

Vector Gaussian Multiterminal Source Coding

Jia Wang and Jun Chen, *Member, IEEE*

Abstract—We derive an outer bound of the rate region of the vector Gaussian L -terminal CEO problem by establishing a lower bound on each supporting hyperplane of the rate region. To this end, we prove a new extremal inequality by exploiting the connection between differential entropy and Fisher information as well as some fundamental estimation-theoretic inequalities. It is shown that the outer bound matches the Berger–Tung inner bound in the high-resolution regime. We then derive a lower bound on each supporting hyperplane of the rate region of the direct vector Gaussian L -terminal source coding problem by coupling it with the CEO problem through a limiting argument. The tightness of this lower bound in the high-resolution regime and the weak-dependence regime is also proved.

Index Terms—Borsuk’s theorem, CEO problem, extremal inequality, Fisher information, MMSE, multiterminal source coding.

I. INTRODUCTION

THE CEO problem, also known as the indirect multiterminal source coding problem, was introduced by Berger, Zhang, and Viswanathan in [1]. A series of papers were devoted to the scalar Gaussian version of the problem [2]–[4], culminating in a complete characterization of the rate region by Prabhakaran *et al.* [5] and Oohama [6] (see also [7] for the solution to a generalized version of this problem). However, the proof technique in [6] and [5] is not completely suitable for the vector Gaussian case due to its heavy reliance on the entropy power inequality. The reason is that the proportionality condition (on the relevant covariance matrices) for the tightness of the entropy power inequality is not necessarily satisfied in the vector source setting, though this technical issue can sometimes be resolved by invoking the entropy power inequality in conjunction with an enhancement argument [8] (see [9] for bounding the sum rate of the vector Gaussian CEO problem via this approach).

Manuscript received June 4, 2012; revised November 3, 2013; accepted May 17, 2014. Date of publication June 26, 2014; date of current version August 14, 2014. J. Wang was supported in part by the National Natural Science Foundation of China under Grant 61271221 and Grant 61221001, in part by the National Key Technology Research and Development Program of China under Grant 2013BAH53F04, and in part by the 973 Program under Grant 2010CB731401 and Grant 2010CB731406. J. Chen was supported in part by the Early Researcher Award through the Province of Ontario and in part by the Natural Science and Engineering Research Council of Canada under a Discovery Grant. This paper was presented at the 2012 IEEE International Symposium on Information Theory and 2012 IEEE Information Theory Workshop.

J. Wang is with the Department of Electronic Engineering, Institute of Image Communication and Network Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: jiawang@sjtu.edu.cn).

J. Chen is with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON L8S 4K1, Canada (e-mail: junchen@ece.mcmaster.ca).

Communicated by Y. Oohama, Associate Editor for Source Coding.
Digital Object Identifier 10.1109/TIT.2014.2333473

A new approach to the Gaussian CEO problem was proposed in [10], which is directly applicable to the vector case due to its estimation-theoretic nature. However, this approach is only effective for bounding the sum rate part of the rate region. In order to treat a general supporting hyperplane of the rate region, one has to resolve the tension among various information-theoretic quantities, which calls for a certain extremal inequality. It turns out that in the two-terminal case [11] the desired extremal inequality is essentially a strengthened version of the Liu-Viswanath inequality [12]. Unfortunately, to handle the general L -terminal case, one has to deal with a certain long Markov chain structure and the extremal inequality needed for that purpose is significantly different from the Liu-Viswanath inequality and its variants (see [13] and [14]). Among other contributions of this paper, we prove such an extremal inequality for the general vector Gaussian L -terminal CEO problem, which yields a lower bound on each supporting hyperplane of the rate region. It is worth mentioning that this lower bound was also derived in a recent independent work by Ekrem and Ulukus [15] through a different method.

It is now well understood that indirect Gaussian multiterminal source coding is intimately connected with its direct coding counterpart. Notably, this connection plays an important role in the breakthrough result by Wagner, Tavildar, and Viswanath [16] on the sum rate of the direct scalar Gaussian two-terminal source coding problem, which, together with Oohama’s earlier work [17], provides a complete characterization of the rate region of this problem. In fact, recent rapid progress in direct Gaussian multiterminal source coding [10], [18], [19] is achieved, to a large extent, through the exploitation of this connection. Following this line of research in this work we establish a lower bound on each supporting hyperplane of the rate region of the direct vector Gaussian L -terminal source coding problem by viewing it as a limiting version of the vector Gaussian CEO problem.

It will be seen that our lower bounds for the vector Gaussian L -terminal CEO problem and the direct vector Gaussian L -terminal source coding problem closely resemble their corresponding Berger-Tung upper bounds. However, the task of deriving a matching condition turns out to be quite intricate. Even in the two-terminal case [11], it requires a significant amount of non-trivial matrix calculation. Furthermore, such kind of brute-force calculation appears to be formidable, if not impossible, in the general L -terminal case. In view of this difficulty, we take a more conceptual approach in the present paper. In particular, our proof of the tightness of the lower bound for the direct vector Gaussian L -terminal source coding problem is based on, among other things, Borsuk’s theorem in nonlinear analysis.

The remainder of this paper is organized as follows. In Section II we derive an outer bound of the rate region of the vector Gaussian L -terminal CEO problem by establishing a lower bound on each supporting hyperplane of the rate region; furthermore, it is shown that the outer bound coincides with the Berger-Tung inner bound in the high-resolution regime. We then, in Section III, derive a lower bound on each supporting hyperplane of the rate region of the direct vector Gaussian L -terminal source coding problem by coupling it with the CEO problem through a limiting argument; the tightness of this lower bound in the high-resolution regime and the weak-dependence regime is also proved. A new extremal inequality, which is the main technical ingredient in the derivation of the lower bound for the vector Gaussian CEO problem, is given in Section IV; the proof of this extremal inequality relies on the connection between differential entropy and Fisher information as well as some fundamental estimation-theoretic inequalities. Section V concludes the paper.

Throughout this paper, for any random object W and $m \times n$ random matrix \mathbf{S} we define $\Sigma_{\mathbf{S}} = \frac{1}{n} \mathbb{E}[\mathbf{S}\mathbf{S}^T]$ and $\Sigma_{\mathbf{S}|W} = \Sigma_{\mathbf{S}} - \mathbb{E}[\mathbf{S}|W]$; we assume the logarithm function is to base e and define $\log^+ x = \max(\log x, 0)$; we denote the $m \times m$ identity matrix by \mathbf{I}_m .

II. THE VECTOR GAUSSIAN L -TERMINAL CEO PROBLEM

A. Problem Definition

Let $\mathbf{X}_0, \mathbf{N}_1, \dots, \mathbf{N}_L$ be mutually independent $m \times 1$ Gaussian random vectors with mean zero and positive definite covariance matrices $\Sigma_{\mathbf{X}_0}, \Sigma_{\mathbf{N}_1}, \dots, \Sigma_{\mathbf{N}_L}$, respectively. Let $\mathbf{X}_i = \mathbf{X}_0 + \mathbf{N}_i$, $i = 1, \dots, L$. We refer to \mathbf{X}_0 as the remote source and \mathbf{X}_i as the noisy observation at Encoder i , $i = 1, \dots, L$.

Let $\{(\mathbf{X}_0(t), \mathbf{N}_1(t), \dots, \mathbf{N}_L(t), \mathbf{X}_1(t), \dots, \mathbf{X}_L(t))\}_{t=1}^{\infty}$ be i.i.d. copies of $(\mathbf{X}_0, \mathbf{N}_1, \dots, \mathbf{N}_L, \mathbf{X}_1, \dots, \mathbf{X}_L)$. The rate region of the vector Gaussian L -terminal CEO problem is defined as follows.

Definition 1: A rate vector (R_1, \dots, R_L) is said to be achievable subject to distortion constraint \mathbf{D} if for all sufficiently large n , there exist encoding functions $f_i^{(n)}: \mathbb{R}^{m \times n} \rightarrow \{1, \dots, K_i\}$, $i = 1, \dots, L$, and decoding function $g^{(n)}: \prod_{i=1}^L \{1, \dots, K_i\} \rightarrow \mathbb{R}^{m \times n}$ such that

$$\begin{aligned} \frac{1}{n} \log K_i &\leq R_i, \quad i = 1, \dots, L, \\ \Sigma_{\hat{\mathbf{X}}_0^n - \hat{\mathbf{X}}_0^n} &\leq \mathbf{D}, \end{aligned} \quad (1)$$

where $\hat{\mathbf{X}}_0^n = g^{(n)}(W_1, \dots, W_L)$ with $W_i = f_i^{(n)}(\mathbf{X}_i^n)$, $i = 1, \dots, L$. The rate region $\mathcal{R}(\mathbf{D})$ is the closure of the set of all achievable rate vectors subject to distortion constraint \mathbf{D} .

Remark: Clearly, there is no loss of optimality in assuming $g^{(n)}(W_1, \dots, W_L) = \mathbb{E}[\mathbf{X}_0^n | W_1, \dots, W_L]$. Therefore, $\mathcal{R}(\mathbf{D})$ is not affected if we replace (1) with $\Sigma_{\mathbf{X}_0^n | W_1, \dots, W_L} \leq \mathbf{D}$.

Since $\mathcal{R}(\mathbf{D})$ is a (closed) convex set, it suffices to characterize its supporting hyperplanes, i.e., to solve the following

optimization problem

$$R(\mathbf{D}, \alpha) \triangleq \inf_{(R_1, \dots, R_L) \in \mathcal{R}(\mathbf{D})} \sum_{i=1}^L \alpha_i R_i,$$

where $\alpha = (\alpha_1, \dots, \alpha_L)$ with $\alpha_i \geq 0$, $i = 1, \dots, L$. Note that one can obtain an outer bound (inner bound) of $\mathcal{R}(\mathbf{D})$ by lower-bounding (upper-bounding) $R(\mathbf{D}, \alpha)$. Without loss of generality, we shall assume $\mathbf{D} \succ \Sigma_{\mathbf{N}}$ and $\alpha_1 \geq \dots \geq \alpha_L \geq 0$, where $\Sigma_{\mathbf{N}} = \Sigma_{\mathbf{X}_0 | \mathbf{X}_1, \dots, \mathbf{X}_L}$. It can be verified that

$$\Sigma_{\mathbf{N}} = \left(\Sigma_{\mathbf{X}_0}^{-1} + \sum_{i=1}^L \Sigma_{\mathbf{N}_i}^{-1} \right)^{-1}.$$

B. Lower and Upper Bounds

The following theorem, which is a generalization of [11, Th. 2] from the two-terminal case to the L -terminal case, provides a lower bound on $R(\mathbf{D}, \alpha)$. The proof is relegated to Appendix A. This result was also derived independently by Ekrem and Ulukus [15] through a different method. The main ingredient of our proof is a new extremal inequality (see Theorem 11 in Section IV), which is of interest in its own right.

Theorem 1: $R(\mathbf{D}, \alpha) \geq \underline{R}(\mathbf{D}, \alpha)$ for α with $\alpha_1 \geq \dots \geq \alpha_L > 0$, where

$$\begin{aligned} \underline{R}(\mathbf{D}, \alpha) &\triangleq \min_{\mathbf{D}_0, \dots, \mathbf{D}_L} \frac{\alpha_L}{2} \log \frac{|\Sigma_{\mathbf{X}_0}|}{|\mathbf{D}_0|} + \sum_{i=1}^L \frac{\alpha_i}{2} \log \frac{|\Sigma_{\mathbf{N}_i}|}{|\mathbf{D}_i|} + \sum_{i=2}^L \frac{\alpha_{i-1} - \alpha_i}{2} \\ &\cdot \log^+ \left(\frac{1}{|\mathbf{D}_0|} \cdot \frac{1}{|\Sigma_{\mathbf{X}_0}^{-1} + \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} - \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} \mathbf{D}_j \Sigma_{\mathbf{N}_j}^{-1}|} \right) \\ &\text{subject to } \Sigma_{\mathbf{N}} \leq \mathbf{D}_0 \leq \Sigma_{\mathbf{X}_0}, \\ &\mathbf{D}_0 \leq \mathbf{D}, \\ &\mathbf{0} \leq \mathbf{D}_i \leq \Sigma_{\mathbf{N}_i}, \quad i = 1, \dots, L, \end{aligned}$$

$$\sum_{i=1}^L \Sigma_{\mathbf{N}_i}^{-1} \mathbf{D}_i \Sigma_{\mathbf{N}_i}^{-1} \leq \Sigma_{\mathbf{N}}^{-1} - \mathbf{D}_0^{-1}.$$

Remark: It is easy to see that if $\mathbf{D} \leq \Sigma_{\mathbf{X}_0}$, then there is no loss of optimality in choosing $\mathbf{D}_0 = \mathbf{D}$ and consequently $\underline{R}(\mathbf{D}, \alpha)$ is given by the minimum value of the following optimization problem

$$\begin{aligned} (\mathbf{P}) \min_{\mathbf{D}_1, \dots, \mathbf{D}_L} &\frac{\alpha_L}{2} \log \frac{|\Sigma_{\mathbf{X}_0}|}{|\mathbf{D}|} + \sum_{i=1}^L \frac{\alpha_i}{2} \log \frac{|\Sigma_{\mathbf{N}_i}|}{|\mathbf{D}_i|} + \sum_{i=2}^L \frac{\alpha_{i-1} - \alpha_i}{2} \\ &\cdot \log^+ \left(\frac{1}{|\mathbf{D}|} \cdot \frac{1}{|\Sigma_{\mathbf{X}_0}^{-1} + \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} - \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} \mathbf{D}_j \Sigma_{\mathbf{N}_j}^{-1}|} \right) \\ &\text{subject to } \mathbf{0} \leq \mathbf{D}_i \leq \Sigma_{\mathbf{N}_i}, \quad i = 1, \dots, L, \\ &\sum_{i=1}^L \Sigma_{\mathbf{N}_i}^{-1} \mathbf{D}_i \Sigma_{\mathbf{N}_i}^{-1} \leq \Sigma_{\mathbf{N}}^{-1} - \mathbf{D}^{-1}. \end{aligned}$$

Corollary 1: If $\alpha_1 \geq \dots \geq \alpha_{L'} > \alpha_{L'+1} = \dots = \alpha_L = 0$ for some $L' < L$, then $R(\mathbf{D}, \alpha) \geq \underline{R}'(\mathbf{D}, \alpha)$, where

$$\begin{aligned} \underline{R}'(\mathbf{D}, \alpha) &\triangleq \min_{\mathbf{D}_0, \dots, \mathbf{D}_{L'}} \frac{\alpha_{L'}}{2} \log \frac{1}{|\mathbf{D}_0| |\Sigma_{\mathbf{X}_0}^{-1} + \sum_{j=L'+1}^L \Sigma_{\mathbf{N}_j}^{-1}|} \\ &+ \sum_{i=1}^{L'} \frac{\alpha_i}{2} \log \frac{|\Sigma_{\mathbf{N}_i}|}{|\mathbf{D}_i|} + \sum_{i=2}^{L'} \frac{\alpha_{i-1} - \alpha_i}{2} \\ &\cdot \log^+ \left(\frac{1}{|\mathbf{D}_0|} \cdot \frac{1}{|\Sigma_{\mathbf{X}_0}^{-1} + \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} - \sum_{j=i}^{L'} \Sigma_{\mathbf{N}_j}^{-1} \mathbf{D}_j \Sigma_{\mathbf{N}_j}^{-1}|} \right) \\ \text{subject to } \Sigma_{\mathbf{N}} &\leq \mathbf{D}_0 \leq \left(\Sigma_{\mathbf{X}_0}^{-1} + \sum_{j=L'+1}^L \Sigma_{\mathbf{N}_j}^{-1} \right)^{-1}, \\ \mathbf{D}_0 &\leq \mathbf{D}, \\ \mathbf{0} &\leq \mathbf{D}_i \leq \Sigma_{\mathbf{N}_i}, \quad i = 1, \dots, L', \\ \sum_{i=1}^{L'} \Sigma_{\mathbf{N}_i}^{-1} \mathbf{D}_i \Sigma_{\mathbf{N}_i}^{-1} &\leq \Sigma_{\mathbf{N}}^{-1} - \mathbf{D}_0^{-1}. \end{aligned}$$

Proof: Since $\alpha_{L'+1} = \dots = \alpha_L = 0$, only the rates of the first L' encoders are relevant. To derive a lower bound on $R(\mathbf{D}, \alpha)$, we provide $(\mathbf{X}_{L'+1}, \dots, \mathbf{X}_L)$ directly to the first L' encoders and the decoder. It can be shown that this problem is equivalent to an L' -terminal CEO problem with $\mathbf{X}'_0 = \mathbf{X}_0 - \mathbb{E}[\mathbf{X}_0 | \mathbf{X}_{L'+1}, \dots, \mathbf{X}_L]$ as the remote source and $\mathbf{X}'_i = \mathbf{X}_i + \mathbf{N}_i$, $i = 1, \dots, L'$, as noisy observations. Note that

$$\Sigma_{\mathbf{X}'_0} = \left(\Sigma_{\mathbf{X}_0}^{-1} + \sum_{j=L'+1}^L \Sigma_{\mathbf{N}_j}^{-1} \right)^{-1}.$$

Now one can readily deduce Corollary 1 from Theorem 1. ■

Remark: Note that if we evaluate $\underline{R}(\mathbf{D}, \alpha)$ for the case where $\alpha_1 \geq \dots \geq \alpha_{L'} > \alpha_{L'+1} = \dots = \alpha_L = 0$ for some $L' < L$, then there is no loss of optimality in choosing $\mathbf{D}_{L'+1} = \dots = \mathbf{D}_L = \mathbf{0}$, which gives

$$\begin{aligned} \underline{R}(\mathbf{D}, \alpha) &= \min_{\mathbf{D}_0, \dots, \mathbf{D}_{L'}} \frac{\alpha_{L'}}{2} \log \frac{1}{|\mathbf{D}_0| |\Sigma_{\mathbf{X}_0}^{-1} + \sum_{j=L'+1}^L \Sigma_{\mathbf{N}_j}^{-1}|} \\ &+ \sum_{i=1}^{L'} \frac{\alpha_i}{2} \log \frac{|\Sigma_{\mathbf{N}_i}|}{|\mathbf{D}_i|} + \sum_{i=2}^{L'} \frac{\alpha_{i-1} - \alpha_i}{2} \\ &\cdot \log^+ \left(\frac{1}{|\mathbf{D}_0|} \cdot \frac{1}{|\Sigma_{\mathbf{X}_0}^{-1} + \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} - \sum_{j=i}^{L'} \Sigma_{\mathbf{N}_j}^{-1} \mathbf{D}_j \Sigma_{\mathbf{N}_j}^{-1}|} \right) \\ \text{subject to } \Sigma_{\mathbf{N}} &\leq \mathbf{D}_0 \leq \Sigma_{\mathbf{X}_0}, \\ \mathbf{D}_0 &\leq \mathbf{D}, \\ \mathbf{0} &\leq \mathbf{D}_i \leq \Sigma_{\mathbf{N}_i}, \quad i = 1, \dots, L', \\ \sum_{i=1}^{L'} \Sigma_{\mathbf{N}_i}^{-1} \mathbf{D}_i \Sigma_{\mathbf{N}_i}^{-1} &\leq \Sigma_{\mathbf{N}}^{-1} - \mathbf{D}_0^{-1}. \end{aligned}$$

It is easy to see that $\underline{R}(\mathbf{D}, \alpha) \leq \underline{R}'(\mathbf{D}, \alpha)$. Therefore, Theorem 1 in fact holds for all α with $\alpha_1 \geq \dots \geq \alpha_L \geq 0$.

One can obtain the following upper bound on $R(\mathbf{D}, \alpha)$, referred to as the Berger-Tung upper bound on $R(\mathbf{D}, \alpha)$, by evaluating the standard Berger-Tung inner bound of

$\mathcal{R}(\mathbf{D})$ [21], [22]. This result is a straightforward generalization of [11, Th. 3] from the two-terminal case to the L -terminal case. The proof is omitted.

Theorem 2: $R(\mathbf{D}, \alpha) \leq \overline{R}(\mathbf{D}, \alpha)$, where

$$\begin{aligned} \overline{R}(\mathbf{D}, \alpha) &\triangleq \min_{\mathbf{D}_0, \dots, \mathbf{D}_L} \frac{\alpha_L}{2} \log \frac{|\Sigma_{\mathbf{X}_0}|}{|\mathbf{D}_0|} + \sum_{i=1}^L \frac{\alpha_i}{2} \log \frac{|\Sigma_{\mathbf{N}_i}|}{|\mathbf{D}_i|} + \sum_{i=2}^L \frac{\alpha_{i-1} - \alpha_i}{2} \\ &\cdot \log^+ \left(\frac{1}{|\mathbf{D}_0|} \cdot \frac{1}{|\Sigma_{\mathbf{X}_0}^{-1} + \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} - \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} \mathbf{D}_j \Sigma_{\mathbf{N}_j}^{-1}|} \right) \\ \text{subject to } \Sigma_{\mathbf{N}} &\leq \mathbf{D}_0 \leq \Sigma_{\mathbf{X}_0}, \\ \mathbf{D}_0 &\leq \mathbf{D}, \\ \mathbf{0} &\leq \mathbf{D}_i \leq \Sigma_{\mathbf{N}_i}, \quad i = 1, \dots, L, \\ \sum_{i=1}^L \Sigma_{\mathbf{N}_i}^{-1} \mathbf{D}_i \Sigma_{\mathbf{N}_i}^{-1} &= \Sigma_{\mathbf{N}}^{-1} - \mathbf{D}_0^{-1}. \end{aligned}$$

Remark: It is easy to verify that

$$\begin{aligned} \overline{R}(\mathbf{D}, \alpha) &= \min_{\mathbf{D}_1, \dots, \mathbf{D}_L} \frac{\alpha_L}{2} \log \left(|\Sigma_{\mathbf{X}_0}| \left| \Sigma_{\mathbf{N}}^{-1} - \sum_{j=1}^L \Sigma_{\mathbf{N}_j}^{-1} \mathbf{D}_j \Sigma_{\mathbf{N}_j}^{-1} \right| \right) \\ &+ \sum_{i=1}^L \frac{\alpha_i}{2} \log \frac{|\Sigma_{\mathbf{N}_i}|}{|\mathbf{D}_i|} + \sum_{i=2}^L \frac{\alpha_{i-1} - \alpha_i}{2} \\ &\cdot \log \frac{|\Sigma_{\mathbf{N}}^{-1} - \sum_{j=1}^L \Sigma_{\mathbf{N}_j}^{-1} \mathbf{D}_j \Sigma_{\mathbf{N}_j}^{-1}|}{|\Sigma_{\mathbf{X}_0}^{-1} + \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} - \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} \mathbf{D}_j \Sigma_{\mathbf{N}_j}^{-1}|} \\ \text{subject to } \mathbf{0} &\leq \mathbf{D}_i \leq \Sigma_{\mathbf{N}_i}, \quad i = 1, \dots, L, \\ \sum_{i=1}^L \Sigma_{\mathbf{N}_i}^{-1} \mathbf{D}_i \Sigma_{\mathbf{N}_i}^{-1} &\leq \Sigma_{\mathbf{N}}^{-1} - \mathbf{D}^{-1}. \end{aligned}$$

C. Matching Conditions

It is easy to show that $\underline{R}(\mathbf{D}, \alpha) = \overline{R}(\mathbf{D}, \alpha)$ in the scalar case (i.e., $m = 1$), which recovers the well-known result in [5] and [6]. Moreover, it is clear that $\underline{R}(\mathbf{D}, \alpha) = \overline{R}(\mathbf{D}, \alpha)$ if there exists an optimal solution $(\hat{\mathbf{D}}_0, \dots, \hat{\mathbf{D}}_L)$ to the minimization problem associated with $\underline{R}(\mathbf{D}, \alpha)$ such that

$$\sum_{i=1}^L \Sigma_{\mathbf{N}_i}^{-1} \hat{\mathbf{D}}_i \Sigma_{\mathbf{N}_i}^{-1} = \Sigma_{\mathbf{N}}^{-1} - \hat{\mathbf{D}}_0^{-1}.$$

The following result provides a simple and explicit matching condition for the vector case.

Theorem 3: $\underline{R}(\mathbf{D}, \alpha) = \overline{R}(\mathbf{D}, \alpha)$ if $\mathbf{D} \leq (\Sigma_{\mathbf{N}}^{-1} - \Sigma_{\mathbf{N}_1}^{-1})^{-1}$.

Proof: Note that $\mathbf{D} \leq (\Sigma_{\mathbf{N}}^{-1} - \Sigma_{\mathbf{N}_1}^{-1})^{-1}$ implies $\mathbf{D} \leq \Sigma_{\mathbf{X}_0}$. It is clear that $\underline{R}(\mathbf{D}, \alpha) = \overline{R}(\mathbf{D}, \alpha)$ if there exists an optimal solution $(\hat{\mathbf{D}}_1, \dots, \hat{\mathbf{D}}_L)$ to (\mathbf{P}) (see the remark after Theorem 1) such that

$$\sum_{i=1}^L \Sigma_{\mathbf{N}_i}^{-1} \hat{\mathbf{D}}_i \Sigma_{\mathbf{N}_i}^{-1} = \Sigma_{\mathbf{N}}^{-1} - \mathbf{D}^{-1}. \quad (2)$$

Let $(\mathbf{D}_1^*, \dots, \mathbf{D}_L^*)$ be an optimal solution to (\mathbf{P}) . Define

$$\hat{\mathbf{D}}_1 = \Sigma_{\mathbf{N}_1} \left(\Sigma_{\mathbf{N}}^{-1} - \mathbf{D}^{-1} - \sum_{i=2}^L \Sigma_{\mathbf{N}_i}^{-1} \mathbf{D}_i^* \Sigma_{\mathbf{N}_i}^{-1} \right) \Sigma_{\mathbf{N}_1}$$

and $\hat{\mathbf{D}}_i = \mathbf{D}_i^*$, $i = 2, \dots, L$. Note that (2) is satisfied by this constructed $(\hat{\mathbf{D}}_1, \dots, \hat{\mathbf{D}}_L)$. Moreover, we have $\mathbf{D}_1^* \preceq \hat{\mathbf{D}}_1$ and

$$\hat{\mathbf{D}}_1 \preceq \Sigma_{\mathbf{N}_1}(\Sigma_{\mathbf{N}}^{-1} - \mathbf{D}^{-1})\Sigma_{\mathbf{N}_1} \preceq \Sigma_{\mathbf{N}_1}, \quad (3)$$

where (3) is due to $\mathbf{D} \preceq (\Sigma_{\mathbf{N}}^{-1} - \Sigma_{\mathbf{N}_1}^{-1})^{-1}$. One can readily see that $(\hat{\mathbf{D}}_1, \dots, \hat{\mathbf{D}}_L)$ must also be an optimal solution to (\mathbf{P}) . This completes the proof. ■

Remark: Theorem 3 implies that the outer bound induced by $\underline{R}(\mathbf{D}, \alpha)$ coincides with the Berger-Tung inner bound if \mathbf{D} satisfies the following condition

$$\mathbf{D} \preceq (\Sigma_{\mathbf{N}}^{-1} - \Sigma_{\mathbf{N}_i}^{-1})^{-1}, \quad i = 1, \dots, L,$$

yielding a complete characterization of $\mathcal{R}(\mathbf{D})$ in the high-resolution regime.

