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 Gaussian broadcast channel with common, private, and confidential messages is considered. The transmitter sends a common message to both users, a confidential message to the User 1 and a private (non-confidential) message to the 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 the 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.

Fig. 1.   MIMO Gaussian BC.

I. 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
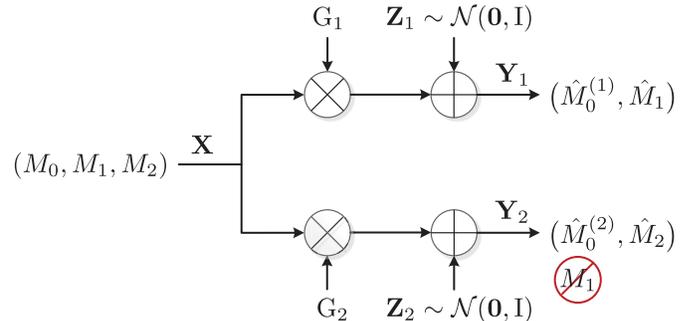
scenarios fall within this framework. One such example is a banking site that simultaneously: (i) broadcasts an advertisement to all online users (modeled by M_0); (ii) offers public information (such as material on different banking programs, reports, forecasts, etc.) that is available only to users that are interested in it (modeled by the private message M_2); and (iii) provides an online banking service, by which users can access their account and perform transactions (this confidential information is modeled by M_1). Furthermore, 5th generation (5G) mobile technology [1] puts significant emphasis on advanced MIMO capabilities and multiuser communication. In particular, schemes supporting multiple users exchanging various kinds of information over a MIMO communication system (and their fundamental limits) are of great interest. The studied MIMO Gaussian BC is an instance of a system where these aspects are jointly incorporated.

In recent years, information-theoretic security over MIMO communication systems has been an active field of research (see [2] for a recent survey of progress in this area). Most noticeably, the secrecy-capacity of the Gaussian wiretap channel (WTC) was characterized in [3]–[5] for the multiple-input single-output scenario, and in [6]–[10] for the MIMO case. The Gaussian MIMO WTC with a common message was studied in [11]. In [12], the secrecy-capacity region for the setting with a degraded message set and an external eavesdropper (from which all messages are concealed) was derived. The MIMO Gaussian BC with confidential messages, in which the private message 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]

TABLE I

MIMO Gaussian BCs With/Without a Common Message and Where None/Some/All of
the Private Messages Are Secret - Summary of Results

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	Ekrem-Ulukus 2012 [14]
—	Secret from User 2	Private	This work
Public	Secret from User 2	Private	This work

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. The limitation of the channel enhancement technique seems to be the difficulty in generalizing it to account for both private and common messages. Attempted adaptations of this technique to scenarios comprising common and private messages include, e.g., [18] and [19], where generally unmatching inner and outer bounds or constant gap-from-capacity results were derived for the MIMO Gaussian BC without and with security requirements, respectively.

Our goal is to fully characterize the secrecy-capacity region of the MIMO Gaussian BC with common, private and confidential messages and show that it is attained by Gaussian inputs. Since channel enhancement arguments are insufficient for this purpose, we adopt the approach of [16] for proving the optimality of Gaussians via factorization of upper concave envelopes (UCEs). We start by characterizing the secrecy-capacity region under an input covariance constraint for the setting with private and confidential messages only (i.e., when no common message is present). The derivation first describes the boundary points of a certain outer bound on the secrecy-capacity region as an UCE of a function of the input distribution. With this result at hand, we show that if this UCE satisfies a specific factorization property, then it is maximized by a Gaussian input distribution. Then, using an

adaptation of dirty-paper coding (DPC) [20], we establish the equivalence of the outer bound to a particular inner bound, thus characterizing the secrecy-capacity region. Interestingly, optimality 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, turns out to be unnecessary. This is in contrast to the case without secrecy requirements [16], in which both variants of DPC are necessary to exhaust the entire region.

We then focus on the MIMO Gaussian BC with common, private and confidential messages (Fig. 1) and derive our main result by characterizing its secrecy-capacity region. Although this is a generalization of the problem without a common message, the secrecy-capacity of the latter setting is solved first. In doing so, we use the result without a common message to show that Gaussian inputs are optimal for a certain portion of the region with a common message. The rest of the region is characterized by extending the tools from [16] and introducing the notion of a double-nested UCE. Gaussian inputs once again are shown to attain optimality. Finally, we visualize our results by a numerical example. Since the obtained regions are described as non-convex matrix optimization problems, we convert them into a computationally efficient form by relying on matrix decomposition properties from [21].

Organization: The UCE factorization method introduced in [16] and further developed in this work relies on rather heavy machinery and many technical functional analysis results. The proofs of our main secrecy-capacity theorems (i.e., without and with a common message) frequently refer to these auxiliary results while also relying on additional information-theoretic arguments (e.g., DPC). In structuring this paper, it was important for us to distill the information-theoretic arguments from the functional analysis aspects of this work for two main reasons. First, this would ease the flow trough the information-theoretic proofs and highlight the usage of the UCE factorization method for showing that Gaussian inputs achieve capacity. Second, by aggregating the machinery behind this method in a separate section we hope to facilitate its application to additional research problems in future work.

In accordance to the above, the rest of the paper is organized as follows. Section II gives definitions and describes the

MIMO Gaussian BC with common, private and confidential messages. In Section III we state our main results: the secrecy-capacity regions of the considered BC without and with 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 $\{n \in \mathbb{N} | \lceil a \rceil \leq n \leq \lfloor b \rfloor\}$; 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., 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 $p_{X,Y} = p_X p_{Y|X}$ are to be understood pointwise, i.e., $p_{X,Y}(x, y) = p_X(x) p_{Y|X}(y|x)$, for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$. Accordingly, when three random variables X, Y and Z satisfy $p_{X|Y,Z} = p_{X|Y}$, they form a Markov chain, which we denote by $X - Y - Z$. 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 $\mathbb{E}X$. When a random variable X is normally distributed we write $X \sim \mathcal{N}(\mu, \sigma^2)$, where $\mu = \mathbb{E}X$ 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 $\boldsymbol{\mu} = \mathbb{E}\mathbf{X} \in \mathbb{R}^n$ and the covariance matrix $\mathrm{K} = \mathbb{E}[(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})^\top]$, for which we write $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \mathrm{K})$. Generally, non-italic capital letters, e.g., A, denote matrices. We use $\mathrm{A} \succeq 0$ to indicate that a matrix A is positive semi-definite, while $\mathrm{A} \preceq \mathrm{B}$ denotes "less than or equal to" in the positive semi-definite ordering, i.e., $\mathrm{B} - \mathrm{A} \succeq 0$. The determinant of a square matrix A is designated by $|\mathrm{A}|$.

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

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

B. Problem Definition

The outputs of a MIMO Gaussian BC at the ith channel use are:
$$\mathbf{Y}_j(i) = \mathrm{G}_j \mathbf{X}(i) + \mathbf{Z}_j(i), \quad j = 1, 2, \quad i \in [n], \qquad (1)$$

where $\mathrm{G}_1, \mathrm{G}_2 \in \mathbb{R}^{t \times t}$ are channel gain matrices (assumed to be known to all parties), $\{\mathbf{Z}_j(i)\}_{i \in [n]}$, 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 $\mathbf{Z}_j(i) = \left[Z_{j,1}(i) \ Z_{j,2}(i) \ \ldots \ Z_{j,t}(i)\right]^\top$ are also i.i.d. Gaussian random variables, whose expected values and variance are specified by the parameters of the normal distribution of $\mathbf{Z}_j(i)$. The input sequence $\{\mathbf{X}(i)\}_{i \in [n]}$ is subject to the covariance constraint
$$\frac{1}{n} \sum_{i=1}^{n} \mathbb{E}\left[\mathbf{X}(i)\mathbf{X}(i)^\top\right] \preceq \mathrm{K}, \qquad (2)$$

where $\mathrm{K} \succeq 0$.

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, Sec. 5]).*

2) *For each $j = 1, 2$, the Gaussian noise vectors, $\{\mathbf{Z}_j(i)\}_{i \in [n]}$, are i.i.d. according to $\mathcal{N}(\mathbf{0}, \mathrm{I})$, 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 (1) 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, which is formally described by the weak-secrecy requirement

$$\frac{1}{n} I(M_1; \mathbf{Y}_2^n) \xrightarrow[n \to \infty]{} 0, \qquad (3)$$

where n is the number of channel uses. In (3), the notation $\mathbf{Y}_2^n \triangleq \begin{bmatrix} \mathbf{Y}_2(1) & \mathbf{Y}_2(2) & \dots & \mathbf{Y}_2(n) \end{bmatrix}^\top$ is used, where for each $i \in [n]$, $\mathbf{Y}_2(i)$ is the output vector (taking values in $\mathbb{R}^t$) observed by User 2 at the ith channel instance. For any covariance constraint $\mathrm{K} \succeq 0$, the secrecy-capacity region $\mathcal{C}_\mathrm{K}$ 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.[1] 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 $\mathrm{K} \succeq 0$, let $\hat{\mathcal{C}}_\mathrm{K}$ be the corresponding secrecy-capacity region, and for any $0 \preceq \mathrm{K}^\star \preceq \mathrm{K}$ set the following shorthand notations:

$$\hat{r}_1(\mathrm{K}^\star) \triangleq \frac{1}{2} \log \frac{\left| \mathrm{I} + \mathrm{G}_1 \mathrm{K}^\star \mathrm{G}_1^\top \right|}{\left| \mathrm{I} + \mathrm{G}_2 \mathrm{K}^\star \mathrm{G}_2^\top \right|} \qquad (4a)$$

$$\hat{r}_2(\mathrm{K}^\star) \triangleq \frac{1}{2} \log \frac{\left| \mathrm{I} + \mathrm{G}_2 \mathrm{K} \mathrm{G}_2^\top \right|}{\left| \mathrm{I} + \mathrm{G}_2 \mathrm{K}^\star \mathrm{G}_2^\top \right|}. \qquad (4b)$$

---

[1] The work [23] derived bounds on the admissible rate region of a BC with privacy leakage constraints. Zero leakage corresponds to secrecy, but since leakage can only be defined in terms of rate the resulting notion of security is weak-secrecy.

Define also

$$\hat{\mathcal{C}}_\mathrm{K}(\mathrm{K}^\star) \triangleq \left\{ (R_1, R_2) \in \mathbb{R}_+^2 \,\middle|\, \begin{matrix} R_1 \leq \hat{r}_1(\mathrm{K}^\star) \\ R_2 \leq \hat{r}_2(\mathrm{K}^\star) \end{matrix} \right\}. \qquad (5)$$

The following theorem characterizes $\hat{\mathcal{C}}_\mathrm{K}$.

Theorem 1 (Secrecy-Capacity Without Common Message): The secrecy-capacity region $\hat{\mathcal{C}}_\mathrm{K}$ of the MIMO Gaussian BC with private and confidential messages under the covariance constraint (2) is

$$\hat{\mathcal{C}}_\mathrm{K} = \bigcup_{0 \preceq \mathrm{K}^\star \preceq \mathrm{K}} \hat{\mathcal{C}}_\mathrm{K}(\mathrm{K}^\star). \qquad (6)$$

The proof of Theorem 1 (given in Section V-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. The secrecy-capacity region is then characterized by evaluating the boundary points of the inner bound under a Gaussian input vector and showing that they coincide with those of the outer bound.

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 $\mathrm{K}^\star$, 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 $\mathrm{K} - \mathrm{K}^\star$ and noise covariance $\mathrm{I} + \mathrm{K}^\star$. 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 $\mathrm{K}^\star$ 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 V-A), the entire secrecy-capacity region $\hat{\mathcal{C}}_\mathrm{K}$ 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 $\mathcal{C}_\mathrm{K}$ (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 $\mathcal{C}_\mathrm{K}$. More specifically, when the private message rate R_2 is larger than the common message rate R_0, the optimizing distribution of $\mathcal{C}_\mathrm{K}$ coincides with that of $\hat{\mathcal{C}}_\mathrm{K}$; 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

$$\frac{1}{n} \sum_{i=1}^{n} \|\mathbf{X}(i)\|^2 \leq P, \qquad (7a)$$

is given by

$$\hat{\mathcal{C}}_P = \bigcup_{0 \preceq K:\ \mathrm{tr}(K) \leq P} \hat{\mathcal{C}}_K. \qquad (7b)$$

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 (6) 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 $\hat{\mathcal{C}}_P$ on a numerical example.

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

To state the secrecy-capacity region $\mathcal{C}_K$ of the MIMO Gaussian BC with common, private and confidential messages as defined in Section II-B, we define

$$r_0^{(j)}(K_1, K_2) \triangleq \frac{1}{2} \log \left| \frac{I + G_j K G_j^\top}{I + G_j(K_1 + K_2)G_j^\top} \right|, \quad j = 1, 2 \quad (8a)$$

$$r_1(K_2) \triangleq \frac{1}{2} \log \left| \frac{I + G_1 K_2 G_1^\top}{I + G_2 K_2 G_2^\top} \right| \qquad (8b)$$

$$r_2(K_1, K_2) \triangleq \frac{1}{2} \log \left| \frac{I + G_2(K_1 + K_2)G_2^\top}{I + G_2 K_2 G_2^\top} \right|, \qquad (8c)$$

and $r_0(K_1, K_2) \triangleq \min\left\{ r_0^{(1)}(K_1, K_2), r_0^{(2)}(K_1, K_2) \right\}$, and set

$$\mathcal{C}_K(K_1, K_2) \triangleq \left\{ (R_0, R_1, R_2) \in \mathbb{R}_+^3 \left| \begin{array}{l} R_0 \leq r_0(K_1, K_2) \\ R_1 \leq r_1(K_2) \\ R_2 \leq r_2(K_1, K_2) \end{array} \right. \right\}. \quad (9)$$

Theorem 2 (Secrecy-Capacity With Common Message): The secrecy-capacity region $\mathcal{C}_K$ of the MIMO Gaussian BC with common, private and confidential messages under the covariance constraint (2) is

$$\mathcal{C}_K = \bigcup_{\substack{0 \preceq K_1, K_2: \\ K_1 + K_2 \preceq K}} \mathcal{C}_K(K_1, K_2). \qquad (10)$$

Theorem 2 is proven in Section V-B.

Remark 7 (Interpretation of Optimal Secrecy Rates): Our interpretation of the structure of $\mathcal{C}_K$ 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 (8c)). 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 (7a) is given by

$$\mathcal{C}_P = \bigcup_{0 \preceq K:\ \mathrm{tr}(K) \leq P} \mathcal{C}_K. \qquad (11)$$

C. Numerical Example

We illustrate the secrecy-capacity region $\hat{\mathcal{C}}_P$ 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 $\hat{\mathcal{C}}_K$, each under a covariance constraint $K \succeq 0$ with $\mathrm{tr}(K) \leq P$. However, $\hat{\mathcal{C}}_K$ 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 $\hat{\mathcal{C}}_K$ by leveraging the decomposition proposed in [21, eq. (10)]: Every positive semi-definite matrix $K^\star \in \mathbb{R}^{t \times t}$ with $K^\star \preceq K$ can be expressed as

$$K^\star = K^{\frac{1}{2}} V D V^\top K^{\frac{1}{2}^\top}, \qquad (12)$$

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

$$V = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}, \quad \theta \in [0, 2\pi]. \qquad (13)$$

Running over all possible diagonal matrices D involves only two parameters, viz. the diagonal entries of D. Finally, note that $K^{\frac{1}{2}}$ is any matrix B satisfying $BB^\top = K$. Obviously, there are many such matrices (in fact if B satisfies $BB^\top = K$, 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 semi-definite matrices K with $\mathrm{tr}(K) = P$. This observation follows because for every K' with $\mathrm{tr}(K') = \pi < P$, there is a K with $\mathrm{tr}(K) = P$, such that

$$\hat{\mathcal{C}}_{K'} \subseteq \hat{\mathcal{C}}_K. \qquad (14)$$
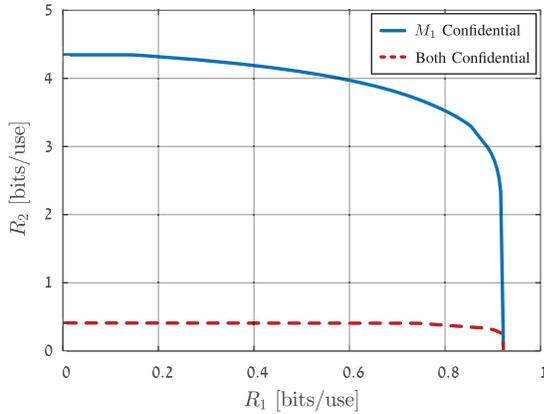
Fig. 2. Secrecy-capacity region under an average total power constraint of the MIMO Gaussian BC without a common message, where: M_1 is confidential and M_2 is private (solid blue) vs. M_1 and M_2 are both confidential (dashed red).

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 $K^\star \preceq K' \preceq K$ and replacing K′ with K in (4) does not alter (4a) and strictly increases (4b).

In the numerical example we set

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

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 of the message to the other user) 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, eq. (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, Secs. 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 is on a new function that was not considered [16, Sec. IV-C], 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, Sec. I-A], which is a special case of a product BC (PBC).[2]

Definition 2 (Product BC): A PBC consists of a sender $(\mathbf{X}_1, \mathbf{X}_2)$ *and two receivers* $(\mathbf{Y}_{11}, \mathbf{Y}_{12})$ *and* $(\mathbf{Y}_{21}, \mathbf{Y}_{22})$, *and is described by a conditional PDF of the form* $q^{(1)}_{\mathbf{Y}_{11}, \mathbf{Y}_{21}|\mathbf{X}_1} \times q^{(2)}_{\mathbf{Y}_{12}, \mathbf{Y}_{22}|\mathbf{X}_2}$.

A MIMO Gaussian PBC can be represented as

$$\begin{bmatrix} \mathbf{Y}_{11} \\ \mathbf{Y}_{12} \end{bmatrix} = \begin{bmatrix} G_{11} & 0 \\ 0 & G_{12} \end{bmatrix} \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_{11} \\ \mathbf{Z}_{12} \end{bmatrix} \quad (16a)$$

$$\begin{bmatrix} \mathbf{Y}_{21} \\ \mathbf{Y}_{22} \end{bmatrix} = \begin{bmatrix} G_{21} & 0 \\ 0 & G_{22} \end{bmatrix} \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_{21} \\ \mathbf{Z}_{22} \end{bmatrix}, \quad (16b)$$

where $\mathbf{Z}_{11}, \mathbf{Z}_{12}, \mathbf{Z}_{21}, \mathbf{Z}_{22} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ are i.i.d. and independent of $(\mathbf{X}_1, \mathbf{X}_2)$. A *two-letter* version of a BC is a PBC in which the components are identical, i.e., $q^{(1)}_{\mathbf{Y}_{11}, \mathbf{Y}_{21}|\mathbf{X}_1} = q^{(2)}_{\mathbf{Y}_{12}, \mathbf{Y}_{22}|\mathbf{X}_2}$. In all subsequent definitions and results, the input covariance constraining matrix $K \succeq 0$ (see (2)) stays fixed.

