MIMO Gaussian Broadcast Channels with Common, Private and Confidential Messages

Ziv Goldfeld, Member, IEEE, and Haim H. Permuter, Senior Member, IEEE

Abstract—The two-user multiple-input multiple-output (MIMO) Gaussian broadcast channel (BC) with common, private and confidential messages is considered. The transmitter sends a common message to both users, a confidential message to User 1 and a private (non-confidential) message to User 2. The secrecy-capacity region is characterized by showing that certain inner and outer bounds coincide and that the boundary points are achieved by Gaussian inputs, which enables the development of a tight converse. The proof relies on factorization of upper concave envelopes and a variant of dirty-paper coding (DPC). It is shown that the entire region is exhausted by using DPC to cancel out the signal of the non-confidential message at Receiver 1, thus making DPC against the signal of the confidential message unnecessary. A numerical example illustrates the secrecy-capacity results.

Index Terms—Additive Gaussian channel, broadcast channel, dirty-paper coding multiple-input multiple-output (MIMO) communications, physical-layer security, upper concave envelopes.

1. INTRODUCTION

Additive Gaussian channels are a common model for wireless communication, whose open nature makes it vulnerable to a variety of security threats, such as eavesdropping. However, eavesdroppers are not always a malicious entity from which all transmissions are concealed. Rather, a legitimate recipient of one message may serve as an eavesdropper for other messages. We encapsulate this notion in a two-user multiple-input multiple-output (MIMO) Gaussian broadcast channel (BC) with common, private and confidential messages (Fig. 1). The common message $M_0$ is intended to both users, while $M_1$ and $M_2$ are private messages that are sent to User 1 and User 2, respectively. Furthermore, $M_1$ is confidential and is kept secret from User 2. Many real-life scenarios fall within this framework. One such example is a banking site that is intended to both users, while $M_1$ and $M_2$ are private messages to each user is kept secret from the opposite user, without and with a common message, was solved in [13] and [14], respectively. As the capacity region of the MIMO Gaussian BC without secrecy requirements was derived in [15] with no common message present, and in [16] with a common message, this work settles the two remaining scenarios concerning secrecy. More specifically, focusing on the two-user MIMO Gaussian BC with or without a common message and where both, either or neither of the private messages are secret, we derive the secrecy-capacity regions of the only two instances that remained unsolved until now. A pointer to each past result and the contribution of this work are found in Table I.

Up until the more recent work of Geng and Nair [16], all the aforementioned results established the optimality of Gaussian inputs based on channel enhancement arguments, originally...
used in [15] to characterize the private message capacity region of the MIMO Gaussian BC (without secrecy constraints). In a nutshell, the idea of [15] was to approximate the actual BC using enhanced BCs, for which the entropy power inequality applies and is invoked to establish the optimality of Gaussian inputs (similarly to the proof for the scalar case by Bergmans [17]). Continuity arguments are then used to characterize the capacity region of the MIMO Gaussian BC of interest.

In doing so, we use the result without a common message, given in Theorems 1 and 2, respectively. Discussions of these results and a numerical example to illustrate the obtained regions are also given in Section III. Section IV presents the various definitions and properties related to UCEs used throughout this work (with proofs relegated to Section VI). Section V contains the proofs of our main secrecy-capacity results based on the technical background supplied in Section IV. Finally, Section VII summarizes the main results and insights of this work.

II. NOTATION AND PRELIMINARIES

A. Notation

We use the following notations. The set of natural numbers (which does not include 0) is denoted by \( \mathbb{N} \), while \( \mathbb{R} \) are the reals. We further define \( \mathbb{R}_+ \triangleq \{ x \in \mathbb{R} | x \geq 0 \} \). Given two real numbers \( a, b \), we denote by \([a : b]\) the set of integers \( \{ n \in \mathbb{N} | a \leq n \leq b \} \); when \( a = 1 \) we use the shorthand \([b]\). Calligraphic letters denote sets, e.g., \( \mathcal{X} \), while \( |\mathcal{X}| \) stands for the cardinality of \( \mathcal{X} \). \( \mathcal{X}^n \) denotes the \( n \)-fold Cartesian product of \( \mathcal{X} \). An element of \( \mathcal{X}^n \) is denoted by \( x^n = (x_1, x_2, \ldots, x_n) \); whenever the dimension \( n \) is clear from the context, vectors (or sequences) are denoted by boldface letters, e.g., \( \mathbf{x} \). The transpose and the Euclidean norm of \( \mathbf{x} \) are denoted by \( \mathbf{x}^\top \) and \( \| \mathbf{x} \| \), respectively. Random variables are denoted by uppercase letters, e.g., \( \mathcal{X} \), with similar conventions for random vectors. All the random variables considered in this work are real valued.

Probability density functions (PDFs) are denoted by the lowercase letters \( p \) or \( q \), with a subscript that identifies the random variable and its possible conditioning. For example, for two jointly continuous random vectors \( \mathbf{X} \) and \( \mathbf{Y} \), let \( p_{\mathbf{X}} \), \( p_{\mathbf{X},\mathbf{Y}} \) and \( p_{\mathbf{X}|\mathbf{Y}} \) denote, respectively, the PDF of \( \mathbf{X} \), the joint PDF of \( \{ \mathbf{X}, \mathbf{Y} \} \) and the conditional PDF of \( \mathbf{X} \) given \( \mathbf{Y} \). Expressions such as \( p_{\mathbf{X},\mathbf{Y}}(x,y) = p_{\mathbf{X}}(x)p_{\mathbf{Y}|\mathbf{X}}(y|x) \) are to be understood pointwise, i.e., \( p_{\mathbf{X},\mathbf{Y}}(x,y) = p_{\mathbf{X}}(x)p_{\mathbf{Y}|\mathbf{X}}(y|x) \), for all \((x, y) \in \mathcal{X} \times \mathcal{Y} \). Accordingly, when three random variables \( \mathbf{X}, \mathbf{Y} \) and \( \mathbf{Z} \) satisfy \( p_{\mathbf{X},\mathbf{Y},\mathbf{Z}} = p_{\mathbf{X}|\mathbf{Y}}p_{\mathbf{Y}|\mathbf{Z}} \), they form a Markov chain, which we denote by \( \mathbf{X} \rightarrow \mathbf{Y} \rightarrow \mathbf{Z} \). The subscripts of a PDF are omitted if its arguments are lowercase versions of the corresponding random variables. General (i.e., not necessarily continuous)
probability distributions are denoted by the uppercase letters $P$ and $Q$, with conventions similar to those used for PDFs. The expectation of a random variable $X$ is $E[X]$. When a random variable $X$ is normally distributed we write $X \sim N(\mu, \sigma^2)$, where $\mu = E[X]$ is the expectation of $X$ and $\sigma^2 = \text{var}(X)$ is its variance. Similarly, an $n$-dimensional Gaussian distribution of dimension is defined by the expectation $\mu = E[X] \in \mathbb{R}^n$ and the covariance matrix $K = E[(X - \mu)(X - \mu)^\top)$, for which we write $X \sim N(\mu, K)$. Generally, non-italic capital letters, e.g., $A$, denote matrices. We use $A \succeq 0$ to indicate that a matrix $A$ is positive semi-definite, while $A \preceq B$ denotes “less than or equal to” in the positive semi-definite ordering, i.e., $B - A \succeq 0$. The determinant of a square matrix $A$ is denoted by $|A|$.

**Definition 1 (Upper Concave Envelope)** Let $f : \mathcal{D} \to \mathbb{R}$ be a function defined on a convex set $\mathcal{D}$. The UCE $\mathcal{E}f : \mathcal{D} \to \mathbb{R}$ of $f$ is the pointwise smallest concave function such that $(\mathcal{E}f)(x) \geq f(x), \forall x \in \mathcal{D}$.

Another representation of the UCE $\mathcal{E}f$ of $f$ relies on the supporting hyperplanes of $f$. Namely, for any $x \in \mathcal{D}$, we have $(\mathcal{E}f)(x) = \sup_{\mathcal{V} : \mathcal{V} = x} \mathcal{E}f(V)$.

**B. Problem Definition**

The outputs of a MIMO Gaussian BC at the $t$th channel use are:

$$Y_j(i) = G_j X(i) + Z_j(i), \quad j = 1, 2, \quad i \in [n],$$

where $G_1, G_2 \in \mathbb{R}^{t \times t}$ are channel gain matrices (assumed to be known to all parties), $\{Z_j(i)\}_{i \in [n]}$, for $j = 1, 2$, is an independent and identically distributed (i.i.d.) sequence of Gaussian random vectors taking values in $\mathbb{R}^{t \times 1}$. For each $j = 1, 2$ and $i \in [n]$, the elements of $Z_j(i) = [Z_{j,1}(i) \quad Z_{j,2}(i) \ldots Z_{j,t}(i)]^\top$ are also i.i.d. Gaussian random variables, whose expected values and variance are specified by the parameters of the normal distribution of $Z_j(i)$. The input sequence $\{X(i)\}_{i \in [n]}$ is subject to the covariance constraint

$$\frac{1}{n} \sum_{i=1}^{n} E[X(i)X(i)^\top] \preceq K,$$

where $K \succeq 0$.

**Remark 1 (Assumptions)** We make the following assumptions on the channel gain matrices and the noise statistics:

1) $G_1$ and $G_2$ are square and invertible. The analysis in this work relies on showing that certain inner and outer bounds on the secrecy-capacity region coincide. These bounds are characterized in terms of mutual information terms between the channel input (or some auxiliary random variables) and the channel outputs. The mutual information terms, and hence the inner and outer bounds, are continuous functions of the channel gain matrices. For square matrices, recall that the set of invertible matrices is a dense open set in the set of all $t \times t$ matrices. Therefore, by continuity of the bounds, the inner and outer bounds coincide for all channel gain matrices. If $G_1$ and $G_2$ are not square, the singular value decomposition (SVD) allows rewriting the original MIMO BC as a MIMO BC with square gain matrices (of size corresponding to the number of transmitting antennas) using only reversible manipulations. These manipulations have no effect on the secrecy-capacity region of the channel (see, e.g., [15, Section 5]).

2) For each $j = 1, 2$, the Gaussian noise vectors, $\{Z_j(i)\}_{i \in [n]}$, are i.i.d. according to $\mathcal{N}(0, I)$, where $I$ is the $t \times t$-identity matrix. This assumption is without loss of generality due to the following reasons: First, the mean of the Gaussian noise does not affect the capacity region. Second, when the covariance matrix is invertible, the noises can be whitened by multiplying by another invertible matrix. On the other hand, if the covariance matrix is non-invertible, the communication scenario degenerates since a suitable linear transformation converts the Gaussian channel into a noiseless channel with infinite capacity.

We study the MIMO Gaussian BC with common, private and confidential messages (Fig. 1). The sender communicates three messages $(M_0, M_1, M_2)$ over the MIMO Gaussian BC from (1). $M_0$ is a common message that is intended to both users, while $M_j$, for $j = 1, 2$, is delivered to user $j$ only. The receivers are to recover their intended messages with arbitrarily small error probability. Moreover, $M_1$ is a confidential message that is to be kept secret from User 2, 

| $M_0$ | $M_1$ | $M_2$ | Solution |
|-------|-------|-------|----------|
| — | Private | Private | Weingarten-Steinberg-Shamai 2006 [15] |
| Public | Private | Private | Geng-Nair 2014 [16] |
| Public | Secret from User 2 | — | Ly-Liu-Liang 2010 [11] |
| — | Secret from User 2 | Secret from User 1 | Liu-Liu-Poor-Shamai 2010 [13] |
| Public | Secret from User 2 | Secret from User 1 | Etemad-Ulukus 2012 [14] |
| — | Secret from User 2 | Private | This work |
| Public | Secret from User 2 | Private | This work |
which is formally described by the weak-secrecy requirement
\[
\frac{1}{n} I(M_1; Y_2^n) \xrightarrow{n \to \infty} 0,
\]
where \( n \) is the number of channel uses. In \([3]\), the notation
\[ Y_2^n \triangleq \left[ Y_2(1) \ Y_2(2) \ldots \ Y_2(n) \right]^\top \]
is used, where for each \( i \in [n] \), \( Y_2(i) \) is the output vector (taking values in \( \mathbb{R}^i \))
observed by User 2 at the \( i \)th channel instance. For any covariance constraint \( K \succeq 0 \), the secrecy-capacity region \( C_K \)
is the closure of all achievable rate triples \((R_0, R_1, R_2) \in \mathbb{R}^3_+\),
where achievability is defined in a standard manner (see, e.g., \([22]\)).

Remark 2 (Weak versus Strong Secrecy) We set up the problem in term of the weak-secrecy metric \([3]\), merely because the general inner and outer bounds we use \([23]\) were originally proven under this paradigm \([3]\). Nonetheless, the results from \([23]\) are readily upgraded to strong secrecy using the approach of Maurer and Wolf from \([24]\), while accounting for the channel being continuous in a manner similar to \([25]\).

Since the focus of this paper is on the optimality of Gaussian inputs and computable secrecy-capacity expressions, we do not dwell on the employed notion of security.

III. SECRECY-CAPACITY RESULTS

A. MIMO Gaussian BCs with Private and Confidential Messages

The MIMO Gaussian BC with private and confidential messages but without a common message is defined as in Section II-B while setting \( R_0 = 0 \). For any covariance constraint \( K \succeq 0 \), let \( \hat{C}_K \) be the corresponding secrecy-capacity region, and for any \( 0 \leq K^* \leq K \) set the following shorthand notations:
\[
\hat{r}_1(K^*) \triangleq \frac{1}{2} \log \left| \frac{I + G_1K^*G_1^\top}{I + G_2K^*G_2^\top} \right|, \\
\hat{r}_2(K^*) \triangleq \frac{1}{2} \log \left| \frac{I + G_2K^*G_1^\top}{I + G_2K^*G_2^\top} \right|.
\]

Define also
\[
\hat{C}_K(K^*) \triangleq \left\{ (R_1, R_2) \in \mathbb{R}^2_+ \left| \begin{array}{c} R_1 \leq \hat{r}_1(K^*) \\
R_2 \leq \hat{r}_2(K^*) \end{array} \right. \right\}. 
\]

The following theorem characterizes \( \hat{C}_K \).

Theorem 1 (Secrecy-Capacity without Common Message)
The secrecy-capacity region \( \hat{C}_K \) of the MIMO Gaussian BC with private and confidential messages under the covariance constraint \( K \) is
\[
\hat{C}_K = \bigcup_{0 \leq K^* \leq K} \hat{C}_K(K^*). 
\]

The proof of Theorem 1 (given in Section IV-A) shows that certain inner and outer bounds of the secrecy-capacity region coincide, and that Gaussian inputs are optimal. First, we show that the boundary points of the outer bound are an UCE of a function of the input distribution. Based on some properties of UCEs (see Sections IV-A and IV-B) we deduce that a Gaussian input distribution maximizes the considered UCE.