A stronger matching condition can be found for the special case $\alpha = \mathbf{1} \triangleq (1, \dots, 1)$, which corresponds to the sum rate.

Theorem 4: $\underline{R}(\mathbf{D}, \mathbf{1}) = \overline{R}(\mathbf{D}, \mathbf{1})$ if $\Sigma_{\mathbf{N}}^{-1} - L\Sigma_{\mathbf{N}_i}^{-1} \preceq \mathbf{D}^{-1}$, $i = 1, \dots, L$.

Proof: It is clear that $\mathbf{D} \preceq \Sigma_{\mathbf{X}_0}$ if $\Sigma_{\mathbf{N}}^{-1} - L\Sigma_{\mathbf{N}_i}^{-1} \preceq \mathbf{D}^{-1}$, $i = 1, \dots, L$. Therefore, in light of the remark after Theorem 1, $\underline{R}(\mathbf{D}, \mathbf{1})$ is given by the minimum value of (\mathbf{P}) with $\alpha = \mathbf{1}$. Note that in this case (\mathbf{P}) is a convex semidefinite programming problem and its Lagrangian is given by

$$\begin{aligned} L(\mathbf{D}_1, \dots, \mathbf{D}_L) &= \frac{1}{2} \log \frac{|\Sigma_{\mathbf{X}_0}|}{|\mathbf{D}|} + \sum_{i=1}^L \frac{1}{2} \log \frac{|\Sigma_{\mathbf{N}_i}|}{|\mathbf{D}_i|} + \sum_{i=1}^L \text{tr}(\Pi_i(\mathbf{D}_i - \Sigma_{\mathbf{N}_i})) \\ &\quad + \text{tr} \left(\Lambda \left(\sum_{i=1}^L \Sigma_{\mathbf{N}_i}^{-1} \mathbf{D}_i \Sigma_{\mathbf{N}_i}^{-1} - \Sigma_{\mathbf{N}}^{-1} + \mathbf{D}^{-1} \right) \right), \end{aligned}$$

where $\Pi_1, \dots, \Pi_L, \Lambda$ are positive semidefinite matrices. Note that $(\hat{\mathbf{D}}_1, \dots, \hat{\mathbf{D}}_L)$ is an optimal solution to (\mathbf{P}) if the following KKT conditions are satisfied

$$\begin{aligned} \nabla_{\mathbf{D}_i} L(\mathbf{D}_1, \dots, \mathbf{D}_L) \Big|_{\mathbf{D}_i = \hat{\mathbf{D}}_i} &= \mathbf{0}, \quad i = 1, \dots, L, \\ \mathbf{0} \preceq \hat{\mathbf{D}}_i \preceq \Sigma_{\mathbf{N}_i}, \quad &i = 1, \dots, L, \end{aligned}$$

$$\sum_{i=1}^L \Sigma_{\mathbf{N}_i}^{-1} \hat{\mathbf{D}}_i \Sigma_{\mathbf{N}_i}^{-1} \preceq \Sigma_{\mathbf{N}}^{-1} - \mathbf{D}^{-1},$$

$$\Pi_i(\hat{\mathbf{D}}_i - \Sigma_{\mathbf{N}_i}) = \mathbf{0}, \quad i = 1, \dots, L,$$

$$\Lambda \left(\sum_{i=1}^L \Sigma_{\mathbf{N}_i}^{-1} \hat{\mathbf{D}}_i \Sigma_{\mathbf{N}_i}^{-1} - \Sigma_{\mathbf{N}}^{-1} + \mathbf{D}^{-1} \right) = \mathbf{0}.$$

It can be verified that

$$\begin{aligned} \nabla_{\mathbf{D}_i} L(\mathbf{D}_1, \dots, \mathbf{D}_L) \Big|_{\mathbf{D}_i = \hat{\mathbf{D}}_i} &= -\frac{1}{2} \hat{\mathbf{D}}_i^{-1} + \Pi_i + \Sigma_{\mathbf{N}_i}^{-1} \Lambda \Sigma_{\mathbf{N}_i}^{-1}, \\ &i = 1, \dots, L. \end{aligned}$$

Now we choose

$$\begin{aligned} \hat{\mathbf{D}}_i &= \frac{1}{L} \Sigma_{\mathbf{N}_i} (\Sigma_{\mathbf{N}}^{-1} - \mathbf{D}^{-1}) \Sigma_{\mathbf{N}_i}, \quad i = 1, \dots, L, \\ \Pi_i &= \mathbf{0}, \quad i = 1, \dots, L, \\ \Lambda &= \frac{L}{2} (\Sigma_{\mathbf{N}}^{-1} - \mathbf{D}^{-1})^{-1}. \end{aligned}$$

It is easy to see that with such a choice the KKT conditions are indeed satisfied. In particular, we have

$$\hat{\mathbf{D}}_i \preceq \Sigma_{\mathbf{N}_i}, \quad i = 1, \dots, L, \quad (4)$$

$$\sum_{i=1}^L \Sigma_{\mathbf{N}_i}^{-1} \hat{\mathbf{D}}_i \Sigma_{\mathbf{N}_i}^{-1} = \Sigma_{\mathbf{N}}^{-1} - \mathbf{D}^{-1}, \quad (5)$$

where (4) is due to $\Sigma_{\mathbf{N}}^{-1} - L\Sigma_{\mathbf{N}_i}^{-1} \preceq \mathbf{D}^{-1}$, $i = 1, \dots, L$. Note that (5) implies $\underline{R}(\mathbf{D}, \mathbf{1}) = \overline{R}(\mathbf{D}, \mathbf{1})$, which completes the proof. ■

III. THE DIRECT VECTOR GAUSSIAN L -TERMINAL SOURCE CODING PROBLEM

A. Problem Definition

Let m_1, \dots, m_L , and m be $L + 1$ positive integers with $m = \sum_{i=1}^L m_i$. Let $n_0 = 0$ and $n_i = \sum_{j=1}^i m_j$, $i = 1, \dots, L$. For any $n_{L'} \times n_{L'}$ matrix $\mathbf{A} = (a_{j,k})$ with $1 \leq L' \leq L$, we denote its i th diagonal submatrix $(a_{j,k})_{n_{i-1}+1 \leq j, k \leq n_i}$ by $(\mathbf{A})_i$, $i = 1, \dots, L'$, and call \mathbf{A} a block diagonal matrix if

$$\begin{aligned} \mathbf{A} &= \text{diag}((\mathbf{A})_1, \dots, (\mathbf{A})_{L'}) \\ &\triangleq \begin{pmatrix} (\mathbf{A})_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & (\mathbf{A})_{L'} \end{pmatrix}. \end{aligned}$$

Let $\mathbf{X} = (\mathbf{X}_1^T, \dots, \mathbf{X}_L^T)^T$ be an $m \times 1$ Gaussian random vector with mean zero and positive definite covariance matrix $\Sigma_{\mathbf{X}}$, where \mathbf{X}_i is an $m_i \times 1$ Gaussian random vector with mean zero and positive definite covariance matrix $\Sigma_{\mathbf{X}_i} = (\Sigma_{\mathbf{X}})_i$, $i = 1, \dots, L$. Let $\{(\mathbf{X}(t), \mathbf{X}_1(t), \dots, \mathbf{X}_L(t))\}_{t=1}^{\infty}$ be i.i.d. copies of $(\mathbf{X}, \mathbf{X}_1, \dots, \mathbf{X}_L)$. The rate region of the direct vector Gaussian L -terminal source coding problem is defined as follows.

Definition 2: A rate vector (R_1, \dots, R_L) is said to be achievable subject to distortion constraint $(\mathbf{D}_1, \dots, \mathbf{D}_L)$ if for all sufficiently large n , there exist encoding functions $f_i^{(n)}: \mathbb{R}^{m_i \times n} \rightarrow \{1, \dots, K_i\}$, $i = 1, \dots, L$, such that

$$\begin{aligned} \frac{1}{n} \log K_i &\leq R_i, \quad i = 1, \dots, L, \\ \Sigma_{\mathbf{X}_i^n | W_1, \dots, W_L} &\preceq \mathbf{D}_i, \quad i = 1, \dots, L, \end{aligned}$$

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

Since $\mathcal{R}(\mathbf{D}_1, \dots, \mathbf{D}_L)$ is a (closed) convex set, it suffices to characterize its supporting hyperplanes, i.e., to solve the following optimization problem

$$R(\mathbf{D}_1, \dots, \mathbf{D}_L, \alpha) \triangleq \inf_{(R_1, \dots, R_L) \in \mathcal{R}(\mathbf{D}_1, \dots, \mathbf{D}_L)} \sum_{i=1}^L \alpha_i R_i,$$

where $\alpha = (\alpha_1, \dots, \alpha_L)$ with $\alpha_i \geq 0$, $i = 1, \dots, L$. Note that one can obtain an outer bound (inner bound) of $\mathcal{R}(\mathbf{D}_1, \dots, \mathbf{D}_L)$ by lower-bounding (upper-bounding) $R(\mathbf{D}_1, \dots, \mathbf{D}_L, \alpha)$. Without loss of generality, we shall assume $\mathbf{D}_i \succ \mathbf{0}$, $i = 1, \dots, L$, and $\alpha_1 \geq \dots \geq \alpha_L \geq 0$.

B. Lower and Upper Bounds

Let $\mathcal{Z}(\Sigma_{\mathbf{X}})$ be the set of positive definite matrices such that $\Sigma_{\mathbf{X}}^{-1} + \Sigma_{\mathbf{Z}}^{-1}$ is a block diagonal matrix for every $\Sigma_{\mathbf{Z}} \in \mathcal{Z}(\Sigma_{\mathbf{X}})$. It is easy to show that $\mathcal{Z}(\Sigma_{\mathbf{X}})$ is non-empty. We obtain the following lower bound on $R(\mathbf{D}_1, \dots, \mathbf{D}_L, \alpha)$, which is a generalization of [11, Th. 5] from the two-terminal case to the L -terminal case, by coupling the direct vector Gaussian L -terminal source coding problem with the vector Gaussian CEO problem through a limiting argument. The proof is given in Appendix B.

Theorem 5: $R(\mathbf{D}_1, \dots, \mathbf{D}_L, \alpha) \geq \underline{R}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{Z}}, \alpha)$ for α with $\alpha_1 \geq \dots \geq \alpha_L > 0$ and $\Sigma_{\mathbf{Z}} \in \mathcal{Z}(\Sigma_{\mathbf{X}})$, where

$$\begin{aligned} \underline{R}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{Z}}, \alpha) & \triangleq \min_{\mathbf{D}, \Gamma_1, \dots, \Gamma_L} \frac{\alpha_L}{2} \log \frac{|\Sigma_{\mathbf{X}} + \Sigma_{\mathbf{Z}}|}{|\mathbf{D} + \Sigma_{\mathbf{Z}}|} + \sum_{i=1}^L \frac{\alpha_i}{2} \log \frac{|\Sigma_{\mathbf{Z}_i}|}{|\Gamma_i|} \\ & + \sum_{i=2}^L \frac{\alpha_{i-1} - \alpha_i}{2} \log^+ \frac{|\Sigma_{\mathbf{Z}}(\Sigma_{\mathbf{Z}} - \bar{\Gamma}_i)^{-1} \Sigma_{\mathbf{Z}}|}{|\mathbf{D} + \Sigma_{\mathbf{Z}}|} \\ & \text{subject to } \mathbf{0} \leq \mathbf{D} \leq \Sigma_{\mathbf{X}}, \\ & (\mathbf{D})_i \leq \mathbf{D}_i, \quad i = 1, \dots, L, \\ & \mathbf{0} \leq \Gamma_i \leq \Sigma_{\mathbf{Z}_i}, \quad i = 1, \dots, L, \\ & \text{diag}(\Gamma_1, \dots, \Gamma_L) \leq (\mathbf{D}^{-1} + \Sigma_{\mathbf{Z}}^{-1})^{-1} \end{aligned}$$

with $\Sigma_{\mathbf{Z}_i} = ((\Sigma_{\mathbf{X}}^{-1} + \Sigma_{\mathbf{Z}}^{-1})^{-1})_i$, $i = 1, \dots, L$, and $\bar{\Gamma}_i = \text{diag}(\Sigma_{\mathbf{Z}_1}, \dots, \Sigma_{\mathbf{Z}_{i-1}}, \Gamma_i, \dots, \Gamma_L)$, $i = 2, \dots, L$.

One can treat the degenerate case $\alpha_1 \geq \dots \geq \alpha_{L'} > \alpha_{L'+1} = \dots = \alpha_L = 0$ for some $L' < L$ through an argument similar to that for Corollary 1. Let $\mathbf{X}'_i = \mathbf{X}_i - \mathbb{E}[\mathbf{X}_i | \mathbf{X}_{L'+1}, \dots, \mathbf{X}_L]$, $i = 1, \dots, L'$, and $\mathbf{X}'^i = ((\mathbf{X}'_1)^T, \dots, (\mathbf{X}'_{L'}^T)^T)^T$. Let $\mathcal{Z}(\Sigma_{\mathbf{X}'})$ be the set of positive definite matrix $\Sigma_{\mathbf{Z}'}$ such that $\Sigma_{\mathbf{X}'}^{-1} + \Sigma_{\mathbf{Z}'}^{-1}$ is a block diagonal matrix. The proof of the following corollary is omitted.

Corollary 2: If $\alpha_1 \geq \dots \geq \alpha_{L'} > \alpha_{L'+1} = \dots = \alpha_L = 0$ for some $L' < L$, then $R(\mathbf{D}_1, \dots, \mathbf{D}_L, \alpha) \geq \underline{R}'(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{Z}'}, \alpha)$ for $\Sigma_{\mathbf{Z}'} \in \mathcal{Z}(\Sigma_{\mathbf{X}'})$, where

$$\begin{aligned} \underline{R}'(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{Z}'}, \alpha) & \triangleq \min_{\mathbf{D}, \Gamma_1, \dots, \Gamma_{L'}} \frac{\alpha_{L'}}{2} \log \frac{|\Sigma_{\mathbf{X}'} + \Sigma_{\mathbf{Z}'}|}{|\mathbf{D} + \Sigma_{\mathbf{Z}'}|} + \sum_{i=1}^{L'} \frac{\alpha_i}{2} \log \frac{|\Sigma_{\mathbf{Z}'_i}|}{|\Gamma_i|} \\ & + \sum_{i=2}^{L'} \frac{\alpha_{i-1} - \alpha_i}{2} \log^+ \frac{|\Sigma_{\mathbf{Z}'}(\Sigma_{\mathbf{Z}'} - \bar{\Gamma}'_i)^{-1} \Sigma_{\mathbf{Z}'}|}{|\mathbf{D} + \Sigma_{\mathbf{Z}'}|} \\ & \text{subject to } \mathbf{0} \leq \mathbf{D} \leq \Sigma_{\mathbf{X}'}, \\ & (\mathbf{D})_i \leq \mathbf{D}_i, \quad i = 1, \dots, L', \\ & \mathbf{0} \leq \Gamma_i \leq \Sigma_{\mathbf{Z}'_i}, \quad i = 1, \dots, L', \\ & \text{diag}(\Gamma_1, \dots, \Gamma_{L'}) \leq (\mathbf{D}^{-1} + \Sigma_{\mathbf{Z}'}^{-1})^{-1} \end{aligned}$$

with $\Sigma_{\mathbf{Z}'_i} = ((\Sigma_{\mathbf{X}'}^{-1} + \Sigma_{\mathbf{Z}'}^{-1})^{-1})_i$, $i = 1, \dots, L'$, and $\bar{\Gamma}'_i = \text{diag}(\Sigma_{\mathbf{Z}'_1}, \dots, \Sigma_{\mathbf{Z}'_{i-1}}, \Gamma_i, \dots, \Gamma_{L'})$, $i = 2, \dots, L'$.

One can obtain the following upper bound on $R(\mathbf{D}_1, \dots, \mathbf{D}_L, \alpha)$, referred to as the Berger-Tung upper bound on $R(\mathbf{D}_1, \dots, \mathbf{D}_L, \alpha)$, by evaluating the standard Berger-Tung inner bound of $\mathcal{R}(\mathbf{D}_1, \dots, \mathbf{D}_L)$ [21], [22]. This result is a straightforward generalization of [11, Th. 6] from

the two-terminal case to the L -terminal case. The proof is omitted.

Theorem 6: $R(\mathbf{D}_1, \dots, \mathbf{D}_L, \alpha) \leq \bar{R}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{Z}}, \alpha)$ for $\Sigma_{\mathbf{Z}} \in \mathcal{Z}(\Sigma_{\mathbf{X}})$, where

$$\begin{aligned} \bar{R}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{Z}}, \alpha) & \triangleq \min_{\mathbf{D}, \Gamma_1, \dots, \Gamma_L} \frac{\alpha_L}{2} \log \frac{|\Sigma_{\mathbf{X}} + \Sigma_{\mathbf{Z}}|}{|\mathbf{D} + \Sigma_{\mathbf{Z}}|} + \sum_{i=1}^L \frac{\alpha_i}{2} \log \frac{|\Sigma_{\mathbf{Z}_i}|}{|\Gamma_i|} \\ & + \sum_{i=2}^L \frac{\alpha_{i-1} - \alpha_i}{2} \log^+ \frac{|\Sigma_{\mathbf{Z}}(\Sigma_{\mathbf{Z}} - \bar{\Gamma}_i)^{-1} \Sigma_{\mathbf{Z}}|}{|\mathbf{D} + \Sigma_{\mathbf{Z}}|} \\ & \text{subject to } \mathbf{0} \leq \mathbf{D} \leq \Sigma_{\mathbf{X}}, \\ & (\mathbf{D})_i \leq \mathbf{D}_i, \quad i = 1, \dots, L, \\ & \mathbf{0} \leq \Gamma_i \leq \Sigma_{\mathbf{Z}_i}, \quad i = 1, \dots, L, \\ & \text{diag}(\Gamma_1, \dots, \Gamma_L) = (\mathbf{D}^{-1} + \Sigma_{\mathbf{Z}}^{-1})^{-1}. \end{aligned} \quad (6)$$

Remark: It can be verified that “log⁺” in (6) can be replaced with “log”; moreover, (7) is implied by (8) and (9). We keep the current form for the purpose of stressing the similarity between $\underline{R}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{Z}}, \alpha)$ and $\bar{R}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{Z}}, \alpha)$.

C. Matching Conditions

Unless specified otherwise, we assume $\alpha_1 \geq \dots \geq \alpha_L > 0$ in this subsection. The degenerate case where $\alpha_1 \geq \dots \geq \alpha_{L'} > \alpha_{L'+1} = \dots = \alpha_L = 0$ for some $L' < L$ can be treated analogously and is left to the interested reader.

Let $\mathcal{D}(\Sigma_{\mathbf{X}})$ be the set of positive definite matrices such that $\mathbf{D}^{-1} - \Sigma_{\mathbf{X}}^{-1}$ is a positive semidefinite block diagonal matrix for every $\mathbf{D} \in \mathcal{D}(\Sigma_{\mathbf{X}})$. Define

$$\mathcal{D}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{X}}) = \{\mathbf{D} : \mathbf{D} \in \mathcal{D}(\Sigma_{\mathbf{X}}), (\mathbf{D})_i = \mathbf{D}_i, i = 1, \dots, L\}.$$

The following two theorems provide a certain characterization of $\mathcal{D}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{X}})$. The proofs are given in Appendices C and D.

Theorem 7: The set $\mathcal{D}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{X}})$ is a singleton if it is nonempty.

Theorem 8: The set $\mathcal{D}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{X}})$ is non-empty if $\mathbf{D}_i \leq \Sigma_{\mathbf{X}_i | \{\mathbf{X}_j\}_{j \neq i}}$, $i = 1, \dots, L$.

Remark: The proof of Theorem 8 is based on Borsuk’s theorem in nonlinear analysis. It will be seen that Theorem 8 plays an important role in our proof of the tightness of $\underline{R}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{Z}}, \alpha)$ and $\bar{R}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{Z}}, \alpha)$ in the high-resolution regime and the weak-dependence regime.

Now we proceed to derive matching conditions for $\underline{R}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{Z}}, \alpha)$ and $\bar{R}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{Z}}, \alpha)$. Let $\mathbf{F} = (\mathbf{F}_1, \dots, \mathbf{F}_L)$ be a block diagonal matrix with $(\mathbf{F})_i = \mathbf{I}_{m_i}$, $i = 1, \dots, L$, where \mathbf{F}_i is an $m \times m_i$ matrix, $i = 1, \dots, L$. The proof of the following lemma can be found in Appendix E.

Lemma 1: If there exist $\mathbf{D}, \Sigma_{\mathbf{Z}}, \Gamma_1, \dots, \Gamma_L$ with $\mathbf{D} \in \mathcal{D}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{X}})$, $\Sigma_{\mathbf{Z}} \in \mathcal{Z}(\Sigma_{\mathbf{X}})$, and $\text{diag}(\Gamma_1, \dots, \Gamma_L) = (\mathbf{D}^{-1} + \Sigma_{\mathbf{Z}}^{-1})^{-1}$ such that one can find a block diagonal

matrix Π satisfying

$$\begin{aligned} \Pi &\succeq \alpha_1(\mathbf{D} + \Sigma_{\mathbf{Z}})^{-1}, \\ (\mathbf{D}(\Pi - \alpha_1(\mathbf{D} + \Sigma_{\mathbf{Z}})^{-1})\mathbf{D})_i &= \alpha_i \Gamma_i + \Gamma_i \Theta_i \Gamma_i, \\ &i = 1, \dots, L, \end{aligned}$$

then $\underline{\mathbf{R}}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{Z}}, \alpha) = \overline{\mathbf{R}}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{Z}}, \alpha)$. Here $\Theta_1 = \mathbf{0}$ and

$$\Theta_i = -\sum_{j=2}^i (\alpha_{j-1} - \alpha_j) \mathbf{F}_i^T (\Sigma_{\mathbf{Z}} - \bar{\Gamma}_j)^{-1} \mathbf{F}_i, \quad i = 2, \dots, L.$$

Remark: It can be verified that for $\mathbf{D}, \Sigma_{\mathbf{Z}}, \Gamma_1, \dots, \Gamma_L$ with $\mathbf{D} \in \mathcal{D}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{X}})$, $\Sigma_{\mathbf{Z}} \in \mathcal{Z}(\Sigma_{\mathbf{X}})$, and $\text{diag}(\Gamma_1, \dots, \Gamma_L) = (\mathbf{D}^{-1} + \Sigma_{\mathbf{Z}}^{-1})^{-1}$, we have

$$\begin{aligned} (\mathbf{D}(\Pi - \alpha_1(\mathbf{D} + \Sigma_{\mathbf{Z}})^{-1})\mathbf{D})_i &= \alpha_i \Gamma_i + \Gamma_i \Theta_i \Gamma_i \\ \Leftrightarrow (\mathbf{D}\Pi\mathbf{D})_i &= \Xi_i, \quad i = 1, \dots, L, \end{aligned} \quad (10)$$

where $\Xi_i = \alpha_1 \mathbf{D}_i - (\alpha_1 - \alpha_i) \Gamma_i + \Gamma_i \Theta_i \Gamma_i$, $i = 1, \dots, L$. Now write \mathbf{D} in the partitioned form

$$\begin{pmatrix} \mathbf{D}_{1,1} & \cdots & \mathbf{D}_{1,L} \\ \vdots & \ddots & \vdots \\ \mathbf{D}_{L,1} & \cdots & \mathbf{D}_{L,L} \end{pmatrix},$$

where $\mathbf{D}_{i,i} = \mathbf{D}_i$, $i = 1, \dots, L$. Note that

$$\begin{pmatrix} \mathbf{D}_{1,1} \otimes \mathbf{D}_{1,1} & \cdots & \mathbf{D}_{1,L} \otimes \mathbf{D}_{1,L} \\ \vdots & \ddots & \vdots \\ \mathbf{D}_{L,1} \otimes \mathbf{D}_{L,1} & \cdots & \mathbf{D}_{L,L} \otimes \mathbf{D}_{L,L} \end{pmatrix} = \mathbf{D} * \mathbf{D},$$

where \otimes is the Kronecker product and $*$ is the Khatri-Rao product [23]. We can express (10) equivalently as

$$(\mathbf{D} * \mathbf{D}) \begin{pmatrix} \text{vec}((\Pi)_1) \\ \vdots \\ \text{vec}((\Pi)_L) \end{pmatrix} = \begin{pmatrix} \text{vec}((\Xi)_1) \\ \vdots \\ \text{vec}((\Xi)_L) \end{pmatrix},$$

where $\text{vec}(\cdot)$ is the vec operator. Since $\mathbf{D} * \mathbf{D}$ is invertible (which is due to the fact that \mathbf{D} is positive definite), it follows that the block diagonal matrix Π is determined by the following equation

$$\begin{pmatrix} \text{vec}((\Pi)_1) \\ \vdots \\ \text{vec}((\Pi)_L) \end{pmatrix} = (\mathbf{D} * \mathbf{D})^{-1} \begin{pmatrix} \text{vec}((\Xi)_1) \\ \vdots \\ \text{vec}((\Xi)_L) \end{pmatrix}.$$

We shall show that the condition specified in Lemma 1 is satisfied in the high-resolution regime (i.e., when $\|\mathbf{D}_1\|, \dots, \|\mathbf{D}_L\|$ are small) and the weak-dependence regime (i.e., when $\|\Sigma_{\mathbf{X}} - \text{diag}(\Sigma_{\mathbf{X}_1}, \dots, \Sigma_{\mathbf{X}_L})\|$ is small), which leads to the following two theorems. The proofs can be found in Appendices F and G.

Theorem 9: For any $\Sigma_{\mathbf{Z}} \in \mathcal{Z}(\Sigma_{\mathbf{X}})$, there exists a $\rho(\Sigma_{\mathbf{X}}, \Sigma_{\mathbf{Z}}, \alpha) > 0$ such that if $\|\mathbf{D}_i\| \leq \rho(\Sigma_{\mathbf{X}}, \Sigma_{\mathbf{Z}}, \alpha)$, $i = 1, \dots, L$, then $\underline{\mathbf{R}}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{Z}}, \alpha) = \overline{\mathbf{R}}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{Z}}, \alpha)$.