A. Difference of Mutual Information Terms

Consider a BC $q_{\mathbf{Y}_1, \mathbf{Y}_2|\mathbf{X}}$. For any $\eta > 1$, let s^q_η be a functional of $\mathbf{X} \sim P_\mathbf{X}$ defined by

$$\mathsf{s}^q_\eta(\mathbf{X}) \triangleq I(\mathbf{X}; \mathbf{Y}_2) - \eta I(\mathbf{X}; \mathbf{Y}_1). \quad (17)$$

Remark 8: The definition of $\mathsf{s}^q_\eta(\mathbf{X})$ *in (17) coincides with that of* $\mathsf{s}^q_\lambda(\mathbf{X}) \triangleq I(\mathbf{X}; \mathbf{Y}_1) - \lambda I(\mathbf{X}; \mathbf{Y}_2)$ *from [16, Sec. 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* $\mathsf{s}^q_\eta(\mathbf{X})$ *established in [16] without providing proofs. Additional attributes of* $\mathsf{s}^q_\eta(\mathbf{X})$ *that were not proven in [16] are rigorously derived.*

For a pair of random variables $(V, \mathbf{X})$ such that $V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$ forms a Markov chain, set

$$\mathsf{s}^q_\eta(\mathbf{X}|V) \triangleq I(\mathbf{X}; \mathbf{Y}_2|V) - \eta I(\mathbf{X}; \mathbf{Y}_1|V), \quad (18)$$

and define the *UCE* of $\mathsf{s}^q_\eta(\mathbf{X})$ as

$$\mathsf{S}^q_\eta(\mathbf{X}) \triangleq \left(\mathfrak{C}\mathsf{s}^q_\eta \right)(\mathbf{X}) = \sup_{\substack{P_{V|\mathbf{X}}: \\ V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)}} \mathsf{s}^q_\eta(\mathbf{X}|V). \quad (19)$$

The second equality in (19) follows directly from Definition 1. For any discrete random variable V we also set $\mathsf{S}^q_\eta(\mathbf{X}|V) \triangleq \sum_v P(v) \mathsf{S}^q_\eta(\mathbf{X}|V = v)$, and naturally extend this definition (merely an expectation) for an arbitrary V.

Proposition 1 (Concave Envelopes Properties): The UCE S^q_η *satisfies:*

1) *If* $V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$ *forms a Markov chain, then* $\mathsf{S}^q_\eta(\mathbf{X}|V) \leq \mathsf{S}^q_\eta(\mathbf{X})$.
2) *If* $W - V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$ *forms a Markov chain, then* $\mathsf{S}^q_\eta(\mathbf{X}|V, W) = \mathsf{S}^q_\eta(\mathbf{X}|V)$.
3) $\mathsf{S}^q_\eta(\mathbf{X})$ *is convex in* η *inside* $(0, 2)$, *for a fixed* $P_\mathbf{X}$, *and therefore it is continuous in* η *at* $\eta = 1$.[3]

---

[2]Henceforth, we omit the time index i.

[3]The 3rd property can be established by considering any bounded, open interval containing 1, and not necessarily $(0, 2)$.

The proof of Proposition 1 is given in Section VI-A.

Definition 3 (Maximized Concave Envelope): For any MIMO Gaussian BC $q_{\mathbf{Y}_1,\mathbf{Y}_2|\mathbf{X}}$, define

$$V_\eta^q(\mathrm{K}) \triangleq \sup_{\mathbf{X}: \, \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq \mathrm{K}} \mathsf{S}_\eta^q(\mathbf{X}) = \sup_{\substack{(V,\mathbf{X}): \, \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq \mathrm{K}, \\ V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)}} \mathsf{s}_\eta^q(\mathbf{X}|V).$$

(20)

We subsequently show that (20) is achieved by a Gaussian input distribution. Central to the proof is a certain factorization property of S_η^q. To formulate this property, we first extend S_η^q to PBCs. For a PBC $q_{\mathbf{Y}_{11},\mathbf{Y}_{21}|\mathbf{X}_1}^{(1)} \times q_{\mathbf{Y}_{12},\mathbf{Y}_{22}|\mathbf{X}_2}^{(2)}$ we set,

$$\mathsf{s}_\eta^{q_1 \times q_2}(\mathbf{X}_1, \mathbf{X}_2)$$
$$\triangleq I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{21}, \mathbf{Y}_{22}) - \eta I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{11}, \mathbf{Y}_{12}), \quad (21)$$

and define the quantities $\mathsf{s}_\eta^{q_1 \times q_2}(\mathbf{X}_1, \mathbf{X}_2|V)$, $\mathsf{S}_\eta^{q_1 \times q_2}(\mathbf{X}_1, \mathbf{X}_2)$ and $\mathsf{S}_\eta^{q_1 \times q_2}(\mathbf{X}_1, \mathbf{X}_2|V)$ analogously to the definitions of $\mathsf{s}_\eta^q(\mathbf{X}|V)$, $\mathsf{S}_\eta^q(\mathbf{X})$ and $\mathsf{S}_\eta^q(\mathbf{X}|V)$ from the above, respectively.

Proposition 2 (Factorization Property): For any PBC $q_{\mathbf{Y}_{11},\mathbf{Y}_{21}|\mathbf{X}_1}^{(1)} \times q_{\mathbf{Y}_{12},\mathbf{Y}_{22}|\mathbf{X}_2}^{(2)}$, the following chain of inequalities holds

$$\mathsf{S}_\eta^{q_1 \times q_2}(\mathbf{X}_1, \mathbf{X}_2) \leq \mathsf{S}_\eta^{q_1}(\mathbf{X}_1|\mathbf{Y}_{22}) + \mathsf{S}_\eta^{q_2}(\mathbf{X}_2|\mathbf{Y}_{11})$$
$$\leq \mathsf{S}_\eta^{q_1}(\mathbf{X}_1) + \mathsf{S}_\eta^{q_2}(\mathbf{X}_2). \quad (22)$$

The proof of Proposition 2 follows by repeating the steps in the proof of [16, Proposition 6], while switching the roles of $\mathbf{Y}_1$ and $\mathbf{Y}_2$.

Theorem 3 (Gaussian Maximizer): Let $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathrm{K})$. There exists a decomposition $\mathbf{X} = \mathbf{X}^\star + \mathbf{X}'$, such that $\mathbf{X}^\star$ and $\mathbf{X}'$ are independent, $\mathbf{X}^\star \sim \mathcal{N}(\mathbf{0}, \mathrm{K}^\star)$, $\mathbf{X}' \sim \mathcal{N}(\mathbf{0}, \mathrm{K} - \mathrm{K}^\star)$, where $\mathrm{K}^\star \preceq \mathrm{K}$, and $\mathsf{S}_\eta^q(\mathbf{X}) = \mathsf{s}_\eta^q(\mathbf{X}^\star) = V_\eta^q(\mathrm{K})$. Furthermore, this decomposition (i.e., the covariance matrix $\mathrm{K}^\star$) is unique.

The proof of Theorem 3 is also omitted as it mimics the proofs [16, Th.1 and Corollary 1].

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 V-A).

For a BC $q_{\mathbf{Y}_1,\mathbf{Y}_2|\mathbf{X}}$, $\eta > 1$, $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$, where $\lambda_j > 0$, $j = 1, 2$, and any $\mathbf{X} \sim P_\mathbf{X}$ define

$$\mathsf{t}_{\boldsymbol{\lambda},\eta}^q(\mathbf{X}) \triangleq \lambda_1 I(\mathbf{X}; \mathbf{Y}_1) - (\lambda_1 + \lambda_2) I(\mathbf{X}; \mathbf{Y}_2) + \lambda_1 \mathsf{S}_\eta^q(\mathbf{X}), \quad (23)$$

where $\mathsf{S}_\eta^q(\mathbf{X})$ is given by (19). As before, for a pair of random variables $(V, \mathbf{X})$ for which $V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$ forms a Markov chain, let

$$\mathsf{t}_{\boldsymbol{\lambda},\eta}^q(\mathbf{X}|V)$$
$$\triangleq \lambda_1 I(\mathbf{X}; \mathbf{Y}_1|V) - (\lambda_1 + \lambda_2) I(\mathbf{X}; \mathbf{Y}_2|V) + \lambda_1 \mathsf{S}_\eta^q(\mathbf{X}|V), \quad (24)$$

and set

$$\mathsf{T}_{\boldsymbol{\lambda},\eta}^q(\mathbf{X}) \triangleq \mathfrak{C}\big(\mathsf{t}_{\boldsymbol{\lambda},\eta}^q(\mathbf{X})\big) = \sup_{\substack{P_{V|\mathbf{X}}: \\ V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)}} \mathsf{t}_{\boldsymbol{\lambda},\eta}^q(\mathbf{X}|V). \quad (25)$$

Define $\mathsf{T}_{\boldsymbol{\lambda},\eta}^q(\mathbf{X}|V) \triangleq \sum_v P(v) \mathsf{T}_{\boldsymbol{\lambda},\eta}^q(\mathbf{X}|V = v)$, 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_η^q stated in Proposition 1, since $\mathsf{T}_{\boldsymbol{\lambda},\eta}^q$ is concave in $P_\mathbf{X}$, Jensen's inequality implies that $\mathsf{T}_{\boldsymbol{\lambda},\eta}^q(\mathbf{X}|V) \leq \mathsf{T}_{\boldsymbol{\lambda},\eta}^q(\mathbf{X})$, for any $(V, \mathbf{X})$ satisfying $V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$. Moreover, if $W - V - \mathbf{X}$ forms a Markov chain, then $\mathsf{T}_{\boldsymbol{\lambda},\eta}^q(\mathbf{X}|W, V) = \mathsf{T}_{\boldsymbol{\lambda},\eta}^q(\mathbf{X}|V)$, because $P_{\mathbf{X}|W,V} = P_{\mathbf{X}|V}$. Finally, $\mathsf{T}_{\boldsymbol{\lambda},\eta}^q(\mathbf{X})$ is convex in η inside $(0, 2)$, for a fixed $P_\mathbf{X}$, and therefore it is continuous as a function of η at $\eta = 1$.

Definition 4 (Maximized Nested Concave Envelope): For any MIMO Gaussian BC $q_{\mathbf{Y}_1,\mathbf{Y}_2|\mathbf{X}}$, define

$$\hat{V}_{\boldsymbol{\lambda},\eta}^q(\mathrm{K}) \triangleq \sup_{\mathbf{X}: \, \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq \mathrm{K}} \mathsf{T}_{\boldsymbol{\lambda},\eta}^q(\mathbf{X}) = \sup_{\substack{(V,\mathbf{X}): \, \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq \mathrm{K}, \\ V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)}} \mathsf{t}_{\boldsymbol{\lambda},\eta}^q(\mathbf{X}|V).$$

(26)

Proposition 3 (Continuity of Maximal Value): For any $\boldsymbol{\lambda}$ as before, $\hat{V}_{\boldsymbol{\lambda},\eta}(\mathrm{K})$ is continuous in η at $\eta = 1$.

The proof of Proposition 3 follows by arguments similar to those in the proof of Property 3 of Proposition 1. Namely, the continuity of $\hat{V}_{\boldsymbol{\lambda},\eta}(\mathrm{K})$ at $\eta = 1$ follows by verifying that $\hat{V}_{\boldsymbol{\lambda},\eta}(\mathrm{K})$ is convex in η inside $(0, 2)$ and using Proposition 17 from [26, Ch. 5].

As before, to state the factorization property for nested UCEs, we extend some of the preceding definitions to PBCs. For a PBC $q_{\mathbf{Y}_{11},\mathbf{Y}_{21}|\mathbf{X}_1}^{(1)} \times q_{\mathbf{Y}_{12},\mathbf{Y}_{22}|\mathbf{X}_2}^{(2)}$, we set

$$\mathsf{t}_{\boldsymbol{\lambda},\eta}^{q_1 \times q_2}(\mathbf{X}_1, \mathbf{X}_2)$$
$$\triangleq \lambda_1 I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{11}, \mathbf{Y}_{12}) - (\lambda_1 + \lambda_2) I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{21}, \mathbf{Y}_{22})$$
$$+ \lambda_1 \mathsf{S}_\eta^{q_1 \times q_2}(\mathbf{X}_1, \mathbf{X}_2). \quad (27)$$

Furthermore, define $\mathsf{t}_{\boldsymbol{\lambda},\eta}^{q_1 \times q_2}(\mathbf{X}_1, \mathbf{X}_2|V)$, $\mathsf{T}_{\boldsymbol{\lambda},\eta}^{q_1 \times q_2}(\mathbf{X}_1, \mathbf{X}_2)$ and $\mathsf{T}_{\boldsymbol{\lambda},\eta}^{q_1 \times q_2}(\mathbf{X}_1, \mathbf{X}_2|V)$ in a similar manner to $\mathsf{t}_{\boldsymbol{\lambda},\eta}^q(\mathbf{X}|V)$, $\mathsf{T}_{\boldsymbol{\lambda},\eta}^q(\mathbf{X})$ and $\mathsf{T}_{\boldsymbol{\lambda},\eta}^q(\mathbf{X}|V)$, respectively. The following proposition states the $\mathsf{T}_{\boldsymbol{\lambda},\eta}^{q_1 \times q_2}(\mathbf{X}_1, \mathbf{X}_2|V)$ factorization property of interest, which is used to prove the existence of a Gaussian maximizer for $\hat{V}_{\boldsymbol{\lambda},\eta}^q(\mathrm{K})$ from (26).

Proposition 4 (Factorization Property): For any PBC $q_{\mathbf{Y}_{11},\mathbf{Y}_{21}|\mathbf{X}_1}^{(1)} \times q_{\mathbf{Y}_{12},\mathbf{Y}_{22}|\mathbf{X}_2}^{(2)}$, the following chain of inequalities holds

$$\mathsf{T}_{\boldsymbol{\lambda},\eta}^{q_1 \times q_2}(\mathbf{X}_1, \mathbf{X}_2) \leq \mathsf{T}_{\boldsymbol{\lambda},\eta}^{q_1}(\mathbf{X}_1|\mathbf{Y}_{22}) + \mathsf{T}_{\boldsymbol{\lambda},\eta}^{q_2}(\mathbf{X}_2|\mathbf{Y}_{11})$$
$$\leq \mathsf{T}_{\boldsymbol{\lambda},\eta}^{q_1}(\mathbf{X}_1) + \mathsf{T}_{\boldsymbol{\lambda},\eta}^{q_2}(\mathbf{X}_2) \quad (28)$$

Furthermore, if the PBC is Gaussian and a triple $(V^\star, \mathbf{X}_1^\star, \mathbf{X}_2^\star)$ satisfies

$$\mathsf{t}_{\boldsymbol{\lambda},\eta}^{q_1 \times q_2}(\mathbf{X}_1^\star, \mathbf{X}_2^\star|V^\star) = \mathsf{T}_{\boldsymbol{\lambda},\eta}^{q_1 \times q_2}(\mathbf{X}_1^\star, \mathbf{X}_2^\star) = \mathsf{T}_{\boldsymbol{\lambda},\eta}^{q_1}(\mathbf{X}_1^\star) + \mathsf{T}_{\boldsymbol{\lambda},\eta}^{q_2}(\mathbf{X}_2^\star)$$

(29)

then $\mathbf{X}_1^\star - V^\star - \mathbf{X}_2^\star$ and $\mathsf{t}_{\boldsymbol{\lambda},\eta}^{q_j}(\mathbf{X}_j^\star|V^\star) = \mathsf{T}_{\boldsymbol{\lambda},\eta}^{q_j}(\mathbf{X}_j^\star)$, for $j = 1, 2$.

See Section VI-B for the proof of Proposition 4. The existence of a Gaussian maximizer for $\hat{V}_{\boldsymbol{\lambda},\eta}^q(\mathrm{K})$ follows by repeating the proofs of Theorem 2 and Corollary 2 in [16] with respect to our definition of $\mathsf{T}_{\boldsymbol{\lambda},\eta}^q$. The existence is stated in the following Theorem, which we give without proof.

Theorem 4 (Gaussian Maximizer): Let $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathrm{K})$. There exists a unique decomposition $\mathbf{X} = \mathbf{X}_1^\star + \mathbf{X}_2^\star + \mathbf{X}'$ into independent random variables $(\mathbf{X}_1^\star, \mathbf{X}_2^\star, \mathbf{X}')$, where $\mathbf{X}_j^\star \sim$

$\mathcal{N}(\mathbf{0}, \mathrm{K}_j)$, $j = 1, 2$, and $\mathbf{X}' \sim \mathcal{N}(\mathbf{0}, \mathrm{K} - (\mathrm{K}_1 + \mathrm{K}_2))$, $\mathrm{K}_1 + \mathrm{K}_2 \preceq \mathrm{K}$, such that

$$\mathsf{T}^q_{\lambda,\eta}(\mathbf{X}) = \mathsf{t}^q_{\lambda,\eta}(\mathbf{X}^\star_1 + \mathbf{X}^\star_2) = \hat{V}^q_{\lambda,\eta}(\mathrm{K}) \tag{30a}$$

$$\mathsf{S}^q_\eta(\mathbf{X}^\star_1 + \mathbf{X}^\star_2) = \mathsf{s}^q_\eta(\mathbf{X}^\star_1) = V^q_\eta(\mathrm{K}_1 + \mathrm{K}_2). \tag{30b}$$

C. Double-Nested Upper Concave 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 II-B). Let $q_{\mathbf{Y}_1,\mathbf{Y}_2|\mathbf{X}}$, $\eta > 1$ be a BC, $\boldsymbol{\lambda_0} = (\lambda_0, \lambda_1, \lambda_2)$, where $\lambda_j > 0$ for $j = 0, 1, 2$ and $\lambda_0 > \lambda_2$, $\alpha \in [0, 1]$ and $\bar{\alpha} = 1 - \alpha$. For any $\mathbf{X} \sim P_{\mathbf{X}}$ define

$$\mathsf{f}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}) \triangleq (\lambda_2 - \bar{\alpha}\lambda_0)I(\mathbf{X}; \mathbf{Y}_2) - \alpha\lambda_0 I(\mathbf{X}; \mathbf{Y}_1) + \mathsf{T}^q_{\lambda,\eta}(\mathbf{X}), \tag{31}$$

where $\mathsf{T}^q_{\lambda,\eta}(\mathbf{X})$ is given by (25) and $\lambda = (\lambda_1, \lambda_2)$.

For $(V, \mathbf{X})$ that satisfy the Markov chain $V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$, we set $\mathsf{f}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}|V)$ in an analogous manner to (24), while $\mathsf{F}^q_{\boldsymbol{\lambda_0},\alpha,\eta} \triangleq \mathfrak{C}\mathsf{f}^q_{\boldsymbol{\lambda_0},\alpha,\eta}$ denotes the UCE of $\mathsf{f}^q_{\boldsymbol{\lambda_0},\alpha,\eta}$. We also set $\mathsf{F}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}|V) \triangleq \sum_v P(v)\mathsf{F}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}|V = v)$ for a discrete V and, as before, consider its natural extension in the case where V is arbitrary.

