysis (Charlie): By recalling that $\{\tilde{\mathsf{x}}_\ell|\ell=1,\ldots,N_A\}$ and $\{\mathsf{u}_\ell|\ell=1,\ldots,N_A\}$ carry the same information, the secrecy analysis is the same as in the proof of Theorem 2 in [10]. ■

IV. Confidential Broadcast Scheme

In view of Scheme 2, the result of Section II has a rather intuitive interpretation: $\mathbf{V}_A$ of the GSVD is the precoding matrix that designs the ratios between $\{b_i\}$ and $\{c_i\}$ to be as large as possible, which corresponds to maximizing the achievable secrecy rate to Bob. In order to achieve Bob's secrecy capacity, only the sub-channels for which the secrecy rate is positive ($b_i>c_i$) need to be utilized.

Allocating the remaining sub-channels to Charlie, on the other hand, attains Charlie's optimal covariance matrix.

Combining the two gives rise to the following scheme, which is a straightforward adaptation of Scheme 2.

Scheme 3 (Confidential Broadcast):

Offline:

- Apply the GSVD decomposition to $\mathbf{G}_B=G(\mathbf{H}_B,\mathbf{K})$ and to $\mathbf{G}_C=G(\mathbf{H}_C,\mathbf{K})$ as in (7).
- Denote the diagonal entries of $\mathbf{T}_B$ and $\mathbf{T}_C$ by $\{b_i\}$ and $\{c_i\}$, respectively.
- Denote further the (first) number of indices for which $b_i>c_i$ by L_B. The remaining $L_C=N_A-L_B$ indices satisfy $c_i\geq b_i$.
- Denote by $\tilde{\mathbf{U}}_B$ the upper-left $N_B\times L_B$ sub-matrix of $\mathbf{U}_B$, and by $\tilde{\mathbf{U}}_C$ — the upper-right $N_C\times L_C$ sub-matrix of $\mathbf{U}_C$.
- Construct N_A good scalar wiretap codes of unit power and length n, denoted by $\tilde{x}_i$ (with the time index omitted to simplify notation), as follows.
  - The first L_B codes are intended for Bob: Codebook $\tilde{x}_i$ ($1\leq i\leq L_B$) is constructed for an AWGN channel to Bob of SNR b_i^2-1 and interference (17):
  $$\sum_{\ell=i+1}^{N_A}T_{B;i,\ell}\tilde{x}_\ell\,,$$
  and for an AWGN channel to Charlie of SNR c_i^2-1 and interference (18):
  $$\sum_{\ell=i+1}^{N_A}T_{C;i,\ell}\tilde{x}_\ell\,.$$
  - The remaining L_C codes are intended for Charlie: Codebook $\tilde{x}_i$ ($L_B+1\leq i\leq N_A$) is constructed for an AWGN channel to Charlie of SNR c_i^2-1 and interference (18):
  $$\sum_{\ell=i+1}^{N_A}T_{C;i,\ell}\tilde{x}_\ell\,,$$
  and for an AWGN channel to Bob of SNR b_i^2-1 and interference (17):
  $$\sum_{\ell=i+1}^{N_A}T_{B;i,\ell}\tilde{x}_\ell\,.$$

Alice: At each time instance:

- Generates $\tilde{x}_i$ from last to first, where $\tilde{x}_i$ is generated according to the messages to be conveyed and the interference signals $\{\tilde{x}_\ell|\ell=i+1,\ldots,N_A\}$.
- Forms $\tilde{\boldsymbol{x}}$ with entries $\{\tilde{x}_i\}$.
- Transmits $\boldsymbol{x}$ according to (13):
$$\boldsymbol{x}=\mathbf{K}^{1/2}\mathbf{V}_A\tilde{\boldsymbol{x}}\,.$$

Bob:

- At each time instance forms
$$\tilde{\boldsymbol{y}}_B=\tilde{\mathbf{U}}_B^\dagger\boldsymbol{y}_B\,.$$
- Decodes codebooks $i=1,\ldots,L_B$ using dirty-paper decoders, where $\tilde{x}_i$ is decoded from $\tilde{y}_{B;i}$.