Theorem 10: For any $(\mathbf{D}_1, \dots, \mathbf{D}_L)$ with $\mathbf{0} \prec \mathbf{D}_i \prec \Sigma_{\mathbf{X}_i}$, $i = 1, \dots, L$, there exists a $\varrho(\Sigma_{\mathbf{X}}, \alpha) > 0$ such that if $\|\Sigma_{\mathbf{X}} - \text{diag}(\Sigma_{\mathbf{X}_1}, \dots, \Sigma_{\mathbf{X}_L})\| \leq \varrho(\Sigma_{\mathbf{X}}, \alpha)$, then

$\underline{\mathbf{R}}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{Z}}, \alpha) = \overline{\mathbf{R}}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{Z}}, \alpha)$ for some $\Sigma_{\mathbf{Z}} \in \mathcal{Z}(\Sigma_{\mathbf{X}})$.

Remark: It can be seen from Theorems 9 and 10 that the matching conditions here in general depend on α . This should be contrasted with the CEO problem for which we have a matching condition that is uniform over all α (see the remark after Theorem 3).

IV. AN EXTREMAL INEQUALITY

Let $\mathbf{X}_0, \mathbf{X}_2, \dots, \mathbf{X}_L, \mathbf{N}_2, \dots, \mathbf{N}_L$ be defined as in Section II-A. For the purpose of subsequent analysis, we define $\mathbf{Y} = \mathbb{E}[\mathbf{X}_0 | \mathbf{X}_2, \dots, \mathbf{X}_L]$, $\tilde{\mathbf{N}} = \mathbf{X}_0 - \mathbf{Y}$, and $\tilde{\mathbf{N}}_i = \mathbf{X}_0 - \mathbb{E}[\mathbf{X}_0 | \mathbf{X}_i]$, $i = 2, \dots, L$. Note that $\tilde{\mathbf{N}}$ is independent of $(\mathbf{X}_2, \dots, \mathbf{X}_L)$, and $\tilde{\mathbf{N}}_i$ is independent of \mathbf{X}_j , $i = 2, \dots, L$. Moreover, it can be verified that

$$\begin{aligned} \Sigma_{\tilde{\mathbf{N}}} &= \left(\Sigma_{\mathbf{X}_0}^{-1} + \sum_{i=2}^L \Sigma_{\mathbf{N}_i}^{-1} \right)^{-1}, \\ \mathbf{Y} &= \sum_{i=2}^L \Sigma_{\tilde{\mathbf{N}}} \Sigma_{\mathbf{N}_i}^{-1} \mathbf{X}_i, \\ \Sigma_{\tilde{\mathbf{N}}_i} &= (\Sigma_{\mathbf{X}_0}^{-1} + \Sigma_{\mathbf{N}_i}^{-1})^{-1}, \quad i = 2, \dots, L. \end{aligned}$$

Let U_2, \dots, U_L, V be jointly distributed with $(\mathbf{X}_0, \mathbf{X}_2, \dots, \mathbf{X}_L)$ such that the joint distribution factors as

$$\begin{aligned} p(\mathbf{x}_0, \mathbf{x}_2, \dots, \mathbf{x}_L, u_2, \dots, u_L, v) \\ = p(\mathbf{x}_0, \mathbf{x}_2, \dots, \mathbf{x}_L) p(v) \prod_{i=2}^L p(u_i | \mathbf{x}_i, v). \end{aligned}$$

Let $\mu_i = (\alpha_i - \alpha_{i+1})\theta_{i+1}$, $i = 1, \dots, L-1$, and $\mu_L = \mu_{L-1} + \alpha_L$, where $\alpha_1 \geq \dots \geq \alpha_L > 0$ and $\theta_i \in [0, 1]$, $i = 2, \dots, L$.

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

Theorem 11: If $\Sigma_{\mathbf{X}_i | \mathbf{X}_0, U_i, V} \leq \mathbf{D}_i \leq \Sigma_{\mathbf{N}_i}$ for some positive definite matrix \mathbf{D}_i , $i = 2, \dots, L$, and there exist matrices $\hat{\mathbf{D}}_i$ and \mathbf{O}_i with $\mathbf{0} \prec \hat{\mathbf{D}}_i \leq \mathbf{D}_i$ and $\mathbf{O}_i \succeq \mathbf{0}$, $i = 2, \dots, L$, such that

$$\begin{aligned} \sum_{j=2}^i \mu_{j-1} \Sigma_{\mathbf{N}_i}^{-1} \left(\Sigma_{\mathbf{X}_0}^{-1} + \sum_{k=j}^L \Sigma_{\mathbf{N}_k}^{-1} - \sum_{k=j}^L \Sigma_{\mathbf{N}_k}^{-1} \hat{\mathbf{D}}_k \Sigma_{\mathbf{N}_k}^{-1} \right)^{-1} \\ \cdot \Sigma_{\mathbf{N}_i}^{-1} + \mathbf{O}_i = \alpha_i \hat{\mathbf{D}}_i^{-1}, \quad i = 2, \dots, L, \end{aligned} \quad (11)$$

$$\mathbf{O}_i (\mathbf{D}_i - \hat{\mathbf{D}}_i) = \mathbf{0}, \quad i = 2, \dots, L, \quad (12)$$

then

$$\begin{aligned} \sum_{i=2}^L \mu_{i-1} h(\mathbf{X}_0 | U_i^L, V) - \sum_{i=2}^L \alpha_i h(\mathbf{X}_i | \mathbf{X}_0, U_i, V) \geq \sum_{i=2}^L \frac{\mu_{i-1}}{2} \\ \cdot \log \left| 2\pi e \left(\Sigma_{\mathbf{X}_0}^{-1} + \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} - \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} \hat{\mathbf{D}}_j \Sigma_{\mathbf{N}_j}^{-1} \right)^{-1} \right| \\ - \sum_{i=2}^L \frac{\alpha_i}{2} \log |2\pi e \hat{\mathbf{D}}_i|. \end{aligned} \quad (13)$$

Remark: The lower bound in (13) is attained when V is independent of $(\mathbf{X}_0, \mathbf{X}_2, \dots, \mathbf{X}_L, U_2, \dots, U_L)$, and U_2, \dots, U_L are jointly Gaussian with $(\mathbf{X}_0, \mathbf{X}_2, \dots, \mathbf{X}_L)$ such that $\Sigma_{\mathbf{X}_i|\mathbf{X}_0, U_i} = \hat{\mathbf{D}}_i$, $i = 2, \dots, L$. It is instructive to note that (11) and (12) are actually the necessary conditions for $(\hat{\mathbf{D}}_2, \dots, \hat{\mathbf{D}}_L)$ to be an optimal solution to the following minimization problem

$$(\mathbf{P}_1) \min_{\hat{\mathbf{D}}_2, \dots, \hat{\mathbf{D}}_L} \sum_{i=2}^L \frac{\mu_{i-1}}{2} \cdot \log \left| 2\pi e \left(\Sigma_{\mathbf{X}_0}^{-1} + \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} - \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} \hat{\mathbf{D}}_j \Sigma_{\mathbf{N}_j}^{-1} \right)^{-1} \right| - \sum_{i=2}^L \frac{\alpha_i}{2} \log |2\pi e \hat{\mathbf{D}}_i|$$

subject to $\mathbf{0} \leq \hat{\mathbf{D}}_i \leq \mathbf{D}_i$, $i = 2, \dots, L$.

Theorem 11 implies that (11) and (12) (together with the conditions $\mathbf{0} < \hat{\mathbf{D}}_i \leq \mathbf{D}_i$ and $\mathbf{O}_i \geq \mathbf{0}$, $i = 2, \dots, L$) are in fact sufficient to guarantee the optimality of $(\hat{\mathbf{D}}_2, \dots, \hat{\mathbf{D}}_L)$ for (\mathbf{P}_1) .

A. An Equivalent Version

Note that

$$\begin{aligned} h(\mathbf{X}_0|U_L, V) &= h(\mathbf{X}_0, \mathbf{X}_L|U_L, V) - h(\mathbf{X}_L|\mathbf{X}_0, U_L, V) \\ &= h(\mathbf{X}_L|U_L, V) + h(\mathbf{X}_0|\mathbf{X}_L, U_L, V) - h(\mathbf{X}_L|\mathbf{X}_0, U_L, V) \\ &= h(\mathbf{X}_L|U_L, V) + h(\mathbf{X}_0|\mathbf{X}_L) - h(\mathbf{X}_L|\mathbf{X}_0, U_L, V) \quad (14) \\ &= h(\mathbf{X}_L|U_L, V) + h(\tilde{\mathbf{N}}_L) - h(\mathbf{X}_L|\mathbf{X}_0, U_L, V) \\ &= h(\mathbf{X}_L|U_L, V) + \frac{1}{2} \log |2\pi e \Sigma_{\tilde{\mathbf{N}}_L}| - h(\mathbf{X}_L|\mathbf{X}_0, U_L, V) \end{aligned}$$

where (14) follows from the fact that $\mathbf{X}_0 \leftrightarrow \mathbf{X}_L \leftrightarrow (U_L, V)$ form a Markov chain. Therefore, we have

$$\begin{aligned} &\sum_{i=2}^L \mu_{i-1} h(\mathbf{X}_0|U_i^L, V) - \sum_{i=2}^L \alpha_i h(\mathbf{X}_i|\mathbf{X}_0, U_i, V) \\ &= \sum_{i=2}^{L-1} \mu_{i-1} h(\mathbf{X}_0|U_i^L, V) - \sum_{i=2}^{L-1} \alpha_i h(\mathbf{X}_i|\mathbf{X}_0, U_i, V) \\ &\quad + \mu_{L-1} h(\mathbf{X}_0|U_L, V) - \alpha_L h(\mathbf{X}_L|\mathbf{X}_0, U_L, V) \\ &= \sum_{i=2}^{L-1} \mu_{i-1} h(\mathbf{X}_0|U_i^L, V) - \sum_{i=2}^{L-1} \alpha_i h(\mathbf{X}_i|\mathbf{X}_0, U_i, V) \\ &\quad + \mu_{L-1} h(\mathbf{X}_L|U_L, V) - \mu_L h(\mathbf{X}_L|\mathbf{X}_0, U_L, V) \\ &\quad + \frac{\mu_{L-1}}{2} \log |2\pi e \Sigma_{\tilde{\mathbf{N}}_L}|. \quad (15) \end{aligned}$$

It is clear that

$$\begin{aligned} &\log |\Sigma_{\mathbf{X}_0}^{-1} + \Sigma_{\mathbf{N}_L}^{-1} - \Sigma_{\mathbf{N}_L}^{-1} \hat{\mathbf{D}}_L \Sigma_{\mathbf{N}_L}^{-1}| - \log |\hat{\mathbf{D}}_L| \\ &= \log |\Sigma_{\tilde{\mathbf{N}}_L}^{-1} - \Sigma_{\mathbf{N}_L}^{-1} \hat{\mathbf{D}}_L \Sigma_{\mathbf{N}_L}^{-1}| - \log |\hat{\mathbf{D}}_L| \end{aligned}$$

$$\begin{aligned} &= \log |\Sigma_{\mathbf{N}_L} \Sigma_{\tilde{\mathbf{N}}_L}^{-1} \Sigma_{\mathbf{N}_L} - \hat{\mathbf{D}}_L| - \log |\hat{\mathbf{D}}_L| - 2 \log |\Sigma_{\mathbf{N}_L}| \\ &= \log |\Sigma_{\mathbf{N}_L} \Sigma_{\tilde{\mathbf{N}}_L}^{-1} \Sigma_{\mathbf{N}_L} \hat{\mathbf{D}}_L^{-1} - \mathbf{I}_m| - 2 \log |\Sigma_{\mathbf{N}_L}| \\ &= \log |\hat{\mathbf{D}}_L^{-1} - \Sigma_{\mathbf{N}_L}^{-1} \Sigma_{\tilde{\mathbf{N}}_L} \Sigma_{\mathbf{N}_L}^{-1}| - \log |\Sigma_{\tilde{\mathbf{N}}_L}|. \end{aligned}$$

As a consequence,

$$\begin{aligned} &\frac{\mu_{L-1}}{2} \log \left| \Sigma_{\mathbf{X}_0}^{-1} + \Sigma_{\mathbf{N}_L}^{-1} - \Sigma_{\mathbf{N}_L}^{-1} \hat{\mathbf{D}}_L \Sigma_{\mathbf{N}_L}^{-1} \right| + \frac{\alpha_L}{2} \log |\hat{\mathbf{D}}_L| \\ &= \frac{\mu_{L-1}}{2} \log |\hat{\mathbf{D}}_L^{-1} - \Sigma_{\mathbf{N}_L}^{-1} \Sigma_{\tilde{\mathbf{N}}_L} \Sigma_{\mathbf{N}_L}^{-1}| + \frac{\mu_L}{2} \log |\hat{\mathbf{D}}_L| \\ &\quad - \frac{\mu_{L-1}}{2} \log |\Sigma_{\tilde{\mathbf{N}}_L}|. \quad (16) \end{aligned}$$

In view of (15) and (16), one can readily show that (13) \Leftrightarrow

$$\begin{aligned} &\sum_{i=2}^{L-1} \mu_{i-1} h(\mathbf{X}_0|U_i^L, V) - \sum_{i=2}^{L-1} \alpha_i h(\mathbf{X}_i|\mathbf{X}_0, U_i, V) \\ &\quad + \mu_{L-1} h(\mathbf{X}_L|U_L, V) - \mu_L h(\mathbf{X}_L|\mathbf{X}_0, U_L, V) \geq \sum_{i=2}^{L-1} \frac{\mu_{i-1}}{2} \\ &\quad \cdot \log \left| 2\pi e \left(\Sigma_{\mathbf{X}_0}^{-1} + \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} - \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} \hat{\mathbf{D}}_j \Sigma_{\mathbf{N}_j}^{-1} \right)^{-1} \right| \\ &\quad - \sum_{i=2}^{L-1} \frac{\alpha_i}{2} \log |2\pi e \hat{\mathbf{D}}_i| \\ &\quad + \frac{\mu_{L-1}}{2} \log \left| 2\pi e (\hat{\mathbf{D}}_L^{-1} - \Sigma_{\mathbf{N}_L}^{-1} \Sigma_{\tilde{\mathbf{N}}_L} \Sigma_{\mathbf{N}_L}^{-1})^{-1} \right| \\ &\quad - \frac{\mu_L}{2} \log |2\pi e \hat{\mathbf{D}}_L|. \quad (17) \end{aligned}$$

Taking the gradient of both sides of the equation in (16) with respect to $\hat{\mathbf{D}}_L$ yields

$$\begin{aligned} &-\frac{\mu_{L-1}}{2} \Sigma_{\mathbf{N}_L}^{-1} \left(\Sigma_{\mathbf{X}_0}^{-1} + \Sigma_{\mathbf{N}_L}^{-1} - \Sigma_{\mathbf{N}_L}^{-1} \hat{\mathbf{D}}_L \Sigma_{\mathbf{N}_L}^{-1} \right)^{-1} \Sigma_{\mathbf{N}_L}^{-1} + \frac{\alpha_L}{2} \hat{\mathbf{D}}_L^{-1} \\ &= -\frac{\mu_{L-1}}{2} \hat{\mathbf{D}}_L^{-1} \left(\hat{\mathbf{D}}_L^{-1} - \Sigma_{\mathbf{N}_L}^{-1} \Sigma_{\tilde{\mathbf{N}}_L} \Sigma_{\mathbf{N}_L}^{-1} \right)^{-1} \hat{\mathbf{D}}_L^{-1} + \frac{\mu_L}{2} \hat{\mathbf{D}}_L^{-1} \end{aligned}$$

which implies that

$$\begin{aligned} &\sum_{j=2}^L \mu_{j-1} \Sigma_{\mathbf{N}_L}^{-1} \left(\Sigma_{\mathbf{X}_0}^{-1} + \sum_{k=j}^L \Sigma_{\mathbf{N}_k}^{-1} - \sum_{k=j}^L \Sigma_{\mathbf{N}_k}^{-1} \hat{\mathbf{D}}_k \Sigma_{\mathbf{N}_k}^{-1} \right)^{-1} \\ &\quad \cdot \Sigma_{\mathbf{N}_L}^{-1} + \mathbf{O}_L = \alpha_L \hat{\mathbf{D}}_L^{-1} \\ &\Leftrightarrow \sum_{j=2}^{L-1} \mu_{j-1} \Sigma_{\mathbf{N}_L}^{-1} \left(\Sigma_{\mathbf{X}_0}^{-1} + \sum_{k=j}^L \Sigma_{\mathbf{N}_k}^{-1} - \sum_{k=j}^L \Sigma_{\mathbf{N}_k}^{-1} \hat{\mathbf{D}}_k \Sigma_{\mathbf{N}_k}^{-1} \right)^{-1} \\ &\quad \cdot \Sigma_{\mathbf{N}_L}^{-1} + \mu_{L-1} \hat{\mathbf{D}}_L^{-1} \left(\hat{\mathbf{D}}_L^{-1} - \Sigma_{\mathbf{N}_L}^{-1} \Sigma_{\tilde{\mathbf{N}}_L} \Sigma_{\mathbf{N}_L}^{-1} \right)^{-1} \hat{\mathbf{D}}_L^{-1} \\ &\quad + \mathbf{O}_L = \mu_L \hat{\mathbf{D}}_L^{-1}. \quad (18) \end{aligned}$$

In view of (17) and (18), we can restate Theorem 11 in the following equivalent form.

Theorem 12: If $\Sigma_{\mathbf{X}_i|\mathbf{X}_0, U_i, V} \leq \mathbf{D}_i \leq \Sigma_{\mathbf{N}_i}$ for some positive definite matrix \mathbf{D}_i , $i = 2, \dots, L$, and there exist matrices $\hat{\mathbf{D}}_i$ and \mathbf{O}_i with $\mathbf{0} < \hat{\mathbf{D}}_i \leq \mathbf{D}_i$ and $\mathbf{O}_i \geq \mathbf{0}$, $i = 2, \dots, L$,

such that

$$\sum_{j=2}^i \mu_{j-1} \Sigma_{\mathbf{N}_i}^{-1} \left(\Sigma_{\mathbf{X}_0}^{-1} + \sum_{k=j}^L \Sigma_{\mathbf{N}_k}^{-1} - \sum_{k=j}^L \Sigma_{\mathbf{N}_k}^{-1} \hat{\mathbf{D}}_k \Sigma_{\mathbf{N}_k}^{-1} \right)^{-1} \cdot \Sigma_{\mathbf{N}_i}^{-1} + \mathbf{O}_i = \alpha_i \hat{\mathbf{D}}_i^{-1}, \quad i = 2, \dots, L-1, \quad (19)$$

$$\sum_{j=2}^{L-1} \mu_{j-1} \Sigma_{\mathbf{N}_L}^{-1} \left(\Sigma_{\mathbf{X}_0}^{-1} + \sum_{k=j}^L \Sigma_{\mathbf{N}_k}^{-1} - \sum_{k=j}^L \Sigma_{\mathbf{N}_k}^{-1} \hat{\mathbf{D}}_k \Sigma_{\mathbf{N}_k}^{-1} \right)^{-1} \cdot \Sigma_{\mathbf{N}_L}^{-1} + \mu_{L-1} \hat{\mathbf{D}}_L^{-1} (\hat{\mathbf{D}}_L^{-1} - \Sigma_{\mathbf{N}_L}^{-1} \Sigma_{\tilde{\mathbf{N}}_L} \Sigma_{\mathbf{N}_L}^{-1})^{-1} \hat{\mathbf{D}}_L^{-1} + \mathbf{O}_L = \mu_L \hat{\mathbf{D}}_L^{-1}, \quad (20)$$

$$\mathbf{O}_i (\mathbf{D}_i - \hat{\mathbf{D}}_i) = \mathbf{0}, \quad i = 2, \dots, L, \quad (21)$$

then

$$\begin{aligned} & \sum_{i=2}^{L-1} \mu_{i-1} h(\mathbf{X}_0 | U_i^L, V) - \sum_{i=2}^{L-1} \alpha_i h(\mathbf{X}_i | \mathbf{X}_0, U_i, V) \\ & + \mu_{L-1} h(\mathbf{X}_L | U_L, V) - \mu_L h(\mathbf{X}_L | \mathbf{X}_0, U_L, V) \geq \sum_{i=2}^{L-1} \frac{\mu_{i-1}}{2} \\ & \cdot \log \left| 2\pi e \left(\Sigma_{\mathbf{X}_0}^{-1} + \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} - \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} \hat{\mathbf{D}}_j \Sigma_{\mathbf{N}_j}^{-1} \right)^{-1} \right| \\ & - \sum_{i=2}^{L-1} \frac{\alpha_i}{2} \log |2\pi e \hat{\mathbf{D}}_i| \\ & + \frac{\mu_{L-1}}{2} \log |2\pi e (\hat{\mathbf{D}}_L^{-1} - \Sigma_{\mathbf{N}_L}^{-1} \Sigma_{\tilde{\mathbf{N}}_L} \Sigma_{\mathbf{N}_L}^{-1})^{-1}| \\ & - \frac{\mu_L}{2} \log |2\pi e \hat{\mathbf{D}}_L|. \end{aligned}$$

B. Fisher Information, Differential Entropy, and Conditional Expectation

We shall prove Theorem 12 by exploiting the connection between Fisher information, differential entropy, and conditional expectation. For any random object W and $m \times 1$ random vector \mathbf{S} , we define the conditional Fisher information matrix of \mathbf{S} given W as

$$\mathbf{J}(\mathbf{S}|W) = \mathbb{E}[\rho(\mathbf{S}|W)\rho^T(\mathbf{S}|W)],$$

where the conditional score function $\rho(\cdot|w)$ is defined as

$$\rho(\mathbf{s}|w) = \nabla_{\mathbf{s}} \log p(\mathbf{s}|w).$$

The following result is the conditional version of the multivariate de Bruijn's Identity [24, (51)].

Lemma 2 (Conditional De Bruijn's Identity): For (\mathbf{S}, W) satisfying $\text{tr}(\Sigma_{\mathbf{S}|W}) < \infty$,

$$\frac{d}{dt} h(\mathbf{H}\mathbf{S} + \sqrt{t}\mathbf{N}|W) = \frac{1}{2} \text{tr}(\mathbf{J}(\mathbf{H}\mathbf{S} + \sqrt{t}\mathbf{N}|W)),$$

where \mathbf{H} is a deterministic matrix, and \mathbf{N} is a random vector, independent of (\mathbf{S}, W) , with independent standard Gaussian entries.

The following two inequalities are straightforward extensions of their unconditional counterparts (see [25] for their proofs).

Lemma 3 (Conditional Fisher Information Matrix Inequality): Let $\mathbf{S}_1, \dots, \mathbf{S}_K$ be $m \times 1$ random vectors conditionally independent given W . For any $m \times m$ matrices $\mathbf{A}_1, \dots, \mathbf{A}_K$ satisfying $\mathbf{A}_1 + \dots + \mathbf{A}_K = \mathbf{I}_m$,

$$\mathbf{J}(\mathbf{S}_1 + \dots + \mathbf{S}_K | W) \preceq \sum_{i=1}^K \mathbf{A}_i \mathbf{J}(\mathbf{S}_i | W) \mathbf{A}_i^T.$$

Lemma 4 (Conditional Cramer-Rao Inequality): For (\mathbf{S}, W) satisfying $\Sigma_{\mathbf{S}|W} > 0$,

$$\mathbf{J}(\mathbf{S}|W) \succeq \Sigma_{\mathbf{S}|W}^{-1}.$$

The following result is the conditional version of [24, (57)].