Remark 10 (Double-Nested Concave Envelopes Properties): The concavity of $\mathsf{F}^q_{\boldsymbol{\lambda_0},\alpha,\eta}$ in $P_{\mathbf{X}}$ and Jensen's inequality imply that for any $(V, \mathbf{X})$ with $V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$, we have $\mathsf{F}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}|V) \leq \mathsf{F}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X})$. If the chain $W - V - \mathbf{X}$ is Markov, then $\mathsf{F}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}|W, V) = \mathsf{F}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}|V)$, because $P_{\mathbf{X}|W,V} = P_{\mathbf{X}|V}$. As a function of η, $\mathsf{F}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X})$ is convex inside $(0, 2)$, for any fixed $\mathbf{X}$, and is thus continuous at $\eta = 1$.

Definition 5 (Maximized Double-Nested Concave Envelope): For any MIMO Gaussian BC $q_{\mathbf{Y}_1,\mathbf{Y}_2|\mathbf{X}}$, define

$$\tilde{V}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathrm{K}) \triangleq \sup_{\mathbf{X}:\, \mathbb{E}[\mathbf{X}\mathbf{X}^\top]\preceq\mathrm{K}} \mathsf{F}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X})$$
$$= \sup_{\substack{(V,\mathbf{X}):\, \mathbb{E}[\mathbf{X}\mathbf{X}^\top]\preceq\mathrm{K},\\ V-\mathbf{X}-(\mathbf{Y}_1,\mathbf{Y}_2)}} \mathsf{f}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}|V). \tag{32}$$

Remark 11 (Continuity of Maximal Value): As before, one readily verifies that as a function of η, $\tilde{V}_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathrm{K})$ is convex inside $(0, 2)$, and deduce its continuity at $\eta = 1$.

The above notions are once again extended to PBCs. Namely, for any PBC $q^{(1)}_{\mathbf{Y}_{11},\mathbf{Y}_{21}|\mathbf{X}_1} \times q^{(2)}_{\mathbf{Y}_{12},\mathbf{Y}_{22}|\mathbf{X}_2}$, we set

$$\mathsf{f}^{q_1 \times q_2}_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}_1, \mathbf{X}_2)$$
$$= (\lambda_2 - \bar{\alpha}\lambda_0)I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{21}, \mathbf{Y}_{22})$$
$$- \alpha\lambda_0 I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{11}, \mathbf{Y}_{12}) + \mathsf{T}^{q_1 \times q_2}_{\lambda,\eta}(\mathbf{X}_1, \mathbf{X}_2), \tag{33}$$

and define $\mathsf{f}^{q_1 \times q_2}_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}_1, \mathbf{X}_2|V)$, $\mathsf{F}^{q_1 \times q_2}_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}_1, \mathbf{X}_2)$ and $\mathsf{F}^{q_1 \times q_2}_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}_1, \mathbf{X}_2|V)$ as the natural extensions to the PBC scenario of $\mathsf{f}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}|V)$, $\mathsf{F}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X})$ and $\mathsf{F}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}|V)$ given above, respectively.

Moving forward, the factorization property of $\mathsf{F}^{q_1 \times q_2}_{\boldsymbol{\lambda_0},\alpha,\eta}$ is stated in Proposition 5, while Proposition 6 establishes the existence of its maximizer.

Proposition 5 (Factorization Property): For any PBC $q^{(1)}_{\mathbf{Y}_{11},\mathbf{Y}_{21}|\mathbf{X}_1} \times q^{(2)}_{\mathbf{Y}_{12},\mathbf{Y}_{22}|\mathbf{X}_2}$, the following chain of inequalities holds

$$\mathsf{F}^{q_1 \times q_2}_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}_1, \mathbf{X}_2) \leq \mathsf{F}^{q_1}_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}_1|\mathbf{Y}_{22}) + \mathsf{F}^{q_2}_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}_2|\mathbf{Y}_{11})$$
$$\leq \mathsf{F}^{q_1}_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}_1) + \mathsf{F}^{q_2}_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}_2) \tag{34}$$

Furthermore, if the PBC is Gaussian and a triple $(V^\star, \mathbf{X}^\star_1, \mathbf{X}^\star_2)$ satisfies

$$\mathsf{f}^{q_1 \times q_2}_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}^\star_1, \mathbf{X}^\star_2|V^\star) = \mathsf{F}^{q_1 \times q_2}_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}^\star_1, \mathbf{X}^\star_2)$$
$$= \mathsf{F}^{q_1}_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}^\star_1) + \mathsf{F}^{q_2}_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}^\star_2), \tag{35}$$

then $\mathbf{X}^\star_1 - V^\star - \mathbf{X}^\star_2$ and $\mathsf{f}^{q_j}_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}^\star_j|V^\star) = \mathsf{F}^{q_j}_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}^\star_j)$, for $j = 1, 2$.

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

Proposition 6 (Existence of a Maximizer): There exists a pair $(V^\star, \mathbf{X}^\star)$ with $|\mathcal{V}^\star| \leq \frac{t(t+1)}{2} + 1$ and $\mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq \mathrm{K}$, such that

$$\tilde{V}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathrm{K}) = \mathsf{f}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}^\star|V^\star). \tag{36}$$

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

The existence of a maximizer and the cardinality bound on $\mathcal{V}^\star$ are proven in Section VI-D. A zero conditional expectation can be assumed because centering conditioned on each $V^\star = v^\star$ does not change the mutual information terms and hence $\mathsf{f}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}^\star_2|V^\star)$ remains unchanged as well. In addition, the centered versions of the input continues to satisfy the covariance constraint.

To show that the distribution that achieves $\tilde{V}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathrm{K})$ is Gaussian, we use the invariance of $\mathsf{f}^q_{\boldsymbol{\lambda_0},\alpha,\eta}$ 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 $(V, \mathbf{X}) \sim P^\star_{V,\mathbf{X}}$ attain $\tilde{V}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathrm{K})$, with $|\mathcal{V}| = m \leq \frac{t(t+1)}{2} + 1$, and let $\mathbf{X}_v$ be a centered random variable (zero mean) distributed according to the conditional PMF $P^\star_{\mathbf{X}|V=v}$. Let $(V_1, \mathbf{X}_1, V_2, \mathbf{X}_2) \sim P^\star_{V,\mathbf{X}} \times P^\star_{V,\mathbf{X}}$ be two i.i.d. copies of $(V^\star, \mathbf{X}^\star)$. Define

$$\tilde{V} = (V_1, V_2)$$
$$\mathbf{X}_{\theta_1}\big|\{\tilde{V} = (v_1, v_2)\} \sim \frac{1}{\sqrt{2}}(\mathbf{X}_{v_1} + \mathbf{X}_{v_2})$$
$$\mathbf{X}_{\theta_2}\big|\{\tilde{V} = (v_1, v_2)\} \sim \frac{1}{\sqrt{2}}(\mathbf{X}_{v_1} - \mathbf{X}_{v_2}),$$

where $\mathbf{X}_{v_1}$ and $\mathbf{X}_{v_2}$ are taken to be independent random variables, i.e., $(\mathbf{X}_{v_1}, \mathbf{X}_{v_2}) \sim P^\star_{\mathbf{X}|V=v_1} \times P^\star_{\mathbf{X}|V=v_2}$. Then $\mathbf{X}_{\theta_1} - \tilde{V} - \mathbf{X}_{\theta_2}$ and $\tilde{V}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathrm{K}) = \mathsf{f}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}_{\theta_j}|\tilde{V})$, for $j = 1, 2$.

The existence of a Gaussian Maximizer for $\tilde{V}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathrm{K})$ is stated next.

Theorem 5 (Existence of Gaussian Maximizer): There exists an $\mathbf{X}^\star \sim \mathcal{N}(\mathbf{0}, \mathrm{K}^\star)$, where $\mathrm{K}^\star \preceq \mathrm{K}$, such that $\tilde{V}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathrm{K}) = \mathsf{f}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}^\star)$. Furthermore, the zero mean maximizer is unique.

See Section VI-F for the proof.

Corollary 3 (Gaussian Maximizer Properties): Let $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathrm{K})$. There is a unique decomposition $\mathbf{X} = \mathbf{X}^\star_1 + \mathbf{X}^\star_2 + \mathbf{X}^\star_3 + \mathbf{X}'$ into independent random variables $(\mathbf{X}^\star_1, \mathbf{X}^\star_2, \mathbf{X}^\star_3, \mathbf{X}')$,

where $\mathbf{X}_j^\star \sim \mathcal{N}(\mathbf{0}, K_j)$, for $j = 1, 2, 3$, and $\mathbf{X}' \sim \mathcal{N}(\mathbf{0}, K - (K_1 + K_2 + K_3))$, with $K_1 + K_2 + K_3 \preceq K$, such that

$$F_{\lambda_0, a, \eta}^q(\mathbf{X}) = f_{\lambda_0, a, \eta}^q(\mathbf{X}_1^\star + \mathbf{X}_2^\star + \mathbf{X}_3^\star) = \tilde{V}_{\lambda_0, a, \eta}^q(K) \tag{37a}$$

$$T_{\lambda, \eta}^q(\mathbf{X}_1^\star + \mathbf{X}_2^\star + \mathbf{X}_3^\star) = t_{\lambda, \eta}^q(\mathbf{X}_1^\star + \mathbf{X}_2^\star) = \hat{V}_{\lambda, \eta}^q(K_1 + K_2 + K_3) \tag{37b}$$

$$S_\eta^q(\mathbf{X}_1^\star + \mathbf{X}_2^\star) = s_\eta^q(\mathbf{X}_1^\star) = V_\eta^q(K_1 + K_2). \tag{37c}$$

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-G.

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 [23, Ths. 1 and 3], respectively. To state the bounds, let $\hat{\mathcal{C}}$ denote the secrecy-capacity region of the corresponding discrete-memoryless (DM) BC.

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

$$R_1 \leq I(U; Y_1|V) - I(U; Y_2|V) \tag{38a}$$
$$R_2 \leq I(V; Y_2) \tag{38b}$$

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

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

$$R_1 \leq I(U; Y_1) - I(U; V) - I(U; Y_2|V) \tag{39a}$$
$$R_2 \leq I(V; Y_2) \tag{39b}$$

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

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

The opposite inclusion, i.e., $\hat{\mathcal{O}}_K \subseteq \hat{\mathcal{I}}_K$, is shown next. The regions $\hat{\mathcal{I}}_K$ and $\hat{\mathcal{O}}_K$ 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 $\hat{\mathcal{O}}_K$ and $\hat{\mathcal{I}}_K$:

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

where

$$\mathcal{H}_1^K \triangleq \max_{\substack{(V,U)-\mathbf{X}-(\mathbf{Y}_1,\mathbf{Y}_2): \\ \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K}} I(U; \mathbf{Y}_1|V) - I(U; \mathbf{Y}_2|V) \tag{41a}$$

$$\mathcal{H}_2^K \triangleq \max_{\mathbf{X}: \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq K} I(\mathbf{X}; \mathbf{Y}_2). \tag{41b}$$

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

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

$$\max_{(R_1,R_2)\in\hat{\mathcal{O}}_K} \lambda_1 R_1 + \lambda_2 R_2 \leq \max_{(R_1,R_2)\in\hat{\mathcal{I}}_K} \lambda_1 R_1 + \lambda_2 R_2, \tag{42}$$

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

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

$$\overset{(a)}{\leq} \sup_{\substack{(V,U)-\mathbf{X}-(\mathbf{Y}_1,\mathbf{Y}_2): \\ \mathbb{E}[\mathbf{X}\mathbf{X}^\top]\preceq K}} \lambda_1\Big[I(U; \mathbf{Y}_1|V) - I(U; \mathbf{Y}_2|V)\Big] + \lambda_2 I(V; \mathbf{Y}_2)$$

$$\overset{(b)}{=} \sup_{\substack{(V,U)-\mathbf{X}-(\mathbf{Y}_1,\mathbf{Y}_2): \\ \mathbb{E}[\mathbf{X}\mathbf{X}^\top]\preceq K}} \lambda_1 I(\mathbf{X}; \mathbf{Y}_1|V) - (\lambda_1 + \lambda_2)I(\mathbf{X}; \mathbf{Y}_2|V) + \lambda_1\Big[I(\mathbf{X}; \mathbf{Y}_2|V, U) - I(\mathbf{X}; \mathbf{Y}_1|V, U)\Big] + \lambda_2 I(\mathbf{X}; \mathbf{Y}_2)$$

$$\overset{(c)}{\leq} \sup_{\substack{V-\mathbf{X}-(\mathbf{Y}_1,\mathbf{Y}_2): \\ \mathbb{E}[\mathbf{X}\mathbf{X}^\top]\preceq K}} \lambda_1 I(\mathbf{X}; \mathbf{Y}_1|V) - (\lambda_1 + \lambda_2)I(\mathbf{X}; \mathbf{Y}_2|V) + \lim_{\eta\downarrow 1} \lambda_1 S_\eta^q(\mathbf{X}|V) + \lambda_2 I(\mathbf{X}; \mathbf{Y}_2)$$

$$\leq \sup_{\mathbb{E}[\mathbf{X}\mathbf{X}^\top]\preceq K} \lambda_2 I(\mathbf{X}; \mathbf{Y}_2) + \sup_{\substack{V-\mathbf{X}-(\mathbf{Y}_1,\mathbf{Y}_2): \\ \mathbb{E}[\mathbf{X}\mathbf{X}^\top]\preceq K}} \lim_{\eta\downarrow 1} t_{\lambda,\eta}(\mathbf{X}|V)$$

$$\overset{(d)}{=} \sup_{\mathbb{E}[\mathbf{X}\mathbf{X}^\top]\preceq K} \lambda_2 I(\mathbf{X}; \mathbf{Y}_2) + \sup_{\mathbb{E}[\mathbf{X}\mathbf{X}^\top]\preceq K} \lim_{\eta\downarrow 1} T_{\lambda,\eta}(\mathbf{X})$$

$$\overset{(e)}{=} \sup_{\mathbb{E}[\mathbf{X}\mathbf{X}^\top]\preceq K} \lambda_2 I(\mathbf{X}; \mathbf{Y}_2) + \lim_{\eta\downarrow 1} \hat{V}_{\lambda,\eta}(K), \tag{43}$$

where:

(a) uses (38);

(b) is because $(V, U) - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$ forms a Markov chain;

(c) follows by the definition of $s_\eta^q(\mathbf{X}|V)$ and since conditioned on V, $U - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$ forms a Markov chain, i.e., it holds that $P_{\mathbf{Y}_1,\mathbf{Y}_2|V,U,\mathbf{X}} = P_{\mathbf{Y}_1,\mathbf{Y}_2|\mathbf{X}}$. Furthermore, (c) uses the continuity of $S_\eta^q(\mathbf{X}|V)$ in η at $\eta = 1$ (see Property 3 of Proposition 1), which implies that for any $(V, \mathbf{X})$

$$S_1^q(\mathbf{X}|V) \triangleq S_{\lim_{\eta\downarrow 1}\eta}^q(\mathbf{X}|V) = \lim_{\eta\downarrow 1} S_\eta^q(\mathbf{X}|V);$$

(d) is by the definition of $T_{\lambda,\eta}^q(\mathbf{X})$, the Markov relation $V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$, and because $T_{\lambda,\eta}^q(\mathbf{X})$ is continuous in η at $\eta = 1$ (see Remark 9);

(e) follows by Proposition 3.

By Theorem 4, for every $\eta > 1$, there exist independent random variables $\mathbf{X}_1^\star \sim \mathcal{N}(\mathbf{0}, K_1)$, $\mathbf{X}_2^\star \sim \mathcal{N}(\mathbf{0}, K_2)$ and $\mathbf{X}' \sim \mathcal{N}(\mathbf{0}, K - (K_1 + K_2))$, $K_1 + K_2 \preceq K$, such that $\hat{V}_\eta^q(K) = t_{\lambda,\eta}^q(\mathbf{X}_1^\star + \mathbf{X}_2^\star)$ and $S_\eta^q(\mathbf{X}_1^\star + \mathbf{X}_2^\star) = s_\eta^q(\mathbf{X}_1^\star)$. Moreover, setting $\mathbf{X} = \mathbf{X}_1^\star + \mathbf{X}_2^\star + \mathbf{X}'$ maximizes $\lambda_2 I(\mathbf{X}; \mathbf{Y}_2)$ and attains $\hat{V}_\eta^q(K)$ simultaneously. In order to conform to the notation in the bounds, let $V^\star = \mathbf{X}'$. Taking the limit as $\eta \downarrow 1$, we have

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

$$\leq \lambda_1 I(\mathbf{X}; \mathbf{Y}_1|V^\star) - (\lambda_1 + \lambda_2)I(\mathbf{X}; \mathbf{Y}_2|V^\star)$$

$$+ \lambda_1 \Big[I(\mathbf{X}; \mathbf{Y}_2 | V^\star, \mathbf{X}_2^\star) - I(\mathbf{X}; \mathbf{Y}_2 | V^\star, \mathbf{X}_2^\star) \Big] + \lambda_2 I(\mathbf{X}; \mathbf{Y}_2)$$

$$\leq \lambda_1 \Big[I(\mathbf{X}_2^\star; \mathbf{Y}_1 | V^\star) - I(\mathbf{X}_2^\star; \mathbf{Y}_2 | V^\star) \Big] + \lambda_2 I(V^\star; \mathbf{Y}_2). \quad (44)$$

The following proposition is used to show that (44) is achievable within $\hat{\mathcal{I}}_{\mathrm{K}}$.