Charlie:

- At each time instance forms
$$\tilde{\boldsymbol{y}}_C=\tilde{\mathbf{U}}_C^\dagger\boldsymbol{y}_C\,.$$

- Decodes codebooks $i = L_B + 1, \ldots, N_A$ using dirty-paper decoders, where $\tilde{x}_i$ is decoded from $\tilde{y}_{C;i}$.

The following theorem proves that this scheme allows both users to attain their respective secrecy capacities *simultaneously*, providing a proof for Theorem 2.

Theorem 4: Let $\epsilon > 0$, however small, and define $\xi = N_A\epsilon$. Then, there exist scalar codebooks intended for Bob of rates $R_i = \log(b_i^2/c_i^2) - \epsilon$, $i = 1, \ldots, L_B$, and scalar codebooks intended for Charlie of rates $R_i = \log(c_i^2/b_i^2) - \epsilon$, $i = L_B + 1, \ldots, N_A$, such that Scheme 3 simultaneously achieves the secrecy rates $C(\mathbf{H}_B, \mathbf{H}_C, \mathbf{K}) - \xi$ and $C(\mathbf{H}_C, \mathbf{H}_B, \mathbf{K}) - \xi$ for Bob and Charlie, respectively.

Proof outline: The proof of the decodability and secrecy analysis for Charlie are the same as in the proof of Theorem 3 (with Charlie being the "legitimate" user). In the treatment for Bob, a small variation is needed: the interference over sub-channel i ($1 \le i \le L_B$) is composed of both messages intended for Charlie, $\tilde{x}_{L_B+1}^{N_A}$, and messages intended for Bob, $\tilde{x}_{i+1}^{L_B}$. Thus, the DPC for Bob is carried w.r.t. both of these interferences, and the decodability and secrecy analysis follow as in the proof of Theorem 3. ∎

Remark 6 (Replacing DPC with SIC): DPC was used in Scheme 3 for both users. However, in the proposed scheme one may use SIC instead of DPC for Charlie, as is done in [10] for the MIMO wiretap problem. Alternatively, by using lower-triangular matrices instead of upper-triangular ones in (7) (which corresponds to switching roles between Bob and Charlie in the construction of the scheme), one can use SIC for Bob and DPC for Charlie. This phenomenon was also observed by Liu et al. [5]. Unfortunately, this scheme does not allow, in general, to avoid DPC for both of the users.

Remark 7 (Other choices of precoding matrix): In [10], different choices of $\mathbf{V}_A$ were proposed for the MIMO wiretap problem: diagonalizing either $\mathbf{T}_B$ or $\mathbf{T}_C$, which corresponds to avoiding SIC by Bob or guaranteeing strong secrecy, respectively; or, by incorporating space–time coding, to design all the resulting SNRs of each of the users to be constant, which allows using the same codebook over all sub-channels and avoiding bit loading / rate allocation. The analog in the case of confidential broadcast is by applying block diagonal unitary operations, in addition to the matrix $\mathbf{V}_A$ that is dictated by the GSVD, where the blocks correspond to the sub-channels that are allocated to Bob and to Charlie. However, whereas we can avoid SIC at Bob's end in Scheme 3 by diagonalizing his channel, we cannot achieve this result for both Charlie and Bob simultaneously, as DPC needs to be employed for at least one of the users.