Remark 3 (Interpretation of Optimal Secrecy Rates) The right-hand side (RHS) of \((4a)\) is the secrecy-capacity of the Gaussian MIMO WTC with input covariance \( K^* \), where User 1 serves as the legitimate party and User 2 as the eavesdropper. The RHS of \((4b)\) is the capacity of the MIMO Gaussian point-to-point channel with input covariance \( K - K^* \) and noise covariance \( I + K^* \). Thus, \( M_1 \) being confidential forces User 1 to treat the second user as an eavesdropper. Then, the transmission rate of \( M_2 \) (to User 2) is maximized by consuming the remaining power, while treating the signal of the first user as noise. The optimization over \( K^* \) corresponds to different choices of user prioritization.

Remark 4 (Relation to Dirty-Paper Coding) As evident from the proof of Theorem 1 (see Proposition 8 in Section IV-A), the entire secrecy-capacity region \( \hat{C}_K \) is achieved by using DPC to cancel out the signal of the non-confidential message \( M_2 \) at Receiver 1 only. The other variant, i.e., DPC against the signal of the confidential message \( M_1 \) at Receiver 2, is unnecessary. This is in contrast to the situation without a secrecy requirement on \( M_1 \) (namely, the private message BC), for which the capacity region is exhausted by taking the convex hull of both variants (DPC against \( M_1 \) and DPC against \( M_2 \)).

Remark 5 (Relation to Common Message Case) Theorem 1 is a special case of the secrecy-capacity region of the MIMO Gaussian BC with common, private and confidential messages \( C_K \) (given in Theorem 2). Nonetheless, we separately state and prove Theorem 1 since it is used as an auxiliary result for the proof of Theorem 2 as it implies the optimality of Gaussian inputs for a certain portion of \( C_K \). More specifically, when the private message rate \( R_0 \) is larger than the common message rate \( R_2 \), the optimizing distribution of \( C_K \) coincides with that of \( \hat{C}_K \); Theorem 1 shows that this distribution is Gaussian.

As a corollary of Theorem 1, we characterize the secrecy-capacity region under the average total power constraint. This is a simple consequence of \([15\) Lemma 1].

Corollary 1 (Average Total Power Constraint) The secrecy-capacity region of the MIMO Gaussian BC with private and confidential messages under the average total power constraint
\[
\frac{1}{n} \sum_{i=1}^{n} \|X(i)\|^2 \leq P,
\]

\([7a]\)
is given by
\[ \hat{C}_P = \bigcup_{0 \leq K: \text{tr}(K) \leq P} \hat{C}_K. \]  
(7b)

Remark 6 (Computing Secrecy-Capacity Region) In general, it is hard to compute (6) and (7b) as they involve non-convex matrix optimization problems. Nonetheless, in Section III-C we show how to convert (5) and (7b) into a computationally efficient form based on matrix decomposition properties from \[21\]. The simplified optimization problem is then used to illustrate the secrecy-capacity region under an average total power constraint \( \hat{C}_P \) on a numerical example.

B. MIMO Gaussian BCs with Common, Private and Confidential Messages

To state the secrecy-capacity region \( C_K \) of the MIMO Gaussian BC with common, private and confidential messages as defined in Section II-B, we define
\[ r_0^{(j)}(K_1, K_2) \triangleq \frac{1}{2} \log \left| \frac{I + G_j K G_j^T}{I + G_j (K_1 + K_2) G_j} \right|, \quad j = 1, 2 \]  
(8a)
\[ r_1(K_2) \triangleq \frac{1}{2} \log \left| \frac{I + G_1 K_2 G_2^T}{I + G_2 K_2 G_2} \right| \]  
(8b)
\[ r_2(K_1, K_2) \triangleq \frac{1}{2} \log \left| \frac{I + G_2 (K_1 + K_2) G_2^T}{I + G_2 K_2 G_2} \right|, \]  
(8c)
and \( r_0(K_1, K_2) = \min \{ r_0^{(1)}(K_1, K_2), r_0^{(2)}(K_1, K_2) \} \), and set
\[ C_K(K_1, K_2) \triangleq \left\{ (R_0, R_1, R_2) \in \mathbb{R}_+^3 \mid \begin{array}{c} R_0 \leq r_0(K_1, K_2) \\ R_1 \leq r_1(K_2) \\ R_2 \leq r_2(K_1, K_2) \end{array} \right\}. \]  
(9)

Theorem 2 (Secrecy-Capacity with Common Message)
The secrecy-capacity region \( C_K \) of the MIMO Gaussian BC with common, private and confidential messages under the covariance constraint (2) is
\[ C_K = \bigcup_{0 \leq K_1, K_2: K_1 + K_2 \leq K} C_K(K_1, K_2). \]  
(10)
Theorem 2 is proven in Section V-B.

Remark 7 (Interpretation of Optimal Secrecy Rates) Our interpretation of the structure of \( C_K \) is reminiscent of Remark 3. First, (8b) indicates that User 1 achieves rates up to the secrecy-capacity of the MIMO Gaussian WTC with input covariance \( K_2 \). The 2nd user treats this signal as an additive Gaussian noise when decoding its private message \( M_2 \), which is transmitted using another (independent) Gaussian signal with covariance \( K_1 \) (see (8a)). According to (8a), the remaining portion of the total covariance matrix, that is, \( K - (K_1 + K_2) \), is employed to encode the common message \( M_0 \), which is decoded by each receiver while treating all other signals as noise. As in the case without a common message, a layered coding scheme, when optimized over the choices of \( K_1 \) and \( K_2 \), exhausts the entire secrecy-capacity region.

As before, Theorem 2 produces a characterization of the secrecy-capacity region under the average total power constraint.

Corollary 2 (Average Total Power Constraint) The secrecy-capacity region of the MIMO Gaussian BC with common, private and confidential messages under the average total power constraint (7b) is given by
\[ C_P = \bigcup_{0 \leq K: \text{tr}(K) \leq P} C_K. \]  
(11)

C. Numerical Example

We illustrate the secrecy-capacity region \( \hat{C}_P \) of the MIMO Gaussian BC with private and confidential messages (without a common message) under an average total power constraint \( P \) (Corollary 1). The region is described in (7b) as the union of all secrecy-capacity regions \( C_K \), each under a covariance constraint \( K \geq 0 \) with \( \text{tr}(K) \leq P \). However, \( \hat{C}_K \) itself is described as matrix optimization problems that is not convex in general, and is therefore, hard to compute.

We overcome the computational inefficiency of \( \hat{C}_K \) by leveraging the decomposition proposed in \[21\] Equation (10): Every positive semi-definite matrix \( K^* \in \mathbb{R}^{t \times t} \) with \( K^* \leq K \) can be expressed as
\[ K^* = K^* D V D^T K^{*\top}, \]  
(12)
where \( V \in \mathbb{R}^{t \times t} \) is a unitary matrix and \( D \in \mathbb{R}^{t \times t} \) is a diagonal matrix whose diagonal values are between 0 and 1. Since in the subsequent example the dimension is \( t = 2 \), a unitary matrix \( V \) is nothing but a rotation matrix, i.e., we set
\[ V = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}, \quad \theta \in [0, 2\pi]. \]  
(13)
Running over all possible diagonal matrices \( D \) involves only two parameters, viz. the diagonal entries of \( D \). Finally, note...
that $K^2$ is any matrix $B$ satisfying $BB^T = K$. Obviously, there are many such matrices (in fact if $B$ satisfies $BB^T = K$, then so does $BU$, for any unitary $U$). However, since the numerical calculation runs over all matrices $V$ from (13) anyway, any choice of $B$ would do. Our simulation uses the Cholesky decomposition of $K$ to calculate $B$.

The region $\hat{\mathcal{C}}_P$ is computed according to (7b), while noting that one may restrict the optimization domain to positive semidefinite matrices $K$ with $\text{tr}(K) = P$. This observation follows because for every $K'$ with $\text{tr}(K') = \pi < P$, there is a $K$ with $\text{tr}(K) = P$, such that

$$
\hat{\mathcal{C}}_{K'} \subseteq \hat{\mathcal{C}}_K. \tag{14}
$$

The matrix $K$ is constructed by increasing the $(1, 1)$-th entry of $K'$ by $P - \pi$, while all other entries of $K'$ remain unchanged. The construction satisfies $K' \preceq K$ and the inclusion in (14) follows because fixing $K^* \preceq K' \preceq K$ and replacing $K'$ with $K$ in (4) does not alter (14) and strictly increases (4b).

In the numerical example we set

$$
G_1 = \begin{bmatrix}
0.3 & 2.5 \\
2.2 & 1.8
\end{bmatrix}, \quad G_2 = \begin{bmatrix}
1.3 & 1.2 \\
1.5 & 3.9
\end{bmatrix} \tag{15}
$$

and $P = 12$. The secrecy-capacity region $\hat{\mathcal{C}}_P$ is given by the solid blue curve in Fig. 2. For comparison, the secrecy-capacity region of the MIMO Gaussian BC with confidential messages [13] (i.e., when each user serves as the eavesdropped channel) is depicted by the dashed red curve. As expected, Fig. 2 shows that imposing a secrecy constraint on $M_2$ at the 1st receiver strictly shrinks the secrecy-capacity region. Although in both regions the maximal value of $R_1$ is the secrecy-capacity of the corresponding MIMO Gaussian WTC (see (4a) and [13] Equation (4)), the achievable values of $R_2$ drop if $M_2$ is also confidential.

**IV. Optimality of Gaussian Inputs via Factorization of Concave Envelopes**

This section provides the mathematical background for characterizing the secrecy-capacity regions of the considered MIMO Gaussian BC without and with a common message (Theorems 1 and 2). In the sequel we define some generic functions and show that they are maximized by Gaussian distributions. These functions are later used to describe the boundary points of certain outer bounds on the secrecy-capacity regions of interest. The properties established in this section are leveraged to show that optimality is achieved by Gaussian inputs, and that the resulting expressions are attainable by a corresponding inner bound.

Sections IV-A and IV-B focus on functions that are reminiscent of those studied in [16] Sections II-B and II-C. Therefore, to avoid verbatim repetition of arguments from [16], we state some of the properties in Sections IV-A and IV-B without proofs. The focus of Section IV-C is on a new function that was not considered in [16], the properties of which we prove in full detail. All the proofs for this section are relegated to Section VI.

Establishing Gaussian inputs as maximizers relies on the notion of two-letter BCs [16], Section I-A, which is a special case of a product BC (PBC).

**Definition 2 (Product BC)** A PBC consists of a sender $(X_1, X_2)$ and two receivers $(Y_{11}, Y_{12})$ and $(Y_{21}, Y_{22})$, and is described by a conditional PDF of the form $q_{Y_{11}, Y_{21}|X_1} = q_{Y_{12}, Y_{22}|X_2}$.

A MIMO Gaussian PBC can be represented as

$$
\begin{bmatrix}
Y_{11} \\
Y_{12} \\
Y_{21} \\
Y_{22}
\end{bmatrix} = \begin{bmatrix}
G_{11} & 0 \\
0 & G_{12} \\
G_{21} & 0 \\
0 & G_{22}
\end{bmatrix} \begin{bmatrix}
X_1 + Z_{11} \\
X_2 + Z_{12} \\
X_1 + Z_{21} \\
X_2 + Z_{22}
\end{bmatrix}, \tag{16a}
$$

where $Z_{11}, Z_{12}, Z_{21}, Z_{22} \sim \mathcal{N}(0, 1)$ are i.i.d. and independent of $(X_1, X_2)$. A two-letter version of a BC is a PBC in which the components are identical, i.e., $q_{Y_{11}, Y_{21}|X_1} = q_{Y_{12}, Y_{22}|X_2}$.

In all subsequent definitions and results, the input covariance constraining matrix $K \succeq 0$ (see (3)) stays fixed.

**A. Difference of Mutual Information Terms**

Consider a BC $q_{Y_1, Y_2|X}$. For any $\eta > 1$, let $s_\eta^q(X)$ be a functional of $X \sim p_X$ defined by

$$
s_\eta^q(X) \triangleq I(X; Y_1) - \eta I(X; Y_2). \tag{17}
$$

**Remark 8** The definition of $s_\eta^q(X)$ in (17) coincides with that of $s_\eta^q(X) \triangleq I(X; Y_1) - \lambda I(X; Y_2)$ from [16] Section II-B (only differing in the ordering of the mutual information terms and the labeling of the parameter). Accordingly, we restate and use some of the properties of $s_\eta^q(X)$ established in [16] without providing proofs. Additional attributes of $s_\eta^q(X)$ that were not proven in [16] are rigorously derived.

For a pair of random variables $(V, X)$ such that $V - X - (Y_1, Y_2)$ forms a Markov chain, set

$$
s_\eta^q(X|V) \triangleq I(X; Y_2|V) - \eta I(X; Y_1|V), \tag{18}
$$

and define the UCE of $s_\eta^q(X)$ as

$$
S_\eta^q(X) \triangleq \max_{V \leftarrow X - (Y_1, Y_2)} \sup_{I_{V|X} : V - X - (Y_1, Y_2)} s_\eta^q(X|V). \tag{19}
$$

The second equality in (13) follows directly from Definition 1. For any discrete random variable $V$ we also set $S_\eta^q(X|V) \triangleq \sum_{v} p(v) s_\eta^q(X|V = v)$, and naturally extend this definition (merely an expectation) for an arbitrary $V$.

**Proposition 1 (Concave Envelopes Properties)** The UCE $S_\eta^q(X)$ satisfies:

1. If $V - X - (Y_1, Y_2)$ forms a Markov chain, then $S_\eta^q(X|V) \leq S_\eta^q(X)$.
2. If $V - Y - X - (Y_1, Y_2)$ forms a Markov chain, then $S_\eta^q(X|V, W) = S_\eta^q(X|V)$.

\(^2\)Henceforth, we omit the time index $i$.\]
3) $S^q_{\eta}(X)$ is convex in $\eta$ inside $(0, 2)$, for a fixed $P_X$, and therefore it is continuous in $\eta$ at $\eta = 1$.\footnote{The 3rd property can be established by considering any bounded, open interval containing 1, and not necessarily (0, 2).}

The proof of Proposition~\ref{prop:convexity} is given in Section~\ref{sec:appendix}.

**Definition 3 (Maximized Concave Envelope) For any MIMO Gaussian BC $q_{Y_1, Y_2|X}$, define**

$$V^q_{\eta}(K) \triangleq \sup_{X: E[XX^T] \leq K} S^q_{\eta}(X) = \sup_{(V, X): E[XX^T] \leq K, V - X - (Y_1, Y_2)} S^q_{\eta}(X|V).$$

(20)

We subsequently show that $\hat{T}^q_{\lambda, \eta}(X|V)$ is achieved by a Gaussian input distribution. Central to the proof is a certain factorization property of $S^q_{\eta}$. To formulate this property, we first extend $S^q_{\eta}$ to PBCs. For a PBC $q_{Y_{11}, Y_{22}|x_1} \times q_{Y_{12}, Y_{22}|x_2}$ we set,

$$s^q_{\eta q_1 q_2}(X_1, X_2) \triangleq I(X_1, X_2; Y_{22}) - \eta I(X_1, X_2; Y_{11}, Y_{12}),$$

(21)

and define the quantities $S^q_{\eta q_1 q_2}(X_1, X_2)$, $S^q_{\eta q_1 q_2}(X_1, X_2)$ and $S^q_{\eta q_1 q_2}(X_1, X_2)$ analogously to the definitions of $S^q_{\eta q_1 q_2}(X|V)$, $S^q_{\eta q_1 q_2}(X)$ and $S^q_{\eta q_1 q_2}(X)$ from the above, respectively.

