Vector Gaussian Two-Terminal Source Coding

Jia Wang and Jun Chen, *Member, IEEE*

Abstract—We derive a lower bound on each supporting line of the rate region of the vector Gaussian two-terminal CEO problem, which is a special case of the indirect vector Gaussian two-terminal source coding problem. The key technical ingredient is a new extremal inequality. It is shown that the lower bound coincides with the Berger–Tung upper bound in the high-resolution regime. Similar results are derived for the direct vector Gaussian two-terminal source coding problem.

Index Terms—CEO problem, extremal inequality, MMSE, multiterminal source coding, rate region.

I. INTRODUCTION

SINCE the introduction of the CEO problem by Berger *et al.* [1], considerable attention has been devoted to characterizing the rate region of the scalar Gaussian version of the problem [2]. In particular, Oohama [3] developed an ingenious method for bounding the rate region by leveraging Shannon's entropy power inequality to relate various information-theoretic quantities. This method was later refined and eventually led to a complete characterization of the rate region of the scalar Gaussian CEO problem [4], [5] and its variant [6]. The scalar Gaussian CEO problem is a special case of the general indirect scalar Gaussian multiterminal source coding problem. In fact, Oohama's method is also applicable to this general problem [7], although in this setting it does not yield a complete solution.

It was observed by Wagner *et al.* [8] that the Gaussian CEO problem is closely related to the direct Gaussian multiterminal source coding problem. By effectively exploiting this link, they settled the longstanding open problem of determining the rate region of the direct scalar Gaussian two-terminal source coding problem.[1] The intimate connection between indirect and direct

[1]The rate region of the direct scalar Gaussian two-terminal source coding problem was partially characterized in an earlier work by Oohama [9].

Gaussian multiterminal source coding was further elucidated in [7]. Through this connection, Oohama's method has become an integral part of a general bounding technique for the direct scalar Gaussian multiterminal source coding problem.

However, this method is not completely suitable for vector Gaussian multiterminal source coding problems because in the vector case the entropy power inequality in general yields a loose bound unless the relevant covariance matrices satisfy a certain proportionality condition. Though this issue can be (partially) resolved for the vector Gaussian CEO problem as far as the sum rate is concerned [10] by combining Oohama's method with an enhancement argument [11], it is unclear whether other weighted sum rates can be treated in a similar way.

Recently, an alternative approach to Gaussian multiterminal source coding problems was proposed in [12] (see [13] for further development). The key idea underlying this approach is that instead of relating information-theoretic quantities via the entropy power inequality, one can relate the corresponding estimation-theoretic quantities by exploiting the semidefinite partial order between the (reduced) error covariance matrices incurred by the MMSE estimator and the (reduced) optimal linear estimator. Due to its estimation-theoretic nature, this approach is directly applicable to the vector case. However, it is only effective for bounding the sum rate. This is because for the sum rate, one can bound information-theoretic quantities in terms of estimation-theoretic quantities in a greedy manner by invoking the simple fact that the Gaussian distribution maximizes the differential entropy under a covariance constraint. On the other hand, for other weighted sum rates, due to the existence of tension among the relevant information-theoretic quantities, the greedy method is not applicable anymore. Toward the end of removing this limitation of the estimation-theoretic approach and making it effective for bounding the whole rate region, in this paper, we establish a new extremal inequality, which can be viewed as a variant of the Liu–Viswanath extremal inequality [14]. It will be seen that this extremal inequality enables us to deal with the aforementioned tension in the two-terminal case.

The rest of this paper is organized as follows. In Section II, we prove a new extremal inequality, which plays an instrumental role in the converse arguments in this paper. In Section III, we derive a lower bound on each supporting line of the rate region of the vector Gaussian two-terminal CEO problem and establish certain sufficient conditions under which the lower bound matches the Berger–Tung upper bound; in particular, we show that the two bounds coincide in the high-resolution regime. Similar results are derived for the direct vector Gaussian two-terminal source coding problem in Section IV. Section V contains a few concluding remarks.

Throughout this paper, the zero matrix and the $k \times k$ identity matrix are denoted by $\mathbf{0}$ and $\mathbf{I}_k$, respectively; for a positive

semidefinite matrix $\boldsymbol{\Sigma}$, its unique positive semidefinite square root is denoted by $\boldsymbol{\Sigma}^{\frac{1}{2}}$; we assume the logarithm function is to base e and define $\log^+ x = \max(\log x, 0)$.

II. EXTREMAL INEQUALITY

We first review some basic definitions and results in [12, Sec. II]. We assume that random matrices are of mean zero unless specified otherwise. For any random matrices $\mathbf{S}, \hat{\mathbf{S}}_1, \ldots, \hat{\mathbf{S}}_k$ with the same column dimension n, define $\boldsymbol{\Sigma}_{\mathbf{S}} = \frac{1}{n}\mathbb{E}[\mathbf{S}\mathbf{S}^T]$, $\boldsymbol{\Sigma}_{\hat{\mathbf{S}}} = \frac{1}{n}\mathbb{E}[\hat{\mathbf{S}}\hat{\mathbf{S}}^T]$, and $\boldsymbol{\Sigma}_{\mathbf{S},\hat{\mathbf{S}}} = \frac{1}{n}\mathbb{E}[\mathbf{S}\hat{\mathbf{S}}^T]$, where $\hat{\mathbf{S}} \triangleq (\hat{\mathbf{S}}_1^T, \ldots, \hat{\mathbf{S}}_k^T)^T$. The (reduced) optimal linear estimator of $\mathbf{S}$ from $\hat{\mathbf{S}}_1, \ldots, \hat{\mathbf{S}}_k$ is defined as $\mathbf{A}\hat{\mathbf{S}}$, where $\mathbf{A}$ is any solution of the equation $\mathbf{A}\boldsymbol{\Sigma}_{\hat{\mathbf{S}}} = \boldsymbol{\Sigma}_{\mathbf{S},\hat{\mathbf{S}}}$ (in particular, $\mathbf{A} = \boldsymbol{\Sigma}_{\mathbf{S},\hat{\mathbf{S}}}\boldsymbol{\Sigma}_{\hat{\mathbf{S}}}^{-1}$ if $\boldsymbol{\Sigma}_{\hat{\mathbf{S}}}$ is invertible), and the incurred (reduced) error covariance matrix $\boldsymbol{\Sigma}_{\mathbf{S}-\mathbf{A}\hat{\mathbf{S}}}$ is given by

$$\boldsymbol{\Sigma}_{\mathbf{S}-\mathbf{A}\hat{\mathbf{S}}} = \boldsymbol{\Sigma}_{\mathbf{S}} - \mathbf{A}\boldsymbol{\Sigma}_{\mathbf{S},\hat{\mathbf{S}}}^T.$$

Note that the (reduced) optimal linear estimator coincides with the standard linear MMSE estimator when $n = 1$. For any random object V jointly distributed with $\mathbf{S}$, we define $\boldsymbol{\Sigma}_{\mathbf{S}|V} = \boldsymbol{\Sigma}_{\mathbf{S}-\mathbb{E}[\mathbf{S}|V]}$ and refer to it as the (reduced) error covariance matrix incurred by the MMSE estimator of $\mathbf{S}$ from V (i.e., $\mathbb{E}[\mathbf{S}|V]$). An important fact is that

$$\boldsymbol{\Sigma}_{\mathbf{S}|V} \preceq \boldsymbol{\Sigma}_{\mathbf{S}-\mathbf{S}'}$$

for any random matrix $\mathbf{S}'$ (of the same size as $\mathbf{S}$) that is a function of V; of particular relevance here is the case $\mathbf{S}' = \mathbf{A}\hat{\mathbf{S}}$ with $\hat{\mathbf{S}}$ being a function of V.

The main result of this section is the following extremal inequality.

Theorem 1: Let $\mathbf{Z}$ be an $m \times n$ random matrix such that each column is an independent copy of an $m \times 1$ Gaussian random vector with mean zero and positive-definite covariance matrix $\boldsymbol{\Sigma}_{\mathbf{Z}}$. Let μ be a real number and $\boldsymbol{\Gamma}$ be a positive-definite matrix satisfying $\mu \geq 1$ and $\boldsymbol{\Gamma} \prec \boldsymbol{\Sigma}_{\mathbf{Z}}$, respectively. Moreover, let $\boldsymbol{\Sigma}^*$ be a positive-definite matrix satisfying $\boldsymbol{\Sigma}^* \preceq \tilde{\boldsymbol{\Gamma}} \triangleq (\boldsymbol{\Gamma}^{-1} - \boldsymbol{\Sigma}_{\mathbf{Z}}^{-1})^{-1}$ for which there exists a positive semidefinite matrix $\mathbf{M}$ such that

$$(\boldsymbol{\Sigma}^*)^{-1} = \mu(\boldsymbol{\Sigma}^* + \boldsymbol{\Sigma}_{\mathbf{Z}})^{-1} + \mathbf{M}, \tag{1}$$

$$\mathbf{M}(\tilde{\boldsymbol{\Gamma}} - \boldsymbol{\Sigma}^*) = \mathbf{0}. \tag{2}$$

Then, for any random object V and $m \times n$ random matrix $\mathbf{S}$, jointly independent of $\mathbf{Z}$, such that $\boldsymbol{\Sigma}_{\mathbf{S}|\mathbf{S}+\mathbf{Z},V} \preceq \boldsymbol{\Gamma}$, we have

$$h(\mathbf{S}|V) - \mu h(\mathbf{S} + \mathbf{Z}|V)$$
$$\leq \frac{n}{2}\log((2\pi e)^m |\boldsymbol{\Sigma}^*|) - \frac{\mu n}{2}\log((2\pi e)^m |\boldsymbol{\Sigma}^* + \boldsymbol{\Sigma}_{\mathbf{Z}}|).$$

Proof: See Appendix A. ∎

Define $\mathcal{G}$ to be the set of joint distributions $p_{V,\mathbf{S}}$ such that $\mathbf{S} - \mathbb{E}[\mathbf{S}|V]$ is independent of V and the columns of $\mathbf{S} - \mathbb{E}[\mathbf{S}|V]$ are independent and identically distributed (i.i.d.) Gaussian random
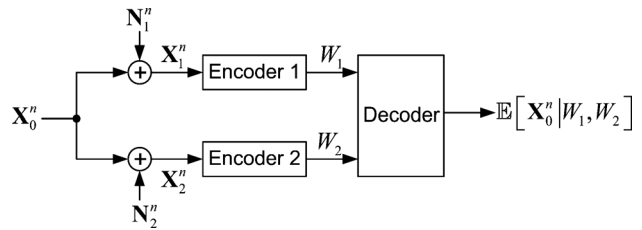

Fig. 1.   Vector Gaussian two-terminal CEO problem.

vectors. It is easy to verify that for $p_{V,\mathbf{S}} \in \mathcal{G}$, $\boldsymbol{\Sigma}_{\mathbf{S}|\mathbf{S}+\mathbf{Z},V} \preceq \boldsymbol{\Gamma}$ if and only if $\boldsymbol{\Sigma}_{\mathbf{S}|V} \preceq \tilde{\boldsymbol{\Gamma}}$. Therefore, we have

$$\max_{p_{V,\mathbf{S}}:\boldsymbol{\Sigma}_{\mathbf{S}|\mathbf{S}+\mathbf{Z},V} \preceq \boldsymbol{\Gamma}} h(\mathbf{S}|V) - \mu h(\mathbf{S} + \mathbf{Z}|V)$$
$$\geq \max_{p_{V,\mathbf{S}} \in \mathcal{G}:\boldsymbol{\Sigma}_{\mathbf{S}|\mathbf{S}+\mathbf{Z},V} \preceq \boldsymbol{\Gamma}} h(\mathbf{S}|V) - \mu h(\mathbf{S} + \mathbf{Z}|V) \tag{3}$$
$$= \max_{p_{V,\mathbf{S}} \in \mathcal{G}:\boldsymbol{\Sigma}_{\mathbf{S}|V} \preceq \tilde{\boldsymbol{\Gamma}}} h(\mathbf{S}|V) - \mu h(\mathbf{S} + \mathbf{Z}|V)$$
$$= \max_{\mathbf{0} \prec \boldsymbol{\Sigma} \preceq \tilde{\boldsymbol{\Gamma}}} \frac{n}{2}\log((2\pi e)^m|\boldsymbol{\Sigma}|) - \frac{\mu n}{2}\log((2\pi e)^m|\boldsymbol{\Sigma}+\boldsymbol{\Sigma}_{\mathbf{Z}}|) \tag{4}$$
$$\geq \frac{n}{2}\log((2\pi e)^m|\boldsymbol{\Sigma}^*|) - \frac{\mu n}{2}\log((2\pi e)^m|\boldsymbol{\Sigma}^* + \boldsymbol{\Sigma}_{\mathbf{Z}}|). \tag{5}$$

In view of Theorem 1, the inequalities in (3) and (5) must be equalities. Note that (1) and (2) are in fact the Karush–Kuhn–Tucker (KKT) conditions that are necessary for a solution of the maximization problem in (4) to be optimal. Theorem 1 implies that these conditions are sufficient for the global optimality. If we replace $\boldsymbol{\Sigma}_{\mathbf{S}|\mathbf{S}+\mathbf{Z},V} \preceq \boldsymbol{\Gamma}$ by $\boldsymbol{\Sigma}_{\mathbf{S}|V} \preceq \tilde{\boldsymbol{\Gamma}}$, then Theorem 1 becomes essentially a conditional version of the extremal inequality by Liu and Viswanath [14, Corollary 5] for the case $\mu \geq 1$. It is easy to show that $\boldsymbol{\Sigma}_{\mathbf{S}|V} \preceq \tilde{\boldsymbol{\Gamma}}$ implies $\boldsymbol{\Sigma}_{\mathbf{S}|\mathbf{S}+\mathbf{Z},V} \preceq \boldsymbol{\Gamma}$ (while they are equivalent for $p_{V,\mathbf{S}} \in \mathcal{G}$) by leveraging the fact[2]

$$\boldsymbol{\Sigma}_{\mathbf{S}|\mathbf{S}+\mathbf{Z},V} \preceq \boldsymbol{\Sigma}_{\mathbf{Z}} - \boldsymbol{\Sigma}_{\mathbf{Z}}(\boldsymbol{\Sigma}_{\mathbf{S}|V} + \boldsymbol{\Sigma}_{\mathbf{Z}})^{-1}\boldsymbol{\Sigma}_{\mathbf{Z}}.$$

As a consequence, Theorem 1 can be viewed as a strengthened version of the Liu–Viswanath extremal inequality.

III. VECTOR GAUSSIAN TWO-TERMINAL CEO PROBLEM

We shall first give a formal definition of the rate region of the vector Gaussian two-terminal CEO problem (see Fig. 1 for the system model). Let $\mathbf{X}_0$, $\mathbf{N}_1$, and $\mathbf{N}_2$ be three mutually independent $m \times 1$ Gaussian random vectors with mean zero and positive-definite covariance matrices $\boldsymbol{\Sigma}_{\mathbf{X}_0}$, $\boldsymbol{\Sigma}_{\mathbf{N}_1}$, and $\boldsymbol{\Sigma}_{\mathbf{N}_2}$, respectively. Let $\mathbf{X}_1 = \mathbf{X}_0 + \mathbf{N}_1$ and $\mathbf{X}_2 = \mathbf{X}_0 + \mathbf{N}_2$. Let $\{(\mathbf{X}_0(t), \mathbf{N}_1(t), \mathbf{N}_2(t), \mathbf{X}_1(t), \mathbf{X}_2(t))\}_{t=1}^{\infty}$ be i.i.d. copies of $(\mathbf{X}_0, \mathbf{N}_1, \mathbf{N}_2, \mathbf{X}_1, \mathbf{X}_2)$.

Definition 1: A rate pair (R_1, R_2) is said to be achievable subject to distortion constraint $\mathbf{D}$ if for all sufficiently large n,

---

[2]Note that $\boldsymbol{\Sigma}_{\mathbf{Z}} - \boldsymbol{\Sigma}_{\mathbf{Z}}(\boldsymbol{\Sigma}_{\mathbf{S}|V} + \boldsymbol{\Sigma}_{\mathbf{Z}})^{-1}\boldsymbol{\Sigma}_{\mathbf{Z}}$ is the (reduced) error covariance matrix incurred by the (reduced) optimal linear estimator of $\mathbf{S}$ from $\mathbb{E}[\mathbf{S}|V]$ and $\mathbf{S} + \mathbf{Z}$.

there exist encoding functions $f_i^{(n)} : \mathbb{R}^{m \times n} \to \{1, 2, \ldots, M_i\}$, $i = 1, 2$, such that

$$\frac{1}{n} \log M_i \leq R_i, \quad i = 1, 2$$
$$\Sigma_{\mathbf{X}_0^n | W_1, W_2} \preceq \mathbf{D}$$

where $W_i = f_i^{(n)}(\mathbf{X}_i^n)$, $i = 1, 2$. The rate region $\mathcal{R}(\mathbf{D})$ is the closure of the set of all achievable rate pairs subject to distortion constraint $\mathbf{D}$.

To facilitate subsequent analysis, we define $\mathbf{Y} = \mathbb{E}[\mathbf{X}_0 | \mathbf{X}_1, \mathbf{X}_2]$, $\mathbf{N} = \mathbf{X}_0 - \mathbf{Y}$, $\tilde{\mathbf{X}}_i = \mathbb{E}[\mathbf{X}_0 | \mathbf{X}_i]$, $\tilde{\mathbf{N}}_i = \mathbf{X}_0 - \tilde{\mathbf{X}}_i$, $i = 1, 2$, and define $\mathbf{Y}(t), \mathbf{N}(t), \tilde{\mathbf{X}}_i(t), \tilde{\mathbf{N}}_i(t)$, $i = 1, 2, t \geq 1$, accordingly. Note that $\mathbf{N}$ is independent of $(\mathbf{X}_1, \mathbf{X}_2)$ and $\tilde{\mathbf{N}}_i$ is independent of $\tilde{\mathbf{X}}_i$, $i = 1, 2$. It can be verified that

$$\Sigma_{\mathbf{N}} = (\Sigma_{\mathbf{X}_0}^{-1} + \Sigma_{\mathbf{N}_1}^{-1} + \Sigma_{\mathbf{N}_2}^{-1})^{-1}$$
$$\mathbf{Y} = \Sigma_{\mathbf{N}} \Sigma_{\mathbf{N}_1}^{-1} \mathbf{X}_1 + \Sigma_{\mathbf{N}} \Sigma_{\mathbf{N}_2}^{-1} \mathbf{X}_2$$
$$\Sigma_{\tilde{\mathbf{N}}_i} = (\Sigma_{\mathbf{X}_0}^{-1} + \Sigma_{\mathbf{N}_i}^{-1})^{-1}, \quad i = 1, 2$$
$$\tilde{\mathbf{X}}_i = \Sigma_{\tilde{\mathbf{N}}_i} \Sigma_{\mathbf{N}_i}^{-1} \mathbf{X}_i, \quad i = 1, 2.$$

In view of the fact that $\Sigma_{\mathbf{N}} \prec \Sigma_{\mathbf{X}_0^n | W_1, W_2} \preceq \Sigma_{\mathbf{X}_0}$, we shall assume $\Sigma_{\mathbf{N}} \prec \mathbf{D} \preceq \Sigma_{\mathbf{X}_0}$. It can be shown through a simple time-sharing argument that $\mathcal{R}(\mathbf{D})$ is a (closed) convex set. Therefore, $\mathcal{R}(\mathbf{D})$ is completely characterized by its supporting lines. Define

$$R(\mathbf{D}, \alpha_1, \alpha_2) = \inf_{(R_1, R_2) \in \mathcal{R}(\mathbf{D})} \alpha_1 R_1 + \alpha_2 R_2$$

where α_1 and α_2 are nonnegative real numbers. We shall bound $\mathcal{R}(\mathbf{D})$ by establishing upper and lower bounds on $R(\mathbf{D}, \alpha_1, \alpha_2)$. Without loss of generality, we assume $\alpha_1 \geq \alpha_2 \geq 0$. Note that the degenerate case $\alpha_2 = 0$ corresponds essentially to a (remote) Wyner–Ziv problem. Specifically, it can be shown (see Appendix B) that

$$R(\mathbf{D}, \alpha_1, 0) = \min_{\tilde{\mathbf{D}}} \frac{\alpha_1}{2} \log \frac{|\Sigma_{\mathbf{N}_1} + \Sigma_{\tilde{\mathbf{N}}_2}|}{|\tilde{\mathbf{D}}|} \qquad (6)$$
$$\text{subject to} \quad \Sigma_{\mathbf{N}} \Sigma_{\mathbf{N}_1}^{-1} \tilde{\mathbf{D}} \Sigma_{\mathbf{N}_1}^{-1} \Sigma_{\mathbf{N}} + \Sigma_{\mathbf{N}} \preceq \mathbf{D},$$
$$\mathbf{0} \prec \tilde{\mathbf{D}} \preceq \Sigma_{\mathbf{N}_1} + \Sigma_{\tilde{\mathbf{N}}_2}.$$

Hence, unless stated otherwise, we shall focus on the case $\alpha_1 \geq \alpha_2 > 0$ in the rest of this section.

A. Lower Bound

The main result of this section is the following theorem, which provides a lower bound on $R(\mathbf{D}, \alpha_1, \alpha_2)$.

Theorem 2: $R(\mathbf{D}, \alpha_1, \alpha_2) \geq \underline{R}(\mathbf{D}, \alpha_1, \alpha_2)$, where

$$\underline{R}(\mathbf{D}, \alpha_1, \alpha_2) \triangleq$$
$$\min_{\mathbf{D}_0, \mathbf{D}_1, \mathbf{D}_2} \frac{\alpha_1 - \alpha_2}{2} \log^+ \frac{1}{|\mathbf{D}_0| |\Sigma_{\tilde{\mathbf{N}}_2}^{-1} - \Sigma_{\mathbf{N}_2}^{-1} \mathbf{D}_2 \Sigma_{\mathbf{N}_2}^{-1}|}$$
$$+ \frac{\alpha_1}{2} \log \frac{|\Sigma_{\mathbf{N}_1}|}{|\mathbf{D}_1|} + \frac{\alpha_2}{2} \log \frac{|\Sigma_{\mathbf{X}_0}| |\Sigma_{\mathbf{N}_2}|}{|\mathbf{D}_0| |\mathbf{D}_2|}$$
$$\text{subject to} \quad \Sigma_{\mathbf{N}_1}^{-1} \mathbf{D}_1 \Sigma_{\mathbf{N}_1}^{-1} + \Sigma_{\mathbf{N}_2}^{-1} \mathbf{D}_2 \Sigma_{\mathbf{N}_2}^{-1} \preceq \Sigma_{\mathbf{N}}^{-1} - \mathbf{D}_0^{-1},$$
$$\Sigma_{\mathbf{N}} \prec \mathbf{D}_0 \preceq \mathbf{D},$$
$$\mathbf{0} \prec \mathbf{D}_i \preceq \Sigma_{\mathbf{N}_i}, \quad i = 1, 2.$$

Proof: Let $f_i^{(n)} : \mathbb{R}^{m \times n} \to \{1, 2, \ldots, M_i\}$, $i = 1, 2$, be two encoding functions such that $\Sigma_{\mathbf{X}_0^n | W_1, W_2} \preceq \mathbf{D}$, where $W_i = f_i^{(n)}(\mathbf{X}_i^n)$, $i = 1, 2$. Denote $\Sigma_{\mathbf{X}_0^n | W_1, W_2}$, $\Sigma_{\mathbf{X}_1^n | \mathbf{X}_0^n, W_1}$, and $\Sigma_{\mathbf{X}_2^n | \mathbf{X}_0^n, W_2}$ by $\mathbf{D}_0$, $\mathbf{D}_1$, and $\mathbf{D}_2$, respectively. Note that

$$\Sigma_{\mathbf{N}} \prec \mathbf{D}_0 \preceq \mathbf{D} \qquad (7)$$
$$\mathbf{0} \prec \mathbf{D}_i \preceq \Sigma_{\mathbf{N}_i}, \quad i = 1, 2. \qquad (8)$$

We shall first leverage the estimation-theoretic approach in [12] to establish a connection between $\mathbf{D}_0$ and $(\mathbf{D}_1, \mathbf{D}_2)$. We only give a sketch of this step here due to its similarity to the derivation of (4) in [12]. It can be verified (cf., [12, eq. (2)]) that the (reduced) error covariance matrix incurred by the MMSE estimator of $\mathbf{Y}^n$ from $\mathbf{X}_0^n$, W_1, and W_2 is $\Sigma_{\mathbf{N}} \Sigma_{\mathbf{N}_1}^{-1} \mathbf{D}_1 \Sigma_{\mathbf{N}_1}^{-1} \Sigma_{\mathbf{N}} + \Sigma_{\mathbf{N}} \Sigma_{\mathbf{N}_2}^{-1} \mathbf{D}_2 \Sigma_{\mathbf{N}_2}^{-1} \Sigma_{\mathbf{N}}$. In addition, it can be shown (cf., [12, eq. (3)]) that the (reduced) error covariance matrix incurred by the (reduced) optimal linear estimator of $\mathbf{Y}^n$ from $\mathbf{X}_0^n$ and $\mathbb{E}[\mathbf{X}_0^n | W_1, W_2]$ is $((\mathbf{D}_0 - \Sigma_{\mathbf{N}})^{-1} + \Sigma_{\mathbf{N}}^{-1})^{-1}$. Therefore, we have

$$\Sigma_{\mathbf{N}} \Sigma_{\mathbf{N}_1}^{-1} \mathbf{D}_1 \Sigma_{\mathbf{N}_1}^{-1} \Sigma_{\mathbf{N}} + \Sigma_{\mathbf{N}} \Sigma_{\mathbf{N}_2}^{-1} \mathbf{D}_2 \Sigma_{\mathbf{N}_2}^{-1} \Sigma_{\mathbf{N}}$$
$$\preceq ((\mathbf{D}_0 - \Sigma_{\mathbf{N}})^{-1} + \Sigma_{\mathbf{N}}^{-1})^{-1}$$

which is equivalent to

$$\Sigma_{\mathbf{N}_1}^{-1} \mathbf{D}_1 \Sigma_{\mathbf{N}_1}^{-1} + \Sigma_{\mathbf{N}_2}^{-1} \mathbf{D}_2 \Sigma_{\mathbf{N}_2}^{-1} \preceq \Sigma_{\mathbf{N}}^{-1} - \mathbf{D}_0^{-1}. \qquad (9)$$