Appendix A
Truncation of Generalized Singular Values

The diagonal variant of the GSVD of $\mathbf{G}_B = G(\mathbf{H}_B, \mathbf{K})$ and $\mathbf{G}_C = G(\mathbf{H}_C, \mathbf{K})$ is given by [8], [9]:

$$\mathbf{G}_B = \mathbf{U}_B \mathbf{D}_B \mathbf{X}^\dagger \tag{20a}$$

$$\mathbf{G}_C = \mathbf{U}_C \mathbf{D}_C \mathbf{X}^\dagger, \tag{20b}$$

where $\mathbf{U}_B$ and $\mathbf{U}_C$ are unitary, $\mathbf{X}$ is invertible, and $\mathbf{D}_B$ and $\mathbf{D}_C$ are diagonal matrices with positive diagonal values satisfying

$$\mathbf{D}_B^2 + \mathbf{D}_C^2 = \mathbf{I}. \tag{21}$$

As in Section II, we take, w.l.o.g., the GSV vector $\boldsymbol{\mu}$, which equals the ratio between the diagonals of $\mathbf{D}_B$ and $\mathbf{D}_C$, to be non-increasing. Denote the number of GSVs that are greater than 1 by L_B, and the rest — by $L_C = N_A - L_B$. Denote further by $\mathbf{I}_B$ the $N_A \, times N_A$ diagonal matrix with L_B ones on its diagonal, followed by L_C zeros.

By applying a QL decomposition to $\mathbf{X}$, we attain

$$\mathbf{G}_B = \mathbf{U}_B \mathbf{D}_B \mathbf{T} \mathbf{V}_A^\dagger$$
$$\mathbf{G}_C = \mathbf{U}_C \mathbf{D}_C \mathbf{T} \mathbf{V}_A^\dagger,$$

where $\mathbf{T}$ is upper-triangular and $\mathbf{V}_A$ is unitary.[9]

Since $\mathbf{V}_A$ is unitary, the following relations hold:

$$\mathbf{G}'_B \triangleq \begin{pmatrix} \mathbf{H}_B \mathbf{K}^{1/2} \mathbf{V}_A \\ \mathbf{I} \end{pmatrix} = \mathbf{U}'_B \mathbf{D}_B \mathbf{T}$$
$$\mathbf{G}'_C \triangleq \begin{pmatrix} \mathbf{H}_C \mathbf{K}^{1/2} \mathbf{V}_A \\ \mathbf{I} \end{pmatrix} = \mathbf{U}'_C \mathbf{D}_C \mathbf{T},$$

where $\mathbf{U}'_B$ and $\mathbf{U}'_C$ are unitary. That is, the GSVD of $\mathbf{G}'_B$ and $\mathbf{G}'_C$ is achieved by applying a QR decomposition to each of them.

Finally, by incorporating $\mathbf{I}_B$ we achieve

$$\tilde{\mathbf{G}}_B \triangleq \begin{pmatrix} \mathbf{H}_B \mathbf{K}^{1/2} \mathbf{V}_A \mathbf{I}_B \\ \mathbf{I} \end{pmatrix} = \tilde{\mathbf{U}}_B \tilde{\mathbf{D}}_B \tilde{\mathbf{T}} \tag{22a}$$

$$\tilde{\mathbf{G}}_C \triangleq \begin{pmatrix} \mathbf{H}_C \mathbf{K}^{1/2} \mathbf{V}_A \mathbf{I}_B \\ \mathbf{I} \end{pmatrix} = \tilde{\mathbf{U}}_C \tilde{\mathbf{D}}_C \tilde{\mathbf{T}}, \tag{22b}$$

where $\tilde{\mathbf{U}}_B$ and $\tilde{\mathbf{U}}_C$ are unitary having the same first L_B columns as $\mathbf{U}'_B$ and $\mathbf{U}'_C$, respectively; $\tilde{\mathbf{T}}$, $\tilde{\mathbf{D}}_B$ and $\tilde{\mathbf{D}}_C$ have the same first L_B columns as $\mathbf{T}$, $\mathbf{D}_B$ and $\mathbf{D}_C$, respectively, whereas the remaining L_C columns are all zero except for the diagonal elements, which are equal to 1:

$$\tilde{D}_{B;i,j} = \tilde{D}_{C;i,j} = \tilde{T}_{i,j} = \begin{cases} 1 & i = j,\ j > L_B \\ 0 & i \ne j,\ j > L_B \end{cases}.$$

The latter is easily seen by noting that the QR decomposition carries a Gram–Schmidt process over the columns of the decomposed matrices, and hence the first L_B columns remain the same after applying $\mathbf{I}_B$, whereas the structure of the remaining columns is trivial due to the nullification of the last L_C columns of $\mathbf{H}_B \mathbf{K}^{1/2} \mathbf{V}_A$.