Lemma 5: For any $k \times m$ deterministic matrix \mathbf{H} and $k \times 1$ random vector \mathbf{N} , independent of (\mathbf{S}, W) , with independent standard Gaussian entries,

$$\mathbf{J}(\sqrt{t}\mathbf{H}\mathbf{S} + \mathbf{N}|W) = \mathbf{I}_k - t\mathbf{H}\Sigma_{\mathbf{S}|\sqrt{t}\mathbf{H}\mathbf{S}+\mathbf{N},W}\mathbf{H}^T.$$

C. Proof of Theorem 12

The proof is based on a monotone path argument [12], [26]. Define

$$\hat{\mathbf{D}}_{\mathbf{X}_L|U_L} = (\hat{\mathbf{D}}_L^{-1} - \Sigma_{\mathbf{N}_L}^{-1} \Sigma_{\tilde{\mathbf{N}}_L} \Sigma_{\mathbf{N}_L}^{-1})^{-1},$$

$$\hat{\mathbf{D}}_{\mathbf{X}_0|U_i^L} = \left(\Sigma_{\mathbf{X}_0}^{-1} + \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} - \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} \hat{\mathbf{D}}_j \Sigma_{\mathbf{N}_j}^{-1} \right)^{-1}, \quad i = 2, \dots, L-1,$$

$$\hat{\mathbf{D}}_{\mathbf{Y}|U_i^L} = \hat{\mathbf{D}}_{\mathbf{X}_0|U_i^L} - \Sigma_{\tilde{\mathbf{N}}}, \quad i = 2, \dots, L-1,$$

$$\hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0, U_i^L} = \Sigma_{\tilde{\mathbf{N}}} \left(\sum_{j=2}^{i-1} \Sigma_{\mathbf{N}_j}^{-1} + \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} \hat{\mathbf{D}}_j \Sigma_{\mathbf{N}_j}^{-1} \right) \Sigma_{\tilde{\mathbf{N}}}, \quad i = 2, \dots, L-1.$$

Now we introduce some auxiliary random objects. Let $\mathbf{X}'_2, \dots, \mathbf{X}'_L, \tilde{\mathbf{X}}'_L, \mathbf{N}'_2, \dots, \mathbf{N}'_{L-2}, \tilde{\mathbf{Y}}'_2, \dots, \tilde{\mathbf{Y}}'_{L-1}$ be zero-mean Gaussian random vectors, independent of $(\mathbf{X}_0, \mathbf{X}_2, \dots, \mathbf{X}_L, U_2, \dots, U_L, V)$, with

$$\Sigma_{\mathbf{X}'_i} = \hat{\mathbf{D}}_i, \quad i = 2, \dots, L,$$

$$\Sigma_{\tilde{\mathbf{X}}'_L} = \hat{\mathbf{D}}_{\mathbf{X}_L|U_L},$$

$$\Sigma_{\mathbf{N}'_i} = \Sigma_{\mathbf{N}_i}, \quad i = 2, \dots, L-2,$$

$$\Sigma_{\tilde{\mathbf{Y}}'_i} = \hat{\mathbf{D}}_{\mathbf{Y}|U_i^L}, \quad i = 2, \dots, L-1.$$

We assume $\mathbf{X}'_2, \dots, \mathbf{X}'_L, \mathbf{N}'_2, \dots, \mathbf{N}'_{L-2}$ are mutually independent. Define

$$\mathbf{Y}'_i = \Sigma_{\tilde{\mathbf{N}}} \left(\sum_{j=2}^{i-1} \Sigma_{\mathbf{N}'_j}^{-1} \mathbf{N}'_j + \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} \mathbf{X}'_j \right), \quad i = 2, \dots, L-1.$$

It can be verified that

$$\Sigma_{\mathbf{Y}'_i} = \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0, U_i^L}, \quad i = 2, \dots, L-1.$$

Note that

$$\begin{aligned}
& \sum_{i=2}^{L-1} \mu_{i-1} h(\mathbf{X}_0 | U_i^L, V) - \sum_{i=2}^{L-1} \alpha_i h(\mathbf{X}_i | \mathbf{X}_0, U_i, V) \\
& \quad + \mu_{L-1} h(\mathbf{X}_L | U_L, V) - \mu_L h(\mathbf{X}_L | \mathbf{X}_0, U_L, V) \\
& = \sum_{i=2}^{L-1} \mu_{i-1} h(\mathbf{Y} | U_i^L, V) - \sum_{i=2}^{L-1} \mu_{i-1} h(\mathbf{Y} | \mathbf{X}_0, U_i^L, V) \\
& \quad - \sum_{i=2}^{L-1} \alpha_i h(\mathbf{X}_i | \mathbf{X}_0, U_i, V) + \mu_{L-1} h(\mathbf{X}_L | U_L, V) \\
& \quad - \mu_L h(\mathbf{X}_L | \mathbf{X}_0, U_L, V) + \sum_{i=2}^{L-1} \mu_{i-1} h(\tilde{\mathbf{N}}).
\end{aligned}$$

For $\lambda \in [0, 1]$, define

$$\begin{aligned}
\tilde{h}_\lambda & = \sum_{i=2}^{L-1} \mu_{i-1} h(\tilde{\mathbf{Y}}_{i,\lambda} | U_i^L, V) \\
& \quad - \sum_{i=2}^{L-1} \mu_{i-1} h(\mathbf{Y}_{i,\lambda} | \mathbf{X}_0, U_i^L, V) \\
& \quad - \sum_{i=2}^{L-1} \alpha_i h(\mathbf{X}_{i,\lambda} | \mathbf{X}_0, U_i, V) \\
& \quad + \mu_{L-1} h(\tilde{\mathbf{X}}_{L,\lambda} | U_L, V) - \mu_L h(\mathbf{X}_{L,\lambda} | \mathbf{X}_0, U_L, V) \\
& \quad + \sum_{i=2}^{L-1} \mu_{i-1} h(\tilde{\mathbf{N}}),
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{X}_{i,\lambda} & = \sqrt{\lambda} \mathbf{X}_i + \sqrt{1-\lambda} \mathbf{X}'_i, \quad i = 2, \dots, L, \\
\tilde{\mathbf{X}}_{L,\lambda} & = \sqrt{\lambda} \mathbf{X}_L + \sqrt{1-\lambda} \tilde{\mathbf{X}}'_L, \\
\mathbf{Y}_{i,\lambda} & = \sqrt{\lambda} \mathbf{Y} + \sqrt{1-\lambda} \mathbf{Y}'_i, \quad i = 2, \dots, L-1, \\
\tilde{\mathbf{Y}}_{i,\lambda} & = \sqrt{\lambda} \mathbf{Y} + \sqrt{1-\lambda} \tilde{\mathbf{Y}}'_i, \quad i = 2, \dots, L-1.
\end{aligned}$$

It can be readily seen that

$$\begin{aligned}
\tilde{h}_\lambda |_{\lambda=0} & = \sum_{i=2}^{L-1} \frac{\mu_{i-1}}{2} \log |2\pi e \hat{\mathbf{D}}_{\mathbf{X}_0 | U_i^L}| - \sum_{i=2}^{L-1} \frac{\alpha_i}{2} \log |2\pi e \hat{\mathbf{D}}_i| \\
& \quad + \frac{\mu_{L-1}}{2} \log |2\pi e \hat{\mathbf{D}}_{\mathbf{X}_L | U_L}| - \frac{\mu_L}{2} \log |2\pi e \hat{\mathbf{D}}_L|, \\
\tilde{h}_\lambda |_{\lambda=1} & = \sum_{i=2}^{L-1} \mu_{i-1} h(\mathbf{X}_0 | U_i^L, V) - \sum_{i=2}^{L-1} \alpha_i h(\mathbf{X}_i | \mathbf{X}_0, U_i, V) \\
& \quad + \mu_{L-1} h(\mathbf{X}_L | U_L, V) - \mu_L h(\mathbf{X}_L | \mathbf{X}_0, U_L, V).
\end{aligned}$$

Therefore, one just need to prove that

$$\frac{d\tilde{h}_\lambda}{d\lambda} \geq 0 \quad (22)$$

for $\lambda \in (0, 1)$. It can be shown by invoking Lemma 2 that

$$\begin{aligned}
\frac{d\tilde{h}_\lambda}{d\lambda} & = \sum_{i=2}^{L-1} \frac{\mu_{i-1}}{2\lambda} \left(m - \text{tr} \left(\hat{\mathbf{D}}_{\mathbf{Y} | U_i^L} \mathbf{J}(\tilde{\mathbf{Y}}_{i,\lambda} | U_i^L, V) \right) \right) \\
& \quad - \sum_{i=2}^{L-1} \frac{\mu_{i-1}}{2\lambda} \left(m - \text{tr} \left(\hat{\mathbf{D}}_{\mathbf{Y} | \mathbf{X}_0, U_i^L} \mathbf{J}(\mathbf{Y}_{i,\lambda} | \mathbf{X}_0, U_i^L, V) \right) \right)
\end{aligned}$$

$$\begin{aligned}
& - \sum_{i=2}^{L-1} \frac{\alpha_i}{2\lambda} \left(m - \text{tr} \left(\hat{\mathbf{D}}_i \mathbf{J}(\mathbf{X}_{i,\lambda} | \mathbf{X}_0, U_i, V) \right) \right) \\
& \quad + \frac{\mu_{L-1}}{2\lambda} \left(m - \text{tr} \left(\hat{\mathbf{D}}_{\mathbf{X}_L | U_L} \mathbf{J}(\tilde{\mathbf{X}}_{L,\lambda} | U_L, V) \right) \right) \\
& \quad - \frac{\mu_L}{2\lambda} \left(m - \text{tr} \left(\hat{\mathbf{D}}_L \mathbf{J}(\mathbf{X}_{L,\lambda} | \mathbf{X}_0, U_L, V) \right) \right)
\end{aligned}$$

for $\lambda \in (0, 1)$. Define

$$\begin{aligned}
f_\lambda & = \sum_{i=2}^{L-1} \mu_{i-1} \text{tr} \left(\mathbf{B}_i \hat{\mathbf{D}}_{\mathbf{Y} | \mathbf{X}_0, U_i^L}^{\frac{1}{2}} \mathbf{J}(\mathbf{Y}_{i,\lambda} | \mathbf{X}_0, U_i^L, V) \hat{\mathbf{D}}_{\mathbf{Y} | \mathbf{X}_0, U_i^L}^{\frac{1}{2}} \right) \\
& \quad - \sum_{i=2}^{L-1} \alpha_i \text{tr} \left(\hat{\mathbf{D}}_i \mathbf{J}(\mathbf{X}_{i,\lambda} | \mathbf{X}_0, U_i, V) \right) \\
& \quad + \mu_{L-1} \text{tr} \left(\hat{\mathbf{D}}_{\mathbf{X}_L | U_L} \mathbf{J}(\tilde{\mathbf{X}}_{L,\lambda} | U_L, V) \right) \\
& \quad - \mu_L \text{tr} \left(\hat{\mathbf{D}}_L \mathbf{J}(\mathbf{X}_{L,\lambda} | \mathbf{X}_0, U_L, V) \right), \\
f_{i,\lambda} & = -\text{tr} \left((\mathbf{B}_i + \mathbf{I}_m) \hat{\mathbf{D}}_{\mathbf{Y} | \mathbf{X}_0, U_i^L}^{\frac{1}{2}} \mathbf{J}(\mathbf{Y}_{i,\lambda} | \mathbf{X}_0, U_i^L, V) \hat{\mathbf{D}}_{\mathbf{Y} | \mathbf{X}_0, U_i^L}^{\frac{1}{2}} \right) \\
& \quad + \text{tr} \left(\hat{\mathbf{D}}_{\mathbf{Y} | U_i^L} \mathbf{J}(\tilde{\mathbf{Y}}_{i,\lambda} | U_i^L, V) \right), \quad i = 2, \dots, L-1,
\end{aligned}$$

where

$$\mathbf{B}_i = \hat{\mathbf{D}}_{\mathbf{Y} | \mathbf{X}_0, U_i^L}^{\frac{1}{2}} \Sigma_{\tilde{\mathbf{N}}}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0 | U_i^L} \Sigma_{\tilde{\mathbf{N}}}^{-1} \hat{\mathbf{D}}_{\mathbf{Y} | \mathbf{X}_0, U_i^L}^{\frac{1}{2}}, \quad i = 2, \dots, L-1.$$

Note that

$$\frac{d\tilde{h}_\lambda}{d\lambda} = -\frac{1}{2\lambda} \left(m \sum_{i=2}^L \alpha_i + f_\lambda + \sum_{i=2}^{L-1} \mu_{i-1} f_{i,\lambda} \right)$$

for $\lambda \in (0, 1)$. As a consequence, for the purpose of proving (22), it suffices to establish the following two lemmas.

Lemma 6: For $\lambda \in (0, 1)$,

$$\sum_{i=2}^{L-1} \mu_{i-1} \text{tr}(\mathbf{B}_i) - m \sum_{i=2}^L \alpha_i - f_\lambda \geq 0.$$

Lemma 7: For $\lambda \in (0, 1)$,

$$\text{tr}(\mathbf{B}_i) + f_{i,\lambda} \leq 0, \quad i = 2, \dots, L-1.$$

D. Proof of Lemma 6

Note that

$$\begin{aligned}
& \mathbf{J}(\mathbf{Y}_{i,\lambda} | \mathbf{X}_0, U_i^L, V) \\
& = \mathbf{J} \left(\sqrt{\lambda} \sum_{j=2}^L \Sigma_{\tilde{\mathbf{N}}} \Sigma_{\mathbf{N}_j}^{-1} \mathbf{X}_j + \sqrt{1-\lambda} \Sigma_{\tilde{\mathbf{N}}} \right. \\
& \quad \cdot \left. \left(\sum_{j=2}^{i-1} \Sigma_{\mathbf{N}_j}^{-1} \mathbf{N}'_j + \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} \mathbf{X}'_j \right) \middle| \mathbf{X}_0, U_i^L, V \right) \\
& \leq \hat{\mathbf{D}}_{\mathbf{Y} | \mathbf{X}_0, U_i^L}^{-1} \left(\sum_{j=2}^{i-1} \Sigma_{\tilde{\mathbf{N}}} \Sigma_{\mathbf{N}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}} \right)
\end{aligned}$$

$$\begin{aligned}
& \cdot \mathbf{J} \left(\sum_{j=2}^{i-1} \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}_j}^{-1} (\sqrt{\lambda} \mathbf{X}_j + \sqrt{1-\lambda} \mathbf{N}_j) \middle| \mathbf{X}_0, U_i^L, V \right) \\
& \cdot \left(\sum_{j=2}^{i-1} \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}} \right) \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0, U_i^L}^{-1} \\
& + \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0, U_i^L}^{-1} \sum_{j=i}^L \left(\Sigma_{\tilde{\mathbf{N}}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \hat{\mathbf{D}}_j \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}} \right. \\
& \cdot \mathbf{J}(\Sigma_{\tilde{\mathbf{N}}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \mathbf{X}_{j,\lambda} | \mathbf{X}_0, U_j, V) \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \hat{\mathbf{D}}_j \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}} \left. \right) \\
& \cdot \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0, U_i^L}^{-1} \tag{23} \\
& = \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0, U_i^L}^{-1} \left(\sum_{j=2}^{i-1} \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}} \right) \mathbf{J} \left(\sum_{j=2}^{i-1} \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \mathbf{N}_j \right) \\
& \cdot \left(\sum_{j=2}^{i-1} \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}} \right) \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0, U_i^L}^{-1} + \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0, U_i^L}^{-1} \\
& \cdot \sum_{j=i}^L \left(\Sigma_{\tilde{\mathbf{N}}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \hat{\mathbf{D}}_j \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}} \mathbf{J}(\Sigma_{\tilde{\mathbf{N}}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \mathbf{X}_{j,\lambda} | \mathbf{X}_0, U_j, V) \right. \\
& \quad \cdot \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \hat{\mathbf{D}}_j \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}} \left. \right) \cdot \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0, U_i^L}^{-1} \\
& = \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0, U_i^L}^{-1} \left(\sum_{j=2}^{i-1} \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}} \right) \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0, U_i^L}^{-1} \\
& + \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0, U_i^L}^{-1} \sum_{j=i}^L \left(\Sigma_{\tilde{\mathbf{N}}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \hat{\mathbf{D}}_j \mathbf{J}(\mathbf{X}_{j,\lambda} | \mathbf{X}_0, U_j, V) \right. \\
& \quad \cdot \hat{\mathbf{D}}_j \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}} \left. \right) \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0, U_i^L}^{-1}, \\
& \quad \quad \quad i = 2, \dots, L-1,
\end{aligned}$$

where (23) is due to Lemma 3. Therefore,

$$\begin{aligned}
& \text{tr} \left(\mathbf{B}_i \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0, U_i^L}^{-1} \mathbf{J}(\mathbf{Y}_{i,\lambda} | \mathbf{X}_0, U_i^L, V) \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0, U_i^L}^{-1} \right) \\
& \leq \sum_{j=2}^{i-1} \text{tr} \left(\Sigma_{\tilde{\mathbf{N}}_j}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_j^L} \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}} \right) \\
& + \sum_{j=i}^L \text{tr} \left(\hat{\mathbf{D}}_j \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_j^L} \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \hat{\mathbf{D}}_j \mathbf{J}(\mathbf{X}_{j,\lambda} | \mathbf{X}_0, U_j, V) \right), \\
& \quad \quad \quad i = 2, \dots, L-1,
\end{aligned}$$

which implies

$$\begin{aligned}
f_\lambda & \leq \sum_{i=2}^{L-1} \sum_{j=2}^{i-1} \mu_{i-1} \text{tr} \left(\Sigma_{\tilde{\mathbf{N}}_j}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_j^L} \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}} \right) \\
& + \sum_{i=2}^{L-1} \sum_{j=i}^L \mu_{i-1} \text{tr} \left(\hat{\mathbf{D}}_j \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_j^L} \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \hat{\mathbf{D}}_j \right. \\
& \quad \quad \quad \cdot \mathbf{J}(\mathbf{X}_{j,\lambda} | \mathbf{X}_0, U_j, V) \left. \right) \\
& - \sum_{i=2}^{L-1} \alpha_i \text{tr} \left(\hat{\mathbf{D}}_i \mathbf{J}(\mathbf{X}_{i,\lambda} | \mathbf{X}_0, U_i, V) \right)
\end{aligned}$$

$$\begin{aligned}
& + \mu_{L-1} \text{tr} \left(\hat{\mathbf{D}}_{\mathbf{X}_L|U_L} \mathbf{J}(\tilde{\mathbf{X}}_{L,\lambda} | U_L, V) \right) \\
& - \mu_L \text{tr} \left(\hat{\mathbf{D}}_L \mathbf{J}(\mathbf{X}_{L,\lambda} | \mathbf{X}_0, U_L, V) \right) \\
& = \sum_{i=2}^{L-1} \sum_{j=2}^{i-1} \mu_{i-1} \text{tr} \left(\Sigma_{\tilde{\mathbf{N}}_j}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_j^L} \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}} \right) \\
& + \sum_{i=2}^{L-1} \sum_{j=i}^L \mu_{i-1} \text{tr} \left(\hat{\mathbf{D}}_j \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_j^L} \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \hat{\mathbf{D}}_j \right. \\
& \quad \quad \quad \cdot \mathbf{J}(\mathbf{X}_{j,\lambda} | \mathbf{X}_0, U_j, V) \left. \right) \\
& + \text{tr} \left(\left(\sum_{i=2}^{L-1} \mu_{i-1} \hat{\mathbf{D}}_L \Sigma_{\tilde{\mathbf{N}}_L}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_i^L} \Sigma_{\tilde{\mathbf{N}}_L}^{-1} \hat{\mathbf{D}}_L - \mu_L \hat{\mathbf{D}}_L \right) \right. \\
& \quad \quad \quad \cdot \mathbf{J}(\mathbf{X}_{L,\lambda} | \mathbf{X}_0, U_L, V) \left. \right) \\
& - \sum_{i=2}^{L-1} \alpha_i \text{tr} \left(\hat{\mathbf{D}}_i \mathbf{J}(\mathbf{X}_{i,\lambda} | \mathbf{X}_0, U_i, V) \right) \\
& + \mu_{L-1} \text{tr} \left(\hat{\mathbf{D}}_{\mathbf{X}_L|U_L} \mathbf{J}(\tilde{\mathbf{X}}_{L,\lambda} | U_L, V) \right) \\
& = \sum_{i=2}^{L-1} \sum_{j=2}^{i-1} \mu_{i-1} \text{tr} \left(\Sigma_{\tilde{\mathbf{N}}_j}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_j^L} \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}} \right) \\
& + \sum_{i=2}^{L-1} \text{tr} \left(\left(\sum_{j=2}^i \mu_{j-1} \hat{\mathbf{D}}_i \Sigma_{\tilde{\mathbf{N}}_i}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_j^L} \Sigma_{\tilde{\mathbf{N}}_i}^{-1} \hat{\mathbf{D}}_i - \alpha_i \hat{\mathbf{D}}_i \right) \right. \\
& \quad \quad \quad \cdot \mathbf{J}(\mathbf{X}_{i,\lambda} | \mathbf{X}_0, U_i, V) \left. \right) \\
& + \text{tr} \left(\left(\sum_{i=2}^{L-1} \mu_{i-1} \hat{\mathbf{D}}_L \Sigma_{\tilde{\mathbf{N}}_L}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_i^L} \Sigma_{\tilde{\mathbf{N}}_L}^{-1} \hat{\mathbf{D}}_L - \mu_L \hat{\mathbf{D}}_L \right) \right. \\
& \quad \quad \quad \cdot \mathbf{J}(\mathbf{X}_{L,\lambda} | \mathbf{X}_0, U_L, V) \left. \right) \\
& + \mu_{L-1} \text{tr} \left(\hat{\mathbf{D}}_{\mathbf{X}_L|U_L} \mathbf{J}(\tilde{\mathbf{X}}_{L,\lambda} | U_L, V) \right). \tag{24}
\end{aligned}$$

We first bound the second term in (24). Note that

$$\begin{aligned}
& \sum_{j=2}^i \mu_{j-1} \hat{\mathbf{D}}_i \Sigma_{\tilde{\mathbf{N}}_i}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_j^L} \Sigma_{\tilde{\mathbf{N}}_i}^{-1} \hat{\mathbf{D}}_i - \alpha_i \hat{\mathbf{D}}_i \\
& = \hat{\mathbf{D}}_i \left(\sum_{j=2}^i \mu_{j-1} \Sigma_{\tilde{\mathbf{N}}_i}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_j^L} \Sigma_{\tilde{\mathbf{N}}_i}^{-1} - \alpha_i \mathbf{D}_i^{-1} \right) \hat{\mathbf{D}}_i \\
& = -\hat{\mathbf{D}}_i \mathbf{O}_i \hat{\mathbf{D}}_i \\
& \leq \mathbf{0}, \quad i = 2, \dots, L-1, \tag{25}
\end{aligned}$$

where (25) is due to (19); moreover, we have

$$\begin{aligned}
& \mathbf{J}(\mathbf{X}_{i,\lambda} | \mathbf{X}_0, U_i, V) \\
& \geq (\lambda \Sigma_{\mathbf{X}_i|\mathbf{X}_0, U_i, V} + (1-\lambda) \hat{\mathbf{D}}_i)^{-1} \\
& \geq (\lambda \mathbf{D}_i + (1-\lambda) \hat{\mathbf{D}}_i)^{-1} \\
& = (\hat{\mathbf{D}}_i + \lambda(\mathbf{D}_i - \hat{\mathbf{D}}_i))^{-1} \tag{26}
\end{aligned}$$

$$\begin{aligned}
&= \hat{\mathbf{D}}_i^{-1} - \hat{\mathbf{D}}_i^{-1} \left(\hat{\mathbf{D}}_i^{-1} + \frac{1}{\lambda} (\mathbf{D}_i - \hat{\mathbf{D}}_i)^{-1} \right)^{-1} \hat{\mathbf{D}}_i^{-1} && \leq \left(\mathbf{D}_L^{-1} + \frac{\lambda}{1-\lambda} \hat{\mathbf{D}}_L^{-1} \right)^{-1}, \quad (29) \\
&= \hat{\mathbf{D}}_i^{-1} - \lambda (\hat{\mathbf{D}}_i + \lambda (\mathbf{D}_i - \hat{\mathbf{D}}_i))^{-1} (\mathbf{D}_i - \hat{\mathbf{D}}_i) \hat{\mathbf{D}}_i^{-1}, \\
& \quad \quad \quad i = 2, \dots, L-1,
\end{aligned}$$

where (26) follows from Lemma 4. Therefore,

$$\begin{aligned}
&\text{tr} \left(\left(\sum_{j=2}^i \mu_{j-1} \hat{\mathbf{D}}_i \Sigma_{\mathbf{N}_i}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_j^L} \Sigma_{\mathbf{N}_i}^{-1} \hat{\mathbf{D}}_i - \alpha_i \hat{\mathbf{D}}_i \right) \right. \\
&\quad \left. \cdot \mathbf{J}(\mathbf{X}_{i,\lambda} | \mathbf{X}_0, U_i, V) \right) \\
&\leq \text{tr} \left(\left(\sum_{j=2}^i \mu_{j-1} \hat{\mathbf{D}}_i \Sigma_{\mathbf{N}_i}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_j^L} \Sigma_{\mathbf{N}_i}^{-1} \hat{\mathbf{D}}_i - \alpha_i \hat{\mathbf{D}}_i \right) \right. \\
&\quad \left. \cdot (\hat{\mathbf{D}}_i^{-1} - \lambda (\hat{\mathbf{D}}_i + \lambda (\mathbf{D}_i - \hat{\mathbf{D}}_i))^{-1} (\mathbf{D}_i - \hat{\mathbf{D}}_i) \hat{\mathbf{D}}_i^{-1}) \right) \\
&= \text{tr} \left(\sum_{j=2}^i \mu_{j-1} \hat{\mathbf{D}}_i \Sigma_{\mathbf{N}_i}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_j^L} \Sigma_{\mathbf{N}_i}^{-1} - \alpha_i \mathbf{I}_m \right) \\
&\quad - \lambda \text{tr} \left(\left(\sum_{j=2}^i \mu_{j-1} \hat{\mathbf{D}}_i \Sigma_{\mathbf{N}_i}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_j^L} \Sigma_{\mathbf{N}_i}^{-1} \hat{\mathbf{D}}_i - \alpha_i \hat{\mathbf{D}}_i \right) \right. \\
&\quad \left. \cdot (\hat{\mathbf{D}}_i + \lambda (\mathbf{D}_i - \hat{\mathbf{D}}_i))^{-1} (\mathbf{D}_i - \hat{\mathbf{D}}_i) \hat{\mathbf{D}}_i^{-1} \right) \\
&= \text{tr} \left(\sum_{j=2}^i \mu_{j-1} \hat{\mathbf{D}}_i \Sigma_{\mathbf{N}_i}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_j^L} \Sigma_{\mathbf{N}_i}^{-1} - \alpha_i \mathbf{I}_m \right) - \lambda \\
&\quad \cdot \text{tr} \left((\mathbf{D}_i - \hat{\mathbf{D}}_i) \left(\sum_{j=2}^i \mu_{j-1} \Sigma_{\mathbf{N}_i}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_j^L} \Sigma_{\mathbf{N}_i}^{-1} - \alpha_i \hat{\mathbf{D}}_i^{-1} \right) \right. \\
&\quad \left. \cdot \hat{\mathbf{D}}_i (\hat{\mathbf{D}}_i + \lambda (\mathbf{D}_i - \hat{\mathbf{D}}_i))^{-1} \right) \\
&= \text{tr} \left(\sum_{j=2}^i \mu_{j-1} \hat{\mathbf{D}}_i \Sigma_{\mathbf{N}_i}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_j^L} \Sigma_{\mathbf{N}_i}^{-1} - \alpha_i \mathbf{I}_m \right) \\
&\quad + \lambda \text{tr} \left((\mathbf{D}_i - \hat{\mathbf{D}}_i) \mathbf{O}_i \hat{\mathbf{D}}_i (\hat{\mathbf{D}}_i + \lambda (\mathbf{D}_i - \hat{\mathbf{D}}_i))^{-1} \right) \quad (27) \\
&= \text{tr} \left(\sum_{j=2}^i \mu_{j-1} \hat{\mathbf{D}}_i \Sigma_{\mathbf{N}_i}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_j^L} \Sigma_{\mathbf{N}_i}^{-1} - \alpha_i \mathbf{I}_m \right), \\
&\quad \quad \quad i = 2, \dots, L-1, \quad (28)
\end{aligned}$$

where (27) and (28) are due to (19) and (21), respectively.

