An Outer Bound for the Vector Gaussian CEO Problem

Ersen Ekrem and Sennur Ulukus, *Member, IEEE*

Abstract—We study the vector Gaussian CEO problem, where there are arbitrary number of agents, each having a noisy observation of a vector Gaussian source. The goal of the agents is to describe the source to a central unit, which wants to reconstruct the source within a given distortion. The rate-distortion region of the vector Gaussian CEO problem is unknown in general. Here, we provide an outer bound for the rate-distortion region of the vector Gaussian CEO problem. We obtain our outer bound by evaluating an outer bound for the multiterminal source coding problem by means of a technique relying on the de Bruijn identity and properties of the Fisher information. Next, we investigate the tightness of our outer bound. Although our outer bound is tight for certain cases, we show that our outer bound does not provide the exact rate-distortion region in general. To this end, we provide an example and show that the rate-distortion region is strictly contained in our outer bound for this example.

Index Terms—CEO problem, Gaussian multi-terminal source coding, entropy power inequality, Fisher information.

I. INTRODUCTION

W E STUDY the vector Gaussian CEO problem, where there is a vector Gaussian source which is observed through some noisy channels by an arbitrary number of agents. The agents process their observations independently and communicate them to a central unit (the so-called CEO unit) through orthogonal and rate-limited links (see Figure 1). The goal of the agents is to describe their observations to the central unit in a way that the central unit can reconstruct the source within a given distortion. The fundamental trade-off between the rate spent by the agents to describe the source and the distortion attained by the central unit is characterized by the rate-distortion region, which is unknown in general.

The CEO problem is introduced in [1], where the authors consider a discrete memoryless setting where the source and the observations of the agents all come from some discrete alphabet. In the setting of [1], the central unit is interested in estimating the source with the minimum expected error frequency which corresponds to the Hamming distance between
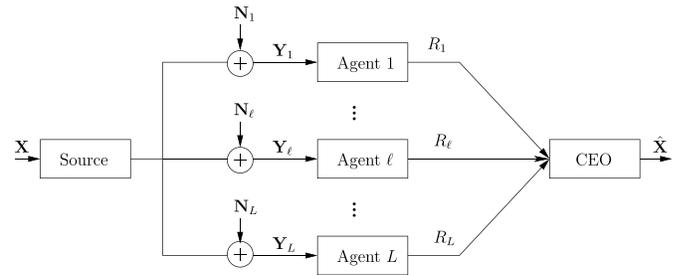
Fig. 1. The vector Gaussian CEO problem.

the source sequence and the central unit's estimation of the source sequence. In [1], the authors consider the decay rate of the error frequency with respect to the rate expenditure of the agents, and obtain the best possible decay rate when the number of agents goes to infinity.

The scalar Gaussian CEO problem is studied in [2], where there is a scalar Gaussian source which is observed through some linear Gaussian channels by the agents. The agents describe their observations to the central unit in a way that the central unit can reconstruct the source within a certain minimum mean square error (MMSE). In [2], the decay rate of the MMSE with respect to the rate expenditure of the agents is considered and shown to be inversely proportional with the rate expenditure of the agents, when the number of agents goes to infinity. The scalar Gaussian CEO problem is further studied in [3] and [4], where instead of the decay rate of the achievable MMSE, the focus was on the entire rate-distortion region. In [3] and [4], the entire rate-distortion region for the scalar Gaussian problem is established. The achievability is shown by using the Berger-Tung inner bound [5], and the converse is established by using the entropy-power inequality. Recently, an alternative proof for the sum-rate of the scalar Gaussian CEO problem is established in [6] without invoking the entropy-power inequality.

As pointed out by several works [7], [8], although entropy-power inequality is a key tool in providing converse proofs for scalar Gaussian problems, it might be restrictive for vector Gaussian problems. For the vector Gaussian CEO problem, this observation is noticed in [9], where the authors provide a lower bound for the sum-rate of the vector Gaussian CEO problem by using the entropy-power inequality. This lower bound is shown to be tight under certain conditions, although it is not tight in general. Recently, [10] provided an outer bound for the rate-distortion region of the vector Gaussian CEO

problem when there are only two agents. They obtain their outer bound by using an extremal inequality, which can be viewed as a generalization of the extremal inequality provided in [11]. The outer bound proposed in [10] holds for two agents, and hence, can be viewed as a special case of our outer bound which holds for an arbitrary number of agents.

In this paper, we consider the vector Gaussian CEO problem for an arbitrary number of agents and provide an outer bound for its rate-distortion region. We first consider the outer bound provided in [12] for the multi-terminal source coding problem, and evaluate it for the vector Gaussian CEO problem at hand. In the evaluation of the outer bound in [12], we use the de Bruijn identity [13], a connection between the differential entropy and the Fisher information, along with the properties of the MMSE and the Fisher information. This evaluation technique which relies on the de Bruijn identity is useful in the sense that it is able to alleviate some shortcomings of the entropy-power inequality in vector Gaussian problems [8], [14].

Next, we investigate the tightness of our outer bound. Despite being tight for certain cases, our outer bound falls short of providing the exact rate-distortion region of the vector Gaussian CEO problem in general. We establish this fact by considering the parallel Gaussian model, for which we obtain the entire rate-distortion region explicitly and show that our outer bound strictly includes this rate-distortion region. In other words, for the parallel Gaussian model, our outer bound is not equal to the rate-distortion region, which shows that our outer bound is not tight in general.

II. PROBLEM STATEMENT AND THE MAIN RESULT

In the CEO problem, there are L sensors, each of which getting a noisy observation of a source. The goal of the sensors is to describe their observations to the CEO unit such that the CEO unit can reconstruct the source within a given distortion. In the vector Gaussian CEO problem, there is an i.i.d. vector Gaussian source $\{\mathbf{X}_i\}_{i=1}^n$ with zero-mean and covariance $\mathbf{K}_X$. Each sensor gets a noisy version of this Gaussian source

$$\mathbf{Y}_{\ell,i} = \mathbf{X}_i + \mathbf{N}_{\ell,i}, \quad \ell = 1, \ldots, L \tag{1}$$

where $\{\mathbf{N}_{\ell,i}\}_{i=1}^n$ is an i.i.d. sequence of Gaussian random vectors with zero-mean and covariance $\mathbf{\Sigma}_\ell$. Moreover, noise among the sensors are independent, i.e., $\mathbf{N}_{1,i}, \ldots, \mathbf{N}_{L,i}$ are independent $\forall i = 1, \ldots, n$. In the vector Gaussian CEO problem, the distortion of the reconstructed vector is measured by its mean square error matrix

$$\frac{1}{n} \sum_{i=1}^n E\left[\left(\mathbf{X}_i - \hat{\mathbf{X}}_i\right)\left(\mathbf{X}_i - \hat{\mathbf{X}}_i\right)^\top\right] \tag{2}$$

where $\hat{\mathbf{X}}^n$ denotes the reconstructed vector.

An $(n, R_1, \ldots, R_L)$ code for the CEO problem consists of an encoding function at each sensor $f_\ell^n : \mathbb{R}^{M \times n} \to \mathcal{B}_\ell^n = \{1, \ldots, 2^{nR_\ell}\}$, i.e., $B_\ell^n = f_\ell^n(\mathbf{Y}_\ell^n)$ where $B_\ell^n \in \mathcal{B}_\ell^n$, $\ell = 1, \ldots, L$, and a decoding function at the CEO unit $g^n : \mathcal{B}_1^n \times \ldots \times \mathcal{B}_L^n \to \mathbb{R}^{M \times n}$, i.e., $\hat{\mathbf{X}}^n = g^n(B_1^n, \ldots, B_L^n)$, where M denotes the size of the vector Gaussian source $\mathbf{X}$.

We note that since the mean square error is minimized by the MMSE estimator, which is the conditional mean, without loss of generality, the decoding function g^n can be chosen as the MMSE estimator. Consequently, we have

$$\hat{\mathbf{X}}_i = E\left[\mathbf{X}_i | B_1^n, \ldots, B_L^n\right] \tag{3}$$

using which in (2), we get

$$\frac{1}{n} \sum_{i=1}^n E\left[\left(\mathbf{X}_i - \hat{\mathbf{X}}_i\right)\left(\mathbf{X}_i - \hat{\mathbf{X}}_i\right)^\top\right] =$$
$$\frac{1}{n} \sum_{i=1}^n \text{mmse}(\mathbf{X}_i | B_1^n, \ldots, B_L^n) \tag{4}$$

In view of (4), a rate tuple $(R_1, \ldots, R_L)$ is said to achieve the distortion $\mathbf{D}$ if there exists an $(n, R_1, \ldots, R_L)$ code such that

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^n \text{mmse}(\mathbf{X}_i | B_1^n, \ldots, B_L^n) \preceq \mathbf{D} \tag{5}$$

where $\mathbf{D}$ is a strictly positive definite matrix. Throughout the paper, we assume that the distortion matrix $\mathbf{D}$ satisfies

$$\left(\mathbf{K}_X^{-1} + \sum_{\ell=1}^L \mathbf{\Sigma}_\ell^{-1}\right)^{-1} \preceq \mathbf{D} \preceq \mathbf{K}_X \tag{6}$$

where the lower bound on the distortion constraint $\mathbf{D}$ corresponds to the MMSE matrix obtained when the CEO unit has direct access to the observations of the agents $\{\mathbf{Y}_\ell\}_{\ell=1}^L$. The derivation of this lower bound is provided in Appendix I, where we also provide insight on the upper bound in (6). In Appendix I, we also show that imposing the lower bound on $\mathbf{D}$ in (6), i.e., imposing $\left(\mathbf{K}_X^{-1} + \sum_{\ell=1}^L \mathbf{\Sigma}_\ell^{-1}\right)^{-1} \preceq \mathbf{D}$, does not incur any loss of generality, while imposing the upper bound on $\mathbf{D}$ in (6), i.e., imposing $\mathbf{D} \preceq \mathbf{K}_X$, might incur some loss of generality.

The rate-distortion region $\mathcal{R}(\mathbf{D})$ of the vector Gaussian CEO problem is defined as the closure of all rate tuples $(R_1, \ldots, R_L)$ that can achieve the distortion $\mathbf{D}$.

The main result of this paper is the following outer bound on the rate-distortion region $\mathcal{R}(\mathbf{D})$ of the vector Gaussian CEO problem stated in the following theorem.

Theorem 1: The rate-distortion region of the Gaussian CEO problem $\mathcal{R}(\mathbf{D})$ is contained in the region $\mathcal{R}^o(\mathbf{D})$ which is given by the union of rate tuples $(R_1, \ldots, R_L)$ satisfying[1]

$$\sum_{\ell \in \mathcal{A}} R_\ell$$
$$\geq \frac{1}{2} \log^+ \frac{\left|\left(\mathbf{K}_X^{-1} + \sum_{\ell \in \mathcal{A}^c} \mathbf{\Sigma}_\ell^{-1} - \sum_{\ell \in \mathcal{A}^c} \mathbf{\Sigma}_\ell^{-1} \mathbf{D}_\ell \mathbf{\Sigma}_\ell^{-1}\right)^{-1}\right|}{|\mathbf{D}|}$$
$$+ \sum_{\ell \in \mathcal{A}} \frac{1}{2} \log \frac{|\mathbf{\Sigma}_\ell|}{|\mathbf{D}_\ell|} \tag{7}$$

---

[1] After we have submitted our work [15], [16], it is brought to our attention that the same result is independently and concurrently obtained in [17].

for all $\mathcal{A} \subseteq \{1,\ldots,L\}$, *where the union is over all positive semi-definite matrices* $\{\mathbf{D}_\ell\}_{\ell=1}^L$ *satisfying the following constraints*

$$\left(\mathbf{K}_X^{-1} + \sum_{\ell=1}^L \boldsymbol{\Sigma}_\ell^{-1} - \sum_{\ell=1}^L \boldsymbol{\Sigma}_\ell^{-1}\mathbf{D}_\ell\boldsymbol{\Sigma}_\ell^{-1}\right)^{-1} \preceq \mathbf{D} \quad (8)$$

$$\mathbf{0} \preceq \mathbf{D}_\ell \preceq \boldsymbol{\Sigma}_\ell, \quad \ell = 1,\ldots,L \quad (9)$$

and $\log^+ x = \max(\log x, 0)$.