We note that (22) is the GSVD of $\tilde{\mathbf{G}}_B$ and $\tilde{\mathbf{G}}_C$ up to the normalization property (21), which has no effect on the GSVs and can be achieved by multiplying by an $N_A \times N_A$ diagonal matrix with its first L_B diagonal entries equal to 1 and the remaining L_C diagonal entries — to $1/\sqrt{2}$.

The desired result is established by noting that $\mathbf{K}_B^{1/2} = \mathbf{K}^{1/2} \mathbf{V}_A \mathbf{I}_B$, and that the first L_B GSVs of $(\tilde{\mathbf{G}}_B, \tilde{\mathbf{G}}_C)$ are equal

[9] Note that, by denoting $\mathbf{T}_B = \mathbf{D}_B \mathbf{T}$ and $\mathbf{T}_C = \mathbf{D}_C \mathbf{T}$, we attain the triangular variant of the GSVD, which is, in turn, a special case of (7).

to the first L_B GSVs of $(\mathbf{G}_B, \mathbf{G}_C)$ (the GSVs that are greater than 1) and the remaining GSVs of $(\tilde{\mathbf{G}}_B, \tilde{\mathbf{G}}_C)$ are equal to 1.

APPENDIX B
PROOF OF LEMMA 1

Consider the diagonal variant of the GSVD of $\mathbf{G}_B = G(\mathbf{H}_B, \mathbf{K})$ and $\mathbf{G}_C = G(\mathbf{H}_C, \mathbf{K})$ (20) and denote the squared GSV vector by $\boldsymbol{\lambda}$, i.e., the vector whose entries satisfy:

$$\lambda_i \triangleq \mu_i^2 \,.$$

The MI difference in terms of $\{\lambda_i\}$ is equal to

$$I(\mathbf{H}_B, \mathbf{H}_C, \mathbf{K}) = \sum \log \lambda_i \,.$$

Proposition 1: For any matrices $\mathbf{G}_B$ and $\mathbf{G}_C$, consider the generalized eigenvalue (GEV) problem:

$$\mathbf{G}_B^\dagger \mathbf{G}_B \boldsymbol{y} = \lambda \mathbf{G}_C^\dagger \mathbf{G}_C \boldsymbol{y} \,.$$

Then, the generalized eigenvalues of $(\mathbf{G}_B^\dagger \mathbf{G}_B, \mathbf{G}_C^\dagger \mathbf{G}_C)$, $\{\lambda_i\}$, are the GSVs of $(\mathbf{G}_B, \mathbf{G}_C)$, $\{\mu_i\}$, and the generalized eigenvectors are the corresponding columns of

$$\mathbf{Y} = \mathbf{X}^{-\dagger}.$$

Furthermore, the differential of the GEV λ in terms of the differentials of $\mathbf{G}_B^\dagger \mathbf{G}_B$ and of $\mathbf{G}_C^\dagger \mathbf{G}_C$ is given by

$$d\lambda = \frac{\boldsymbol{y}^\dagger \Big(d(\mathbf{G}_B^\dagger \mathbf{G}_B) - \lambda d(\mathbf{G}_C^\dagger \mathbf{G}_C) \Big) \boldsymbol{y}}{\boldsymbol{y}^\dagger \mathbf{G}_C^\dagger \mathbf{G}_C \boldsymbol{y}} \,. \tag{23}$$