The next step is to derive a lower bound on $\frac{\alpha_1}{n} \log M_1 + \frac{\alpha_2}{n} \log M_2$ in terms of $\mathbf{D}_0$, $\mathbf{D}_1$, and $\mathbf{D}_2$. Note that

$$\frac{\alpha_1}{n} \log M_1 + \frac{\alpha_2}{n} \log M_2$$
$$\geq \frac{\alpha_1}{n} H(W_1) + \frac{\alpha_2}{n} H(W_2)$$
$$\geq \frac{\alpha_1 - \alpha_2}{n} H(W_1 | W_2) + \frac{\alpha_2}{n} H(W_1, W_2)$$
$$= \frac{\alpha_1 - \alpha_2}{n} I(\mathbf{X}_0^n, \mathbf{X}_1^n; W_1 | W_2)$$
$$\quad + \frac{\alpha_2}{n} I(\mathbf{X}_0^n, \mathbf{X}_1^n, \mathbf{X}_2^n; W_1, W_2)$$
$$= \frac{\alpha_1 - \alpha_2}{n} (I(\mathbf{X}_0^n; W_1 | W_2) + I(\mathbf{X}_1^n; W_1 | \mathbf{X}_0^n))$$
$$\quad + \frac{\alpha_2}{n} (I(\mathbf{X}_0^n; W_1, W_2) + I(\mathbf{X}_1^n; W_1 | \mathbf{X}_0^n)$$
$$\quad + I(\mathbf{X}_2^n; W_2 | \mathbf{X}_0^n))$$
$$= \frac{\alpha_1 - \alpha_2}{n} I(\mathbf{X}_0^n; W_1 | W_2) + \frac{\alpha_1}{n} I(\mathbf{X}_1^n; W_1 | \mathbf{X}_0^n)$$
$$\quad + \frac{\alpha_2}{n} (I(\mathbf{X}_0^n; W_1, W_2) + I(\tilde{\mathbf{X}}_2^n; W_2 | \mathbf{X}_0^n)). \qquad (10)$$

For reasons that will become clear, we introduce the following minimization problem:

$$\min_{\mathbf{0} \prec \Sigma \preceq \tilde{\mathbf{D}}_2} \frac{\alpha_1 - \alpha_2}{2} \log^+ \frac{|\Sigma + \Sigma_{\tilde{\mathbf{N}}_2}|}{|\mathbf{D}_0|} + \frac{\alpha_2}{2} \log \frac{|\Sigma + \Sigma_{\tilde{\mathbf{N}}_2}|}{|\Sigma|},$$

where $\tilde{\mathbf{D}}_2 = (\Sigma_{\tilde{\mathbf{N}}_2}^{-1} \Sigma_{\mathbf{N}_2} \mathbf{D}_2^{-1} \Sigma_{\mathbf{N}_2} \Sigma_{\tilde{\mathbf{N}}_2}^{-1} - \Sigma_{\tilde{\mathbf{N}}_2}^{-1})^{-1}$. It can be shown (cf., [17, Proposition 3.3.11]) that for any optimal so-

lution $\boldsymbol{\Sigma}^*$ to this minimization problem, there exist $\theta \in [0,1]$ and $\mathbf{M} \succeq \mathbf{0}$ such that

$$\alpha_2 (\boldsymbol{\Sigma}^*)^{-1} = ((\alpha_1 - \alpha_2)\theta + \alpha_2)(\boldsymbol{\Sigma}^* + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2})^{-1} + \mathbf{M} \quad (11)$$

$$\mathbf{M}(\tilde{\mathbf{D}}_2 - \boldsymbol{\Sigma}^*) = \mathbf{0} \quad (12)$$

$$\frac{(\alpha_1 - \alpha_2)\theta}{2} \log \frac{|\boldsymbol{\Sigma}^* + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}|}{|\mathbf{D}_0|} + \frac{\alpha_2}{2} \log \frac{|\boldsymbol{\Sigma}^* + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}|}{|\boldsymbol{\Sigma}^*|}$$
$$= \min_{\mathbf{0} \prec \boldsymbol{\Sigma} \preceq \tilde{\mathbf{D}}_2} \frac{\alpha_1 - \alpha_2}{2} \log^+ \frac{|\boldsymbol{\Sigma} + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}|}{|\mathbf{D}_0|}$$
$$+ \frac{\alpha_2}{2} \log \frac{|\boldsymbol{\Sigma} + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}|}{|\boldsymbol{\Sigma}|}. \quad (13)$$

Now continuing from (10), we have

$$\frac{\alpha_1}{n} \log M_1 + \frac{\alpha_2}{n} \log M_2$$
$$\geq \frac{(\alpha_1 - \alpha_2)\theta}{n} (I(\mathbf{X}_0^n; W_1, W_2) - I(\mathbf{X}_0^n; W_2))$$
$$+ \frac{\alpha_1}{n} I(\mathbf{X}_1^n; W_1 | \mathbf{X}_0^n) + \frac{\alpha_2}{n} (I(\mathbf{X}_0^n; W_1, W_2)$$
$$+ I(\tilde{\mathbf{X}}_2^n; W_2 | \mathbf{X}_0^n))$$
$$\geq \frac{(\alpha_1 - \alpha_2)\theta}{2} \log \frac{|\boldsymbol{\Sigma}_{\mathbf{X}_0}|}{|\mathbf{D}_0|} - \frac{(\alpha_1 - \alpha_2)\theta}{2} \log((2\pi e)^m |\boldsymbol{\Sigma}_{\mathbf{X}_0}|)$$
$$+ \frac{(\alpha_1 - \alpha_2)\theta}{n} h(\mathbf{X}_0^n | W_2) + \frac{\alpha_1}{2} \log \frac{|\boldsymbol{\Sigma}_{\mathbf{N}_1}|}{|\mathbf{D}_1|}$$
$$+ \frac{\alpha_2}{2} \log \frac{|\boldsymbol{\Sigma}_{\mathbf{X}_0}|}{|\mathbf{D}_0|} + \frac{\alpha_2}{n} h(\tilde{\mathbf{X}}_2^n | \mathbf{X}_0^n) - \frac{\alpha_2}{n} h(\tilde{\mathbf{X}}_2^n | \mathbf{X}_0^n, W_2)$$
$$= \frac{(\alpha_1 - \alpha_2)\theta}{2} \log \frac{|\boldsymbol{\Sigma}_{\mathbf{X}_0}|}{|\mathbf{D}_0|} - \frac{(\alpha_1 - \alpha_2)\theta}{2} \log((2\pi e)^m |\boldsymbol{\Sigma}_{\mathbf{X}_0}|)$$
$$+ \frac{(\alpha_1 - \alpha_2)\theta}{n} h(\mathbf{X}_0^n | W_2) + \frac{\alpha_1}{2} \log \frac{|\boldsymbol{\Sigma}_{\mathbf{N}_1}|}{|\mathbf{D}_1|}$$
$$+ \frac{\alpha_2}{2} \log \frac{|\boldsymbol{\Sigma}_{\mathbf{X}_0}|}{|\mathbf{D}_0|} + \frac{\alpha_2}{n} h(\tilde{\mathbf{X}}_2^n | \mathbf{X}_0^n)$$
$$- \frac{\alpha_2}{n} (h(\tilde{\mathbf{X}}_2^n | W_2) + h(\mathbf{X}_0^n | \tilde{\mathbf{X}}_2^n) - h(\mathbf{X}_0^n | W_2))$$
$$= \frac{(\alpha_1 - \alpha_2)\theta}{2} \log \frac{|\boldsymbol{\Sigma}_{\mathbf{X}_0}|}{|\mathbf{D}_0|} - \frac{(\alpha_1 - \alpha_2)\theta}{2} \log((2\pi e)^m |\boldsymbol{\Sigma}_{\mathbf{X}_0}|)$$
$$+ \frac{\alpha_1}{2} \log \frac{|\boldsymbol{\Sigma}_{\mathbf{N}_1}|}{|\mathbf{D}_1|} + \frac{\alpha_2}{2} \log \frac{|\boldsymbol{\Sigma}_{\mathbf{X}_0}|}{|\mathbf{D}_0|} + \frac{\alpha_2}{2} \log \frac{|\boldsymbol{\Sigma}_{\tilde{\mathbf{X}}_2 | \mathbf{X}_0}|}{|\boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}|}$$
$$+ \frac{(\alpha_1 - \alpha_2)\theta + \alpha_2}{n} h(\mathbf{X}_0^n | W_2) - \frac{\alpha_2}{n} h(\tilde{\mathbf{X}}_2^n | W_2). \quad (14)$$

In view of (11), (12), and the fact that

$$\boldsymbol{\Sigma}_{\tilde{\mathbf{X}}_2^n | \mathbf{X}_0^n, W_2} = \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2} \boldsymbol{\Sigma}_{\mathbf{N}_2}^{-1} \mathbf{D}_2 \boldsymbol{\Sigma}_{\mathbf{N}_2}^{-1} \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}$$

one can readily show by leveraging Theorem 1 that

$$\frac{(\alpha_1 - \alpha_2)\theta + \alpha_2}{n} h(\mathbf{X}_0^n | W_2) - \frac{\alpha_2}{n} h(\tilde{\mathbf{X}}_2^n | W_2)$$
$$\geq \frac{(\alpha_1 - \alpha_2)\theta + \alpha_2}{2} \log((2\pi e)^m |\boldsymbol{\Sigma}^* + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}|)$$
$$- \frac{\alpha_2}{2} \log((2\pi e)^m |\boldsymbol{\Sigma}^*|). \quad (15)$$

Substituting (15) into (14) gives

$$\frac{\alpha_1}{n} \log M_1 + \frac{\alpha_2}{n} \log M_2$$
$$\geq \frac{(\alpha_1 - \alpha_2)\theta}{2} \log \frac{|\boldsymbol{\Sigma}^* + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}|}{|\mathbf{D}_0|} + \frac{\alpha_1}{2} \log \frac{|\boldsymbol{\Sigma}_{\mathbf{N}_1}|}{|\mathbf{D}_1|}$$
$$+ \frac{\alpha_2}{2} \log \frac{|\boldsymbol{\Sigma}_{\mathbf{X}_0}||\boldsymbol{\Sigma}_{\tilde{\mathbf{X}}_2 | \mathbf{X}_0}||\boldsymbol{\Sigma}^* + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}|}{|\mathbf{D}_0||\boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}||\boldsymbol{\Sigma}^*|}$$

which, together with (13), implies

$$\frac{\alpha_1}{n} \log M_1 + \frac{\alpha_2}{n} \log M_2$$
$$\geq \min_{\mathbf{0} \prec \boldsymbol{\Sigma} \preceq \tilde{\mathbf{D}}_2} \frac{\alpha_1 - \alpha_2}{2} \log^+ \frac{|\boldsymbol{\Sigma} + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}|}{|\mathbf{D}_0|} + \frac{\alpha_1}{2} \log \frac{|\boldsymbol{\Sigma}_{\mathbf{N}_1}|}{|\mathbf{D}_1|}$$
$$+ \frac{\alpha_2}{2} \log \frac{|\boldsymbol{\Sigma}_{\mathbf{X}_0}||\boldsymbol{\Sigma}_{\tilde{\mathbf{X}}_2 | \mathbf{X}_0}||\boldsymbol{\Sigma} + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}|}{|\mathbf{D}_0||\boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}||\boldsymbol{\Sigma}|}.$$

Define $\mathbf{D}_2' = \boldsymbol{\Sigma}_{\mathbf{N}_2} \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} (\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1})^{-1} \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} \boldsymbol{\Sigma}_{\mathbf{N}_2}$. Note that $\mathbf{0} \prec \boldsymbol{\Sigma} \preceq \tilde{\mathbf{D}}_2$ is equivalent to $\mathbf{0} \prec \mathbf{D}_2' \preceq \mathbf{D}_2$; moreover, we have

$$|\boldsymbol{\Sigma} + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}| = |\boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} - \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} (\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1})^{-1} \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1}|^{-1}$$
$$= |\boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} - \boldsymbol{\Sigma}_{\mathbf{N}_2}^{-1} \mathbf{D}_2' \boldsymbol{\Sigma}_{\mathbf{N}_2}^{-1}|^{-1}$$

and

$$\frac{|\boldsymbol{\Sigma}_{\tilde{\mathbf{X}}_2 | \mathbf{X}_0}||\boldsymbol{\Sigma} + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}|}{|\boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}||\boldsymbol{\Sigma}|} = \frac{|\boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}||\boldsymbol{\Sigma}_{\mathbf{X}_2 | \mathbf{X}_0}||\boldsymbol{\Sigma} + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}|}{|\boldsymbol{\Sigma}_{\mathbf{N}_2}|^2 |\boldsymbol{\Sigma}|}$$
$$= \frac{|\boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}||\boldsymbol{\Sigma} + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}|}{|\boldsymbol{\Sigma}_{\mathbf{N}_2}||\boldsymbol{\Sigma}|}$$
$$= \frac{|\boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}|^2 |\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1}|}{|\boldsymbol{\Sigma}_{\mathbf{N}_2}|}$$
$$= \frac{|\boldsymbol{\Sigma}_{\mathbf{N}_2}|}{|\mathbf{D}_2'|}.$$

As a consequence

$$\frac{\alpha_1}{n} \log M_1 + \frac{\alpha_2}{n} \log M_2$$
$$\geq \min_{\mathbf{0} \prec \mathbf{D}_2' \preceq \mathbf{D}_2} \frac{\alpha_1 - \alpha_2}{2} \log^+ \frac{1}{|\mathbf{D}_0||\boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} - \boldsymbol{\Sigma}_{\mathbf{N}_2}^{-1} \mathbf{D}_2' \boldsymbol{\Sigma}_{\mathbf{N}_2}^{-1}|}$$
$$+ \frac{\alpha_1}{2} \log \frac{|\boldsymbol{\Sigma}_{\mathbf{N}_1}|}{|\mathbf{D}_1|} + \frac{\alpha_2}{2} \log \frac{|\boldsymbol{\Sigma}_{\mathbf{X}_0}||\boldsymbol{\Sigma}_{\mathbf{N}_2}|}{|\mathbf{D}_0||\mathbf{D}_2'|}. \quad (16)$$

Combining (7)–(9) and (16) yields $R(\mathbf{D}, \alpha_1, \alpha_2) \geq \underline{R}'(\mathbf{D}, \alpha_1, \alpha_2)$, where

$$\underline{R}'(\mathbf{D}, \alpha_1, \alpha_2) \triangleq$$
$$\min_{\mathbf{D}_0, \mathbf{D}_1, \mathbf{D}_2, \mathbf{D}_2'} \frac{\alpha_1 - \alpha_2}{2} \log^+ \frac{1}{|\mathbf{D}_0||\boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} - \boldsymbol{\Sigma}_{\mathbf{N}_2}^{-1} \mathbf{D}_2' \boldsymbol{\Sigma}_{\mathbf{N}_2}^{-1}|}$$
$$+ \frac{\alpha_1}{2} \log \frac{|\boldsymbol{\Sigma}_{\mathbf{N}_1}|}{|\mathbf{D}_1|} + \frac{\alpha_2}{2} \log \frac{|\boldsymbol{\Sigma}_{\mathbf{X}_0}||\boldsymbol{\Sigma}_{\mathbf{N}_2}|}{|\mathbf{D}_0||\mathbf{D}_2'|}$$
$$\text{subject to} \quad \boldsymbol{\Sigma}_{\mathbf{N}_1}^{-1} \mathbf{D}_1 \boldsymbol{\Sigma}_{\mathbf{N}_1}^{-1} + \boldsymbol{\Sigma}_{\mathbf{N}_2}^{-1} \mathbf{D}_2 \boldsymbol{\Sigma}_{\mathbf{N}_2}^{-1} \preceq \boldsymbol{\Sigma}_{\mathbf{N}}^{-1} - \mathbf{D}_0^{-1},$$
$$\boldsymbol{\Sigma}_{\mathbf{N}} \prec \mathbf{D}_0 \preceq \mathbf{D},$$
$$\mathbf{0} \prec \mathbf{D}_i \preceq \boldsymbol{\Sigma}_{\mathbf{N}_i}, \quad i = 1, 2$$
$$\mathbf{0} \prec \mathbf{D}_2' \preceq \mathbf{D}_2.$$

Without loss of optimality, one can set $\mathbf{D}'_2 = \mathbf{D}_2$ in the above minimization problem, which implies that $\underline{R}'(\mathbf{D}, \alpha_1, \alpha_2) = \underline{R}(\mathbf{D}, \alpha_1, \alpha_2)$. This completes the proof of Theorem 2. ∎

B. Upper Bound

The main result of this section is the following upper bound on $R(\mathbf{D}, \alpha_1, \alpha_2)$, which will be referred to as the Berger–Tung upper bound since it is a simple consequence of the Berger–Tung inner bound of $\mathcal{R}(\mathbf{D})$.

Theorem 3: $R(\mathbf{D}, \alpha_1, \alpha_2) \leq \overline{R}(\mathbf{D}, \alpha_1, \alpha_2)$, where

$$\overline{R}(\mathbf{D}, \alpha_1, \alpha_2) \triangleq$$
$$\min_{\mathbf{D}_0, \mathbf{D}_1, \mathbf{D}_2} \frac{\alpha_1 - \alpha_2}{2} \log^+ \frac{1}{|\mathbf{D}_0| |\mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} - \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} \mathbf{D}_2 \mathbf{\Sigma}_{\mathbf{N}_2}^{-1}|}$$
$$+ \frac{\alpha_1}{2} \log \frac{|\mathbf{\Sigma}_{\mathbf{N}_1}|}{|\mathbf{D}_1|} + \frac{\alpha_2}{2} \log \frac{|\mathbf{\Sigma}_{\mathbf{X}_0}| |\mathbf{\Sigma}_{\mathbf{N}_2}|}{|\mathbf{D}_0| |\mathbf{D}_2|}$$
$$\text{subject to} \quad \mathbf{\Sigma}_{\mathbf{N}_1}^{-1} \mathbf{D}_1 \mathbf{\Sigma}_{\mathbf{N}_1}^{-1} + \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} \mathbf{D}_2 \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} = \mathbf{\Sigma}_{\mathbf{N}}^{-1} - \mathbf{D}_0^{-1},$$
$$\mathbf{\Sigma}_{\mathbf{N}} \prec \mathbf{D}_0 \preceq \mathbf{D},$$
$$\mathbf{0} \prec \mathbf{D}_i \preceq \mathbf{\Sigma}_{\mathbf{N}_i}, \quad i = 1, 2.$$

Proof: Let $\mathbf{W}_1$ and $\mathbf{W}_2$ be generated by $\mathbf{X}_1$ and $\mathbf{X}_2$, respectively, via paralleled Gaussian test channels, i.e., $\mathbf{W}_i = \mathbf{A}_i \mathbf{X}_i + \mathbf{U}_i$, $i = 1, 2$, where $\mathbf{A}$ is an $m \times m$ matrix and $\mathbf{U}_i$ is an $m \times 1$ Gaussian random vector with independent zero-mean unit-variance entries; moreover, we assume $\mathbf{U}_1$, $\mathbf{U}_2$, and $(\mathbf{X}_0, \mathbf{X}_1, \mathbf{X}_2)$ are mutually independent. Denote $\mathbf{\Sigma}_{\mathbf{X}_0|\mathbf{W}_1, \mathbf{W}_2}$, $\mathbf{\Sigma}_{\mathbf{X}_1|\mathbf{X}_0, \mathbf{W}_1}$, and $\mathbf{\Sigma}_{\mathbf{X}_2|\mathbf{X}_0, \mathbf{W}_2}$ by $\mathbf{D}_0$, $\mathbf{D}_1$, and $\mathbf{D}_2$, respectively. It is clear that

$$\mathbf{D}_0 \succ \mathbf{\Sigma}_{\mathbf{N}}. \tag{17}$$

Moreover, we can set $\mathbf{D}_i$ to be any matrix satisfying

$$\mathbf{0} \prec \mathbf{D}_i \preceq \mathbf{\Sigma}_{\mathbf{N}_i} \tag{18}$$

by suitably choosing $\mathbf{A}_i$, $i = 1, 2$. It can also be verified (cf., [12, eq. (5)]) that

$$\mathbf{\Sigma}_{\mathbf{N}_1}^{-1} \mathbf{D}_1 \mathbf{\Sigma}_{\mathbf{N}_1}^{-1} + \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} \mathbf{D}_2 \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} = \mathbf{\Sigma}_{\mathbf{N}}^{-1} - \mathbf{D}_0^{-1}. \tag{19}$$

Let $R_1 = I(\mathbf{X}_1; \mathbf{W}_1 | \mathbf{W}_2)$ and $R_2 = I(\mathbf{X}_2; \mathbf{W}_2)$. According to the well-known Berger–Tung inner bound [15], [16], we have $(R_1, R_2) \in \mathcal{R}(\mathbf{D})$ if

$$\mathbf{D}_0 \preceq \mathbf{D}. \tag{20}$$

Note that

$$\alpha_1 R_1 + \alpha_2 R_2$$
$$= \alpha_1 I(\mathbf{X}_1; \mathbf{W}_1 | \mathbf{W}_2) + \alpha_2 I(\mathbf{X}_2; \mathbf{W}_2)$$
$$= (\alpha_1 - \alpha_2) I(\mathbf{X}_1; \mathbf{W}_1 | \mathbf{W}_2) + \alpha_2 I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{W}_1, \mathbf{W}_2)$$
$$= (\alpha_1 - \alpha_2) I(\mathbf{X}_0, \mathbf{X}_1; \mathbf{W}_1 | \mathbf{W}_2)$$
$$\quad + \alpha_2 I(\mathbf{X}_0, \mathbf{X}_1, \mathbf{X}_2; \mathbf{W}_1, \mathbf{W}_2)$$
$$= (\alpha_1 - \alpha_2) I(\mathbf{X}_0; \mathbf{W}_1 | \mathbf{W}_2) + \alpha_1 I(\mathbf{X}_1; \mathbf{W}_1 | \mathbf{X}_0)$$
$$\quad + \alpha_2 (I(\mathbf{X}_0; \mathbf{W}_1, \mathbf{W}_2) + I(\mathbf{X}_2; \mathbf{W}_2 | \mathbf{X}_0))$$

$$= \frac{\alpha_1 - \alpha_2}{2} \log \frac{|\mathbf{\Sigma}_{\mathbf{X}_0|\mathbf{W}_2}|}{|\mathbf{D}_0|} + \frac{\alpha_1}{2} \log \frac{|\mathbf{\Sigma}_{\mathbf{N}_1}|}{|\mathbf{D}_1|}$$
$$+ \frac{\alpha_2}{2} \log \frac{|\mathbf{\Sigma}_{\mathbf{X}_0}| |\mathbf{\Sigma}_{\mathbf{N}_2}|}{|\mathbf{D}_0| |\mathbf{D}_2|}$$
$$= \frac{\alpha_1 - \alpha_2}{2} \log \frac{1}{|\mathbf{D}_0| |\mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} - \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} \mathbf{D}_2 \mathbf{\Sigma}_{\mathbf{N}_2}^{-1}|}$$
$$+ \frac{\alpha_1}{2} \log \frac{|\mathbf{\Sigma}_{\mathbf{N}_1}|}{|\mathbf{D}_1|} + \frac{\alpha_2}{2} \log \frac{|\mathbf{\Sigma}_{\mathbf{X}_0}| |\mathbf{\Sigma}_{\mathbf{N}_2}|}{|\mathbf{D}_0| |\mathbf{D}_2|} \tag{21}$$

where the last equality is due to

$$\mathbf{\Sigma}_{\mathbf{X}_0|\mathbf{W}_2} = (\mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} - \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} \mathbf{D}_2 \mathbf{\Sigma}_{\mathbf{N}_2}^{-1})^{-1}.$$