We obtain this outer bound by evaluating the outer bound given in [12]. In this evaluation, we use the de Bruijn identity, a differential connection between the Fisher information and the differential entropy, as well as the semi-definite orders between the MMSE and the Fisher information. In this regard, our outer bound is reminiscent of the earlier works on the CEO problem [3], [10], where the authors also evaluate an outer bound similar to the single-letter one in [12], by using either the entropy-power inequality or an extremal inequality. The difference of our work is the way we evaluate the existing single-letter outer bound, where, in this evaluation, we use de Bruijn identity. The proof of Theorem 1 is given in Section V. Next, we provide the following inner bound for the rate-distortion region $\mathcal{R}(\mathbf{D})$.

Theorem 2: An inner bound for the rate-distortion region of the vector Gaussian CEO problem is given by the region $\mathcal{R}^i(\mathbf{D})$ *which is described by the union of rate tuples* $(R_1,\ldots,R_L)$ *satisfying*

$$\sum_{\ell \in \mathcal{A}} R_\ell$$

$$\geq \frac{1}{2}\log \frac{\left|\left(\mathbf{K}_X^{-1} + \sum_{\ell \in \mathcal{A}^c} \boldsymbol{\Sigma}_\ell^{-1} - \sum_{\ell \in \mathcal{A}^c} \boldsymbol{\Sigma}_\ell^{-1}\mathbf{D}_\ell\boldsymbol{\Sigma}_\ell^{-1}\right)^{-1}\right|}{\left|\left(\mathbf{K}_X^{-1} + \sum_{\ell=1}^L \boldsymbol{\Sigma}_\ell^{-1} - \sum_{\ell=1}^L \boldsymbol{\Sigma}_\ell^{-1}\mathbf{D}_\ell\boldsymbol{\Sigma}_\ell^{-1}\right)^{-1}\right|}$$

$$+ \sum_{\ell \in \mathcal{A}} \frac{1}{2}\log\frac{|\boldsymbol{\Sigma}_\ell|}{|\mathbf{D}_\ell|} \quad (10)$$

for all $\mathcal{A} \subseteq \{1,\ldots,L\}$, *where the union is over all positive semi-definite matrices* $\{\mathbf{D}_\ell\}_{\ell=1}^L$ *satisfying*

$$\left(\mathbf{K}_X^{-1} + \sum_{\ell=1}^L \boldsymbol{\Sigma}_\ell^{-1} - \sum_{\ell=1}^L \boldsymbol{\Sigma}_\ell^{-1}\mathbf{D}_\ell\boldsymbol{\Sigma}_\ell^{-1}\right)^{-1} \preceq \mathbf{D} \quad (11)$$

$$\mathbf{0} \preceq \mathbf{D}_\ell \preceq \boldsymbol{\Sigma}_\ell, \quad \ell = 1,\ldots,L \quad (12)$$

We obtain this inner bound by evaluating the Berger-Tung inner bound [5] by jointly Gaussian auxiliary random variables. The proof of Theorem 2 is given in Appendix VII.

III. ALTERNATIVE CHARACTERIZATIONS OF THE BOUNDS

In this section, we provide alternative characterizations for the outer and inner bounds given in Theorem 1 and Theorem 2, respectively. To this end, we note that since the rate-distortion region $\mathcal{R}(\mathbf{D})$ is convex, it can be characterized by the tangent hyperplanes to it, i.e., by solving the following optimization problem

$$\min_{(R_1,\ldots,R_L) \in \mathcal{R}(\mathbf{D})} \sum_{\ell=1}^L \mu_\ell R_\ell \quad (13)$$

for all $\mu_\ell \geq 0, \ell = 1,\ldots,L$. Hence, the outer and inner bounds in Theorem 1 and 2 provide lower and upper bounds for the optimization problem in (13), respectively. Since both the outer and inner bounds are also convex, they can also be described by the tangent hyperplanes to them. In particular, the outer and inner bounds can be described by the following optimization problems

$$\min_{(R_1,\ldots,R_L) \in \mathcal{R}^o(\mathbf{D})} \sum_{\ell=1}^L \mu_\ell R_\ell \quad (14)$$

$$\min_{(R_1,\ldots,R_L) \in \mathcal{R}^i(\mathbf{D})} \sum_{\ell=1}^L \mu_\ell R_\ell \quad (15)$$

respectively, where $\mu_\ell \geq 0$, $\ell = 1,\ldots,L$. We note that the optimization problem in (14) corresponds to the alternative characterization of the outer bound in Theorem 1, and hence, provides a lower bound for the optimization problem in (13) that characterizes the rate-distortion region of the vector Gaussian CEO problem. Similarly, the second optimization problem in (15) corresponds to the alternative characterization of the inner bound in Theorem 2, and hence, provides an upper bound for the optimization problem in (13). Now, by using the contra-polymatroid structure of the achievable rate region and the outer bound [3], we state the explicit form of the optimization problems in (14)-(15) starting with the one for the outer bound.

Theorem 3: Assume $\mu_1 \geq \cdots \geq \mu_L \geq 0$. *The explicit form of the optimization problem in (14) is given by (16) (on the next page) where* $\{\mathbf{D}_\ell\}_{\ell=1}^L$ *are subject to the following constraints*

$$\left(\mathbf{K}_X^{-1} + \sum_{\ell=1}^L \boldsymbol{\Sigma}_\ell^{-1} - \sum_{\ell=1}^L \boldsymbol{\Sigma}_\ell^{-1}\mathbf{D}_\ell\boldsymbol{\Sigma}_\ell^{-1}\right)^{-1} \preceq \mathbf{D} \quad (17)$$

$$\mathbf{0} \preceq \mathbf{D}_\ell \preceq \boldsymbol{\Sigma}_\ell, \quad \ell = 1,\ldots,L \quad (18)$$

Next, we provide the explicit form of the optimization problem in (15) as follows.

Theorem 4: Assume $\mu_1 \geq \ldots \geq \mu_L \geq 0$. *The explicit form of the optimization problem in (14) is given by (19) (on the next page) where* $\{\mathbf{D}_\ell\}_{\ell=1}^L$ *are subject to the following constraints*

$$\left(\mathbf{K}_X^{-1} + \sum_{\ell=1}^L \boldsymbol{\Sigma}_\ell^{-1} - \sum_{\ell=1}^L \boldsymbol{\Sigma}_\ell^{-1}\mathbf{D}_\ell\boldsymbol{\Sigma}_\ell^{-1}\right)^{-1} \preceq \mathbf{D} \quad (20)$$

$$\mathbf{0} \preceq \mathbf{D}_\ell \preceq \boldsymbol{\Sigma}_\ell, \quad \ell = 1,\ldots,L \quad (21)$$

Next, we provide some remarks about the outer bound given in Theorem 3 and the inner bound given in Theorem 4. First, we note that in both cases, the bounds are to be optimized over the positive semi-definite matrices $\{\mathbf{D}_\ell\}_{\ell=1}^L$, and the feasible sets for both cases are identical as seen through (17)-(18) and (20)-(21). On the other hand, rate bounds differ as seen through (16) and (19). Despite this difference, there are cases where the outer and inner bounds match, providing a complete characterization of the rate-distortion region. Here, we note a general *sufficient* condition under which the outer and inner bounds coincide. If the minimum in Theorem 3 is achieved by positive semi-definite matrices $\{\mathbf{D}_\ell^*\}_{\ell=1}^L$ which

$$\min_{(R_1,\ldots,R_L)\in\mathcal{R}(\mathbf{D})} \sum_{\ell=1}^{L} \mu_\ell R_\ell \geq \min_{(R_1,\ldots,R_L)\in\mathcal{R}^o(\mathbf{D})} \sum_{\ell=1}^{L} \mu_\ell R_\ell$$

$$= \min_{\{\mathbf{D}_\ell\}_{\ell=1}^L} \sum_{\ell=1}^{L-1} \frac{\mu_\ell - \mu_{\ell+1}}{2} \log^+ \frac{\left|\left(\mathbf{K}_X^{-1} + \sum_{j=\ell+1}^{L} \boldsymbol{\Sigma}_j^{-1}(\boldsymbol{\Sigma}_j - \mathbf{D}_j)\boldsymbol{\Sigma}_j^{-1}\right)^{-1}\right|}{|\mathbf{D}|} + \sum_{\ell=1}^{L} \frac{\mu_\ell}{2} \log \frac{|\boldsymbol{\Sigma}_\ell|}{|\mathbf{D}_\ell|} + \frac{\mu_L}{2} \log \frac{|\mathbf{K}_X|}{|\mathbf{D}|} \quad (16)$$

$$\min_{(R_1,\ldots,R_L)\in\mathcal{R}(\mathbf{D})} \sum_{\ell=1}^{L} \mu_\ell R_\ell \leq \min_{(R_1,\ldots,R_L)\in\mathcal{R}^i(\mathbf{D})} \sum_{\ell=1}^{L} \mu_\ell R_\ell$$

$$= \min_{\{\mathbf{D}_\ell\}_{\ell=1}^L} \sum_{\ell=1}^{L-1} \frac{\mu_\ell - \mu_{\ell+1}}{2} \log \frac{\left|\left(\mathbf{K}_X^{-1} + \sum_{j=\ell+1}^{L} \boldsymbol{\Sigma}_j^{-1}(\boldsymbol{\Sigma}_j - \mathbf{D}_j)\boldsymbol{\Sigma}_j^{-1}\right)^{-1}\right|}{\left|\left(\mathbf{K}_X^{-1} + \sum_{j=1}^{L} \boldsymbol{\Sigma}_j^{-1}(\boldsymbol{\Sigma}_j - \mathbf{D}_j)\boldsymbol{\Sigma}_j^{-1}\right)^{-1}\right|} + \sum_{\ell=1}^{L} \frac{\mu_\ell}{2} \log \frac{|\boldsymbol{\Sigma}_\ell|}{|\mathbf{D}_\ell|}$$

$$+ \frac{\mu_L}{2} \log \frac{|\mathbf{K}_X|}{\left|\left(\mathbf{K}_X^{-1} + \sum_{j=1}^{L} \boldsymbol{\Sigma}_j^{-1}(\boldsymbol{\Sigma}_j - \mathbf{D}_j)\boldsymbol{\Sigma}_j^{-1}\right)^{-1}\right|} \quad (19)$$

attain the distortion constraint in (17) with equality, then the optimization problems in Theorem 3 and Theorem 4 yield identical results, implying the tightness of the outer bound. One particular example where the outer and inner bounds match is the scalar Gaussian model considered next.