Proposition 8 (Partial Dirty-Paper Coding (P-DPC)): Fix a covariance matrix K *and let* $\mathbf{X} = \mathbf{X}_1^\star + \mathbf{X}_2^\star + V^\star$, *where* $\mathbf{X}_1^\star$, $\mathbf{X}_2^\star$ *and* V *are independent Gaussian random vectors with covariance matrices* K_1, K_2 *and* $\mathrm{K} - (\mathrm{K}_1 + \mathrm{K}_2)$, *respectively, for some* $0 \preceq \mathrm{K}_1, \mathrm{K}_2 \preceq \mathrm{K}$ *with* $\mathrm{K}_1 + \mathrm{K}_2 \preceq \mathrm{K}$. *Let* $\mathbf{Y}_j = \mathrm{G}_j \mathbf{X} + \mathbf{Z}_j$, *for* $j = 1, 2$, *where* $\mathbf{Z}_j \sim \mathcal{N}(\mathbf{0}, \mathrm{I})$ *is independent of* $(\mathbf{X}_1^\star, \mathbf{X}_2^\star, V^\star)$. *Set* $U = \mathbf{X}_2^\star + \mathrm{A} V^\star$, *where* $\mathrm{A} = \mathrm{K}_2 \tilde{\mathrm{G}}_1^\top \Big[\mathrm{I} + \tilde{\mathrm{G}}_1 \mathrm{K}_2 \tilde{\mathrm{G}}_1^\top \Big]^{-1}$ *and* $\tilde{\mathrm{G}}_1 = \Big[\mathrm{I} + \mathrm{G}_1 \mathrm{K}_1 \mathrm{G}_1^\top \Big]^{-\frac{1}{2}} \mathrm{G}_1$. *Then*

$$I(\mathbf{X}_2^\star; \mathbf{Y}_1 | V^\star) - I(\mathbf{X}_2^\star; \mathbf{Y}_2 | V^\star)$$
$$= I(U^\star; \mathbf{Y}_1) - I(U^\star; V^\star) - I(\mathbf{X}_2^\star; \mathbf{Y}_2 | V^\star). \quad (45)$$

Proof: We first write

$$\begin{aligned} \mathbf{Y}_1 &= \mathrm{G}_1 \mathbf{X} + \mathbf{Z}_1 \\ &= \mathrm{G}_1 (\mathbf{X}_1^\star + \mathbf{X}_2^\star + V^\star) + \mathbf{Z}_1 \\ &= \mathrm{G}_1 (\mathbf{X}_2^\star + V^\star) + (\mathrm{G}_1 \mathbf{X}_1^\star + \mathbf{Z}_1) \\ &\stackrel{(a)}{=} \mathrm{G}_1 \tilde{\mathbf{X}} + \mathbf{Z}_1', \end{aligned} \quad (46)$$

where (a) follows by setting $\tilde{\mathbf{X}} \triangleq \mathbf{X}_2^\star + V^\star$ and $\mathbf{Z}_1' \triangleq \mathrm{G}_1 \mathbf{X}_1^\star + \mathbf{Z}_1$. By the independence of $\mathbf{X}_1^\star$, $\mathbf{X}_2^\star$, $V^\star$ and $\mathbf{Z}_1$, we have that $\tilde{\mathbf{X}}$ and $\mathbf{Z}_1'$ are also independent. Moreover, $\mathbf{Z}_1' \sim \mathcal{N}\big(\mathbf{0}, \mathrm{I} + \mathrm{G}_1 \mathrm{K}_1 \mathrm{G}_1^\top\big)$, where the covariance matrix $\mathrm{I} + \mathrm{G}_1 \mathrm{K}_1 \mathrm{G}_1^\top$ is diagonalizable (due to its symmetry) and invertible (because it is positive definite). Denoting $\Sigma \triangleq \mathrm{I} + \mathrm{G}_1 \mathrm{K}_1 \mathrm{G}_1^\top$, gives

$$\Sigma = \mathrm{Q} \Lambda \mathrm{Q}^\top, \quad (47)$$

where Q is a unitary matrix and Λ is diagonal, and furthermore $\Sigma^{-\frac{1}{2}} = \mathrm{Q} \Lambda^{-\frac{1}{2}} \mathrm{Q}^\top$. By defining $\tilde{\mathbf{Y}}_1 = \Sigma^{-\frac{1}{2}} \mathbf{Y}_1$, we have

$$\tilde{\mathbf{Y}}_1 = \tilde{\mathrm{G}}_1 \tilde{\mathbf{X}} + \tilde{\mathbf{Z}}_1, \quad (48)$$

where $\tilde{\mathrm{G}}_1 = \Sigma^{-\frac{1}{2}} \mathrm{G}_1$, $\tilde{\mathbf{Z}}_1 = \Sigma^{-\frac{1}{2}} \mathbf{Z}_1'$ and $\tilde{\mathbf{Z}}_1 \sim \mathcal{N}(\mathbf{0}, \mathrm{I})$. Setting $U^\star$ as above and invoking the classic Dirty-Paper Coding Theorem (here we use the formulation from [16, Proposition 12]), we have

$$I(\tilde{\mathbf{X}}; \tilde{\mathbf{Y}}_1 | V^\star) = I(U^\star; \tilde{\mathbf{Y}}_1) - (U^\star; V^\star). \quad (49)$$

Furthermore,

$$I(\mathbf{X}_2^\star; \mathbf{Y}_2 | V^\star) = I(\mathbf{X}_2^\star + \mathrm{A} V^\star; \mathbf{Y}_2 | V^\star) = I(U^\star; \mathbf{Y}_2 | V^\star). \quad (50)$$

Note that $\mathbf{Y}_1 \mapsto \Sigma^{-\frac{1}{2}} \mathbf{Y}_1$ is an invertible mapping, and as such, preserves mutual information. We conclude the proof as follows:

$$\begin{aligned} I(\mathbf{X}_2^\star; \mathbf{Y}_1 | V^\star) &- I(\mathbf{X}_2^\star; \mathbf{Y}_2 | V^\star) \\ &\stackrel{(a)}{=} I(\tilde{\mathbf{X}}; \mathbf{Y}_1 | V^\star) - I(U^\star; \mathbf{Y}_2 | V^\star) \\ &\stackrel{(b)}{=} I(\tilde{\mathbf{X}}; \tilde{\mathbf{Y}}_1 | V^\star) - I(U^\star; \mathbf{Y}_2 | V^\star) \\ &\stackrel{(c)}{=} I(U^\star; \tilde{\mathbf{Y}}_1) - I(U^\star; V^\star) - I(U^\star; \mathbf{Y}_2 | V^\star) \\ &\stackrel{(d)}{=} I(U^\star; \mathbf{Y}_1) - I(U^\star; V^\star) - I(U^\star; \mathbf{Y}_2 | V^\star), \end{aligned} \quad (51)$$

where (a) is because $\tilde{\mathbf{X}} = \mathbf{X}_2^\star + V^\star$ and by (50), (b) and (d) follow since $\mathbf{Y}_1 \mapsto \Sigma^{-\frac{1}{2}} \mathbf{Y}_1$ preserves mutual information, while (c) uses (49). ∎

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

$$\max_{(R_1, R_2) \in \hat{\mathcal{O}}_{\mathrm{K}}} \lambda_1 R_1 + \lambda_2 R_2$$
$$\leq \lambda_1 \Big[I(\mathbf{X}_2^\star; \mathbf{Y}_1 | V^\star) - I(\mathbf{X}_2^\star; \mathbf{Y}_2 | V^\star) \Big] + \lambda_2 I(V^\star; \mathbf{Y}_2)$$
$$= \lambda_1 \Big[I(U^\star; \mathbf{Y}_1) - I(U^\star; V^\star) - I(U^\star; \mathbf{Y}_2 | V^\star) \Big] + \lambda_2 I(V^\star; \mathbf{Y}_2)$$
$$\stackrel{(a)}{\leq} \max_{(R_1, R_2) \in \hat{\mathcal{I}}_{\mathrm{K}}} \lambda_1 R_1 + \lambda_2 R_2, \quad (52)$$

where (a) follows since $(U^\star, V^\star) - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$ forms a Markov chain and $\mathbb{E}\big[\mathbf{X}\mathbf{X}^\top\big] \preceq \mathrm{K}$ is satisfied, which implies that the rate pair $R_1 = I(U^\star; \mathbf{Y}_1) - I(U^\star; V^\star) - I(U^\star; \mathbf{Y}_2 | V^\star)$ and $R_2 = I(V^\star; \mathbf{Y}_2)$ belongs to $\hat{\mathcal{I}}_{\mathrm{K}}$. Concluding, we see that $\hat{\mathcal{I}}_{\mathrm{K}} = \hat{\mathcal{C}}_{\mathrm{K}} = \hat{\mathcal{O}}_{\mathrm{K}}$, which characterizes the secrecy-capacity region of the MIMO Gaussian BC with private and confidential messages.

Furthermore, equality (and hence the extreme points of $\hat{\mathcal{C}}_{\mathrm{K}}$) is achieved by Gaussian inputs as stated in Proposition 8, thus making the region computable. By evaluating $\hat{\mathcal{I}}_{\mathrm{K}}$ (or, equivalently $\hat{\mathcal{O}}_{\mathrm{K}}$) with respect to this input distribution, we describe the secrecy-capacity region $\hat{\mathcal{C}}_{\mathrm{K}}$ as the union of rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ satisfying:

$$\begin{aligned} R_1 &\leq \frac{1}{2} \log \left| \frac{\mathrm{I} + \mathrm{G}_1 (\mathrm{K}_1 + \mathrm{K}_2) \mathrm{G}_1^\top}{\mathrm{I} + \mathrm{G}_1 \mathrm{K}_1 \mathrm{G}_1^\top} \right| \\ &\quad - \frac{1}{2} \log \left| \frac{\mathrm{I} + \mathrm{G}_2 (\mathrm{K}_1 + \mathrm{K}_2) \mathrm{G}_2^\top}{\mathrm{I} + \mathrm{G}_2 \mathrm{K}_1 \mathrm{G}_2^\top} \right| \end{aligned} \quad (53a)$$

$$R_2 \leq \frac{1}{2} \log \left| \frac{\mathrm{I} + \mathrm{G}_2 \mathrm{K} \mathrm{G}_2^\top}{\mathrm{I} + \mathrm{G}_2 (\mathrm{K}_1 + \mathrm{K}_2) \mathrm{G}_2^\top} \right|, \quad (53b)$$

where the union is over all positive semi-definite matrices $\mathrm{K}_1, \mathrm{K}_2$, such that $\mathrm{K}_1 + \mathrm{K}_2 \preceq \mathrm{K}$. 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 $\mathrm{K}_1 = 0$ (see [7]–[9]). Further note that $\mathrm{K}_1 = 0$ cannot decrease the RHS of (53b). Thus, by relabeling $\mathrm{K}_2 \triangleq \mathrm{K}^\star$ we establish (6).

B. Proof of Theorem 2

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

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

$$R_0 \leq \min \big\{ I(W; Y_1), I(W; Y_2) \big\} \quad (54a)$$
$$R_1 \leq I(U; Y_1 | W, V) - I(U; Y_2 | W, V) \quad (54b)$$
$$R_0 + R_2 \leq I(V; Y_2 | W) + \min \big\{ I(W; Y_1), I(W; Y_2) \big\} \quad (54c)$$

over all $(W, V, U) - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$. *Then* $\mathcal{C} \subseteq \mathcal{O}$.

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:

$$R_0 \leq \min\left\{I(W; Y_1), I(W; Y_2)\right\} \tag{55a}$$

$$R_1 \leq I(U; Y_1|W) - I(U; V|W) - I(U; Y_2|W, V) \tag{55b}$$

$$R_2 \leq I(V; Y_2|W) \tag{55c}$$

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

The proofs of Bounds 3 and 4 are relegated to Appendix A. Denoting by $\mathcal{O}_K$ 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 $\mathcal{I}_K \subseteq \mathcal{C}_K \subseteq \mathcal{O}_K$.

Next, we use the factorization of UCEs method to show that the opposite inclusion $\mathcal{O}_K \subseteq \mathcal{I}_K$ 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 $\max_{(R_0,R_1,R_2)\in\mathcal{C}_K} \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2$, for $\lambda_j > 0$, $j = 1, 2, 3$. Further note that it suffices to restrict attention to the case where $\lambda_0 > \lambda_2$. This follows from the following observation: If a rate triple (R_0, R_1, R_2) is in $\mathcal{C}_K$ 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

$$\lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2 \leq 0 \cdot R_0 + \lambda_1 R_1 + \lambda_2(R_0 + R_2), \tag{56}$$

and therefore,

$$\begin{aligned}
&\max_{(R_0,R_1,R_2)\in\mathcal{C}_K} \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2 \\
&= \max_{(0,R_1,R_2)\in\mathcal{C}_K} \lambda_1 R_1 + \lambda_2 R_2 \\
&= \max_{(R_1,R_2)\in\hat{\mathcal{C}}_K} \lambda_1 R_1 + \lambda_2 R_2,
\end{aligned} \tag{57}$$

where $\hat{\mathcal{C}}_K$ is the secrecy-capacity region without a common message that was characterized in Section III-A.

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

$$\begin{aligned}
\max_{(R_0,R_1,R_2)\in\mathcal{O}_K} &\lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2 \\
&\leq \max_{(R_0,R_1,R_2)\in\mathcal{I}_K} \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2.
\end{aligned} \tag{58}$$

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

$$\begin{aligned}
&\max_{(R_0,R_1,R_2)\in\mathcal{O}_K} \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2 \\
&\overset{(a)}{\leq} \sup_{\substack{(W,V,U)-X-(Y_1,Y_2): \\ \mathbb{E}[XX^\top]\preceq K}} \lambda_0\Big[\alpha I(W; Y_1) + \bar{\alpha} I(W; Y_2)\Big] \\
&\quad + \lambda_1\Big[I(U; Y_1|W, V) - I(U; Y_2|W, V)\Big] \\
&\quad + \lambda_2 I(V; Y_2|W) \\
&\overset{(b)}{=} \sup_{\substack{(W,V,U)-X-(Y_1,Y_2): \\ \mathbb{E}[XX^\top]\preceq K}} \alpha\lambda_0 I(X; Y_1) + \bar{\alpha}\lambda_0 I(X; Y_2) \\
&\quad + (\lambda_2 - \bar{\alpha}\lambda_0)I(X; Y_2|W) - \alpha\lambda_0 I(X; Y_1|W) \\
&\quad + \lambda_1 I(X; Y_1|W, V) - (\lambda_1 + \lambda_2)I(X; Y_2|W, V) \\
&\quad + \lambda_1 I(X; Y_2|W, V, U) - \lambda_1 I(X; Y_1|W, V, U) \\
&\overset{(c)}{\leq} \sup_{\substack{(W,V)-X-(Y_1,Y_2): \\ \mathbb{E}[XX^\top]\preceq K}} \alpha\lambda_0 I(X; Y_1) + \bar{\alpha}\lambda_0 I(X; Y_2)
\end{aligned}$$

$$\begin{aligned}
&\quad + (\lambda_2 - \bar{\alpha}\lambda_0)I(X; Y_2|W) - \alpha\lambda_0 I(X; Y_1|W) \\
&\quad + \lambda_1 I(X; Y_1|W, V) - (\lambda_1 + \lambda_2)I(X; Y_2|W, V) \\
&\quad + \lim_{\eta\downarrow 1} \lambda_1 \mathsf{S}^q_\eta(X|W, V) \\
&\overset{(d)}{\leq} \sup_{\substack{W-X-(Y_1,Y_2): \\ \mathbb{E}[XX^\top]\preceq K}} \alpha\lambda_0 I(X; Y_1) + \bar{\alpha}\lambda_0 I(X; Y_2) \\
&\quad + (\lambda_2 - \bar{\alpha}\lambda_0)I(X; Y_2|W) - \alpha\lambda_0 I(X; Y_1|W) \\
&\quad + \lim_{\eta\downarrow 1} \mathsf{T}^q_{\lambda,\eta}(X|W) \\
&\leq \sup_{\mathbb{E}[XX^\top]\preceq K} \alpha\lambda_0 I(X; Y_1) + \bar{\alpha}\lambda_0 I(X; Y_2) \\
&\quad + \sup_{\substack{W-X-(Y_1,Y_2): \\ \mathbb{E}[XX^\top]\preceq K}} \lim_{\eta\downarrow 1} \mathsf{f}^q_{\lambda_0,\alpha,\eta}(X|W) \\
&\overset{(e)}{\leq} \sup_{\mathbb{E}[XX^\top]\preceq K} \alpha\lambda_0 I(X; Y_1) + \bar{\alpha}\lambda_0 I(X; Y_2) \\
&\quad + \sup_{\mathbb{E}[XX^\top]\preceq K} \lim_{\eta\downarrow 1} \mathsf{F}^q_{\lambda_0,\alpha,\eta}(X) \\
&\overset{(f)}{\leq} \sup_{\mathbb{E}[XX^\top]\preceq K} \alpha\lambda_0 I(X; Y_1) + \bar{\alpha}\lambda_0 I(X; Y_2) + \lim_{\eta\downarrow 1} \tilde{V}^q_{\lambda_0,\alpha,\eta}(K),
\end{aligned} \tag{59}$$

where:

(a) is by (54);

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

(c) is by the definition of $\mathsf{S}^q_\eta(X|W, V)$ since conditioned on (W, V), $U - X - (Y_1, Y_2)$ forms a Markov chain, and because $\mathsf{S}^q_\eta(X|V, W)$ is continuous in η at $\eta = 1$ (Property 3 of Proposition 1);

(d) follows by the definition of $\mathsf{T}^q_{\lambda,\eta}(X|W)$ since conditioned on V, $W - X - (Y_1, Y_2)$ forms a Markov chain. Furthermore, the continuity of $\mathsf{T}^q_{\lambda,\eta}(X|W)$ in η at $\eta = 1$ (see Remark 9) is also exploited;

(e) is by the definition of $\mathsf{F}^q_{\lambda_0,\alpha,\eta}(X)$ (while noting that $W - X - (Y_1, Y_2)$ forms a Markov chain), and because $\mathsf{F}^q_{\lambda,\eta}(X)$ is continuous at $\eta = 1$ (Remark 10);

(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 $X^\star_j \sim \mathcal{N}(0, K_j)$, $j = 1, 2, 3$, and $X' \sim \mathcal{N}(0, K - (K_1 + K_2 + K_3))$, such that (37) is satisfied. Furthermore, setting $X = X^\star_1 + X^\star_2 + X^\star_3 + X'$ not only attains $\tilde{V}^q_\eta(K)$, but it also simultaneously maximizes $\alpha\lambda_0 I(X; Y_1)$ and $\bar{\alpha}\lambda_0 I(X; Y_2)$. Relabeling $W^\star = X'$ and $V^\star = X^\star_3$ while taking the limit as $\eta \downarrow 1$, we have

$$\begin{aligned}
&\max_{(R_0,R_1,R_2)\in\mathcal{O}_K} \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2 \\
&\leq \alpha\lambda_0 I(X; Y_1) + \bar{\alpha}\lambda_0 I(X; Y_2) \\
&\quad + (\lambda_2 - \bar{\alpha}\lambda_0)I(X; Y_2|W^\star) - \alpha\lambda_0 I(X; Y_1|W^\star) \\
&\quad + \lambda_1 I(X; Y_1|W^\star, V^\star) - (\lambda_1 + \lambda_2)I(X; Y_2|W^\star, V^\star) \\
&\quad + \lambda_1 I(X; Y_2|W^\star, V^\star, X^\star_2) - \lambda_1 I(X; Y_1|W^\star, V^\star, X^\star_2) \\
&\leq \lambda_0\Big[\alpha I(W^\star; Y_1) + \bar{\alpha} I(W^\star; Y_2)\Big] \\
&\quad + \lambda_1\Big[I(X^\star_2; Y_1|W^\star, V^\star) - I(X^\star_2; Y_2|W^\star, V^\star)\Big] \\
&\quad + \lambda_2 I(V^\star; Y_2|W^\star).
\end{aligned} \tag{60}$$

Using Proposition 8, we set $U = \mathbf{X}_2^\star + \tilde{A}V^\star$ as before and obtain

$$I(\mathbf{X}_2^\star; \mathbf{Y}_1|W^\star, V^\star) - I(\mathbf{X}_2^\star; \mathbf{Y}_2|W^\star, V^\star)$$
$$= I(U^\star; \mathbf{Y}_1|W^\star) - I(U^\star; V^\star|W^\star) - I(\mathbf{X}_2^\star; \mathbf{Y}_2|W^\star, V^\star). \tag{61}$$