**Proposition 2 (Factorization Property) For any PBC $q_{Y_{11}, Y_{22}|x_1} \times q_{Y_{12}, Y_{22}|x_2}$, the following chain of inequalities holds**

$$S^q_{\eta q_1 q_2}(X_1, X_2) \leq S^q_{\eta q_1 q_2}(X_1|Y_{22}) + S^q_{\eta q_2}(X_2|Y_{11}) \leq S^q_{\eta q_1 q_2}(X_1) + S^q_{\eta q_2}(X_2).$$

(22)

The proof of Proposition~\ref{prop:factorization} follows by repeating the steps in the proof of [16] Proposition 6, while switching the roles of $Y_1$ and $Y_2$.

**Theorem 3 (Gaussian Maximizer) Let $X \sim N(0, K)$. There exists a decomposition $X = X^* + X'$, such that $X^*$ and $X'$ are independent, $X^* \sim N(0, K^*)$, $X' \sim N(0, K - K^*)$, where $K^* \preceq K$, and $S^q_{\eta q_1 q_2}(X) = S^q_{\eta q_1 q_2}(X^*) = V^q_{\eta}(K)$. Furthermore, this decomposition (i.e., the covariance matrix $K^*$) is unique.**

The proof of Theorem~\ref{thm:gaussian_maximizer} is also omitted as it mimics the proofs of Theorem 1 and Corollary 1 in [16].

**B. Nested Upper Concave Envelopes**

The function considered in this subsection is used to derive the secrecy-capacity region of the considered MIMO Gaussian BC without a common message (see Section~\ref{sec:concavity}).

For a BC $q_{Y_{1j}, Y_{2j}|X}, \eta > 1$, $\lambda = (\lambda_1, \lambda_2), \lambda_j > 0$, $j = 1, 2$, and any $X \sim P_X$ define

$$T^q_{\lambda, \eta}(X) \triangleq \lambda_1 I(X; Y_1) - (\lambda_1 + \lambda_2) I(X; Y_2) + \lambda_1 S^q_{\eta}(X),$$

(23)

where $S^q_{\eta}(X)$ is given by \[19\]. As before, for a pair of random variables $(V, X)$ for which $V - X - (Y_1, Y_2)$ forms a Markov chain, let

$$t^q_{\lambda, \eta}(X|V) \triangleq \lambda_1 I(X; Y_1|V) - (\lambda_1 + \lambda_2) I(X; Y_2|V) + \lambda_1 S^q_{\eta}(X|V).$$

(24)

Define $T^q_{\lambda, \eta}(X|V) \triangleq \sum_{v} P(v) T^q_{\lambda, \eta}(X|V = v)$, for a $V$ with a countable alphabet and consider its natural extension when $V$ is an arbitrary random variable.

**Remark 9 (Nested Concave Envelopes Properties)**

Similarly to the properties of $S^q_{\eta}$ stated in Proposition~\ref{prop:convexity} since $T^q_{\lambda, \eta}$ is concave in $P_X$, Jensen’s inequality implies that $T^q_{\lambda, \eta}(X|V) \leq T^q_{\lambda, \eta}(X)$, for any $(V, X)$ satisfying $V - X - (Y_1, Y_2)$. Moreover, if $W - V - X$ forms a Markov chain, then $T^q_{\lambda, \eta}(X|W, V) = T^q_{\lambda, \eta}(X|V)$, because $P_{X|W, V} = P_{X|V}$. Finally, $T^q_{\lambda, \eta}(X)$ is convex in $\eta$ inside $(0, 2)$, for a fixed $P_X$, and therefore it is continuous as a function of $\eta$ at $\eta = 1$.

**Definition 4 (Maximized Nested Concave Envelope) For any MIMO Gaussian BC $q_{Y_1, Y_2|X}$, define**

$$V^q_{\lambda, \eta}(K) \triangleq \sup_{X: E[XX^T] \leq K} T^q_{\lambda, \eta}(X) = \sup_{(V, X): E[XX^T] \leq K, V - X - (Y_1, Y_2)} t^q_{\lambda, \eta}(X).$$

(26)

**Proposition 3 (Continuity of Maximal Value) For any $\lambda$ as before, $V^\lambda_{\lambda, \eta}(K)$ is continuous in $\eta$ at $\eta = 1$.**

The proof of Proposition~\ref{prop:continuity} follows by arguments similar to those in the proof of Property 3 of Proposition~\ref{prop:convexity}.

As before, to state the factorization property for nested UCEs, we extend some of the preceding definitions to PBCs. For a PBC $q_{Y_{11}, Y_{22}|x_1} \times q_{Y_{12}, Y_{22}|x_2}$, we set

$$t^q_{\lambda, \eta}(X_1, X_2) \triangleq \lambda_1 I(X_1, X_2; Y_{11}, Y_{12}) - (\lambda_1 + \lambda_2) I(X_1, X_2; Y_{21}, Y_{22}) + \lambda_1 S^q_{\eta q_1 q_2}(X_1, X_2).$$

(27)

Furthermore, define $t^q_{\lambda, \eta}(X_1, X_2|V)$, $T^q_{\lambda, \eta}(X_1, X_2|V)$ and $T^q_{\lambda, \eta}(X_1, X_2)$ in a similar manner to $t^q_{\lambda, \eta}(X|V)$, $T^q_{\lambda, \eta}(X)$ and $T^q_{\lambda, \eta}(X)$, respectively. The following proposition states the $T^q_{\lambda, \eta}(X_1, X_2|V)$ factorization property of interest, which is used to prove the existence of a Gaussian maximizer for $V^\lambda_{\lambda, \eta}(K)$ from (26).

**Proposition 4 (Factorization Property) For any PBC $q_{Y_{1j}, Y_{2j}|x_1} \times q_{Y_{12}, Y_{22}|x_2}$, the following chain of inequalities holds**

$$T^q_{\lambda, \eta}(X_1, X_2) \leq T^q_{\lambda, \eta}(X_1|Y_{22}) + T^q_{\lambda, \eta}(X_2|Y_{11}) \leq T^q_{\lambda, \eta}(X_1) + T^q_{\lambda, \eta}(X_2).$$

(28)

Furthermore, if the PBC is Gaussian and a triple $(V^*, X_1^*, X_2^*)$ satisfies

$$t^q_{\lambda, \eta}(X_1^*, X_2^*|V^*) = T^q_{\lambda, \eta}(X_1^*) + T^q_{\lambda, \eta}(X_2^*)$$

(29)
then $X_1^* - V^* - X_2^*$ and $t_{\lambda,\eta}^{(2)}(X_j^*|V^*) = T_{\lambda,\eta}^{q_j}(X_j^*)$, for $j = 1, 2$.

See Section VI-B for the proof of Proposition 4. The existence of a Gaussian maximizer for $V_{\lambda,\eta}(K)$ follows by repeating the proofs of Theorem 2 and Corollary 2 in [10] with respect to our definition of $T_{\lambda,\eta}^q$. The existence is stated in the following Theorem, which we give without proof.

**Theorem 4 (Gaussian Maximizer)** Let $X \sim N(0, K)$. There exists a unique decomposition $X = X_1^* + X_2^* + X'$ into independent random variables $(X_1^*, X_2^*, X')$, where $X_j^* \sim N(0, K_j)$, $j = 1, 2$, and $X' \sim N(0, K - (K_1 + K_2))$. There exists a unique decomposition $f_{\lambda,\eta}(X) = f_{\lambda,\eta}(X_1^*) + T_{\lambda,\eta}^{q_1}(X_1^*, X_2^*)$, and $f_{\lambda,\eta}(X') = V_{\lambda,\eta}(K_1 + K_2)$.

**C. Double-Nested Upper Convex Envelopes**

The definitions and properties in this section are used to derive the secrecy-capacity region of the cooperative BC with common, private and confidential messages (as defined in Section I-C). Let $q_y$, $y_1, y_2|X, \eta > 1$ be a BC, $\lambda_0 = (\lambda_0, \lambda_1, \lambda_2)$, where $\lambda_j > 0$ for $j = 0, 1, 2$, and $\lambda_0 > \lambda_2$, $\alpha \in [0, 1]$ and $\alpha = 1 - \alpha$. For any $X \sim P_X$ define

$$f^{(1)}_{\lambda_0,\alpha,\eta}(X) = (\lambda_2 - \lambda_0)I(X; Y_2) - \alpha \lambda_0 I(X; Y_1) + T_{\lambda,\eta}^{q_{(1)}}(X),$$

and define $f^{(2)}_{\lambda_0,\alpha,\eta}(X_1, X_2|V^*) = F^{\lambda_0,\alpha,\eta}(X_1, X_2|V)$ and $F^{\lambda_0,\alpha,\eta}(X_1, X_2|V)$ as the natural extensions to the PBC scenario of $f^{\lambda_0,\alpha,\eta}(X_1, X_2|V)$.

Moving forward, the factorization property of $F^{\lambda_0,\alpha,\eta}$ is stated in Proposition 5, while Proposition 6 establishes the existence of its maximizer.

**Proposition 5 (Factorization Property)** For any PBC $q_{Y_1, Y_2|X_1} \times q_{Y_1, Y_2|X_2}$, the following chain of inequalities holds

$$F^{\lambda_0,\alpha,\eta}(X_1, X_2) \leq F^{\lambda_0,\alpha,\eta}(X_1|Y_{22}) + F^{\lambda_0,\alpha,\eta}(X_2|Y_{11})$$

and

$$F^{\lambda_0,\alpha,\eta}(X_1, X_2|V^*) = F^{\lambda_0,\alpha,\eta}(X_1^*) + F^{\lambda_0,\alpha,\eta}(X_2^*),$$

for $j = 1, 2$.

See Section VI-C for the proof of Proposition 5.

**Proposition 6 (Existence of a Maximizer)** There exists a pair $(V^*, X^*)$ with $|V^*| \leq \frac{\mu(K)}{2} + 1$ and $\mathbb{E}[XX^*] \leq K$, such that

$$\tilde{V}_{\lambda_0,\alpha,\eta}(K) = f^{\tilde{V}_{\lambda_0,\alpha,\eta}(K)}_{\lambda_0,\alpha,\eta}(X^*|V^*).$$

Furthermore, one may assume that $\mathbb{E}[XX^*|V^* = v^*] = 0$, for every $v^* \in V^*$.

The existence of a maximizer and the cardinality bound on $V^*$ are proven in Section VI-D. A zero conditional expectation is assumed because centering conditioned on each $V^* = v^*$ does not change the mutual information terms and hence $f^{\tilde{V}_{\lambda_0,\alpha,\eta}(X_2^*|V^*)}$ remains unchanged as well. In addition, the centered versions of the input continue to satisfy the covariance constraint.

To show that the distribution that achieves $\tilde{V}_{\lambda_0,\alpha,\eta}(K)$ is Gaussian, we use the invariance of $f^{(1)}_{\lambda_0,\alpha,\eta}(K)$ to rotation. This invariance property is stated in the context of the next Proposition, which is proven in Section VI-E.

**Proposition 7 (Invariance to Rotation)** Let $(V, X) \sim P^*_{V,X}$ attain $\tilde{V}_{\lambda_0,\alpha,\eta}(K)$, with $|V| = m \leq \frac{\mu(K)}{2} + 1$, and let $X_0$ be a centered random variable (zero mean) distributed according to the conditional PMF $P_{X|V = v^*}$. Let $(V_1, X_1, V_2, X_2) \sim P^*_{V,X} \times P^*_{V,X}$ be two i.i.d. copies of $(V^*, X^*)$. Define

$$\tilde{V} = (V_1, V_2),$$

$$X_{v_1}, \{V = (v_1, v_2)\} \sim \frac{1}{\sqrt{2}}(X_{v_1} + X_{v_2}),$$

$$X_{v_2}, \{V = (v_1, v_2)\} \sim \frac{1}{\sqrt{2}}(X_{v_1} - X_{v_2}),$$

where $X_{v_1}$ and $X_{v_2}$ are taken to be independent random variables, i.e., $(X_{v_1}, X_{v_2}) \sim P^*_{V|V = v^*} \times P^*_{V|V = v^*}$. Then $X_{v_1} - V - X_{v_2} \sim V_{\lambda_0,\alpha,\eta}(K) = f^{\lambda_0,\alpha,\eta}(X|V)$, for $j = 1, 2$.
The existence of a Gaussian Maximizer for $\hat{V}^q_{\lambda_o, o, \eta}(K)$ is stated next.

**Theorem 5 (Existence of Gaussian Maximizer)** There exists an $X^* \sim \mathcal{N}(0, K^*)$, where $K^* \preceq K$, such that $\hat{V}^q_{\lambda_o, o, \eta}(K) = f^q_{\lambda_o, o, \eta}(X^*)$. Furthermore, the zero mean maximizer is unique.

See Section VI-F for the proof.

**Corollary 3 (Gaussian Maximizer Properties)** Let $X \sim \mathcal{N}(0, K)$. There is a unique decomposition $X = X_1^* + X_2^* + X_3^*$ into independent random variables $(X_1^*, X_2^*, X_3^*, X')$, where $X_j^* \sim \mathcal{N}(0, K_j)$, for $j = 1, 2, 3$, and $X' \sim \mathcal{N}(0, K - (K_1 + K_2 + K_3))$, with $K_1 + K_2 + K_3 \preceq K$, such that

$$f^q_{\lambda_o, o, \eta}(X) = f^q_{\lambda_o, o, \eta}(X_1^* + X_2^* + X_3^*) = \hat{V}^q_{\lambda_o, o, \eta}(K) \quad (37a)$$

$$T^q_{\lambda_o, \eta}(X_1^* + X_2^* + X_3^*) = \lambda_1^q(X_1^* + X_2^*) = \hat{V}^q_{\lambda_1, \eta}(K_1 + K_2 + K_3) \quad (37b)$$

$$S^q_{\eta}(X_1^* + X_2^*) = S^q_{\eta}(X_1^*) = V^q_{\eta}(K_1 + K_2). \quad (37c)$$

Corollary 3, which is a consequence of Theorem 5, is our main tool for characterizing the secrecy-capacity region of the MIMO Gaussian BC with common, private and confidential messages. The proof of the Corollary is provided in Section VI-C.