A. Scalar Gaussian Model

Here, we consider the scalar Gaussian model:

$$Y_{\ell,i} = X_i + N_{\ell,i}, \quad \ell = 1,\ldots,L \quad (22)$$

where X_i is an i.i.d. Gaussian source with zero-mean and variance σ_X^2. The noise at the ℓth sensor $N_{\ell,i}$ is also an i.i.d. Gaussian random variable sequence with variance σ_ℓ^2. For the scalar model (scalar Gaussian CEO problem), our outer bound in Theorem 1 reduces to the following form.

Corollary 1: The rate-distortion region of the scalar Gaussian CEO problem $\mathcal{R}(D)$ is contained in the region $\mathcal{R}^o(D)$ which is given by the union of rate tuples $(R_1,\ldots,R_L)$ satisfying

$$\sum_{\ell\in\mathcal{A}} R_\ell \geq \frac{1}{2} \log^+ \frac{1}{D} \left(\frac{1}{\sigma_X^2} + \sum_{\ell\in\mathcal{A}^c} \frac{\sigma_\ell^2 - D_\ell}{\sigma_\ell^4}\right)^{-1}$$
$$+ \sum_{\ell\in\mathcal{A}} \frac{1}{2} \log \frac{\sigma_\ell^2}{D_\ell} \quad (23)$$

for all $\mathcal{A} \subseteq \{1,\ldots,L\}$, where the union is over all $\{D_\ell\}_{\ell=1}^L$ satisfying the following constraints

$$\left(\frac{1}{\sigma_X^2} + \sum_{\ell=1}^{L} \frac{\sigma_\ell^2 - D_\ell}{\sigma_\ell^4}\right)^{-1} \leq D \quad (24)$$

$$0 \leq D_\ell \leq \sigma_\ell^2, \quad \ell = 1,\ldots,L \quad (25)$$

Using Theorem 3, our outer bound for the scalar Gaussian model can be expressed in the following alternative form

$$\min_{(R_1,\ldots,R_L)\in\mathcal{R}(\mathbf{D})} \sum_{\ell=1}^{L} \mu_\ell R_\ell \geq \min_{(R_1,\ldots,R_L)\in\mathcal{R}^o(\mathbf{D})} \sum_{\ell=1}^{L} \mu_\ell R_\ell \quad (26)$$

$$= \min_{\{D_\ell\}_{\ell=1}^L} \sum_{\ell=1}^{L-1} \frac{\mu_\ell - \mu_{\ell+1}}{2} \log^+ \frac{1}{D} \left(\frac{1}{\sigma_X^2} + \sum_{j=\ell+1}^{L} \frac{\sigma_\ell^2 - D_\ell}{\sigma_\ell^4}\right)^{-1}$$
$$+ \sum_{\ell=1}^{L} \frac{\mu_\ell}{2} \log \frac{\sigma_\ell^2}{D_\ell} + \frac{\mu_L}{2} \log \frac{\sigma_X^2}{D} \quad (27)$$

where $\{D_\ell\}_{\ell=1}^L$ are subject to the constraints in (24)-(25), and we assume $\mu_1 \geq \ldots \geq \mu_L \geq 0$. In [3], it is shown that the optimal $\{D_\ell^*\}_{\ell=1}^L$ that minimizes (27) satisfies the constraint in (24) with equality, i.e., for this optimal $\{D_\ell^*\}_{\ell=1}^L$, we have

$$\left(\frac{1}{\sigma_X^2} + \sum_{\ell=1}^{L} \frac{\sigma_\ell^2 - D_\ell^*}{\sigma_\ell^4}\right)^{-1} = D \quad (28)$$

As we pointed out in the previous section, when, for the outer bound, the distortion constraint is satisfied with equality, then the outer bound in Theorem 1 and the inner bound in Theorem 2 match; yielding the rate-distortion region. Hence, in view of (28), we have the entire rate-distortion region for the scalar Gaussian CEO problem.

Theorem 5 ([3], [4]): The rate-distortion region of the scalar Gaussian CEO problem $\mathcal{R}(D)$ is given by the union of rate tuples $(R_1,\ldots,R_L)$ satisfying

$$\sum_{\ell\in\mathcal{A}} R_\ell \geq \frac{1}{2} \log \frac{1}{D} \left(\frac{1}{\sigma_X^2} + \sum_{\ell\in\mathcal{A}^c} \frac{\sigma_\ell^2 - D_\ell}{\sigma_\ell^4}\right)^{-1} + \sum_{\ell\in\mathcal{A}} \frac{1}{2} \log \frac{\sigma_\ell^2}{D_\ell} \quad (29)$$

for all $\mathcal{A} \subseteq \{1, \ldots, L\}$, *where the union is over all* $\{D_\ell\}_{\ell=1}^{L}$ *satisfying the following constraints*

$$\left(\frac{1}{\sigma_X^2} + \sum_{\ell=1}^{L} \frac{\sigma_\ell^2 - D_\ell}{\sigma_\ell^4} \right)^{-1} = D \tag{30}$$

$$0 \leq D_\ell \leq \sigma_\ell^2, \quad \ell = 1, \ldots, L \tag{31}$$

We note that since the distortion constraint in (30) is satisfied with equality, we do not need the positivity operator in (29).

IV. PARALLEL GAUSSIAN MODEL AND A COUNTER-EXAMPLE

In this section, first, we consider the parallel Gaussian model, and obtain its rate-distortion region. Next, we consider a specific parallel Gaussian model and show that our outer bound in Theorem 1 is not tight. In other words, we show that, in general, there are rate tuples $(R_1, \ldots, R_L)$ that lie inside our outer bound and are not contained in the rate-distortion region, i.e., in general, our outer bound strictly contains the rate-distortion region.

In the parallel Gaussian model, the Gaussian source $\mathbf{X}_i$ has a diagonal covariance matrix. In particular, we have $\mathbf{X}_i = [X_{1,i} \ \ldots \ X_{M,i}]$ where $\{X_{m,i}\}_{m=1}^{M}$ are independent Gaussian random variables with zero-mean and variance $\{\sigma_m^2\}_{m=1}^{M}$, respectively. Moreover, the noise at the ℓth sensor $\mathbf{N}_{\ell,i}$ also has a diagonal covariance matrix. In particular, we have $\mathbf{N}_{\ell,i} = [N_{\ell 1,i} \ \ldots \ N_{\ell M,i}]$, where $\{N_{\ell m,i}\}_{m=1}^{M}$ are independent Gaussian random variables with zero-mean with variance $\{\sigma_{\ell m}^2\}_{m=1}^{M}$, respectively. In the parallel Gaussian model, there is a separate-distortion constraint on each component of the source as follows

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \text{mmse}(X_{m,i}|B_1^n, \ldots, B_L^n) \leq D_m, \ m=1, \ldots, M \tag{32}$$

where we have the following constraints on $\{D_m\}_{m=1}^{M}$

$$\left(\frac{1}{\sigma_m^2} + \sum_{\ell=1}^{L} \frac{1}{\sigma_{\ell m}^2} \right)^{-1} \leq D_m \leq \sigma_m^2, \quad m = 1, \ldots, M \tag{33}$$

We note that the constraints on D_m in (33) are the scalar versions of the constraints in (6) that we impose for the vector Gaussian model. For the parallel Gaussian model, we establish the rate-distortion region $\mathcal{R}^p(\{D_m\}_{m=1}^{M})$ as stated in the following theorem.

Theorem 6: The rate-distortion region $\mathcal{R}^p(\{D_m\}_{m=1}^{M})$ *of the parallel Gaussian CEO problem is given by the union of rate tuples* $(R_1, \ldots, R_L)$ *satisfying*

$$\sum_{\ell \in \mathcal{A}} R_\ell \geq \sum_{m=1}^{M} \frac{1}{2} \log \frac{1}{D_m} \left(\frac{1}{\sigma_m^2} + \sum_{\ell \in \mathcal{A}^c} \frac{\sigma_{\ell m}^2 - D_{\ell m}}{\sigma_{\ell m}^4} \right)^{-1} \\ + \sum_{m=1}^{M} \sum_{\ell \in \mathcal{A}} \frac{1}{2} \log \frac{\sigma_{\ell m}^2}{D_{\ell m}} \tag{34}$$

for all $\mathcal{A} \subseteq \{1, \ldots, L\}$, *where the union is over all* $\{D_{\ell m}\}_{\forall \ell, \forall m}$ *satisfying the following constraints*

$$\left(\frac{1}{\sigma_m^2} + \sum_{\ell=1}^{L} \frac{\sigma_{\ell m}^2 - D_{\ell m}}{\sigma_{\ell m}^4} \right)^{-1} = D_m, \ m = 1, \ldots, M \tag{35}$$

$$0 \leq D_{\ell m} \leq \sigma_{\ell m}^2, \quad \ell = 1, \ldots, L, \\ m = 1, \ldots, M \tag{36}$$

We note that since the distortion constraints in (35) are met with equality, the first $\log(\cdot)$ in (34) is always positive, and hence, we do not need a positivity operator. We obtain the rate-distortion region of the parallel Gaussian CEO problem in two steps. In the first step, we specialize the outer bound in [12] to the parallel model. In the second step, we evaluate the outer bound we obtain in the first step, and show that it matches the inner bound given in Theorem 2. The details of the proof are given in Appendix II.

Next, we consider the case $L = M = 2$, and provide an example where our outer bound strictly contains the rate-distortion region, i.e., our outer bound includes rate pairs which are outside of the rate-distortion region. In the example we provide, we assume that the following conditions hold[2]:

$$\frac{\mu_2}{\mu_1} \frac{1}{\sigma_{12}^2} < \frac{1}{\sigma_2^2} + \frac{1}{\sigma_{22}^2} - \frac{1}{D_2} \tag{37}$$

$$\frac{\mu_2}{\mu_1 - \mu_2} \frac{1}{\sigma_2^2} < \frac{1}{\sigma_{22}^2} \tag{38}$$

$$\frac{\mu_1}{\mu_1 - \mu_2} \frac{1}{D_2} < \frac{1}{\sigma_2^2} + \frac{1}{\sigma_{22}^2} \tag{39}$$

$$\frac{1}{D_1} \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_{21}^2} \right)^{-1} > \frac{\mu_1 - \mu_2}{\mu_1} D_2 \left(\frac{1}{\sigma_2^2} + \frac{1}{\sigma_{22}^2} \right) \tag{40}$$

Under the constraints in (37)-(40), the rate-distortion region $\mathcal{R}^p(D_1, D_2)$ can be characterized as follows.