Combining (17)–(21) yields $R(\mathbf{D}, \alpha_1, \alpha_2) \leq \overline{R}'(\mathbf{D}, \alpha_1, \alpha_2)$, where

$$\overline{R}'(\mathbf{D}, \alpha_1, \alpha_2) \triangleq$$
$$\min_{\mathbf{D}_0, \mathbf{D}_1, \mathbf{D}_2} \frac{\alpha_1 - \alpha_2}{2} \log \frac{1}{|\mathbf{D}_0| |\mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} - \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} \mathbf{D}_2 \mathbf{\Sigma}_{\mathbf{N}_2}^{-1}|}$$
$$+ \frac{\alpha_1}{2} \log \frac{|\mathbf{\Sigma}_{\mathbf{N}_1}|}{|\mathbf{D}_1|} + \frac{\alpha_2}{2} \log \frac{|\mathbf{\Sigma}_{\mathbf{X}_0}| |\mathbf{\Sigma}_{\mathbf{N}_2}|}{|\mathbf{D}_0| |\mathbf{D}_2|}$$
$$\text{subject to} \quad \mathbf{\Sigma}_{\mathbf{N}_1}^{-1} \mathbf{D}_1 \mathbf{\Sigma}_{\mathbf{N}_1}^{-1} + \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} \mathbf{D}_2 \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} \preceq \mathbf{\Sigma}_{\mathbf{N}}^{-1} - \mathbf{D}_0^{-1},$$
$$\mathbf{\Sigma}_{\mathbf{N}} \prec \mathbf{D}_0 \preceq \mathbf{D},$$
$$\mathbf{0} \prec \mathbf{D}_i \preceq \mathbf{\Sigma}_{\mathbf{N}_i}, \quad i = 1, 2.$$

Since $\mathbf{\Sigma}_{\mathbf{X}_0|\mathbf{W}_1, \mathbf{W}_2} \preceq \mathbf{\Sigma}_{\mathbf{X}_0|\mathbf{W}_2}$, it follows that

$$\log \frac{1}{|\mathbf{D}_0| |\mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} - \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} \mathbf{D}_2 \mathbf{\Sigma}_{\mathbf{N}_2}^{-1}|}$$
$$= \log^+ \frac{1}{|\mathbf{D}_0| |\mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} - \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} \mathbf{D}_2 \mathbf{\Sigma}_{\mathbf{N}_2}^{-1}|}$$

and the proof is complete. ∎

C. Matching Conditions

In this section, we shall investigate the conditions under which $\underline{R}(\mathbf{D}, \alpha_1, \alpha_2) = \overline{R}(\mathbf{D}, \alpha_1, \alpha_2)$. It is easy to see that without loss of optimality one can set $\mathbf{D}_0 = \mathbf{D}$ in the minimization problem associated with $\underline{R}(\mathbf{D}, \alpha_1, \alpha_2)$. Therefore, $\underline{R}(\mathbf{D}, \alpha_1, \alpha_2)$ can be defined equivalently as

(**P**) $\quad \underline{R}(\mathbf{D}, \alpha_1, \alpha_2) \triangleq$
$$\min_{\mathbf{D}_1, \mathbf{D}_2} \frac{\alpha_1 - \alpha_2}{2} \log^+ \frac{1}{|\mathbf{D}| |\mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} - \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} \mathbf{D}_2 \mathbf{\Sigma}_{\mathbf{N}_2}^{-1}|}$$
$$+ \frac{\alpha_1}{2} \log \frac{|\mathbf{\Sigma}_{\mathbf{N}_1}|}{|\mathbf{D}_1|} + \frac{\alpha_2}{2} \log \frac{|\mathbf{\Sigma}_{\mathbf{X}_0}| |\mathbf{\Sigma}_{\mathbf{N}_2}|}{|\mathbf{D}| |\mathbf{D}_2|}$$
$$\text{subject to} \quad \mathbf{\Sigma}_{\mathbf{N}_1}^{-1} \mathbf{D}_1 \mathbf{\Sigma}_{\mathbf{N}_1}^{-1} + \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} \mathbf{D}_2 \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} \preceq \mathbf{\Sigma}_{\mathbf{N}}^{-1} - \mathbf{D}^{-1},$$
$$\mathbf{0} \prec \mathbf{D}_i \preceq \mathbf{\Sigma}_{\mathbf{N}_i}, \quad i = 1, 2.$$

Moreover, it is clear that if an optimal solution $(\hat{\mathbf{D}}_1, \hat{\mathbf{D}}_2)$ to (**P**) satisfies

$$\mathbf{\Sigma}_{\mathbf{N}_1}^{-1} \hat{\mathbf{D}}_1 \mathbf{\Sigma}_{\mathbf{N}_1}^{-1} + \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} \hat{\mathbf{D}}_2 \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} = \mathbf{\Sigma}_{\mathbf{N}}^{-1} - \mathbf{D}^{-1} \tag{22}$$

then $(\hat{\mathbf{D}}_0, \hat{\mathbf{D}}_1, \hat{\mathbf{D}}_2)$ with $\hat{\mathbf{D}}_0 = \mathbf{D}$ is an optimal solution to the minimization problem associated with $\overline{R}(\mathbf{D}, \alpha_1, \alpha_2)$, and con-

sequently $\underline{R}(\mathbf{D}, \alpha_1, \alpha_2) = \overline{R}(\mathbf{D}, \alpha_1, \alpha_2)$. It can be verified that the matching condition (22) is satisfied in the scalar case (i.e., $m = 1$).

To facilitate analysis, we replace $(\mathbf{P})$ with the following minimization problem:

$$(\tilde{\mathbf{P}}) \qquad \tilde{\underline{R}}(\mathbf{D}, \alpha_1, \alpha_2) \triangleq$$
$$\min_{\mathbf{D}_1, \mathbf{D}_2} \frac{\alpha_1 - \alpha_2}{2} \log \frac{1}{|\mathbf{D}||\mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} - \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} \mathbf{D}_2 \mathbf{\Sigma}_{\mathbf{N}_2}^{-1}|}$$
$$+ \frac{\alpha_1}{2} \log \frac{|\mathbf{\Sigma}_{\mathbf{N}_1}|}{|\mathbf{D}_1|} + \frac{\alpha_2}{2} \log \frac{|\mathbf{\Sigma}_{\mathbf{X}_0}||\mathbf{\Sigma}_{\mathbf{N}_2}|}{|\mathbf{D}||\mathbf{D}_2|}$$
$$\text{subject to} \quad \mathbf{\Sigma}_{\mathbf{N}_1}^{-1} \mathbf{D}_1 \mathbf{\Sigma}_{\mathbf{N}_1}^{-1} + \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} \mathbf{D}_2 \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} \preceq \mathbf{\Sigma}_{\mathbf{N}}^{-1} - \mathbf{D}^{-1},$$
$$\mathbf{0} \prec \mathbf{D}_i \preceq \mathbf{\Sigma}_{\mathbf{N}_i}, \quad i = 1, 2.$$

Note that $(\tilde{\mathbf{P}})$ provides a lower bound on $\underline{R}(\mathbf{D}, \alpha_1, \alpha_2)$; furthermore, if an optimal solution $(\hat{\mathbf{D}}_1, \hat{\mathbf{D}}_2)$ to $(\tilde{\mathbf{P}})$ satisfies (22), then it is also an optimal solution to $(\mathbf{P})$ and consequently $\underline{R}(\mathbf{D}, \alpha_1, \alpha_2) = \overline{R}(\mathbf{D}, \alpha_1, \alpha_2)$. Ignoring the constant terms, one can write the Lagrangian of $(\tilde{\mathbf{P}})$ as

$$L(\mathbf{D}_1, \mathbf{D}_2)$$
$$= -(\alpha_1 - \alpha_2) \log |\mathbf{\Sigma}_{\mathbf{N}_2} \mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} \mathbf{\Sigma}_{\mathbf{N}_2} - \mathbf{D}_2|$$
$$\quad - \alpha_1 \log |\mathbf{D}_1| - \alpha_2 \log |\mathbf{D}_2|$$
$$\quad + \text{tr}[\mathbf{\Lambda}(\mathbf{\Sigma}_{\mathbf{N}_1}^{-1} \mathbf{D}_1 \mathbf{\Sigma}_{\mathbf{N}_1}^{-1} + \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} \mathbf{D}_2 \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} - \mathbf{\Sigma}_{\mathbf{N}}^{-1} + \mathbf{D}^{-1})]$$
$$\quad + \text{tr}[\mathbf{\Pi}_1(\mathbf{D}_1 - \mathbf{\Sigma}_{\mathbf{N}_1})] + \text{tr}[\mathbf{\Pi}_2(\mathbf{D}_2 - \mathbf{\Sigma}_{\mathbf{N}_2})]$$

where $\mathbf{\Lambda}$, $\mathbf{\Pi}_1$, and $\mathbf{\Pi}_2$ are positive semidefinite matrices. Since $(\tilde{\mathbf{P}})$ is a convex semidefinite programming problem, $(\hat{\mathbf{D}}_1, \hat{\mathbf{D}}_2)$ is an optimal solution to $(\tilde{\mathbf{P}})$ if it satisfies the following KKT conditions:

$$\nabla_{\mathbf{D}_1} L(\mathbf{D}_1, \mathbf{D}_2)|_{\mathbf{D}_1 = \hat{\mathbf{D}}_1}$$
$$= -\alpha_1 \hat{\mathbf{D}}_1^{-1} + \mathbf{\Sigma}_{\mathbf{N}_1}^{-1} \mathbf{\Lambda} \mathbf{\Sigma}_{\mathbf{N}_1}^{-1} + \mathbf{\Pi}_1 = \mathbf{0},$$
$$\nabla_{\mathbf{D}_2} L(\mathbf{D}_1, \mathbf{D}_2)|_{\mathbf{D}_2 = \hat{\mathbf{D}}_2}$$
$$= (\alpha_1 - \alpha_2)(\mathbf{\Sigma}_{\mathbf{N}_2} \mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} \mathbf{\Sigma}_{\mathbf{N}_2} - \hat{\mathbf{D}}_2)^{-1}$$
$$\quad - \alpha_2 \hat{\mathbf{D}}_2^{-1} + \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} \mathbf{\Lambda} \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} + \mathbf{\Pi}_2 = \mathbf{0},$$
$$\mathbf{\Lambda}(\mathbf{\Sigma}_{\mathbf{N}_1}^{-1} \hat{\mathbf{D}}_1 \mathbf{\Sigma}_{\mathbf{N}_1}^{-1} + \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} \hat{\mathbf{D}}_2 \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} - \mathbf{\Sigma}_{\mathbf{N}}^{-1} + \mathbf{D}^{-1}) = \mathbf{0},$$
$$\mathbf{\Pi}_i(\hat{\mathbf{D}}_i - \mathbf{\Sigma}_{\mathbf{N}_i}) = \mathbf{0}, \quad i = 1, 2,$$
$$\mathbf{\Sigma}_{\mathbf{N}_1}^{-1} \hat{\mathbf{D}}_1 \mathbf{\Sigma}_{\mathbf{N}_1}^{-1} + \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} \hat{\mathbf{D}}_2 \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} \preceq \mathbf{\Sigma}_{\mathbf{N}}^{-1} - \mathbf{D}^{-1},$$
$$\mathbf{0} \prec \hat{\mathbf{D}}_i \preceq \mathbf{\Sigma}_{\mathbf{N}_i}, \quad i = 1, 2.$$

By setting $\mathbf{\Lambda} = \alpha_1 \mathbf{\Sigma}_{\mathbf{N}_1} \hat{\mathbf{D}}_1^{-1} \mathbf{\Sigma}_{\mathbf{N}_1}$ and $\mathbf{\Pi}_1 = \mathbf{\Pi}_2 = \mathbf{0}$, we can simplify the KKT conditions as

$$(\alpha_1 - \alpha_2)(\mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} - \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} \hat{\mathbf{D}}_2 \mathbf{\Sigma}_{\mathbf{N}_2}^{-1})^{-1} + \alpha_1 \mathbf{\Sigma}_{\mathbf{N}_1} \hat{\mathbf{D}}_1^{-1} \mathbf{\Sigma}_{\mathbf{N}_1}$$
$$- \alpha_2 \mathbf{\Sigma}_{\mathbf{N}_2} \hat{\mathbf{D}}_2^{-1} \mathbf{\Sigma}_{\mathbf{N}_2} = \mathbf{0} \qquad (23)$$
$$\mathbf{\Sigma}_{\mathbf{N}_1}^{-1} \hat{\mathbf{D}}_1 \mathbf{\Sigma}_{\mathbf{N}_1}^{-1} + \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} \hat{\mathbf{D}}_2 \mathbf{\Sigma}_{\mathbf{N}_2}^{-1} - \mathbf{\Sigma}_{\mathbf{N}}^{-1} + \mathbf{D}^{-1} = \mathbf{0} \qquad (24)$$
$$\mathbf{0} \prec \hat{\mathbf{D}}_i \preceq \mathbf{\Sigma}_{\mathbf{N}_i}, \quad i = 1, 2. \qquad (25)$$

It can be readily shown by comparing (22) and (24) that $\underline{R}(\mathbf{D}, \alpha_1, \alpha_2) = \overline{R}(\mathbf{D}, \alpha_1, \alpha_2)$ if there exists $(\hat{\mathbf{D}}_1, \hat{\mathbf{D}}_2)$ satisfying (23)–(25). For the special case $\alpha_1 = \alpha_2 = 1$ (which

corresponds to the sum rate), one can verify that the solution to the matrix (23) and (24) is given by

$$\hat{\mathbf{D}}_i = \frac{1}{2} \mathbf{\Sigma}_{\mathbf{N}_i}(\mathbf{\Sigma}_{\mathbf{N}}^{-1} - \mathbf{D}^{-1}) \mathbf{\Sigma}_{\mathbf{N}_i}, \quad i = 1, 2.$$

Substituting this solution into (25), we see that $\underline{R}(\mathbf{D}, 1, 1) = \overline{R}(\mathbf{D}, 1, 1)$ if $\mathbf{\Sigma}_{\mathbf{N}}^{-1} - \mathbf{D}^{-1} \preceq 2\mathbf{\Sigma}_{\mathbf{N}_i}^{-1}, i = 1, 2$. This matching condition was first derived in [10]. Now consider the general case $\alpha_1 \geq \alpha_2 > 0$. Define $\mathbf{A} = \mathbf{\Sigma}_{\mathbf{N}}^{-1} - \mathbf{D}^{-1}$, $\mathbf{B} = \mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1}$, and $\bar{\mathbf{D}}_i = \mathbf{\Sigma}_{\mathbf{N}_i}^{-1} \hat{\mathbf{D}}_i \mathbf{\Sigma}_{\mathbf{N}_i}^{-1}, i = 1, 2$. We can rewrite (23)–(25) as

$$(\alpha_1 - \alpha_2)(\mathbf{B} - \bar{\mathbf{D}}_2)^{-1} + \alpha_1 \bar{\mathbf{D}}_1^{-1} - \alpha_2 \bar{\mathbf{D}}_2^{-1} = \mathbf{0} \qquad (26)$$
$$\bar{\mathbf{D}}_1 + \bar{\mathbf{D}}_2 - \mathbf{A} = \mathbf{0} \qquad (27)$$
$$\mathbf{0} \prec \bar{\mathbf{D}}_i \preceq \mathbf{\Sigma}_{\mathbf{N}_i}^{-1}, \quad i = 1, 2. \qquad (28)$$

Combining (26) and (27) gives

$$(\alpha_1 - \alpha_2)(\mathbf{B} - \bar{\mathbf{D}}_2)^{-1} + \alpha_1(\mathbf{A} - \bar{\mathbf{D}}_2)^{-1} - \alpha_2 \bar{\mathbf{D}}_2^{-1} = \mathbf{0}.$$

Note that

$$(\alpha_1 - \alpha_2)(\mathbf{B} - \bar{\mathbf{D}}_2)^{-1} + \alpha_1(\mathbf{A} - \bar{\mathbf{D}}_2)^{-1} - \alpha_2 \bar{\mathbf{D}}_2^{-1} = \mathbf{0}$$
$$\Longleftrightarrow \quad (\alpha_1 - \alpha_2)(\mathbf{B}\bar{\mathbf{D}}_2^{-1} - \mathbf{I}_m)^{-1} + \alpha_1(\mathbf{A}\bar{\mathbf{D}}_2^{-1} - \mathbf{I}_m)^{-1}$$
$$\qquad - \alpha_2 \mathbf{I}_m = \mathbf{0}$$
$$\Longleftrightarrow \quad (\alpha_1 - \alpha_2)(\mathbf{A}\bar{\mathbf{D}}_2^{-1} - \mathbf{I}_m) + \alpha_1(\mathbf{B}\bar{\mathbf{D}}_2^{-1} - \mathbf{I}_m)$$
$$\qquad - \alpha_2(\mathbf{A}\bar{\mathbf{D}}_2^{-1} - \mathbf{I}_m)(\mathbf{B}\bar{\mathbf{D}}_2^{-1} - \mathbf{I}_m) = \mathbf{0}$$
$$\Longleftrightarrow \quad (\alpha_1 - \alpha_2)(\mathbf{A}\bar{\mathbf{D}}_2^{-1} - \mathbf{I}_m) + \alpha_1(\mathbf{B}\bar{\mathbf{D}}_2^{-1} - \mathbf{I}_m)$$
$$\qquad - \alpha_2(\mathbf{A}\bar{\mathbf{D}}_2^{-1}\mathbf{B}\bar{\mathbf{D}}_2^{-1} - \mathbf{B}\bar{\mathbf{D}}_2^{-1} - \mathbf{A}\bar{\mathbf{D}}_2^{-1} + \mathbf{I}_m) = \mathbf{0}$$
$$\Longleftrightarrow \quad (\alpha_1 - \alpha_2)(\mathbf{B}^{\frac{1}{2}}\bar{\mathbf{D}}_2^{-1}\mathbf{B}^{\frac{1}{2}} - \mathbf{B}^{\frac{1}{2}}\mathbf{A}^{-1}\mathbf{B}^{\frac{1}{2}})$$
$$\qquad + \alpha_1(\mathbf{B}^{\frac{1}{2}}\mathbf{A}^{-1}\mathbf{B}\bar{\mathbf{D}}_2^{-1}\mathbf{B}^{\frac{1}{2}} - \mathbf{B}^{\frac{1}{2}}\mathbf{A}^{-1}\mathbf{B}^{\frac{1}{2}})$$
$$\qquad - \alpha_2 \left(\mathbf{B}^{\frac{1}{2}}\bar{\mathbf{D}}_2^{-1}\mathbf{B}\bar{\mathbf{D}}_2^{-1}\mathbf{B}^{\frac{1}{2}} - \mathbf{B}^{\frac{1}{2}}\mathbf{A}^{-1}\mathbf{B}\bar{\mathbf{D}}_2^{-1}\mathbf{B}^{\frac{1}{2}} \right.$$
$$\qquad \left. - \mathbf{B}^{\frac{1}{2}}\bar{\mathbf{D}}_2^{-1}\mathbf{B}^{\frac{1}{2}} + \mathbf{B}^{\frac{1}{2}}\mathbf{A}^{-1}\mathbf{B}^{\frac{1}{2}} \right) = \mathbf{0}.$$

Setting $\mathbf{G} = \mathbf{B}^{\frac{1}{2}}\bar{\mathbf{D}}_2^{-1}\mathbf{B}^{\frac{1}{2}}$, $\mathbf{C} = \frac{\alpha_1}{\alpha_2}\mathbf{I}_m + (1 + \frac{\alpha_1}{\alpha_2})\mathbf{B}^{\frac{1}{2}}\mathbf{A}^{-1}\mathbf{B}^{\frac{1}{2}}$, and $\mathbf{Q} = \frac{2\alpha_1}{\alpha_2}\mathbf{B}^{\frac{1}{2}}\mathbf{A}^{-1}\mathbf{B}^{\frac{1}{2}}$, we obtain the following quadratic matrix equation:

$$\mathbf{G}^2 - \mathbf{C}\mathbf{G} + \mathbf{Q} = \mathbf{0}. \qquad (29)$$

For the purpose of solving (29), we consider a closely related matrix equation

$$\mathbf{G}^2 - \frac{1}{2}\mathbf{C}\mathbf{G} - \frac{1}{2}\mathbf{G}\mathbf{C} + \mathbf{Q} = \mathbf{0}$$

which can be written equivalently as

$$\left(\mathbf{G} - \frac{1}{2}\mathbf{C}\right)^2 = \frac{1}{4}\mathbf{C}^2 - \mathbf{Q}. \qquad (30)$$
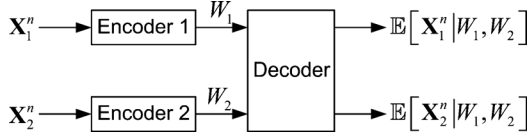
Fig. 2. Direct vector Gaussian two-terminal source coding problem.

Let $\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T$ be the eigenvalue decomposition of $\mathbf{B}^{\frac{1}{2}}\mathbf{A}^{-1}\mathbf{B}^{\frac{1}{2}}$, where $\mathbf{U}$ is a unitary matrix and $\boldsymbol{\Lambda} \triangleq \mathrm{diag}(\lambda_1,\ldots,\lambda_m)$ is a positive-definite diagonal matrix. Note that

$$\frac{1}{4}\mathbf{C}^2 - \mathbf{Q}$$
$$= \mathbf{U}\left[\frac{1}{4}(\frac{\alpha_1}{\alpha_2}\mathbf{I}_m + (1+\frac{\alpha_1}{\alpha_2})\boldsymbol{\Lambda})^2 - \frac{2\alpha_1}{\alpha_2}\boldsymbol{\Lambda}\right]\mathbf{U}^T$$
$$= \mathbf{U}\left[\frac{1}{4}(1+\frac{\alpha_1}{\alpha_2})^2\boldsymbol{\Lambda}^2 + \frac{1}{2}\frac{\alpha_1}{\alpha_2}(\frac{\alpha_1}{\alpha_2}-3)\boldsymbol{\Lambda} + \frac{1}{4}(\frac{\alpha_1}{\alpha_2})^2\mathbf{I}_m\right]\mathbf{U}^T.$$

Since $\frac{\alpha_1}{\alpha_2} \geq 1$, it follows that

$$\frac{1}{4}(1+\frac{\alpha_1}{\alpha_2})^2\lambda^2 + \frac{1}{2}\frac{\alpha_1}{\alpha_2}(\frac{\alpha_1}{\alpha_2}-3)\lambda + \frac{1}{4}(\frac{\alpha_1}{\alpha_2})^2 \geq 0$$

for all λ. Therefore, $\frac{1}{4}\mathbf{C}^2 - \mathbf{Q}$ is positive semidefinite and $(\frac{1}{4}\mathbf{C}^2 - \mathbf{Q})^{\frac{1}{2}}$ is well defined. Note that $(\frac{1}{4}\mathbf{C}^2 - \mathbf{Q})^{\frac{1}{2}} = \mathbf{U}\boldsymbol{\Theta}\mathbf{U}^T$, where $\boldsymbol{\Theta} = \mathrm{diag}(\theta_1,\ldots,\theta_m)$ with

$$\theta_i = \left[\frac{1}{4}(1+\frac{\alpha_1}{\alpha_2})^2\lambda_i^2 + \frac{1}{2}\frac{\alpha_1}{\alpha_2}(\frac{\alpha_1}{\alpha_2}-3)\lambda_i + \frac{1}{4}(\frac{\alpha_1}{\alpha_2})^2\right]^{\frac{1}{2}},$$
$$i = 1,\ldots,m.$$

Let $\hat{\mathbf{G}} = \frac{1}{2}\mathbf{C} + (\frac{1}{4}\mathbf{C}^2 - \mathbf{Q})^{\frac{1}{2}}$. It is easy to see that $\hat{\mathbf{G}}$ is a solution to (30). Moreover, it can be verified that $\mathbf{C}\hat{\mathbf{G}} = \hat{\mathbf{G}}\mathbf{C}$; as a consequence, $\hat{\mathbf{G}}$ is also a solution to (29). Let $\bar{\mathbf{D}}_2 = \mathbf{B}^{\frac{1}{2}}\hat{\mathbf{G}}^{-1}\mathbf{B}^{\frac{1}{2}}$ and $\bar{\mathbf{D}}_1 = \mathbf{A} - \bar{\mathbf{D}}_2$. Since $\hat{\mathbf{G}} \succeq \frac{1}{2}\mathbf{C} \succ \mathbf{B}^{\frac{1}{2}}\mathbf{A}^{-1}\mathbf{B}^{\frac{1}{2}}$, it follows that $\bar{\mathbf{D}}_i \succ \mathbf{0}$, $i = 1,2$. Therefore, if $\bar{\mathbf{D}}_i \preceq \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_i}^{-1}$, $i = 1,2$, then (26)–(28) are all satisfied, and consequently, $\underline{R}(\mathbf{D},\alpha_1,\alpha_2) = \overline{R}(\mathbf{D},\alpha_1,\alpha_2)$. A simple sufficient condition is $\mathbf{A} \preceq \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_i}^{-1}$, $i = 1,2$. In view of the fact that $\mathbf{A} = \boldsymbol{\Sigma}_{\mathbf{N}}^{-1} - \mathbf{D}^{-1}$, this sufficient condition is satisfied in the high-resolution regime[3]

$$\mathbf{D} \preceq (\boldsymbol{\Sigma}_{\mathbf{N}}^{-1} - \boldsymbol{\Sigma}_{\mathbf{N}_i}^{-1})^{-1}, \quad i = 1,2. \tag{31}$$

It is worth noting that (31) does not depend on (α_1,α_2) and thus provides a condition[4] under which the rate region $\mathcal{R}(\mathbf{D})$ is completely characterized. We summarize this result in the following theorem.