V. PROOFS OF SECRECY-CAPACITY RESULTS

A. Proof of Theorem 1

We establish the secrecy-capacity region of the MIMO Gaussian BC with private and confidential messages by showing that certain outer bound and inner bounds match. In particular, we consider special cases of the inner and outer bounds from Theorems 1 and 3 of [23], respectively. To state the bounds, let $\hat{C}$ denote the secrecy-capacity region of the corresponding discrete-memoryless (DM) BC.

**Bound 1 (Outer Bound)** Let $\hat{C}$ be the closure of the union of rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ satisfying:

$$R_1 \leq I(U; Y_1 | V) - I(U; Y_2 | V) \quad (38a)$$

$$R_2 \leq I(V; Y_2) \quad (38b)$$

over all $(V, U) - X = (Y_1, Y_2)$. Then $\hat{C} \subseteq \hat{C}$.

**Bound 2 (Inner Bound)** Let $\tilde{C}$ be the closure of the union of rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ satisfying:

$$R_1 \leq I(U; Y_1) - I(U; V) - I(U; Y_2 | V) \quad (39a)$$

$$R_2 \leq I(V; Y_2) \quad (39b)$$

over all $(V, U) - X = (Y_1, Y_2)$. Then $\tilde{C} \subseteq \hat{C}$.

The reader is referred to Appendix A for the proofs of Bounds 1 and 2. Let $\hat{C}_K$, $\tilde{C}_K$ and $\tilde{C}_K$ denote the secrecy-capacity region, the outer bound and the inner bound for a MIMO Gaussian BC computed under a covariance input constraint $\mathbb{E}[XX^T] \preceq K$. Accordingly, we have $\tilde{C}_K \subseteq \hat{C}_K \subseteq \hat{C}_K$.

The opposite inclusion, i.e., $\hat{C}_K \subseteq \tilde{C}_K$, is shown next. The regions $\tilde{C}_K$ and $\hat{C}_K$ are closed, convex and bounded subsets of the first quadrant, and therefore, are characterized by the intersection of their supporting hyperplanes.

**Lemma 1 (Supporting Hyperplanes)** The following are supporting hyperplanes of $\hat{C}_K$ and $\tilde{C}_K$:

$$R_1 \geq 0, \quad R_1 \leq H^{K}_1, \quad R_2 \geq 0, \quad R_2 \leq H^{K}_2, \quad (40)$$

where

$$H^{K}_1 \triangleq \max_{(V, U) - X = (Y_1, Y_2)} I(U; Y_1 | V) - I(U; Y_2 | V) \quad (41a)$$

$$H^{K}_2 \triangleq \max_{X \in [XX^T] \preceq K} I(X; Y_2). \quad (41b)$$

Furthermore, $(H^{K}_1, 0)$ and $(0, H^{K}_2)$ are boundary points of $\hat{C}_K$ and $\tilde{C}_K$.

Lemma 1 is proven in Appendix B. It shows that if we set $R^*_1 \triangleq \max \{ R_1 | (R_1, 0) \in \hat{C}_K \}$ and $R^*_2 \triangleq \max \{ R_2 | (0, R_2) \in \hat{C}_K \}$, then $(R^*_1, 0)$ and $(0, R^*_2)$ are attainable in $\tilde{C}_K$. Consequently, to show that the regions coincide, it suffices to establish

$$\max \left\{ \lambda_1 R_1 + \lambda_2 R_2 : (R_1, R_2) \in \hat{C}_K \right\} \leq \max \left\{ \lambda_1 R_1 + \lambda_2 R_2 : (R_1, R_2) \in \tilde{C}_K \right\} \leq \lambda_1 R_1 + \lambda_2 R_2, \quad (42)$$

for $\lambda_1, \lambda_2 > 0$. Observe that

$$\max_{(R_1, R_2) \in \hat{C}_K} \lambda_1 R_1 + \lambda_2 R_2 \leq \sup_{(V, U) - X = (Y_1, Y_2)} \lambda_1 I(U; Y_1 | V) - I(U; Y_2 | V) + \lambda_2 I(V; Y_2) \quad (a)$$

$$\leq \sup_{(V, U) - X = (Y_1, Y_2)} \lambda_1 I(X; Y_1 | V) - (\lambda_1 + \lambda_2) I(X; Y_2 | V) + \lambda_2 I(X; Y_2) \quad (b)$$

$$\leq \sup_{X \in [XX^T] \preceq K} \lambda_1 I(X; Y_1 | V) - (\lambda_1 + \lambda_2) I(X; Y_2 | V) \quad (c)$$

$$\leq \sup_{X \in [XX^T] \preceq K} \lambda_2 I(X; Y_2) + \sup_{Y \in \mathbb{Y}} \max_{X \in [XX^T] \preceq K} \lambda_1 S^q_{\eta}(X | V) + \lambda_2 I(X; Y_2) \quad (d)$$

$$\leq \sup_{X \in [XX^T] \preceq K} \lambda_2 I(X; Y_2) + \sup_{Y \in \mathbb{Y}} \lim_{\eta \downarrow 1} \lambda_1 S^q_{\eta}(X | V) + \lambda_2 I(X; Y_2) \quad (e)$$

where:

(a) uses (38);

(b) is because $(V, U) - X = (Y_1, Y_2)$ forms a Markov chain;

(c) follows by the definition of $S^q_{\eta}(X | V)$ and since conditioned on $V$, $U - X = (Y_1, Y_2)$ forms a Markov chain, i.e., it holds that $P_{Y_1, Y_2 | V, U, X} = P_{Y_1, Y_2 | X}$. Furthermore, (c) uses the continuity of $S^q_{\eta}(X | V)$ in $\eta$ at $\eta = 1$ (see Property 3 of Proposition 1), which implies that for any $(V, X)$

$$S^q_{\eta}(X | V) \triangleq S^q_{\lim_{\eta \downarrow 1}}(X | V) = \lim_{\eta \downarrow 1} S^q_{\eta}(X | V);$$
(d) is by the definition of $T_{\eta}^{\lambda}(X)$, the Markov relation $V - X - (Y_1, Y_2)$, and because $T_{\lambda,\eta}^{\lambda}(X)$ is continuous in $\eta$ at $\eta = 1$ (see Remark 2).

(e) follows by Proposition 3.

By Theorem 4, for every $\eta > 1$, there exist independent random variables $X_1^* \sim \mathcal{N}(0, K_1)$, $X_2^* \sim \mathcal{N}(0, K_2)$, and $X' \sim \mathcal{N}(0, K - (K_1 + K_2))$, $K_1 + K_2 \leq K$, such that $V_2^{\eta}(K_1) = T_{\eta}^{\lambda}(X_1^* + X_2^*)$ and $S_{\eta}^{\lambda}(X_1^* + X_2^*) = S_{\eta}^{\lambda}(X_1^*)$. Moreover, setting $X = X_1^* + X_2^* + X'$ maximizes $\lambda_2 I(X; Y_2|2)$ and attains $V_2^{\eta}(K)$ simultaneously. In order to conform to the notation in the bounds, let $V^* = X'$. Taking the limit as $\eta \downarrow 1$, we have

$$\max_{(R_1, R_2) \in \mathcal{O}_K} \lambda_1 R_1 + \lambda_2 R_2 \leq \lambda_1 I(X; Y_1|V^*) - (\lambda_1 + \lambda_2) I(X; Y_2|V^*) + \lambda_1 I(X; Y_2|V^*, X_2^*) - I(X; Y_2|V^*, X_2) + \lambda_2 I(X; Y_2)$$

$$\leq \lambda_1 I(X_2^*; Y_1|V^*) - I(X_2^*; Y_2|V^*) + \lambda_2 I(V^*; Y_2).$$

(44)

The following proposition is used to show that (44) is achievable within $\bar{I}_K$.

**Proposition 8 (Partial Dirty-Paper Coding (P-DPC))** Fix a covariance matrix $K$ and let $X = X_1^* + X_2^* + V^*$, where $X_1^*$, $X_2^*$ and $V^*$ are independent Gaussian random vectors with covariance matrices $K_1$, $K_2$ and $K - (K_1 + K_2)$, respectively, for some $0 \leq K_1, K_2 \leq K$ with $K_1 + K_2 \leq K$. Let $Y_j = G_j X + Z_j$, for $j = 1, 2$, where $Z_j \sim \mathcal{N}(0, I)$ is independent of $(X_1^*, X_2^*, V^*)$. Set $U = X_2^* + AV^*$, where $A = K_2 G_1^* \left[1 + \tilde{G}_1 K_2 G_1^* \right]^{-1}$ and $\tilde{G}_1 = \left[I + G_1 K_1 G_1^* \right]^{-\frac{1}{2}} G_1$. Then

$$I(X_2^*; Y_1|V^*) - I(X_2^*; Y_2|V^*) = I(U^*; Y_1) - I(U^*; V^*) - I(X_2^*; Y_2|V^*).$$

(45)

**Proof:** We first write

$$Y_1 = G_1 X + Z_1$$

$$= G_1 (X_1^* + X_2^* + V^*) + Z_1$$

$$= G_1 (X_2^* + V^*) + (G_1 X_1^* + Z_1)$$

(46)

(a) follows by setting $\tilde{X} \triangleq X_2^* + V^*$ and $Z_1' \triangleq G_1 X_1^* + Z_1$. By the independence of $X_1^*$, $X_2^*$, $V^*$ and $Z_1$, we have that $\tilde{X}$ and $Z_1'$ are also independent. Moreover, $Z_1' \sim \mathcal{N}(0, I + G_1 K_1 G_1^*)$, where the covariance matrix $I + G_1 K_1 G_1^*$ is diagonalizable (due to its symmetry) and invertible (because it is positive definite). Denoting $\Sigma \triangleq I + G_1 K_1 G_1^*$, gives

$$\Sigma = QAQ^T,$$

(47)

where $Q$ is a unitary matrix and $A$ is diagonal, and furthermore $\Sigma^{-\frac{1}{2}} = QA^{-\frac{1}{2}}Q^T$. By defining $\tilde{Y}_1 = \Sigma^{-\frac{1}{2}} Y_1$, we have

$$\tilde{Y}_1 = \tilde{G}_1 \tilde{X} + \tilde{Z}_1,$$

(48)

where $\tilde{G}_1 = \Sigma^{-\frac{1}{2}} G_1$, $\tilde{Z}_1 = \Sigma^{-\frac{1}{2}} Z_1'$ and $\tilde{Z}_1 \sim \mathcal{N}(0, I)$.

Setting $U^*$ as above and invoking the classic Dirty-Paper Coding Theorem (here we use the formulation from Proposition 12), we have

$$I(\tilde{X}; \tilde{Y}_1|V^*) = I(U^*; \tilde{Y}_1) - (U^*; V^*).$$

(49)

Furthermore,

$$I(X_2^*; Y_2|V^*) = I(X_2^* + AV^*; Y_2|V^*) = I(U^*; Y_2|V^*).$$

(50)

Note that $Y_1 \mapsto \Sigma^{-\frac{1}{2}} Y_1$ is an invertible mapping, and as such, preserves mutual information. We conclude the proof as follows:

$$I(X_2^*; Y_1|V^*) - I(X_2^*; Y_2|V^*)$$

$$= I(\tilde{X}; \tilde{Y}_1|V^*) - I(U^*; \tilde{Y}_1) - I(U^*; Y_2|V^*)$$

$$= I(U^*; Y_1) - I(U^*; V^*) - I(U^*; Y_2|V^*) + \lambda_2 I(V^*; Y_2)$$

$$= I(U^*; Y_1) - I(U^*; V^*) - I(U^*; Y_2|V^*),$$

(51)

where (a) is because $\tilde{X} = X_2^* + V^*$ and by (50), (b) and (d) follow since $Y_1 \mapsto \Sigma^{-\frac{1}{2}} Y_1$ preserves mutual information, while (c) uses (49).

Inserting $U^*$ as stated in Proposition 8 into the RHS of (44), we obtain

$$\max_{(R_1, R_2) \in \mathcal{O}_K} \lambda_1 R_1 + \lambda_2 R_2$$

$$\leq \lambda_1 I(X_2^*; Y_1|V^*) - I(X_2^*; Y_2|V^*) + \lambda_2 I(V^*; Y_2)$$

$$= \lambda_1 I(U^*; Y_1) - I(U^*; V^*) - I(U^*; Y_2|V^*) + \lambda_2 I(V^*; Y_2)$$

(52)

where (a) follows since $(U^*, V^*) - X - (Y_1, Y_2)$ forms a Markov chain and $E[XX^T] \preceq K$ is satisfied, which implies that the rate pair $R_1 = I(U^*; Y_1) - I(U^*; V^*) - I(U^*; Y_2|V^*)$ and $R_2 = I(V^*; Y_2)$ belongs to $\bar{I}_K$. Concluding, we see that $\bar{I}_K = \tilde{C}_K = C_K$, which characterizes the secrecy-capacity region of the MIMO Gaussian BC with private and confidential messages.

Furthermore, equality (and hence the extreme points of $\tilde{C}_K$) is achieved by Gaussian inputs as stated in Proposition 8, thus making the region computable. By evaluating $\bar{I}_K$ (or, equivalently $\tilde{C}_K$) with respect to this input distribution, we describe the secrecy-capacity region $\tilde{C}_K$ as the union of rate pairs $(R_1, R_2) \in \mathcal{R}_K$ satisfying:

$$R_1 \leq \frac{1}{2} \log \left[ \frac{I + G_1 (K_1 + K_2) G_1^T}{I + G_1 K_1 G_1^T} \right]$$

(53a)

$$R_2 \leq \frac{1}{2} \log \left[ \frac{I + G_2 (K_1 + K_2) G_2^T}{I + G_2 K_1 G_2^T} \right]$$

(53b)

where the union is over all positive semi-definite matrices $K_1, K_2$, such that $K_1 + K_2 \preceq K$. We further simplify (53) by noting that the RHS of (53a) is the secrecy-capacity of the MIMO Gaussian WTC as derived in [16, Appendix III].
which is maximized by setting $K_1 = 0$ (see (7–9)). Further note that $K_1 = 0$ cannot decrease the RHS of (53b). Thus, by relabeling $K_2 \triangleq K^*$ we establish (9).

B. Proof of Theorem 2

As in the case without a common message, the secrecy-capacity region $C_K$ is derived by showing that certain outer bound and inner bounds on $C_K$ coincide. Denoting by $C$ the region of the DM-BC with common, private and confidential messages, we bound it as follows.

**Bound 3 (Outer Bound)** Let $O$ be the closure of the union of rate triples $(R_0, R_1, R_2) \in \mathbb{R}^+_3$ satisfying:

\begin{align*}
R_0 &\leq \min \{ I(W; Y_1), I(W; Y_2) \} \quad (54a) \\
R_1 &\leq I(U; Y_1 | W, V) - I(U; Y_2 | W, V) \quad (54b) \\
R_0 + R_2 &\leq I(V; Y_2 | W) + \min \{ I(W; Y_1), I(W; Y_2) \} \quad (54c)
\end{align*}

over all $(W, V, U) - X - (Y_1, Y_2)$. Then $C \subseteq O$.