Corollary 2: Assume that (37)-(40) hold. Then, we have

$$\text{T}^p = \min_{(R_1, R_2) \in \mathcal{R}^p(D_1, D_2)} \mu_1 R_1 + \mu_2 R_2 \tag{41}$$

$$= \min_{(D_{11}, D_{21}) \in \mathcal{D}_1} f_1(D_{11}, D_{21}) + \frac{\mu_2}{2} \log \frac{\sigma_2^2}{D_2} \\ + \frac{\mu_2}{2} \log \frac{1}{\sigma_{22}^2} \left(\frac{1}{\sigma_2^2} + \frac{1}{\sigma_{22}^2} - \frac{1}{D_2} \right)^{-1} \tag{42}$$

where $f_1(D_{11}, D_{21})$ *is given by*

$$f_1(D_{11}, D_{21}) = \sum_{\ell=1}^{2} \frac{\mu_\ell}{2} \log \frac{\sigma_{\ell 1}^2}{D_{\ell 1}} + \frac{\mu_2}{2} \log \frac{\sigma_1^2}{D_1} \\ + \frac{\mu_1 - \mu_2}{2} \log \frac{1}{D_1} \left(\frac{1}{\sigma_1^2} + \frac{\sigma_{21}^2 - D_{21}}{\sigma_{21}^4} \right)^{-1} \tag{43}$$

and the set $\mathcal{D}_1$ *consists of* (D_{11}, D_{21}) *pairs satisfying*

$$\frac{1}{\sigma_1^2} + \sum_{\ell=1}^{2} \frac{\sigma_{\ell 1}^2 - D_{\ell 1}}{\sigma_{\ell 1}^4} = \frac{1}{D_1} \tag{44}$$

$$0 \leq D_{\ell 1} \leq \sigma_{\ell 1}^2, \quad \ell = 1, 2 \tag{45}$$

---

[2]We note that if one selects $\sigma_m^2 = \sigma_{\ell m}^2 = \sigma^2$, $D_1 = 2/5\sigma^2$, $D_2 = 4/5\sigma^2$ and $\mu_1/\mu_2 = 4$, the four assumptions in (37)-(40) hold in addition to the original constraints on (D_1, D_2) given in (33).

The proof of Corollary 2 is given in Appendix III. Next, we find an upper bound for our outer bound in Theorem 1 as follows.

Corollary 3: Assume that (37)-(40) hold. Then, we have

$$T^+ = \min_{(R_1, R_2) \in \mathcal{R}^o(D_1, D_2)} \mu_1 R_1 + \mu_2 R_2 \tag{46}$$

$$\leq \min_{(D_{11}, D_{21}) \in \mathcal{D}_1} f_1(D_{11}, D_{21}) + \frac{\mu_2}{2} \log \frac{\mu_1}{\mu_2} \frac{1}{\sigma_{22}^2} \left(\frac{1}{\sigma_2^2} + \frac{1}{\sigma_{22}^2} \right)^{-1}$$
$$+ \frac{\mu_2}{2} \log \frac{\sigma_2^2}{D_2} + \frac{\mu_1 - \mu_2}{2} \log \frac{\mu_1}{\mu_1 - \mu_2} \frac{1}{D_2} \left(\frac{1}{\sigma_2^2} + \frac{1}{\sigma_{22}^2} \right)^{-1} \tag{47}$$

where the function $f_1(D_{11}, D_{21})$ is given by (43) and the set $\mathcal{D}_1$ is given by the union of (D_{11}, D_{21}) satisfying the constraints in (44)-(45).

The proof of Corollary 3 is given in Appendix IV.

Now, we are ready to compare our outer bound with the rate-distortion region for the parallel Gaussian model. Using Corollary 2 and Corollary 3, we have

$$T^+ - T^p \leq \frac{\mu_2}{2} \log \frac{\mu_1}{\mu_2} \frac{1}{\sigma_{22}^2} \left(\frac{1}{\sigma_2^2} + \frac{1}{\sigma_{22}^2} \right)^{-1}$$
$$+ \frac{\mu_1 - \mu_2}{2} \log \frac{\mu_1}{\mu_1 - \mu_2} \frac{1}{D_2} \left(\frac{1}{\sigma_2^2} + \frac{1}{\sigma_{22}^2} \right)^{-1}$$
$$- \frac{\mu_2}{2} \log \frac{1}{\sigma_{22}^2} \left(\frac{1}{\sigma_2^2} + \frac{1}{\sigma_{22}^2} - \frac{1}{D_2} \right)^{-1} \tag{48}$$
$$= \frac{\mu_2}{2} \log \frac{\mu_1}{\mu_2} \left(1 - \frac{1}{D_2} \left(\frac{1}{\sigma_2^2} + \frac{1}{\sigma_{22}^2} \right)^{-1} \right)$$
$$+ \frac{\mu_1 - \mu_2}{2} \log \frac{\mu_1}{\mu_1 - \mu_2} \frac{1}{D_2} \left(\frac{1}{\sigma_2^2} + \frac{1}{\sigma_{22}^2} \right)^{-1} \tag{49}$$
$$< \frac{\mu_1}{2} \log 1 \tag{50}$$
$$= 0 \tag{51}$$

where (50) follows from the facts that $\log(\cdot)$ is strictly concave, and we have

$$\frac{\mu_1}{\mu_2} \left(1 - \frac{1}{D_2} \left(\frac{1}{\sigma_2^2} + \frac{1}{\sigma_{22}^2} \right)^{-1} \right) \neq \frac{\mu_1}{\mu_1 - \mu_2} \frac{1}{D_2} \left(\frac{1}{\sigma_2^2} + \frac{1}{\sigma_{22}^2} \right)^{-1} \tag{52}$$

which is due to the assumption in (39). Equation (51) implies that there are some rate pairs (R_1, R_2) in our outer bound which are outside of the rate-distortion region of the parallel Gaussian model. Hence, our outer bound strictly contains the rate-distortion region of the vector Gaussian CEO problem. In other words, our outer bound is not tight in general.

V. PROOF OF THEOREM 1

The following theorem provides an outer bound for the rate-distortion region of the CEO problem.

Theorem 7 ([12], Theorem 1): The rate region of the CEO problem $\mathcal{R}(\mathbf{D})$ is contained in the union of rate tuples $(R_1, \ldots, R_L)$ satisfying

$$\sum_{\ell \in \mathcal{A}} R_\ell \geq I(\mathbf{X}; \{U_\ell\}_{\ell \in \mathcal{A}} | \{U_\ell\}_{\ell \in \mathcal{A}^c}) + \sum_{\ell \in \mathcal{A}} I(\mathbf{Y}_\ell; U_\ell | \mathbf{X}, W) \tag{53}$$

$\forall \mathcal{A} \subseteq \{1, \ldots, L\}$, *where the union is over all joint distributions $p(\mathbf{x}, \{\mathbf{y}_\ell, u_\ell\}_{\ell=1}^L, w)$ that can be factorized as*

$$p(\mathbf{x}, \{\mathbf{y}_\ell, u_\ell\}_{\ell=1}^L, w) = p(\mathbf{x})p(w) \prod_{\ell=1}^L p(\mathbf{y}_\ell | \mathbf{x}) p(u_\ell | \mathbf{y}_\ell, w) \tag{54}$$

and satisfies

$$\text{mmse}(\mathbf{X} | U_1, \ldots, U_L) \preceq \mathbf{D} \tag{55}$$

In [12], the outer bound is stated in a slightly different form, where there is a time-sharing random variable T involved in the description of the outer bound. However, as pointed out by [12], this time-sharing random variable T can be combined with other auxiliary random variables $(W, U_1, \ldots, U_L)$ to obtain the form of the outer bound we stated here.

We now evaluate this outer bound for the vector Gaussian CEO problem. To this end, we first provide some background information which will be used in the proof.

A. Background

Lemma 1 ([8]): Let $(U, \mathbf{X})$ be an arbitrarily correlated random vector with well-defined densities. We assume that $\text{mmse}(\mathbf{X} | U) \succ \mathbf{0}$. Then, we have

$$\mathbf{J}(\mathbf{X} | U) \succeq \text{mmse}^{-1}(\mathbf{X} | U) \tag{56}$$

which is satisfied with equality if $(U, \mathbf{X})$ is jointly Gaussian.

Next, we note the following lemma which will be used subsequently.

Lemma 2 ([18], [19]): Let $(U, \mathbf{X})$ be an arbitrary random vector, where the conditional Fisher information of $\mathbf{X}$, conditioned on U, exists. Then, we have

$$\frac{1}{2} \log |(2\pi e) \mathbf{J}^{-1}(\mathbf{X} | U)| \leq h(\mathbf{X} | U) \tag{57}$$

We also need the following lemma in the upcoming proof.

Lemma 3 ([13]): Let $(\mathbf{V}_1, \mathbf{V}_2)$ be an arbitrary random vector with finite second moments, and $\mathbf{N}$ be a zero-mean Gaussian random vector with covariance $\mathbf{\Sigma}_N$. Assume $(\mathbf{V}_1, \mathbf{V}_2)$ and $\mathbf{N}$ are independent. We have

$$\text{mmse}(\mathbf{V}_2 | \mathbf{V}_1, \mathbf{V}_2 + \mathbf{N}) = \mathbf{\Sigma}_N - \mathbf{\Sigma}_N \mathbf{J}(\mathbf{V}_2 + \mathbf{N} | \mathbf{V}_1) \mathbf{\Sigma}_N \tag{58}$$

B. Proof

Here, we consider the rate bounds in (53) and obtain a lower bound for them for a given $(W, U_1, \ldots, U_L)$. First, we consider the following mutual information terms

$$I(\mathbf{Y}_\ell; U_\ell | \mathbf{X}, W) = h(\mathbf{Y}_\ell | \mathbf{X}, W) - h(\mathbf{Y}_\ell | \mathbf{X}, W, U_\ell) \tag{59}$$
$$= h(\mathbf{Y}_\ell | \mathbf{X}) - h(\mathbf{Y}_\ell | \mathbf{X}, W, U_\ell) \tag{60}$$
$$= \frac{1}{2} \log |(2\pi e) \mathbf{\Sigma}_\ell| - h(\mathbf{Y}_\ell | \mathbf{X}, W, U_\ell) \tag{61}$$

Using Lemma 2 and the fact that jointly Gaussian $(\mathbf{X}, W, U_\ell, \mathbf{Y}_\ell)$ maximizes $h(\mathbf{Y}_\ell | \mathbf{X}, W, U_\ell)$, we have the following bounds for the second term in (61)

$$\frac{1}{2} \log |(2\pi e) \mathbf{J}^{-1}(\mathbf{Y}_\ell | \mathbf{X}, W, U_\ell)| \leq h(\mathbf{Y}_\ell | \mathbf{X}, W, U_\ell)$$
$$\leq \frac{1}{2} \log |(2\pi e) \text{mmse}(\mathbf{Y}_\ell | \mathbf{X}, W, U_\ell)| \tag{62}$$

Next, we define the function $\mathbf{D}_\ell(\alpha_\ell)$ as follows

$$\mathbf{D}_\ell(\alpha_\ell) = \alpha_\ell \mathbf{J}^{-1}(\mathbf{Y}_\ell|\mathbf{X}, W, U_\ell) + \bar{\alpha}_\ell \mathrm{mmse}(\mathbf{Y}_\ell|\mathbf{X}, W, U_\ell) \tag{63}$$

where $\alpha_\ell = 1 - \bar{\alpha}_\ell \in [0, 1]$. Using the function in (63), the bounds in (62) can be expressed as follows

$$\frac{1}{2}\log|(2\pi e)\mathbf{D}_\ell(1)| \le h(\mathbf{Y}_\ell|\mathbf{X}, W, U_\ell) \le \frac{1}{2}\log|(2\pi e)\mathbf{D}_\ell(0)| \tag{64}$$

Since $\log|(2\pi e)\mathbf{D}_\ell(\alpha_\ell)|$ is continuous in α_ℓ, due to the intermediate value theorem, there exists an $\alpha_\ell^* = 1 - \bar{\alpha}_\ell^* \in [0, 1]$ such that

$$h(\mathbf{Y}_\ell|\mathbf{X}, W, U_\ell) = \frac{1}{2}\log|(2\pi e)\mathbf{D}_\ell(\alpha_\ell^*)| \tag{65}$$

Hence, using (65) in (61), we have

$$I(\mathbf{Y}_\ell; U_\ell|\mathbf{X}, W) = \frac{1}{2}\log\frac{|\mathbf{\Sigma}_\ell|}{|\mathbf{D}_\ell(\alpha_\ell^*)|}, \quad \ell = 1, \dots, L \tag{66}$$

We note the following bounds on $\mathbf{D}_\ell(\alpha_\ell^*)$

$$\mathbf{J}^{-1}(\mathbf{Y}_\ell|\mathbf{X}, W, U_\ell) \preceq \mathbf{D}_\ell(\alpha_\ell^*) \preceq \mathrm{mmse}(\mathbf{Y}_\ell|\mathbf{X}, W, U_\ell) \tag{67}$$
$$\preceq \mathrm{mmse}(\mathbf{Y}_\ell|\mathbf{X}) \tag{68}$$
$$= \mathbf{\Sigma}_\ell \tag{69}$$

where (67) is due to Lemma 1 and (68) comes from the fact that conditioning reduces the MMSE matrix in the positive semi-definite ordering sense.