Next we bound the last two terms in (24). Define

$$\begin{aligned}
\mathbf{D}_{L,\lambda} &= \Sigma_{\mathbf{X}_L | \mathbf{X}_0, \mathbf{X}_{L,\lambda}, U_L, V}, \\
\tilde{\mathbf{D}}_{L,\lambda} &= \Sigma_{\mathbf{X}_L | \tilde{\mathbf{X}}_{L,\lambda}, U_L, V}.
\end{aligned}$$

We have

$$\mathbf{D}_{L,\lambda} \leq \left(\Sigma_{\mathbf{X}_L | \mathbf{X}_0, U_L, V}^{-1} + \frac{\lambda}{1-\lambda} \hat{\mathbf{D}}_L^{-1} \right)^{-1}$$

where the first “ \leq ” follows from the fact that its right-hand side is the error covariance matrix incurred by the linear MMSE estimator of \mathbf{X}_L from $(\mathbb{E}[\mathbf{X}_L | \mathbf{X}_0, U_L, V], \mathbf{X}_{L,\lambda})$ (which is a function of $(\mathbf{X}_0, \mathbf{X}_{L,\lambda}, U_L, V)$). Moreover, it is shown in Appendix H that

$$\tilde{\mathbf{D}}_{L,\lambda} \geq \left(\mathbf{D}_{L,\lambda}^{-1} - \frac{1}{1-\lambda} \Sigma_{\mathbf{N}_L}^{-1} \Sigma_{\tilde{\mathbf{N}}_L} \Sigma_{\mathbf{N}_L}^{-1} \right)^{-1}. \quad (30)$$

Note that

$$\begin{aligned}
&\mathbf{J}(\mathbf{X}_{L,\lambda} | \mathbf{X}_0, U_L, V) \\
&= \mathbf{J} \left(\sqrt{1-\lambda} \hat{\mathbf{D}}_L^{\frac{1}{2}} \right. \\
&\quad \left. \cdot \left(\sqrt{\frac{\lambda}{1-\lambda}} \hat{\mathbf{D}}_L^{-\frac{1}{2}} \mathbf{X}_L + \hat{\mathbf{D}}_L^{-\frac{1}{2}} \mathbf{X}'_L \right) \middle| \mathbf{X}_0, U_L, V \right) \\
&= \frac{1}{1-\lambda} \hat{\mathbf{D}}_L^{-\frac{1}{2}} \\
&\quad \cdot \mathbf{J} \left(\sqrt{\frac{\lambda}{1-\lambda}} \hat{\mathbf{D}}_L^{-\frac{1}{2}} \mathbf{X}_L + \hat{\mathbf{D}}_L^{-\frac{1}{2}} \mathbf{X}'_L \middle| \mathbf{X}_0, U_L, V \right) \hat{\mathbf{D}}_L^{-\frac{1}{2}} \\
&= \frac{1}{1-\lambda} \hat{\mathbf{D}}_L^{-\frac{1}{2}} \left(\mathbf{I}_m - \frac{\lambda}{1-\lambda} \hat{\mathbf{D}}_L^{-\frac{1}{2}} \right. \\
&\quad \left. \cdot \Sigma_{\mathbf{X}_L | \mathbf{X}_0, \sqrt{\frac{\lambda}{1-\lambda}} \hat{\mathbf{D}}_L^{-\frac{1}{2}} \mathbf{X}_L + \hat{\mathbf{D}}_L^{-\frac{1}{2}} \mathbf{X}'_L, U_L, V} \hat{\mathbf{D}}_L^{-\frac{1}{2}} \right) \hat{\mathbf{D}}_L^{-\frac{1}{2}} \quad (31)
\end{aligned}$$

$$= \frac{1}{1-\lambda} \hat{\mathbf{D}}_L^{-1} - \frac{\lambda}{(1-\lambda)^2} \hat{\mathbf{D}}_L^{-1} \mathbf{D}_{L,\lambda} \hat{\mathbf{D}}_L^{-1}, \quad (32)$$

where (31) follows from Lemma 5; similarly, we have

$$\begin{aligned}
&\mathbf{J}(\tilde{\mathbf{X}}_{L,\lambda} | U_L, V) \\
&= \frac{1}{1-\lambda} \hat{\mathbf{D}}_{\mathbf{X}_L | U_L}^{-1} - \frac{\lambda}{(1-\lambda)^2} \hat{\mathbf{D}}_{\mathbf{X}_L | U_L}^{-1} \tilde{\mathbf{D}}_{L,\lambda} \hat{\mathbf{D}}_{\mathbf{X}_L | U_L}^{-1}.
\end{aligned}$$

As a consequence,

$$\begin{aligned}
&\text{tr} \left(\left(\sum_{i=2}^{L-1} \mu_{i-1} \hat{\mathbf{D}}_L \Sigma_{\mathbf{N}_L}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_i^L} \Sigma_{\mathbf{N}_L}^{-1} \hat{\mathbf{D}}_L - \mu_L \hat{\mathbf{D}}_L \right) \right. \\
&\quad \left. \cdot \mathbf{J}(\mathbf{X}_{L,\lambda} | \mathbf{X}_0, U_L, V) \right) \\
&\quad + \mu_{L-1} \text{tr} \left(\hat{\mathbf{D}}_{\mathbf{X}_L | U_L} \mathbf{J}(\tilde{\mathbf{X}}_{L,\lambda} | U_L, V) \right) \\
&= \frac{1}{1-\lambda} \text{tr} \left(\sum_{i=2}^{L-1} \mu_{i-1} \hat{\mathbf{D}}_L \Sigma_{\mathbf{N}_L}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_i^L} \Sigma_{\mathbf{N}_L}^{-1} - \alpha_L \mathbf{I}_m \right) \\
&\quad - \frac{\lambda}{(1-\lambda)^2} \text{tr} \left(\left(\sum_{i=2}^{L-1} \mu_{i-1} \Sigma_{\mathbf{N}_L}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_i^L} \Sigma_{\mathbf{N}_L}^{-1} - \mu_L \hat{\mathbf{D}}_L^{-1} \right) \right. \\
&\quad \left. \cdot \mathbf{D}_{L,\lambda} + \mu_{L-1} \hat{\mathbf{D}}_{\mathbf{X}_L | U_L}^{-1} \tilde{\mathbf{D}}_{L,\lambda} \right) \\
&\leq \frac{1}{1-\lambda} \text{tr} \left(\sum_{i=2}^{L-1} \mu_{i-1} \hat{\mathbf{D}}_L \Sigma_{\mathbf{N}_L}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_i^L} \Sigma_{\mathbf{N}_L}^{-1} - \alpha_L \mathbf{I}_m \right)
\end{aligned}$$

$$\begin{aligned}
& -\frac{\lambda}{(1-\lambda)^2} \\
& \cdot \text{tr} \left(\left(\sum_{i=2}^{L-1} \mu_{i-1} \Sigma_{\mathbf{N}_L}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_i^L} \Sigma_{\mathbf{N}_L}^{-1} - \mu_L \hat{\mathbf{D}}_L^{-1} \right) \mathbf{D}_{L,\lambda} \right) \\
& + \mu_{L-1} \hat{\mathbf{D}}_{\mathbf{X}_L|U_L}^{-1} \left(\mathbf{D}_{L,\lambda}^{-1} - \frac{1}{1-\lambda} \Sigma_{\mathbf{N}_L}^{-1} \Sigma_{\tilde{\mathbf{N}}_L} \Sigma_{\mathbf{N}_L}^{-1} \right)^{-1}, \quad (33)
\end{aligned}$$

where (33) is due to (30). Now consider the following optimization problem

$$\begin{aligned}
(\mathbf{P}_2) \min_{\mathbf{D}_{L,\lambda}} & \text{tr} \left(\left(\sum_{i=2}^{L-1} \mu_{i-1} \Sigma_{\mathbf{N}_L}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_i^L} \Sigma_{\mathbf{N}_L}^{-1} - \mu_L \hat{\mathbf{D}}_L^{-1} \right) \mathbf{D}_{L,\lambda} \right) \\
& + \mu_{L-1} \hat{\mathbf{D}}_{\mathbf{X}_L|U_L}^{-1} \left(\mathbf{D}_{L,\lambda}^{-1} - \frac{1}{1-\lambda} \Sigma_{\mathbf{N}_L}^{-1} \Sigma_{\tilde{\mathbf{N}}_L} \Sigma_{\mathbf{N}_L}^{-1} \right)^{-1} \\
\text{subject to } & \mathbf{0} \leq \mathbf{D}_{L,\lambda} \leq \left(\mathbf{D}_L^{-1} + \frac{\lambda}{1-\lambda} \hat{\mathbf{D}}_L^{-1} \right)^{-1}, \quad (34)
\end{aligned}$$

where the second “ \leq ” in (34) is due to (29). We show in Appendix I that the minimum value of (\mathbf{P}_2) is given by

$$(1-\lambda) \text{tr} \left(\sum_{i=2}^{L-1} \mu_{i-1} \hat{\mathbf{D}}_L \Sigma_{\mathbf{N}_L}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_i^L} \Sigma_{\mathbf{N}_L}^{-1} - \alpha_L \mathbf{I}_m \right). \quad (35)$$

In view of (33) and (35), we have

$$\begin{aligned}
& \text{tr} \left(\left(\sum_{i=2}^{L-1} \mu_{i-1} \hat{\mathbf{D}}_L \Sigma_{\mathbf{N}_L}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_i^L} \Sigma_{\mathbf{N}_L}^{-1} \hat{\mathbf{D}}_L - \mu_L \hat{\mathbf{D}}_L \right) \right. \\
& \quad \left. \cdot \mathbf{J}(\mathbf{X}_{L,\lambda} | \mathbf{X}_0, U_L, V) \right) \\
& + \mu_{L-1} \text{tr} \left(\hat{\mathbf{D}}_{\mathbf{X}_L|U_L} \mathbf{J}(\tilde{\mathbf{X}}_{L,\lambda} | U_L, V) \right) \\
& \leq \text{tr} \left(\sum_{i=2}^{L-1} \mu_{i-1} \hat{\mathbf{D}}_L \Sigma_{\mathbf{N}_L}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_i^L} \Sigma_{\mathbf{N}_L}^{-1} - \alpha_L \mathbf{I}_m \right). \quad (36)
\end{aligned}$$

Substituting (28) and (36) into (24) yields

$$\begin{aligned}
f\lambda & \leq \sum_{i=2}^{L-1} \sum_{j=2}^{i-1} \mu_{i-1} \text{tr} \left(\Sigma_{\tilde{\mathbf{N}}_j}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_i^L} \Sigma_{\mathbf{N}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}_j} \right) \\
& + \sum_{i=2}^{L-1} \text{tr} \left(\sum_{j=2}^i \mu_{j-1} \hat{\mathbf{D}}_i \Sigma_{\mathbf{N}_i}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_j^L} \Sigma_{\mathbf{N}_i}^{-1} - \alpha_i \mathbf{I}_m \right) \\
& + \text{tr} \left(\sum_{i=2}^{L-1} \mu_{i-1} \hat{\mathbf{D}}_L \Sigma_{\mathbf{N}_L}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_i^L} \Sigma_{\mathbf{N}_L}^{-1} - \alpha_L \mathbf{I}_m \right) \\
& = \sum_{i=2}^{L-1} \sum_{j=2}^{i-1} \mu_{i-1} \text{tr} \left(\Sigma_{\tilde{\mathbf{N}}_j}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_i^L} \Sigma_{\mathbf{N}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}_j} \right) \\
& + \sum_{j=2}^{L-1} \sum_{i=2}^j \mu_{i-1} \text{tr} \left(\hat{\mathbf{D}}_j \Sigma_{\mathbf{N}_j}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_i^L} \Sigma_{\mathbf{N}_j}^{-1} \right) \\
& + \sum_{i=2}^{L-1} \mu_{i-1} \text{tr} \left(\hat{\mathbf{D}}_L \Sigma_{\mathbf{N}_L}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_i^L} \Sigma_{\mathbf{N}_L}^{-1} \right) - m \sum_{i=2}^L \alpha_i
\end{aligned}$$

$$\begin{aligned}
& = \sum_{i=2}^{L-1} \sum_{j=2}^{i-1} \mu_{i-1} \text{tr} \left(\Sigma_{\tilde{\mathbf{N}}_j}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_i^L} \Sigma_{\mathbf{N}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}_j} \right) \\
& + \sum_{i=2}^{L-1} \sum_{j=i}^{L-1} \mu_{i-1} \text{tr} \left(\hat{\mathbf{D}}_j \Sigma_{\mathbf{N}_j}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_i^L} \Sigma_{\mathbf{N}_j}^{-1} \right) \\
& + \sum_{i=2}^{L-1} \mu_{i-1} \text{tr} \left(\hat{\mathbf{D}}_L \Sigma_{\mathbf{N}_L}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_i^L} \Sigma_{\mathbf{N}_L}^{-1} \right) - m \sum_{i=2}^L \alpha_i \\
& = \sum_{i=2}^{L-1} \mu_{i-1} \text{tr} \left(\sum_{j=2}^{i-1} \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_i^L} \Sigma_{\mathbf{N}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}_j} \right. \\
& \quad \left. + \sum_{j=i}^L \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_i^L} \Sigma_{\mathbf{N}_j}^{-1} \hat{\mathbf{D}}_j \Sigma_{\mathbf{N}_j}^{-1} \Sigma_{\tilde{\mathbf{N}}_j} \right) - m \sum_{i=2}^L \alpha_i \\
& = \sum_{i=2}^{L-1} \mu_{i-1} \text{tr} \left(\Sigma_{\tilde{\mathbf{N}}_i}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_i^L} \Sigma_{\tilde{\mathbf{N}}_i}^{-1} \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0, U_i^L} \right) - m \sum_{i=2}^L \alpha_i \\
& = \sum_{i=2}^{L-1} \mu_{i-1} \text{tr}(\mathbf{B}_i) - m \sum_{i=2}^L \alpha_i,
\end{aligned}$$

which proves Lemma 6.

E. Proof of Lemma 7

Define

$$\mathbf{D}_{\mathbf{Y},i,\lambda} = \Sigma_{\mathbf{Y}|\mathbf{X}_0, \mathbf{Y}_{i,\lambda}, U_i^L, V}, \quad i = 2, \dots, L-1,$$

$$\tilde{\mathbf{D}}_{\mathbf{Y},i,\lambda} = \Sigma_{\mathbf{Y}|\tilde{\mathbf{Y}}_{i,\lambda}, U_i^L, V}, \quad i = 2, \dots, L-1.$$

Following steps similar to those in Appendix H, one can show that

$$\mathbf{D}_{\mathbf{Y},i,\lambda} \leq \Theta_{i,\lambda} \triangleq \left(\tilde{\mathbf{D}}_{\mathbf{Y},i,\lambda}^{-1} + \frac{1}{1-\lambda} \Sigma_{\tilde{\mathbf{N}}_i}^{-1} \right)^{-1}, \quad i = 2, \dots, L-1. \quad (37)$$

Moreover, it can be verified (cf. the derivation of (32)) that

$$\begin{aligned}
& \mathbf{J}(\mathbf{Y}_{i,\lambda} | \mathbf{X}_0, U_i^L, V) \\
& = \frac{1}{1-\lambda} \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0, U_i^L}^{-1} - \frac{\lambda}{(1-\lambda)^2} \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0, U_i^L}^{-1} \mathbf{D}_{\mathbf{Y},i,\lambda} \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0, U_i^L}^{-1}, \\
& \quad i = 2, \dots, L-1, \quad (38)
\end{aligned}$$

$$\begin{aligned}
& \mathbf{J}(\tilde{\mathbf{Y}}_{i,\lambda} | U_i^L, V) \\
& = \frac{1}{1-\lambda} \hat{\mathbf{D}}_{\mathbf{Y}|U_i^L}^{-1} - \frac{\lambda}{(1-\lambda)^2} \hat{\mathbf{D}}_{\mathbf{Y}|U_i^L}^{-1} \tilde{\mathbf{D}}_{\mathbf{Y},i,\lambda} \hat{\mathbf{D}}_{\mathbf{Y}|U_i^L}^{-1}, \\
& \quad i = 2, \dots, L-1. \quad (39)
\end{aligned}$$

In view of (37)-(39), we have

$$\begin{aligned}
f_{i,\lambda} & = -\frac{1}{1-\lambda} \text{tr}(\mathbf{B}_i) - \frac{\lambda}{(1-\lambda)^2} \\
& \quad \cdot \text{tr} \left(\tilde{\mathbf{D}}_{\mathbf{Y},i,\lambda} \hat{\mathbf{D}}_{\mathbf{Y}|U_i^L}^{-1} - (\mathbf{B}_i + \mathbf{I}_m) \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0, U_i^L}^{-\frac{1}{2}} \mathbf{D}_{\mathbf{Y},i,\lambda} \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0, U_i^L}^{-\frac{1}{2}} \right) \\
& \leq -\frac{1}{1-\lambda} \text{tr}(\mathbf{B}_i) - \frac{\lambda}{(1-\lambda)^2} g_{i,\lambda}, \quad i = 2, \dots, L-1,
\end{aligned}$$

where

$$g_{i,\lambda} = \text{tr} \left(\tilde{\mathbf{D}}_{\mathbf{Y},i,\lambda} \hat{\mathbf{D}}_{\mathbf{Y}|U_i^L}^{-1} - \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0, U_i^L}^{-\frac{1}{2}} (\mathbf{B}_i + \mathbf{I}_m) \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0, U_i^L}^{-\frac{1}{2}} \Theta_{i,\lambda} \right), \quad i = 2, \dots, L-1.$$

Therefore, in order to prove Lemma 7, one just need to show that

$$g_{i,\lambda} \geq -(1-\lambda)\text{tr}(\mathbf{B}_i), \quad i = 2, \dots, L-1.$$

Now consider the following convex optimization problem

$$(\mathbf{P}_3) \quad \min_{\tilde{\mathbf{D}}_{\mathbf{Y},i,\lambda} \geq \mathbf{0}} g_{i,\lambda}, \quad i = 2, \dots, L-1.$$

Note that

$$\begin{aligned} \nabla_{\tilde{\mathbf{D}}_{\mathbf{Y},i,\lambda}} g_{i,\lambda} &= \hat{\mathbf{D}}_{\mathbf{Y}|U_i^L}^{-1} - \tilde{\mathbf{D}}_{\mathbf{Y},i,\lambda}^{-1} \Theta_{i,\lambda} \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0,U_i^L}^{-\frac{1}{2}} \\ &\quad \cdot (\mathbf{B}_i + \mathbf{I}_m) \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0,U_i^L}^{-\frac{1}{2}} \Theta_{i,\lambda} \tilde{\mathbf{D}}_{\mathbf{Y},i,\lambda}^{-1}, \\ &\quad i = 2, \dots, L-1. \end{aligned}$$

By leveraging the fact that $\hat{\mathbf{D}}_{\mathbf{Y}|U_i^L}^{-1} + \Sigma_{\tilde{\mathbf{N}}}^{-1} = \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0,U_i^L}^{-1}$ and

$$\begin{aligned} &\hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0,U_i^L}^{-\frac{1}{2}} (\mathbf{B}_i + \mathbf{I}_m) \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0,U_i^L}^{-\frac{1}{2}} \\ &= \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0,U_i^L}^{-1} \left(\hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0,U_i^L}^{-1} - \Sigma_{\tilde{\mathbf{N}}}^{-1} \right)^{-1} \hat{\mathbf{D}}_{\mathbf{Y}|\mathbf{X}_0,U_i^L}^{-1}, \\ &\quad i = 2, \dots, L-1, \end{aligned}$$

one can readily verify that

$$\nabla_{\tilde{\mathbf{D}}_{\mathbf{Y},i,\lambda}} g_{i,\lambda} \Big|_{\tilde{\mathbf{D}}_{\mathbf{Y},i,\lambda} = (1-\lambda)\hat{\mathbf{D}}_{\mathbf{Y}|U_i^L}} = \mathbf{0}, \quad i = 2, \dots, L-1.$$

Therefore, the minimum value of (\mathbf{P}_3) is attained at $\tilde{\mathbf{D}}_{\mathbf{Y},i,\lambda} = (1-\lambda)\hat{\mathbf{D}}_{\mathbf{Y}|U_i^L}$, $i = 2, \dots, L-1$. A direct evaluation shows that

$$g_{i,\lambda} \Big|_{\tilde{\mathbf{D}}_{\mathbf{Y},i,\lambda} = (1-\lambda)\hat{\mathbf{D}}_{\mathbf{Y}|U_i^L}} = -(1-\lambda)\text{tr}(\mathbf{B}_i), \quad i = 2, \dots, L-1,$$

which completes the proof.

V. CONCLUSION

A new extremal inequality is established and is leveraged to derive a lower bound on each supporting hyperplane of the rate region of the vector Gaussian CEO problem. We have also derived a lower bound on each supporting hyperplane of the rate region of the direct vector Gaussian multiterminal source coding problem by coupling it with the CEO problem through a limiting argument. The tightness of our lower bounds is established under certain conditions. It is expected that combining our techniques with those recently developed in [30] might lead to improved lower bounds.

APPENDIX A

PROOF OF THEOREM 1

Let $f_i^{(n)} : \mathbb{R}^{m_i \times n} \rightarrow \{1, \dots, 2^{nR_i}\}$, $i = 1, \dots, L$, be L encoding functions such that $\Sigma_{\mathbf{X}_i^n | W_1, \dots, W_L} \preceq \mathbf{D}$, where $W_i = f_i^{(n)}(\mathbf{X}_i^n)$, $i = 1, \dots, L$. It follows from [20, Th. 1] that

$$\begin{aligned} \sum_{i \in \mathcal{A}} R_i &\geq I(\mathbf{X}_0; \{U_i\}_{i \in \mathcal{A}} | \{U_i\}_{i \notin \mathcal{A}}, T) \\ &\quad + \sum_{i \in \mathcal{A}} I(\mathbf{X}_i; U_i | \mathbf{X}_0, W, T), \quad \emptyset \subset \mathcal{A} \subseteq \{1, \dots, L\}, \\ \mathbf{D} &\succeq \Sigma_{\mathbf{X}_0 | U_1^L, T} \end{aligned}$$

for some (U_1, \dots, U_L, W, T) whose joint distribution with $(\mathbf{X}_0, \mathbf{X}_1, \dots, \mathbf{X}_L)$ factors as

$$\begin{aligned} &p(\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_L, u_1, \dots, u_L, w, t) \\ &= p(\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_L) p(w, t) \prod_{i=1}^L p(u_i | \mathbf{x}_i, w, t). \end{aligned}$$

Define $\mathbf{D}_0 = \Sigma_{\mathbf{X}_0 | U_1^L, T}$ and $\mathbf{D}_i = \Sigma_{\mathbf{X}_i | \mathbf{X}_0, U_i, W, T}$, $i = 1, \dots, L$. Clearly,

$$\Sigma_{\mathbf{N}} \prec \mathbf{D}_0 \preceq \Sigma_{\mathbf{X}_0}, \quad (40)$$

$$\mathbf{D}_0 \preceq \mathbf{D}, \quad (41)$$

$$\mathbf{0} \prec \mathbf{D}_i \preceq \Sigma_{\mathbf{N}_i}, \quad i = 1, \dots, L. \quad (42)$$

Moreover, it can be shown via an estimation-theoretic argument (see the derivation of [10, (4)] and [11, (9)]) that

$$\sum_{i=1}^L \Sigma_{\mathbf{N}_i}^{-1} \mathbf{D}_i \Sigma_{\mathbf{N}_i}^{-1} \preceq \Sigma_{\mathbf{N}}^{-1} - \mathbf{D}_0^{-1}. \quad (43)$$

Note that

$$\begin{aligned} \sum_{i=1}^L \alpha_i R_i &= \sum_{i=2}^L (\alpha_{i-1} - \alpha_i) \sum_{j=1}^{i-1} R_j + \alpha_L \sum_{i=1}^L R_i \\ &\geq \sum_{i=2}^L (\alpha_{i-1} - \alpha_i) \left(I(\mathbf{X}_0; U_1^{i-1} | U_i^L, T) \right. \\ &\quad \left. + \sum_{j=1}^{i-1} I(\mathbf{X}_j; U_j | \mathbf{X}_0, W, T) \right) \\ &\quad + \alpha_L \left(I(\mathbf{X}_0; U_1^L | T) + \sum_{i=1}^L I(\mathbf{X}_i; U_i | \mathbf{X}_0, W, T) \right) \\ &= \sum_{i=2}^L (\alpha_{i-1} - \alpha_i) I(\mathbf{X}_0; U_1^{i-1} | U_i^L, T) \\ &\quad + \sum_{i=1}^L \alpha_i I(\mathbf{X}_i; U_i | \mathbf{X}_0, W, T) + \alpha_L I(\mathbf{X}_0; U_1^L | T) \\ &\geq \sum_{i=2}^L (\alpha_{i-1} - \alpha_i) \theta_i I(\mathbf{X}_0; U_1^{i-1} | U_i^L, T) \\ &\quad + \sum_{i=1}^L \alpha_i I(\mathbf{X}_i; U_i | \mathbf{X}_0, W, T) + \alpha_L I(\mathbf{X}_0; U_1^L | T), \end{aligned} \quad (44)$$

where $\theta_i \in [0, 1]$, $i = 2, \dots, L$. It is easy to verify that

$$I(\mathbf{X}_0; U_1^{i-1} | U_i^L, T) \geq h(\mathbf{X}_0 | U_i^L, W, T) - \frac{1}{2} \log |2\pi e \mathbf{D}_0|, \quad i = 2, \dots, L, \quad (45)$$

$$I(\mathbf{X}_1; U_1 | \mathbf{X}_0, W, T) \geq \frac{1}{2} \log \frac{|\Sigma_{\mathbf{N}_1}|}{|\mathbf{D}_1|}, \quad (46)$$

$$\begin{aligned} I(\mathbf{X}_i; U_i | \mathbf{X}_0, W, T) &= \frac{1}{2} \log |2\pi e \Sigma_{\mathbf{N}_i}| \\ &\quad - h(\mathbf{X}_i | \mathbf{X}_0, U_i, W, T), \quad i = 2, \dots, L, \end{aligned} \quad (47)$$

$$I(\mathbf{X}_0; U_1^L | T) \geq \frac{1}{2} \log \frac{|\Sigma_{\mathbf{X}_0}|}{|\mathbf{D}_0|}. \quad (48)$$

To specify $(\theta_2, \dots, \theta_L)$, we consider the following optimization problem

$$\eta \triangleq \min_{\mathbf{D}'_2, \dots, \mathbf{D}'_L} \sum_{i=2}^L \frac{\alpha_i}{2} \log \frac{|\Sigma_{\mathbf{N}_i}|}{|\mathbf{D}'_i|} + \sum_{i=2}^L \frac{\alpha_{i-1} - \alpha_i}{2} \cdot \log^+ \frac{1}{|\mathbf{D}_0| |\Sigma_{\mathbf{X}_0}^{-1} + \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} - \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} \mathbf{D}'_j \Sigma_{\mathbf{N}_j}^{-1}|}$$

subject to $\mathbf{0} \leq \mathbf{D}'_i \leq \mathbf{D}_i$, $i = 2, \dots, L$.