Inserting (61) into (60), yields

$$\max_{(R_0,R_1,R_2)\in\mathcal{O}_K} \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2$$
$$\leq \lambda_0\Big[\alpha I(W^\star; \mathbf{Y}_1) + \bar{\alpha} I(W^\star; \mathbf{Y}_2)\Big]$$
$$+ \lambda_1\Big[I(U^\star; \mathbf{Y}_1|W^\star) - I(U^\star; V^\star|W^\star) - I(U^\star; \mathbf{Y}_2|W^\star, V^\star)\Big]$$
$$+ \lambda_2 I(V^\star; \mathbf{Y}_2|W^\star)$$
$$\leq \sup_{\substack{(W,V,U)-\mathbf{X}-(\mathbf{Y}_1,\mathbf{Y}_2):\\ \mathbb{E}[\mathbf{XX}^\top]\preceq K}} \lambda_0\Big[\alpha I(W; \mathbf{Y}_1) + \bar{\alpha} I(W; \mathbf{Y}_2)\Big]$$
$$+ \lambda_1\Big[I(U; \mathbf{Y}_1|W) - I(U; V|W) - I(U; \mathbf{Y}_2|W, V)\Big]$$
$$+ \lambda_2 I(V; \mathbf{Y}_2|W). \tag{62}$$

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

$$\max_{(R_0,R_1,R_2)\in\mathcal{O}_K} \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2$$
$$\leq \min_{\alpha\in[0,1]} \sup_{\substack{(W,V,U)-\mathbf{X}-(\mathbf{Y}_1,\mathbf{Y}_2):\\ \mathbb{E}[\mathbf{XX}^\top]\preceq K}} \lambda_0\Big[\alpha I(W; \mathbf{Y}_1) + \bar{\alpha} I(W; \mathbf{Y}_2)\Big]$$
$$+ \lambda_1\Big[I(U; \mathbf{Y}_1|W) - I(U; V|W) - I(U; \mathbf{Y}_2|W, V)\Big]$$
$$+ \lambda_2 I(V; \mathbf{Y}_2|W). \tag{63}$$

Having (63), the desired equality $\mathcal{O}_K = \mathcal{I}_K$ is a consequence of the following Proposition.[4]

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

$$\min_{\alpha\in[0,1]} \sup_{\substack{(W,V,U)-\mathbf{X}-(\mathbf{Y}_1,\mathbf{Y}_2):\\ \mathbb{E}[\mathbf{XX}^\top]\preceq K}} \lambda_0\Big[\alpha I(W; \mathbf{Y}_1) + \bar{\alpha} I(W; \mathbf{Y}_2)\Big]$$
$$+ \lambda_1\Big[I(U; \mathbf{Y}_1|W) - I(U; V|W) - I(U; \mathbf{Y}_2|W, V)\Big]$$
$$+ \lambda_2 I(V; \mathbf{Y}_2|W)$$
$$= \sup_{\substack{(W,V,U)-\mathbf{X}-(\mathbf{Y}_1,\mathbf{Y}_2):\\ \mathbb{E}[\mathbf{XX}^\top]\preceq K}} \min_{\alpha\in[0,1]} \lambda_0\Big[\alpha I(W; \mathbf{Y}_1) + \bar{\alpha} I(W; \mathbf{Y}_2)\Big]$$
$$+ \lambda_1\Big[I(U; \mathbf{Y}_1|W) - I(U; V|W) - I(U; \mathbf{Y}_2|W, V)\Big]$$
$$+ \lambda_2 I(V; \mathbf{Y}_2|W)$$
$$= \sup_{\substack{(W,V,U)-\mathbf{X}-(\mathbf{Y}_1,\mathbf{Y}_2):\\ \mathbb{E}[\mathbf{XX}^\top]\preceq K}} \lambda_0 \cdot \min\{I(W; \mathbf{Y}_1), I(W; \mathbf{Y}_2)\}$$
$$+ \lambda_1\Big[I(U; \mathbf{Y}_1|W) - I(U; V|W) - I(U; \mathbf{Y}_2|W, V)\Big]$$
$$+ \lambda_2 I(V; \mathbf{Y}_2|W). \tag{64}$$

The proof of Proposition 9 is given in Appendix D. Now, noting that $(W, V, U) - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$ forms a Markov chain and $\mathbb{E}[\mathbf{XX}^\top] \preceq K$, we see that the triple

$$R_0 = \min\{I(W; \mathbf{Y}_1), I(W; \mathbf{Y}_2)\}$$
$$R_1 = I(U; \mathbf{Y}_1|W) - I(U; V|W) - I(U; \mathbf{Y}_2|W, V)$$

[4]Proposition 9 bears strong resemblance [16, Proposition 13] that was originally established in [27]

$$R_2 = I(V; \mathbf{Y}_2|W)$$

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

$$\max_{(R_0,R_1,R_2)\in\mathcal{O}_K} \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2$$
$$\leq \sup_{\substack{(W,V,U)-\mathbf{X}-(\mathbf{Y}_1,\mathbf{Y}_2):\\ \mathbb{E}[\mathbf{XX}^\top]\preceq K}} \lambda_0\Big[\alpha I(W; \mathbf{Y}_1) + \bar{\alpha} I(W; \mathbf{Y}_2)\Big]$$
$$+ \lambda_1\Big[I(U; \mathbf{Y}_1|W) - I(U; V|W) - I(U; \mathbf{Y}_2|W, V)\Big]$$
$$+ \lambda_2 I(V; \mathbf{Y}_2|W)$$
$$\leq \max_{(R_0,R_1,R_2)\in\mathcal{I}_K} \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2, \tag{65}$$

which implies that $\hat{\mathcal{I}}_K = \hat{\mathcal{C}}_K = \hat{\mathcal{O}}_K$ 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

$$\mathbf{X} = \mathbf{X}_1^\star + \mathbf{X}_2^\star + V^\star + W^\star, \tag{66}$$

where $\mathbf{X}_j^\star \sim \mathcal{N}(\mathbf{0}, K_j)$, $j = 1, 2$, $V^\star \sim \mathcal{N}(\mathbf{0}, K_3)$ and $W^\star \sim \mathcal{N}(\mathbf{0}, K - (K_1 + K_2 + K_3))$ are independent of each other, and $U = \mathbf{X}_2^\star + \tilde{A}V^\star$, where $\tilde{A}$ is the P-DPC matrix from Proposition 8. For the case when $\lambda_0 \leq \lambda_2$, (56)-(57) 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 V-A). 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 $\mathcal{C}_K$ as the union of rate triples $(R_0, R_1, R_2) \in \mathbb{R}_+^3$ satisfying:

$$R_0 \leq \min\Bigg\{\frac{1}{2}\log\left|\frac{I + G_1 K G_1^\top}{I + G_1(K_1 + K_2 + K_3)G_1^\top}\right|,$$
$$\frac{1}{2}\log\left|\frac{I + G_1 K G_1^\top}{I + G_1(K_1 + K_2 + K_3)G_1^\top}\right|\Bigg\} \tag{67a}$$

$$R_1 \leq \frac{1}{2}\log\left|\frac{I + G_1(K_1 + K_2)G_1^\top}{I + G_1 K_1 G_1^\top}\right|$$
$$- \frac{1}{2}\log\left|\frac{I + G_2(K_1 + K_2)G_2^\top}{I + G_2 K_1 G_2^\top}\right| \tag{67b}$$

$$R_2 \leq \frac{1}{2}\log\left|\frac{I + G_2(K_1 + K_2 + K_3)G_2^\top}{I + G_2(K_1 + K_2)G_2^\top}\right|, \tag{67c}$$

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

The region from (67) is further simplified using reasoning similar to this from Section V-A. First, (67b) indicates that the signal to User 1 is a sum of two independent zero mean Gaussian random vector with covariance K_1 and K_2. The signal that corresponds to K_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 an 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 1

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

$$
\begin{aligned}
&S_{\alpha\eta_1+\bar{\alpha}\eta_2}^q(\mathbf{X}) \\
&= \sup_{\substack{P_{V|\mathbf{X}}: \\ V-\mathbf{X}-(\mathbf{Y}_1,\mathbf{Y}_2)}} I(\mathbf{X}; \mathbf{Y}_2|V) - (\alpha\eta_1 + \bar{\alpha}\eta_2)I(\mathbf{X}; \mathbf{Y}_1|V) \\
&\leq \alpha \cdot \sup_{\substack{P_{V|\mathbf{X}}: \\ V-\mathbf{X}-(\mathbf{Y}_1,\mathbf{Y}_2)}} I(\mathbf{X}; \mathbf{Y}_2|V) - \eta_1 I(\mathbf{X}; \mathbf{Y}_1|V) \\
&\quad + \bar{\alpha} \cdot \sup_{\substack{P_{V|\mathbf{X}}: \\ V-\mathbf{X}-(\mathbf{Y}_1,\mathbf{Y}_2)}} I(\mathbf{X}; \mathbf{Y}_2|V) - \eta_2 I(\mathbf{X}; \mathbf{Y}_1|V) \\
&= \alpha S_{\eta_1}^q(\mathbf{X}) + \bar{\alpha} S_{\eta_2}^q(\mathbf{X}).
\end{aligned}
$$

Clearly, $S_\eta^q(\mathbf{X})$ is also bounded for every $\eta \in (0, 2)$ and by invoking Proposition 17 from [26, Ch. 5], we have that $S_\eta^q(\mathbf{X})$ is continuous inside every closed subinterval of $(0, 2)$, and in particular, at $\eta = 1$.

B. Proof of Proposition 4

Let $V - (\mathbf{X}_1, \mathbf{X}_2) - (\mathbf{Y}_{11}, \mathbf{Y}_{12}, \mathbf{Y}_{21}, \mathbf{Y}_{22})$ form a Markov chain. We have,

$$
\begin{aligned}
&t_{\lambda,\eta}^{q_1 \times q_2}(\mathbf{X}_1, \mathbf{X}_2|V) \\
&= \lambda_1 I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{11}, \mathbf{Y}_{12}|V) \\
&\quad - (\lambda_1+\lambda_2)I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{21}, \mathbf{Y}_{22}|V) + \lambda_1 S_\eta^{q_1 \times q_2}(\mathbf{X}_1, \mathbf{X}_2|V) \\
&= \lambda_1 \Big[I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{11}|V) + I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{12}|V, \mathbf{Y}_{11}) \Big] \\
&\quad - (\lambda_1+\lambda_2)\Big[I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{22}|V) + I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{21}|V, \mathbf{Y}_{22}) \Big] \\
&\quad + \lambda_1 S_\eta^{q_1 \times q_2}(\mathbf{X}_1, \mathbf{X}_2|V) \\
&\overset{(a)}{=} \lambda_1 \Big[I(\mathbf{X}_1; \mathbf{Y}_{11}|V, \mathbf{Y}_{22}) + I(\mathbf{X}_2; \mathbf{Y}_{12}|V, \mathbf{Y}_{11}) \\
&\quad + I(\mathbf{Y}_{11}; \mathbf{Y}_{22}|\mathbf{Y}_{11}) \Big] - (\lambda_1+\lambda_2)\Big[I(\mathbf{X}_2; \mathbf{Y}_{22}|V, \mathbf{Y}_{11}) \\
&\quad + I(\mathbf{X}_1; \mathbf{Y}_{21}|V, \mathbf{Y}_{22}) + I(\mathbf{Y}_{11}; \mathbf{Y}_{22}|V) \Big] \\
&\quad + \lambda_1 S_\eta^{q_1 \times q_2}(\mathbf{X}_1, \mathbf{X}_2|V) \\
&\overset{(b)}{\leq} \lambda_1 I(\mathbf{X}_1; \mathbf{Y}_{11}|V, \mathbf{Y}_{22}) - (\lambda_1+\lambda_2)I(\mathbf{X}_1; \mathbf{Y}_{21}|V, \mathbf{Y}_{22}) \\
&\quad + \lambda_1 S_\eta^{q_1}(\mathbf{X}_1|V, \mathbf{Y}_{22}) + \lambda_1 I(\mathbf{X}_2; \mathbf{Y}_{12}|V, \mathbf{Y}_{11}) \\
&\quad - (\lambda_1+\lambda_2)I(\mathbf{X}_2; \mathbf{Y}_{22}|V, \mathbf{Y}_{11}) + \lambda_1 S_\eta^{q_2}(\mathbf{X}_2|V, \mathbf{Y}_{11}) \\
&\quad - \lambda_2 I(\mathbf{Y}_{11}; \mathbf{Y}_{22}|V) \\
&\overset{(c)}{\leq} T_{\lambda,\eta}^{q_1}(\mathbf{X}_1|\mathbf{Y}_{22}) + T_{\lambda,\eta}^{q_2}(\mathbf{X}_2|\mathbf{Y}_{11}) - \lambda_2 I(\mathbf{Y}_{11}; \mathbf{Y}_{22}|V)
\end{aligned}
$$

$$
\begin{aligned}
&\overset{(d)}{\leq} T_{\lambda,\eta}^{q_1}(\mathbf{X}_1) + T_{\lambda,\eta}^{q_2}(\mathbf{X}_2) - \lambda_2 I(\mathbf{Y}_{11}; \mathbf{Y}_{22}|V) \\
&\leq T_{\lambda,\eta}^{q_1}(\mathbf{X}_1) + T_{\lambda,\eta}^{q_2}(\mathbf{X}_2), \tag{68}
\end{aligned}
$$

where:
(a) is since given V we have the Markov chain $(\mathbf{Y}_{11}, \mathbf{Y}_{21}) - \mathbf{X}_1 - \mathbf{X}_2 - (\mathbf{Y}_{12}, \mathbf{Y}_{22})$;
(b) follows from Proposition 2 by the definition of $S_\eta^q(\cdot|\cdot)$;
(c) is because $(V, \mathbf{Y}_{22}) - \mathbf{X}_1 - (\mathbf{Y}_{11}, \mathbf{Y}_{21})$ and $(V, \mathbf{Y}_{11}) - \mathbf{X}_2 - (\mathbf{Y}_{12}, \mathbf{Y}_{22})$ form Markov chains;
(d) follows by Remark 9 due to the Markov chains $\mathbf{Y}_{22} - \mathbf{X}_1 - (\mathbf{Y}_{11}, \mathbf{Y}_{21})$ and $\mathbf{Y}_{11} - \mathbf{X}_2 - (\mathbf{Y}_{12}, \mathbf{Y}_{22})$.

Now for $(V^\star, \mathbf{X}_1^\star, \mathbf{X}_2^\star)$, an end-to-end equality holds in (68). In particular, this implies that $I(\mathbf{Y}_{11}^\star; \mathbf{Y}_{22}^\star|V^\star) = 0$, i.e., that $\mathbf{Y}_{11}^\star - V^\star - \mathbf{Y}_{22}^\star$ forms a Markov chain. By [16, Proposition 2], we have that $\mathbf{X}_1^\star - V^\star - \mathbf{X}_2^\star$, which further implies the Markov chain

$$
(\mathbf{Y}_{11}^\star, \mathbf{Y}_{21}^\star) - \mathbf{X}_1^\star - V^\star - \mathbf{X}_2^\star - (\mathbf{Y}_{12}^\star, \mathbf{Y}_{22}^\star). \tag{69}
$$

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

$$
\begin{aligned}
&T_{\lambda,\eta}^{q_1}(\mathbf{X}_1^\star) \\
&= \lambda_1 I(\mathbf{X}_1^\star; \mathbf{Y}_{11}^\star|V^\star, \mathbf{Y}_{22}^\star) - (\lambda_1+\lambda_2)I(\mathbf{X}_1^\star; \mathbf{Y}_{21}^\star|V^\star, \mathbf{Y}_{22}^\star) \\
&\quad + \lambda_1 S_\eta^{q_1}(\mathbf{X}_1^\star|V^\star, \mathbf{Y}_{22}^\star) \\
&\overset{(a)}{=} \lambda_1 I(\mathbf{X}_1^\star; \mathbf{Y}_{11}^\star|V^\star) - (\lambda_1+\lambda_2)I(\mathbf{X}_1^\star; \mathbf{Y}_{21}^\star|V^\star) \\
&\quad + \lambda_1 S_\eta^{q_1}(\mathbf{X}_1^\star|V^\star), \tag{70}
\end{aligned}
$$

where (a) follows because (69) implies that the chain $\mathbf{Y}_{22}^\star - V^\star - \mathbf{X}_1^\star - (\mathbf{Y}_{11}^\star, \mathbf{Y}_{21}^\star)$ is Markov. Similarly, it can be shown that $T_{\lambda,\eta}^{q_2}(\mathbf{X}_2^\star) = t_{\lambda,\eta}^{q_2}(\mathbf{X}_2^\star|V^\star)$.

C. Proof of Proposition 5

Let $V - (\mathbf{X}_1, \mathbf{X}_2) - (\mathbf{Y}_{11}, \mathbf{Y}_{12}, \mathbf{Y}_{21}, \mathbf{Y}_{22})$ be a Markov chain. We have,

$$
\begin{aligned}
&f_{\boldsymbol{\lambda_0},\alpha,\eta}^{q_1 \times q_2}(\mathbf{X}_1, \mathbf{X}_2|V) \\
&= (\lambda_2 - \bar{\alpha}\lambda_0)I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{21}, \mathbf{Y}_{22}) \\
&\quad - \alpha\lambda_0 I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_{11}, \mathbf{Y}_{12}) + T_{\lambda,\eta}^{q_1 \times q_2}(\mathbf{X}_1, \mathbf{X}_2|V) \\
&\overset{(a)}{=} (\lambda_2 - \bar{\alpha}\lambda_0)I(\mathbf{X}_1; \mathbf{Y}_{21}|V, \mathbf{Y}_{22}) - \alpha\lambda_0 I(\mathbf{X}_1; \mathbf{Y}_{11}|V, \mathbf{Y}_{22}) \\
&\quad - \alpha\lambda_0 I(\mathbf{Y}_{11}; \mathbf{Y}_{22}|V) + (\lambda_2 - \bar{\alpha}\lambda_0)I(\mathbf{X}_2; \mathbf{Y}_{22}|V, \mathbf{Y}_{11}) \\
&\quad - \alpha\lambda_0 I(\mathbf{X}_2; \mathbf{Y}_{21}|V, \mathbf{Y}_{11}) + (\lambda_2 - \bar{\alpha}\lambda_0)I(\mathbf{Y}_{11}; \mathbf{Y}_{22}|V) \\
&\quad + T_{\lambda,\eta}^{q_1 \times q_2}(\mathbf{X}_1, \mathbf{X}_2|V) \\
&\overset{(b)}{\leq} (\lambda_2 - \bar{\alpha}\lambda_0)I(\mathbf{X}_1; \mathbf{Y}_{21}|V, \mathbf{Y}_{22}) - \alpha\lambda_0 I(\mathbf{X}_1; \mathbf{Y}_{11}|V, \mathbf{Y}_{22}) \\
&\quad + T_{\lambda,\eta}^{q_1}(\mathbf{X}_1|V, \mathbf{Y}_{22}) + (\lambda_2 - \bar{\alpha}\lambda_0)I(\mathbf{X}_2; \mathbf{Y}_{22}|V, \mathbf{Y}_{11}) \\
&\quad - \alpha\lambda_0 I(\mathbf{X}_2; \mathbf{Y}_{21}|V, \mathbf{Y}_{11}) + T_{\lambda,\eta}^{q_2}(\mathbf{X}_2|V, \mathbf{Y}_{11}) \\
&\quad - (\lambda_0 - \lambda_2)I(\mathbf{Y}_{11}; \mathbf{Y}_{22}|V) \\
&\overset{(c)}{\leq} F_{\boldsymbol{\lambda_0},\alpha,\eta}^{q_1}(\mathbf{X}_1|\mathbf{Y}_{22}) + F_{\boldsymbol{\lambda_0},\alpha,\eta}^{q_2}(\mathbf{X}_2|\mathbf{Y}_{11}) \\
&\quad - (\lambda_0 - \lambda_2)I(\mathbf{Y}_{11}; \mathbf{Y}_{22}|V) \\
&\overset{(d)}{\leq} F_{\boldsymbol{\lambda_0},\alpha,\eta}^{q_1}(\mathbf{X}_1) + F_{\boldsymbol{\lambda_0},\alpha,\eta}^{q_2}(\mathbf{X}_2) - (\lambda_0 - \lambda_2)I(\mathbf{Y}_{11}; \mathbf{Y}_{22}|V) \\
&\leq F_{\boldsymbol{\lambda_0},\alpha,\eta}^{q_1}(\mathbf{X}_1) + F_{\boldsymbol{\lambda_0},\alpha,\eta}^{q_2}(\mathbf{X}_2), \tag{71}
\end{aligned}
$$

where:
(a) is similar to step (a) in the proof of Proposition 4;

(b) uses Proposition 4 and the definition of $\mathsf{T}^q_{\lambda,\eta}(\cdot|\cdot)$;

(c) is because $(V, \mathbf{Y}_{22}) - \mathbf{X}_1 - (\mathbf{Y}_{11}, \mathbf{Y}_{21})$ and $(\tilde{V}, \mathbf{Y}_{11}) - \mathbf{X}_2 - (\mathbf{Y}_{12}, \mathbf{Y}_{22})$ form Markov chains;

(d) follows by Remark 10 due to the Markov chains $\mathbf{Y}_{22} - \mathbf{X}_1 - (\mathbf{Y}_{11}, \mathbf{Y}_{21})$ and $\mathbf{Y}_{11} - \mathbf{X}_2 - (\mathbf{Y}_{12}, \mathbf{Y}_{22})$.