Next, we consider the following mutual information term

$$I(\mathbf{X}; \{U_\ell\}_{\ell \in \mathcal{A}}|\{U_\ell\}_{\ell \in \mathcal{A}^c})$$
$$= h(\mathbf{X}|\{U_\ell\}_{\ell \in \mathcal{A}^c}) - h(\mathbf{X}|U_1, \dots, U_L) \tag{70}$$
$$\ge h(\mathbf{X}|\{U_\ell\}_{\ell \in \mathcal{A}^c}) - \frac{1}{2}\log|(2\pi e)\mathrm{mmse}(\mathbf{X}|U_1, \dots, U_L)| \tag{71}$$
$$\ge h(\mathbf{X}|\{U_\ell\}_{\ell \in \mathcal{A}^c}) - \frac{1}{2}\log|(2\pi e)\mathbf{D}| \tag{72}$$
$$\ge h(\mathbf{X}|\{U_\ell\}_{\ell \in \mathcal{A}^c}, W) - \frac{1}{2}\log|(2\pi e)\mathbf{D}| \tag{73}$$

where (71) comes from the fact that $h(\mathbf{X}|U_1, \dots, U_L)$ is maximized by jointly Gaussian $(\mathbf{X}, U_1, \dots, U_L)$, (72) follows from the monotonicity of $\log|\cdot|$ function in positive semi-definite matrices in conjunction with the distortion constraint in (55), and (73) comes from the fact that conditioning cannot increase entropy.

Next, we obtain a lower bound for $h(\mathbf{X}|\{U_\ell\}_{\ell \in \mathcal{A}^c}, W)$. To this end, in view of Lemma 2, we note the following lower bound on $h(\mathbf{X}|\{U_\ell\}_{\ell \in \mathcal{A}^c}, W)$

$$h(\mathbf{X}|\{U_\ell\}_{\ell \in \mathcal{A}^c}, W) \ge \frac{1}{2}\log|(2\pi e)\mathbf{J}^{-1}(\mathbf{X}|\{U_\ell\}_{\ell \in \mathcal{A}^c}, W)| \tag{74}$$

which implies that a lower bound on $\mathbf{J}^{-1}(\mathbf{X}|\{U_\ell\}_{\ell \in \mathcal{A}^c}, W)$ will yield a lower bound for $h(\mathbf{X}|\{U_\ell\}_{\ell \in \mathcal{A}^c}, W)$. To obtain a lower bound for $\mathbf{J}^{-1}(\mathbf{X}|\{U_\ell\}_{\ell \in \mathcal{A}^c}, W)$, we will use the connection between the Fisher information and the MMSE given in Lemma 3. To this end, we note that $\mathbf{X}$ can be decomposed as (see (107) in Appendix I-A)

$$\mathbf{X} = \sum_{\ell \in \mathcal{A}^c} \mathbf{A}_\ell \mathbf{Y}_\ell + \mathbf{N}_{\mathcal{A}^c} \tag{75}$$

where the matrices $\{\mathbf{A}_\ell\}_{\ell \in \mathcal{A}^c}$ are given by (see (109) in Appendix I-A)

$$\mathbf{A}_\ell = \mathbf{\Sigma}_{\mathcal{A}^c}\mathbf{\Sigma}_\ell^{-1}, \quad \ell \in \mathcal{A}^c \tag{76}$$

In (75), $\mathbf{N}_{\mathcal{A}^c}$ is a zero-mean Gaussian vector with covariance matrix (see (108) in Appendix I-A)

$$\mathbf{\Sigma}_{\mathcal{A}^c} = \left(\mathbf{K}_X^{-1} + \sum_{\ell \in \mathcal{A}^c} \mathbf{\Sigma}_\ell^{-1}\right)^{-1} \tag{77}$$

We also note that $\mathbf{N}_{\mathcal{A}^c}$ is independent of $(\{\mathbf{Y}_\ell, U_\ell\}_{\ell \in \mathcal{A}^c}, W)$ which implies the following Markov chain

$$\{U_\ell\}_{\ell \in \mathcal{A}^c}, W \to \sum_{\ell \in \mathcal{A}^c} \mathbf{A}_\ell \mathbf{Y}_\ell \to \mathbf{X} = \sum_{\ell \in \mathcal{A}^c} \mathbf{A}_\ell \mathbf{Y}_\ell + \mathbf{N}_{\mathcal{A}^c} \tag{78}$$

In view of this Markov chain, due to Lemma 3, we have

$$\mathrm{mmse}(\mathbf{S}_{\mathcal{A}^c}|\mathbf{X}, \{U_\ell\}_{\ell \in \mathcal{A}^c}, W)$$
$$= \mathbf{\Sigma}_{\mathcal{A}^c} - \mathbf{\Sigma}_{\mathcal{A}^c}\mathbf{J}(\mathbf{X}|\{U_\ell\}_{\ell \in \mathcal{A}^c}, W)\mathbf{\Sigma}_{\mathcal{A}^c} \tag{79}$$

where we define $\mathbf{S}_{\mathcal{A}^c}$ as follows

$$\mathbf{S}_{\mathcal{A}^c} = \sum_{\ell \in \mathcal{A}^c} \mathbf{A}_\ell \mathbf{Y}_\ell \tag{80}$$

Next, we obtain the MMSE matrix in (79) in terms of the individual MMSE matrices $\{\mathrm{mmse}(\mathbf{Y}_\ell|\mathbf{X}, U_\ell, W)\}_{\ell \in \mathcal{A}^c}$ as given in the following lemma.

Lemma 4: Under the current conditions, we have

$$\mathrm{mmse}(\mathbf{S}_{\mathcal{A}^c}|\mathbf{X}, \{U_\ell\}_{\ell \in \mathcal{A}^c}, W)$$
$$= \sum_{\ell \in \mathcal{A}^c} \mathbf{A}_\ell \mathrm{mmse}(\mathbf{Y}_\ell|\mathbf{X}, W, U_\ell)\mathbf{A}_\ell^\top \tag{81}$$

The proof of this lemma is given in Appendix V.

Hence, using Lemma 4 in (79), we get

$$\mathbf{\Sigma}_{\mathcal{A}^c} - \mathbf{\Sigma}_{\mathcal{A}^c}\mathbf{J}(\mathbf{X}|\{U_\ell\}_{\ell \in \mathcal{A}^c}, W)\mathbf{\Sigma}_{\mathcal{A}^c}$$
$$= \sum_{\ell \in \mathcal{A}^c} \mathbf{A}_\ell \mathrm{mmse}(\mathbf{Y}_\ell|\mathbf{X}, W, U_\ell)\mathbf{A}_\ell^\top \tag{82}$$
$$\succeq \sum_{\ell \in \mathcal{A}^c} \mathbf{A}_\ell \mathbf{D}_\ell(\alpha_\ell^*)\mathbf{A}_\ell^\top \tag{83}$$
$$= \mathbf{\Sigma}_{\mathcal{A}^c}\left(\sum_{\ell \in \mathcal{A}^c} \mathbf{\Sigma}_\ell^{-1}\mathbf{D}_\ell(\alpha_\ell^*)\mathbf{\Sigma}_\ell^{-1}\right)\mathbf{\Sigma}_{\mathcal{A}^c} \tag{84}$$

where (83) is due to (67), and in (84), we use the definition of $\mathbf{A}_\ell$ given in (76). We note that (84) implies

$$\mathbf{J}^{-1}(\mathbf{X}|\{U_\ell\}_{\ell \in \mathcal{A}^c}, W) \succeq \left(\mathbf{\Sigma}_{\mathcal{A}^c}^{-1} - \sum_{\ell \in \mathcal{A}^c} \mathbf{\Sigma}_\ell^{-1}\mathbf{D}_\ell(\alpha_\ell^*)\mathbf{\Sigma}_\ell^{-1}\right)^{-1} \tag{85}$$
$$= \left(\mathbf{K}_X^{-1} + \sum_{\ell \in \mathcal{A}^c} \mathbf{\Sigma}_\ell^{-1} - \sum_{\ell \in \mathcal{A}^c} \mathbf{\Sigma}_\ell^{-1}\mathbf{D}_\ell(\alpha_\ell^*)\mathbf{\Sigma}_\ell^{-1}\right)^{-1} \tag{86}$$

where (86) comes from the definition of $\mathbf{\Sigma}_{\mathcal{A}^c}$ in (77). In view of (74) and (86), we have the following lower bound

for $h(\mathbf{X}|\{U_\ell\}_{\ell\in\mathcal{A}^c}, W)$ as follows

$$
h(\mathbf{X}|\{U_\ell\}_{\ell\in\mathcal{A}^c}, W) \geq \frac{1}{2} \log \left| (2\pi e) \left(\mathbf{K}_X^{-1} + \sum_{\ell\in\mathcal{A}^c} \boldsymbol{\Sigma}_\ell^{-1} \right.\right.
$$
$$
\left.\left. - \sum_{\ell\in\mathcal{A}^c} \boldsymbol{\Sigma}_\ell^{-1} \mathbf{D}_\ell(\alpha_\ell^*) \boldsymbol{\Sigma}_\ell^{-1} \right)^{-1} \right| \tag{87}
$$

Hence, using (87) in (73), we get

$$
I(\mathbf{X}; \{U_\ell\}_{\ell\in\mathcal{A}}|\{U_\ell\}_{\ell\in\mathcal{A}^c}) \geq \frac{1}{2} \log
$$
$$
\frac{\left| \left(\mathbf{K}_X^{-1} + \sum_{\ell\in\mathcal{A}^c} \boldsymbol{\Sigma}_\ell^{-1} - \sum_{\ell\in\mathcal{A}^c} \boldsymbol{\Sigma}_\ell^{-1} \mathbf{D}_\ell(\alpha_\ell^*) \boldsymbol{\Sigma}_\ell^{-1} \right)^{-1} \right|}{|\mathbf{D}|} \tag{88}
$$