For any optimal solution $(\hat{\mathbf{D}}_2, \dots, \hat{\mathbf{D}}_L)$ to this minimization problem, we must have $\hat{\mathbf{D}}_i > \mathbf{0}$, $i = 2, \dots, L$; moreover, it can be shown (see [27, Proposition 3.3.11, p. 327]) that there exist $\theta_i^* \in [0, 1]$ and $\mathbf{O}_i \geq \mathbf{0}$, $i = 2, \dots, L$, such that

$$\sum_{j=2}^i (\alpha_{j-1} - \alpha_j) \theta_j^* \Sigma_{\mathbf{N}_j}^{-1} \cdot \left(\Sigma_{\mathbf{X}_0}^{-1} + \sum_{k=j}^L \Sigma_{\mathbf{N}_k}^{-1} - \sum_{k=j}^L \Sigma_{\mathbf{N}_k}^{-1} \hat{\mathbf{D}}_k \Sigma_{\mathbf{N}_k}^{-1} \right)^{-1} \Sigma_{\mathbf{N}_i}^{-1} + \mathbf{O}_i = \alpha_i \hat{\mathbf{D}}_i^{-1},$$

$$i = 2, \dots, L, \quad (49)$$

$$\mathbf{O}_i (\mathbf{D}_i - \hat{\mathbf{D}}_i) = \mathbf{0}, \quad i = 2, \dots, L, \quad (50)$$

$$\sum_{i=2}^L \frac{\alpha_i}{2} \log \frac{|\Sigma_{\mathbf{N}_i}|}{|\hat{\mathbf{D}}_i|} + \sum_{i=2}^L \frac{(\alpha_{i-1} - \alpha_i) \theta_i^*}{2} \cdot \log \frac{1}{|\mathbf{D}_0| |\Sigma_{\mathbf{X}_0}^{-1} + \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} - \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} \hat{\mathbf{D}}_j \Sigma_{\mathbf{N}_j}^{-1}|} = \eta. \quad (51)$$

Substituting (45)-(48) into (44) and setting $\theta_i = \theta_i^*$, $i = 2, \dots, L$, gives

$$\begin{aligned} \sum_{i=1}^L \alpha_i R_i &\geq - \sum_{i=2}^L \frac{(\alpha_{i-1} - \alpha_i) \theta_i^*}{2} \log |2\pi e \mathbf{D}_0| \\ &+ \frac{\alpha_1}{2} \log \frac{|\Sigma_{\mathbf{N}_1}|}{|\mathbf{D}_1|} + \sum_{i=2}^L \frac{\alpha_i}{2} \log |2\pi e \Sigma_{\mathbf{N}_i}| \\ &+ \frac{\alpha_L}{2} \log \frac{|\Sigma_{\mathbf{X}_0}|}{|\mathbf{D}_0|} \\ &+ \sum_{i=2}^L (\alpha_{i-1} - \alpha_i) \theta_i^* h(\mathbf{X}_0 | U_i^L, W, T) \\ &- \sum_{i=2}^L \alpha_i h(\mathbf{X}_i | \mathbf{X}_0, U_i, W, T). \end{aligned} \quad (52)$$

In view of (49) and (50), it follows from Theorem 11 that

$$\begin{aligned} &\sum_{i=2}^L (\alpha_{i-1} - \alpha_i) \theta_i^* h(\mathbf{X}_0 | U_i^L, W, T) \\ &- \sum_{i=2}^L \alpha_i h(\mathbf{X}_i | \mathbf{X}_0, U_i, W, T) \\ &\geq \sum_{i=2}^L \frac{(\alpha_{i-1} - \alpha_i) \theta_i^*}{2} \end{aligned}$$

$$\begin{aligned} &\cdot \log \left| 2\pi e \left(\Sigma_{\mathbf{X}_0}^{-1} + \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} - \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} \hat{\mathbf{D}}_j \Sigma_{\mathbf{N}_j}^{-1} \right)^{-1} \right| \\ &- \sum_{i=2}^L \frac{\alpha_i}{2} \log |2\pi e \hat{\mathbf{D}}_i|. \end{aligned} \quad (53)$$

Combining (52) and (53) yields

$$\begin{aligned} \sum_{i=1}^L \alpha_i R_i &\geq \frac{\alpha_L}{2} \log \frac{|\Sigma_{\mathbf{X}_0}|}{|\mathbf{D}_0|} + \frac{\alpha_1}{2} \log \frac{|\Sigma_{\mathbf{N}_1}|}{|\mathbf{D}_1|} \\ &+ \sum_{i=2}^L \frac{\alpha_i}{2} \log \frac{|\Sigma_{\mathbf{N}_i}|}{|\hat{\mathbf{D}}_i|} + \sum_{i=2}^L \frac{(\alpha_{i-1} - \alpha_i) \theta_i^*}{2} \\ &\cdot \log \frac{1}{|\mathbf{D}_0| |\Sigma_{\mathbf{X}_0}^{-1} + \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} - \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} \hat{\mathbf{D}}_j \Sigma_{\mathbf{N}_j}^{-1}|}. \end{aligned} \quad (54)$$

In view of (40)-(43), (51), and (54), we obtain the following lower bound

$$\begin{aligned} &\sum_{i=1}^L \alpha_i R_i \\ &\geq \min_{\mathbf{D}_0, \dots, \mathbf{D}_L, \mathbf{D}'_2, \dots, \mathbf{D}'_L} \frac{\alpha_L}{2} \log \frac{|\Sigma_{\mathbf{X}_0}|}{|\mathbf{D}_0|} + \frac{\alpha_1}{2} \log \frac{|\Sigma_{\mathbf{N}_1}|}{|\mathbf{D}_1|} \\ &+ \sum_{i=2}^L \frac{\alpha_i}{2} \log \frac{|\Sigma_{\mathbf{N}_i}|}{|\mathbf{D}'_i|} + \sum_{i=2}^L \frac{\alpha_{i-1} - \alpha_i}{2} \\ &\cdot \log^+ \frac{1}{|\mathbf{D}_0| |\Sigma_{\mathbf{X}_0}^{-1} + \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} - \sum_{j=i}^L \Sigma_{\mathbf{N}_j}^{-1} \mathbf{D}'_j \Sigma_{\mathbf{N}_j}^{-1}|} \end{aligned}$$

$$\text{subject to } \Sigma_{\mathbf{N}} \prec \mathbf{D}_0 \leq \Sigma_{\mathbf{X}_0},$$

$$\mathbf{D}_0 \leq \mathbf{D},$$

$$\mathbf{0} \prec \mathbf{D}_i \leq \Sigma_{\mathbf{N}_i}, \quad i = 1, \dots, L,$$

$$\sum_{i=1}^L \Sigma_{\mathbf{N}_i}^{-1} \mathbf{D}_i \Sigma_{\mathbf{N}_i}^{-1} \leq \Sigma_{\mathbf{N}}^{-1} - \mathbf{D}_0^{-1},$$

$$\mathbf{0} \leq \mathbf{D}'_i \leq \mathbf{D}_i, \quad i = 2, \dots, L.$$

It is clear that this lower bound is equivalent to $\underline{R}(\mathbf{D}, \alpha)$, which completes the proof.

APPENDIX B

PROOF OF THEOREM 5

Let $f_i^{(n)} : \mathbb{R}^{m_i \times n} \rightarrow \{1, \dots, 2^{nR_i}\}$, $i = 1, \dots, L$, be L encoding functions such that $\Sigma_{\mathbf{X}_i^n | W_1, \dots, W_L} \leq \mathbf{D}_i$, $i = 1, \dots, L$, where $W_i = f_i^{(n)}(\mathbf{X}_i^n)$, $i = 1, \dots, L$. Denote $\Sigma_{\mathbf{X}^n | W_1, \dots, W_L}$ by \mathbf{D} . Let \mathbf{Z} be a Gaussian random vector, independent of \mathbf{X} , with mean zero and covariance matrix $\Sigma_{\mathbf{Z}} \in \mathcal{Z}(\Sigma_{\mathbf{X}})$. Define $\tilde{\mathbf{X}}_0 = \mathbb{E}[\mathbf{X} | \mathbf{X} + \mathbf{Z}]$ and $\mathbf{Z}_i = \mathbf{X}_i - \mathbb{E}[\mathbf{X}_i | \mathbf{X} + \mathbf{Z}]$, $i = 1, \dots, L$. Clearly, $\tilde{\mathbf{X}}_0$ is independent of $(\mathbf{Z}_1, \dots, \mathbf{Z}_L)$. It is also easy to verify that

$$\begin{aligned} \tilde{\mathbf{X}}_0 &= \Sigma_{\mathbf{X}} (\Sigma_{\mathbf{X}} + \Sigma_{\mathbf{Z}})^{-1} (\mathbf{X} + \mathbf{Z}), \\ \Sigma_{\tilde{\mathbf{X}}_0} &= \Sigma_{\mathbf{X}} (\Sigma_{\mathbf{X}} + \Sigma_{\mathbf{Z}})^{-1} \Sigma_{\mathbf{X}}. \end{aligned}$$

Moreover, it follows from the definition of $\mathcal{Z}(\Sigma_{\mathbf{X}})$ that $\mathbf{Z}_1, \dots, \mathbf{Z}_L$ are mutually independent and

$$\text{diag}(\Sigma_{\mathbf{Z}_1}, \dots, \Sigma_{\mathbf{Z}_L}) = (\Sigma_{\mathbf{X}}^{-1} + \Sigma_{\mathbf{Z}}^{-1})^{-1}.$$

We shall derive a lower bound on $\sum_{i=1}^L \alpha_i R_i$ by coupling the direct vector Gaussian L -terminal source coding problem with the CEO problem. Let $\tilde{\mathbf{N}}_i$ be an $m \times 1$ Gaussian random vector with mean zero and covariance matrix $\Sigma_{\tilde{\mathbf{N}}_i}$, $i = 1, \dots, L$. We assume that $\Sigma_{\tilde{\mathbf{N}}_i}$ is a block diagonal matrix with $(\Sigma_{\tilde{\mathbf{N}}_i})_i = \Sigma_{\mathbf{Z}_i}$ and $(\Sigma_{\tilde{\mathbf{N}}_i})_j = \beta \mathbf{I}_{m_j}$ for $j \neq i$; moreover, $\tilde{\mathbf{X}}_0, \tilde{\mathbf{N}}_1, \dots, \tilde{\mathbf{N}}_L$ are mutually independent. Define $\tilde{\mathbf{X}}_i = \tilde{\mathbf{X}}_0 + \tilde{\mathbf{N}}_i$, $i = 1, \dots, L$. Consider the CEO problem with $\tilde{\mathbf{X}}_0$ as the remote source and $\tilde{\mathbf{X}}_1, \dots, \tilde{\mathbf{X}}_L$ as noisy observations. It can be readily seen that $\sum_{i=1}^L \alpha_i R_i$ is lower-bounded by $R(\tilde{\mathbf{D}}, \alpha)$ associated with this CEO problem for all $\beta > 0$, where

$$\tilde{\mathbf{D}} = \Sigma_{\mathbf{X}}(\Sigma_{\mathbf{X}} + \Sigma_{\mathbf{Z}})^{-1}(\mathbf{D} + \Sigma_{\mathbf{Z}})(\Sigma_{\mathbf{X}} + \Sigma_{\mathbf{Z}})^{-1}\Sigma_{\mathbf{X}}.$$

Therefore, we have

$$\begin{aligned} \sum_{i=1}^L \alpha_i R_i &\geq \min_{\tilde{\mathbf{D}}_1, \dots, \tilde{\mathbf{D}}_L} \phi(\tilde{\mathbf{D}}_1, \dots, \tilde{\mathbf{D}}_L, \beta) \\ \text{subject to } \mathbf{0} &\leq \tilde{\mathbf{D}}_i \leq \Sigma_{\tilde{\mathbf{N}}_i}, \quad i = 1, \dots, L, \end{aligned} \quad (55)$$

$$\sum_{i=1}^L \Sigma_{\tilde{\mathbf{N}}_i}^{-1} \tilde{\mathbf{D}}_i \Sigma_{\tilde{\mathbf{N}}_i}^{-1} \leq \Sigma_{\tilde{\mathbf{X}}_0}^{-1} + \sum_{i=1}^L \Sigma_{\tilde{\mathbf{N}}_i}^{-1} - \tilde{\mathbf{D}}^{-1} \quad (56)$$

where

$$\begin{aligned} \phi(\tilde{\mathbf{D}}_1, \dots, \tilde{\mathbf{D}}_L, \beta) &= \frac{\alpha_L}{2} \log \frac{|\Sigma_{\tilde{\mathbf{X}}_0}|}{|\tilde{\mathbf{D}}|} + \sum_{i=1}^L \frac{\alpha_i}{2} \log \frac{|\Sigma_{\tilde{\mathbf{N}}_i}|}{|\tilde{\mathbf{D}}_i|} + \sum_{i=2}^L \frac{\alpha_{i-1} - \alpha_i}{2} \\ &\quad \cdot \log^+ \frac{1}{|\tilde{\mathbf{D}}| |\Sigma_{\tilde{\mathbf{X}}_0}^{-1} + \sum_{j=i}^L \Sigma_{\tilde{\mathbf{N}}_j}^{-1} - \sum_{j=i}^L \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \tilde{\mathbf{D}}_j \Sigma_{\tilde{\mathbf{N}}_j}^{-1}|}. \end{aligned}$$

Define $\Gamma_i = (\tilde{\mathbf{D}}_i)_i$, $i = 1, \dots, L$. Note that (55) implies

$$\mathbf{0} \leq \Gamma_i \leq \Sigma_{\mathbf{Z}_i}, \quad i = 1, \dots, L. \quad (57)$$

In view of the definition of $\Sigma_{\tilde{\mathbf{N}}_i}$ and the fact that $\Sigma_{\tilde{\mathbf{N}}_i}^{-1} \tilde{\mathbf{D}}_i \Sigma_{\tilde{\mathbf{N}}_i}^{-1} \leq \Sigma_{\tilde{\mathbf{N}}_i}^{-1}$, it is easy to show that $\Sigma_{\tilde{\mathbf{N}}_i}^{-1} \tilde{\mathbf{D}}_i \Sigma_{\tilde{\mathbf{N}}_i}^{-1}$ converges uniformly (over $\tilde{\mathbf{D}}_i$) to \mathbf{A}_i as $\beta \rightarrow \infty$, where \mathbf{A}_i is a block diagonal matrix with $(\mathbf{A}_i)_i = \Sigma_{\mathbf{Z}_i}^{-1} \Gamma_i \Sigma_{\mathbf{Z}_i}^{-1}$ and $(\mathbf{A}_i)_j = \mathbf{0}$ for $j \neq i$, $i = 1, \dots, L$. As a consequence, $\sum_{i=1}^L \Sigma_{\tilde{\mathbf{N}}_i}^{-1} \tilde{\mathbf{D}}_i \Sigma_{\tilde{\mathbf{N}}_i}^{-1}$ converges uniformly (over $\tilde{\mathbf{D}}_1, \dots, \tilde{\mathbf{D}}_L$) to $(\Sigma_{\mathbf{X}}^{-1} + \Sigma_{\mathbf{Z}}^{-1}) \text{diag}(\Gamma_1, \dots, \Gamma_L) (\Sigma_{\mathbf{X}}^{-1} + \Sigma_{\mathbf{Z}}^{-1})$ as $\beta \rightarrow \infty$. It can also be verified that

$$\begin{aligned} \Sigma_{\tilde{\mathbf{X}}_0}^{-1} + \sum_{i=1}^L \Sigma_{\tilde{\mathbf{N}}_i}^{-1} - \tilde{\mathbf{D}}^{-1} &= \Sigma_{\mathbf{X}}^{-1}(\Sigma_{\mathbf{X}} + \Sigma_{\mathbf{Z}})\Sigma_{\mathbf{X}}^{-1} + (\Sigma_{\mathbf{X}}^{-1} + \Sigma_{\mathbf{Z}}^{-1}) + \frac{L-1}{\beta} \mathbf{I}_m \\ &\quad - \Sigma_{\mathbf{X}}^{-1}(\Sigma_{\mathbf{X}} + \Sigma_{\mathbf{Z}})(\mathbf{D} + \Sigma_{\mathbf{Z}})^{-1}(\Sigma_{\mathbf{X}} + \Sigma_{\mathbf{Z}})\Sigma_{\mathbf{X}}^{-1} \end{aligned}$$

$$\begin{aligned} &= \Sigma_{\mathbf{X}}^{-1}(\Sigma_{\mathbf{X}} + \Sigma_{\mathbf{Z}})(\Sigma_{\mathbf{X}} + \Sigma_{\mathbf{Z}})^{-1}(\Sigma_{\mathbf{X}} + \Sigma_{\mathbf{Z}})\Sigma_{\mathbf{X}}^{-1} \\ &\quad + (\Sigma_{\mathbf{X}}^{-1} + \Sigma_{\mathbf{Z}}^{-1}) + \frac{L-1}{\beta} \mathbf{I}_m \\ &\quad - \Sigma_{\mathbf{X}}^{-1}(\Sigma_{\mathbf{X}} + \Sigma_{\mathbf{Z}})(\mathbf{D} + \Sigma_{\mathbf{Z}})^{-1}(\Sigma_{\mathbf{X}} + \Sigma_{\mathbf{Z}})\Sigma_{\mathbf{X}}^{-1} \\ &= (\Sigma_{\mathbf{X}}^{-1} + \Sigma_{\mathbf{Z}}^{-1})\Sigma_{\mathbf{Z}}(\Sigma_{\mathbf{X}} + \Sigma_{\mathbf{Z}})^{-1}\Sigma_{\mathbf{Z}}(\Sigma_{\mathbf{X}}^{-1} + \Sigma_{\mathbf{Z}}^{-1}) \\ &\quad + (\Sigma_{\mathbf{X}}^{-1} + \Sigma_{\mathbf{Z}}^{-1}) + \frac{L-1}{\beta} \mathbf{I}_m \\ &\quad - (\Sigma_{\mathbf{X}}^{-1} + \Sigma_{\mathbf{Z}}^{-1})\Sigma_{\mathbf{Z}}(\mathbf{D} + \Sigma_{\mathbf{Z}})^{-1}\Sigma_{\mathbf{Z}}(\Sigma_{\mathbf{X}}^{-1} + \Sigma_{\mathbf{Z}}^{-1}) \\ &= (\Sigma_{\mathbf{X}}^{-1} + \Sigma_{\mathbf{Z}}^{-1})(\Sigma_{\mathbf{Z}}(\Sigma_{\mathbf{X}} + \Sigma_{\mathbf{Z}})^{-1}\Sigma_{\mathbf{Z}} + (\Sigma_{\mathbf{X}}^{-1} + \Sigma_{\mathbf{Z}}^{-1})^{-1} \\ &\quad - \Sigma_{\mathbf{Z}}(\mathbf{D} + \Sigma_{\mathbf{Z}})^{-1}\Sigma_{\mathbf{Z}})(\Sigma_{\mathbf{X}}^{-1} + \Sigma_{\mathbf{Z}}^{-1}) + \frac{L-1}{\beta} \mathbf{I}_m \\ &= (\Sigma_{\mathbf{X}}^{-1} + \Sigma_{\mathbf{Z}}^{-1})(\Sigma_{\mathbf{Z}} - \Sigma_{\mathbf{Z}}(\mathbf{D} + \Sigma_{\mathbf{Z}})^{-1}\Sigma_{\mathbf{Z}})(\Sigma_{\mathbf{X}}^{-1} + \Sigma_{\mathbf{Z}}^{-1}) \\ &\quad + \frac{L-1}{\beta} \mathbf{I}_m \\ &= (\Sigma_{\mathbf{X}}^{-1} + \Sigma_{\mathbf{Z}}^{-1})(\mathbf{D}^{-1} + \Sigma_{\mathbf{Z}}^{-1})^{-1}(\Sigma_{\mathbf{X}}^{-1} + \Sigma_{\mathbf{Z}}^{-1}) + \frac{L-1}{\beta} \mathbf{I}_m. \end{aligned}$$

Therefore, (56) implies that

$$\text{diag}(\Gamma_1, \dots, \Gamma_L) \leq (\mathbf{D}^{-1} + \Sigma_{\mathbf{Z}}^{-1})^{-1} + \mathbf{A}(\beta), \quad (58)$$

where $\mathbf{A}(\beta)$ is a positive definite matrix that does not depend on $(\tilde{\mathbf{D}}_1, \dots, \tilde{\mathbf{D}}_L)$, and $\mathbf{A}(\beta)$ tends to $\mathbf{0}$ as $\beta \rightarrow \infty$. Similarly, it can be shown that

$$\begin{aligned} \Sigma_{\tilde{\mathbf{X}}_0}^{-1} + \sum_{j=i}^L \Sigma_{\tilde{\mathbf{N}}_j}^{-1} - \sum_{j=i}^L \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \tilde{\mathbf{D}}_j \Sigma_{\tilde{\mathbf{N}}_j}^{-1} \\ \leq (\Sigma_{\mathbf{X}}^{-1} + \Sigma_{\mathbf{Z}}^{-1})(\Sigma_{\mathbf{Z}} - \tilde{\Gamma}_i + \mathbf{B}_i(\beta))(\Sigma_{\mathbf{X}}^{-1} + \Sigma_{\mathbf{Z}}^{-1}), \end{aligned}$$

where $\mathbf{B}_i(\beta)$ is a positive definite matrix that does not depend on $(\tilde{\mathbf{D}}_1, \dots, \tilde{\mathbf{D}}_L)$, and $\mathbf{B}_i(\beta)$ tends to $\mathbf{0}$ as $\beta \rightarrow \infty$, $i = 2, \dots, L$. Moreover, in view of the fact that $|\Sigma_{\tilde{\mathbf{N}}_i}| = (\prod_{j \neq i} \beta^{m_j}) |\Sigma_{\mathbf{Z}_i}|$ and $|\tilde{\mathbf{D}}_i| \leq (\prod_{j \neq i} \beta^{m_j}) |\Gamma_i|$, we have

$$\frac{\alpha_i}{2} \log \frac{|\Sigma_{\tilde{\mathbf{N}}_i}|}{|\tilde{\mathbf{D}}_i|} \geq \frac{\alpha_i}{2} \log \frac{|\Sigma_{\mathbf{Z}_i}|}{|\Gamma_i|}, \quad i = 1, \dots, L.$$

Now one can readily verify that

$$\phi(\tilde{\mathbf{D}}_1, \dots, \tilde{\mathbf{D}}_L, \beta) \geq \varphi(\Gamma_1, \dots, \Gamma_L, \beta), \quad (59)$$

where

$$\begin{aligned} \varphi(\Gamma_1, \dots, \Gamma_L, \beta) &= \frac{\alpha_L}{2} \log \frac{|\Sigma_{\mathbf{X}} + \Sigma_{\mathbf{Z}}|}{|\mathbf{D} + \Sigma_{\mathbf{Z}}|} + \sum_{i=1}^L \frac{\alpha_i}{2} \log \frac{|\Sigma_{\mathbf{Z}_i}|}{|\Gamma_i|} \\ &\quad + \sum_{i=2}^L \frac{\alpha_{i-1} - \alpha_i}{2} \log^+ \frac{|\Sigma_{\mathbf{Z}}(\Sigma_{\mathbf{Z}} - \tilde{\Gamma}_i + \mathbf{B}_i(\beta))^{-1}\Sigma_{\mathbf{Z}}|}{|\mathbf{D} + \Sigma_{\mathbf{Z}}|}. \end{aligned}$$

Combining (57), (58), and (59) gives

$$\begin{aligned} \sum_{i=1}^L \alpha_i R_i &\geq \min_{\Gamma_1, \dots, \Gamma_L} \varphi(\Gamma_1, \dots, \Gamma_L, \beta) \\ \text{subject to } \mathbf{0} &\leq \Gamma_i \leq \Sigma_{\mathbf{Z}_i}, \quad i = 1, \dots, L, \\ \text{diag}(\Gamma_1, \dots, \Gamma_L) &\leq (\mathbf{D}^{-1} + \Sigma_{\mathbf{Z}}^{-1})^{-1} + \mathbf{A}(\beta). \end{aligned}$$

Note that this lower bound is valid for all $\beta > 0$, and, as $\beta \rightarrow \infty$, it converges to

$$\begin{aligned} \min_{\Gamma_1, \dots, \Gamma_L} \frac{\alpha_L}{2} \log \frac{|\Sigma_{\mathbf{X}} + \Sigma_{\mathbf{Z}}|}{|\mathbf{D} + \Sigma_{\mathbf{Z}}|} + \sum_{i=1}^L \frac{\alpha_i}{2} \log \frac{|\Sigma_{\mathbf{Z}_i}|}{|\Gamma_i|} \\ + \sum_{i=2}^L \frac{\alpha_{i-1} - \alpha_i}{2} \log^+ \frac{|\Sigma_{\mathbf{Z}}(\Sigma_{\mathbf{Z}} - \bar{\Gamma}_i)^{-1} \Sigma_{\mathbf{Z}}|}{|\mathbf{D} + \Sigma_{\mathbf{Z}}|} \\ \text{subject to } \mathbf{0} \leq \Gamma_i \leq \Sigma_{\mathbf{Z}_i}, \quad i = 1, \dots, L, \\ \text{diag}(\Gamma_1, \dots, \Gamma_L) \leq (\mathbf{D}^{-1} + \Sigma_{\mathbf{Z}}^{-1})^{-1}. \end{aligned}$$

The proof is complete in view of the fact that $\mathbf{0} \leq \mathbf{D} = \Sigma_{\mathbf{X}^n|W_1, \dots, W_L} \leq \Sigma_{\mathbf{X}}$ and $(\mathbf{D})_i = \Sigma_{\mathbf{X}_i^n|W_1, \dots, W_L} \leq \mathbf{D}_i$, $i = 1, \dots, L$.

APPENDIX C

PROOF OF THEOREM 7

The following lemma [10, Lemma 10] is needed for the proof of Theorem 7.

Lemma 8: Let \mathbf{A} be a positive definite matrix. For any diagonal matrix \mathbf{S} , if $\mathbf{A} + \mathbf{S}$ is positive definite and the diagonal entries of $(\mathbf{A} + \mathbf{S})^{-1} - \mathbf{A}^{-1}$ are equal to zero, then $\mathbf{S} = \mathbf{0}$.