For $(V^\star, \mathbf{X}_1^\star, \mathbf{X}_2^\star)$ that satisfy (35), an end-to-end equality holds in (71), implying that $I(\mathbf{Y}_{11}^\star; \mathbf{Y}_{22}^\star|V^\star) = 0$. Invoking [16, Proposition 2], we deduce that $\mathbf{X}_1^\star - V^\star - \mathbf{X}_2^\star$ forms a Markov chain, which further implies the Markov chain

$$(\mathbf{Y}_{11}^\star, \mathbf{Y}_{21}^\star) - \mathbf{X}_1^\star - V^\star - \mathbf{X}_2^\star - (\mathbf{Y}_{12}^\star, \mathbf{Y}_{22}^\star). \tag{72}$$

By the end-to-end equality in (68), we also have

$$\begin{aligned}
&\mathsf{F}^{q_1}_{\lambda_0,\alpha,\eta}(\mathbf{X}_1^\star)\\
&= (\lambda_2 - \bar{a}\lambda_0)I(\mathbf{X}_1^\star; \mathbf{Y}_{21}^\star|V^\star, \mathbf{Y}_{22}^\star) - \alpha\lambda_0 I(\mathbf{X}_1^\star; \mathbf{Y}_{11}^\star|V^\star, \mathbf{Y}_{22}^\star)\\
&\quad + \mathsf{T}^{q_1}_{\lambda,\alpha,\eta}(\mathbf{X}_1^\star|V^\star, \mathbf{Y}_{22}^\star)\\
&\overset{(a)}{=} (\lambda_2 - \bar{a}\lambda_0)I(\mathbf{X}_1^\star; \mathbf{Y}_{21}^\star|V^\star) - \alpha\lambda_0 I(\mathbf{X}_1^\star; \mathbf{Y}_{11}^\star|V^\star)\\
&\quad + \mathsf{T}^{q_1}_{\lambda,\alpha,\eta}(\mathbf{X}_1^\star|V^\star), \tag{73}
\end{aligned}$$

where (a) uses (72), which implies that $\mathbf{Y}_{22}^\star - V^\star - \mathbf{X}_1^\star - (\mathbf{Y}_{11}^\star, \mathbf{Y}_{21}^\star)$ forms a Markov chain. Similarly, it can be shown that $\mathsf{F}^{q_2}_{\lambda_0,\alpha,\eta}(\mathbf{X}_2^\star) = \mathsf{f}^{q_2}_{\lambda_0,\alpha,\eta}(\mathbf{X}_2^\star|V^\star)$.

D. Proof of Proposition 6

A key arguments in the proof of Proposition 6 is the continuity of the nested UCE $\mathsf{T}^q_{\lambda,\eta}(\mathbf{X}) = \left(\mathfrak{C}\mathsf{t}^q_{\lambda,\eta}\right)(\mathbf{X})$ in $P_{\mathbf{X}}$. We shall establish this continuity using [16, Proposition 21], which we reproduced as follows.

Proposition 10 (Boundedness and Continuity of UCE): Consider the space of all Borel probability distributions on $\mathbb{R}^t$ endowed with the topology induced by weak convergence.[5] Let $\{\mathbf{X}_n\}_{n \in \mathbb{N}}$ be a sequence of random variable that satisfies the following two properties: (i) $\exists\ p > 1$, $B \in \mathbb{R}$ such that $\mathbb{E}\|\mathbf{X}_n\|^p \leq B$, $\forall n \in \mathbb{N}$ (i.e., the sequence has a uniformly bounded pth moment); (ii) $\mathbf{X}_n \xrightarrow[n\to\infty]{\mathcal{D}} \mathbf{X}^\star$. If $g : \mathbb{R}^t \to \mathbb{R}$ is a bounded real-valued function that satisfies $g(\mathbf{X}_n) \xrightarrow[n\to\infty]{} g(\mathbf{X}^\star)$, then its UCE $G = \mathfrak{C}g$ is bounded and satisfies $G(\mathbf{X}_n) \xrightarrow[n\to\infty]{} G(\mathbf{X}^\star)$.

Before proving Proposition 6 at the end of this subsection, we first verify that Proposition 10 applies to $\mathsf{T}^q_{\lambda,\eta}(\mathbf{X})$. This is done by showing that $\mathsf{t}^q_{\lambda,\eta}(\mathbf{X})$ is bounded and continuous, as stated in the subsequent Lemma 2 and 3. 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 $\mathsf{t}^q_{\lambda,\eta}(\mathbf{X}) \leq B_{\lambda,\eta}$, for all $P_{\mathbf{X}}$.

Lemma 3 (Continuity of Nested Concave Envelopes): Consider the space of all Borel probability distributions on $\mathbb{R}^t$ endowed with the topology induced by weak convergence and let $\{\mathbf{X}_n\}_{n \in \mathbb{N}}$ be a sequence of random variable that satisfies:

[5] A sequence $\{X_n\}_{n \in \mathbb{N}}$ of real-valued random variables is said to *converge weakly* or, equivalently, converge 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 $X_n \xrightarrow[n\to\infty]{\mathcal{D}} X$

(i) $\exists\ p > 1$, $B \in \mathbb{R}$ such that $\mathbb{E}\|\mathbf{X}_n\|^p \leq B$, $\forall n$; (ii) $\mathbf{X}_n \xrightarrow[n\to\infty]{\mathcal{D}} \mathbf{X}^\star$. Then $\mathsf{t}^q_{\lambda,\eta}(\mathbf{X}_n) \xrightarrow[n\to\infty]{} \mathsf{t}^q_{\lambda,\eta}(\mathbf{X}^\star)$.

Based on Lemmas 2 and 3, [16, Proposition 21] states that $\mathsf{T}^q_{\lambda,\eta}(\mathbf{X}_n)$ is bounded and that it satisfies $\mathsf{T}^q_{\lambda,\eta}(\mathbf{X}_n) \xrightarrow[n\to\infty]{} \mathsf{T}^q_{\lambda,\eta}(\mathbf{X}^\star)$. The existence of a unique maximizer of $\tilde{V}_{\lambda_0,\alpha,\eta}(\mathbf{K})$ is established as follows. Let $\hat{\mathbf{K}} \succeq 0$ and define

$$\begin{aligned}
\tilde{\mathsf{v}}_{\lambda_0,\alpha,\eta}(\hat{\mathbf{K}}) &\triangleq \sup_{\mathbf{X}:\ \mathbb{E}[\mathbf{XX}^\top]=\hat{\mathbf{K}}} \mathsf{f}_{\lambda_0,\alpha,\eta}(\mathbf{X})\\
&= \sup_{\mathbf{X}:\ \mathbb{E}[\mathbf{XX}^\top]=\hat{\mathbf{K}}} (\lambda_2 - \bar{a}\lambda_0)I(\mathbf{X}; \mathbf{Y}_2)\\
&\quad - \alpha\lambda_0 I(\mathbf{X}; \mathbf{Y}_1) + \mathsf{T}^q_{\lambda,\eta}(\mathbf{X}). \tag{74}
\end{aligned}$$

Let $\{\mathbf{X}_n\}_{n \in \mathbb{N}}$ be a sequence of random variables with $\mathbb{E}[\mathbf{X}_n\mathbf{X}_n^\top] = \hat{\mathbf{K}}$, such that $\mathsf{f}_{\lambda_0,\alpha,\eta}(\mathbf{X}_n) \uparrow \tilde{\mathsf{v}}_{\lambda_0,\alpha,\eta}(\hat{\mathbf{K}})$, as $n \to \infty$. By [16, Proposition 17] and since $\mathbb{E}[\mathbf{X}_n\mathbf{X}_n^\top] = \hat{\mathbf{K}}$ for every $n \in \mathbb{N}$, we have that $\{\mathbf{X}_n\}_{n \in \mathbb{N}}$ is a tight sequence,[6] and that there exist an $\mathbf{X}^\star_{\hat{\mathbf{K}}}$ and a convergent subsequence $\{\mathbf{X}_{n_m}\}_{m \in \mathbb{N}}$ such that $\mathbf{X}_{n_m} \xrightarrow[m\to\infty]{\mathcal{D}} \mathbf{X}^\star_{\hat{\mathbf{K}}}$. Invoking [16, Proposition 18] once more we have that $h\left(\mathbf{Y}_{j,n_m}\right) \xrightarrow[m\to\infty]{} h\left(\mathbf{Y}^\star_{j,\hat{\mathbf{K}}}\right)$, for $j = 1, 2$, where $\mathbf{Y}_{1,n_m}, \mathbf{Y}_{2,n_m}, \mathbf{Y}^\star_{1,\hat{\mathbf{K}}}$ and $\mathbf{Y}^\star_{2,\hat{\mathbf{K}}}$ are the corresponding outputs. Thus,

$$\mathsf{f}_{\lambda_0,\alpha,\eta}(\mathbf{X}^\star_{\hat{\mathbf{K}}}) = \tilde{\mathsf{v}}_{\lambda_0,\alpha,\eta}(\hat{\mathbf{K}}). \tag{75}$$

By the definition of $\tilde{V}^q_{\lambda_0,\alpha,\eta}(\mathbf{K})$, we write

$$\begin{aligned}
\tilde{V}^q_{\lambda_0,\alpha,\eta}(\mathbf{K}) &= \sup_{\substack{(V,\mathbf{X}):\ \mathbb{E}[\mathbf{XX}^\top]\preceq\mathbf{K},\\ V-\mathbf{X}-(\mathbf{Y}_1,\mathbf{Y}_2)}} \mathsf{f}^q_{\lambda_0,\alpha,\eta}(\mathbf{X}|V)\\
&= \sup_{\substack{(V,\mathbf{X}):\ \mathbb{E}[\mathbf{XX}^\top]\preceq\mathbf{K},\\ V-\mathbf{X}-(\mathbf{Y}_1,\mathbf{Y}_2)}} \sum_v P(v)\mathsf{f}^q_{\lambda_0,\alpha,\eta}(\mathbf{X}|V=v).
\end{aligned}$$
$$\tag{76}$$

Since $\tilde{V}^q_{\lambda_0,\alpha,\eta}(\mathbf{K})$ 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 $\mathbf{X}^\star_{\mathbf{K}}$, for $\mathbf{K} \succeq 0$. Thus,

$$\tilde{V}^q_{\lambda_0,\alpha,\eta}(\mathbf{K}) = \sup_{\substack{\{\alpha_i\},\{\hat{\mathbf{K}}_i\}:\ \alpha_i\geq 0\\ \sum_i \alpha_i=1,\ \sum_i \alpha_i\hat{\mathbf{K}}_i\preceq\mathbf{K}}} \sum_i \alpha_i \tilde{\mathsf{v}}^q_{\lambda_0,\alpha,\eta}(\hat{\mathbf{K}}). \tag{77}$$

It takes $\frac{t(t+1)}{2}$ constraints to preserve the covariance matrix (due to its symmetry) and one other constraint to preserve $\sum_i \alpha_i \tilde{\mathsf{v}}^q_{\lambda_0,\alpha,\eta}(\hat{\mathbf{K}})$. Hence, by using the Bunt-Caratheodory theorem [28], we can restrict ourselves to convex combinations of at most $m \triangleq \frac{t(t+1)}{2} + 1$ points, i.e.,

$$\tilde{V}^q_{\lambda_0,\alpha,\eta}(\mathbf{K}) = \sup_{\substack{\{\alpha_i\},\{\hat{\mathbf{K}}_i\}:\ \alpha_i\geq 0\\ \sum_{i=1}^m \alpha_i=1,\ \sum_{i=1}^m \alpha_i\hat{\mathbf{K}}_i\preceq\mathbf{K}}} \sum_{i=1}^m \alpha_i \tilde{\mathsf{v}}^q_{\lambda_0,\alpha,\eta}(\hat{\mathbf{K}}). \tag{78}$$

Consider any sequence of convex combinations $\left\{\{\alpha_i^{(n)}\}_{i\in[m]}, \{\mathbf{K}_i^{(n)}\}_{i\in[m]}\right\}_{n\in\mathbb{N}}$ that approaches the supremum

[6] As defined in [16], a sequence of random variables $\{\mathbf{X}_n\}_{n\in\mathbb{N}}$ taking values in $\mathbb{R}^t$ is *tight* if for every $\epsilon > 0$ there exists a compact set $\mathcal{C}_\epsilon \subset \mathbb{R}^t$, such that $\mathbb{P}(\mathbf{X}_n \notin \mathcal{C}_\epsilon) \leq \epsilon$, $\forall n \in \mathbb{N}$.

as $n \to \infty$. The compactness of the m-dimensional simplex implies that $\alpha_i^{(n)} \xrightarrow[n \to \infty]{} \alpha_i^\star$, for all $i \in [m]$. Furthermore, we have the following property of the limiting points $\alpha_i^\star$; see Appendix C-C for the proof.

Lemma 4: For any $i \in [m]$, if $\alpha_i^\star = 0$ then $\alpha_i^{(n)} \tilde{v}_{\boldsymbol{\lambda_0}, \alpha, \eta}^q \left(K_i^{(n)} \right) \xrightarrow[n \to \infty]{} 0$.

Based on Lemma 4, we assume that $\alpha^\star \triangleq \min_{i \in [m]} \alpha_i^\star > 0$, which implies that $K_i^{(n)} \preceq \frac{2}{\alpha^\star} K$ uniformly in $i \in [m]$, for large enough n. Hence, for each $i \in [m]$ we can find a convergent subsequence $\left\{ K_i^{(n_k)} \right\}_{k \in \mathbb{N}}$, such that $K_i^{(n_k)} \xrightarrow[k \to \infty]{} K_i^\star$. Putting these together, we have

$$\tilde{V}_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(K) = \sum_{i=1}^m \alpha_i^\star \tilde{v}_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(K_i^\star), \tag{79}$$

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

E. Proof of Proposition 7

Denote $q_{\mathbf{Y}_1, \mathbf{Y}_2 | \mathbf{X}} \triangleq q$ and consider the two-letter BC $q(\mathbf{y}_{11}, \mathbf{y}_{21} | \mathbf{x}_1) \times q(\mathbf{y}_{12}, \mathbf{y}_{22} | \mathbf{x}_2)$. We have

$$2 \tilde{V}_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(K) \overset{(a)}{=} f_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(\mathbf{X}_1 | V_1) + f_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(\mathbf{X}_2 | V_2)$$

$$\overset{(b)}{=} f_{\boldsymbol{\lambda_0}, \alpha, \eta}^{q \times q}(\mathbf{X}_1, \mathbf{X}_2 | V_1, V_2)$$

$$\overset{(c)}{=} f_{\boldsymbol{\lambda_0}, \alpha, \eta}^{q \times q}(\mathbf{X}_{\theta_1}, \mathbf{X}_{\theta_2} | \tilde{V})$$

$$\overset{(d)}{\leq} F_{\boldsymbol{\lambda_0}, \alpha, \eta}^{q \times q}(\mathbf{X}_{\theta_1}, \mathbf{X}_{\theta_2})$$

$$\overset{(e)}{\leq} F_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(\mathbf{X}_{\theta_1}) + F_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(\mathbf{X}_{\theta_2})$$

$$\overset{(f)}{\leq} \tilde{V}_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(K) + \tilde{V}_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(K) = 2 \tilde{V}_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(K), \tag{80}$$

where:
(a) is because $P_{V, \mathbf{X}}^\star$ achieves $\tilde{V}_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(K)$;
(b) uses the independence of $(V_1, \mathbf{X}_1)$ and $(V_2, \mathbf{X}_2)$;
(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 $\tilde{V}_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(K)$, while noting that the independence of $\mathbf{X}_{v_1}$ and $\mathbf{X}_{v_2}$, for every $v_1 \neq v_2 \in \mathcal{V}$, implies that

$$\mathbb{E}\left[\mathbf{X}_{\theta_1} \mathbf{X}_{\theta_1}^\top \right] = \mathbb{E}\left[\frac{1}{2} \mathbb{E}\left[\mathbf{X}_{V_1} \mathbf{X}_{V_1}^\top \Big| \tilde{V} \right] + \frac{1}{2} \mathbb{E}\left[\mathbf{X}_{V_2} \mathbf{X}_{V_2}^\top \Big| \tilde{V} \right] \right]$$

$$= \mathbb{E}\left[\mathbf{X}_{\theta_2} \mathbf{X}_{\theta_2}^\top \right] \tag{81}$$

and

$$\mathbb{E}\left[\mathbf{X}_{\theta_2} \mathbf{X}_{\theta_2}^\top \right] = \sum_v P_V^\star(v) K_v = \mathbb{E}\left[\mathbb{E}\left[\mathbf{X} \mathbf{X}^\top | V \right] \right] \preceq K,$$

where we have denoted $K_v \triangleq \mathbb{E}\left[\mathbf{X}_v \mathbf{X}_v^\top \right]$.