Moreover, using the non-negativity of the mutual information, we can improve this lower bound as follows

$$
I(\mathbf{X}; \{U_\ell\}_{\ell\in\mathcal{A}}|\{U_\ell\}_{\ell\in\mathcal{A}^c}) \geq \frac{1}{2} \log^+
$$
$$
\frac{\left| \left(\mathbf{K}_X^{-1} + \sum_{\ell\in\mathcal{A}^c} \boldsymbol{\Sigma}_\ell^{-1} - \sum_{\ell\in\mathcal{A}^c} \boldsymbol{\Sigma}_\ell^{-1} \mathbf{D}_\ell(\alpha_\ell^*) \boldsymbol{\Sigma}_\ell^{-1} \right)^{-1} \right|}{|\mathbf{D}|} \tag{89}
$$

where $\log^+ x = \max(\log x, 0)$. Using (66) and (89) in the rate bounds given in (53), we get

$$
\sum_{\ell\in\mathcal{A}} R_\ell \geq \frac{1}{2} \log^+
$$
$$
\frac{\left| \left(\mathbf{K}_X^{-1} + \sum_{\ell\in\mathcal{A}^c} \boldsymbol{\Sigma}_\ell^{-1} - \sum_{\ell\in\mathcal{A}^c} \boldsymbol{\Sigma}_\ell^{-1} \mathbf{D}_\ell(\alpha_\ell^*) \boldsymbol{\Sigma}_\ell^{-1} \right)^{-1} \right|}{|\mathbf{D}|}
$$
$$
+ \sum_{\ell\in\mathcal{A}} \frac{1}{2} \log \frac{|\boldsymbol{\Sigma}_\ell|}{|\mathbf{D}_\ell(\alpha_\ell^*)|} \tag{90}
$$

Next, we establish a connection between $\mathbf{D}$ and $(\mathbf{D}_1(\alpha_1^*), \ldots, \mathbf{D}_L(\alpha_L^*))$. To this end, by taking $\mathcal{A}^c = \{1, \ldots, L\}$ in (86), we get

$$
\left(\mathbf{K}_X^{-1} + \sum_{\ell=1}^L \boldsymbol{\Sigma}_\ell^{-1} - \sum_{\ell=1}^L \boldsymbol{\Sigma}_\ell^{-1} \mathbf{D}_\ell(\alpha_\ell^*) \boldsymbol{\Sigma}_\ell^{-1} \right)^{-1}
$$
$$
\preceq \mathbf{J}^{-1}(\mathbf{X}|\{U_\ell\}_{\ell=1}^L, W) \tag{91}
$$
$$
\preceq \text{mmse}(\mathbf{X}|\{U_\ell\}_{\ell=1}^L, W) \tag{92}
$$
$$
\preceq \text{mmse}(\mathbf{X}|\{U_\ell\}_{\ell=1}^L) \tag{93}
$$
$$
\preceq \mathbf{D} \tag{94}
$$

where (92) is due to Lemma 1, (93) comes from the fact that conditioning reduces the MMSE matrix in the positive semi-definite ordering sense, and (94) follows from the distortion constraint in (55). Hence, in view of (90) and (94), we show that the rate region of the vector Gaussian CEO problem is

included in the union of rate tuples $(R_1, \ldots, R_L)$ satisfying

$$
\sum_{\ell\in\mathcal{A}} R_\ell \geq \frac{1}{2} \log^+
$$
$$
\frac{\left| \left(\mathbf{K}_X^{-1} + \sum_{\ell\in\mathcal{A}^c} \boldsymbol{\Sigma}_\ell^{-1} - \sum_{\ell\in\mathcal{A}^c} \boldsymbol{\Sigma}_\ell^{-1} \mathbf{D}_\ell(\alpha_\ell^*) \boldsymbol{\Sigma}_\ell^{-1} \right)^{-1} \right|}{|\mathbf{D}|}
$$
$$
+ \sum_{\ell\in\mathcal{A}} \frac{1}{2} \log \frac{|\boldsymbol{\Sigma}_\ell|}{|\mathbf{D}_\ell(\alpha_\ell^*)|} \tag{95}
$$

for all $\mathcal{A} \subseteq \{1, \ldots, L\}$, where the union is over all positive semi-definite matrices $\mathbf{D}_1(\alpha_1^*), \ldots, \mathbf{D}_L(\alpha_L^*)$ satisfying the following orders

$$
\left(\mathbf{K}_X^{-1} + \sum_{\ell=1}^L \boldsymbol{\Sigma}_\ell^{-1} - \sum_{\ell=1}^L \boldsymbol{\Sigma}_\ell^{-1} \mathbf{D}_\ell(\alpha_\ell^*) \boldsymbol{\Sigma}_\ell^{-1} \right)^{-1} \preceq \mathbf{D} \tag{96}
$$
$$
\mathbf{0} \preceq \mathbf{D}_\ell(\alpha_\ell^*) \preceq \boldsymbol{\Sigma}_\ell \tag{97}
$$

where $\ell = 1, \ldots, L$. The orders in (97) follow from (69). The region given in Theorem 1 can be obtained from the outer bound described in (95)-(97) by setting $\mathbf{D}_\ell(\alpha_\ell^*) = \mathbf{D}_\ell$, which completes the proof of Theorem 1.

VI. GENERALIZATION OF THE BOUNDS

In this section, we consider the most general form of the vector Gaussian CEO problem, and generalize the outer and the inner bounds in Theorem 1 and Theorem 2, respectively. In the most general form of the vector Gaussian CEO problem, the observations at the sensors are given by

$$
\mathbf{Y}_\ell = \mathbf{H}_\ell \mathbf{X} + \mathbf{N}_\ell, \quad \ell = 1, \ldots, L \tag{98}
$$

where $\{\mathbf{N}_\ell\}_{\ell=1}^L$ are i.i.d. zero-mean Gaussian random vectors with identity covariance matrices. We note that the general form for the observations in (98) cover the model in (1) we studied so far. All definitions we introduced in Section II hold for the general model defined by (98) except for the distortion constraints in (6). In the general model, the distortion $\mathbf{D}$ is assumed to satisfy

$$
\left(\mathbf{K}_X^{-1} + \sum_{\ell=1}^L \mathbf{H}_\ell^\top \mathbf{H}_\ell \right)^{-1} \preceq \mathbf{D} \preceq \mathbf{K}_X \tag{99}
$$

where the left hand-side is the MMSE matrix obtained when the CEO unit has access to all observations in (98). Similar to the model given by (1), here also, imposing the lower bound constraint on $\mathbf{D}$ in (99) does not incur any loss of generality, while the upper bound constraint on $\mathbf{D}$ in (99) might incur some loss of generality.

Now, we provide an outer bound for the rate-distortion region $\mathcal{R}(\mathbf{D})$ for the general model given by (98), which, in fact, corresponds to the generalization of the outer bound in Theorem 1 to the most general form of the vector Gaussian CEO problem.

Theorem 8: An outer bound for the rate-distortion region of the general vector Gaussian CEO problem is given by the

union of rate tuples $(R_1, \ldots, R_L)$ satisfying

$$\sum_{\ell \in \mathcal{A}} R_\ell \geq \frac{1}{2} \log^+ \frac{\left| \left(\mathbf{K}_X^{-1} + \sum_{\ell \in \mathcal{A}^c} \mathbf{H}_\ell^\top (\mathbf{I} - \mathbf{D}_\ell) \mathbf{H}_\ell \right)^{-1} \right|}{|\mathbf{D}|}$$
$$+ \sum_{\ell \in \mathcal{A}} \frac{1}{2} \log \frac{1}{|\mathbf{D}_\ell|} \qquad (100)$$

for all $\mathcal{A} \subseteq \{1, \ldots, L\}$, where the union is over all positive semi-definite matrices $\{\mathbf{D}_\ell\}_{\ell=1}^L$ satisfying the following constraints

$$\left(\mathbf{K}_X^{-1} + \sum_{\ell=1}^L \mathbf{H}_\ell^\top (\mathbf{I} - \mathbf{D}_\ell) \mathbf{H}_\ell \right)^{-1} \preceq \mathbf{D} \qquad (101)$$

$$\mathbf{0} \preceq \mathbf{D}_\ell \preceq \mathbf{I}, \quad \ell = 1, \ldots, L \qquad (102)$$

We prove Theorem 8 in two steps. In the first step, we enhance (improve) the observations at the sensors in a way that the enhanced observations are in a similar form given by (1). In the next step, we use Theorem 1 to obtain an outer bound for the enhanced model, and from this outer bound, we obtain Theorem 8 by using some limiting arguments. The details of the proof can be found in Appendix VI.

Now, we introduce an inner bound for the rate-distortion region $\mathcal{R}(\mathbf{D})$ for the general model given by (98), which, in fact, corresponds to the generalization of the inner bound in Theorem 2 to the most general form of the vector Gaussian CEO problem.

Theorem 9: An inner bound for the rate-distortion region of the general vector Gaussian CEO problem is given by the union of rate tuples $(R_1, \ldots, R_L)$ satisfying

$$\sum_{\ell \in \mathcal{A}} R_\ell \geq \frac{1}{2} \log \frac{\left| \left(\mathbf{K}_X^{-1} + \sum_{\ell \in \mathcal{A}^c} \mathbf{H}_\ell^\top (\mathbf{I} - \mathbf{D}_\ell) \mathbf{H}_\ell \right)^{-1} \right|}{\left| \left(\mathbf{K}_X^{-1} + \sum_{\ell=1}^L \mathbf{H}_\ell^\top (\mathbf{I} - \mathbf{D}_\ell) \mathbf{H}_\ell \right)^{-1} \right|}$$
$$+ \sum_{\ell \in \mathcal{A}} \frac{1}{2} \log \frac{1}{|\mathbf{D}_\ell|} \qquad (103)$$

for all $\mathcal{A} \subseteq \{1, \ldots, L\}$, where the union is over all positive semi-definite matrices $\{\mathbf{D}_\ell\}_{\ell=1}^L$ satisfying

$$\left(\mathbf{K}_X^{-1} + \sum_{\ell=1}^L \mathbf{H}_\ell^\top (\mathbf{I} - \mathbf{D}_\ell) \mathbf{H}_\ell \right)^{-1} \preceq \mathbf{D} \qquad (104)$$

$$\mathbf{0} \preceq \mathbf{D}_\ell \preceq \mathbf{I}, \quad \ell = 1, \ldots, L \qquad (105)$$

The proof of Theorem 9 is given in Appendix VII. We obtain this inner bound by evaluating the Berger-Tung inner bound [5] by jointly Gaussian auxiliary random variables.

We note that since the outer and the inner bounds in Theorem 8 and Theorem 9 correspond to the generalizations of the outer and inner bounds in Theorem 1 and Theorem 2, respectively, our previous comments and remarks about Theorem 1 and Theorem 2 hold for Theorem 8 and Theorem 9 as well. In particular, similar to Theorem 1 and Theorem 2, we can provide alternative characterizations for Theorem 8 and Theorem 9 as well. Moreover, similar to Theorem 1 and Theorem 2, the bounds in Theorem 8 and Theorem 9 match when the boundary of the outer bound in Theorem 8 can be

described by the matrices $\{\mathbf{D}_\ell\}_{\ell=1}^L$ that satisfy the distortion constraint in (101) with equality.