Theorem 4: $\underline{R}(\mathbf{D},\alpha_1,\alpha_2) = \overline{R}(\mathbf{D},\alpha_1,\alpha_2)$ if $\mathbf{D} \preceq (\boldsymbol{\Sigma}_{\mathbf{N}}^{-1} - \boldsymbol{\Sigma}_{\mathbf{N}_i}^{-1})^{-1}$, $i = 1,2$.

IV. DIRECT VECTOR GAUSSIAN TWO-TERMINAL SOURCE CODING PROBLEM

In this section, we shall apply the new bounding technique to the direct vector Gaussian two-terminal source coding problem (see Fig. 2 for the system model).

Let m, m_1, and m_2 be three positive integers with $m = m_1 + m_2$. For any $m \times m$ matrix $\mathbf{A} = (a_{j,k})$, we denote its

[3]It is easy to verify that $(\boldsymbol{\Sigma}_{\mathbf{N}}^{-1} - \boldsymbol{\Sigma}_{\mathbf{N}_i}^{-1})^{-1} = \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_i} = \boldsymbol{\Sigma}_{\mathbf{x}_0|\mathbf{x}_i}$, $i = 1,2$.

[4]Due to its inherent symmetry, this condition does not rely on the assumption $\alpha_1 \geq \alpha_2$.

first diagonal submatrix $(a_{j,k})_{1 \leq j,k \leq m_1}$ by $(\mathbf{A})_1$ and its second diagonal submatrix $(a_{j,k})_{m_1+1 \leq j,k \leq m}$ by $(\mathbf{A})_2$, and call $\mathbf{A}$ an (m_1,m_2) block diagonal matrix if

$$\mathbf{A} = \mathrm{diag}((\mathbf{A})_1,(\mathbf{A})_2) \triangleq \begin{pmatrix} (\mathbf{A})_1 & 0 \\ 0 & (\mathbf{A})_2 \end{pmatrix}.$$

Let $\mathbf{X} = (\mathbf{X}_1^T, \mathbf{X}_2^T)^T$ be an $m \times 1$ Gaussian random vector with mean zero and positive-definite covariance matrix $\boldsymbol{\Sigma}_{\mathbf{X}}$, where $\mathbf{X}_i$ is an $m_i \times 1$ Gaussian random vector with mean zero and positive-definite covariance matrix $\boldsymbol{\Sigma}_{\mathbf{X}_i} = (\boldsymbol{\Sigma}_{\mathbf{X}})_i$, $i = 1,2$. Let $\{(\mathbf{X}(t),\mathbf{X}_1(t),\mathbf{X}_2(t))\}_{t=1}^{\infty}$ be i.i.d. copies of $(\mathbf{X},\mathbf{X}_1,\mathbf{X}_2)$.

Definition 2: A rate pair (R_1,R_2) is said to be achievable subject to distortion constraint $(\mathbf{D}_1,\mathbf{D}_2)$ if for all sufficiently large n, there exist encoding functions $f_i^{(n)} : \mathbb{R}^{m_i \times n} \to \{1,2,\ldots,M_i\}$, $i = 1,2$, such that

$$\frac{1}{n}\log M_i \leq R_i, \quad i = 1,2$$
$$\boldsymbol{\Sigma}_{\mathbf{X}_i^n|W_1,W_2} \preceq \mathbf{D}_i, \quad i = 1,2$$

where $W_i = f_i^{(n)}(\mathbf{X}_i^n)$, $i = 1,2$. The rate region $\mathcal{R}(\mathbf{D}_1,\mathbf{D}_2)$ is the closure of the set of all achievable rate pairs subject to distortion constraint $(\mathbf{D}_1,\mathbf{D}_2)$.

We shall assume $\mathbf{0} \prec \mathbf{D}_i \preceq \boldsymbol{\Sigma}_{\mathbf{X}_i}$, $i = 1,2$, which is justified by the fact that $\mathbf{0} \prec \boldsymbol{\Sigma}_{\mathbf{X}_i^n|W_1,W_2} \preceq \boldsymbol{\Sigma}_{\mathbf{X}_i}$, $i = 1,2$. Since $\mathcal{R}(\mathbf{D})$ is a (closed) convex set, it is completely characterized by its supporting lines. Define

$$R(\mathbf{D}_1,\mathbf{D}_2,\alpha_1,\alpha_2) = \inf_{(R_1,R_2) \in \mathcal{R}(\mathbf{D}_1,\mathbf{D}_2)} \alpha_1 R_1 + \alpha_2 R_2$$

where α_1 and α_2 are nonnegative real numbers. Following the proof strategy for the CEO problem, we shall bound $\mathcal{R}(\mathbf{D}_1,\mathbf{D}_2)$ by establishing upper and lower bounds on $R(\mathbf{D}_1,\mathbf{D}_2,\alpha_1,\alpha_2)$. Without loss of generality, we assume $\alpha_1 \geq \alpha_2 \geq 0$. For the degenerate case $\alpha_2 = 0$, it can be shown [cf., the proof of (6)] that

$$R(\mathbf{D}_1,\mathbf{D}_2,\alpha_1,0) = \min_{\tilde{\mathbf{D}}} \frac{\alpha_1}{2}\log\frac{|\boldsymbol{\Sigma}_{\mathbf{X}_1|\mathbf{X}_2}|}{|\tilde{\mathbf{D}}|}$$
$$\text{subject to} \quad \tilde{\mathbf{D}} \preceq \mathbf{D}_1,$$
$$\mathbf{0} \prec \tilde{\mathbf{D}} \preceq \boldsymbol{\Sigma}_{\mathbf{X}_1|\mathbf{X}_2}.$$

Hence, we shall focus on the case $\alpha_1 \geq \alpha_2 > 0$ in the rest of this section.

For the purpose of subsequent analysis, we augment the probability space by introducing random objects which are not necessarily functions of the existing ones. Specifically, let $\mathbf{N}$ be an $m \times 1$ Gaussian random vector, independent of $\mathbf{X}$, with mean zero and positive-definite covariance matrix $\boldsymbol{\Sigma}_{\mathbf{N}}$ and let $\{(\mathbf{X}(t),\mathbf{X}_1(t),\mathbf{X}_2(t),\mathbf{N}(t))\}_{t=1}^{\infty}$ be i.i.d. copies of $(\mathbf{X},\mathbf{X}_1,\mathbf{X}_2,\mathbf{N})$. Furthermore, we define $\mathbf{Y} = \mathbf{X} + \mathbf{N}$, $\hat{\mathbf{X}}_i = \mathbb{E}[\mathbf{X}_i|\mathbf{Y}]$, $\hat{\mathbf{N}}_i = \mathbf{X}_i - \hat{\mathbf{X}}_i$, $\tilde{\mathbf{X}}_i = \mathbb{E}[\hat{\mathbf{X}}_i|\mathbf{X}_i]$, $\tilde{\mathbf{N}}_i = \hat{\mathbf{X}}_i - \tilde{\mathbf{X}}_i$, $i = 1,2$, and define $\mathbf{N}(t),\mathbf{Y}(t),\hat{\mathbf{X}}_i(t)$, $\hat{\mathbf{N}}_i(t),\tilde{\mathbf{X}}_i(t),\tilde{\mathbf{N}}_i(t)$, $i = 1,2$, $t \geq 1$, accordingly. It is clear that $\tilde{\mathbf{X}}_i$ and $\hat{\mathbf{N}}_i$ are mutually independent, $i = 1,2$; similarly, $\tilde{\mathbf{X}}_i$ and $\tilde{\mathbf{N}}_i$ are mutually independent, $i = 1,2$. Let $\mathbb{N}(\boldsymbol{\Sigma}_{\mathbf{X}})$ be the set of positive-definite matrices $\boldsymbol{\Sigma}$ such that $\boldsymbol{\Sigma}_{\mathbf{X}}^{-1} + \boldsymbol{\Sigma}^{-1}$ is an (m_1,m_2) block diagonal matrix. Note that if $\boldsymbol{\Sigma}_{\mathbf{N}} \in \mathbb{N}(\boldsymbol{\Sigma}_{\mathbf{X}})$,

then $(\mathbf{X}_1, \tilde{\mathbf{X}}_1) \leftrightarrow \hat{\mathbf{X}}_1 \leftrightarrow \mathbf{Y} \leftrightarrow \hat{\mathbf{X}}_2 \leftrightarrow (\mathbf{X}_2, \tilde{\mathbf{X}}_2)$ form a Markov chain.

A. Lower and Upper Bounds

We shall establish a lower bound and an upper bound on $R(\mathbf{D}_1, \mathbf{D}_2, \alpha_1, \alpha_2)$ in this section.

Theorem 5: $R(\mathbf{D}_1, \mathbf{D}_2, \alpha_1, \alpha_2) \geq \underline{R}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma}_{\mathbf{N}}, \alpha_1, \alpha_2)$ for any $\mathbf{\Sigma}_{\mathbf{N}} \in \mathbb{N}(\mathbf{\Sigma}_{\mathbf{X}})$, where

$$\underline{R}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma}_{\mathbf{N}}, \alpha_1, \alpha_2) \triangleq$$
$$\min_{\mathbf{D}, \mathbf{\Gamma}_1, \mathbf{\Gamma}_2} \frac{\alpha_1 - \alpha_2}{2} \log^+ \frac{|\mathbf{\Sigma}_{\mathbf{X}} + \mathbf{\Sigma}_{\mathbf{N}}|}{|\mathbf{D} + \mathbf{\Sigma}_{\mathbf{N}}||\mathbf{\Sigma}_{\hat{\mathbf{X}}_2}||\mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} - \mathbf{\Sigma}_{\hat{\mathbf{N}}_2}^{-1} \mathbf{\Gamma}_2 \mathbf{\Sigma}_{\hat{\mathbf{N}}_2}^{-1}|}$$
$$+ \frac{\alpha_1}{2} \log \frac{|\mathbf{\Sigma}_{\hat{\mathbf{N}}_1}|}{|\mathbf{\Gamma}_1|} + \frac{\alpha_2}{2} \log \frac{|\mathbf{\Sigma}_{\mathbf{X}} + \mathbf{\Sigma}_{\mathbf{N}}||\mathbf{\Sigma}_{\hat{\mathbf{N}}_2}|}{|\mathbf{D} + \mathbf{\Sigma}_{\mathbf{N}}||\mathbf{\Gamma}_2|}$$
$$\text{subject to} \quad \text{diag}(\mathbf{\Gamma}_1, \mathbf{\Gamma}_2) \preceq (\mathbf{D}^{-1} + \mathbf{\Sigma}_{\mathbf{N}}^{-1})^{-1}, \quad (32)$$
$$0 \prec \mathbf{D} \preceq \mathbf{\Sigma}_{\mathbf{X}},$$
$$(\mathbf{D})_i \preceq \mathbf{D}_i, \quad i = 1, 2,$$
$$0 \prec \mathbf{\Gamma}_i \preceq \mathbf{\Sigma}_{\hat{\mathbf{N}}_i}, \quad i = 1, 2.$$

Proof: The proof is similar to that of Theorem 2 and is relegated to Appendix C. ∎

Theorem 6: $R(\mathbf{D}_1, \mathbf{D}_2, \alpha_1, \alpha_2) \leq \overline{R}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma}_{\mathbf{N}}, \alpha_1, \alpha_2)$ for any $\mathbf{\Sigma}_{\mathbf{N}} \in \mathbb{N}(\mathbf{\Sigma}_{\mathbf{X}})$, where

$$\overline{R}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma}_{\mathbf{N}}, \alpha_1, \alpha_2) \triangleq$$
$$\min_{\mathbf{D}, \mathbf{\Gamma}_1, \mathbf{\Gamma}_2} \frac{\alpha_1 - \alpha_2}{2} \log^+ \frac{|\mathbf{\Sigma}_{\mathbf{X}} + \mathbf{\Sigma}_{\mathbf{N}}|}{|\mathbf{D} + \mathbf{\Sigma}_{\mathbf{N}}||\mathbf{\Sigma}_{\hat{\mathbf{X}}_2}||\mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} - \mathbf{\Sigma}_{\hat{\mathbf{N}}_2}^{-1} \mathbf{\Gamma}_2 \mathbf{\Sigma}_{\hat{\mathbf{N}}_2}^{-1}|}$$
$$+ \frac{\alpha_1}{2} \log \frac{|\mathbf{\Sigma}_{\hat{\mathbf{N}}_1}|}{|\mathbf{\Gamma}_1|} + \frac{\alpha_2}{2} \log \frac{|\mathbf{\Sigma}_{\mathbf{X}} + \mathbf{\Sigma}_{\mathbf{N}}||\mathbf{\Sigma}_{\hat{\mathbf{N}}_2}|}{|\mathbf{D} + \mathbf{\Sigma}_{\mathbf{N}}||\mathbf{\Gamma}_2|}$$
$$\text{subject to} \quad \text{diag}(\mathbf{\Gamma}_1, \mathbf{\Gamma}_2) = (\mathbf{D}^{-1} + \mathbf{\Sigma}_{\mathbf{N}}^{-1})^{-1}, \quad (33)$$
$$0 \prec \mathbf{D} \preceq \mathbf{\Sigma}_{\mathbf{X}},$$
$$(\mathbf{D})_i \preceq \mathbf{D}_i, \quad i = 1, 2,$$
$$0 \prec \mathbf{\Gamma}_i \preceq \mathbf{\Sigma}_{\hat{\mathbf{N}}_i}, \quad i = 1, 2.$$

Proof: Due to the similarity to the proof of Theorem 3, we only give an abbreviated proof here. Let $\mathbf{W}_1$ and $\mathbf{W}_2$ be generated by $\mathbf{X}_1$ and $\mathbf{X}_2$, respectively, via paralleled Gaussian test channels. Specifically, $\mathbf{W}_i = \mathbf{A}_i \mathbf{X}_i + \mathbf{U}_i$, where $\mathbf{A}_i$ is an $m_i \times m_i$ matrix and $\mathbf{U}_i$ is an $m_i \times 1$ Gaussian random vector with independent zero-mean unit-variance entries, $i = 1, 2$; moreover, it is assumed that $\mathbf{U}_1$, $\mathbf{U}_2$, and $\mathbf{X}$ are mutually independent. Now pick an arbitrary $\mathbf{\Sigma}_{\mathbf{N}}$ from $\mathbb{N}(\mathbf{\Sigma}_{\mathbf{X}})$ and denote $\mathbf{\Sigma}_{\mathbf{X}|\mathbf{W}_1, \mathbf{W}_2}, \mathbf{\Sigma}_{\mathbf{X}_1|\mathbf{Y}, \mathbf{W}_1}, \mathbf{\Sigma}_{\mathbf{X}_2|\mathbf{Y}, \mathbf{W}_2}$ by $\mathbf{D}, \mathbf{\Gamma}_1, \mathbf{\Gamma}_2$, respectively. It can be verified (cf., [12, eq. (11)]) that

$$\text{diag}(\mathbf{\Gamma}_1, \mathbf{\Gamma}_2) = (\mathbf{D}^{-1} + \mathbf{\Sigma}_{\mathbf{N}}^{-1})^{-1}.$$

Furthermore, we have

$$\alpha_1 I(\mathbf{X}_1; \mathbf{W}_1 | \mathbf{W}_2) + \alpha_2 I(\mathbf{X}_2; \mathbf{W}_2)$$
$$= \frac{\alpha_1 - \alpha_2}{2} \log^+ \frac{|\mathbf{\Sigma}_{\mathbf{X}} + \mathbf{\Sigma}_{\mathbf{N}}|}{|\mathbf{D} + \mathbf{\Sigma}_{\mathbf{N}}||\mathbf{\Sigma}_{\hat{\mathbf{X}}_2}||\mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} - \mathbf{\Sigma}_{\hat{\mathbf{N}}_2}^{-1} \mathbf{\Gamma}_2 \mathbf{\Sigma}_{\hat{\mathbf{N}}_2}^{-1}|}$$
$$+ \frac{\alpha_1}{2} \log \frac{|\mathbf{\Sigma}_{\hat{\mathbf{N}}_1}|}{|\mathbf{\Gamma}_1|} + \frac{\alpha_2}{2} \log \frac{|\mathbf{\Sigma}_{\mathbf{X}} + \mathbf{\Sigma}_{\mathbf{N}}||\mathbf{\Sigma}_{\hat{\mathbf{N}}_2}|}{|\mathbf{D} + \mathbf{\Sigma}_{\mathbf{N}}||\mathbf{\Gamma}_2|}.$$

The desired result follows by invoking the fact that $(I(\mathbf{X}_1; \mathbf{W}_1 | \mathbf{W}_2), I(\mathbf{X}_2; \mathbf{W}_2)) \in \mathcal{R}(\mathbf{D}_1, \mathbf{D}_2)$ if $\mathbf{\Sigma}_{\mathbf{X}_i | \mathbf{W}_1, \mathbf{W}_2} \preceq \mathbf{D}_i$ (i.e., $(\mathbf{D})_i \preceq \mathbf{D}_i$), $i = 1, 2$, and the fact that one can set $\mathbf{\Gamma}_i$ to be an arbitrary matrix satisfying $0 \prec \mathbf{\Gamma}_i \preceq \mathbf{\Sigma}_{\hat{\mathbf{N}}_i}$ by suitably choosing $\mathbf{A}_i$, $i = 1, 2$. ∎

B. Matching Conditions

It is easy to see that the minimization problems associated with $\underline{R}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma}_{\mathbf{N}}, \alpha_1, \alpha_2)$ and $\overline{R}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma}_{\mathbf{N}}, \alpha_1, \alpha_2)$ are almost identical except that "$\preceq$" in (32) is replaced by "$=$" in (33). Indeed, we shall show that the two bounds coincide under certain conditions.

Let $\mathbb{D}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma}_{\mathbf{X}})$ be the set of positive-definite matrices $\mathbf{D}$ such that $(\mathbf{D})_i = \mathbf{D}_i$, $i = 1, 2$, and $\mathbf{D}^{-1} - \mathbf{\Sigma}_{\mathbf{X}}^{-1}$ is a positive semidefinite (m_1, m_2) block diagonal matrix.

Lemma 1: If for some $\mathbf{D}$, $\mathbf{\Sigma}_{\mathbf{N}}$, $\mathbf{\Gamma}_1$, and $\mathbf{\Gamma}_2$ with $\mathbf{D} \in \mathbb{D}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma}_{\mathbf{X}})$, $\mathbf{\Sigma}_{\mathbf{N}} \in \mathbb{N}(\mathbf{\Sigma}_{\mathbf{X}})$, and $\text{diag}(\mathbf{\Gamma}_1, \mathbf{\Gamma}_2) = (\mathbf{D}^{-1} + \mathbf{\Sigma}_{\mathbf{N}}^{-1})^{-1}$, there exists an (m_1, m_2) block diagonal matrix $\mathbf{\Pi}$ satisfying

$$\mathbf{\Pi} \succeq \alpha_1 (\mathbf{D} + \mathbf{\Sigma}_{\mathbf{N}})^{-1},$$
$$(\mathbf{D}(\mathbf{\Pi} - \alpha_1 (\mathbf{D} + \mathbf{\Sigma}_{\mathbf{N}})^{-1})\mathbf{D})_1 = \alpha_1 \mathbf{\Gamma}_1,$$
$$(\mathbf{D}(\mathbf{\Pi} - \alpha_1 (\mathbf{D} + \mathbf{\Sigma}_{\mathbf{N}})^{-1})\mathbf{D})_2 =$$
$$\alpha_2 \mathbf{\Gamma}_2 - (\alpha_1 - \alpha_2)\mathbf{\Gamma}_2(\mathbf{\Sigma}_{\hat{\mathbf{N}}_2} \mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} \mathbf{\Sigma}_{\hat{\mathbf{N}}_2} - \mathbf{\Gamma}_2)^{-1} \mathbf{\Gamma}_2,$$

then $\underline{R}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma}_{\mathbf{N}}, \alpha_1, \alpha_2) = \overline{R}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma}_{\mathbf{N}}, \alpha_1, \alpha_2)$. ∎

Proof: See Appendix D. ∎

Note that for $\mathbf{D}$, $\mathbf{\Sigma}_{\mathbf{N}}$, $\mathbf{\Gamma}_1$, and $\mathbf{\Gamma}_2$ with $\mathbf{D} \in \mathbb{D}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma}_{\mathbf{X}})$, $\mathbf{\Sigma}_{\mathbf{N}} \in \mathbb{N}(\mathbf{\Sigma}_{\mathbf{X}})$, and $\text{diag}(\mathbf{\Gamma}_1, \mathbf{\Gamma}_2) = (\mathbf{D}^{-1} + \mathbf{\Sigma}_{\mathbf{N}}^{-1})^{-1}$, we have

$$(\mathbf{D}(\mathbf{\Pi} - \alpha_1 (\mathbf{D} + \mathbf{\Sigma}_{\mathbf{N}})^{-1})\mathbf{D})_1 = \alpha_1 \mathbf{\Gamma}_1$$
$$\iff (\mathbf{D}\mathbf{\Pi}\mathbf{D})_1 = \alpha_1 (\mathbf{D} - (\mathbf{D}^{-1} + \mathbf{\Sigma}_{\mathbf{N}}^{-1})^{-1})_1 + \alpha_1 \mathbf{\Gamma}_1$$
$$\iff (\mathbf{D}\mathbf{\Pi}\mathbf{D})_1 = \mathbf{\Omega}_1 \triangleq \alpha_1 \mathbf{D}_1 \quad (34)$$

and

$$(\mathbf{D}(\mathbf{\Pi} - \alpha_1 (\mathbf{D} + \mathbf{\Sigma}_{\mathbf{N}})^{-1})\mathbf{D})_2$$
$$= \alpha_2 \mathbf{\Gamma}_2 - (\alpha_1 - \alpha_2)\mathbf{\Gamma}_2(\mathbf{\Sigma}_{\hat{\mathbf{N}}_2} \mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} \mathbf{\Sigma}_{\hat{\mathbf{N}}_2} - \mathbf{\Gamma}_2)^{-1} \mathbf{\Gamma}_2$$
$$\iff (\mathbf{D}\mathbf{\Pi}\mathbf{D})_2 = \alpha_1 (\mathbf{D} - (\mathbf{D}^{-1} + \mathbf{\Sigma}_{\mathbf{N}}^{-1})^{-1})_2$$
$$+ \alpha_2 \mathbf{\Gamma}_2 - (\alpha_1 - \alpha_2)\mathbf{\Gamma}_2(\mathbf{\Sigma}_{\hat{\mathbf{N}}_2} \mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} \mathbf{\Sigma}_{\hat{\mathbf{N}}_2} - \mathbf{\Gamma}_2)^{-1} \mathbf{\Gamma}_2$$
$$\iff (\mathbf{D}\mathbf{\Pi}\mathbf{D})_2 = \alpha_1 \mathbf{D}_2$$
$$- (\alpha_1 - \alpha_2)(\mathbf{\Gamma}_2 + \mathbf{\Gamma}_2(\mathbf{\Sigma}_{\hat{\mathbf{N}}_2} \mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} \mathbf{\Sigma}_{\hat{\mathbf{N}}_2} - \mathbf{\Gamma}_2)^{-1} \mathbf{\Gamma}_2)$$
$$\iff (\mathbf{D}\mathbf{\Pi}\mathbf{D})_2 = \mathbf{\Omega}_2 \triangleq \alpha_1 \mathbf{D}_2$$
$$- (\alpha_1 - \alpha_2)((\mathbf{D}^{-1} + \mathbf{\Sigma}_{\mathbf{N}}^{-1})_2 - \mathbf{\Sigma}_{\hat{\mathbf{N}}_2}^{-1} \mathbf{\Sigma}_{\tilde{\mathbf{N}}_2} \mathbf{\Sigma}_{\hat{\mathbf{N}}_2}^{-1})^{-1}. \quad (35)$$

In fact, there is a unique (m_1, m_2) block diagonal matrix $\mathbf{\Pi}$ satisfying (34) and (35). To see this, we partition $\mathbf{D}$ into the form

$$\begin{pmatrix} \mathbf{D}_{1,1} & \mathbf{D}_{1,2} \\ \mathbf{D}_{2,1} & \mathbf{D}_{2,2} \end{pmatrix}$$

where $\mathbf{D}_{i,i} = \mathbf{D}_i$, $i = 1, 2$. Now (34) and (35) can be written equivalently as

$$\sum_{j=1}^{2} \mathbf{D}_{i,j}(\mathbf{\Pi})_j \mathbf{D}_{i,j}^T = \mathbf{\Omega}_i, \quad i = 1, 2$$

which can be further written as

$$\begin{pmatrix} \mathbf{D}_{1,1} \otimes \mathbf{D}_{1,1} & \mathbf{D}_{1,2} \otimes \mathbf{D}_{1,2} \\ \mathbf{D}_{2,1} \otimes \mathbf{D}_{2,1} & \mathbf{D}_{2,2} \otimes \mathbf{D}_{2,2} \end{pmatrix} \begin{pmatrix} \mathrm{vec}((\mathbf{\Pi})_1) \\ \mathrm{vec}((\mathbf{\Pi})_2) \end{pmatrix}$$
$$= \begin{pmatrix} \mathrm{vec}(\mathbf{\Omega}_1) \\ \mathrm{vec}(\mathbf{\Omega}_2) \end{pmatrix}$$

where $\otimes$ is the Kronecker product and $\mathrm{vec}(\cdot)$ is the vec operator. Note that

$$\begin{pmatrix} \mathbf{D}_{1,1} \otimes \mathbf{D}_{1,1} & \mathbf{D}_{1,2} \otimes \mathbf{D}_{1,2} \\ \mathbf{D}_{2,1} \otimes \mathbf{D}_{2,1} & \mathbf{D}_{2,2} \otimes \mathbf{D}_{2,2} \end{pmatrix} = \mathbf{D} * \mathbf{D}$$

where $*$ is the Khatri–Rao product [18]. Since $\mathbf{D} \succ \mathbf{0}$, it follows that $\mathbf{D} * \mathbf{D}$ is positive definite and thus invertible. As a consequence, $\mathbf{\Pi}$ is uniquely determined by the following equation:

$$\begin{pmatrix} \mathrm{vec}((\mathbf{\Pi})_1) \\ \mathrm{vec}((\mathbf{\Pi})_2) \end{pmatrix} = (\mathbf{D} * \mathbf{D})^{-1} \begin{pmatrix} \mathrm{vec}(\mathbf{\Omega}_1) \\ \mathrm{vec}(\mathbf{\Omega}_2) \end{pmatrix}. \tag{36}$$

Therefore, we can reexpress Lemma 1 in the following form.