Proof: The first part of the proposition easily follows from

$$\mathbf{G}_B^\dagger \mathbf{G}_B \mathbf{Y} = \mathbf{X} \mathbf{D}_B^2$$
$$\mathbf{G}_C^\dagger \mathbf{G}_C \mathbf{Y} = \mathbf{X} \mathbf{D}_C^2 \,.$$

The proof of the differential identity (23) can be derived by standard eigenvalue perturbation analysis; see, e.g., [20]. ■

Lemma 2: The differential of GSV λ_i ($i = 1, \ldots, N_A$), in terms of the differential of the covariance matrix $\mathbf{K}$, is given by

$$c_i^2 d\lambda_i = (\lambda_i - 1) \boldsymbol{y}_i^\dagger \mathbf{B}^{-1} (d\mathbf{K}) \mathbf{B}^{-\dagger} \boldsymbol{y}_i \,,$$

where $\mathbf{B} = \mathbf{K}^{1/2}$, $\boldsymbol{c}$ is the diagonal of $\mathbf{D}_C$, and $\boldsymbol{y}_i$ is the corresponding generalized eigenvector corresponding to λ_i.

Proof: By specializing $\mathbf{G}_B^\dagger \mathbf{G}_B$ and $\mathbf{G}_C^\dagger \mathbf{G}_C$ to the matrices in (20), and differentiating w.r.t. $\mathbf{K}$, we obtain

$$2d(\mathbf{G}_B^\dagger \mathbf{G}_B) = \mathbf{B}^{-1} (d\mathbf{K}) \mathbf{H}_B^\dagger \mathbf{H}_B \mathbf{B} + \mathbf{B}^\dagger \mathbf{H}_B^\dagger \mathbf{H}_B (d\mathbf{K}) \mathbf{B}^{-\dagger}, \tag{24a}$$

$$2d(\mathbf{G}_C^\dagger \mathbf{G}_C) = \mathbf{B}^{-1} (d\mathbf{K}) \mathbf{H}_C^\dagger \mathbf{H}_C \mathbf{B} + \mathbf{B}^\dagger \mathbf{H}_C^\dagger \mathbf{H}_C (d\mathbf{K}) \mathbf{B}^{-\dagger}. \tag{24b}$$

Substituting (24) in (23), gives rise to

$$\begin{aligned} 2c_i^2 d\lambda_i &= \boldsymbol{y}_i^\dagger \Big(\mathbf{B}^{-1} (d\mathbf{K}) (\mathbf{H}_B^\dagger \mathbf{H}_B - \lambda_i \mathbf{H}_C^\dagger \mathbf{H}_C) \mathbf{B} \\ &\quad + \mathbf{B}^\dagger (\mathbf{H}_B^\dagger \mathbf{H}_B - \lambda_i \mathbf{H}_C^\dagger \mathbf{H}_C) (d\mathbf{K}) \mathbf{B}^{-\dagger} \Big) \boldsymbol{y}_i \\ &= \boldsymbol{y}_i^\dagger \Big(\mathbf{B}^{-1} (d\mathbf{K}) \mathbf{B}^{-\dagger} \mathbf{B}^\dagger (\mathbf{H}_B^\dagger \mathbf{H}_B - \lambda_i \mathbf{H}_C^\dagger \mathbf{H}_C) \mathbf{B} \\ &\quad + \mathbf{B}^\dagger (\mathbf{H}_B^\dagger \mathbf{H}_B - \lambda_i \mathbf{H}_C^\dagger \mathbf{H}_C) \mathbf{B} \mathbf{B}^{-1} (d\mathbf{K}) \mathbf{B}^{-\dagger} \Big) \boldsymbol{y}_i \\ &= 2(\lambda_i - 1) \boldsymbol{y}_i^\dagger \mathbf{B}^{-1} (d\mathbf{K}) \mathbf{B}^{-\dagger} \boldsymbol{y}_i \,, \end{aligned}$$

as desired. ■

Corollary 1: If $d\mathbf{K}$ is positive semi-definite, then the sign of $d\lambda_i$ equals the sign of $\lambda_i - 1$.