**Bound 4 (Inner Bound)** Let $\mathcal{I}$ be the closure of the union of rate triples $(R_0, R_1, R_2) \in \mathbb{R}^+_3$ satisfying:

\begin{align*}
R_0 &\leq \min \{ I(W; Y_1), I(W; Y_2) \} \quad (55a) \\
R_1 &\leq I(U; Y_1 | W) - I(U; Y_2 | W) - I(U; Y_2 | W, V) \quad (55b) \\
R_2 &\leq I(V; Y_2 | W) \quad (55c)
\end{align*}

over all $(W, V, U) - X - (Y_1, Y_2)$. Then $\mathcal{I} \subseteq C$.

The proofs of Bounds 3 and 4 are relegated to Appendix A. Denoting by $O_K$ and $\mathcal{I}_K$ the adaptations of Bounds 3 and 4 to a MIMO Gaussian BC with a covariance input constraint $\mathbb{E}[XX^\top] \preceq K$, we have $\mathcal{I}_K \subseteq C_K \subseteq O_K$.

Next, we use the factorization of UCEs method to show that the opposite inclusion $C_K \subseteq \mathcal{I}_K$ also holds. Given the supporting hyperplanes characterization of bounded and closed convex sets, using a similar reasoning as in Section V-A (see Lemma 1), it suffices to study $\max_{(R_0, R_1, R_2) \in C_K} \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2$, for $\lambda_j > 0$, $j = 1, 2, 3$. Further note that it suffices to restrict attention to the case where $\lambda_0 > 2$. This follows from the following observation: If a rate triple $(R_0, R_1, R_2)$ is in $C_K$ then so does the triple $(0, R_1, R_2 + R_0)$, since one may always treat the common message as part of the (non-confidential) private message to Receiver 2. Assuming $\lambda_0 \leq \lambda_2$, we have

$$\lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2 \leq 0 \cdot R_0 + \lambda_1 R_1 + \lambda_2 (R_0 + R_2),$$

and therefore,

$$\begin{align*}
\max_{(R_0, R_1, R_2) \in C_K} \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2 &= \max_{(0, R_1, R_2) \in C_K} \lambda_0 R_1 + \lambda_2 R_2 \\
&= \max_{(R_0, R_1, R_2) \in C_K} \lambda_1 R_1 + \lambda_2 R_2,
\end{align*}$$

where $C_K$ is the secrecy-capacity region without a common message that was characterized in Section V-A.

Hence, it suffices to show that for all $\lambda_j > 0$, $j = 1, 2, 3$, with $\lambda_0 > \lambda_2$, we have

$$\begin{align*}
\max_{(R_0, R_1, R_2) \in O_K} \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2 &\leq \max_{(R_0, R_1, R_2) \in \mathcal{I}_K} \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2. \quad (58)
\end{align*}$$

Now, for any $\alpha \in [0, 1]$ set $\bar{\alpha} = 1 - \alpha$, and consider the following.

$$\begin{align*}
\max_{(R_0, R_1, R_2) \in \mathcal{I}_K} &\lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2 \\
\leq& \sup_{(W, V, U) - X - (Y_1, Y_2)} \lambda_0 \left[ \alpha I(W; Y_1) + \bar{\alpha} I(W; Y_2) \right] \\
&+ \lambda_1 \left[ I(U; Y_1 | W, V) - I(U; Y_2 | W, V) \right] + \lambda_2 I(V; Y_2 | W) \quad (a)
\end{align*}$$

(b) is by the definition of $S_q^q(X|W, V)$ since conditioned on $(W, V), U - X - (Y_1, Y_2)$ forms a Markov chain, and hence $S_q^q(X|W, V)$ is continuous in $\eta$ at $\eta = 1$ (Property 3 of Proposition 1).

(d) follows by the definition of $T_{X, \eta}(X|W)$ since conditioned on $V, W - X - (Y_1, Y_2)$ forms a Markov chain. Furthermore, the continuity of $T_{X, \eta}(X|W)$ in $\eta$ at $\eta = 1$ (see Remark 2) is also exploited.

(e) is by the definition of $F_{X, \eta}(X)$ (while noting that $W - X - (Y_1, Y_2)$ forms a Markov chain), and because $F_{X, \eta}(X)$ is continuous at $\eta = 1$ (Remark 10).
Having \( (63) \), the desired equality \( X_\bar{\alpha} (f) \) makes use of the continuity argument from Remark \( 11 \).

Recall that for any \( \eta > 1 \), \( \lambda_j > 0 \), for \( j = 0, 1, 2 \), and \( \lambda_0 > \lambda_2 \). Corollary \( 3 \) implies that there exist independent random variables \( X_j^* \sim \mathcal{N}(0, K_j) \), \( j = 1, 2, 3 \), and \( X' \sim \mathcal{N}(0, K - (K_1 + K_2 + K_3)) \), such that \( (37) \) is satisfied. Furthermore, setting \( X = X_1^* + X_2^* + X_3^* + X' \) not only attains \( V_0^* \) (K), but it also simultaneously maximizes \( \alpha_0 \lambda(X; Y_1) \) and \( \bar{\alpha}_0 \lambda(X; Y_2) \). Relabeling \( W^* = X' \) and \( V^* = X_3^* \) while taking the limit as \( \eta \downarrow 1 \), we have

\[
\begin{align*}
\max_{(\rho_0, \rho_1, \rho_2) \in \mathcal{O}_K} & \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2 \\
& \leq \alpha_0 \lambda(I(X; Y_1)) + \bar{\alpha}_0 \lambda(I(X; Y_2)) \\
& + \lambda_0 I(X; Y_1 | W^*, V^*) - \lambda_0 I(X; Y_1 | W^*) \\
& + \lambda_1 I(X; Y_2 | W^*, V^*) - \lambda_1 I(X; Y_2 | W^*) \\
& + \lambda_2 I(V^*; Y_2 | W^*) \\
& \leq \lambda_0 \alpha(I(W^*; Y_1) + \bar{\alpha}(I(W^*; Y_2)) \\
& + \lambda_1 [I(U^*; Y_1 | W^*) - I(U^*; V^* | W^*) - I(U^*; Y_2 | W^*)] \\
& + \lambda_2 I(V^*; Y_2 | W^*). 
\end{align*}
\]

Using Proposition \( 8 \) we set \( U = X_1^* + \tilde{A} V^* \) as before and obtain

\[
I(X_2^*; Y_1 | W^*, V^*) - I(X_2^*; Y_2 | W^*, V^*) = I(U^*; Y_1 | W^*) - I(U^*; V^* | W^*) - I(X_2^*; Y_2 | W^*, V^*). 
\]

Inserting \( (61) \) into \( (60) \), yields

\[
\begin{align*}
\max_{(\rho_0, \rho_1, \rho_2) \in \mathcal{O}_K} & \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2 \\
& \leq \lambda_0 \alpha(I(W^*; Y_1) + \bar{\alpha}(I(W^*; Y_2)) \\
& + \lambda_1 [I(U^*; Y_1 | W^*) - I(U^*; V^* | W^*) - I(U^*; Y_2 | W^*)] \\
& + \lambda_2 I(V^*; Y_2 | W^*) \\
& \leq \lambda_0 \alpha(I(W; Y_1) + \bar{\alpha}(I(W; Y_2)) \\
& + \lambda_1 [I(U; Y_1 | W) - I(U; V | W) - I(U; Y_2 | W, V)] \\
& + \lambda_2 I(V; Y_2 | W). 
\end{align*}
\]

Since \( (62) \) holds for all \( \alpha \in [0, 1] \), we have

\[
\begin{align*}
\max_{(\rho_0, \rho_1, \rho_2) \in \mathcal{O}_K} & \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2 \\
& \leq \min_{\alpha \in [0, 1]} \sup_{(W, V, U) \in \mathcal{X}, \mathcal{Y}, \mathcal{Z}} \lambda_0 \alpha(I(W; Y_1) + \bar{\alpha}(I(W; Y_2)) \\
& + \lambda_1 [I(U; Y_1 | W) - I(U; V | W) - I(U; Y_2 | W, V)] \\
& + \lambda_2 I(V; Y_2 | W). 
\end{align*}
\]

Having \( (63) \), the desired equality \( \mathcal{O}_K = \mathcal{Z}_K \) is a consequence of the following Proposition \( 8 \).

**Proposition 9 (Max-Min Interchanging)** The following max-min interchanging holds

\[
\begin{align*}
\min_{\alpha \in [0, 1]} & \sup_{(W, V, U) \in \mathcal{X}, \mathcal{Y}, \mathcal{Z}} \lambda_0 \alpha(I(W; Y_1) + \bar{\alpha}(I(W; Y_2)) \\
& + \lambda_1 [I(U; Y_1 | W) - I(U; V | W) - I(U; Y_2 | W, V)] \\
& + \lambda_2 I(V; Y_2 | W) \\
& = \sup_{(W, V, U) \in \mathcal{X}, \mathcal{Y}, \mathcal{Z}} \min_{\alpha \in [0, 1]} \lambda_0 \alpha(I(W; Y_1) + \bar{\alpha}(I(W; Y_2)) \\
& + \lambda_1 [I(U; Y_1 | W) - I(U; V | W) - I(U; Y_2 | W, V)] \\
& + \lambda_2 I(V; Y_2 | W) \\
& = \sup_{(W, V, U) \in \mathcal{X}, \mathcal{Y}, \mathcal{Z}} \lambda_0 \cdot \min \{I(W; Y_1), I(W; Y_2)\} \\
& + \lambda_1 [I(U; Y_1 | W) - I(U; V | W) - I(U; Y_2 | W, V)] \\
& + \lambda_2 I(V; Y_2 | W). 
\end{align*}
\]

The proof of Proposition \( 9 \) is given in Appendix \( D \). Now, noting that \( (W, V, U) - X - (Y_1, Y_2) \) forms a Markov chain and \( E[XX^T] \leq K \), we see that the triple

\[
\begin{align*}
R_0 &= \min \{I(W; Y_1), I(W; Y_2)\} \\
R_1 &= I(U; Y_1 | W) - I(U; V | W) - I(U; Y_2 | W, V) \\
R_2 &= I(V; Y_2 | W) \\
\end{align*}
\]

is inside the inner bound \( \mathcal{I}_K \). Hence

\[
\begin{align*}
\max_{(\rho_0, \rho_1, \rho_2) \in \mathcal{O}_K} & \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2 \\
& \leq \sup_{(W, V, U) \in \mathcal{X}, \mathcal{Y}, \mathcal{Z}} \lambda_0 \alpha(I(W; Y_1) + \bar{\alpha}(I(W; Y_2)) \\
& + \lambda_1 [I(U; Y_1 | W) - I(U; V | W) - I(U; Y_2 | W, V)] \\
& + \lambda_2 I(V; Y_2 | W) \\
& \leq \max_{(\rho_0, \rho_1, \rho_2) \in \mathcal{I}_K} \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2, 
\end{align*}
\]

which implies that \( \mathcal{I}_K = \mathcal{C}_K = \mathcal{O}_K \) and characterizes the secrecy-capacity region of the MIMO Gaussian BC with common, private and confidential messages.

To obtain the description of \( \mathcal{C}_K \) stated in \( (10) \), note that when \( \lambda_0 > \lambda_2 \), equality (and hence the extreme points of \( \mathcal{C}_K \)) is achieved by setting

\[
X = X_1^* + X_2^* + V^* + W^*,
\]

where \( X_j^* \sim \mathcal{N}(0, K_j) \), \( j = 1, 2 \), \( V^* \sim \mathcal{N}(0, K_3) \) and \( W^* \sim \mathcal{N}(0, K - (K_1 + K_2 + K_3)) \) are independent of each other, and \( U = X_2^* + \tilde{A} V^* \), where \( \tilde{A} \) is the P-DPC matrix from Proposition \( 8 \). For the case when \( \lambda_0 \leq \lambda_2 \), \( (64) - (67) \) imply that the boundary-achieving input distribution corresponds to the one that achieves the secrecy-capacity region when there is no common message (see Section \( 12 \)). Setting \( K_3 = K - (K_1 + K_2) \) recovers the optimal input distribution for the case without common message.

By evaluating \( \mathcal{I}_K \) (or, equivalently \( \mathcal{O}_K \)) with respect to \( (66) \),
we characterize the secrecy-capacity region $C_K$ as the union of rate triples $(R_0, R_1, R_2) \in \mathbb{R}^+_3$ satisfying:

$$R_0 \leq \min \left\{ \frac{1}{2} \log \frac{I + G_1 K G_1^T}{I + G_1 (K_1 + K_2 + K_3) G_1} \right\}$$

$$R_1 \leq \frac{1}{2} \log \frac{I + G_1 (K_1 + K_2) G_1^T}{I + G_1 K_1 G_1^T}$$

$$R_2 \leq \frac{1}{2} \log \frac{I + G_2 (K_1 + K_2) G_2^T}{I + G_2 (K_1 + K_2 + K_3) G_2}$$

where the union is over all positive semi-definite matrices $K_1, K_2, K_3$, such that $K_1 + K_2 + K_3 \leq K$.

The region from (67a) is further simplified using reasoning similar to that from Section V-A. First, (67b) indicates that the signal to User 1 is a sum of two independent zero mean Gaussian random vectors with covariance $K_1$ and $K_2$. The signal that corresponds to User 2 carries the confidential message $M_1$, while the $K_1$ signal is an artificial noise sent (on purpose) to confuse User 2 (which serves as a eavesdropper of $M_1$). The lack of structure in the artificial noise adds to the noise floor at both receivers (see also [13] Remark 4). However, since the RHS of (67b) is the secrecy-capacity of the MIMO Gaussian WTC, it is maximized by setting $K_1 = 0$. Furthermore, since in both (67a) and (67c) $K_1$ serves as noise (i.e., it is not used to encode any of the messages), setting $K_1 = 0$ achieves optimality. This is since $K_1 = 0$ corresponds to revealing the $K_1$ signal to both receivers, which can only increase the transmission rates. Taking $K_1 = 0$ and recasting $K_3$ as $K_1$, recovers (10).

VI. PROOFS OF UPPER CONCAVE ENVELOPES PROPERTIES

A. Proof of Proposition 4

Property 1 follows by Jensen’s inequality since $S^q_\alpha$ is concave in $P_\lambda$, while for Property 2 we use the fact that $P_{X|W,V} = P_{X|V}$. To prove Property 3, fix $P_X$ and let $\eta_1, \eta_2 \in (0, 2)$, $\alpha \in [0, 1]$ and $\bar{\alpha} = 1 - \alpha$. Observe that

$$S^q_{\alpha \eta_1 + \bar{\alpha} \eta_2}(X) = \sup_{P_{X|Y_1} \in \mathcal{P}_{\alpha \eta_1 + \bar{\alpha} \eta_2}(X)} I(X; Y_2|V) - (\alpha \eta_1 + \bar{\alpha} \eta_2) I(X; Y_1|V) \leq \alpha \cdot \sup_{P_{X|Y_1} \in \mathcal{P}_{\alpha \eta_1}} I(X; Y_2|V) - \eta_1 I(X; Y_1|V)$$

$$+ \bar{\alpha} \cdot \sup_{P_{X|Y_1} \in \mathcal{P}_{\bar{\alpha} \eta_2}} I(X; Y_2|V) - \eta_2 I(X; Y_1|V) = \alpha S^q_{\eta_1}(X) + \bar{\alpha} S^q_{\eta_2}(X).$$

Clearly, $S^q_\eta(X)$ is also bounded for every $\eta \in (0, 2)$ and by invoking Proposition 17 from [28] Chapter 5, we have that $S^q_\eta(X)$ is continuous inside every closed subinterval of $(0, 2)$, and in particular, at $\eta = 1$.