VII. CONCLUSIONS

In this paper, we study the vector Gaussian CEO problem and provide an outer bound for its rate-distortion region. We obtain our outer bound by evaluating the rather general outer bound in [12]. We accomplish this evaluation by using a technique that relies on the de Bruijn identity along with the properties of the MMSE and Fisher information. Next, we investigate the tightness of outer bound. Despite being tight for certain cases, we show that our outer bound does not provide the exact rate-distortion region in general. We show this by providing an example where our outer bound strictly includes the rate-distortion region.

APPENDIX I
DISTORTION LIMITS

In this appendix, we first note some facts about Gaussian random vectors that are used throughout the paper.

A. Gaussian Random Vectors

Let $\mathbf{T}$ be a zero-mean Gaussian random vector with covariance matrix $\mathbf{\Sigma}_T \succ \mathbf{0}$. We define the Gaussian random vectors $\{\mathbf{T}_\ell\}_{\ell=1}^L$ as

$$\mathbf{T}_\ell = \mathbf{H}_\ell \mathbf{T} + \mathbf{N}_\ell \qquad (106)$$

where $\{\mathbf{N}_\ell\}_{\ell=1}^L$ are zero-mean independent Gaussian random vectors with covariance matrices $\{\mathbf{\Sigma}_\ell\}_{\ell=1}^L$, which are also independent of $\mathbf{T}$. We assume $\mathbf{\Sigma}_\ell \succ \mathbf{0}, \ \ell = 1, \ldots, L$.

For any subset $\mathcal{A} \subseteq \{1, \ldots, L\}$, we have

$$\mathbf{T} = \sum_{\ell \in \mathcal{A}} \mathbf{A}_\ell \mathbf{T}_\ell + \mathbf{N}_\mathcal{A} \qquad (107)$$

where $\mathbf{N}_\mathcal{A}$ is a zero-mean Gaussian random vector with covariance matrix $\mathbf{\Sigma}_\mathcal{A}$ given by

$$\mathbf{\Sigma}_\mathcal{A} = \left(\mathbf{\Sigma}_T^{-1} + \sum_{\ell \in \mathcal{A}} \mathbf{H}_\ell^\top \mathbf{\Sigma}_\ell^{-1} \mathbf{H}_\ell \right)^{-1} \qquad (108)$$

and is independent of $\{\mathbf{T}_\ell\}_{\ell \in \mathcal{A}}$. The matrices $\{\mathbf{A}_\ell\}_{\ell \in \mathcal{A}}$ are given by

$$\mathbf{A}_\ell = \mathbf{\Sigma}_\mathcal{A} \mathbf{H}_\ell^\top \mathbf{\Sigma}_\ell^{-1}, \quad \ell \in \mathcal{A} \qquad (109)$$

The decomposition in (107) follows from the MMSE estimation of Gaussian random vectors, which is equivalent to the linear MMSE estimation. In particular, we have

$$\hat{\mathbf{T}} = E\left[\mathbf{T} | \{\mathbf{T}_\ell\}_{\ell \in \mathcal{A}} \right] = \sum_{\ell \in \mathcal{A}} \mathbf{A}_\ell \mathbf{T}_\ell \qquad (110)$$

which is the MMSE, equivalently the linear MMSE, estimator of $\mathbf{T}$ from $\{\mathbf{T}_\ell\}_{\ell=1}^L$. The error in estimation is $\mathbf{N}_\mathcal{A}$, and the MMSE matrix is

$$\mathrm{mmse}(\mathbf{T} | \{\mathbf{T}_\ell\}_{\ell \in \mathcal{A}}) = \mathbf{\Sigma}_\mathcal{A} \qquad (111)$$

B. Regarding (6)

We first obtain the lower bound on the distortion constraint $\mathbf{D}$ in (6) as follows

$$\text{mmse}(\mathbf{X}_i|B_1^n,\ldots,B_L^n)$$
$$\succeq \text{mmse}(\mathbf{X}_i|B_1^n,\ldots,B_L^n,\mathbf{Y}_1^n,\ldots,\mathbf{Y}_L^n) \quad (112)$$
$$= \text{mmse}(\mathbf{X}_i|\mathbf{Y}_1^n,\ldots,\mathbf{Y}_L^n) \quad (113)$$
$$= \text{mmse}(\mathbf{X}_i|\mathbf{Y}_{1,i},\ldots,\mathbf{Y}_{L,i}) \quad (114)$$
$$= \left(\mathbf{K}_X^{-1} + \sum_{\ell=1}^{L}\mathbf{\Sigma}_\ell^{-1}\right)^{-1} \quad (115)$$

where (112) follows from the fact that conditioning reduces the MMSE matrix in the positive semi-definite ordering sense, (113) is due to the fact that B_ℓ^n is a function of $\mathbf{Y}_\ell^n$, (114) comes from the independence of $(\mathbf{X}_i, \mathbf{Y}_{1,i},\ldots,\mathbf{Y}_{L,i})$ across time, and (115) is due to (108) and (111). Hence, (115) implies that imposing the constraint $\mathbf{D} \succeq \left(\mathbf{K}_X^{-1} + \sum_{\ell=1}^{L}\mathbf{\Sigma}_\ell^{-1}\right)^{-1}$ does not incur any loss of generality.

Next, we consider the upper bound on the distortion constraint in (55). To this end, we note the following order

$$\text{mmse}(\mathbf{X}_i|B_1^n,\ldots,B_L^n) \preceq \text{mmse}(\mathbf{X}) = \mathbf{K}_X \quad (116)$$

where we use the fact that conditioning reduces the MMSE matrix in the positive semi-definite ordering sense. Equation (116) implies that all $(n, R_1,\ldots,R_L)$ codes achieve a distortion which is smaller than $\mathbf{K}_X$. In other words, if $\hat{\mathbf{D}}$ is the distortion achieved by a specific code, we always have $\hat{\mathbf{D}} \preceq \mathbf{K}_X$. In spite of this fact, we still cannot impose the constraint $\mathbf{D} \preceq \mathbf{K}_X$ without loss of generality. To demonstrate this point, assume that $\mathbf{K}_X - \mathbf{D}$ is indefinite. Hence, to be able to impose the constraint $\mathbf{D} \preceq \mathbf{K}_X$, we should find a new distortion constraint $\mathbf{D}'$ which satisfies $\mathbf{D}' \preceq \{\mathbf{D}, \mathbf{K}_X\}$ and the rate-distortion regions $\mathcal{R}(\mathbf{D})$ and $\mathcal{R}(\mathbf{D}')$ are identical. In other words, there needs to be a distortion matrix $\mathbf{D}' \preceq \{\mathbf{D}, \mathbf{K}_X\}$, and for any code achieving a distortion $\hat{\mathbf{D}} \preceq \mathbf{D}$, we also have $\hat{\mathbf{D}} \preceq \mathbf{D}'$. However, as we will show now, this is not possible in general. Assume that there exist two codes achieving the distortion $\hat{\mathbf{D}}_j$, $j = 1, 2$, where $\hat{\mathbf{D}}_j \preceq \{\mathbf{D}, \mathbf{K}_X\}$. Hence, $\mathbf{D}'$ needs to satisfy

$$\{\hat{\mathbf{D}}_1, \hat{\mathbf{D}}_2\} \preceq \mathbf{D}' \preceq \{\mathbf{D}, \mathbf{K}_X\} \quad (117)$$

However, there are cases where it is impossible to find a matrix $\mathbf{D}'$ satisfying the order in (117) as shown in [20, Appendix I] by a counter-example. Consequently, imposing the constraint $\mathbf{D} \preceq \mathbf{K}_X$ might incur some loss of generality.

APPENDIX II
PROOF OF THEOREM 6

We prove Theorem 6 in two steps. In the first step, we specialize the outer bound in [12] to the parallel model defined by the following joint distribution

$$p(x^M, \{y_\ell^M\}_{\ell=1}^L) = \prod_{m=1}^{M} p(x_m) \prod_{\ell=1}^{L} p(y_{\ell m}|x_m) \quad (118)$$

Next, we evaluate the outer bound we obtain in the first step, and show that it can be attained by the inner bound provided in Theorem 2.

A. A General Outer Bound

First, we restate the outer bound in [12] for the parallel model satisfying (118) as follows.

Theorem 10 ([12], Theorem 1): We have $\mathcal{R}^p(\{D_m\}_{m=1}^M) \subseteq \mathcal{R}^{p-o}(\{D_m\}_{m=1}^M)$, where $\mathcal{R}^{p-o}(\{D_m\}_{m=1}^M)$ is given by the union of rate tuples $(R_1,\ldots,R_L)$ satisfying

$$\sum_{\ell\in\mathcal{A}} R_\ell \geq I(X^M; \{U_\ell\}_{\ell\in\mathcal{A}}|\{U_\ell\}_{\ell\in\mathcal{A}^c})$$
$$+ \sum_{\ell\in\mathcal{A}} I(U_\ell; Y_\ell^M|X^M, W) \quad (119)$$

for all $\mathcal{A} \subseteq \{1,\ldots,L\}$, where the union is over all $\{U_\ell\}_{\ell=1}^L$ satisfying

$$p(x^M, \{y_\ell^M\}_{\ell=1}^L, \{u_\ell\}_{\ell=1}^L, w)$$
$$= p(w)\prod_{m=1}^{M} p(x_m) \prod_{\ell=1}^{L} p(y_{\ell m}|x_m)p(u_\ell|w, y_\ell^M) \quad (120)$$

and

$$\text{mmse}(X_m|U_1,\ldots,U_L) \leq D_m, \quad m = 1,\ldots,M \quad (121)$$

Next, we define the following auxiliary random variables

$$U_{\ell m} = U_\ell X^{m-1}, \quad \ell = 1,\ldots,L, \quad m = 1,\ldots,M \quad (122)$$
$$W_m = W X^{m-1}\{Y_{\ell,m+1}^M\}_{\ell=1}^L, \quad m = 1,\ldots,M \quad (123)$$

Using these auxiliary random variables, we will find lower bounds for the rate constraints in (119). We start with the following term

$$I(X^M; \{U_\ell\}_{\ell\in\mathcal{A}}|\{U_\ell\}_{\ell\in\mathcal{A}^c})$$
$$= \sum_{m=1}^{M} I(X_m; \{U_\ell\}_{\ell\in\mathcal{A}}|\{U_\ell\}_{\ell\in\mathcal{A}^c}, X^{m-1}) \quad (124)$$
$$= \sum_{m=1}^{M} I(X_m; \{U_{\ell m}\}_{\ell\in\mathcal{A}}|\{U_{\ell m}\}_{\ell\in\mathcal{A}^c}) \quad (125)$$