If there exist non-identical $\mathbf{D}, \tilde{\mathbf{D}} \in \mathcal{D}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{X}})$, then one can find a non-zero block diagonal matrix $\Delta = \text{diag}(\Delta_1, \dots, \Delta_L)$ such that $\mathbf{D} = (\tilde{\mathbf{D}}^{-1} + \Delta)^{-1}$. Let the eigenvalue decomposition of Δ_i be given by $\Delta_i = \mathbf{U}_i \Lambda_i \mathbf{U}_i^T$, where \mathbf{U}_i is a unitary matrix and Λ_i is a diagonal matrix, $i = 1, \dots, L$. Note that $\Delta = \mathbf{U} \Lambda \mathbf{U}^T$, where $\mathbf{U} = \text{diag}(\mathbf{U}_1, \dots, \mathbf{U}_L)$ and $\Lambda = \text{diag}(\Lambda_1, \dots, \Lambda_L)$. Let $\mathbf{B} = \mathbf{U}^T \mathbf{D} \mathbf{U}$ and $\tilde{\mathbf{B}} = \mathbf{U}^T \tilde{\mathbf{D}} \mathbf{U}$. Since $(\mathbf{D})_i = (\tilde{\mathbf{D}})_i$, $i = 1, \dots, L$, it follows that $((\tilde{\mathbf{B}}^{-1} + \Lambda)^{-1})_i = (\tilde{\mathbf{B}})_i$, $i = 1, \dots, L$, and consequently the diagonal entries of $(\tilde{\mathbf{B}}^{-1} + \Lambda)^{-1} - \tilde{\mathbf{B}}$ are equal to zero. This leads to a contradiction with Lemma 8 since $\tilde{\mathbf{B}} > \mathbf{0}$ and $\Lambda \neq \mathbf{0}$.

APPENDIX D

PROOF OF THEOREM 8

Let $\Omega \subseteq \mathbb{R}^n$ be an open, bounded, convex set that is symmetric with respect to its center x_0 (i.e., $x_0 + y \in \Omega$ if and only if $x_0 - y \in \Omega$) and let $\text{cl}(\Omega)$ denote the closure of Ω . Then Borsuk's theorem [28, Theorem 2.1] (see also [29]) can be stated as follows.

Theorem 13: Assume that $f : \text{cl}(\Omega) \rightarrow \mathbb{R}^n$ is continuous and that for all $x_0 + y \in \partial\Omega$, the topological boundary of Ω , we have $f(x_0 + y) \neq \lambda f(x_0 - y)$ for all $\lambda > 0$. Then f has a zero in $\text{cl}(\Omega)$.

We first consider the case $\mathbf{D}_i < \Sigma_{\mathbf{X}_i|\{\mathbf{X}_j\}_{j \neq i}}$, $i = 1, \dots, L$. Let $\tilde{\mathbf{X}}_i = \Sigma_i^{-\frac{1}{2}} \mathbf{X}_i$ and $\tilde{\mathbf{D}}_i = \Sigma_i^{-\frac{1}{2}} \mathbf{D}_i \Sigma_i^{-\frac{1}{2}}$, where $\Sigma_i = \Sigma_{\mathbf{X}_i|\{\mathbf{X}_j\}_{j \neq i}}$, $i = 1, \dots, L$. Define $\tilde{\mathbf{X}} = (\tilde{\mathbf{X}}_1^T, \dots, \tilde{\mathbf{X}}_L^T)^T$. It is clear that $\mathcal{D}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{X}})$ is non-empty if and only if $\mathcal{D}(\tilde{\mathbf{D}}_1, \dots, \tilde{\mathbf{D}}_L, \Sigma_{\tilde{\mathbf{X}}})$ is non-empty. Note that $\tilde{\mathbf{D}}_i < \Sigma_{\tilde{\mathbf{X}}_i|\{\tilde{\mathbf{X}}_j\}_{j \neq i}} = \mathbf{I}_{m_i}$, $i = 1, \dots, L$.

Let \mathcal{C}_i denote the set of all symmetric $m_i \times m_i$ matrices \mathbf{C}_i such that $\beta \mathbf{I}_{m_i} \pm \mathbf{C}_i > \mathbf{0}$, $i = 1, \dots, L$, where β is a positive number. Let $\Omega_i = \beta \mathbf{I}_i + \mathcal{C}_i$, $i = 1, \dots, L$, and $\Omega = \Omega_1 \times \dots \times \Omega_L$. Now we proceed to show that $\mathcal{D}(\tilde{\mathbf{D}}_1, \dots, \tilde{\mathbf{D}}_L, \Sigma_{\tilde{\mathbf{X}}})$

is non-empty. To this end, it suffices to prove that, when β is sufficiently large, the function

$$f(\Lambda) \triangleq (\mathbf{K}_1 - \tilde{\mathbf{D}}_1, \dots, \mathbf{K}_L - \tilde{\mathbf{D}}_L)$$

has a zero in $\text{cl}(\Omega)$, where $\mathbf{K}_i = ((\Sigma_{\tilde{\mathbf{X}}}^{-1} + \Lambda)^{-1})_i$, $i = 1, \dots, L$.

Note that for any $\Lambda = \text{diag}(\Lambda_1, \dots, \Lambda_L) \in \text{cl}(\Omega)$, we can write $\Lambda_i = \beta \mathbf{I}_{m_i} + \mathbf{C}_i$, where \mathbf{C}_i is an $m_i \times m_i$ symmetric matrix satisfying $\beta \mathbf{I}_{m_i} \pm \mathbf{C}_i \geq \mathbf{0}$, $i = 1, \dots, L$. Suppose the eigenvalue decomposition of \mathbf{C}_i is given by $\mathbf{C}_i = \mathbf{U}_i \Delta_i \mathbf{U}_i^T$, where \mathbf{U}_i is a unitary matrix and $\Delta_i = \text{diag}(\delta_{i,1}, \dots, \delta_{i,m_i})$, $i = 1, \dots, L$. Without loss of generality, we assume $|\Lambda_1| = |\beta \mathbf{I}_{m_1} + \mathbf{C}_1| = 0$ for $\Lambda \in \partial\Omega$, and further assume that $\delta_{1,1} = -\beta$. Let $\tilde{\Lambda} = \text{diag}(\tilde{\Lambda}_1, \dots, \tilde{\Lambda}_L)$, where $\tilde{\Lambda}_i = \beta \mathbf{I}_{m_i} - \mathbf{C}_i$, $i = 1, \dots, L$. Moreover, let $\tilde{\mathbf{K}}_i = ((\Sigma_{\tilde{\mathbf{X}}}^{-1} + \tilde{\Lambda})^{-1})_i$, $i = 1, \dots, L$. We shall show that when β is sufficiently large, there is no positive λ such that $f(\Lambda) = \lambda f(\tilde{\Lambda})$. Note that

$$\begin{aligned} \mathbf{K}_1 &\geq (\mathbf{I}_{m_1} + \Lambda_1)^{-1}, \\ \tilde{\mathbf{K}}_1 &\leq (\Sigma_{\tilde{\mathbf{X}}_1}^{-1} + \tilde{\Lambda}_1)^{-1}. \end{aligned}$$

Let $\mathbf{E}_1 = \beta \mathbf{I}_{m_1} + \Delta_1$ and $\tilde{\mathbf{E}}_1 = \beta \mathbf{I}_{m_1} - \Delta_1$. We have

$$\begin{aligned} (\mathbf{I}_{m_1} + \Lambda_1)^{-1} &= \mathbf{U}_1 (\mathbf{I}_{m_1} + \mathbf{E}_1)^{-1} \mathbf{U}_1^T, \\ (\Sigma_{\tilde{\mathbf{X}}_1}^{-1} + \tilde{\Lambda}_1)^{-1} &= \mathbf{U}_1 (\tilde{\Sigma}_1^{-1} + \tilde{\mathbf{E}}_1)^{-1} \mathbf{U}_1^T, \end{aligned}$$

where $\tilde{\Sigma}_1 = \mathbf{U}_1^T \Sigma_{\tilde{\mathbf{X}}_1} \mathbf{U}_1$. If $\mathbf{K}_1 - \tilde{\mathbf{D}}_1 = \lambda(\tilde{\mathbf{K}}_1 - \tilde{\mathbf{D}}_1)$ for some positive λ , then we must have

$$\begin{aligned} \mathbf{U}_1 (\mathbf{I}_{m_1} + \mathbf{E}_1)^{-1} \mathbf{U}_1^T - \tilde{\mathbf{D}}_1 \\ \leq \lambda (\mathbf{U}_1 (\tilde{\Sigma}_1^{-1} + \tilde{\mathbf{E}}_1)^{-1} \mathbf{U}_1^T - \tilde{\mathbf{D}}_1), \end{aligned}$$

which further implies

$$\begin{aligned} (\mathbf{I}_{m_1} + \mathbf{E}_1)^{-1} - \mathbf{U}_1^T \tilde{\mathbf{D}}_1 \mathbf{U}_1 \\ \leq \lambda ((\tilde{\Sigma}_1^{-1} + \tilde{\mathbf{E}}_1)^{-1} - \mathbf{U}_1^T \tilde{\mathbf{D}}_1 \mathbf{U}_1). \end{aligned} \quad (60)$$

Note that the first diagonal entry of $\mathbf{U}_1^T \tilde{\mathbf{D}}_1 \mathbf{U}_1$ is in the interval $[\tilde{d}_{1,\min}, \tilde{d}_{1,\max}]$, where $\tilde{d}_{1,\min}$ ($\tilde{d}_{1,\max}$) is the smallest (largest) eigenvalue of $\tilde{\mathbf{D}}_1$. Since $\mathbf{0} < \tilde{\mathbf{D}}_1 < \Sigma_{\tilde{\mathbf{X}}_1|\{\tilde{\mathbf{X}}_j\}_{j \neq 1}} = \mathbf{I}_{m_1}$, it follows that $0 < \tilde{d}_{1,\min} \leq \tilde{d}_{1,\max} < 1$. Moreover, in view of the fact that $\delta_{1,1} = -\beta$, it can be readily shown that the first diagonal entry of $(\mathbf{I}_{m_1} + \mathbf{E}_1)^{-1}$ equals one while the first diagonal entry of $(\tilde{\Sigma}_1^{-1} + \tilde{\mathbf{E}}_1)^{-1}$ is less than $\frac{1}{2\beta}$ (which tends to zero as $\beta \rightarrow \infty$). Consequently, the first diagonal entry of the left-hand side of (60) is positive while the first diagonal entry of the right-hand side of (60) is negative when β is sufficiently large, which leads to a contradiction. This completes the proof for the case $\mathbf{D}_i < \Sigma_{\mathbf{X}_i|\{\mathbf{X}_j\}_{j \neq i}}$, $i = 1, \dots, L$.

Now consider the general case $\mathbf{D}_i \leq \Sigma_{\mathbf{X}_i|\{\mathbf{X}_j\}_{j \neq i}}$, $i = 1, \dots, L$. Define $\mathbf{D}_i(k) = \mathbf{D}_i - \frac{1}{k} \mathbf{I}_{m_i}$, $i = 1, \dots, L$. We assume $k \geq k^*$, where k^* is chosen such that $\mathbf{D}_i(k^*) > \mathbf{0}$, $i = 1, \dots, L$. Since $\mathbf{D}_i(k) < \Sigma_{\mathbf{X}_i|\{\mathbf{X}_j\}_{j \neq i}}$, $i = 1, \dots, L$, it follows that there exists a positive semidefinite block diagonal matrix $\Lambda(k) = \text{diag}(\Lambda_1(k), \dots, \Lambda_L(k))$ such that $((\Sigma_{\tilde{\mathbf{X}}}^{-1} + \Lambda(k))^{-1})_i = \mathbf{D}_i(k)$, $i = 1, \dots, L$. Note that $\mathbf{0} \leq \Lambda_i(k) \leq \mathbf{D}_i^{-1}(k^*) - \Sigma_{\tilde{\mathbf{X}}_i}^{-1}$, $i = 1, \dots, L$. Since $\{\Lambda(k)\}$ is defined over a compact set, it must contain a converging subsequence. Denote the limit of this converging subsequence by Λ^* . It can be

readily shown that Λ^* is a positive semidefinite block diagonal matrix satisfying $((\Sigma_{\mathbf{X}}^{-1} + \Lambda^*)^{-1})_i = \mathbf{D}_i$, $i = 1, \dots, L$. Therefore, $\mathcal{D}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{X}})$ is non-empty.

APPENDIX E

PROOF OF LEMMA 1

Consider the following convex semidefinite programming problem (which is essentially a relaxed version of the minimization problem associated with $\underline{R}(\mathbf{D}_1, \dots, \mathbf{D}_L, \alpha)$)

$$\begin{aligned} (\tilde{\mathbf{P}}) \quad & \min_{\mathbf{D}, \Gamma_1, \dots, \Gamma_L} \frac{\alpha_L}{2} \log \frac{|\Sigma_{\mathbf{X}} + \Sigma_{\mathbf{Z}}|}{|\mathbf{D} + \Sigma_{\mathbf{Z}}|} + \sum_{i=1}^L \frac{\alpha_i}{2} \log \frac{|\Sigma_{\mathbf{Z}_i}|}{|\Gamma_i|} \\ & + \sum_{i=2}^L \frac{\alpha_{i-1} - \alpha_i}{2} \log \frac{|\Sigma_{\mathbf{Z}}(\Sigma_{\mathbf{Z}} - \bar{\Gamma}_i)^{-1} \Sigma_{\mathbf{Z}}|}{|\mathbf{D} + \Sigma_{\mathbf{Z}}|} \end{aligned}$$

subject to $\mathbf{0} \preceq \mathbf{D}$,

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

$$\mathbf{0} \preceq \Gamma_i, \quad i = 1, \dots, L,$$

$$\text{diag}(\Gamma_1, \dots, \Gamma_L) \preceq (\mathbf{D}^{-1} + \Sigma_{\mathbf{Z}}^{-1})^{-1},$$

where \mathbf{E}_i is a block diagonal matrix with $(\mathbf{E}_i)_i = \mathbf{I}_{m_i}$ and $(\mathbf{E}_i)_j = \mathbf{0}$ for $j \neq i$, $i = 1, \dots, L$. Its Lagrangian is given by

$$\begin{aligned} L(\mathbf{D}, \Gamma_1, \dots, \Gamma_L) &= \frac{\alpha_L}{2} \log \frac{|\Sigma_{\mathbf{X}} + \Sigma_{\mathbf{Z}}|}{|\mathbf{D} + \Sigma_{\mathbf{Z}}|} + \sum_{i=1}^L \frac{\alpha_i}{2} \log \frac{|\Sigma_{\mathbf{Z}_i}|}{|\Gamma_i|} \\ &+ \sum_{i=2}^L \frac{\alpha_{i-1} - \alpha_i}{2} \log \frac{|\Sigma_{\mathbf{Z}}(\Sigma_{\mathbf{Z}} - \bar{\Gamma}_i)^{-1} \Sigma_{\mathbf{Z}}|}{|\mathbf{D} + \Sigma_{\mathbf{Z}}|} \\ &+ \sum_{i=1}^L \text{tr}(\Pi_i (\mathbf{E}_i \mathbf{D} \mathbf{E}_i - \mathbf{F}_i \mathbf{D}_i \mathbf{F}_i^T)) \\ &+ \text{tr} \left(\Lambda \left(\text{diag}(\Gamma_1, \dots, \Gamma_L) - (\mathbf{D}^{-1} + \Sigma_{\mathbf{Z}}^{-1})^{-1} \right) \right), \end{aligned}$$

where $\Pi_1, \dots, \Pi_L, \Lambda$ are positive semidefinite matrices. Note that $(\hat{\mathbf{D}}, \hat{\Gamma}_1, \dots, \hat{\Gamma}_L)$ is an optimal solution to $(\tilde{\mathbf{P}})$ if it satisfies the following KKT conditions

$$\nabla_{\mathbf{D}} L(\mathbf{D}, \Gamma_1, \dots, \Gamma_L)|_{\mathbf{D}=\hat{\mathbf{D}}} = \mathbf{0}, \quad (61)$$

$$\nabla_{\Gamma_i} L(\mathbf{D}, \Gamma_1, \dots, \Gamma_L)|_{\Gamma_i=\hat{\Gamma}_i} = \mathbf{0}, \quad i = 1, \dots, L, \quad (62)$$

$$\mathbf{0} \preceq \hat{\mathbf{D}}, \quad (63)$$

$$\mathbf{E}_i \hat{\mathbf{D}} \mathbf{E}_i \preceq \mathbf{F}_i \mathbf{D}_i \mathbf{F}_i^T, \quad i = 1, \dots, L, \quad (64)$$

$$\mathbf{0} \preceq \hat{\Gamma}_i, \quad i = 1, \dots, L, \quad (65)$$

$$\Pi_i (\mathbf{E}_i \hat{\mathbf{D}} \mathbf{E}_i - \mathbf{F}_i \mathbf{D}_i \mathbf{F}_i^T) = \mathbf{0}, \quad i = 1, \dots, L, \quad (66)$$

$$\Lambda \left(\text{diag}(\hat{\Gamma}_1, \dots, \hat{\Gamma}_L) - (\hat{\mathbf{D}}^{-1} + \Sigma_{\mathbf{Z}}^{-1})^{-1} \right) = \mathbf{0}. \quad (67)$$

It can be verified that

$$\begin{aligned} & \nabla_{\mathbf{D}} \text{tr} \left(\Lambda (\mathbf{D}^{-1} + \Sigma_{\mathbf{Z}}^{-1})^{-1} \right) \\ &= \mathbf{D}^{-1} (\mathbf{D}^{-1} + \Sigma_{\mathbf{Z}}^{-1})^{-1} \Lambda (\mathbf{D}^{-1} + \Sigma_{\mathbf{Z}}^{-1})^{-1} \mathbf{D}^{-1}. \end{aligned}$$

Furthermore, in view of the fact that

$$\bar{\Gamma}_j = \sum_{k=1}^{j-1} \mathbf{F}_k \Sigma_{\mathbf{Z}_k} \mathbf{F}_k^T + \sum_{k=j}^L \mathbf{F}_k \Gamma_k \mathbf{F}_k^T, \quad j = 2, \dots, L,$$

we have

$$\begin{aligned} \nabla_{\Gamma_i} \left(\sum_{j=2}^L (\alpha_{j-1} - \alpha_j) \log |\Sigma_{\mathbf{Z}} - \bar{\Gamma}_j| \right) &= \Theta_i, \\ & i = 1, \dots, L, \end{aligned}$$

where $\Theta_1 = \mathbf{0}$ and

$$\begin{aligned} \Theta_i &= - \sum_{j=2}^i (\alpha_{j-1} - \alpha_j) \mathbf{F}_i^T (\Sigma_{\mathbf{Z}} - \bar{\Gamma}_j)^{-1} \mathbf{F}_i, \\ & i = 2, \dots, L. \end{aligned}$$

Therefore, (61) and (62) can be rewritten as

$$\begin{aligned} & \hat{\mathbf{D}} \left(\sum_{i=1}^L \mathbf{E}_i \Pi_i \mathbf{E}_i - \alpha_1 (\hat{\mathbf{D}} + \Sigma_{\mathbf{Z}})^{-1} \right) \hat{\mathbf{D}} \\ &= (\hat{\mathbf{D}}^{-1} + \Sigma_{\mathbf{Z}}^{-1})^{-1} \Lambda (\hat{\mathbf{D}}^{-1} + \Sigma_{\mathbf{Z}}^{-1})^{-1}, \quad (68) \end{aligned}$$

$$(\Lambda)_i = \alpha_i \hat{\Gamma}_i^{-1} + \Theta_i, \quad i = 1, \dots, L. \quad (69)$$

Now pick $\hat{\mathbf{D}}$ and $\Sigma_{\mathbf{Z}}$ with $\hat{\mathbf{D}} \in \mathcal{D}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{X}})$ and $\Sigma_{\mathbf{Z}} \in \mathcal{Z}(\Sigma_{\mathbf{X}})$. Since $\hat{\mathbf{D}}^{-1} + \Sigma_{\mathbf{Z}}^{-1}$ is a block diagonal matrix, we can find $\hat{\Gamma}_1, \dots, \hat{\Gamma}_L$ such that

$$\text{diag}(\hat{\Gamma}_1, \dots, \hat{\Gamma}_L) = (\hat{\mathbf{D}}^{-1} + \Sigma_{\mathbf{Z}}^{-1})^{-1}. \quad (70)$$

Note that our choice of $(\hat{\mathbf{D}}, \hat{\Gamma}_1, \dots, \hat{\Gamma}_L)$ satisfies (63)-(67). Furthermore, (68) and (69) are satisfied if and only if there exists a block diagonal matrix Π such that

$$\begin{aligned} \Pi &\succeq \alpha_1 (\hat{\mathbf{D}} + \Sigma_{\mathbf{Z}})^{-1}, \\ (\hat{\mathbf{D}}(\Pi - \alpha_1 (\hat{\mathbf{D}} + \Sigma_{\mathbf{Z}})^{-1})\hat{\mathbf{D}})_i &= \alpha_i \hat{\Gamma}_i + \hat{\Gamma}_i \Theta_i \hat{\Gamma}_i, \\ & i = 1, \dots, L. \end{aligned}$$

It is clear that if our choice of $(\hat{\mathbf{D}}, \hat{\Gamma}_1, \dots, \hat{\Gamma}_L)$ is an optimal solution to $(\tilde{\mathbf{P}})$, then it is also an optimal solution to the minimization problem associated with $\underline{R}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{Z}}, \alpha)$. Therefore, one can readily complete the proof by noticing that (70) implies $\underline{R}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{Z}}, \alpha) = \bar{R}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{Z}}, \alpha)$.

APPENDIX F

PROOF OF THEOREM 9

In view of Theorems 7 and 8, it can be readily shown that when $\|\mathbf{D}_1\|, \dots, \|\mathbf{D}_L\|$ are sufficiently small, there exists a unique element $\mathbf{D} \in \mathcal{D}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{X}})$ such that $\mathbf{D} \prec \Sigma_{\mathbf{X}}$. Let \mathbf{G} be a positive definite block diagonal matrix defined as $\mathbf{G} = (\mathbf{D}^{-1} - \Sigma_{\mathbf{X}}^{-1})^{-1}$. It can be verified that

$$\mathbf{D} = \mathbf{G}^{\frac{1}{2}} (\mathbf{I}_m - \mathbf{A}) \mathbf{G}^{\frac{1}{2}},$$

where

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{1,1} & \cdots & \mathbf{A}_{1,L} \\ \vdots & \ddots & \vdots \\ \mathbf{A}_{L,1} & \cdots & \mathbf{A}_{L,L} \end{pmatrix} = \mathbf{G}^{\frac{1}{2}} (\mathbf{G} + \Sigma_{\mathbf{X}})^{-1} \mathbf{G}^{\frac{1}{2}}$$

with $\mathbf{A}_{i,i} = (\mathbf{A})_i$, $i = 1, \dots, L$. It is clear that \mathbf{D} , \mathbf{G} , and \mathbf{A} all tend to $\mathbf{0}$ as $\|\mathbf{D}_i\| \rightarrow \mathbf{0}$, $i = 1, \dots, L$.

Now choose an arbitrary $\Sigma_{\mathbf{Z}}$ from $\mathcal{Z}(\Sigma_{\mathbf{X}})$ and let $\Gamma_i = ((\mathbf{D}^{-1} + \Sigma_{\mathbf{Z}}^{-1})^{-1})_i$, $i = 1, \dots, L$. Note that

$$\begin{aligned} (\mathbf{D} \Pi \mathbf{D})_i &= \Xi_i \\ \Leftrightarrow ((\mathbf{I}_m - \mathbf{A}) \mathbf{H} (\mathbf{I}_m - \mathbf{A}))_i &= (\mathbf{G})_i^{-\frac{1}{2}} \Xi_i (\mathbf{G})_i^{-\frac{1}{2}}, \\ & i = 1, \dots, L, \quad (71) \end{aligned}$$

where \mathbf{H} is a block diagonal matrix defined as $\mathbf{H} = \mathbf{G}^{\frac{1}{2}} \Pi \mathbf{G}^{\frac{1}{2}}$. Let $\mathbf{B} = (\mathbf{B}_{i,j})$ with $\mathbf{B}_{i,i} = (\mathbf{I}_m - \mathbf{A})_i$, $i = 1, \dots, L$, and $\mathbf{B}_{i,j} = -\mathbf{A}_{i,j}$ for $i \neq j$. We can express (71) equivalently as

$$(\mathbf{B} * \mathbf{B}) \begin{pmatrix} \text{vec}((\mathbf{H})_1) \\ \vdots \\ \text{vec}((\mathbf{H})_L) \end{pmatrix} = \begin{pmatrix} \text{vec}((\mathbf{G})_1^{-\frac{1}{2}} \Xi_1 (\mathbf{G})_1^{-\frac{1}{2}}) \\ \vdots \\ \text{vec}((\mathbf{G})_L^{-\frac{1}{2}} \Xi_L (\mathbf{G})_L^{-\frac{1}{2}}) \end{pmatrix}.$$

Clearly, $\mathbf{B} * \mathbf{B} \rightarrow \mathbf{I}_q$ as $\|\mathbf{D}_i\| \rightarrow \mathbf{0}$, $i = 1, \dots, L$, where $q = \sum_{i=1}^L m_i^2$. Note that

$$\begin{aligned} (\mathbf{G})_1^{-\frac{1}{2}} \Xi_1 (\mathbf{G})_1^{-\frac{1}{2}} &= \alpha_1 (\mathbf{G})_1^{-\frac{1}{2}} \mathbf{D}_1 (\mathbf{G})_1^{-\frac{1}{2}} \\ &= \alpha_1 (\mathbf{G}^{-\frac{1}{2}} \mathbf{D} \mathbf{G}^{-\frac{1}{2}})_1 \\ &= \alpha_1 (\mathbf{I}_m - \mathbf{A})_1 \end{aligned}$$

and

$$\begin{aligned} &(\mathbf{G})_i^{-\frac{1}{2}} \Xi_i (\mathbf{G})_i^{-\frac{1}{2}} \\ &= \alpha_1 (\mathbf{G})_i^{-\frac{1}{2}} \mathbf{D}_i (\mathbf{G})_i^{-\frac{1}{2}} - (\alpha_1 - \alpha_i) (\mathbf{G})_i^{-\frac{1}{2}} \Gamma_i (\mathbf{G})_i^{-\frac{1}{2}} \\ &\quad - \sum_{j=2}^i (\alpha_{j-1} - \alpha_i) (\mathbf{G})_i^{-\frac{1}{2}} \Gamma_i \mathbf{F}_i^T (\Sigma_{\mathbf{Z}} - \bar{\Gamma}_j)^{-1} \mathbf{F}_i \Gamma_i (\mathbf{G})_i^{-\frac{1}{2}} \\ &= \alpha_1 (\mathbf{G}^{-\frac{1}{2}} \mathbf{D} \mathbf{G}^{-\frac{1}{2}})_i - (\alpha_1 - \alpha_i) \mathbf{Q}_i \\ &\quad - \sum_{j=2}^i (\alpha_{j-1} - \alpha_i) \mathbf{Q}_i (\mathbf{G})_i^{\frac{1}{2}} \mathbf{F}_i^T (\Sigma_{\mathbf{Z}} - \bar{\Gamma}_j)^{-1} \mathbf{F}_i (\mathbf{G})_i^{-\frac{1}{2}} \mathbf{Q}_i \\ &= \alpha_1 (\mathbf{I}_m - \mathbf{A})_i - (\alpha_1 - \alpha_i) \mathbf{Q}_i \\ &\quad - \sum_{j=2}^i (\alpha_{j-1} - \alpha_i) \mathbf{Q}_i (\mathbf{G})_i^{\frac{1}{2}} \mathbf{F}_i^T (\Sigma_{\mathbf{Z}} - \bar{\Gamma}_j)^{-1} \mathbf{F}_i (\mathbf{G})_i^{\frac{1}{2}} \mathbf{Q}_i, \end{aligned}$$

$i = 2, \dots, L$,

where

$$\begin{aligned} \mathbf{Q}_i &= (\mathbf{G})_i^{-\frac{1}{2}} \Gamma_i (\mathbf{G})_i^{-\frac{1}{2}} \\ &= (\mathbf{G}^{-\frac{1}{2}} (\mathbf{D}^{-1} + \Sigma_{\mathbf{Z}}^{-1})^{-1} \mathbf{G}^{-\frac{1}{2}})_i \\ &= ((\mathbf{I}_m - \mathbf{A})^{-1} + \mathbf{G}^{\frac{1}{2}} \Sigma_{\mathbf{Z}}^{-1} \mathbf{G}^{\frac{1}{2}})^{-1}_i, \quad i = 2, \dots, L. \end{aligned}$$

Since $\mathbf{0} \leq \bar{\Gamma}_i \leq (\Sigma_{\mathbf{X}}^{-1} + \Sigma_{\mathbf{Z}}^{-1})^{-1} \prec \Sigma_{\mathbf{Z}}$, it follows that

$$\Sigma_{\mathbf{Z}}^{-1} \leq (\Sigma_{\mathbf{Z}} - \bar{\Gamma}_i)^{-1} \leq (\Sigma_{\mathbf{Z}} - (\Sigma_{\mathbf{X}}^{-1} + \Sigma_{\mathbf{Z}}^{-1})^{-1})^{-1},$$

$i = 2, \dots, L$.