B. Proof of Proposition 2

Let $V - (X_1, X_2) - (Y_{11}, Y_{12}, Y_{21}, Y_{22})$ form a Markov chain. We have,

$$t^{q_1, q_2}_{\eta_1}(X_1, X_2|V) = \lambda_1 I(X_1, X_2; Y_{11}, Y_{12}|V) - (\lambda_1 + \lambda_2) I(X_1, X_2; Y_{21}, Y_{22}|V) + \lambda_1 S^{q_1, q_2}_{\eta_1}(X_1, X_2|V)$$

$$= \lambda_1 \left[ I(X_1, X_2; Y_{11}|V) + I(X_1, X_2; Y_{12}|V, Y_{11}) \right]$$

$$+ \lambda_1 I(X_1, X_2; Y_{22}|V, Y_{21}) + \lambda_1 S^{q_1, q_2}_{\eta_1}(X_1, X_2|V)$$

$$= \lambda_1 \left[ I(X_1; Y_{11}|V, Y_{22}) + I(X_2; Y_{12}|V, Y_{11}) + I(Y_{11}; Y_{22}|Y_{11}) \right]$$

+ \lambda_1 S^{q_1, q_2}_{\eta_1}(X_1, X_2|V)$$

$$\leq \lambda_1 I(X_1; Y_{11}|V, Y_{22}) - (\lambda_1 + \lambda_2) I(X_1; Y_{21}|V, Y_{22})$$

$$+ \lambda_1 S^{q_1}_{\eta_1}(X_1|V, Y_{22}) + \lambda_1 I(X_2; Y_{12}|V, Y_{11}) - (\lambda_1 + \lambda_2) I(X_2; Y_{22}|V, Y_{11}) + \lambda_1 S^{q_1, q_2}_{\eta_1}(X_1, X_2|V)$$

$$\leq T^{q_1}_{X_1; \eta_1}(X_1|Y_{22}) + T^{q_2}_{X_1; \eta_1}(X_1|Y_{11}) - \lambda_2 I(Y_{11}; Y_{22}|V)$$

$$\leq T^{q_1}_{X_1; \eta_1}(X_1) + T^{q_2}_{X_1; \eta_1}(X_1) - \lambda_2 I(Y_{11}; Y_{22}|V)$$

Now for $(V^*, X_1^*, X_2^*)$, an end-to-end equality holds in (68). In particular, this implies that $I(Y_{11}^*; Y_{22}^*|V^*) = 0$, i.e., that $Y_{11}^* = V^* - Y_{22}^*$ forms a Markov chain. By Proposition 2 in [16], we have that $X_1^* = V^* - X_2^*$, which further implies the Markov chain

$$(Y_{11}^*, Y_{21}^*) - X_1^* = V^* - X_2^* - (Y_{12}^*, Y_{22}^*).$$

The end-to-end equality in (68) also gives

$$T^{q_1}_{\eta_1}(X_1^*) = \lambda_1 I(X_1^*; Y_{11}^*|V^*); Y_{22}^*) - (\lambda_1 + \lambda_2) I(X_1^*; Y_{21}^*|V^*); Y_{22}^*) + \lambda_1 S^{q_1}_{\eta_1}(X_1^*|V^*)$$

$$\leq \lambda_1 I(X_1^*; Y_{11}^*|V^*) - (\lambda_1 + \lambda_2) I(X_1^*; Y_{21}^*|V^*) + \lambda_1 S^{q_1}_{\eta_1}(X_1^*|V^*)$$

$$= \lambda_1 S^{q_1}_{\eta_1}(X_1^*) = T^{q_2}_{X_1^*; \eta_1}(X_1^*|V^*)$$.
C. Proof of Proposition 3

Let $V - (X_1, X_2) - (Y_{11}, Y_{12}, Y_{21}, Y_{22})$ be a Markov chain. We have,

$$f_{\lambda, \alpha, \eta}^{q_1,q_2}(X_1, X_2 | V) = (\lambda_2 - \tilde{\alpha}_\lambda) I(X_1; Y_1 | V, Y_1) + \tilde{\alpha}_\lambda I(X_1; Y_2 | V, Y_1) + \tilde{\alpha}_\lambda I(X_1; Y_2 | V, Y_2)$$

$$= (\lambda_2 - \tilde{\alpha}_\lambda) I(X_1; Y_1 | V) + \tilde{\alpha}_\lambda I(X_1; Y_2 | V) + \tilde{\alpha}_\lambda I(X_1; Y_2 | V)$$

Proposition 10 (Boundedness and Continuity of UCE)

Consider the space of all Borel probability distributions on $\mathbb{R}^4$ endowed with the topology induced by weak convergence. \(\mathcal{D}\) Let \(\{X_n\}_{n \in \mathbb{N}}\) be a sequence of random variable that satisfies the following two properties: (i) \(\exists p > 1, B \in \mathbb{R}\) such that \(E[|X_n|^p] \leq B, \forall n \in \mathbb{N}\) (i.e., the sequence has a uniformly bounded \(p\)th moment); (ii) \(X_n \overset{D}{\longrightarrow} X^*\). If \(g : \mathbb{R}^4 \to \mathbb{R}\) is a bounded real-valued function that satisfies \(g(X_n) \overset{D}{\longrightarrow} g(X^*)\), then its UCE $G = g\mathcal{E}_3$ is bounded and satisfies \(G(\mathbb{X}_n) \overset{n}{\longrightarrow} G(\mathbb{X}^*)\).

Before proving Proposition 6 at the end of this subsection, we first verify that Proposition 10 applies to \(T_{\lambda, \alpha, \eta}^q(X)\). This is done by showing that \(T_{\lambda, \alpha, \eta}^q(X)\) is bounded and continuous, as stated in the subsequent Lemma 2. The lemmas are proven, respectively, in Appendices C-A and C-B.

Lemma 2 (Boundedness of Nested Concave Envelopes)

For \(\eta > 1\) and \(\lambda_1, \lambda_2 > 0\) there is a \(B_{\lambda, \eta} \in \mathbb{R}\) such that \(T_{\lambda, \alpha, \eta}^q(X) \leq B_{\lambda, \eta}\) for all \(P_X\).

Lemma 3 (Continuity of Nested Concave Envelopes)

Consider the space of all Borel probability distributions on $\mathbb{R}^4$ endowed with the topology induced by weak convergence and let \(\{X_n\}_{n \in \mathbb{N}}\) be a sequence of random variable that satisfies the following two properties: (i) \(\exists p > 1, B \in \mathbb{R}\) such that \(E[|X_n|^p] \leq B, \forall n \in \mathbb{N}\); (ii) \(X_n \overset{D}{\longrightarrow} X^*\). Then \(T_{\lambda, \alpha, \eta}^q(X_n) \overset{n}{\longrightarrow} T_{\lambda, \alpha, \eta}^q(X^*)\).

Based on Lemmas 2 and 3, Proposition 21 from [16] states that \(T_{\lambda, \alpha, \eta}^q(X_n)\) is bounded and that it satisfies \(T_{\lambda, \alpha, \eta}^q(X_n) \overset{n}{\longrightarrow} T_{\lambda, \alpha, \eta}^q(X^*)\). The existence of a unique maximizer of \(V_{\lambda, \alpha, \eta}(K)\) is established as follows. Let \(K \geq 0\) and define \(V_{\lambda, \alpha, \eta}(K) = \sup_{X \in \mathcal{D}} f_{\lambda, \alpha, \eta}(X)\).

Let \(\{X_n\}_{n \in \mathbb{N}}\) be a sequence of random variables with $E[X_n X_n^*] = K$, such that $f_{\lambda, \alpha, \eta}(X_n) \uparrow V_{\lambda, \alpha, \eta}(K)$, as $n \to \infty$. By Proposition 17 and since $E[X_n X_n^*] = K$ for every $n \in \mathbb{N}$, we have that \(\{X_n\}_{n \in \mathbb{N}}\) is a tight sequence. Further, there exist an $X_K$ and a convergent subsequence \(\{X_{n_m}\}_{m \in \mathbb{N}}\) such that $X_{n_m} \overset{D}{\longrightarrow} X_K$. Invoking Proposition 18 once more we have that $h(Y_{j,n_m}) \overset{m}{\longrightarrow} h(Y_{j,K})$.

A sequence \(\{X_n\}_{n \in \mathbb{N}}\) of real-valued random variables is said to converge weakly or, equivalently, convert in distribution to a random variable \(X\) if $\lim_{n \to \infty} F_{X_n}(x) = F(x)$, for every $x \in \mathbb{R}$ for which $F_X$ is continuous, where $F_{X_n}$ and $F_X$ are the cumulative distribution functions of $X_n$ and $X$, respectively. This notion of convergence is denoted by $X_n \overset{D}{\longrightarrow} X$.

As defined in Proposition 21, a sequence of random variables \(\{X_n\}_{n \in \mathbb{N}}\) taking values in $\mathbb{R}^4$ is tight if for every $\epsilon > 0$ there exists a compact set $C \subset \mathbb{R}^4$, such that $P(X_n \not\in C) \leq \epsilon$, $\forall n \in \mathbb{N}$.
\( h\left(\mathbf{Y}_{j,j}^*\right) \), for \( j = 1, 2 \), where \( \mathbf{Y}_{1, n_1}, \mathbf{Y}_{2, n_2}, \mathbf{Y}_{1, \tilde{K}}^* \) and \( \mathbf{Y}_{2, \tilde{K}}^* \) are the corresponding outputs. Thus,

\[
f_{\lambda_0, \alpha, \eta}(\mathbf{X}_{\tilde{K}}^*) = \tilde{v}_{\lambda_0, \alpha, \eta}(\tilde{K}).
\]

(75)

By the definition of \( \tilde{V}_{\lambda_0, \alpha, \eta}(\tilde{K}) \), we write

\[
\tilde{V}_{\lambda_0, \alpha, \eta}(\tilde{K}) = \max \left\{ V : \mathbb{E}[XX^\top] \preceq \tilde{K}, V \preceq X - (Y_1, Y_2) \right\} = \max \left\{ V : \mathbb{E}[XX^\top] \preceq \tilde{K}, V \preceq X - (Y_1, Y_2) \right\}
\]

\[
\sum_{v} P(v)f^q_{\lambda_0, \alpha, \eta}(X|V = v).
\]

(76)

Since \( V_{\lambda_0, \alpha, \eta}(\tilde{K}) \) is a convex combination as above, to obtain the maximizer subject to the covariance constraint it suffices to restrict attention to the family of maximizers \( \mathbf{X}_{K_i}^* \) for \( K_i \geq 0 \). Thus,

\[
\tilde{V}_{\lambda_0, \alpha, \eta}(K_i) = \max \left\{ \alpha_1 V_1, \ldots, \alpha_m V_m \right\} \quad \text{subject to} \quad \sum_{i=1}^{m} \alpha_i = 1, \quad \sum_{i=1}^{m} \alpha_i K_i \leq K.
\]

(77)

It takes \( \left(\frac{(t+1)}{2}\right) \) constraints to preserve the covariance matrix (due to its symmetry) and one other constraint to preserve \( \sum_{i=1}^{m} \alpha_i K_i \leq K \). Hence, by using the Bunt-Carathedory theorem \([28]\), we can restrict ourselves to convex combinations of at most \( m \geq \frac{(t+1)}{2} \) + 1 points, i.e.,

\[
\tilde{V}_{\lambda_0, \alpha, \eta}(K_i) = \max \left\{ \alpha_1 V_1, \ldots, \alpha_m V_m \right\} \quad \text{subject to} \quad \sum_{i=1}^{m} \alpha_i = 1, \quad \sum_{i=1}^{m} \alpha_i K_i \leq K.
\]

(78)

Consider any sequence of convex combinations \( \{K_i^{(n)}\}_{n \in \mathbb{N}} \) that approaches the supremum as \( n \to \infty \). The compactness of the \( m \)-dimensional simplex implies that \( \alpha_i^{(n)} \to \alpha_i^* \), for all \( i \in [m] \). Furthermore, we have the following property of the limiting points \( \alpha_i^* \); see Appendix C for the proof.

**Lemma 4** For any \( i \in [m] \), if \( \alpha_i^* = 0 \) then \( \alpha_i^{(n)} \to \alpha_i^* \) as \( n \to \infty \).

Based on Lemma 4, we assume that \( \alpha^* = \min_{i \in [m]} \alpha_i^* > 0 \), which implies that \( K_i^{(n)} \rightarrow K_i^* \) uniformly in \( i \in [m] \), for large enough values of \( n \). Hence, for each \( i \in [m] \) we can find a convergent subsequence \( \{K_i^{(n_k)}\}_{k \in \mathbb{N}} \) such that \( K_i^{(n_k)} \rightarrow K_i^* \). Putting these together, we have

\[
\tilde{V}_{\lambda_0, \alpha, \eta}(K) = \max \left\{ \alpha_1 V_1, \ldots, \alpha_m V_m \right\} \quad \text{subject to} \quad \sum_{i=1}^{m} \alpha_i = 1, \quad \sum_{i=1}^{m} \alpha_i K_i \leq K.
\]

(79)

i.e., one can always find a pair of random variables \( (V^*, X^*) \) with \( |V^*| \leq \frac{(t+1)}{2} + 1 \), such that \( \tilde{V}_{\lambda_0, \alpha, \eta}(K) = f_{\lambda_0, \alpha, \eta}(X^*|V^*) \).

**E. Proof of Proposition 7**

Denote \( q_{Y_{1}, Y_{2}|X} \) and consider the two-letter BC \( q(Y_{11}, Y_{21}|X_11) \times q(Y_{12}, Y_{22}|X_22) \). We have

\[
2\tilde{V}_{\lambda_0, \alpha, \eta}(K) \preceq f_{\lambda_0, \alpha, \eta}(X_1, X_2|V_1, V_2) = f_{\lambda_0, \alpha, \eta}(X_1, X_2|V_1, V_2)
\]

(80)

where:

(a) is because \( P_{\ast X}^v \) achieves \( \tilde{V}_{\lambda_0, \alpha, \eta}(K) \);

(b) uses the independence of \( (V_1, X_1) \) and \( (V_2, X_2) \);

(c) is a consequence of \([16]\) Proposition 1) (namely, the invariance with respect to rotation of the mutual information between the input and output of an additive Gaussian channel);

(d) follows by the definition of the double-nested UCE;

(e) uses Proposition 5.