The result of Lemma 1 follows immediately from this corollary.

REFERENCES

[1] R. Liu and H. V. Poor, "Secrecy capacity region of a multiple-antenna Gaussian broadcast channel with confidential messages," *IEEE Trans. Info. Theory*, vol. 55, pp. 1235–1249, 2009.

[2] S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian wiretap channel," *IEEE Trans. Info. Theory*, vol. 24, pp. 451–456, 1978.

[3] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas—part II: The MIMOME wiretap channel," *IEEE Trans. Info. Theory*, vol. 56, pp. 5515–5532, 2010.

[4] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," *IEEE Trans. Info. Theory*, vol. 57, pp. 4961–4972, 2011.

[5] R. Liu, T. Liu, H. V. Poor, and S. Shamai, "Multiple-input multiple-output Gaussian broadcast channels with confidential messages," *IEEE Trans. Info. Theory*, vol. 56, pp. 4215–4227, 2010.

[6] T. Liu and S. Shamai, "A note on the secrecy capacity of the multiple-antenna wiretap channel," *IEEE Trans. Info. Theory*, vol. 55, pp. 2547–2553, 2009.

[7] R. Bustin, R. Liu, H. V. Poor, and S. Shamai, "An MMSE approach to the secrecy capacity of the MIMO Gaussian wiretap channel," *EURASIP Journal on Wireless Comm. and Networking. Special Issue on Wireless Physical Security*, 2009.

[8] C. F. Van Loan, "Generalizing the singular value decomposition," *SIAM J. Numer.*, vol. 13, pp. 76–83, 1976.

[9] G. H. Golub and C. F. Van Loan, *Matrix Computations, 3rd ed.* Baltimore: Johns Hopkins University Press, 1996.

[10] A. Khina, Y. Kochman, and A. Khisti, "Decomposing the MIMO wiretap channel," in *Proc. IEEE Int. Symp. on Info. Theory (ISIT)*, Honolulu, HI, USA, June/July 2014, pp. 206–210.

[11] A. Khina, Y. Kochman, and U. Erez, "Joint unitary triangularization for MIMO networks," *IEEE Trans. Sig. Proc.*, vol. 60, pp. 326–336, 2012.

[12] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Info. Theory*, vol. 52, pp. 3936–3964, Sept. 2006.

[13] Y. Jiang, W. Hager, and J. Li, "Uniform channel decomposition for MIMO communications," *IEEE Trans. Sig. Proc.*, vol. 53, pp. 4283–4294, 2005.

[14] B. Hassibi, "An efficient square-root algorithm for BLAST," in *Proc. IEEE Int. Conf. on Acoust. Speech and Sig. Proc. (ICASSP)*, vol. 2, Istanbul, Turkey, June 2000, pp. 737–740.

[15] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. URSI Int. Symp. Sig., Sys., Elect. (ISSSE)*, Sep./Oct. 1998, pp. 295–300.

[16] J. M. Cioffi and G. D. Forney Jr., "Generalized decision-feedback equalization for packet transmission with ISI and Gaussian noise," in *Comm., Comp., Cont. and Sig. Proc.*, 1997, pp. 79–127.

[17] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Info. Theory*, vol. 29, pp. 439–441, May 1983.

[18] A. S. Cohen and A. Lapidoth, "The Gaussian watermarking game," *IEEE Trans. Info. Theory*, pp. 1639–1667, 2002.

[19] U. Erez, S. Shamai, and R. Zamir, "Capacity and lattice strategies for cancelling known interference," *IEEE Trans. Info. Theory*, pp. 3820–3833, Nov. 2005.

[20] J. de Leeuw, "Derivatives of generalized eigen systems with applications," Preprint Series 528, Department of Statistics, UCLA, Sep. 2007.