Now one can readily show that

$$(\mathbf{G})_i^{-\frac{1}{2}} \Xi_i (\mathbf{G})_i^{-\frac{1}{2}} \rightarrow \alpha_i \mathbf{I}_{m_i}, \quad i = 1, \dots, L,$$

and consequently

$$\mathbf{H} \rightarrow \text{diag}(\alpha_1 \mathbf{I}_{m_1}, \dots, \alpha_L \mathbf{I}_{m_L})$$

as $\|\mathbf{D}_i\| \rightarrow \mathbf{0}$, $i = 1, \dots, L$. Let $\mathbf{S} = \mathbf{G}^{\frac{1}{2}} \Sigma_{\mathbf{Z}}^{-1} \mathbf{G}^{\frac{1}{2}}$. We have

$$\begin{aligned} &\mathbf{G}^{\frac{1}{2}} (\Pi - \alpha_1 (\mathbf{D} + \Sigma_{\mathbf{Z}})^{-1}) \mathbf{G}^{\frac{1}{2}} \\ &= \mathbf{H} - \alpha_1 (\mathbf{I}_m - \mathbf{A} + \mathbf{S}^{-1})^{-1} \\ &= \mathbf{H} - \alpha_1 (\mathbf{S} - \mathbf{S}((\mathbf{I}_m - \mathbf{A})^{-1} + \mathbf{S})^{-1} \mathbf{S}) \\ &\rightarrow \text{diag}(\alpha_1 \mathbf{I}_{m_1}, \dots, \alpha_L \mathbf{I}_{m_L}) \end{aligned}$$

as $\|\mathbf{D}_i\| \rightarrow \mathbf{0}$, $i = 1, \dots, L$. Therefore, $\Pi - \alpha_1 (\mathbf{D} + \Sigma_{\mathbf{Z}})^{-1}$ must be positive definite when $\|\mathbf{D}_1\|, \dots, \|\mathbf{D}_L\|$ are sufficiently small, which completes the proof.

APPENDIX G

PROOF OF THEOREM 10

Throughout the proof we assume that $\Sigma_{\mathbf{X}_1}, \dots, \Sigma_{\mathbf{X}_L}$ are fixed. In view of Theorems 7 and 8, it can be readily shown that when $\|\Sigma_{\mathbf{X}} - \text{diag}(\Sigma_{\mathbf{X}_1}, \dots, \Sigma_{\mathbf{X}_L})\|$ is sufficiently small, there exists a unique element $\mathbf{D} \in \mathcal{D}(\mathbf{D}_1, \dots, \mathbf{D}_L, \Sigma_{\mathbf{X}})$. We choose β such that $\beta > 0$ and $(\alpha_1 - \alpha_i)\beta \leq \alpha_1$, $i = 1, \dots, L$. Let

$$\Sigma_{\mathbf{Z}} = (\beta \text{diag}(\Sigma_{\mathbf{X}_1}^{-1}, \dots, \Sigma_{\mathbf{X}_L}^{-1}) - \Sigma_{\mathbf{X}}^{-1})^{-1}.$$

Note that

$$\Sigma_{\mathbf{Z}} \rightarrow \frac{1}{\beta - 1} \text{diag}(\Sigma_{\mathbf{X}_1}, \dots, \Sigma_{\mathbf{X}_L})$$

as $\|\Sigma_{\mathbf{X}} - \text{diag}(\Sigma_{\mathbf{X}_1}, \dots, \Sigma_{\mathbf{X}_L})\| \rightarrow 0$, which, together with the fact that $\Sigma_{\mathbf{X}}^{-1} + \Sigma_{\mathbf{Z}}^{-1} = \beta \text{diag}(\Sigma_{\mathbf{X}_1}, \dots, \Sigma_{\mathbf{X}_L})$, implies $\Sigma_{\mathbf{Z}} \in \mathcal{Z}(\Sigma_{\mathbf{X}})$ when $\|\Sigma_{\mathbf{X}} - \text{diag}(\Sigma_{\mathbf{X}_1}, \dots, \Sigma_{\mathbf{X}_L})\|$ is sufficiently small. Let $\Gamma_i = ((\mathbf{D}^{-1} + \Sigma_{\mathbf{Z}}^{-1})^{-1})_i$, $i = 1, \dots, L$. It is easy to see that

$$\mathbf{D} \rightarrow \text{diag}(\mathbf{D}_1, \dots, \mathbf{D}_L),$$

$$\Gamma_i \rightarrow (\mathbf{D}_i^{-1} + (\beta - 1) \Sigma_{\mathbf{X}_i}^{-1})^{-1}, \quad i = 1, \dots, L,$$

$$\Theta_i \rightarrow$$

$$-(\alpha_1 - \alpha_i) \left(\frac{1}{\beta - 1} \Sigma_{\mathbf{X}_i} - (\mathbf{D}_i^{-1} + (\beta - 1) \Sigma_{\mathbf{X}_i}^{-1})^{-1} \right)^{-1},$$

$i = 2, \dots, L$,

as $\|\Sigma_{\mathbf{X}} - \text{diag}(\Sigma_{\mathbf{X}_1}, \dots, \Sigma_{\mathbf{X}_L})\| \rightarrow 0$. Therefore, we have

$$\begin{aligned} \Xi_i &\rightarrow \alpha_1 \mathbf{D}_i - (\alpha_1 - \alpha_i) (\mathbf{D}_i^{-1} + (\beta - 1) \Sigma_{\mathbf{X}_i}^{-1})^{-1} \\ &\quad - (\alpha_1 - \alpha_i) (\mathbf{D}_i^{-1} + (\beta - 1) \Sigma_{\mathbf{X}_i}^{-1})^{-1} \\ &\quad \cdot \left(\frac{1}{\beta - 1} \Sigma_{\mathbf{X}_i} - (\mathbf{D}_i^{-1} + (\beta - 1) \Sigma_{\mathbf{X}_i}^{-1})^{-1} \right)^{-1} \\ &\quad \cdot (\mathbf{D}_i^{-1} + (\beta - 1) \Sigma_{\mathbf{X}_i}^{-1})^{-1} \\ &= \alpha_1 \mathbf{D}_i - (\alpha_1 - \alpha_i) (\mathbf{D}_i^{-1} + (\beta - 1) \Sigma_{\mathbf{X}_i}^{-1})^{-1} \\ &\quad - (\alpha_1 - \alpha_i) \left(\frac{1}{\beta - 1} (\mathbf{D}_i^{-1} + (\beta - 1) \Sigma_{\mathbf{X}_i}^{-1}) \right. \\ &\quad \cdot \Sigma_{\mathbf{X}_i} (\mathbf{D}_i^{-1} + (\beta - 1) \Sigma_{\mathbf{X}_i}^{-1}) \\ &\quad \left. - (\mathbf{D}_i^{-1} + (\beta - 1) \Sigma_{\mathbf{X}_i}^{-1}) \right)^{-1} \\ &= \alpha_1 \mathbf{D}_i - (\alpha_1 - \alpha_i) (\mathbf{D}_i^{-1} + (\beta - 1) \Sigma_{\mathbf{X}_i}^{-1})^{-1} \\ &\quad - (\alpha_1 - \alpha_i) \left(\frac{1}{\beta - 1} \mathbf{D}_i^{-1} \Sigma_{\mathbf{X}_i} \mathbf{D}_i^{-1} + \mathbf{D}_i^{-1} \right)^{-1} \\ &= \alpha_i \mathbf{D}_i, \quad i = 2, \dots, L, \end{aligned}$$

as $\|\Sigma_{\mathbf{X}} - \text{diag}(\Sigma_{\mathbf{X}_1}, \dots, \Sigma_{\mathbf{X}_L})\| \rightarrow 0$. Also note that $\Xi_1 = \alpha_1 \mathbf{D}_1$. Now one can readily show that

$$\Pi \rightarrow \text{diag}(\alpha_1 \mathbf{D}_1^{-1}, \dots, \alpha_L \mathbf{D}_L^{-1})$$

and consequently

$$\begin{aligned} & \Pi - \alpha_1 (\mathbf{D} + \Sigma_{\mathbf{Z}})^{-1} \\ & \rightarrow \text{diag}(\alpha_1 \mathbf{D}_1 - \alpha_1 (\beta - 1) ((\beta - 1) \mathbf{D}_1 + \Sigma_{\mathbf{X}_1})^{-1}, \\ & \quad \dots, \alpha_L \mathbf{D}_L - \alpha_1 (\beta - 1) ((\beta - 1) \mathbf{D}_L + \Sigma_{\mathbf{X}_L})^{-1}) \end{aligned}$$

as $\|\Sigma_{\mathbf{X}} - \text{diag}(\Sigma_{\mathbf{X}_1}, \dots, \Sigma_{\mathbf{X}_L})\| \rightarrow 0$. Since

$$\begin{aligned} & \alpha_i \mathbf{D}_i^{-1} - \alpha_1 (\beta - 1) ((\beta - 1) \mathbf{D}_i + \Sigma_{\mathbf{X}_i})^{-1} \\ & > \alpha_i \mathbf{D}_i^{-1} - \alpha_1 (\beta - 1) (\beta \mathbf{D}_i)^{-1} \\ & = \frac{\alpha_1 - (\alpha_1 - \alpha_i) \beta}{\beta} \mathbf{D}_i^{-1} \\ & \geq \mathbf{0}, \quad i = 1, \dots, L, \end{aligned}$$

it follows that $\Pi - \alpha_1 (\mathbf{D} + \Sigma_{\mathbf{Z}})^{-1}$ is positive definite when $\|\Sigma_{\mathbf{X}} - \text{diag}(\Sigma_{\mathbf{X}_1}, \dots, \Sigma_{\mathbf{X}_L})\|$ is sufficiently small, which completes the proof.

APPENDIX H PROOF OF (30)

Without loss of generality, we assume $\tilde{\mathbf{X}}'_L = \mathbf{X}'_L + \mathbf{M}$, where \mathbf{M} is a Gaussian random vector, independent of $(\mathbf{X}_0, \mathbf{X}_L, \mathbf{X}'_L, U_L, V)$, with covariance matrix $\Sigma_{\mathbf{M}} = \hat{\mathbf{D}}_{\mathbf{X}_L|U_L} - \hat{\mathbf{D}}_L$. Note that

$$\begin{aligned} \mathbf{X}_{L,\lambda} &= \mathbb{E}[\mathbf{X}_{L,\lambda} | \mathbf{X}_L, \tilde{\mathbf{X}}_{L,\lambda}] + \Delta \\ &= \sqrt{\lambda} \mathbf{X}_L + \sqrt{1 - \lambda} \mathbb{E}[\mathbf{X}'_L | \tilde{\mathbf{X}}'_L] + \Delta \\ &= \sqrt{\lambda} \mathbf{X}_L + \sqrt{1 - \lambda} \hat{\mathbf{D}}_L \hat{\mathbf{D}}_{\mathbf{X}'_L|U_L}^{-1} \tilde{\mathbf{X}}'_L + \Delta \\ &= \sqrt{\lambda} \hat{\mathbf{D}}_L \Sigma_{\mathbf{N}_L}^{-1} \Sigma_{\tilde{\mathbf{N}}_L} \Sigma_{\mathbf{N}_L}^{-1} \mathbf{X}_L \\ & \quad + (\mathbf{I}_m - \hat{\mathbf{D}}_L \Sigma_{\mathbf{N}_L}^{-1} \Sigma_{\tilde{\mathbf{N}}_L} \Sigma_{\mathbf{N}_L}^{-1}) \tilde{\mathbf{X}}_{L,\lambda} + \Delta, \end{aligned}$$

where Δ is a Gaussian random vector, independent of $(\mathbf{X}_0, \mathbf{X}_L, \tilde{\mathbf{X}}_{L,\lambda}, U_L, V)$, with covariance matrix $\Sigma_{\Delta} = (1 - \lambda) \hat{\mathbf{D}}_L \Sigma_{\mathbf{N}_L}^{-1} \Sigma_{\tilde{\mathbf{N}}_L} \Sigma_{\mathbf{N}_L}^{-1} \hat{\mathbf{D}}_L$. Define $\Delta' = \frac{1}{\sqrt{\lambda}} \Sigma_{\mathbf{N}_L} \Sigma_{\tilde{\mathbf{N}}_L}^{-1} \Sigma_{\mathbf{N}_L} \hat{\mathbf{D}}_L^{-1} \Delta$. It is clear that $\Sigma_{\Delta'} = \frac{1 - \lambda}{\lambda} \Sigma_{\mathbf{N}_L} \Sigma_{\tilde{\mathbf{N}}_L}^{-1} \Sigma_{\mathbf{N}_L}$. We have

$$\begin{aligned} \Sigma_{\mathbf{X}_L | \mathbf{X}_{L,\lambda}, U_L, V} &= \Sigma_{\mathbf{X}_L | \mathbf{X}_{L,\lambda}, \tilde{\mathbf{X}}_{L,\lambda}, U_L, V} \\ &= \Sigma_{\mathbf{X}_L | \mathbf{X}_L + \Delta', \tilde{\mathbf{X}}_{L,\lambda}, U_L, V}. \end{aligned}$$

Now consider the linear MMSE estimator of \mathbf{X}_L from $(\mathbf{X}_L + \Delta', \mathbb{E}[\mathbf{X}_L | \tilde{\mathbf{X}}_{L,\lambda}, U_L, V])$, which is a function of $(\mathbf{X}_L + \Delta', \tilde{\mathbf{X}}_{L,\lambda}, U_L, V)$. It can be verified that the resulting error covariance matrix is $(\tilde{\mathbf{D}}_{L,\lambda}^{-1} + \Sigma_{\Delta'}^{-1})^{-1}$. As a consequence,

$$\Sigma_{\mathbf{X}_L | \mathbf{X}_{L,\lambda}, U_L, V} \preceq (\tilde{\mathbf{D}}_{L,\lambda}^{-1} + \Sigma_{\Delta'}^{-1})^{-1}. \quad (72)$$

To bound $\mathbf{D}_{L,\lambda}$ in terms of $\Sigma_{\mathbf{X}_L | \mathbf{X}_{L,\lambda}, U_L, V}$, we consider the linear MMSE estimator of \mathbf{X}_L from $(\mathbf{X}_0, \mathbb{E}[\mathbf{X}_L | \mathbf{X}_{L,\lambda}, U_L, V])$. It can be shown that the error covariance matrix incurred by this estimator is

$$\left(\Sigma_{\mathbf{X}_L | \mathbf{X}_{L,\lambda}, U_L, V}^{-1} + \Sigma_{\mathbf{N}_L}^{-1} \Sigma_{\tilde{\mathbf{N}}_L} \Sigma_{\mathbf{N}_L}^{-1} \right)^{-1}.$$

Therefore, we have

$$\mathbf{D}_{L,\lambda} \preceq \left(\Sigma_{\mathbf{X}_L | \mathbf{X}_{L,\lambda}, U_L, V}^{-1} + \Sigma_{\mathbf{N}_L}^{-1} \Sigma_{\tilde{\mathbf{N}}_L} \Sigma_{\mathbf{N}_L}^{-1} \right)^{-1}. \quad (73)$$

Combining (72) and (73) gives

$$\begin{aligned} \mathbf{D}_{L,\lambda} &\preceq \left(\tilde{\mathbf{D}}_{L,\lambda}^{-1} + \Sigma_{\Delta'}^{-1} + \Sigma_{\mathbf{N}_L}^{-1} \Sigma_{\tilde{\mathbf{N}}_L} \Sigma_{\mathbf{N}_L}^{-1} \right)^{-1} \\ &= \left(\tilde{\mathbf{D}}_{L,\lambda}^{-1} + \frac{1}{1 - \lambda} \Sigma_{\mathbf{N}_L}^{-1} \Sigma_{\tilde{\mathbf{N}}_L} \Sigma_{\mathbf{N}_L}^{-1} \right)^{-1}, \end{aligned}$$

which completes the proof of (30).

APPENDIX I

PROOF OF (35)

We shall show that $\hat{\mathbf{D}}_{L,\lambda} \triangleq (1 - \lambda) \hat{\mathbf{D}}_L$ is an optimal solution to (\mathbf{P}_2) , from which (35) follows immediately. Since (\mathbf{P}_2) is a convex semidefinite programming problem, it suffices to verify the KKT conditions. Clearly, $\hat{\mathbf{D}}_{L,\lambda}$ satisfies the constraints in (34); therefore, one just need to show that there exists a positive semidefinite matrix Λ such that

$$\begin{aligned} & \bar{\mathbf{B}} - \mu_L \hat{\mathbf{D}}_L^{-1} + \mu_{L-1} \hat{\mathbf{D}}_{L,\lambda}^{-1} \left(\hat{\mathbf{D}}_{L,\lambda}^{-1} - \frac{1}{1 - \lambda} \Xi^{-1} \right)^{-1} \hat{\mathbf{D}}_{\mathbf{X}'_L|U_L}^{-1} \\ & \quad \cdot \left(\hat{\mathbf{D}}_{L,\lambda}^{-1} - \frac{1}{1 - \lambda} \Xi^{-1} \right)^{-1} \hat{\mathbf{D}}_{L,\lambda}^{-1} + \Lambda = \mathbf{0}, \quad (74) \\ & \Lambda \left(\hat{\mathbf{D}}_{L,\lambda} - \left(\mathbf{D}_L^{-1} + \frac{\lambda}{1 - \lambda} \hat{\mathbf{D}}_L^{-1} \right)^{-1} \right) = \mathbf{0}, \end{aligned}$$

where

$$\begin{aligned} \bar{\mathbf{B}} &= \sum_{i=2}^{L-1} \mu_{i-1} \Sigma_{\mathbf{N}_L}^{-1} \hat{\mathbf{D}}_{\mathbf{X}_0|U_i} \Sigma_{\mathbf{N}_L}^{-1}, \\ \Xi &= \Sigma_{\mathbf{N}_L} \Sigma_{\tilde{\mathbf{N}}_L}^{-1} \Sigma_{\mathbf{N}_L}. \end{aligned}$$

According to (74),

$$\Lambda = -\bar{\mathbf{B}} + \mu_L \hat{\mathbf{D}}_L^{-1} - \mu_{L-1} \hat{\mathbf{D}}_L^{-1} \hat{\mathbf{D}}_{\mathbf{X}'_L|U_L} \hat{\mathbf{D}}_L^{-1},$$

which, together with (20), implies that $\Lambda = \mathbf{O}_L \geq \mathbf{0}$. As a consequence, it suffices to show

$$\mathbf{O}_L \left(\hat{\mathbf{D}}_{L,\lambda} - \left(\mathbf{D}_L^{-1} + \frac{\lambda}{1 - \lambda} \hat{\mathbf{D}}_L^{-1} \right)^{-1} \right) = \mathbf{0}. \quad (75)$$

Note that

$$\begin{aligned} & \hat{\mathbf{D}}_{L,\lambda} - \left(\mathbf{D}_L^{-1} + \frac{\lambda}{1 - \lambda} \hat{\mathbf{D}}_L^{-1} \right)^{-1} \\ &= (1 - \lambda) \left(\hat{\mathbf{D}}_L - ((1 - \lambda) \mathbf{D}_L^{-1} + \lambda \hat{\mathbf{D}}_L^{-1})^{-1} \right) \\ &= (1 - \lambda) \left(\hat{\mathbf{D}}_L - \mathbf{D}_L ((1 - \lambda) \hat{\mathbf{D}}_L + \lambda \mathbf{D}_L)^{-1} \hat{\mathbf{D}}_L \right) \\ &= (1 - \lambda) \left((1 - \lambda) \hat{\mathbf{D}}_L + \lambda \mathbf{D}_L - \mathbf{D}_L \right) \left((1 - \lambda) \hat{\mathbf{D}}_L \right. \\ & \quad \left. + \lambda \mathbf{D}_L \right)^{-1} \hat{\mathbf{D}}_L \\ &= (1 - \lambda)^2 \left(\hat{\mathbf{D}}_L - \mathbf{D}_L \right) \left((1 - \lambda) \hat{\mathbf{D}}_L + \lambda \mathbf{D}_L \right)^{-1} \hat{\mathbf{D}}_L. \quad (76) \end{aligned}$$

Combining (21) and (76) proves (75).

REFERENCES

- [1] T. Berger, Z. Zhang, and H. Viswanathan, "The CEO problem [multiterminal source coding]," *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] J. Chen, X. Zhang, T. Berger, and S. B. Wicker, "An upper bound on the sum-rate distortion function and its corresponding rate allocation schemes for the CEO problem," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 977–987, Aug. 2004.
- [5] V. Prabhakaran, D. Tse, and K. Ramchandran, "Rate region of the quadratic Gaussian CEO problem," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Chicago, IL, USA, Jun./Jul. 2004, p. 117.
- [6] 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.

- [7] S. Tavildar, P. Viswanath, and A. B. Wagner, "The Gaussian many-helper-one distributed source coding problem," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 564–581, Jan. 2010.
- [8] H. Weingarten, Y. Steinberg, and S. 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.
- [9] S. Tavildar and P. Viswanath, "On the sum-rate of the vector Gaussian CEO problem," in *Proc. 39th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Oct./Nov. 2005, pp. 3–7.
- [10] 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. 8, pp. 3946–3960, Aug. 2010.
- [11] J. Wang and J. Chen, "Vector Gaussian two-terminal source coding," *IEEE Trans. Inf. Theory*, vol. 59, no. 6, pp. 3693–3708, Jun. 2013.
- [12] 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.
- [13] H. Weingarten, T. Liu, S. Shamai, Y. Steinberg, and P. Viswanath, "The capacity region of the degraded multiple-input multiple-output compound broadcast channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5011–5023, Nov. 2009.
- [14] R. Liu, T. Liu, H. V. Poor, and S. Shamai, "A vector generalization of Costa's entropy-power inequality with applications," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1865–1879, Apr. 2010.
- [15] E. Ekrem and S. Ulukus, "An outer bound for the vector Gaussian CEO problem," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Jul. 2012, pp. 576–580.
- [16] 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.
- [17] Y. Oohama, "Gaussian multiterminal source coding," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1912–1923, Nov. 1997.
- [18] 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.
- [19] Y. Oohama, *Distributed Source Coding of Correlated Gaussian Sources* [Online]. Available: <http://arxiv.org/abs/1007.4418>
- [20] A. B. Wagner and V. Anantharam, "An improved outer bound for multiterminal source coding," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 1919–1937, May 2008.
- [21] T. Berger, "Multiterminal source coding," in *The Information Theory Approach to Communications* (CISM International Centre for Mechanical Sciences), vol. 229, G. Longo, Ed. New York, NY, USA: Springer-Verlag, 1978, pp. 171–231.
- [22] S.-Y. Tung, "Multiterminal source coding," Ph.D. dissertation, School Electr. Eng., Cornell Univ., Ithaca, NY, USA, 1978.
- [23] S. Liu, "Matrix results on the Khatri–Rao and Tracy–Singh products," *Linear Algebra Appl.*, vol. 289, nos. 1–3, pp. 267–277, Mar. 1999.
- [24] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1261–1282, Apr. 2005.
- [25] E. Ekrem and S. Ulukus, "The secrecy capacity region of the Gaussian MIMO multi-receiver wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2083–2114, Apr. 2011.
- [26] A. Dembo, T. M. Cover, and J. A. Thomas, "Information theoretic inequalities," *IEEE Trans. Inf. Theory*, vol. 37, no. 6, pp. 1501–1518, Nov. 1991.
- [27] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA, USA: Athena Scientific, 1999.
- [28] A. Frommer and B. Lang, "Existence tests for solutions of nonlinear equations using Borsuk's theorem," *SIAM J. Numer. Anal.*, vol. 43, no. 3, pp. 1348–1361, 2005.
- [29] K. Deimling, *Nonlinear Functional Analysis*. Berlin, Germany: Springer-Verlag, 1985.
- [30] Md. S. Rahman and A. B. Wagner, "Rate region of the vector Gaussian one-helper source-coding problem," in *Proc. 49th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Sep. 2011, pp. 1710–1717.

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 September 2005 to July 2006, and a Postdoctoral Fellow at the IBM Thomas J. Watson Research Center, Yorktown Heights, NY, from July 2006 to August 2007. Since September 2007 he has been with the Department of Electrical and Computer Engineering at McMaster University, Hamilton, ON, Canada, where he is currently an Associate Professor. 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. He is currently serving as an Associate Editor for Shannon Theory for the IEEE TRANSACTIONS ON INFORMATION THEORY.