Since the extremes of the chain of inequalities in (80) match, all inequalities are, in fact, equalities. Equality in step (d) implies that $P_{\tilde{V} | \mathbf{X}_{\theta_1}, \mathbf{X}_{\theta_2}}$ achieves $F_{\boldsymbol{\lambda_0}, \alpha, \eta}^{q \times q}(\mathbf{X}_{\theta_1}, \mathbf{X}_{\theta_2})$. Furthermore, by Proposition 5, since (d) and (e) are equalities

we have that $\mathbf{X}_{\theta_1} - \tilde{V} - \mathbf{X}_{\theta_2}$, and that $P_{\tilde{V} | \mathbf{X}_{\theta_1}}$ and $P_{\tilde{V} | \mathbf{X}_{\theta_2}}$ achieve $F_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(\mathbf{X}_{\theta_1})$ and $F_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(\mathbf{X}_{\theta_1})$, respectively. Finally, equality in (f) means that $\tilde{V}_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(K) = F_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(\mathbf{X}_{\theta_j} | \tilde{V}) = f_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(\mathbf{X}_{\theta_j} | \tilde{V})$, for $j = 1, 2$.

F. Proof of Theorem 5

As a consequence of Proposition 7, for any fixed $(v_1, v_2) \in \mathcal{V}^2$, $\mathbf{X}_{v_1} + \mathbf{X}_{v_2}$ and $\mathbf{X}_{v_1} - \mathbf{X}_{v_2}$ are independent. Combined with $\mathbf{X}_{v_1}$ and $\mathbf{X}_{v_2}$ being independent zero mean random variables, Corollary 3 in Appendix I-A of [16] implies that $\mathbf{X}_{v_1}$ and $\mathbf{X}_{v_2}$ are Gaussian random vectors with the same covariance matrix. Since the pair $(v_1, v_2) \in \mathcal{V}^2$ is arbitrary, we see that all Gaussian vectors $\{\mathbf{X}_v\}_{v \in \mathcal{V}}$ have the same covariance matrix. Furthermore, we may assume that $\{\mathbf{X}_v\}_{v \in \mathcal{V}}$ 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 $K^\star$, it clearly satisfies $K^\star \preceq K$. Letting $\mathbf{X}^\star \sim \mathcal{N}(\mathbf{0}, K^\star)$, we have

$$\tilde{V}_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(K) \overset{(a)}{=} f_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(\mathbf{X} | V)$$

$$\overset{(b)}{=} \sum_{i=1}^m P_V^\star(v_i) f_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(\mathbf{X}_{v_i})$$

$$\overset{(c)}{=} \sum_{i=1}^m P_V^\star(v_i) f_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(\mathbf{X}^\star)$$

$$= f_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(\mathbf{X}^\star), \tag{82}$$

where (a) follows since $(V, \mathbf{X}) \sim P_{V, \mathbf{X}}^\star$ attains $\tilde{V}_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(K)$ (Proposition 6), (b) follows by the definition of $\mathbf{X}_v$ in the statement of Proposition 7, while (c) follows since $\mathbf{X}^\star$ and $\mathbf{X}_{v_i}$ 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 $\mathbf{X}$ is a maximizer, i.e., $\tilde{V}_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(K) = f_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(\mathbf{X})$, it must be Gaussian. Let $\mathbf{X}_1$ and $\mathbf{X}_1$ be two i.i.d. copies of $\mathbf{X}$. Applying Proposition 7 while taking V to be a constant, we obtain that $\mathbf{X}_1 + \mathbf{X}_2$ and $\mathbf{X}_1 - \mathbf{X}_2$ are also independent. Hence, by [16, Corollary 3], $\mathbf{X}$ is Gaussian.

Next, suppose that $\tilde{V}_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(K)$ has two independent Gaussian maximizers denoted by $\mathbf{A}_1 \sim \mathcal{N}(\mathbf{0}, K_1)$ and $\mathbf{A}_2 \sim \mathcal{N}(\mathbf{0}, K_2)$, such that $K_1, K_2 \preceq$ and $K_1 \neq K_2$. Let $(V, \mathbf{X})$ be a pair of random variables, such that $V \sim \text{Ber}\left(\frac{1}{2}\right)$ on $\mathcal{V} = \{1, 2\}$, $\mathbf{X} | \{V = 1\} \sim \mathcal{N}(\mathbf{0}, K_1)$ and $\mathbf{X} | \{V = 2\} \sim \mathcal{N}(\mathbf{0}, K_2)$. Note that $(V, \mathbf{X})$ also attains $\tilde{V}_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(K)$. Taking $v_1 = 1$ and $v_2 = 2$, Proposition 7 implies that $\mathbf{A}_1 + \mathbf{A}_2$ and $\mathbf{A}_1 - \mathbf{A}_2$ are independent, which contradict Corollary 3 from [16] as $K_1 \neq K_2$.

G. Proof of Corollary 3

By Theorem 5, there is an $\mathbf{X}^\star \sim \mathcal{N}(\mathbf{0}, K^\star)$, such that $K^\star \preceq K$ and $f_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(\mathbf{X}^\star) = \tilde{V}_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(K)$. Let $\mathbf{X}' \sim \mathcal{N}(\mathbf{0}, K - K^\star)$ be independent of $\mathbf{X}^\star$ and set $\mathbf{X} = \mathbf{X}^\star + \mathbf{X}'$. Thus $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, K)$ and by definition we have $F_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(\mathbf{X}) \leq \tilde{V}_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(K)$.

On the other hand,

$$F_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(\mathbf{X}) = \sup_{\substack{P_{V | \mathbf{X}}: \\ V - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)}} f_{\boldsymbol{\lambda_0}, \alpha, \eta}^q(\mathbf{X} | V)$$

$$\overset{(a)}{\geq} \mathsf{f}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}|\mathbf{X}')$$

$$\overset{(b)}{=} \mathsf{f}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}^\star)$$

$$= \tilde{V}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathrm{K}), \tag{83}$$

where (a) follows since $\mathbf{X}' - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$ forms a Markov chain, while (b) follows because $\mathbf{X}|\{\mathbf{X}' = \mathbf{x}'\} \sim \mathbf{X}^\star + \mathbf{x}'$. Thus,

$$\begin{aligned}
\tilde{V}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathrm{K}) &= \mathsf{F}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}) \\
&= \mathsf{f}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}^\star) \\
&= (\lambda_2 - \bar{a}\lambda_0)I(\mathbf{X}^\star; \mathbf{Y}_2) - \alpha\lambda_0 I(\mathbf{X}^\star; \mathbf{Y}_1) \\
&\quad + \mathsf{T}^q_{\boldsymbol{\lambda},\eta}(\mathbf{X}^\star).
\end{aligned} \tag{84}$$

By Theorem 4, one can decompose $\mathbf{X}^\star$ into independent $\mathbf{X}_1^\star \sim \mathcal{N}(\mathbf{0}, \mathrm{K}_1)$ and $\mathbf{X}_2^\star \sim \mathcal{N}(\mathbf{0}, \mathrm{K}_2)$, with $\mathrm{K}_1 + \mathrm{K}_2 \preceq \mathrm{K}^\star$, such that

$$\mathsf{T}^q_{\boldsymbol{\lambda},\eta}(\mathbf{X}^\star) = \mathsf{t}^q_{\boldsymbol{\lambda},\eta}(\mathbf{X}_1^\star + \mathbf{X}_2^\star) = \hat{V}^q_{\boldsymbol{\lambda},\eta}(\mathrm{K}^\star), \tag{85}$$

and

$$\mathsf{S}^q_\eta(\mathbf{X}_1^\star + \mathbf{X}_2^\star) = \mathsf{s}^q_\eta(\mathbf{X}_1^\star) = V^q_\eta(\mathrm{K}_1 + \mathrm{K}_2). \tag{86}$$

The proof of existence is concluded by setting $\mathbf{X}_3^\star \sim \mathcal{N}(\mathbf{0}, \mathrm{K}_3)$, where $\mathrm{K}_3 = \mathrm{K}^\star - (\mathrm{K}_1 + \mathrm{K}_2)$ and noting that $\mathbf{X} = \mathbf{X}_1^\star + \mathbf{X}_2^\star + \mathbf{X}_3^\star + \mathbf{X}' \sim \mathcal{N}(\mathbf{0}, \mathrm{K})$ and $\mathbf{X}^\star = \mathbf{X}_1^\star + \mathbf{X}_2^\star + \mathbf{X}_3^\star \sim \mathcal{N}(\mathbf{0}, \mathrm{K}^\star)$, which implies that (37) 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 [21] 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).

APPENDIX A
DERIVATION OF INNER AND OUTER BOUNDS

A. Outer Bounds 1 and 3

We first establish Bound 3 as an outer bound on the secrecy-capacity region of the setting with a common message, and then use it to establish Bound 1 as an outer bound on the region without a common message.

The result of [23, Th. 3] characterizes an outer bound $\mathcal{R}_O(L_1, L_2)$ on the (L_1, L_2)-leakage-capacity region of a DM-BC with common and private messages, for some leakage thresholds $(L_1, L_2) \in \mathbb{R}^2_+$. Setting $L_1 = 0$ and letting $L_2 \to \infty$ in $\mathcal{R}_O(L_1, L_2)$ (which corresponds to M_1 being confidential and M_2 not being subject to any secrecy requirements), we have that the closure of the union of rate triples $(R_0, R_1, R_2) \in \mathbb{R}^3_+$ satisfying:

$$R_0 \leq \min\{I(W; Y_1), I(W; Y_2)\} \tag{87a}$$

$$R_1 \leq I(U; Y_1|W, V) - I(U; Y_2|W, V) \tag{87b}$$

$$R_1 \leq I(U; Y_1|W) - I(U; Y_2|W) \tag{87c}$$

$$R_0 + R_2 \leq I(V; Y_2|W) + \min\{I(W; Y_1), I(W; Y_2)\} \tag{87d}$$

$$R_0 + R_1 + R_2 \leq I(U; Y_1|W) + I(V; Y_2|W, U) \\ + \min\{I(W; Y_1), I(W; Y_2)\} \tag{87e}$$

over all $(W, U, V) - X - (Y_1, Y_2)$, is an outer bound on $\mathcal{C}$. By removing the rate bounds in (87c) and (87e) from $\mathcal{R}_O(0, \infty)$, one recovers the region $\mathcal{O}$ from (54). Clearly $\mathcal{R}_O(0, \infty) \subseteq \mathcal{O}$, which shows that $\mathcal{C} \subseteq \mathcal{O}$ and establishes Bound 3.

When there is no common message, one obtains Bound 1, i.e., that $\hat{\mathcal{C}} \subseteq \hat{\mathcal{O}}$, by setting $R_0 = 0$ into Bound 3. This follows by noting that

$$I(V; Y_2|W) + \min\{I(W; Y_1), I(W; Y_2)\} \leq I(W, V; Y_2), \tag{88}$$

and defining $\tilde{V} = (W, V)$.

B. Inner Bounds 2 and 4

Referring to [23, Th. 1], we have $\mathcal{R}_I(0, \infty)$ as an inner bound on $\mathcal{C}$, where $\mathcal{R}_I(0, \infty)$ is the closure of the union of rate triples $(R_0, R_1, R_2) \in \mathbb{R}^3_+$ satisfying:

$$R_0 \leq \min\{I(W; Y_1), I(W; Y_2)\} \tag{89a}$$

$$R_1 \leq I(U; Y_1|W) - I(U; V|W) - I(U; Y_2|W, V) \tag{89b}$$

$$R_0 + R_1 \leq I(U; Y_1|W) + \min\{I(W; Y_1), I(W; Y_2)\} \tag{89c}$$

$$R_0 + R_2 \leq I(V; Y_2|W) + \min\{I(W; Y_1), I(W; Y_2)\} \tag{89d}$$

$$R_0 + R_1 + R_2 \leq I(U; Y_1|W) + I(V; Y_2|W) - I(U; V|W) \tag{89e}$$

$$+ \min\{I(W; Y_1), I(W; Y_2)\} \tag{89f}$$

over all $(W, U, V) - X - (Y_1, Y_2)$. The inclusion $\mathcal{I} \subseteq \mathcal{R}_I(0, \infty)$ immediately follows by noting that if (R_1, R_2) satisfy (55) then they also satisfy (89). A simple consequence of the above is that Bound 2 is an inner bound on the secrecy-capacity region without a common message, which follows by setting $R_0 = 0$ and $W = 0$ into Bound 4.

APPENDIX B
PROOF OF LEMMA 1

First notice that (40) are supporting hyperplanes of $\hat{\mathcal{O}}_K$ and that the points $(\mathcal{H}_1^K, 0)$ and $(0, \mathcal{H}_2^K)$ are on its boundary. Furthermore, $R_1 \geq 0$, $R_2 \geq 0$ and $R_2 \leq \mathcal{H}_2^K$ are also supporting hyperplanes of $\hat{\mathcal{I}}_K$, and since $(0, \mathcal{H}_2^K) \in \hat{\mathcal{I}}_K$, it is a boundary point of $\hat{\mathcal{I}}_K$. Note that $\mathcal{H}_1^K$ describes the secrecy-capacity of the MIMO Gaussian WTC, where User 1 serves as the legitimate receiver and User 2 as the eavesdropper. Therefore (see [7]–[9]),

$$\mathcal{H}_1^K = \frac{1}{2} \max_{0 \preceq K^\star \preceq K} \log \left| I + G_1 K^\star G_1^\top \right| - \log \left| I + G_2 K^\star G_2^\top \right|. \tag{90}$$

To see that $(\mathcal{H}_1^K, 0)$ is also in $\hat{\mathcal{I}}_K$ consider the following. For every $0 \preceq K^\star \preceq K$, let X_1 and X_2 be independent Gaussian random vectors with covariances $K^\star$ and $K - K^\star$, respectively. Set

$$U = X_1 + A X_2$$
$$V = X_2$$
$$X = X_1 + X_2,$$

where $A = K^\star G_1^\top \left[I + G_1 K^\star G_1^\top \right]^{-1}$ is the precoding matrix for suppressing V from Y_1 [20, Th. 1]. Evaluating the mutual information terms on the RHS of (39a), we first have

$$I(U; Y_1) - I(U; V) = I(X; Y_1 | V) = \frac{1}{2} \log \left| I + G_1 K^\star G_1^\top \right|. \tag{91}$$

Moreover,

$$\begin{aligned} I(U; Y_2 | V) &= I\left(X_1 + A X_2; G_2(X_1 + X_2) + Z_2 \big| X_2\right) \\ &= I(X_1; G_2 X_1 + Z_2 | X_2) \\ &\overset{(a)}{=} I(X_1; G_2 X_1 + Z_2) \\ &= \frac{1}{2} \log \left| I + G_2 K^\star G_2^\top \right|, \end{aligned} \tag{92}$$

where (a) follows because (X_1, Z_2) and X_2 are independent. Combining (91) with (92) yields

$$\begin{aligned} &I(U; Y_1) - I(U; V) - I(U; Y_2 | V) \\ &= \frac{1}{2} \log \left| I + G_1 K^\star G_1^\top \right| - \frac{1}{2} \log \left| I + G_2 K^\star G_2^\top \right|, \end{aligned} \tag{93}$$

which implies that $(\mathcal{H}_1^K, 0) \in \hat{\mathcal{I}}_K$. Furthermore, since $(\mathcal{H}_1^K, 0)$ is on the boundary of $\hat{\mathcal{O}}_K$ and $\hat{\mathcal{I}}_K \subseteq \hat{\mathcal{O}}_K$, $(\mathcal{H}_1^K, 0)$ must also be a boundary point of $\hat{\mathcal{I}}_K$, and therefore, $R_1 \leq \mathcal{H}_1^K$ is a supporting hyperplane of $\hat{\mathcal{I}}_K$.

APPENDIX C
PROOF OF AUXILIARY RESULTS FOR
PROPOSITION 6 - LEMMAS 2 AND 3

A. Proof of Lemma 2

By Theorem 4, we have that if $\mathbb{E}[XX^\top] \preceq K$, then

$$\hat{V}_{\lambda, \eta}^q(K) = t_{\lambda, \eta}^q(X^\star), \tag{94}$$

where $X^\star = X_1^\star + X_2^\star$, and $X_1^\star$ and $X_2^\star$ are independent random variables with $X_j^\star \sim \mathcal{N}(0, K_j)$, for $j = 1, 2$, such that $K^\star \triangleq$

$K_1 + K_2 \preceq K$. Furthermore, by the definition of the UCE and the definition of $\hat{V}_{\lambda, \eta}^q(K)$, (94) implies that for every $X \sim P_X$ with $\mathbb{E}[XX^\top] \preceq K$, we have

$$t_{\lambda, \eta}^q(X) \leq T_{\lambda, \eta}^q(X) \leq \hat{V}_{\lambda, \eta}^q(K) = t_{\lambda, \eta}^q(X^\star). \tag{95}$$

Thus,

$$\begin{aligned} &\sup_{X: \, \mathbb{E}[XX^\top] \preceq K} t_{\lambda, \eta}^q(X) \\ &\quad \leq \lambda_1 I(X^\star; Y_1) - (\lambda_1 + \lambda_2) I(X^\star; Y_2) + \lambda_1 S_\eta^q(X^\star) \\ &\quad \overset{(a)}{=} \lambda_1 I(X^\star; Y_1) - (\lambda_1 + \lambda_2) I(X^\star; Y_2) + \lambda_1 s_\eta^q(X_1^\star) \\ &\quad \overset{(b)}{\leq} \lambda_1 I(X^\star; Y_1) - (\lambda_1 + \lambda_2) I(X^\star; Y_2) + \lambda_1 C_\eta, \end{aligned} \tag{96}$$

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

$$\begin{aligned} &\sup_X t_{\lambda, \eta}^q(X) \\ &\leq \sup_{0 \preceq K: \, X \sim \mathcal{N}(0, K)} \lambda_1 I(X; Y_1) - (\lambda_1 + \lambda_2) I(X; Y_2) + \lambda_1 C_\eta. \end{aligned} \tag{97}$$

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

$$\begin{aligned} &2\lambda_1 I(X; Y_1) - 2(\lambda_1 + \lambda_2) I(X; Y_2) \\ &= \lambda_1 \log |I + G_1 K G_1^\top| - (\lambda_1 + \lambda_2) \log |I + G_2 K G_2^\top| \\ &= \lambda_1 \log |I + K G_1 G_1^\top| - (\lambda_1 + \lambda_2) \log |I + K G_2 G_2^\top| \\ &\overset{(a)}{=} -\lambda_1 \log |\Sigma_1| + (\lambda_1 + \lambda_2) \log |\Sigma_2| \\ &\qquad + \lambda_1 \left(\log |\Sigma_1 + K| - \lambda \log |\Sigma_2 + K| \right), \end{aligned} \tag{98}$$

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 ith smallest eigenvalue of the symmetric matrix $A \in \mathbb{R}^{t \times t}$, we have

$$\mu_i(A) = \min_{L_i} \max_{0 \neq u \in L_i} \frac{u^\top A u}{u^\top u} = \max_{L_{t+1-i}} \min_{0 \neq u \in L_{t+1-i}} \frac{u^\top A u}{u^\top u}, \tag{99}$$