(f) is the definition of \( \tilde{V}_{\lambda_0, \alpha, \eta}(K) \), while noting that the independence of \( X_{v_1} \) and \( X_{v_2} \), for every \( v_1 \neq v_2 \), implies that

\[
\mathbb{E}[X_{\theta_1}X_{\theta_2}^\top] = \mathbb{E}\left[\frac{1}{2}X_{V_1}X_{V_1}^\top + \frac{1}{2}X_{V_2}X_{V_2}^\top\right] = \mathbb{E}[X_{\theta_1}X_{\theta_2}^\top]
\]

(81)

and

\[
\mathbb{E}[X_{\theta_j}X_{\theta_j}^\top] = \sum_v P_v(v)K_v = \mathbb{E}\left[\left[XX^\top|V\right]\right] \leq K,
\]

where we have denoted \( K_v = \mathbb{E}[X_{v}X_{v}^\top] \).

Since the extremes of the chain of inequalities in \([80]\) match, all inequalities are, in fact, equalities. Equality in step (d) implies that \( P_{V|X_1, X_2} \) achieves \( F_{\lambda_0, \alpha, \eta}(X_1, X_2) \). Furthermore, by Proposition 5 since (d) and (e) are equalities we have that \( X_{\theta_1} = V - X_{\theta_2} \), and that \( P_{V|X_1, X_2} \) and \( P_{V|X_{\theta_2}} \) achieve \( F_{\lambda_0, \alpha, \eta}(X_1) \) and \( F_{\lambda_0, \alpha, \eta}(X_2) \), respectively. Finally, equality in (f) means that \( \tilde{V}_{\lambda_0, \alpha, \eta}(K) = f_{\lambda_0, \alpha, \eta}(X_{\theta_j}|V) = f_{\lambda_0, \alpha, \eta}(X_{\theta_j}|V_j) \), for \( j = 1, 2 \).

**F. Proof of Theorem 5**

As a consequence of Proposition 7 for any fixed \( (v_1, v_2) \in V^2 \), \( X_{v_1} + X_{v_2} \) and \( X_{v_1} - X_{v_2} \) are independent. Combined with \( X_{v_1} \) and \( X_{v_2} \) being independent zero mean random variables, Corollary 3 in Appendix A of \([16]\) implies that \( X_{v_1} \) and \( X_{v_2} \) are Gaussian random vectors with the same covariance matrix. Since the pair \( (v_1, v_2) \in V^2 \) is arbitrary, we see that all Gaussian vectors \( \{X_v\}_{v \in V} \) have the same covariance matrix. Furthermore, we may assume that \( \{X_v\}_{v \in V} \) are all centered, and therefore, we get that this is an i.i.d. set of Gaussian random variables. Denoting this common covariance matrix
by $K^*$, it clearly satisfies $K^* \leq K$. Letting $X^* \sim \mathcal{N}(0, K^*)$, we have
\[
\tilde{V}^q_{\lambda_0, \alpha, \eta}(K) = \frac{f^q_{\lambda_0, \alpha, \eta}(X|V)}{\mathbb{E}_{V \sim \mathcal{N}(0, K^*)} f^q_{\lambda_0, \alpha, \eta}(X|V)},
\]
where (a) follows since $(V, X) \sim P^*_{V, X}$, (b) follows by the definition of $X_v$ in the statement of Proposition 1 while (c) follows since $X^*$ and $X_v$ are identically distributed, for every $i \in [m]$.

To account for the uniqueness of the zero-mean maximizer we first show that if a zero mean random vector $X$ is a maximizer, i.e., $\tilde{V}^q_{\lambda_0, \alpha, \eta}(K)$ has two independent Gaussian maximizers denoted by $A_1 \sim \mathcal{N}(0, K_1)$ and $A_2 \sim \mathcal{N}(0, K_2)$, such that $K_1, K_2 \leq K$ and $K_1 \neq K_2$. Let $(V, X)$ be a pair of random variables, such that $V \sim \text{Ber}(\frac{1}{2})$ on $V = \{1, 2\}$, $X|V = 1 \sim \mathcal{N}(0, K_1)$ and $X|V = 2 \sim \mathcal{N}(0, K_2)$. Note that $(V, X)$ also attains $\tilde{V}^q_{\lambda_0, \alpha, \eta}(K)$. Taking $v_1 = 1$ and $v_2 = 2$, Proposition 1 implies that $A_1 + A_2$ and $A_1 - A_2$ are independent, which contradict Corollary 3 from 10 as $K_1 \neq K_2$.

G. Proof of Corollary 3

By Theorem 4 there is an $X^* \sim \mathcal{N}(0, K^*)$, such that $K^* \leq K$ and $f^q_{\lambda_0, \alpha, \eta}(X^*) = \tilde{V}^q_{\lambda_0, \alpha, \eta}(K)$. Let $X' \sim \mathcal{N}(0, K - K^*)$ be independent of $X^*$ and set $X = X^* + X'$. Thus $X \sim \mathcal{N}(0, K)$ and by definition we have $f^q_{\lambda_0, \alpha, \eta}(X) \leq \tilde{V}^q_{\lambda_0, \alpha, \eta}(K)$.

On the other hand,
\[
f^q_{\lambda_0, \alpha, \eta}(X) = \sup_{P_{V|X}} f^q_{\lambda_0, \alpha, \eta}(X|V)
\]
\[
\geq f^q_{\lambda_0, \alpha, \eta}(X|X')
\]
\[
= \tilde{V}^q_{\lambda_0, \alpha, \eta}(K),
\]
where (a) follows since $X' - X - (Y_1, Y_2)$ forms a Markov chain, while (b) follows because $X|\{X' = x'\} \sim \mathcal{N}(x^* + x')$. Thus,
\[
\tilde{V}^q_{\lambda_0, \alpha, \eta}(K) = f^q_{\lambda_0, \alpha, \eta}(X) = f^q_{\lambda_0, \alpha, \eta}(X^*) - \lambda^*(X^*, Y_2) - \lambda^*(X^*, Y_1) + T^q_{\lambda_0, \eta}(X^*).
\]

By Theorem 4 one can decompose $X$ into independent $X_1^* \sim \mathcal{N}(0, K_1)$ and $X_2^* \sim \mathcal{N}(0, K_2)$, with $K_1 + K_2 \leq K^*$, such that
\[
T^q_{\lambda_0, \eta}(X^*) = T^q_{\lambda_0, \eta}(X_1^* + X_2^*) = \tilde{V}^q_{\lambda_0, \eta}(K^*),
\]
and
\[
S^q(\tilde{X}_1^* + \tilde{X}_2^*) = S^q(X_1^*) = V^q(K_1 + K_2).
\]

The proof of existence is concluded by setting $X_1^* \sim \mathcal{N}(0, K_1)$, where $K_3 = K - (K_1 + K_2)$ and noting that $X = X_1^* + X_2^* + X_3^* \sim \mathcal{N}(0, K)$ and $X^* = X_1^* + X_2^* + X_3^* \sim \mathcal{N}(0, K^*)$, which implies that (7) holds. The uniqueness of the decomposition (i.e., of the covariance matrices $K_1, K_2$ and $K_3$) is a direct consequence of Theorems 3, 4 and 5.

VII. SUMMARY AND CONCLUDING REMARKS

The two-user MIMO Gaussian BC with common, private and confidential messages was studied. The private message to Receiver 1 is confidential and kept secret from Receiver 2. The secrecy-capacity region without a common message was characterized first and Gaussian inputs were shown to achieve optimality. The proof relied on establishing an equivalence between certain inner and outer bounds using factorization of UCEs 16 and a variant of DPC 20. Our results showed that using DPC to cancel out the signal of the non-confidential message at Receiver 1 exhausts the entire region, making DPC against the signal of the confidential message unnecessary.

This secrecy-capacity region without a common message was then used to characterize a portion of the region with a common message. The rest of the region was found using double-nested UCEs. The secrecy-capacity region without a common message was illustrated using a numerical example. To make the region (efficiently) computable, matrix decomposition properties from 22 were leveraged. The region was shown to be strictly larger than the secrecy-capacity region of the MIMO Gaussian BC with confidential messages (in which each private message is kept secret from the opposite user).

ACKNOWLEDGEMENTS

The authors would like to thank Anatoly Khina for insightful discussions about DPC and his advice regrading implementation of the numerical example. We also thank the Associate Editor and the anonymous reviewers for their suggestions and comments that helped us improve the presentation of this paper.

APPENDIX A

DERIVATION OF INNER AND OUTER BOUNDS

A. Outer Bounds 4 and 3

We first establish Bound 4 as an outer bound on the secrecy-capacity region of the setting with a common message, and then use it to establish Bound 3 as an outer bound on the region without a common message.

The result of 23 (Theorem 3) characterizes an outer bound $R_0(L_1, L_2)$ on the $(L_1, L_2)$-leakage-capacity region of a DM-BE with common and private messages, for some leakage thresholds $(L_1, L_2) \in \mathbb{R}_+$. Setting $L_1 = 0$ and letting
as the legitimate receiver and User 2 as the eavesdropper. Therefore (see [7]–[9]),
\[ H_1^K = \frac{1}{2} \max_{0 \leq K \leq K} \log |I + G_1 K^* G_1^T| - \log |I + G_2 K^* G_2^T| . \]

To see that \((H_1^K, 0)\) is also in \(\hat{I}_K\) consider the following. For every \(0 \leq K \leq K\) let \(X_1\) and \(X_2\) be independent Gaussian random vectors with covariances \(K^*\) and \(K - K^*\), respectively. Set
\[ U = X_1 + AX_2, \]
\[ V = X_2, \]
\[ X = X_1 + X_2, \]
where \(A = K^* G_1^T [I + G_1 K^* G_1^T]^{-1}\) is the precoding matrix for suppressing \(V\) from \(Y_1\) [20] Theorem 1]. Evaluating the mutual information terms on the RHS of (93a), we first have
\[ I(U; Y_1) - I(U; V) = I(X; Y_1 | V) = \frac{1}{2} \log |I + G_1 K^* G_1^T| . \]

Moreover,
\[ I(U; Y_2 | V) = I(X_1 + AX_2; G_2(X_1 + X_2) + Z_2 | X_2) = I(X_1; G_2 X_1 + Z_2 | X_2) \]
\[ = \frac{1}{2} \log |I + G_2 K^* G_2^T|, \]
where the equality in (a) follows because \((X_1, Z_2)\) and \(X_2\) are independent. Combining (91) with (92) yields
\[ I(U; Y_1) - I(U; V) - I(U; Y_2 | V) = \frac{1}{2} \log |I + G_1 K^* G_1^T| - \frac{1}{2} \log |I + G_2 K^* G_2^T|, \]
which implies that \((H_1^K, 0)\) is in \(\hat{I}_K\). Furthermore, since \((H_1^K, 0)\) is on the boundary of \(\hat{I}_K\) and \(\hat{I}_K \subseteq \hat{I}_K\), \((H_1^K, 0)\) must also be a boundary point of \(\hat{I}_K\) and therefore, \(R_1 \leq H_1^K\) is a supporting hyperplane of \(\hat{I}_K\).

\section*{Appendix C}
\textbf{Proof of Auxiliary Results for Proposition 6: Lemmas 2 and 3}

\subsection*{A. Proof of Lemma 2}
By Theorem 4 we have that if \(E[XX^T] \preceq K\), then
\[ \hat{V}_q^{\eta}(K) = t_{\lambda, \eta}^q(X^*), \]
where \(X^* = X_1^* + X_2^*\), and \(X_1^*\) and \(X_2^*\) are independent random variables with \(X_j^* \sim \mathcal{N}(0, K_j)\), for \(j = 1, 2\), such that \(K^* \triangleq K_1 + K_2 \preceq K\). Furthermore, by the definition of the UCE and the definition of \(V_q^{\eta}(K)\), \(24\) implies that for every \(X \sim P_X\) with \(E[XX^T] \preceq K\), we have
\[ t_{\lambda, \eta}^{q}(X) \leq T_{\lambda, \eta}^q(X) \leq \hat{V}_q^{\eta}(K) = t_{\lambda, \eta}^q(X^*). \]

Thus,
\[ \sup_{X: E[XX^T] \preceq K} t_{\lambda, \eta}^{q}(X) \]
\[ \leq \lambda_1 I(X^*; Y_1) - (\lambda_1 + \lambda_2) I(X^*; Y_2) + \lambda_1 s_\eta^2(X^*) \]

(a) \[ \lambda_1 I(X^*; Y_1) - (\lambda_1 + \lambda_2) I(X^*; Y_2) + \lambda_1 s_\eta^2(X^*) \]

(b) \[ \lambda_1 I(X^*; Y_1) - (\lambda_1 + \lambda_2) I(X^*; Y_2) + \lambda_1 C_\eta, \quad (96) \]

where (a) is by Theorem 4, while (b) follows from an adaptation of [16, Proposition 19] to the function \( s_\eta^2(X) \) as defined in [17], which implies that for \( \eta > 1 \) there exists a \( C_\eta \), such that \( s_\eta^2(X) \leq C_\eta \), for all \( P_X \) (see Remark 8). By (96), we have

\[ \sup_{X \sim \mathcal{N}(0, K)} t_{\lambda, \eta}^2(X) \leq \sup_{0 \leq \eta \leq 1} \lambda_1 I(X; Y_1) - (\lambda_1 + \lambda_2) I(X; Y_2) + \lambda_1 C_\eta. \]

(97)

Let \( \Sigma_j = (G_j^T G_j)^{-1} \), \( j = 1, 2 \). For \( X \sim \mathcal{N}(0, K) \), we write

\[ 2\lambda_1 I(X; Y_1) - 2(\lambda_1 + \lambda_2) I(X; Y_2) = \lambda_1 \log(1 + G_1 K G_1^T) - (\lambda_1 + \lambda_2) \log(1 + G_2 K G_2^T) \]

\[ = \lambda_1 \log(1 + K G_1 G_1^T) - (\lambda_1 + \lambda_2) \log(1 + K G_2 G_2^T) \]

\[ = -\lambda_1 \log(\Sigma_1) + (\lambda_1 + \lambda_2) \log(\Sigma_2) \]

\[ \leq -\lambda_1 \log(\Sigma_1) + (\lambda_1 + \lambda_2) \log(\Sigma_2), \quad (98) \]

where (a) is by setting \( \lambda \triangleq \frac{\lambda_1 + \lambda_2}{\lambda_1} > 1 \). To bound the last two terms, we use the min-max theorem on eigenvalues: Let \( \mu_i(A) \) be the \( i \)th smallest eigenvalue of the symmetric matrix \( A \in \mathbb{R}^{t \times t} \), we have

\[ \mu_i(A) = \min_{L_i} \max_{u \neq 0 \in L_i} \frac{u^T A u}{u^T u}, \]

\[ \mu_i(A) = \max_{L_i} \min_{u \neq 0 \in L_i} \frac{u^T A u}{u^T u}, \]

where \( L_i \) is an \( i \)-dimensional subspace of \( \mathbb{R}^t \). Since the \( i \)-dimensional subspace of \( \mathbb{R}^t \) is unique (that is, \( L_i = \mathbb{R}^t \)), we obtain

\[ \mu_1(A) = \min_{L_i} \max_{u \neq 0 \in L_i} \frac{u^T A u}{u^T u} = \min_{u \neq 0} \frac{u^T A u}{u^T u}, \]

\[ \mu_t(A) = \max_{L_i} \min_{u \neq 0 \in L_i} \frac{u^T A u}{u^T u} = \max_{u \neq 0} \frac{u^T A u}{u^T u}, \]