Lemma 2: If for some $\mathbf{D} \in \mathbb{D}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma_X})$ and $\mathbf{\Sigma_N} \in \mathbb{N}(\mathbf{\Sigma_X})$, the unique (m_1, m_2) block diagonal matrix $\mathbf{\Pi}$ determined by (34) and (35) [or, equivalently, by (36)] satisfies $\mathbf{\Pi} \succeq \alpha_1(\mathbf{D} + \mathbf{\Sigma_N})^{-1}$, then $\underline{R}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma_N}, \alpha_1, \alpha_2) = \overline{R}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma_N}, \alpha_1, \alpha_2)$.

Define $\mathcal{D} = \{(\mathbf{D}_1, \mathbf{D}_2) : \mathbf{0} \prec \mathbf{D}_i \preceq \mathbf{\Sigma}_{\mathbf{X}_i}, i = 1, 2\}$ and $\mathcal{D}_0 = \{(\mathbf{D}_1, \mathbf{D}_2) \in \mathcal{D} : \mathbf{\Sigma}_{\mathbf{X}_1,\mathbf{X}_2}^T(\mathbf{\Sigma}_{\mathbf{X}_1}^{-1}\mathbf{D}_1\mathbf{\Sigma}_{\mathbf{X}_1}^{-1} - \mathbf{\Sigma}_{\mathbf{X}_1}^{-1})\mathbf{\Sigma}_{\mathbf{X}_1,\mathbf{X}_2} + \mathbf{\Sigma}_{\mathbf{X}_2} \succeq \mathbf{D}_2, \mathbf{\Sigma}_{\mathbf{X}_1,\mathbf{X}_2}(\mathbf{\Sigma}_{\mathbf{X}_2}^{-1}\mathbf{D}_2\mathbf{\Sigma}_{\mathbf{X}_2}^{-1} - \mathbf{\Sigma}_{\mathbf{X}_2}^{-1})\mathbf{\Sigma}_{\mathbf{X}_1,\mathbf{X}_2}^T + \mathbf{\Sigma}_{\mathbf{X}_1} \succeq \mathbf{D}_1\}$.

Lemma 3: $\mathbb{D}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma_X})$ is a singleton for any $(\mathbf{D}_1, \mathbf{D}_2) \in \mathcal{D}_0$.

Proof: See Appendix E. ∎

The following two theorems (i.e., Theorems 7 and 8) indicate that $\underline{R}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma_N}, \alpha_1, \alpha_2)$ and $\overline{R}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma_N}, \alpha_1, \alpha_2)$ coincide in the high-resolution regime (i.e., when $\|\mathbf{D}_1\|$ and $\|\mathbf{D}_2\|$ are small) and the weak-dependence regime (i.e., when $\|\mathbf{\Sigma}_{\mathbf{X}_1,\mathbf{X}_2}\|$ is small), where $\|\cdot\|$ is the Frobenius norm. The proofs can be found in Appendixes F and G.

Theorem 7: There exists a $\rho(\alpha_1, \alpha_2) > 0$ such that if $\|\mathbf{D}_i\| \le \rho(\alpha_1, \alpha_2)$, $i = 1, 2$, then $\underline{R}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma_N}, \alpha_1, \alpha_2) = \overline{R}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma_N}, \alpha_1, \alpha_2)$ for some $\mathbf{\Sigma_N} \in \mathbb{N}(\mathbf{\Sigma_X})$.

Theorem 8: For any $(\mathbf{D}_1, \mathbf{D}_2)$ with $\mathbf{0} \prec \mathbf{D}_i \prec \mathbf{\Sigma}_{\mathbf{X}_i}$, $i = 1, 2$, there exists a $\varrho(\alpha_1, \alpha_2) > 0$ such that if $\|\mathbf{\Sigma}_{\mathbf{X}_1,\mathbf{X}_2}\| \le \varrho(\alpha_1, \alpha_2)$, then $\underline{R}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma_N}, \alpha_1, \alpha_2) = \overline{R}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma_N}, \alpha_1, \alpha_2)$ for some $\mathbf{\Sigma_N} \in \mathbb{N}(\mathbf{\Sigma_X})$.

We shall proceed to derive a more explicit matching condition for the case $(\mathbf{D}_1, \mathbf{D}_2) \in \mathcal{D}_0$. According to Lemma 3, $\mathbb{D}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma_X})$ is a singleton when $(\mathbf{D}_1, \mathbf{D}_2) \in \mathcal{D}_0$. Denote its element by $\mathbf{D}$, which is of the form

$$\begin{pmatrix} \mathbf{D}_1 & \mathbf{\Theta} \\ \mathbf{\Theta}^T & \mathbf{D}_2 \end{pmatrix}.$$

Now we choose

$$\mathbf{\Sigma_N} = \begin{pmatrix} \mathbf{D}_1 & -\mathbf{\Theta} \\ -\mathbf{\Theta}^T & \mathbf{D}_2 \end{pmatrix}.$$

Clearly, $\mathbf{D} \succ \mathbf{0}$ implies $\mathbf{\Sigma_N} \succ \mathbf{0}$. Note that

$$\mathbf{D} + \mathbf{\Sigma_N} = \begin{pmatrix} 2\mathbf{D}_1 & \mathbf{0} \\ \mathbf{0} & 2\mathbf{D}_2 \end{pmatrix}$$

$$\mathbf{D}^{-1} + \mathbf{\Sigma_N}^{-1}$$
$$= \begin{pmatrix} 2(\mathbf{D}_1 - \mathbf{\Theta}\mathbf{D}_2^{-1}\mathbf{\Theta}^T)^{-1} & \mathbf{0} \\ \mathbf{0} & 2(\mathbf{D}_2 - \mathbf{\Theta}^T\mathbf{D}_1^{-1}\mathbf{\Theta})^{-1} \end{pmatrix}.$$

Since both $\mathbf{\Sigma_X}^{-1} - \mathbf{D}^{-1}$ and $\mathbf{D}^{-1} + \mathbf{\Sigma_N}^{-1}$ are (m_1, m_2) block diagonal matrices, it follows that $\mathbf{\Sigma_N} \in \mathbb{N}(\mathbf{\Sigma_X})$. In view of the fact that $\mathbf{D} + \mathbf{\Sigma_N}$ is an (m_1, m_2) block diagonal matrix, the condition of Lemma 1 is satisfied if we can find a positive semidefinite (m_1, m_2) block diagonal matrix $\mathbf{\Psi}$ such that

$$\mathbf{D}_1(\mathbf{\Psi})_1\mathbf{D}_1 + \mathbf{\Theta}(\mathbf{\Psi})_2\mathbf{\Theta}^T = \alpha_1\mathbf{\Gamma}_1 \tag{37}$$
$$\mathbf{D}_2(\mathbf{\Psi})_2\mathbf{D}_2 + \mathbf{\Theta}^T(\mathbf{\Psi})_1\mathbf{\Theta} = \alpha_2\mathbf{\Gamma}_2$$
$$- (\alpha_1 - \alpha_2)\mathbf{\Gamma}_2(\mathbf{\Sigma}_{\hat{\mathbf{N}}_2}\mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1}\mathbf{\Sigma}_{\hat{\mathbf{N}}_2} - \mathbf{\Gamma}_2)^{-1}\mathbf{\Gamma}_2 \tag{38}$$

where $\mathbf{\Gamma}_1 = \frac{1}{2}(\mathbf{D}_1 - \mathbf{\Theta}\mathbf{D}_2^{-1}\mathbf{\Theta}^T)$ and $\mathbf{\Gamma}_2 = \frac{1}{2}(\mathbf{D}_2 - \mathbf{\Theta}^T\mathbf{D}_1^{-1}\mathbf{\Theta})$. Define $\mathbf{G}_i = \mathbf{D}_i^{\frac{1}{2}}(\mathbf{\Psi})_i\mathbf{D}_i^{\frac{1}{2}}$, $i = 1, 2$, $\mathbf{H} = \mathbf{D}_2^{-\frac{1}{2}}\mathbf{\Theta}^T\mathbf{D}_1^{-\frac{1}{2}}$, $\mathbf{Q}_1 = \mathbf{H}^T\mathbf{H}$, and $\mathbf{Q}_2 = \mathbf{H}\mathbf{H}^T$. Clearly, $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are positive semidefinite. Moreover, since $\mathbf{D}$ is positive definite, it follows that $\mathbf{D}_1 \succ \mathbf{\Theta}\mathbf{D}_2^{-1}\mathbf{\Theta}^T$, $\mathbf{D}_2 \succ \mathbf{\Theta}^T\mathbf{D}_1^{-1}\mathbf{\Theta}$, and consequently, $\mathbf{Q}_i \prec \mathbf{I}_{m_i}$, $i = 1, 2$. Define

$$\mathbf{C} = \left[\mathbf{D}_2^{-\frac{1}{2}}\mathbf{\Sigma}_{\hat{\mathbf{N}}_2}\mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1}\mathbf{\Sigma}_{\hat{\mathbf{N}}_2}\mathbf{D}_2^{-\frac{1}{2}} - \frac{1}{2}\mathbf{I}_{m_2} + \frac{1}{2}\mathbf{Q}_2 \right]^{-1}.$$

Through some algebraic manipulations, we can convert (37) and (38) to the following matrix equations:

$$\mathbf{G}_i - \mathbf{Q}_i\mathbf{G}_i\mathbf{Q}_i = \mathbf{B}_i, \quad i = 1, 2 \tag{39}$$

where

$$\mathbf{B}_1 = \frac{1}{2}(\alpha_1\mathbf{I}_{m_1} - \alpha_2\mathbf{Q}_1)(\mathbf{I}_{m_1} - \mathbf{Q}_1)$$
$$- \frac{\alpha_1 - \alpha_2}{4}\mathbf{H}^T(\mathbf{I}_{m_2} - \mathbf{Q}_2)\mathbf{C}(\mathbf{I}_{m_2} - \mathbf{Q}_2)\mathbf{H}$$
$$\mathbf{B}_2 = \frac{1}{2}(\alpha_2\mathbf{I}_{m_2} - \alpha_1\mathbf{Q}_2)(\mathbf{I}_{m_2} - \mathbf{Q}_2)$$
$$- \frac{\alpha_1 - \alpha_2}{4}(\mathbf{I}_{m_2} - \mathbf{Q}_2)\mathbf{C}(\mathbf{I}_{m_2} - \mathbf{Q}_2).$$

Let $\mathbf{U}_i\mathbf{\Lambda}_i\mathbf{U}_i^T$ be the eigenvalue decomposition of $\mathbf{Q}_i$, where $\mathbf{U}_i$ is a unitary matrix and $\mathbf{\Lambda}_i \triangleq \mathrm{diag}(\lambda_{i,1}, \ldots, \lambda_{i,m_i})$ is a diagonal matrix (with $\lambda_{i,j} \in [0, 1)$, $j = 1, \ldots, m_i$), $i = 1, 2$. Define $\tilde{\mathbf{G}}_i = \mathbf{U}_i^T\mathbf{G}_i\mathbf{U}_i$ and $\tilde{\mathbf{B}}_i = \mathbf{U}_i^T\mathbf{B}_i\mathbf{U}_i$, $i = 1, 2$. We can write (39) equivalently as

$$\tilde{\mathbf{G}}_i - \mathbf{\Lambda}_i\tilde{\mathbf{G}}_i\mathbf{\Lambda}_i = \tilde{\mathbf{B}}_i, \quad i = 1, 2$$

which can be solved to obtain

$$\tilde{\mathbf{G}}_i = \mathbf{\Xi}_i \odot \tilde{\mathbf{B}}_i, \quad i = 1, 2 \tag{40}$$

where $\mathbf{\Xi}_1 \triangleq (\xi_{j,k})$ with $\xi_{j,k} = \frac{1}{1 - \lambda_{1,j}\lambda_{1,k}}$, $j, k = 1, \ldots, m_1$, and $\mathbf{\Xi}_2 \triangleq (\xi'_{j,k})$ with $\xi'_{j,k} = \frac{1}{1 - \lambda_{2,j}\lambda_{2,k}}$, $j, k = 1, \ldots, m_2$, and $\odot$ is the Hadamard product. Note that if $\tilde{\mathbf{G}}_1$ and $\tilde{\mathbf{G}}_2$ given by (40) are positive semidefinite, then the resulting $\mathbf{\Psi}$ is also positive semidefinite, and thus, the condition of Lemma 1 is satisfied.

For the special case $\alpha_1 = \alpha_2 = 1$ (which corresponds to the sum rate), one can readily verify that (40) yields

$$\tilde{\mathbf{G}}_i = \mathrm{diag}\left(\frac{1 - \lambda_{i,1}}{2(1 + \lambda_{i,1})}, \ldots, \frac{1 - \lambda_{i,m_i}}{2(1 + \lambda_{m_i})}\right), \quad i = 1, 2$$

which are positive definite. As a consequence, we have $\underline{R}(\mathbf{D}_1, \mathbf{D}_2, \boldsymbol{\Sigma}_\mathbf{N}, 1, 1) = \overline{R}(\mathbf{D}_1, \mathbf{D}_2, \boldsymbol{\Sigma}_\mathbf{N}, 1, 1)$. Furthermore, it can be shown that both $\underline{R}(\mathbf{D}_1, \mathbf{D}_2, \boldsymbol{\Sigma}_\mathbf{N}, 1, 1)$ and $\overline{R}(\mathbf{D}_1, \mathbf{D}_2, \boldsymbol{\Sigma}_\mathbf{N}, 1, 1)$ are equal to $\frac{1}{2} \log \frac{|\boldsymbol{\Sigma}_\mathbf{X}|}{|\mathbf{D}|}$. Therefore, we obtain the following theorem.

Theorem 9: If $(\mathbf{D}_1, \mathbf{D}_2) \in \mathcal{D}_0$, then

$$R(\mathbf{D}_1, \mathbf{D}_2, 1, 1) = \frac{1}{2} \log \frac{|\boldsymbol{\Sigma}_\mathbf{X}|}{|\mathbf{D}|}$$

where $\mathbf{D}$ is the unique element of $\mathbb{D}(\mathbf{D}_1, \mathbf{D}_2, \boldsymbol{\Sigma}_\mathbf{X})$.

Define $\mathcal{D}_1 = \{(\mathbf{D}_1, \mathbf{D}_2) \in \mathcal{D} : \boldsymbol{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}^T (\boldsymbol{\Sigma}_{\mathbf{X}_1}^{-1} \mathbf{D}_1 \boldsymbol{\Sigma}_{\mathbf{X}_1}^{-1} - \boldsymbol{\Sigma}_{\mathbf{X}_1}^{-1}) \boldsymbol{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2} + \boldsymbol{\Sigma}_{\mathbf{X}_2} \preceq \mathbf{D}_2\}$, $\mathcal{D}_2 = \{(\mathbf{D}_1, \mathbf{D}_2) \in \mathcal{D} : \boldsymbol{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}(\boldsymbol{\Sigma}_{\mathbf{X}_2}^{-1} \mathbf{D}_2 \boldsymbol{\Sigma}_{\mathbf{X}_2}^{-1} - \boldsymbol{\Sigma}_{\mathbf{X}_2}^{-1}) \boldsymbol{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}^T + \boldsymbol{\Sigma}_{\mathbf{X}_1} \preceq \mathbf{D}_1\}$, and $\mathcal{D}_3 = \mathcal{D} - \mathcal{D}_0 \cup \mathcal{D}_1 \cup \mathcal{D}_2$. It is easy to show that if $(\mathbf{D}_1, \mathbf{D}_2) \in \mathcal{D}_i$, then

$$R(\mathbf{D}_1, \mathbf{D}_2, 1, 1) = \frac{1}{2} \log \frac{|\boldsymbol{\Sigma}_{\mathbf{X}_i}|}{|\mathbf{D}_i|}, \quad i = 1, 2.$$

The characterization of $R(\mathbf{D}_1, \mathbf{D}_2, 1, 1)$ for $(\mathbf{D}_1, \mathbf{D}_2) \in \mathcal{D}_3$ remains an open problem. It is worth noting that $\mathcal{D}_3 = \emptyset$ in the scalar case (i.e., $m_1 = m_2 = 1$).

V. Conclusion

We have derived a lower bound on each supporting line of the rate region of the vector Gaussian two-terminal CEO problem by leveraging the estimation-theoretic method developed in [12] together with a variant of the Liu–Viswanath extremal inequality. It is shown that the lower bound coincides with the standard Berger–Tung upper bound under certain conditions. Note that the extremal inequality in Theorem 1 is only suitable for the two-terminal case, while the estimation-theoretic method does not have such a limitation. As far as the sum rate is concerned, the lower bound in this study can be directly extended to the general multiterminal case since the extremal inequality is needed only when the rates are weighted differently. Furthermore, due to the resemblance between our lower bound and the Berger–Tung upper bound, one can easily conjecture the form of the desired multiterminal generalization of our lower bound. However, a more sophisticated extremal inequality is needed to establish such a lower bound for a general weighted sum rate in the multiterminal setting [19].

We have also studied the direct vector Gaussian two-terminal source coding problem and obtained similar results. Again, generalizing these results to the multiterminal setting is not completely straightforward. In particular, certain brute-force calculations in this study (see, e.g., the proof of Lemma 3) need to be handled in a more conceptual manner. It is worth mentioning that recently the vector Gaussian one-helper source coding problem, which is a special case of the direct vector Gaussian two-terminal source coding problem, has been completely solved [20], [21].

It should be noted that our bounding technique is applicable to the vector version of the indirect Gaussian two-terminal source coding problem (which contains the vector Gaussian two-terminal CEO problem as a special case) studied in [7]. In fact, albeit somewhat implicitly, our treatment of the direct vector Gaussian two-terminal source coding problem is to a large extent based on a careful analysis of an indirect source coding problem in which $\mathbf{Y}$ is the remote source.

Appendix A
Proof of Theorem 1

If $\mu = 1$, then (1) and (2) imply that $\boldsymbol{\Sigma}^* = \tilde{\boldsymbol{\Gamma}}$. Since

$$h(\mathbf{S}, \mathbf{S} + \mathbf{Z}|V) = h(\mathbf{S}|V) + h(\mathbf{Z}),$$
$$h(\mathbf{S}, \mathbf{S} + \mathbf{Z}|V) = h(\mathbf{S} + \mathbf{Z}|V) + h(\mathbf{S}|\mathbf{S} + \mathbf{Z}, V),$$

it follows that

$$h(\mathbf{S}|V) - h(\mathbf{S} + \mathbf{Z}|V)$$
$$= h(\mathbf{S}|\mathbf{S} + \mathbf{Z}, V) - h(\mathbf{Z})$$
$$\leq \frac{n}{2} \log((2\pi e)^m |\boldsymbol{\Gamma}|) - \frac{n}{2} \log((2\pi e)^m |\boldsymbol{\Sigma}_\mathbf{Z}|)$$
$$= \frac{n}{2} \log((2\pi e)^m |\tilde{\boldsymbol{\Gamma}} - \tilde{\boldsymbol{\Gamma}}(\tilde{\boldsymbol{\Gamma}} + \boldsymbol{\Sigma}_\mathbf{Z})^{-1} \tilde{\boldsymbol{\Gamma}}|)$$
$$\quad - \frac{n}{2} \log((2\pi e)^m |\boldsymbol{\Sigma}_\mathbf{Z}|)$$
$$= \frac{n}{2} \log((2\pi e)^m |\tilde{\boldsymbol{\Gamma}}| |(\tilde{\boldsymbol{\Gamma}} + \boldsymbol{\Sigma}_\mathbf{Z})^{-1}| |(\tilde{\boldsymbol{\Gamma}} + \boldsymbol{\Sigma}_\mathbf{Z}) - \tilde{\boldsymbol{\Gamma}}|)$$
$$\quad - \frac{n}{2} \log((2\pi e)^m |\boldsymbol{\Sigma}_\mathbf{Z}|)$$
$$= \frac{n}{2} \log((2\pi e)^m |\tilde{\boldsymbol{\Gamma}}|) - \frac{n}{2} \log((2\pi e)^m |\tilde{\boldsymbol{\Gamma}} + \boldsymbol{\Sigma}_\mathbf{Z}|)$$
$$= \frac{n}{2} \log((2\pi e)^m |\boldsymbol{\Sigma}^*|) - \frac{n}{2} \log((2\pi e)^m |\boldsymbol{\Sigma}^* + \boldsymbol{\Sigma}_\mathbf{Z}|).$$

So it suffices to consider the case $\mu > 1$.

If $\mathbf{M} = \mathbf{0}$, then (1) implies

$$\boldsymbol{\Sigma}^* = \frac{1}{\mu - 1} \boldsymbol{\Sigma}_\mathbf{Z}.$$

Consequently

$$\frac{1}{2} \log((2\pi e)^m |\boldsymbol{\Sigma}^*|) - \frac{\mu}{2} \log((2\pi e)^m |\boldsymbol{\Sigma}^* + \boldsymbol{\Sigma}_\mathbf{Z}|)$$
$$= -\frac{\mu - 1}{2} \log((2\pi e)^m |\boldsymbol{\Sigma}_\mathbf{Z}|)$$
$$\quad + \frac{(\mu - 1)m}{2} \log(\mu - 1) - \frac{\mu m}{2} \log \mu. \tag{41}$$

By the conditional version of the entropy power inequality, we have

$$h(\mathbf{S}|V) - \mu h(\mathbf{S} + \mathbf{Z}|V) \leq h(\mathbf{S}|V)$$
$$- \frac{\mu m n}{2} \log \left(\exp\left(\frac{2}{mn} h(\mathbf{S}|V)\right) + \exp\left(\frac{2}{mn} h(\mathbf{Z})\right) \right)$$

where the right-hand side of the inequality, as a function of $h(\mathbf{S}|V)$, is maximized at

$$h(\mathbf{S}|V) = h(\mathbf{Z}) - \frac{mn}{2} \log(\mu - 1)$$

and the corresponding maximum value is equal to

$$-(\mu - 1)h(\mathbf{Z}) + \frac{(\mu - 1)mn}{2} \log(\mu - 1) - \frac{\mu mn}{2} \log \mu. \quad (42)$$

A comparison of (41) and (42) completes the proof for the case $\mathbf{M} = \mathbf{0}$.