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

$$\mu_1(A) = \max_{L_t} \min_{0 \neq u \in L_t} \frac{u^\top A u}{u^\top u} = \min_{0 \neq u \in L_t} \frac{u^\top A u}{u^\top u} \tag{100a}$$

$$\mu_t(A) = \min_{L_t} \max_{0 \neq u \in L_t} \frac{u^\top A u}{u^\top u} = \max_{0 \neq u \in L_t} \frac{u^\top A u}{u^\top u}. \tag{100b}$$

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

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

$$\begin{cases} \geq \min_{L_i} \max_{0 \neq \mathbf{u} \in L_i} \left(\frac{\mathbf{u}^\top K \mathbf{u}}{\mathbf{u}^\top \mathbf{u}} + \mu_1(\Sigma_j) \right) = \mu_i(K) + \mu_1(\Sigma_j), \\ \leq \min_{L_i} \max_{0 \neq \mathbf{u} \in L_i} \left(\frac{\mathbf{u}^\top K \mathbf{u}}{\mathbf{u}^\top \mathbf{u}} + \mu_t(\Sigma_j) \right) = \mu_i(K) + \mu_t(\Sigma_j). \end{cases}$$
(101)

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

$$\mu_i(K) + \mu_1(\Sigma_j) \leq \mu_i(K + \Sigma_j) \leq \mu_i(K) + \mu_t(\Sigma_j), \quad (102)$$

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

$$\begin{aligned} &\log |\Sigma_1 + K| - \lambda \log |\Sigma_2 + K| \\ &= \sum_{i=1}^{t} \log \left(\frac{\mu_i(K + \Sigma_1)}{\mu_i(K + \Sigma_2)^\lambda} \right) \\ &\stackrel{(a)}{\leq} \sum_{i=1}^{t} \log \left(\frac{\mu_i(K) + \mu_t(\Sigma_1)}{\left(\mu_i(K) + \mu_1(\Sigma_2)\right)^\lambda} \right) \\ &\leq t \cdot \max_i \log \left(\frac{\mu_i(K) + \mu_t(\Sigma_1)}{\left(\mu_i(K) + \mu_1(\Sigma_2)\right)^\lambda} \right) \\ &\stackrel{(b)}{\leq} t \cdot \log \left(\frac{\mu^\star + \mu_t(\Sigma_1)}{\left(\mu^\star + \mu_1(\Sigma_2)\right)^\lambda} \right), \end{aligned}$$
(103)

where (a) follows from by (101), and (b) is by setting $\mu^\star = \max \left\{ 0, \frac{1}{1-\lambda} \left(\mu_1(\Sigma_2) - \lambda \mu_t(\Sigma_1) \right) \right\}$ and noting that $\mu_i(K) \geq 0$ and that the derivative of the function $c(x) \triangleq \log \left(x + \mu_t(\Sigma_1) \right) - \lambda \log \left(x + \mu_1(\Sigma_2) \right)$ is zero at $x = \mu^\star$, negative when $x > \mu^\star$ and positive when $x < \mu^\star$. Since $\mu_t(\Sigma_1), \mu_1(\Sigma_2) > 0$ (which holds since the positive semi-definite matrices Σ_1 and Σ_2 are invertible), we conclude that for every $\mathbf{X} \sim P_\mathbf{X}$

$$\begin{aligned} t^q_{\lambda,\eta}(\mathbf{X}) &\leq -\lambda_1 \log |\Sigma_1| + (\lambda_1 + \lambda_2) \log |\Sigma_2| \\ &\quad + \lambda_1 \cdot t \cdot \log \left(\frac{\mu^\star + \mu_t(\Sigma_1)}{\left(\mu^\star + \mu_1(\Sigma_2)\right)^{\frac{\lambda_1+\lambda_2}{\lambda_1}}} \right) \\ &\triangleq B_{\lambda,\eta} < \infty. \end{aligned}$$
(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, Th. 5], which, respectively, imply that $S^q_\eta(\mathbf{X}_n) \xrightarrow[n \to \infty]{} S^q_\eta(\mathbf{X}^\star)$, and that $h(\mathbf{Y}_{j,n}) \xrightarrow[n \to \infty]{} h(\mathbf{Y}_j^\star)$, for $j = 1, 2$. Here $\mathbf{Y}_{1,n}$ and $\mathbf{Y}_{2,n}$ are the outputs of the MIMO Gaussian BC with input $\mathbf{X}_n$.

C. Proof of Lemma 4

Let $i \in [m]$ be such that $\alpha_i^\star = 0$ and note that for every $\hat{K} \succeq 0$, we have

$$\begin{aligned} \tilde{v}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\hat{K}) &\stackrel{(a)}{=} f^q_{\boldsymbol{\lambda_0},\alpha,\eta}(\mathbf{X}^\star_{\hat{K}}) \\ &\stackrel{(b)}{\leq} \lambda_2 I(\mathbf{X}^\star_{\hat{K}}; \mathbf{Y}_2) + \hat{B}_{\boldsymbol{\lambda_0},\alpha,\eta} \\ &\leq \frac{\lambda_2}{2} \log |I + G_2 \hat{K} G_2^\top| + \hat{B}_{\boldsymbol{\lambda_0},\alpha,\eta}, \quad (105) \end{aligned}$$

where (a) follows from (75) and the non-negativity of mutual information, while (b) is since $T^q_{\lambda,\eta}(\mathbf{X})$ 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 4 is immediate). Since $\alpha_i^{(n_k)} K_i^{(n_k)} \preceq K$ and $\alpha_i^{(n_k)} > 0$, we obtain

$$K_i^{(n_k)} \preceq \frac{1}{\alpha_i^{(n_k)}} K, \quad \forall k \in \mathbb{N}. \quad (106)$$

We thus conclude that

$$\begin{aligned} &\alpha_i^{(n_k)} \tilde{v}^q_{\boldsymbol{\lambda_0},\alpha,\eta}(K_i^{(n_k)}) \\ &\stackrel{(a)}{\leq} \alpha_i^{(n_k)} \left(\frac{\lambda_2}{2} \log |I + G_2 K_i^{(n_k)} G_2^\top| + \hat{B}_{\boldsymbol{\lambda_0},\alpha,\eta} \right) \\ &\stackrel{(b)}{\leq} \alpha_i^{(n_k)} \left(\frac{\lambda_2}{2} \log \left| I + G_2 \frac{K}{\alpha_i^{(n_k)}} G_2^\top \right| + \hat{B}_{\boldsymbol{\lambda_0},\alpha,\eta} \right) \\ &\stackrel{(c)}{=} \alpha_i^{(n_k)} \left(\frac{\lambda_2}{2} \sum_{j=1}^{t} \log \left(1 + \frac{\mu_j}{\alpha_i^{(n_k)}} \right) + \hat{B}_{\boldsymbol{\lambda_0},\alpha,\eta} \right) \xrightarrow[n \to \infty]{} 0, \end{aligned}$$
(107)

where (a) and (b) follow from (105) and (106), respectively, while (c) follows by denoting the eigenvalues of $G_2 K G_2^\top$ 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 $\Lambda_d \triangleq \left\{ \boldsymbol{\lambda} \in \mathbb{R}^d_+ \big| \sum_{i=1}^d \lambda_i = 1 \right\}$ 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 functionals such that

$$\mathcal{A} \triangleq \left\{ \mathbf{a} \in \mathbb{R}^d \big| \forall i \in [d], \exists P_U \in \mathcal{P}, a_i \leq g_i(P_U) \right\} \quad (108)$$

is a convex set. Then

$$\sup_{P_U \in \mathcal{P}} \min_{\boldsymbol{\lambda} \in \Lambda_d} \sum_{i=1}^{d} \lambda_i g_i(P_U) = \min_{\boldsymbol{\lambda} \in \Lambda_d} \sup_{P_U \in \mathcal{P}} \sum_{i=1}^{d} \lambda_i g_i(P_U). \quad (109)$$

Let $d = 2$ and $\mathcal{P}$ be the set of PDFs $P_{W,V,U,\mathbf{X}}$ that satisfy $\mathbb{E}[\mathbf{XX}^\top] \preceq K$. Set

$$\begin{aligned} &g_1\left(P_{W,V,U,\mathbf{X}}\right) \\ &= \lambda_0 I(W; \mathbf{Y}_1) + \lambda_2 I(V; \mathbf{Y}_2|W) \\ &\quad + \lambda_1 \Big[I(U; \mathbf{Y}_1|W) - I(U; V|W) - I(U; \mathbf{Y}_2|W, V) \Big] \end{aligned}$$
(110a)

$$\begin{aligned} &g_2\left(P_{W,V,U,\mathbf{X}}\right) \\ &= \lambda_0 I(W; \mathbf{Y}_2) + \lambda_2 I(V; \mathbf{Y}_2|W) \\ &\quad + \lambda_1 \Big[I(U; \mathbf{Y}_1|W) - I(U; V|W) - I(U; \mathbf{Y}_2|W, V) \Big], \end{aligned}$$
(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

$$P(w, v, u, \mathbf{x}) = \alpha P_a(w, v, u, \mathbf{x}) + \bar{\alpha} P_b(w, v, u, \mathbf{x}), \quad (111)$$

for all $(w, v, u, \mathbf{x}) \in \mathcal{W} \times \mathcal{V} \times \mathcal{U} \times \mathcal{X}^n$. Equivalently, P can be represented by setting $\tilde{W} = (Q, W)$, where $Q \sim \mathsf{Ber}(\alpha)$, and denoting $P \triangleq P_{\tilde{W}, V, U, \mathbf{X}} = P_{(Q,W), V, U, \mathbf{X}}$, for which

$$P_{(Q,W), V, U, \mathbf{X}}\big((0, w), v, u, \mathbf{x}\big) = \alpha P_a(w, v, u, \mathbf{x}) \quad (112a)$$
$$P_{(Q,W), V, U, \mathbf{X}}\big((1, w), v, u, \mathbf{x}\big) = \bar{\alpha} P_b(w, v, u, \mathbf{x}), \quad (112b)$$

for all $(w, v, u, \mathbf{x}) \in \mathcal{W} \times \mathcal{V} \times \mathcal{U} \times \mathcal{X}^n$. First note that

$$\mathbb{E}_P\big[\mathbf{X}\mathbf{X}^T\big] = \alpha \mathbb{E}_{P_a}\big[\mathbf{X}\mathbf{X}^T\big] + \bar{\alpha} \mathbb{E}_{P_b}\big[\mathbf{X}\mathbf{X}^T\big] \preceq \mathsf{K}, \quad (113)$$

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

$$\begin{aligned}
g_i(P) &= \lambda_0 I_P(\tilde{W}; \mathbf{Y}_i) + \lambda_1 \Big[I_P(U; \mathbf{Y}_1 | \tilde{W}) - I_P(U; V | \tilde{W}) \\
&\quad - I_P(U; \mathbf{Y}_2 | \tilde{W}, V) \Big] + \lambda_2 I_P(V; \mathbf{Y}_2 | \tilde{W}) \\
&= \lambda_0 I_P(Q; \mathbf{Y}_i) + \alpha g_i(P_a) + \bar{\alpha} g_i(P_b) \\
&\geq \alpha a_i + \bar{\alpha} b_i, \quad (114)
\end{aligned}$$

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 $(W, V, U, \mathbf{X})$, with $(W, V, U) - \mathbf{X} - (\mathbf{Y}_1, \mathbf{Y}_2)$ and $\mathbb{E}\big[\mathbf{X}\mathbf{X}^\top\big] \preceq \mathsf{K}$, the minimum

$$\begin{aligned}
\min_{\alpha \in [0,1]} \Big\{ &\lambda_0 \Big[\alpha I(W; \mathbf{Y}_1) + \bar{\alpha} I(W; \mathbf{Y}_2) \Big] \\
&+ \lambda_1 \Big[I(U; \mathbf{Y}_1 | W) - I(U; V | W) - I(U; \mathbf{Y}_2 | W, V) \Big] \\
&\qquad\qquad\qquad\qquad\qquad + \lambda_2 I(V; \mathbf{Y}_2 | W) \Big\}
\end{aligned}$$

is attained by either $\alpha = 0$ or $\alpha = 1$.

ACKNOWLEDGEMENTS

REFERENCES

[1] A. Gupta and R. K. Jha, "A survey of 5G network: Architecture and emerging technologies," *IEEE Access*, vol. 3, pp. 1206–1232, 2015.

[2] Y. Liang, H. V. Poor, and S. Shamai (Shitz). "Information theoretic security," in *Foundations and Trends in Communications and Information Theory*, vol. 5. Boston, MA, USA: Now, 2008, pp. 355–580.

[3] Z. Li, W. Trappe, and R. Yates, "Secret communication via multi-antenna transmission," in *Proc. 41st Annu. Conf. Inf. Sci. Syst.*, Baltimore, MD, USA, Mar. 2007, pp. 905–910.

[4] S. Shafiee, N. Liu, and S. Ulukus, "Towards the secrecy capacity of the Gaussian MIMO wire-tap channel: The 2-2-1 channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4033–4039, Sep. 2009.

[5] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas I: The MISOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3088–3104, Jul. 2010.

[6] R. Bustin, R. Liu, H. V. Poor, and S. Shamai (Shitz), "An MMSE approach to the secrecy capacity of the MIMO Gaussian wiretap channel," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Seoul, South Korea, Jun./Jul. 2009, pp. 2602–2606.

[7] T. Liu and S. Shamai (Shitz), "A note on the secrecy capacity of the multiple-antenna wiretap channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2547–2553, Jun. 2009.

[8] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas—Part II: The MIMOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5515–5532, Nov. 2010.

[9] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4961–4972, Aug. 2011.

[10] E. Ekrem and S. Ulukus, "The secrecy capacity region of the Gaussian MIMO multi-receiver wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2083–2114, Apr. 2011.

[11] H. D. Ly, T. Liu, and Y. Liang, "Multiple-input multiple-output Gaussian broadcast channels with common and confidential messages," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5477–5487, Nov. 2010.

[12] H.-F. Chong and Y.-C. Liang, "Secrecy capacity region of a class of two-user Gaussian MIMO BC with degraded message sets," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Istanbul, Turkey, Jul. 2013, pp. 2009–2013.

[13] R. Liu, T. Liu, H. V. Poor, and S. Shamai (Shitz), "Multiple-input multiple-output Gaussian broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4215–4227, Sep. 2010.

[14] E. Ekrem and S. Ulukus, "Capacity region of Gaussian MIMO broadcast channels with common and confidential messages," *IEEE Trans. Inf. Theory*, vol. 58, no. 9, pp. 5669–5680, Sep. 2012.

[15] H. Weingarten, Y. Steinberg, and S. Shamai (Shitz), "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.

[16] Y. Geng and C. Nair, "The capacity region of the two-receiver Gaussian vector broadcast channel with private and common messages," *IEEE Trans. Inf. Theory*, vol. 60, no. 4, pp. 2087–2104, Apr. 2014.

[17] P. P. Bergmans, "A simple converse for broadcast channels with additive white Gaussian noise," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 2, pp. 279–280, Mar. 1974.

[18] E. Ekrem and S. Ulukus, "An outer bound for the Gaussian MIMO broadcast channel with common and private messages," *IEEE Trans. Inf. Theory*, vol. 58, no. 11, pp. 6766–6772, Nov. 2012.

[19] E. Ekrem and S. Ulukus, "Multi-receiver wiretap channel with public and confidential messages," *IEEE Trans. Inf. Theory*, vol. 59, no. 4, pp. 2165–2177, Apr. 2013.

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

[21] A. Khina, Y. Kochman, and A. Khisti, "The confidential MIMO broadcast capacity: A simple derivation," in *Proc. Int. Symp. Inf. Theory (ISIT)*, Hong Kong, Jun. 2015, pp. 1981–1985.

[22] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[23] Z. Goldfeld, G. Kramer, and H. H. Permuter, "Broadcast channels with privacy leakage constraints," *IEEE Trans. Inf. Theory*, vol. 63, no. 8, pp. 5138–5161, Aug. 2017.

[24] U. Maurer and S. Wolf, "Information-theoretic key agreement: From weak to strong secrecy for free," in *Advances in Cryptology—EUROCRYPT* (Lecture Notes in Computer Science). Berlin, Germany: Springer, 2000, pp. 351–368.

[25] J. Barros and M. Bloch, "Strong secrecy for wireless channels," in *Proc. Int. Conf. Inf. Theoretic Secur.* Berlin, Germany: Springer, 2008, pp. 40–53.

[26] H. L. Royden, *Real Analysis*. New York, NY, USA: Macmillan, 1988.

[27] Y. Geng, A. Gohari, C. Nair, and Y. Yu, "On Marton's inner bound and its optimality for classes of product broadcast channels," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 22–41, Jan. 2014.

[28] L. Bunt, "Bijdrage tot de theorie der convexe puntverzamelingen," Ph.D. dissertation, University of Groningen, Amsterdam, The Netherlands, 1934.

Ziv Goldfeld (S'13–M'17) is currently a postdoctoral fellow at the Laboratory for Information and Decision Systems (LIDS) at MIT. He graduated with a B.Sc. *(summa cum laude)*, an M.Sc. *(summa cum laude)*, and a Ph.D. in Electrical and Computer Engineering from Ben-Gurion University, Israel, in 2012, 2015 and 2018, respectively. His research interest include theoretical machine learning, information theory, complex dynamical systems, high-dimensional and nonparametric statistics and applied probability. Honors include the received his B.Sc. *(summa cum laude)*, M.Sc. *(summa cum laude)* and Ph.D. degrees in Electrical and Computer Engineering from the Ben-Gurion University, Israel, in 2012, 2014 and 2017, respectively. He is currently a postdoctoral fellow at the Laboratory for Information and Decision Systems (LIDS) at MIT.

Ziv is a recipient of several awards, among them the are Rothschild postdoctoral fellowship, the Feder Award, a best student paper award in the IEEE 28-th Convention of Electrical and Electronics Engineers in Israel, the Basor fellowship for outstanding students in the direct Ph.D. program, the Lev-Zion fellowship and the Minerva Short-Term Research Grant (MRG).

Haim H. Permuter (M'08–SM'13) received his B.Sc. *(summa cum laude)* and M.Sc. *(summa cum laude)* degrees in Electrical and Computer Engineering from the Ben-Gurion University, Israel, in 1997 and 2003, respectively, and a Ph.D. degree in Electrical Engineering from Stanford University, California in 2008.

Between 1997 and 2004, he was an officer at a research and development unit of the Israeli Defense Forces. Since 2009 he is with the department of Electrical and Computer Engineering at Ben-Gurion University where he is currently an associate professor.

Prof. Permuter is a recipient of several awards, among them are the Fullbright Fellowship, the Stanford Graduate Fellowship (SGF), Allon Fellowship, and the U.S.-Israel Binational Science Foundation Bergmann Memorial Award. Haim is currently serving on the editorial board of the IEEE TRANSACTIONS ON INFORMATION THEORY.