The RHSs of (100a) imply that for every non-zero \( u \in \mathbb{R}^t \) we have \( \mu_1(A) \leq \frac{u^T A u}{u^T u} \leq \mu_t(A) \). We upper and lower bound the \( i \)th eigenvalue of \( K + \Sigma_j \), for \( j = 1, 2 \), as follows

\[ \mu_i(K + \Sigma_j) = \min_{L_i} \max_{u \neq 0 \in L_i} \left( \frac{u^T K u}{u^T u} + \frac{u^T \Sigma_j u}{u^T u} \right) \]

\[ \geq \min_{L_i} \max_{u \neq 0 \in L_i} \left( \frac{u^T K u}{u^T u} + \mu_i(\Sigma_j) \right) = \mu_i(K) + \mu_i(\Sigma_j), \]

\[ \leq \min_{L_i} \max_{u \neq 0 \in L_i} \left( \frac{u^T K u}{u^T u} + \mu_t(\Sigma_j) \right) = \mu_i(K) + \mu_t(\Sigma_j). \]

\[ (101) \]

Hence the eigenvalues of \( K + \Sigma_j \), \( j = 1, 2 \), satisfy

\[ \mu_i(K) + \mu_1(\Sigma_j) \leq \mu_i(K + \Sigma_j) \leq \mu_i(K) + \mu_t(\Sigma_j), \]

(102)

where \( i \in [t] \). We now bound the last two terms in (98) as

\[ \log(\Sigma_1 + K) \leq \log(\Sigma_1 + K) \]

\[ = \sum_{i=1}^{t} \log \left( \frac{\mu_i(K + \Sigma_1)}{\mu_i(\Sigma_1)} \right) \]

\[ \leq \sum_{i=1}^{t} \log \left( \frac{\mu_i(K) + \mu_i(\Sigma_1)}{\mu_i(\Sigma_1)} \right) \]

\[ \leq t \cdot \max_i \log \left( \frac{\mu_i(K) + \mu_i(\Sigma_1)}{\mu_i(\Sigma_1)} \right) \]

\[ \leq t \cdot \log \left( \frac{\mu^* + \mu_1(\Sigma_1)}{(\mu^* + \mu_1(\Sigma_2))^\frac{1}{\lambda_1}} \right), \quad (103) \]

where (a) follows from by (101), and (b) is by setting \( \mu^* = \max \left\{ 0, \frac{1}{\lambda_1}(\mu_1(\Sigma_2) - \mu_1(\Sigma_1)) \right\} \) and noting that \( \mu_i(K) \geq 0 \) and that the derivative of the function \( c(x) \triangleq \log(\mu_i(K + \Sigma_1) - \mu_i(\Sigma_1)) \) is zero at \( x = \mu^* \), negative when \( x > \mu^* \) and positive when \( x < \mu^* \). Since \( \mu_i(\Sigma_1), \mu_i(\Sigma_2) > 0 \) (which holds since the positive semi-definite matrices \( \Sigma_1 \) and \( \Sigma_2 \) are invertible), we conclude that for every \( X \sim P_X \)

\[ t_{\lambda, \eta}^2(X) \]

\[ \leq -\lambda_1 \log(\Sigma_1) + (\lambda_1 + \lambda_2) \log(\Sigma_2) \]

\[ + \mu_1(\Sigma_1) \cdot t \cdot \log \left( \frac{\mu^* + \mu_1(\Sigma_1)}{(\mu^* + \mu_1(\Sigma_2))^\frac{1}{\lambda_1}} \right) \]

\[ \triangleq B_{\lambda, \eta} < \infty. \quad (104) \]

B. Proof of Lemma 3

The proof of Lemma 3 follows immediately by an adaptation of [16, Proposition 20] (see Remark 8) and by [16, Theorem 5], which, respectively, imply that \( S_\eta^2(X_n) \xrightarrow{n \to \infty} S_\eta^2(X^*) \), and that \( h(Y_{j,n}) \xrightarrow{n \to \infty} h(Y_j^*) \), for \( j = 1, 2 \). Here \( Y_{1,n} \) and \( Y_{2,n} \) are the outputs of the MIMO Gaussian BC with input \( X_n \).

C. Proof of Lemma 4

Let \( i \in [m] \) be such that \( \alpha_i^* = 0 \) and note that for every \( \bar{K} \geq 0 \), we have

\[ \varphi_{\lambda, \alpha, \eta}^2(\bar{K}) \triangleq \min_{\lambda_m = 1} \max_{\lambda_m \neq 0} \frac{\lambda_m}{\lambda_m} \]

\[ \leq \frac{\lambda_2}{2} \log(1 + G_2 K G_2^T) + B_{\lambda, \alpha, \eta} \]

\[ \leq \frac{\lambda_2}{2} \log(1 + G_2 K G_2^T) + B_{\lambda, \alpha, \eta}, \]

(105)

where (a) follows from (75) and the non-negativity of mutual information, while (b) is since \( T_{\lambda, \alpha, \eta}(X) \) is bounded. Let \( \{\alpha_i^{(n_k)}\}_{k \in \mathbb{N}} \) be a subsequence of \( \{\alpha_i^{(n)}\}_{n \in \mathbb{N}} \), such that \( \alpha_i^{(n_k)} > 0 \) for every \( k \in \mathbb{N} \) (if there is no such subsequence, the result of Lemma 3 is immediate). Since \( \alpha_i^{(n_k)} K_i^{(n_k)} \leq \bar{K} \) and \( \alpha_i^{(n_k)} > 0 \), we obtain

\[ K_i^{(n_k)} \leq \frac{1}{\alpha_i^{(n_k)}} \bar{K}, \quad \forall k \in \mathbb{N}. \]

(106)
We thus conclude that
\[
\alpha_i^{(n_k)} \alpha_{\lambda_0, \alpha, \eta}(K_i^{(n_k)}) \\
\leq \alpha_i^{(n_k)} \left( \frac{2}{\lambda_2} \log \left| I + G_2 K_i^{(n_k)} G_2^T \right| + \bar{B}_{\lambda_0, \alpha, \eta} \right) \\
\leq \alpha_i^{(n_k)} \left( \frac{\lambda_2}{2} \log \left| I + G_2 K_i^{(n_k)} G_2^T \right| + \bar{B}_{\lambda_0, \alpha, \eta} \right) \\
\leq \alpha_i^{(n_k)} \left( \frac{\lambda_2}{2} \sum_{j=1}^t \log \left( 1 + \frac{\lambda_j}{\alpha_i^{(n_k)}} \right) + \bar{B}_{\lambda_0, \alpha, \eta} \right) \xrightarrow{n \to \infty} 0,
\]
(107)
where (a) and (b) follow from (105) and (106), respectively, while (c) follows by denoting the eigenvalues of $G_2 K_i G_2^T$ by $\{\mu_j\}_{j=1}^t$.

**APPENDIX D**

**PROOF OF PROPOSITION 9**

The proof relies on a result from [27 Corollary 2], which we reproduce in the following.

**Proposition 11 (Min-Max Interchange)** Let $\Lambda_d = \{ \lambda \in \mathbb{R}_+^d, \sum_{i=1}^d \lambda_i = 1 \}$ be the $d$-dimensional simplex. Let $\mathcal{P}$ be a set of distribution $P_U$ over a set $\mathcal{U}$. Let $\{g_i : \mathcal{P} \to \mathbb{R}\}_{i \in [d]}$ be a set of functions such that
\[
\mathcal{A} \triangleq \left\{ a \in \mathbb{R}^d : \forall i \in [d], \exists P_U \in \mathcal{P}, \ a_i \leq g_i(P_U) \right\}
\]
(108)
is a convex set. Then
\[
\sup_{P_U \in \mathcal{P}} \min_{\lambda \in \Lambda_d} \sum_{i=1}^d \lambda_i g_i(P_U) = \min_{\lambda \in \Lambda_d} \sup_{P_U \in \mathcal{P}} \sum_{i=1}^d \lambda_i g_i(P_U).
\]
(109)

Let $d = 2$ and $\mathcal{P}$ be the set of PDFs $P_{W,V,U,X}$ that satisfy $\mathbb{E}[XX^T] \preceq K$. Set
\[
g_1(P_{W,V,U,X}) = \lambda_0 I(W; Y_1) + \lambda_2 I(V; Y_2 | W) + \lambda_1 \left[ I(U; Y_1 | W) - I(U; V | W) - I(U; Y_2 | W, V) \right]
\]
(110a)
\[
g_2(P_{W,V,U,X}) = \lambda_0 I(W; Y_2) + \lambda_2 I(V; Y_2 | W) + \lambda_1 \left[ I(U; Y_1 | W) - I(U; V | W) - I(U; Y_2 | W, V) \right],
\]
(110b)
and consider the corresponding set $\mathcal{A}$ from (108). To show that $\mathcal{A}$ is convex, let $(a_1, a_2), (b_1, b_2) \in \mathcal{A}$ and $P_a, P_b \in \mathcal{P}$ be two distributions, such that $a_i \leq g_i(P_a)$ and $b_i \leq g_i(P_b)$, for $i = 1, 2$. Fix $\alpha \in [0, 1]$ and consider a distribution $P$ given by
\[
P(w, v, u, x) = \alpha P_a(w, v, u, x) + \bar{\alpha} P_b(w, v, u, x),
\]
(111)
for all $(w, v, u, x) \in \mathcal{W} \times \mathcal{V} \times \mathcal{U} \times \mathcal{X}^n$. Equivalently, $P$ can be represented by setting $\bar{W} = (Q, W)$, where $Q \sim \text{Ber}(\alpha)$, and denoting $P \triangleq P_{W,V,U,X} = P_{(Q,W),V,U,X}$, for which
\[
P_{(Q,W),V,U,X}((0, w), v, u, x) = \alpha P_a(w, v, u, x)
\]
(112a)
\[
P_{(Q,W),V,U,X}((1, w), v, u, x) = \bar{\alpha} P_b(w, v, u, x),
\]
(112b)
for all $(w, v, u, x) \in \mathcal{W} \times \mathcal{V} \times \mathcal{U} \times \mathcal{X}^n$. First note that
\[
\mathbb{E}_P[XX^T] = \alpha \mathbb{E}_{P_a}[XX^T] + \bar{\alpha} \mathbb{E}_{P_b}[XX^T] \preceq K,
\]
(113)
where $\mathbb{E}_Q$ denotes that an expectation is taken with respect to $Q$. This implies that $P \in \mathcal{P}$.

Next, by evaluating $g_i$, $i = 1, 2$, with respect to $P$, we have
\[
g_i(P) = \lambda_0 I_P(\bar{W}; Y_i) + \lambda_1 \left[ I_P(U; Y_1 | \bar{W}) - I_P(U; V | \bar{W}) \right]
\]
\[
- I_P(U; Y_2 | \bar{W}, V)
\]
\[
+ \lambda_2 I_P(V; Y_2 | \bar{W})
\]
\[
= \lambda_0 I_P(Q; Y_i) + \alpha g_i(P_a) + \bar{\alpha} g_i(P_b)
\]
\[
\geq \alpha a_i + \bar{\alpha} b_i,
\]
(114)
implying that $\alpha(a_1, a_2) + \bar{\alpha}(b_1, b_2) \in \mathcal{A}$, which establishes the convexity of $\mathcal{A}$. In the above, $I_P$ indicates that a mutual information term is taken with respect to an underlying distribution $P$. The proof of Proposition 9 is completed by invoking Proposition 11 while noting that for every tuple of random variables $(W, V, U, X)$, with $(W, V, U) - X - (Y_1, Y_2)$ and $\mathbb{E}[XX^T] \preceq K$, the minimum
\[
\min_{\alpha \in [0, 1]} \left\{ \lambda_0 \left[ \alpha I(W; Y_1) + \bar{\alpha} I(W; Y_2) \right]
\right.
\]
\[
+ \lambda_1 \left[ I(U; Y_1 | W) - I(U; V | W) - I(U; Y_2 | W, V) \right]
\]
\[
\left. + \lambda_2 I(V; Y_2 | W) \right\}
\]
is attained by either $\alpha = 0$ or $\alpha = 1$. 

First note that
\[
\mathbb{E}_P[XX^T] = \alpha \mathbb{E}_{P_a}[XX^T] + \bar{\alpha} \mathbb{E}_{P_b}[XX^T] \preceq K,
\]
(113)
where $\mathbb{E}_Q$ denotes that an expectation is taken with respect to $Q$. This implies that $P \in \mathcal{P}$.

Next, by evaluating $g_i$, $i = 1, 2$, with respect to $P$, we have
\[
g_i(P) = \lambda_0 I_P(\bar{W}; Y_i) + \lambda_1 \left[ I_P(U; Y_1 | \bar{W}) - I_P(U; V | \bar{W}) \right]
\]
\[
- I_P(U; Y_2 | \bar{W}, V)
\]
\[
+ \lambda_2 I_P(V; Y_2 | \bar{W})
\]
\[
= \lambda_0 I_P(Q; Y_i) + \alpha g_i(P_a) + \bar{\alpha} g_i(P_b)
\]
\[
\geq \alpha a_i + \bar{\alpha} b_i,
\]
(114)
implying that $\alpha(a_1, a_2) + \bar{\alpha}(b_1, b_2) \in \mathcal{A}$, which establishes the convexity of $\mathcal{A}$. In the above, $I_P$ indicates that a mutual information term is taken with respect to an underlying distribution $P$. The proof of Proposition 9 is completed by invoking Proposition 11 while noting that for every tuple of random variables $(W, V, U, X)$, with $(W, V, U) - X - (Y_1, Y_2)$ and $\mathbb{E}[XX^T] \preceq K$, the minimum
\[
\min_{\alpha \in [0, 1]} \left\{ \lambda_0 \left[ \alpha I(W; Y_1) + \bar{\alpha} I(W; Y_2) \right]
\right.
\]
\[
+ \lambda_1 \left[ I(U; Y_1 | W) - I(U; V | W) - I(U; Y_2 | W, V) \right]
\]
\[
\left. + \lambda_2 I(V; Y_2 | W) \right\}
\]
is attained by either $\alpha = 0$ or $\alpha = 1$. 