Now consider the case $\mathbf{M} \neq \mathbf{0}$. Let $\boldsymbol{\Sigma}_{\tilde{\mathbf{Z}}} = (\mu - 1)\boldsymbol{\Sigma}^*$. We have

$$\mu(\boldsymbol{\Sigma}^* + \boldsymbol{\Sigma}_{\tilde{\mathbf{Z}}})^{-1} = (\boldsymbol{\Sigma}^*)^{-1} = \mu(\boldsymbol{\Sigma}^* + \boldsymbol{\Sigma}_{\mathbf{Z}})^{-1} + \mathbf{M}.$$

It follows from [11, Lemma 12] that $\mathbf{0} \prec \boldsymbol{\Sigma}_{\tilde{\mathbf{Z}}} \preceq \boldsymbol{\Sigma}_{\mathbf{Z}}$ and

$$\frac{|\boldsymbol{\Sigma}^* + \boldsymbol{\Sigma}_{\tilde{\mathbf{Z}}}|}{|\boldsymbol{\Sigma}^* + \boldsymbol{\Sigma}_{\mathbf{Z}}|} = \frac{|\tilde{\boldsymbol{\Gamma}} + \boldsymbol{\Sigma}_{\tilde{\mathbf{Z}}}|}{|\tilde{\boldsymbol{\Gamma}} + \boldsymbol{\Sigma}_{\mathbf{Z}}|}. \quad (43)$$

Let $\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T$ be the eigenvalue decomposition of $\boldsymbol{\Sigma}_{\mathbf{Z}} - \boldsymbol{\Sigma}_{\tilde{\mathbf{Z}}}$, where $\mathbf{U}$ is a unitary matrix and $\boldsymbol{\Lambda} \triangleq \mathrm{diag}(\lambda_1, \ldots, \lambda_r, 0, \ldots, 0)$ is a positive semidefinite diagonal matrix of rank r. Note that $\mathbf{M} \neq \mathbf{0}$ implies $r \geq 1$. Let $\tilde{\mathbf{Z}}$ be an $m \times n$ random matrix such that each column is an independent copy of an $m \times 1$ Gaussian random vector with mean zero and covariance matrix $\boldsymbol{\Sigma}_{\tilde{\mathbf{Z}}}$. Similarly, let $\hat{\mathbf{Z}}$ be an $r \times n$ random matrix such that each column is an independent copy of an $r \times 1$ Gaussian random vector with mean zero and covariance matrix $\boldsymbol{\Sigma}_{\hat{\mathbf{Z}}}$, where $\boldsymbol{\Sigma}_{\hat{\mathbf{Z}}} = \mathrm{diag}(\lambda_1, \ldots, \lambda_r)$. We assume $\tilde{\mathbf{Z}}$, $\hat{\mathbf{Z}}$, and $(\mathbf{S}, V)$ are mutually independent. Define an $m \times r$ matrix

$$\mathbf{G} = \mathbf{U} \begin{pmatrix} \mathbf{I}_r \\ \mathbf{0} \end{pmatrix}.$$

It is easy to see that the joint distribution of $(\mathbf{Z}, \mathbf{S}, V)$ is preserved if we let

$$\mathbf{Z} = \tilde{\mathbf{Z}} + \mathbf{G}\hat{\mathbf{Z}}.$$

Note that

$$\begin{pmatrix} \mathbf{S} + \mathbf{Z} \\ \hat{\mathbf{Z}} \end{pmatrix} = \tilde{\mathbf{G}} \begin{pmatrix} \mathbf{S} + \tilde{\mathbf{Z}} \\ \hat{\mathbf{Z}} \end{pmatrix}$$

where

$$\tilde{\mathbf{G}} = \begin{pmatrix} \mathbf{I}_m & \mathbf{G} \\ \mathbf{0} & \mathbf{I}_r \end{pmatrix}.$$

Since $|\tilde{\mathbf{G}}| = 1$, it follows that

$$h(\mathbf{S} + \mathbf{Z}, \hat{\mathbf{Z}}|V) = h(\mathbf{S} + \tilde{\mathbf{Z}}, \hat{\mathbf{Z}}|V).$$

In view of the fact that

$$h(\mathbf{S} + \mathbf{Z}, \hat{\mathbf{Z}}|V) = h(\mathbf{S} + \mathbf{Z}|V) + h(\hat{\mathbf{Z}}|\mathbf{S} + \mathbf{Z}, V)$$
$$h(\mathbf{S} + \tilde{\mathbf{Z}}, \hat{\mathbf{Z}}|V) = h(\mathbf{S} + \tilde{\mathbf{Z}}|V) + h(\hat{\mathbf{Z}})$$

we have

$$h(\mathbf{S} + \mathbf{Z}|V) = h(\mathbf{S} + \tilde{\mathbf{Z}}|V) + h(\hat{\mathbf{Z}}) - h(\hat{\mathbf{Z}}|\mathbf{S} + \mathbf{Z}, V). \quad (44)$$

Let $\boldsymbol{\Delta} = \hat{\mathbf{Z}} - \mathbf{A}\mathbf{Z}$, where $\mathbf{A}\mathbf{Z}$ is the (reduced) optimal linear estimator of $\hat{\mathbf{Z}}$ from $\mathbf{Z}$. One can readily show that $\boldsymbol{\Delta}$ is independent of $(\mathbf{Z}, \mathbf{S}, V)$. Moreover, it is clear that $\boldsymbol{\Sigma}_{\boldsymbol{\Delta}} \succ \mathbf{0}$. Note that

$$\hat{\mathbf{Z}} = \mathbf{A}\mathbf{Z} + \boldsymbol{\Delta} = -\mathbf{A}\mathbf{S} + \mathbf{A}(\mathbf{S} + \mathbf{Z}) + \boldsymbol{\Delta}.$$

Since $\boldsymbol{\Sigma}_{\mathbf{S}|\mathbf{S}+\mathbf{Z},V} \preceq \boldsymbol{\Gamma}$, it follows that $\boldsymbol{\Sigma}_{\hat{\mathbf{Z}}|\mathbf{S}+\mathbf{Z},V} \preceq \mathbf{A}\boldsymbol{\Gamma}\mathbf{A}^T + \boldsymbol{\Sigma}_{\boldsymbol{\Delta}}$, and consequently

$$h(\hat{\mathbf{Z}}|\mathbf{S} + \mathbf{Z}, V) \leq \frac{n}{2} \log((2\pi e)^r |\mathbf{A}\boldsymbol{\Gamma}\mathbf{A}^T + \boldsymbol{\Sigma}_{\boldsymbol{\Delta}}|) \quad (45)$$

which, together with (44), implies

$$h(\mathbf{S}|V) - \mu h(\mathbf{S} + \mathbf{Z}|V) \leq h(\mathbf{S}|V) - \mu h(\mathbf{S} + \tilde{\mathbf{Z}}|V)$$
$$- \mu h(\hat{\mathbf{Z}}) + \frac{\mu n}{2} \log((2\pi e)^r |\mathbf{A}\boldsymbol{\Gamma}\mathbf{A}^T + \boldsymbol{\Sigma}_{\boldsymbol{\Delta}}|). \quad (46)$$

It can be shown by invoking the conditional version of the entropy power inequality (cf., the case $\mathbf{M} = \mathbf{0}$) that

$$h(\mathbf{S}|V) - \mu h(\mathbf{S} + \tilde{\mathbf{Z}}|V) \leq \frac{n}{2} \log((2\pi e)^m |\boldsymbol{\Sigma}^*|)$$
$$- \frac{\mu n}{2} \log((2\pi e)^m |\boldsymbol{\Sigma}^* + \boldsymbol{\Sigma}_{\tilde{\mathbf{Z}}}|). \quad (47)$$

Substituting (47) into (46), we obtain

$$h(\mathbf{S}|V) - \mu h(\mathbf{S} + \mathbf{Z}|V)$$
$$\leq \frac{n}{2} \log((2\pi e)^m |\boldsymbol{\Sigma}^*|) - \frac{\mu n}{2} \log((2\pi e)^m |\boldsymbol{\Sigma}^* + \boldsymbol{\Sigma}_{\tilde{\mathbf{Z}}}|$$
$$- \mu h(\hat{\mathbf{Z}}) + \frac{\mu n}{2} \log((2\pi e)^r |\mathbf{A}\boldsymbol{\Gamma}\mathbf{A}^T + \boldsymbol{\Sigma}_{\boldsymbol{\Delta}}|). \quad (48)$$

Let $\tilde{\mathbf{S}}$ be an $m \times n$ random matrix, independent of $(\mathbf{Z}, \tilde{\mathbf{Z}}, \hat{\mathbf{Z}})$, such that each column is an independent copy of an $m \times 1$ Gaussian random vector with mean zero and covariance matrix $\tilde{\boldsymbol{\Gamma}}$. It can be verified [cf., (44) and (45)] that

$$h(\tilde{\mathbf{S}} + \tilde{\mathbf{Z}}) + h(\hat{\mathbf{Z}}) - h(\tilde{\mathbf{S}} + \mathbf{Z})$$
$$= h(\hat{\mathbf{Z}}|\tilde{\mathbf{S}} + \mathbf{Z})$$
$$= \frac{n}{2} \log((2\pi e)^r |\mathbf{A}\boldsymbol{\Gamma}\mathbf{A}^T + \boldsymbol{\Sigma}_{\boldsymbol{\Delta}}|). \quad (49)$$

In view of (43) and (49), we have

$$\frac{n}{2} \log((2\pi e)^m |\boldsymbol{\Sigma}^* + \boldsymbol{\Sigma}_{\tilde{\mathbf{Z}}}|) + h(\hat{\mathbf{Z}})$$
$$- \frac{n}{2} \log((2\pi e)^m |\boldsymbol{\Sigma}^* + \boldsymbol{\Sigma}_{\mathbf{Z}}|)$$
$$= \frac{n}{2} \log((2\pi e)^m |\tilde{\boldsymbol{\Gamma}} + \boldsymbol{\Sigma}_{\tilde{\mathbf{Z}}}|) + h(\hat{\mathbf{Z}})$$
$$- \frac{n}{2} \log((2\pi e)^m |\tilde{\boldsymbol{\Gamma}} + \boldsymbol{\Sigma}_{\mathbf{Z}}|)$$
$$= h(\tilde{\mathbf{S}} + \tilde{\mathbf{Z}}) + h(\hat{\mathbf{Z}}) - h(\tilde{\mathbf{S}} + \mathbf{Z})$$
$$= \frac{n}{2} \log((2\pi e)^r |\mathbf{A}\boldsymbol{\Gamma}\mathbf{A}^T + \boldsymbol{\Sigma}_{\boldsymbol{\Delta}}|)$$

which implies

$$\frac{n}{2} \log((2\pi e)^m |\boldsymbol{\Sigma}^*|) - \frac{\mu n}{2} \log((2\pi e)^m |\boldsymbol{\Sigma}^* + \boldsymbol{\Sigma}_{\tilde{\mathbf{Z}}}|)$$
$$- \mu h(\hat{\mathbf{Z}}) + \frac{\mu n}{2} \log((2\pi e)^r |\mathbf{A}\boldsymbol{\Gamma}\mathbf{A}^T + \boldsymbol{\Sigma}_{\boldsymbol{\Delta}}|)$$
$$= \frac{n}{2} \log((2\pi e)^m |\boldsymbol{\Sigma}^*|) - \frac{\mu n}{2} \log((2\pi e)^m |\boldsymbol{\Sigma}^* + \boldsymbol{\Sigma}_{\mathbf{Z}}|). \quad (50)$$

Combining (48) and (50) completes the proof.

APPENDIX B
CHARACTERIZATION OF $R(\mathbf{D}, \alpha_1, 0)$

Let $f_i^{(n)} : \mathbb{R}^{m \times n} \to \{1, 2, \dots, M_i\}$, $i = 1, 2$, be two encoding functions such that $\boldsymbol{\Sigma}_{\mathbf{X}_0^n | W_1, W_2} \preceq \mathbf{D}$, where $W_i = f_i^{(n)}(\mathbf{X}_i^n)$, $i = 1, 2$. Note that

$$\mathbf{0} \prec \boldsymbol{\Sigma}_{\mathbf{X}_1^n | W_1, \mathbf{X}_2^n} \preceq \boldsymbol{\Sigma}_{\mathbf{X}_1^n | \mathbf{X}_2^n} = \boldsymbol{\Sigma}_{\mathbf{N}_1} + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2} \qquad (51)$$

and

$$\begin{aligned} \mathbf{D} &\succeq \boldsymbol{\Sigma}_{\mathbf{X}_0^n | W_1, W_2} \\ &= \boldsymbol{\Sigma}_{\mathbf{Y}^n | W_1, W_2} + \boldsymbol{\Sigma}_{\mathbf{N}} \\ &\succeq \boldsymbol{\Sigma}_{\mathbf{Y}^n | W_1, \mathbf{X}_2^n} + \boldsymbol{\Sigma}_{\mathbf{N}} \\ &= \boldsymbol{\Sigma}_{\mathbf{N}} \boldsymbol{\Sigma}_{\mathbf{N}_1}^{-1} \boldsymbol{\Sigma}_{\mathbf{X}_1^n | W_1, \mathbf{X}_2^n} \boldsymbol{\Sigma}_{\mathbf{N}_1}^{-1} \boldsymbol{\Sigma}_{\mathbf{N}} + \boldsymbol{\Sigma}_{\mathbf{N}}. \end{aligned} \qquad (52)$$

Moreover, we have

$$\begin{aligned} \frac{1}{n} \log M_1 &\geq \frac{1}{n} I(\mathbf{X}_1^n; W_1 | \mathbf{X}_2^n) \\ &= \frac{1}{n} h(\mathbf{X}_1^n | \mathbf{X}_2^n) - h(\mathbf{X}_1^n | W_1, \mathbf{X}_2^n) \\ &\geq \frac{1}{2} \log \frac{|\boldsymbol{\Sigma}_{\mathbf{N}_1} + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}|}{|\boldsymbol{\Sigma}_{\mathbf{X}_1^n | W_1, \mathbf{X}_2^n}|}. \end{aligned} \qquad (53)$$

Combining (51)–(53) and denoting $\boldsymbol{\Sigma}_{\mathbf{X}_1^n | W_1, \mathbf{X}_2^n}$ by $\tilde{\mathbf{D}}$, one can readily see that

$$\begin{aligned} R(\mathbf{D}, \alpha_1, 0) &\geq \min_{\tilde{\mathbf{D}}} \frac{\alpha_1}{2} \log \frac{|\boldsymbol{\Sigma}_{\mathbf{N}_1} + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}|}{|\tilde{\mathbf{D}}|} \\ \text{subject to} \quad & \boldsymbol{\Sigma}_{\mathbf{N}} \boldsymbol{\Sigma}_{\mathbf{N}_1}^{-1} \tilde{\mathbf{D}} \boldsymbol{\Sigma}_{\mathbf{N}_1}^{-1} \boldsymbol{\Sigma}_{\mathbf{N}} + \boldsymbol{\Sigma}_{\mathbf{N}} \preceq \mathbf{D}, \\ & \mathbf{0} \prec \tilde{\mathbf{D}} \preceq \boldsymbol{\Sigma}_{\mathbf{N}_1} + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}. \end{aligned}$$

Now we proceed to show that this lower bound is in fact tight. Let $\mathbf{W}_1$ and $\mathbf{W}_2$ be generated by $\mathbf{X}_1$ and $\mathbf{X}_2$, respectively, via paralleled Gaussian test channels, i.e., $\mathbf{W}_i = \mathbf{A}_i \mathbf{X}_i + \mathbf{U}_i$, $i = 1, 2$, where $\mathbf{A}_i$ is an $m \times m$ matrix and $\mathbf{U}_i$ is an $m \times 1$ Gaussian random vector with independent zero-mean unit-variance entries; moreover, we assume $\mathbf{U}_1$, $\mathbf{U}_2$, and $(\mathbf{X}_0, \mathbf{X}_1, \mathbf{X}_2)$ are mutually independent. Denote $\boldsymbol{\Sigma}_{\mathbf{X}_1 | \mathbf{W}_1, \mathbf{W}_2}$ by $\tilde{\mathbf{D}}'$. It is clear that we can set $\tilde{\mathbf{D}}'$ to be any matrix satisfying

$$\mathbf{0} \prec \tilde{\mathbf{D}}' \preceq \boldsymbol{\Sigma}_{\mathbf{X}_1 | \mathbf{W}_2} \qquad (54)$$

by suitably choosing $\mathbf{A}_1$. Let $\mathbf{B}_1 \mathbf{X}_1 + \mathbf{B}_2 \mathbf{W}_2$ be the linear MMSE estimator of $\mathbf{X}_0$ from $\mathbf{X}_1$ and $\mathbf{W}_2$. Note that

$$\boldsymbol{\Sigma}_{\mathbf{X}_0 | \mathbf{W}_1, \mathbf{W}_2} = \mathbf{B}_1 \tilde{\mathbf{D}}' \mathbf{B}_1^T + \boldsymbol{\Sigma}_{\mathbf{X}_0 | \mathbf{X}_1, \mathbf{W}_2}. \qquad (55)$$

According to the well-known Berger–Tung inner bound [15], [16], we have $(I(\mathbf{X}_1; \mathbf{W}_1 | \mathbf{W}_2), I(\mathbf{X}_2; \mathbf{W}_2)) \in \mathcal{R}(\mathbf{D})$ if $\boldsymbol{\Sigma}_{\mathbf{X}_0 | \mathbf{W}_1, \mathbf{W}_2} \preceq \mathbf{D}$, where

$$I(\mathbf{X}_1; \mathbf{W}_1 | \mathbf{W}_2) = \frac{1}{2} \log \frac{|\boldsymbol{\Sigma}_{\mathbf{X}_1 | \mathbf{W}_2}|}{|\tilde{\mathbf{D}}'|}. \qquad (56)$$

Combining (54)–(56) yields the following upper bound on $R(\mathbf{D}, \alpha_1, 0)$:

$$\begin{aligned} R(\mathbf{D}, \alpha_1, 0) &\leq \min_{\tilde{\mathbf{D}}'} \frac{\alpha_1}{2} \log \frac{|\boldsymbol{\Sigma}_{\mathbf{X}_1 | \mathbf{W}_2}|}{|\tilde{\mathbf{D}}'|} \\ \text{subject to} \quad & \mathbf{B}_1 \tilde{\mathbf{D}}' \mathbf{B}_1^T + \boldsymbol{\Sigma}_{\mathbf{X}_0 | \mathbf{X}_1, \mathbf{W}_2} \preceq \mathbf{D}, \\ & \mathbf{0} \prec \tilde{\mathbf{D}}' \preceq \boldsymbol{\Sigma}_{\mathbf{X}_1 | \mathbf{W}_2}. \end{aligned}$$

Let $\mathbf{A}_2 = \kappa \mathbf{I}_m$. In view of the fact that $\boldsymbol{\Sigma}_{\mathbf{X}_1 | \mathbf{W}_2} \to \boldsymbol{\Sigma}_{\mathbf{X}_1 | \mathbf{X}_2} = \boldsymbol{\Sigma}_{\mathbf{N}_1} + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}$, $\mathbf{B}_1 \to \boldsymbol{\Sigma}_{\mathbf{N}} \boldsymbol{\Sigma}_{\mathbf{N}_1}^{-1}$, and $\boldsymbol{\Sigma}_{\mathbf{X}_0 | \mathbf{X}_1, \mathbf{W}_2} \to \boldsymbol{\Sigma}_{\mathbf{N}}$ as $\kappa \to \infty$, the proof can be completed by a continuity argument.

APPENDIX C
PROOF OF THEOREM 5

Let $f_i^{(n)} : \mathbb{R}^{m \times n} \to \{1, 2, \dots, M_i\}$, $i = 1, 2$, be two encoding functions such that $\boldsymbol{\Sigma}_{\mathbf{X}_i^n | W_1, W_2} \preceq \mathbf{D}_i$, $i = 1, 2$, where $W_i = f_i^{(n)}(\mathbf{X}_i^n)$, $i = 1, 2$. We pick an arbitrary $\boldsymbol{\Sigma}_{\mathbf{N}}$ from $\mathbb{N}(\boldsymbol{\Sigma}_{\mathbf{X}})$ and denote $\boldsymbol{\Sigma}_{\mathbf{X}^n | W_1, W_2}, \boldsymbol{\Sigma}_{\mathbf{X}_1^n | \mathbf{Y}^n, W_1}, \boldsymbol{\Sigma}_{\mathbf{X}_2^n | \mathbf{Y}^n, W_2}$ by $\mathbf{D}$, $\boldsymbol{\Gamma}_1, \boldsymbol{\Gamma}_2$, respectively. Note that

$$\mathbf{0} \prec \mathbf{D} \preceq \boldsymbol{\Sigma}_{\mathbf{X}} \qquad (57)$$

$$(\mathbf{D})_i \preceq \mathbf{D}_i, \quad i = 1, 2 \qquad (58)$$

$$\mathbf{0} \prec \boldsymbol{\Gamma}_i \preceq \boldsymbol{\Sigma}_{\hat{\mathbf{N}}_i}, \quad i = 1, 2. \qquad (59)$$

Moreover, one can show by following the derivation of (10) in [12] that

$$\text{diag}(\boldsymbol{\Gamma}_1, \boldsymbol{\Gamma}_2) \preceq (\mathbf{D}^{-1} + \boldsymbol{\Sigma}_{\mathbf{N}}^{-1})^{-1}. \qquad (60)$$

Specifically, $\text{diag}(\boldsymbol{\Gamma}_1, \boldsymbol{\Gamma}_2)$ is the (reduced) error covariance matrix incurred by the MMSE estimator of $\mathbf{X}^n$ from $\mathbf{Y}^n$, W_1, and W_2, while $(\mathbf{D}^{-1} + \boldsymbol{\Sigma}_{\mathbf{N}}^{-1})^{-1}$ is the (reduced) error covariance matrix incurred by the (reduced) optimal linear estimator of $\mathbf{X}^n$ from $\mathbf{Y}^n$ and $\mathbb{E}[\mathbf{X}^n | W_1, W_2]$.

Now we proceed to derive a lower bound on $\frac{\alpha_1}{n} \log M_1 + \frac{\alpha_2}{n} \log M_2$ in terms of $\mathbf{D}$, $\boldsymbol{\Gamma}_1$, and $\boldsymbol{\Gamma}_2$. Let $\boldsymbol{\Sigma}^*$ be an optimal solution to the following minimization problem:

$$\begin{aligned} \min_{\mathbf{0} \prec \boldsymbol{\Sigma} \preceq \tilde{\boldsymbol{\Gamma}}_2} \quad & \frac{\alpha_1 - \alpha_2}{2} \log^+ \frac{|\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{N}}||\boldsymbol{\Sigma} + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}|}{|\mathbf{D} + \boldsymbol{\Sigma}_{\mathbf{N}}||\boldsymbol{\Sigma}_{\hat{\mathbf{X}}_2}|} \\ & + \frac{\alpha_2}{2} \log \frac{|\boldsymbol{\Sigma} + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}|}{|\boldsymbol{\Sigma}|} \end{aligned}$$

where $\tilde{\boldsymbol{\Gamma}}_2 = (\boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} \boldsymbol{\Sigma}_{\hat{\mathbf{N}}_2} \boldsymbol{\Gamma}_2^{-1} \boldsymbol{\Sigma}_{\hat{\mathbf{N}}_2} \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} - \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1})^{-1}$. It can be shown [cf., (11)–(13)] that there exist $\theta \in [0, 1]$ and $\mathbf{M} \succeq \mathbf{0}$ such that

$$\alpha_2(\boldsymbol{\Sigma}^*)^{-1} = ((\alpha_1 - \alpha_2)\theta + \alpha_2)(\boldsymbol{\Sigma}^* + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2})^{-1} + \mathbf{M} \qquad (61)$$

$$\mathbf{M}(\tilde{\boldsymbol{\Gamma}}_2 - \boldsymbol{\Sigma}^*) = \mathbf{0} \qquad (62)$$

$$\begin{aligned} & \frac{(\alpha_1 - \alpha_2)\theta}{2} \log^+ \frac{|\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{N}}||\boldsymbol{\Sigma}^* + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}|}{|\mathbf{D} + \boldsymbol{\Sigma}_{\mathbf{N}}||\boldsymbol{\Sigma}_{\hat{\mathbf{X}}_2}|} \\ & \quad + \frac{\alpha_2}{2} \log \frac{|\boldsymbol{\Sigma}^* + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}|}{|\boldsymbol{\Sigma}^*|} \\ & = \min_{\mathbf{0} \prec \boldsymbol{\Sigma} \preceq \tilde{\boldsymbol{\Gamma}}_2} \frac{\alpha_1 - \alpha_2}{2} \log^+ \frac{|\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{N}}||\boldsymbol{\Sigma} + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}|}{|\mathbf{D} + \boldsymbol{\Sigma}_{\mathbf{N}}||\boldsymbol{\Sigma}_{\hat{\mathbf{X}}_2}|} \\ & \quad + \frac{\alpha_2}{2} \log \frac{|\boldsymbol{\Sigma} + \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2}|}{|\boldsymbol{\Sigma}|}. \end{aligned} \qquad (63)$$

Note that

$$
\begin{aligned}
&\frac{\alpha_1}{n}\log M_1 + \frac{\alpha_2}{n}\log M_2 \\
&\geq \frac{\alpha_1 - \alpha_2}{n} I(\mathbf{Y}^n, \mathbf{X}_1^n; W_1 | W_2) \\
&\quad + \frac{\alpha_2}{n} I(\mathbf{Y}^n, \mathbf{X}_1^n, \mathbf{X}_2^n; W_1, W_2) \\
&= \frac{\alpha_1 - \alpha_2}{n} I(\mathbf{Y}^n; W_1 | W_2) + \frac{\alpha_1}{n} I(\mathbf{X}_1^n; W_1 | \mathbf{Y}^n) \\
&\quad + \frac{\alpha_2}{n}\left(I(\mathbf{Y}^n; W_1, W_2) + I(\mathbf{X}_2^n; W_2 | \mathbf{Y}^n)\right) \\
&\geq \frac{(\alpha_1 - \alpha_2)\theta}{n}\left(I(\mathbf{Y}^n; W_1, W_2) - I(\hat{\mathbf{X}}_2^n; W_2)\right) \\
&\quad + \frac{\alpha_1}{n} I(\mathbf{X}_1^n; W_1 | \mathbf{Y}^n) + \frac{\alpha_2}{n}\left(I(\mathbf{Y}^n; W_1, W_2)\right. \\
&\quad \left. + I(\tilde{\mathbf{X}}_2^n; W_2 | \hat{\mathbf{X}}_2^n)\right) \\
&\geq \frac{(\alpha_1 - \alpha_2)\theta}{2}\log\frac{|\mathbf{\Sigma_X} + \mathbf{\Sigma_N}|}{|\mathbf{D} + \mathbf{\Sigma_N}|} \\
&\quad - \frac{(\alpha_1 - \alpha_2)\theta}{2}\log\left((2\pi e)^{m_2}|\mathbf{\Sigma}_{\hat{\mathbf{X}}_2}|\right) + \frac{\alpha_1}{2}\log\frac{|\mathbf{\Sigma}_{\hat{\mathbf{N}}_1}|}{|\mathbf{\Gamma}_1|} \\
&\quad + \frac{\alpha_2}{2}\log\frac{|\mathbf{\Sigma_X} + \mathbf{\Sigma_N}|}{|\mathbf{D} + \mathbf{\Sigma_N}|} + \frac{\alpha_2}{2}\log\frac{|\mathbf{\Sigma}_{\tilde{\mathbf{X}}_2 | \hat{\mathbf{X}}_2}|}{|\mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}|} \\
&\quad + \frac{(\alpha_1 - \alpha_2)\theta + \alpha_2}{n} h(\hat{\mathbf{X}}_2^n | W_2) - \frac{\alpha_2}{n} h(\tilde{\mathbf{X}}_2^n | W_2). \quad (64)
\end{aligned}
$$

In view of (61), (62), and the fact that $\mathbf{\Sigma}_{\tilde{\mathbf{X}}_2^n | \hat{\mathbf{X}}_2^n, W_2} = \mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}\mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1}\mathbf{\Gamma}_2\mathbf{\Sigma}_{\hat{\mathbf{N}}_2}^{-1}\mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}$, Theorem 1 can be readily applied to show that

$$
\begin{aligned}
&\frac{(\alpha_1 - \alpha_2)\theta + \alpha_2}{n} h(\hat{\mathbf{X}}_2^n | W_2) - \frac{\alpha_2}{n} h(\tilde{\mathbf{X}}_2^n | W_2) \\
&\geq \frac{(\alpha_1 - \alpha_2)\theta + \alpha_2}{2}\log\left((2\pi e)^m |\mathbf{\Sigma}^* + \mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}|\right) \\
&\quad - \frac{\alpha_2}{2}\log\left((2\pi e)^m |\mathbf{\Sigma}^*|\right). \quad (65)
\end{aligned}
$$

Substituting (65) into (64) and invoking (63) yields

$$
\begin{aligned}
&\frac{\alpha_1}{n}\log M_1 + \frac{\alpha_2}{n}\log M_2 \geq \\
&\min_{\mathbf{0} \prec \mathbf{\Sigma} \preceq \tilde{\mathbf{\Gamma}}_2} \frac{\alpha_1 - \alpha_2}{2}\log^+\frac{|\mathbf{\Sigma_X} + \mathbf{\Sigma_N}||\mathbf{\Sigma} + \mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}|}{|\mathbf{D} + \mathbf{\Sigma_N}||\mathbf{\Sigma}_{\hat{\mathbf{X}}_2}|} \\
&+ \frac{\alpha_1}{2}\log\frac{|\mathbf{\Sigma}_{\hat{\mathbf{N}}_1}|}{|\mathbf{\Gamma}_1|} + \frac{\alpha_2}{2}\log\frac{|\mathbf{\Sigma_X} + \mathbf{\Sigma_N}||\mathbf{\Sigma}_{\tilde{\mathbf{X}}_2 | \hat{\mathbf{X}}_2}||\mathbf{\Sigma} + \mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}|}{|\mathbf{D} + \mathbf{\Sigma_N}||\mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}||\mathbf{\Sigma}|}
\end{aligned}
$$

which can be shown, through algebraic manipulations, to be equivalent to

$$
\begin{aligned}
&\frac{\alpha_1}{n}\log M_1 + \frac{\alpha_2}{n}\log M_2 \geq \\
&\min_{\mathbf{0} \prec \mathbf{\Gamma}_2' \preceq \mathbf{\Gamma}_2} \frac{\alpha_1}{2}\log\frac{|\mathbf{\Sigma}_{\hat{\mathbf{N}}_1}|}{|\mathbf{\Gamma}_1|} + \frac{\alpha_2}{2}\log\frac{|\mathbf{\Sigma_X} + \mathbf{\Sigma_N}||\mathbf{\Sigma}_{\hat{\mathbf{N}}_2}|}{|\mathbf{D} + \mathbf{\Sigma_N}||\mathbf{\Gamma}_2'|} \\
&+ \frac{\alpha_1 - \alpha_2}{2}\log^+\frac{|\mathbf{\Sigma_X} + \mathbf{\Sigma_N}|}{|\mathbf{D} + \mathbf{\Sigma_N}||\mathbf{\Sigma}_{\hat{\mathbf{X}}_2}||\mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} - \mathbf{\Sigma}_{\hat{\mathbf{N}}_2}^{-1}\mathbf{\Gamma}_2'\mathbf{\Sigma}_{\hat{\mathbf{N}}_2}^{-1}|}.
\end{aligned}
$$
$$(66)$$

It is clear that (57)–(60) and (66) together imply $R(\mathbf{D}_1, \mathbf{D}_2, \alpha_1, \alpha_2) \geq \underline{R}'(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma_N}, \alpha_1, \alpha_2)$, where

$$
\begin{aligned}
&\underline{R}'(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma_N}, \alpha_1, \alpha_2) \triangleq \\
&\min_{\mathbf{D}, \mathbf{\Gamma}_1, \mathbf{\Gamma}_2, \mathbf{\Gamma}_2'} \frac{\alpha_1}{2}\log\frac{|\mathbf{\Sigma}_{\hat{\mathbf{N}}_1}|}{|\mathbf{\Gamma}_1|} + \frac{\alpha_2}{2}\log\frac{|\mathbf{\Sigma_X} + \mathbf{\Sigma_N}||\mathbf{\Sigma}_{\hat{\mathbf{N}}_2}|}{|\mathbf{D} + \mathbf{\Sigma_N}||\mathbf{\Gamma}_2'|} \\
&\quad + \frac{\alpha_1 - \alpha_2}{2}\log^+\frac{|\mathbf{\Sigma_X} + \mathbf{\Sigma_N}|}{|\mathbf{D} + \mathbf{\Sigma_N}||\mathbf{\Sigma}_{\hat{\mathbf{X}}_2}||\mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} - \mathbf{\Sigma}_{\hat{\mathbf{N}}_2}^{-1}\mathbf{\Gamma}_2'\mathbf{\Sigma}_{\hat{\mathbf{N}}_2}^{-1}|} \\
&\text{subject to} \quad \mathrm{diag}(\mathbf{\Gamma}_1, \mathbf{\Gamma}_2) \preceq (\mathbf{D}^{-1} + \mathbf{\Sigma_N}^{-1})^{-1}, \\
&\qquad\qquad \mathbf{0} \prec \mathbf{D} \preceq \mathbf{\Sigma_X}, \\
&\qquad\qquad (\mathbf{D})_i \preceq \mathbf{D}_i, \quad i = 1, 2, \\
&\qquad\qquad \mathbf{0} \prec \mathbf{\Gamma}_i \preceq \mathbf{\Sigma}_{\hat{\mathbf{N}}_i}, \quad i = 1, 2, \\
&\qquad\qquad \mathbf{0} \prec \mathbf{\Gamma}_2' \preceq \mathbf{\Gamma}_2.
\end{aligned}
$$

The desired result follows by recognizing the fact that $\underline{R}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma_N}, \alpha_1, \alpha_2) = \underline{R}'(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma_N}, \alpha_1, \alpha_2)$.

APPENDIX D
PROOF OF LEMMA 1

Consider the following minimization problem (which is essentially a relaxed version of the minimization problem associated with $\underline{R}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma_N}, \alpha_1, \alpha_2)$)

$$
\begin{aligned}
(\hat{\mathbf{P}}) \quad &\min_{\mathbf{D}, \mathbf{\Gamma}_1, \mathbf{\Gamma}_2} - \alpha_1 \log|\mathbf{D} + \mathbf{\Sigma_N}| \\
&\quad - (\alpha_1 - \alpha_2)\log|\mathbf{\Sigma}_{\hat{\mathbf{N}}_2}\mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1}\mathbf{\Sigma}_{\hat{\mathbf{N}}_2} - \mathbf{\Gamma}_2| \\
&\quad - \alpha_1 \log|\mathbf{\Gamma}_1| - \alpha_2 \log|\mathbf{\Gamma}_2| \\
&\text{subject to} \quad \mathrm{diag}(\mathbf{\Gamma}_1, \mathbf{\Gamma}_2) \preceq (\mathbf{D}^{-1} + \mathbf{\Sigma_N}^{-1})^{-1}, \\
&\qquad\qquad \mathbf{0} \prec \mathbf{D}, \\
&\qquad\qquad \mathbf{E}_i \mathbf{D} \mathbf{E}_i \preceq \mathbf{F}_i \mathbf{D}_i \mathbf{F}_i^T, \quad i = 1, 2, \\
&\qquad\qquad \mathbf{0} \prec \mathbf{\Gamma}_i, \quad i = 1, 2,
\end{aligned}
$$

where $\mathbf{E}_1 = \mathrm{diag}(\mathbf{I}_{m_1}, \mathbf{0})$, $\mathbf{E}_2 = \mathrm{diag}(\mathbf{0}, \mathbf{I}_{m_2})$, $\mathbf{F}_1 = (\mathbf{I}_{m_1}, \mathbf{0})^T$, and $\mathbf{F}_2 = (\mathbf{0}, \mathbf{I}_{m_2})^T$. The Lagrangian of $(\hat{\mathbf{P}})$ is given by

$$
\begin{aligned}
&L(\mathbf{D}, \mathbf{\Gamma}_1, \mathbf{\Gamma}_2) \\
&= -\alpha_1 \log|\mathbf{D} + \mathbf{\Sigma_N}| - (\alpha_1 - \alpha_2)\log|\mathbf{\Sigma}_{\hat{\mathbf{N}}_2}\mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1}\mathbf{\Sigma}_{\hat{\mathbf{N}}_2} - \mathbf{\Gamma}_2| \\
&\quad - \alpha_1 \log|\mathbf{\Gamma}_1| - \alpha_2 \log|\mathbf{\Gamma}_2| \\
&\quad + \mathrm{tr}(\mathbf{\Lambda}(\mathrm{diag}(\mathbf{\Gamma}_1, \mathbf{\Gamma}_2) - (\mathbf{D}^{-1} + \mathbf{\Sigma_N}^{-1})^{-1})) \\
&\quad + \mathrm{tr}(\mathbf{\Pi}_1 \mathbf{E}_1 \mathbf{D} \mathbf{E}_1) + \mathrm{tr}(\mathbf{\Pi}_2 \mathbf{E}_2 \mathbf{D} \mathbf{E}_2)
\end{aligned}
$$

where $\mathbf{\Lambda}$, $\mathbf{\Pi}_1$, and $\mathbf{\Pi}_2$ are positive semidefinite matrices. Since $(\hat{\mathbf{P}})$ is a convex semidefinite programming problem,

$(\hat{\mathbf{D}}, \hat{\mathbf{\Gamma}}_1, \hat{\mathbf{\Gamma}}_2)$ is an optimal solution to $(\hat{\mathbf{P}})$ if it satisfies the following KKT conditions:

$$\nabla_{\mathbf{D}} L(\mathbf{D}, \mathbf{\Gamma}_1, \mathbf{\Gamma}_2)|_{\mathbf{D}=\hat{\mathbf{D}}} = \mathbf{0}$$

$$\nabla_{\mathbf{\Gamma}_i} L(\mathbf{D}, \mathbf{\Gamma}_1, \mathbf{\Gamma}_2)|_{\mathbf{\Gamma}_i=\hat{\mathbf{\Gamma}}_i} = \mathbf{0}, \quad i = 1, 2$$

$$\mathbf{0} \prec \hat{\mathbf{D}}$$

$$\mathbf{E}_i \hat{\mathbf{D}} \mathbf{E}_i \preceq \mathbf{F}_i \mathbf{D}_i \mathbf{F}_i^T, \quad i = 1, 2$$

$$\mathbf{0} \prec \hat{\mathbf{\Gamma}}_i, \quad i = 1, 2$$

$$\mathbf{\Lambda}(\operatorname{diag}(\hat{\mathbf{\Gamma}}_1, \hat{\mathbf{\Gamma}}_2) - (\hat{\mathbf{D}}^{-1} + \mathbf{\Sigma}_{\mathbf{N}}^{-1})^{-1}) = \mathbf{0}$$

$$\mathbf{\Pi}_i(\mathbf{E}_i \hat{\mathbf{D}} \mathbf{E}_i - \mathbf{F}_i \mathbf{D}_i \mathbf{F}_i^T) = \mathbf{0}, \quad i = 1, 2$$

which can be written equivalently as

$$\hat{\mathbf{D}}(\mathbf{E}_1 \mathbf{\Pi}_1 \mathbf{E}_1 + \mathbf{E}_2 \mathbf{\Pi}_2 \mathbf{E}_2 - \alpha_1(\hat{\mathbf{D}} + \mathbf{\Sigma}_{\mathbf{N}})^{-1})\hat{\mathbf{D}}$$
$$= (\hat{\mathbf{D}}^{-1} + \mathbf{\Sigma}_{\mathbf{N}}^{-1})^{-1} \mathbf{\Lambda} (\hat{\mathbf{D}}^{-1} + \mathbf{\Sigma}_{\mathbf{N}}^{-1})^{-1} \tag{67}$$

$$(\mathbf{\Lambda})_1 = \alpha_1 \hat{\mathbf{\Gamma}}_1^{-1} \tag{68}$$

$$(\mathbf{\Lambda})_2 = \alpha_2 \hat{\mathbf{\Gamma}}_2^{-1} - (\alpha_1 - \alpha_2)(\mathbf{\Sigma}_{\hat{\mathbf{N}}_2} \mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} \mathbf{\Sigma}_{\hat{\mathbf{N}}_2} - \hat{\mathbf{\Gamma}}_2)^{-1} \tag{69}$$

$$\mathbf{0} \prec \hat{\mathbf{D}} \tag{70}$$

$$\mathbf{E}_i \hat{\mathbf{D}} \mathbf{E}_i \preceq \mathbf{F}_i \mathbf{D}_i \mathbf{F}_i^T, \quad i = 1, 2 \tag{71}$$

$$\mathbf{0} \prec \hat{\mathbf{\Gamma}}_i, \quad i = 1, 2 \tag{72}$$

$$\mathbf{\Lambda}(\operatorname{diag}(\hat{\mathbf{\Gamma}}_1, \hat{\mathbf{\Gamma}}_2) - (\hat{\mathbf{D}}^{-1} + \mathbf{\Sigma}_{\mathbf{N}}^{-1})^{-1}) = \mathbf{0} \tag{73}$$

$$\mathbf{\Pi}_i(\mathbf{E}_i \hat{\mathbf{D}} \mathbf{E}_i - \mathbf{F}_i \mathbf{D}_i \mathbf{F}_i^T) = \mathbf{0}, \quad i = 1, 2. \tag{74}$$

Now assume $\hat{\mathbf{D}} \in \mathbb{D}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma}_{\mathbf{X}})$ and $\mathbf{\Sigma}_{\mathbf{N}} \in \mathbb{N}(\mathbf{\Sigma}_{\mathbf{X}})$. Since $\hat{\mathbf{D}}^{-1} + \mathbf{\Sigma}_{\mathbf{N}}^{-1}$ is an (m_1, m_2) block diagonal matrix, we can construct $\hat{\mathbf{\Gamma}}_1$ and $\hat{\mathbf{\Gamma}}_2$ such that

$$\operatorname{diag}(\hat{\mathbf{\Gamma}}_1, \hat{\mathbf{\Gamma}}_2) = (\hat{\mathbf{D}}^{-1} + \mathbf{\Sigma}_{\mathbf{N}}^{-1})^{-1}. \tag{75}$$

It is easy to show that if $(\hat{\mathbf{D}}, \hat{\mathbf{\Gamma}}_1, \hat{\mathbf{\Gamma}}_2)$ is an optimal solution to $(\hat{\mathbf{P}})$, then it is also an optimal solution to the minimization problem associated with $\underline{R}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma}_{\mathbf{N}}, \alpha_1, \alpha_2)$, which, together with (75), implies $\underline{R}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma}_{\mathbf{N}}, \alpha_1, \alpha_2) = \overline{R}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma}_{\mathbf{N}}, \alpha_1, \alpha_2)$. The proof is completed by noticing the fact that our choice of $(\hat{\mathbf{D}}, \hat{\mathbf{\Gamma}}_1, \hat{\mathbf{\Gamma}}_2)$ satisfies the conditions (70)–(74) and the fact that the existence of positive semidefinite matrices $\mathbf{\Lambda}$, $\mathbf{\Pi}_1$, and $\mathbf{\Pi}_2$ satisfying (67)–(69) is equivalent to the existence of an (m_1, m_2) block diagonal matrix $\mathbf{\Pi}$ satisfying

$$\mathbf{\Pi} \succeq \alpha_1(\hat{\mathbf{D}} + \mathbf{\Sigma}_{\mathbf{N}})^{-1}$$
$$(\hat{\mathbf{D}}(\mathbf{\Pi} - \alpha_1(\hat{\mathbf{D}} + \mathbf{\Sigma}_{\mathbf{N}})^{-1})\hat{\mathbf{D}})_1 = \alpha_1 \hat{\mathbf{\Gamma}}_1$$
$$(\hat{\mathbf{D}}(\mathbf{\Pi} - \alpha_1(\hat{\mathbf{D}} + \mathbf{\Sigma}_{\mathbf{N}})^{-1})\hat{\mathbf{D}})_2$$
$$= \alpha_2 \hat{\mathbf{\Gamma}}_2 - (\alpha_1 - \alpha_2)\hat{\mathbf{\Gamma}}_2(\mathbf{\Sigma}_{\hat{\mathbf{N}}_2} \mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} \mathbf{\Sigma}_{\hat{\mathbf{N}}_2} - \hat{\mathbf{\Gamma}}_2)^{-1} \hat{\mathbf{\Gamma}}_2.$$

APPENDIX E
PROOF OF LEMMA 3

We shall first show that there exists a unique positive-definite matrix $\mathbf{D}$ such that $(\mathbf{D})_i = \mathbf{D}_i$, $i = 1, 2$, and $\mathbf{D}^{-1} - \mathbf{\Sigma}_{\mathbf{X}}^{-1}$ is an (m_1, m_2) block diagonal matrix. Let $\mathbf{D}$ be a positive-definite matrix of the form

$$\begin{pmatrix} \mathbf{D}_1 & \mathbf{\Theta} \\ \mathbf{\Theta}^T & \mathbf{D}_2 \end{pmatrix}.$$

Note that

$$\mathbf{\Sigma}_{\mathbf{X}}^{-1} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_3 \\ \mathbf{A}_3^T & \mathbf{A}_2 \end{pmatrix}$$

$$\mathbf{D}^{-1} = \begin{pmatrix} \mathbf{B}_1 & \mathbf{B}_3 \\ \mathbf{B}_3^T & \mathbf{B}_2 \end{pmatrix}$$

where

$$\mathbf{A}_1 = \mathbf{\Sigma}_{\mathbf{X}_1}^{-1} + \mathbf{\Sigma}_{\mathbf{X}_1}^{-1} \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2} \mathbf{A}_2 \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}^T \mathbf{\Sigma}_{\mathbf{X}_1}^{-1}$$

$$\mathbf{A}_2 = (\mathbf{\Sigma}_{\mathbf{X}_2} - \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}^T \mathbf{\Sigma}_{\mathbf{X}_1}^{-1} \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2})^{-1}$$

$$\mathbf{A}_3 = -\mathbf{\Sigma}_{\mathbf{X}_1}^{-1} \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2} \mathbf{A}_2$$

$$\mathbf{B}_1 = \mathbf{D}_1^{-1} + \mathbf{D}_1^{-1} \mathbf{\Theta} \mathbf{B}_2 \mathbf{\Theta}^T \mathbf{D}_1^{-1}$$

$$\mathbf{B}_2 = (\mathbf{D}_2 - \mathbf{\Theta}^T \mathbf{D}_1^{-1} \mathbf{\Theta})^{-1}$$

$$\mathbf{B}_3 = -\mathbf{D}_1^{-1} \mathbf{\Theta} \mathbf{B}_2.$$

Clearly, $\mathbf{D}^{-1} - \mathbf{\Sigma}_{\mathbf{X}}^{-1}$ is an (m_1, m_2) block diagonal matrix if and only if $\mathbf{A}_3 = \mathbf{B}_3$, i.e.,

$$\mathbf{\Sigma}_{\mathbf{X}_1}^{-1} \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2} (\mathbf{\Sigma}_{\mathbf{X}_2} - \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}^T \mathbf{\Sigma}_{\mathbf{X}_1}^{-1} \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2})^{-1}$$
$$= \mathbf{D}_1^{-1} \mathbf{\Theta} (\mathbf{D}_2 - \mathbf{\Theta}^T \mathbf{D}_1^{-1} \mathbf{\Theta})^{-1}. \tag{76}$$

We shall show that $\mathbf{\Theta}$ is uniquely determined by (76). Define $\mathbf{H} = \mathbf{D}_1^{-\frac{1}{2}} \mathbf{\Theta} \mathbf{D}_2^{-\frac{1}{2}}$ and

$$\mathbf{M} = \mathbf{D}_1^{\frac{1}{2}} \mathbf{\Sigma}_{\mathbf{X}_1}^{-1} \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2} (\mathbf{\Sigma}_{\mathbf{X}_2} - \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}^T \mathbf{\Sigma}_{\mathbf{X}_1}^{-1} \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2})^{-1} \mathbf{D}_2^{\frac{1}{2}}.$$

Since $\mathbf{D}$ is positive definite, it follows that $\mathbf{I}_{m_2} - \mathbf{H}^T \mathbf{H} \succ \mathbf{0}$. We can write (76) equivalently as

$$\mathbf{M}(\mathbf{I}_{m_2} - \mathbf{H}^T \mathbf{H}) = \mathbf{H}. \tag{77}$$

Let $\mathbf{C} = \mathbf{H} \mathbf{M}^T$. Note that (77) implies

$$\mathbf{C}^T \mathbf{C} + \mathbf{C} - \mathbf{M} \mathbf{M}^T = \mathbf{0}$$

i.e., $\mathbf{C} = \mathbf{M}(\mathbf{I}_{m_2} - \mathbf{H}^T \mathbf{H})\mathbf{M}^T$. Therefore, $\mathbf{C}$ is symmetric and $\mathbf{C} \succeq \mathbf{0}$. Now it can be readily verified that

$$\mathbf{C}^T \mathbf{C} + \mathbf{C} - \mathbf{M} \mathbf{M}^T = \mathbf{0}$$
$$\Rightarrow \mathbf{C}^2 + \mathbf{C} - \mathbf{M} \mathbf{M}^T = \mathbf{0}$$
$$\Rightarrow (\frac{1}{2}\mathbf{I}_{m_1} + \mathbf{C})^2 = \frac{1}{4}\mathbf{I}_{m_1} + \mathbf{M} \mathbf{M}^T$$
$$\Rightarrow \mathbf{C} = (\frac{1}{4}\mathbf{I}_{m_1} + \mathbf{M} \mathbf{M}^T)^{\frac{1}{2}} - \frac{1}{2}\mathbf{I}_{m_1}. \tag{78}$$

Let $\mathbf{U} \mathbf{\Lambda} \mathbf{V}^T$ be the singular value decomposition of $\mathbf{M}$, where $\mathbf{U}$ and $\mathbf{V}$ are, respectively, $m_1 \times m_1$ and $m_2 \times m_2$ unitary matrices, and $\mathbf{\Lambda}$ is an $m_1 \times m_2$ diagonal matrix. Without loss of generality, we assume that $\mathbf{\Lambda}$ is of the form

$$\begin{pmatrix} \operatorname{diag}(\lambda_1, \ldots, \lambda_r) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

where $\lambda_1, \ldots, \lambda_r$ are positive. In view of (78), we have

$$\mathbf{U}^T \mathbf{H} \mathbf{V} \mathbf{\Lambda}^T = \begin{pmatrix} \operatorname{diag}(\mu_1, \ldots, \mu_r) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

where $\mu_i = (\frac{1}{4} + \lambda_i^2)^{\frac{1}{2}} - \frac{1}{2}, i = 1, \ldots, r$. Note that (77), together with the fact $\mathbf{I}_{m_2} - \mathbf{H}^T \mathbf{H} \succ \mathbf{0}$, implies that $\mathbf{H}$ and $\mathbf{M}$ have the same rank. As a consequence, $\mathbf{H}$ must be of the form

$$\mathbf{U} \begin{pmatrix} \operatorname{diag}(\frac{\mu_1}{\lambda_1}, \ldots, \frac{\mu_r}{\lambda_r}) & \mathbf{\Phi} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{V}^T.$$

Again by (77), we must have

$$\mathbf{\Phi} = -\operatorname{diag}(\mu_1, \ldots, \mu_r) \mathbf{\Phi}$$

which implies $\mathbf{\Phi} = \mathbf{0}$. Therefore, $\mathbf{\Theta}$ is uniquely given by

$$\mathbf{\Theta} = \mathbf{D}_1^{\frac{1}{2}} \mathbf{U} \begin{pmatrix} \operatorname{diag}(\frac{\mu_1}{\lambda_1}, \ldots, \frac{\mu_1}{\lambda_1}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{V}^T \mathbf{D}_2^{\frac{1}{2}}. \qquad (79)$$

It can be verified that the resulting $\mathbf{D}$ is positive definite and $\mathbf{D}^{-1} - \mathbf{\Sigma}_{\mathbf{X}}^{-1}$ is an (m_1, m_2) block diagonal matrix.

Now we proceed to show that for this $\mathbf{D}$, if $(\mathbf{D}_1, \mathbf{D}_2) \in \mathcal{D}_0$, then $\mathbf{B}_i \succeq \mathbf{A}_i, i = 1, 2$. By symmetry, it suffices to prove that $\mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}^T (\mathbf{\Sigma}_{\mathbf{X}_1}^{-1} \mathbf{D}_1 \mathbf{\Sigma}_{\mathbf{X}_1}^{-1} - \mathbf{\Sigma}_{\mathbf{X}_1}^{-1}) \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2} + \mathbf{\Sigma}_{\mathbf{X}_2} \succeq \mathbf{D}_2$ implies $\mathbf{B}_2 \succeq \mathbf{A}_2$, i.e.,

$$(\mathbf{D}_2 - \mathbf{\Theta}^T \mathbf{D}_1^{-1} \mathbf{\Theta})^{-1} \succeq (\mathbf{\Sigma}_{\mathbf{X}_2} - \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}^T \mathbf{\Sigma}_{\mathbf{X}_1}^{-1} \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2})^{-1}.$$

Note that

$$\mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}^T (\mathbf{\Sigma}_{\mathbf{X}_1}^{-1} \mathbf{D}_1 \mathbf{\Sigma}_{\mathbf{X}_1}^{-1} - \mathbf{\Sigma}_{\mathbf{X}_1}^{-1}) \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2} + \mathbf{\Sigma}_{\mathbf{X}_2} \succeq \mathbf{D}_2$$
$$\Rightarrow \mathbf{D}_2^{-\frac{1}{2}} (\mathbf{\Sigma}_{\mathbf{X}_2} - \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}^T \mathbf{\Sigma}_{\mathbf{X}_1}^{-1} \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}) \mathbf{D}_2^{-\frac{1}{2}}$$
$$+ \mathbf{D}_2^{-\frac{1}{2}} \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}^T \mathbf{\Sigma}_{\mathbf{X}_1}^{-1} \mathbf{D}_1 \mathbf{\Sigma}_{\mathbf{X}_1}^{-1} \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2} \mathbf{D}_2^{-\frac{1}{2}} \succeq \mathbf{I}_{m_2}. \quad (80)$$

Define $\mathbf{G} = \mathbf{D}_2^{-\frac{1}{2}} (\mathbf{\Sigma}_{\mathbf{X}_2} - \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}^T \mathbf{\Sigma}_{\mathbf{X}_1}^{-1} \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}) \mathbf{D}_2^{-\frac{1}{2}}$, $\mathbf{Q} = (\mathbf{M}^T \mathbf{M})^{\frac{1}{2}}$, and $\mathbf{Q}(\epsilon) = \mathbf{Q} + \epsilon \mathbf{I}_{m_2}$, where $\epsilon > 0$. Continuing from (80), we have

$$\mathbf{D}_2^{-\frac{1}{2}} (\mathbf{\Sigma}_{\mathbf{X}_2} - \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}^T \mathbf{\Sigma}_{\mathbf{X}_1}^{-1} \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}) \mathbf{D}_2^{-\frac{1}{2}}$$
$$+ \mathbf{D}_2^{-\frac{1}{2}} \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}^T \mathbf{\Sigma}_{\mathbf{X}_1}^{-1} \mathbf{D}_1 \mathbf{\Sigma}_{\mathbf{X}_1}^{-1} \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2} \mathbf{D}_2^{-\frac{1}{2}} \succeq \mathbf{I}_{m_2}$$
$$\Rightarrow \mathbf{G} + \mathbf{G} \mathbf{Q}^2 \mathbf{G} \succeq \mathbf{I}_{m_2}$$
$$\Rightarrow \mathbf{G} + \mathbf{G} \mathbf{Q}^2(\epsilon) \mathbf{G} \succeq \mathbf{I}_{m_2}$$
$$\Rightarrow \mathbf{Q}(\epsilon) \mathbf{G} \mathbf{Q}(\epsilon) + \mathbf{Q}(\epsilon) \mathbf{G} \mathbf{Q}^2(\epsilon) \mathbf{G} \mathbf{Q}(\epsilon) \succeq \mathbf{Q}^2(\epsilon)$$
$$\Rightarrow (\frac{1}{2} \mathbf{I}_{m_2} + \mathbf{Q}(\epsilon) \mathbf{G} \mathbf{Q}(\epsilon))^2 \succeq \frac{1}{4} \mathbf{I}_{m_2} + \mathbf{Q}^2(\epsilon)$$
$$\Rightarrow \mathbf{Q}(\epsilon) \mathbf{G} \mathbf{Q}(\epsilon) \succeq (\frac{1}{4} \mathbf{I}_{m_2} + \mathbf{Q}^2(\epsilon))^{\frac{1}{2}} - \frac{1}{2} \mathbf{I}_{m_2}$$
$$\Rightarrow \mathbf{G} \succeq \mathbf{Q}^{-1}(\epsilon) \left[(\frac{1}{4} \mathbf{I}_{m_2} + \mathbf{Q}^2(\epsilon))^{\frac{1}{2}} - \frac{1}{2} \mathbf{I}_{m_2} \right] \mathbf{Q}^{-1}(\epsilon)$$
$$\Rightarrow \mathbf{G}^{-1} \preceq \mathbf{Q}(\epsilon) \left[(\frac{1}{4} \mathbf{I}_{m_2} + \mathbf{Q}^2(\epsilon))^{\frac{1}{2}} - \frac{1}{2} \mathbf{I}_{m_2} \right]^{-1} \mathbf{Q}(\epsilon).$$

Since $\mathbf{Q}(\epsilon) = (\mathbf{M}^T \mathbf{M})^{\frac{1}{2}} + \epsilon \mathbf{I}_{m_2} = \mathbf{V}((\mathbf{\Lambda}^T \mathbf{\Lambda})^{\frac{1}{2}} + \epsilon \mathbf{I}_{m_2}) \mathbf{V}^T$, it follows that

$$\mathbf{Q}(\epsilon) \left[(\frac{1}{4} \mathbf{I}_{m_2} + \mathbf{Q}^2(\epsilon))^{\frac{1}{2}} - \frac{1}{2} \mathbf{I}_{m_2} \right]^{-1} \mathbf{Q}(\epsilon)$$
$$= \mathbf{V} \operatorname{diag}(\nu_1(\epsilon), \ldots, \nu_{m_2}(\epsilon)) \mathbf{V}^T$$

where $\nu_i(\epsilon) = (\lambda_i + \epsilon)^2 ((\frac{1}{4} + (\lambda_i + \epsilon)^2)^{\frac{1}{2}} - \frac{1}{2})^{-1}$ for $i = 1, \ldots, r$, and $\nu_i(\epsilon) = \epsilon^2 ((\frac{1}{4} + \epsilon^2)^{\frac{1}{2}} - \frac{1}{2})^{-1}$ for $i = r + 1, \ldots, m_2$. Note that $\nu_i(\epsilon) \to \nu_i, i = 1, \ldots, m_2$, as $\epsilon \to 0$, where $\nu_i = \lambda_i^2 ((\frac{1}{4} + \lambda_i^2)^{\frac{1}{2}} - \frac{1}{2})^{-1}$ for $i = 1, \ldots, r$, and $\nu_i = 1$ for $i = r + 1, \ldots, m_2$. Therefore, we have $\mathbf{G}^{-1} \preceq \mathbf{V} \operatorname{diag}(\nu_1, \ldots, \nu_{m_2}) \mathbf{V}^T$, i.e.,

$$(\mathbf{\Sigma}_{\mathbf{X}_2} - \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}^T \mathbf{\Sigma}_{\mathbf{X}_1}^{-1} \mathbf{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2})^{-1}$$
$$\preceq \mathbf{D}_2^{-\frac{1}{2}} \mathbf{V} \operatorname{diag}(\nu_1, \ldots, \nu_{m_2}) \mathbf{V}^T \mathbf{D}_2^{-\frac{1}{2}}.$$

It can be shown by leveraging (79) that

$$(\mathbf{D}_2 - \mathbf{\Theta}^T \mathbf{D}_1^{-1} \mathbf{\Theta})^{-1} = \mathbf{D}_2^{-\frac{1}{2}} \mathbf{V} \operatorname{diag}(\nu_1, \ldots, \nu_{m_2}) \mathbf{V}^T \mathbf{D}_2^{-\frac{1}{2}}$$

which completes the proof.

APPENDIX F
PROOF OF THEOREM 7

We shall show that the condition of Lemma 2 is satisfied in the high-resolution regime. To this end, it suffices to consider the case where $\mathbb{D}(\mathbf{D}_1, \mathbf{D}_2, \mathbf{\Sigma}_{\mathbf{X}})$ is a singleton[5] and its element $\mathbf{D}$ satisfies $\mathbf{D} \prec \mathbf{\Sigma}_{\mathbf{X}}$. Let $\mathbf{G} = (\mathbf{D}^{-1} - \mathbf{\Sigma}_{\mathbf{X}}^{-1})^{-1}$ (which is a positive-definite (m_1, m_2) block diagonal matrix) and

$$\mathbf{A} \triangleq \begin{pmatrix} (\mathbf{A})_1 & \mathbf{A}_{1,2} \\ \mathbf{A}_{2,1} & (\mathbf{A})_2 \end{pmatrix} = \mathbf{G}^{\frac{1}{2}} (\mathbf{G} + \mathbf{\Sigma}_{\mathbf{X}})^{-1} \mathbf{G}^{\frac{1}{2}}.$$

Clearly, $\mathbf{D}, \mathbf{G}$, and $\mathbf{A}$ all converge to $\mathbf{0}$ as $\|\mathbf{D}_i\| \to \mathbf{0}, i = 1, 2$. It can also be verified that

$$\mathbf{D} = (\mathbf{G}^{-1} + \mathbf{\Sigma}_{\mathbf{X}}^{-1})^{-1}$$
$$= \mathbf{G}^{\frac{1}{2}} (\mathbf{I}_m + \mathbf{G}^{\frac{1}{2}} \mathbf{\Sigma}_{\mathbf{X}}^{-1} \mathbf{G}^{\frac{1}{2}})^{-1} \mathbf{G}^{\frac{1}{2}}$$
$$= \mathbf{G}^{\frac{1}{2}} (\mathbf{I}_m - \mathbf{A}) \mathbf{G}^{\frac{1}{2}}.$$

Now choose an arbitrary $\mathbf{\Sigma}_{\mathbf{N}}$ from $\mathbb{N}(\mathbf{\Sigma}_{\mathbf{X}})$. Note that

$$(\mathbf{D} \mathbf{\Pi} \mathbf{D})_1 = \alpha_1 \mathbf{D}_1$$
$$\iff ((\mathbf{I}_m - \mathbf{A}) \mathbf{H} (\mathbf{I}_m - \mathbf{A}))_1 = \tilde{\mathbf{\Omega}}_1 \qquad (81)$$

and

$$(\mathbf{D} \mathbf{\Pi} \mathbf{D})_2 = \alpha_1 \mathbf{D}_2 - (\alpha_1 - \alpha_2) \mathbf{K}$$
$$\iff ((\mathbf{I}_m - \mathbf{A}) \mathbf{H} (\mathbf{I}_m - \mathbf{A}))_2 = \tilde{\mathbf{\Omega}}_2 \qquad (82)$$

where $\mathbf{K} = ((\mathbf{D}^{-1} + \mathbf{\Sigma}_{\mathbf{N}}^{-1})_2 - \mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1} \mathbf{\Sigma}_{\tilde{\mathbf{N}}_2} \mathbf{\Sigma}_{\tilde{\mathbf{N}}_2}^{-1})^{-1}$, $\mathbf{H} = \mathbf{G}^{\frac{1}{2}} \mathbf{\Pi} \mathbf{G}^{\frac{1}{2}}$ (which is an (m_1, m_2) block diagonal matrix), $\tilde{\mathbf{\Omega}}_1 = \alpha_1 (\mathbf{I}_m - \mathbf{A})_1$, and

$$\tilde{\mathbf{\Omega}}_2 = \alpha_1 (\mathbf{I}_m - \mathbf{A})_2 - (\alpha_1 - \alpha_2) (\mathbf{G})_2^{-\frac{1}{2}} \mathbf{K} (\mathbf{G})_2^{-\frac{1}{2}}.$$

Define $\mathbf{B} = (\mathbf{B}_{i,j})$ with $\mathbf{B}_{i,i} = (\mathbf{I}_m - \mathbf{A})_i, i = 1, 2$, and $\mathbf{B}_{i,j} = -\mathbf{A}_{i,j}$ for $i \neq j$. We can rewrite (81) and (82) as

$$(\mathbf{B} * \mathbf{B}) \begin{pmatrix} \operatorname{vec}((\mathbf{H})_1) \\ \operatorname{vec}((\mathbf{H})_2) \end{pmatrix} = \begin{pmatrix} \operatorname{vec}(\tilde{\mathbf{\Omega}}_1) \\ \operatorname{vec}(\tilde{\mathbf{\Omega}}_2) \end{pmatrix}.$$

[5]Note that the condition $(\mathbf{D}_1, \mathbf{D}_2) \in \mathcal{D}_0$ in Lemma 3 is satisfied when $\|\mathbf{D}_1\|$ and $\|\mathbf{D}_2\|$ are sufficiently small.

Clearly, $\mathbf{B} * \mathbf{B} \to \mathbf{I}_{m_1^2 + m_2^2}$ and $\tilde{\boldsymbol{\Omega}}_1 \to \alpha_1 \mathbf{I}_{m_1}$ as $\|\mathbf{D}_i\| \to 0$, $i = 1, 2$; moreover

$$\tilde{\boldsymbol{\Omega}}_2 = \alpha_1 (\mathbf{I}_m - \mathbf{A})_2 - (\alpha_1 - \alpha_2)(\mathbf{G})_2^{-\frac{1}{2}} \mathbf{K}(\mathbf{G})_2^{-\frac{1}{2}}$$
$$= \alpha_1 (\mathbf{I}_m - \mathbf{A})_2$$
$$- (\alpha_1 - \alpha_2)\left(((\mathbf{I}_m - \mathbf{A})^{-1} + \mathbf{G}^{\frac{1}{2}}\boldsymbol{\Sigma}_{\mathbf{N}}^{-1}\mathbf{G}^{\frac{1}{2}})_2\right.$$
$$\left. - (\mathbf{G})_2^{\frac{1}{2}} \boldsymbol{\Sigma}_{\hat{\mathbf{N}}_2}^{-1} \boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2} \boldsymbol{\Sigma}_{\hat{\mathbf{N}}_2}^{-1}(\mathbf{G})_2^{\frac{1}{2}}\right)^{-1}$$

which converges to $\alpha_2 \mathbf{I}_{m_2}$ as $\|\mathbf{D}_i\| \to 0$, $i = 1, 2$. As a consequence, we must have $\mathbf{H} \to \text{diag}(\alpha_1 \mathbf{I}_{m_1}, \alpha_2 \mathbf{I}_{m_2})$ as $\|\mathbf{D}_i\| \to 0$, $i = 1, 2$. Since $\boldsymbol{\Pi} = \mathbf{G}^{-\frac{1}{2}} \mathbf{H} \mathbf{G}^{-\frac{1}{2}}$, it follows that $\boldsymbol{\Pi} - \alpha_1 (\mathbf{D} + \boldsymbol{\Sigma}_{\mathbf{N}})^{-1}$ is positive definite when $\|\mathbf{D}_i\|$, $i = 1, 2$, are sufficiently small.

APPENDIX G
PROOF OF THEOREM 8

We shall show that the condition of Lemma 2 is satisfied in the weak-dependence regime. In view of Lemma 3 and the fact that $(\mathbf{D}_1, \mathbf{D}_2) \in \mathcal{D}_0$ when $\|\boldsymbol{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}\|$ is sufficiently small, it suffices to consider the case where $\mathbb{D}(\mathbf{D}_1, \mathbf{D}_2, \boldsymbol{\Sigma}_{\mathbf{X}})$ is a singleton. Denote the element of $\mathbb{D}(\mathbf{D}_1, \mathbf{D}_2, \boldsymbol{\Sigma}_{\mathbf{X}})$ by $\mathbf{D}$. It is easy to see that $\mathbf{D} \to \text{diag}(\mathbf{D}_1, \mathbf{D}_2)$ as $\|\boldsymbol{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}\| \to 0$. Let $\boldsymbol{\Sigma}_{\mathbf{N}} = (\beta \text{diag}(\boldsymbol{\Sigma}_{\mathbf{X}_1}^{-1}, \boldsymbol{\Sigma}_{\mathbf{X}_2}^{-1}) - \boldsymbol{\Sigma}_{\mathbf{X}}^{-1})^{-1}$ for some β satisfying $\beta > 1$ and $(\alpha_1 - \alpha_2)\beta \leq \alpha_1$. Due to the fact that $\boldsymbol{\Sigma}_{\mathbf{N}} \to \frac{1}{\beta-1} \text{diag}(\boldsymbol{\Sigma}_{\mathbf{X}_1}, \boldsymbol{\Sigma}_{\mathbf{X}_2})$ (which is a positive-definite matrix) as $\|\boldsymbol{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}\| \to 0$ and the fact that $\boldsymbol{\Sigma}_{\mathbf{X}}^{-1} + \boldsymbol{\Sigma}_{\mathbf{N}}^{-1} = \beta \text{diag}(\boldsymbol{\Sigma}_{\mathbf{X}_1}^{-1}, \boldsymbol{\Sigma}_{\mathbf{X}_2}^{-1})$ (which is an (m_1, m_2) block diagonal matrix), we have $\boldsymbol{\Sigma}_{\mathbf{N}} \in \mathbb{N}(\boldsymbol{\Sigma}_{\mathbf{X}})$ when $\|\boldsymbol{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}\|$ is sufficiently small. Since $(\mathbf{D}^{-1} + \boldsymbol{\Sigma}_{\mathbf{N}}^{-1})_2 \to \mathbf{D}_2^{-1} + (\beta-1)\boldsymbol{\Sigma}_{\mathbf{X}_2}^{-1}$, $\boldsymbol{\Sigma}_{\hat{\mathbf{N}}_2} \to \frac{1}{\beta}\boldsymbol{\Sigma}_{\mathbf{X}_2}$, and $\boldsymbol{\Sigma}_{\tilde{\mathbf{N}}_2} \to \frac{\beta-1}{\beta^2}\boldsymbol{\Sigma}_{\mathbf{X}_2}$ as $\|\boldsymbol{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}\| \to 0$, it follows that $\boldsymbol{\Omega}_2 \to \alpha_2 \mathbf{D}_2$ as $\|\boldsymbol{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}\| \to 0$. Now it can be readily shown that for the (m_1, m_2) block diagonal matrix $\boldsymbol{\Pi}$ determined by (34) and (35) [or, equivalently, by (36)], we must have $\boldsymbol{\Pi} \to \text{diag}(\alpha_1 \mathbf{D}_1^{-1}, \alpha_2 \mathbf{D}_2^{-1})$ and consequently

$$\boldsymbol{\Pi} - \alpha_1 (\mathbf{D} + \boldsymbol{\Sigma}_{\mathbf{N}})^{-1}$$
$$\to \text{diag}\left(\alpha_1 \mathbf{D}_1^{-1} - \alpha_1(\beta-1)((\beta-1)\mathbf{D}_1 + \boldsymbol{\Sigma}_{\mathbf{X}_1})^{-1},\right.$$
$$\left. \alpha_2 \mathbf{D}_2^{-1} - \alpha_1(\beta-1)((\beta-1)\mathbf{D}_2 + \boldsymbol{\Sigma}_{\mathbf{X}_2})^{-1}\right)$$

as $\|\boldsymbol{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}\| \to 0$. Note that $\alpha_1 \mathbf{D}_1^{-1} - \alpha_1(\beta-1)((\beta-1)\mathbf{D}_1 + \boldsymbol{\Sigma}_{\mathbf{X}_1})^{-1} \succ \mathbf{0}$ and

$$\alpha_2 \mathbf{D}_2^{-1} - \alpha_1(\beta-1)((\beta-1)\mathbf{D}_2 + \boldsymbol{\Sigma}_{\mathbf{X}_2})^{-1}$$
$$\succ \alpha_2 \mathbf{D}_2^{-1} - \alpha_1(\beta-1)(\beta \mathbf{D}_2)^{-1}$$
$$= \frac{\alpha_1 - (\alpha_1 - \alpha_2)\beta}{\beta} \mathbf{D}_2^{-1}$$
$$\succeq \mathbf{0}.$$

Therefore, $\boldsymbol{\Pi} - \alpha_1 (\mathbf{D} + \boldsymbol{\Sigma}_{\mathbf{N}})^{-1}$ is positive definite when $\|\boldsymbol{\Sigma}_{\mathbf{X}_1, \mathbf{X}_2}\|$ is sufficiently small.

REFERENCES

[1] T. Berger, Z. Zhang, and H. Viswanathan, "The CEO problem," *IEEE Trans. Inf. Theory*, vol. 42, no. 3, pp. 887–902, May 1996.
[2] H. Viswanathan and T. Berger, "The quadratic Gaussian CEO problem," *IEEE Trans. Inf. Theory*, vol. 43, no. 5, pp. 1549–1559, Sep. 1997.
[3] Y. Oohama, "The rate-distortion function for the quadratic Gaussian CEO problem," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 1057–1070, May 1998.
[4] Y. Oohama, "Rate-distortion theory for Gaussian multiterminal source coding systems with several side informations at the decoder," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2577–2593, Jul. 2005.
[5] V. Prabhakaran, D. Tse, and K. Ramchandran, "Rate region of the quadratic Gaussian CEO problem," in *Proc. IEEE Symp. Inf. Theory*, Chicago, IL, USA, Jun./Jul. 2004, p. 119.
[6] S. Tavildar, P. Viswanath, and A. B. Wagner, "The Gaussian many-help-one distributed source coding problem," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 564–581, Jan. 2010.
[7] Y. Oohama, Distributed source coding of correlated Gaussian sources [Online]. Available: arXiv:1007.4418v2
[8] A. B. Wagner, S. Tavildar, and P. Viswanath, "Rate region of the quadratic Gaussian two-encoder source-coding problem," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 1938–1961, May 2008.
[9] Y. Oohama, "Gaussian multiterminal source coding," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1912–1923, Nov. 1997.
[10] S. Tavildar and P. Viswanath, "On the sum-rate of the vector Gaussian CEO problem," in *Proc. 39th Asilomar Conf. Signals, Syst. Comput.*, 2005, pp. 3–7.
[11] H. Weingarten and Y. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.
[12] J. Wang, J. Chen, and X. Wu, "On the sum rate of Gaussian multiterminal source coding: New proofs and results," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 3946–3960, Aug. 2010.
[13] Y. Yang, Y. Zhang, and Z. Xiong, "A new sufficient condition for sum-rate tightness in quadratic Gaussian multiterminal source coding," *IEEE Trans. Inf. Theory* vol. 59, no. 1, pp. 408–423, Jan. 2013.
[14] T. Liu and P. Viswanath, "An extremal inequality motivated by multiterminal information-theoretic problems," *IEEE Trans. Inf. Theory*, vol. 53, no. 5, pp. 1839–1851, May 2007.
[15] T. Berger, "Multiterminal source coding," in *The Information Theory Approach to Communications*, ser. CISM Courses and Lectures, G. Longo, Ed. New York, NY, USA: Springer-Verlag, 1978, vol. 229, pp. 171–231.
[16] S.-Y. Tung, "Multiterminal source coding," Ph.D. dissertation, School Electr. Eng., Cornell Univ., Ithaca, NY, USA, 1978.
[17] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA, USA: Athena Scientific, 1999.
[18] S. Liu, "Matrix results on the Khatri-Rao and Tracy-Singh products," *Linear Algebra Appl.*, vol. 289, pp. 267–277, Mar. 1999.
[19] J. Wang and J. Chen, "On the vector Gaussian L-terminal CEO problem," in *Proc. IEEE Symp. Inf. Theory*, Cambridge, MA, USA, Jul. 1–6, 2012, pp. 571–575.
[20] Md. S. Rahman and A. B. Wagner, "Rate region of the Gaussian scalar-help-vector source-coding problem," *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 172–188, Jan. 2012.
[21] Md. S. Rahman and A. B. Wagner, Rate region of the vector Gaussian one-helper source-coding problem [Online]. Available: arXiv:1112.6367v1

Jia Wang received the B.Sc. degree in electronic engineering, the M.S. degree in pattern recognition and intelligence control, and the Ph.D. degree in electronic engineering from Shanghai Jiao Tong University, China, in 1997, 1999, and 2002, respectively.

He is currently an Associate Professor of the Institute of Image Communication and Information Processing, Shanghai Jiao Tong University, and a member of Shanghai Key Laboratory of Digital Media Processing and Transmission. His research interests include multiuser information theory and its application in video coding.

Jun Chen (S'03–M'06) received the B.E. degree with honors in communication engineering from Shanghai Jiao Tong University, Shanghai, China, in 2001 and the M.S. and Ph.D. degrees in electrical and computer engineering from Cornell University, Ithaca, NY, in 2004 and 2006, respectively.

He was a Postdoctoral Research Associate in the Coordinated Science Laboratory at the University of Illinois at Urbana-Champaign, Urbana, IL, from 2005 to 2006, and a Postdoctoral Fellow at the IBM Thomas J. Watson Research Center, Yorktown Heights, NY, from 2006 to 2007. He is currently an Assistant Professor of Electrical and Computer Engineering at McMaster University, Hamilton, ON, Canada. He holds the Barber-Gennum Chair in Information Technology. His research interests include information theory, wireless communications, and signal processing.

He received several awards for his research, including the Josef Raviv Memorial Postdoctoral Fellowship in 2006, the Early Researcher Award from the Province of Ontario in 2010, and the IBM Faculty Award in 2010.