Next, we consider the following term

$$I(U_\ell; Y_\ell^M|X^M, W)$$
$$= h(U_\ell|X^M, W) - h(U_\ell|X^M, W, Y_\ell^M) \quad (126)$$
$$\geq h(U_\ell|X^M, W, \{Y_j^M\}_{j=1,j\neq\ell}^L)$$
$$- h(U_\ell|X^M, W, Y_\ell^M) \quad (127)$$
$$= h(U_\ell|X^M, W, \{Y_j^M\}_{j=1,j\neq\ell}^L)$$
$$- h(U_\ell|X^M, W, \{Y_j^M\}_{j=1}^L) \quad (128)$$
$$= I(U_\ell; Y_\ell^M|X^M, W, \{Y_j^M\}_{j=1,j\neq\ell}^L) \quad (129)$$
$$= \sum_{m=1}^{M} I(U_\ell; Y_{\ell m}|X^M, W, \{Y_j^M\}_{j=1,j\neq\ell}^L, Y_{\ell,m+1}^M) \quad (130)$$
$$= \sum_{m=1}^{M} h(Y_{\ell m}|X^M, W, \{Y_j^M\}_{j=1,j\neq\ell}^L, Y_{\ell,m+1}^M)$$
$$- h(Y_{\ell m}|X^M, W, \{Y_j^M\}_{j=1,j\neq\ell}^L, Y_{\ell,m+1}^M, U_\ell) \quad (131)$$

$$= \sum_{m=1}^{M} h(Y_{\ell m} | X^m, W, \{Y_{j,m+1}^M\}_{j=1}^L)$$
$$\quad - h(Y_{\ell m} | X^M, W, \{Y_j^M\}_{j=1, j\neq \ell}^L, Y_{\ell,m+1}^M, U_\ell) \quad (132)$$

$$\geq \sum_{m=1}^{M} h(Y_{\ell m} | X^m, W, \{Y_{j,m+1}^M\}_{j=1}^L)$$
$$\quad - h(Y_{\ell m} | X^m, W, \{Y_{j,m+1}^M\}_{j=1}^L, U_\ell) \quad (133)$$

$$= \sum_{m=1}^{M} I(U_\ell; Y_{\ell m} | X^m, W, \{Y_{j,m+1}^M\}_{j=1}^L) \quad (134)$$

$$= \sum_{m=1}^{M} I(U_{\ell m}; Y_{\ell m} | X_m, W_m) \quad (135)$$

where (127) follows from the fact that conditioning cannot increase entropy, (128) and (132) come from the following Markov chains

$$U_\ell \to W, Y_\ell^M \to X^M, \{Y_j^M\}_{j=1, j\neq \ell}^L \quad (136)$$
$$Y_{\ell m} \to X_m \to W, X^{m-1}, X_{m+1}^M, \{Y_j^M\}_{j=1, j\neq \ell}^L, Y_{\ell,m+1}^M \quad (137)$$

respectively, which are consequences of the joint distribution in (120), and (133) is due to the fact that conditioning cannot increase entropy.

Next, we consider the distortion constraints in (121) as follows

$$D_m \geq \mathrm{mmse}(X_m | \{U_\ell\}_{\ell=1}^L) \quad (138)$$
$$\geq \mathrm{mmse}(X_m | \{U_\ell\}_{\ell=1}^L, X^{m-1}) \quad (139)$$
$$= \mathrm{mmse}(X_m | \{U_{\ell m}\}_{\ell=1}^L) \quad (140)$$

where we use the fact that conditioning reduces MMSE.

Hence, using (125) and (135), the rate constraints in Theorem 10 can be expressed as

$$\sum_{\ell \in \mathcal{A}} R_\ell \geq \sum_{m=1}^{M} I(X_m; \{U_{\ell m}\}_{\ell\in\mathcal{A}} | \{U_{\ell m}\}_{\ell\in\mathcal{A}^c})$$
$$\quad + \sum_{m=1}^{M} \sum_{\ell \in \mathcal{A}} I(U_{\ell m}; Y_{\ell m} | X_m, W_m) \quad (141)$$

and the distortion constraints in Theorem 10 are

$$\mathrm{mmse}(X_m | \{U_{\ell m}\}_{\ell=1}^L) \leq D_m \quad (142)$$

We note that the random variable tuples

$$\{(X_m, \{Y_{\ell m}, U_{\ell m}\}_{\ell=1}^L, W_m)\}_{m=1}^M \quad (143)$$

might be correlated over the index m. However, neither the expressions in the rate bounds given by (141) nor the distortion constraints in (142) depend on the entire joint distribution of $\{(X_m, \{Y_{\ell m}, U_{\ell m}\}_{\ell=1}^L, W_m)\}_{m=1}^M$. Instead, both the expressions in the rate bounds given by (141) and the distortion constraints in (142) depend only on the distribution of $(X_m, \{Y_{\ell m}, U_{\ell m}\}_{\ell=1}^L, W_m)$ for each m involved. Hence, without loss of generality, we can assume that

$$(X_m, \{Y_{\ell m}, U_{\ell m}\}_{\ell=1}^L, W_m) \quad (144)$$

and

$$\{(X_j, \{Y_{\ell j}, U_{\ell j}\}_{\ell=1}^L, W_j)\}_{j=1, j\neq m}^M \quad (145)$$

are independent for all $m = 1, \ldots, M$. Next, we note that the joint distribution of $(X_m, \{Y_{\ell m}, U_{\ell m}\}_{\ell=1}^L, W_m)$ can be factorized as follows

$$p(x_m, \{y_{\ell m}, u_{\ell m}\}_{\ell=1}^L, w_m)$$
$$= p(x_m)p(w_m) \prod_{\ell=1}^L p(y_{\ell m}|x_m)p(u_{\ell m}|y_{\ell m}, w_m) \quad (146)$$

whose proof is given in Appendix II-C. In view of (141)-(142) and (146), we obtain the following outer bound for the parallel model.

Theorem 11: We have $\mathcal{R}^p(\{D_m\}_{m=1}^M) \subseteq \mathcal{R}^{p-o}(\{D_m\}_{m=1}^M)$, where $\mathcal{R}^{p-o}(\{D_m\}_{m=1}^M)$ is given by the union of rate tuples $(R_1, \ldots, R_L)$ satisfying

$$\sum_{\ell\in\mathcal{A}} R_\ell \geq \sum_{m=1}^M I(X_m; \{U_{\ell m}\}_{\ell\in\mathcal{A}} | \{U_{\ell m}\}_{\ell\in\mathcal{A}^c})$$
$$\quad + \sum_{m=1}^M \sum_{\ell\in\mathcal{A}} I(U_{\ell m}; Y_{\ell m} | X_m, W_m) \quad (147)$$

for all $\mathcal{A} \subseteq \{1, \ldots, L\}$, *where the union is over all* $\{U_{\ell m}\}_{\forall \ell, \forall m}$ *satisfying*

$$p(x^M, \{y_\ell^M\}_{\ell=1}^L, \{u_\ell\}_{\ell=1}^L, w)$$
$$= \prod_{m=1}^M p(x_m)p(w_m) \prod_{\ell=1}^L p(y_{\ell m}|x_m)p(u_{\ell m}|w_m, y_{\ell m}) \quad (148)$$

and

$$\mathrm{mmse}(X_m | \{U_{\ell m}\}_{\ell=1}^L) \leq D_m, \quad m = 1, \ldots, M \quad (149)$$

B. Evaluation of the Outer Bound

Now, we evaluate the outer bound in Theorem 11 for the parallel Gaussian model, and show that it is attainable by the inner bound given in Theorem 2. To this end, we note that following the analysis in Section V, one can evaluate the outer bound in Theorem 11 yielding the following outer bound for the parallel Gaussian model.

Theorem 12: An outer bound for the rate-distortion region $\mathcal{R}^p(\{D_m\}_{m=1}^M)$ *of the parallel Gaussian model is given by* $\tilde{\mathcal{R}}^p(\{D_m\}_{m=1}^M)$ *which corresponds to the union of rate tuples* $(R_1, \ldots, R_L)$ *satisfying*

$$\sum_{\ell\in\mathcal{A}} R_\ell \geq \sum_{m=1}^M \frac{1}{2} \log^+ \frac{1}{D_m} \left(\frac{1}{\sigma_m^2} + \sum_{\ell\in\mathcal{A}^c} \frac{\sigma_{\ell m}^2 - D_{\ell m}}{\sigma_{\ell m}^4} \right)^{-1}$$
$$\quad + \sum_{m=1}^M \sum_{\ell\in\mathcal{A}} \frac{1}{2} \log \frac{\sigma_{\ell m}^2}{D_{\ell m}} \quad (150)$$

for all $\mathcal{A} \subseteq \{1, \ldots, L\}$, *where the union is over all* $\{D_{\ell m}\}_{\forall \ell, \forall m}$ *satisfying the following constraints*

$$\left(\frac{1}{\sigma_m^2} + \sum_{\ell=1}^L \frac{\sigma_{\ell m}^2 - D_{\ell m}}{\sigma_{\ell m}^4} \right)^{-1} \leq D_m, \quad m = 1, \ldots, M \quad (151)$$
$$0 \leq D_{\ell m} \leq \sigma_{\ell m}^2, \quad \ell = 1, \ldots, L, \quad m = 1, \ldots, M \quad (152)$$

Next, we show that there is no loss of generality to assume that the constraints in (151) are satisfied with equality. To prove this, we consider an alternative description of the outer bound in Theorem 12 by means of the tangent hyperplanes. In other words, we consider the following optimization problem

$$\min_{(R_1,\ldots,R_L)\in\tilde{\mathcal{R}}^P(\{D_m\}_{m=1}^M)} \sum_{\ell=1}^{L} \mu_\ell R_\ell \qquad (153)$$

where we assume $\mu_1 \geq \cdots \geq \mu_L \geq 0$. Using the fact that the outer bound is a contra-polymatroid [3], we can express the optimization problem in (153) as follows

$$\min_{(R_1,\ldots,R_L)\in\tilde{\mathcal{R}}^P(\{D_m\}_{m=1}^M)} \sum_{\ell=1}^{L} \mu_\ell R_\ell$$
$$= \min_{\{D_{\ell m}\}_{\forall \ell, \forall m}} \sum_{\ell=1}^{L} \frac{\mu_\ell}{2} \sum_{m=1}^{M} \log \frac{\sigma_{\ell m}^2}{D_{\ell m}} + \frac{\mu_L}{2} \sum_{m=1}^{M} \log \frac{\sigma_m^2}{D_m}$$
$$+ \sum_{\ell=1}^{L-1} \frac{\mu_\ell - \mu_{\ell+1}}{2} \sum_{m=1}^{M} \log^+$$
$$\frac{1}{D_m} \left(\frac{1}{\sigma_m^2} + \sum_{j=\ell+1}^{L} \frac{\sigma_{jm}^2 - D_{jm}}{\sigma_{jm}^4} \right)^{-1} \qquad (154)$$
$$= \min_{\{D_{\ell m}\}_{\forall \ell, \forall m}} \sum_{m=1}^{M} f_m(\{D_{\ell m}\}_{\ell=1}^L) \qquad (155)$$
$$= \sum_{m=1}^{M} \min_{\{D_{\ell m}\}_{\ell=1}^L} f_m(\{D_{\ell m}\}_{\ell=1}^L) \qquad (156)$$

where we define the function $f_m(\{D_{\ell m}\}_{\ell=1}^L)$ as follows

$$f_m(\{D_{\ell m}\}_{\ell=1}^L) = \sum_{\ell=1}^{L} \frac{\mu_\ell}{2} \log \frac{\sigma_{\ell m}^2}{D_{\ell m}} + \frac{\mu_L}{2} \log \frac{\sigma_m^2}{D_m}$$
$$+ \sum_{\ell=1}^{L-1} \frac{\mu_\ell - \mu_{\ell+1}}{2} \log^+ \frac{1}{D_m} \left(\frac{1}{\sigma_m^2} + \sum_{j=\ell+1}^{L} \frac{\sigma_{jm}^2 - D_{jm}}{\sigma_{jm}^4} \right)^{-1} \qquad (157)$$

and the feasible set of the minimizations in (154)-(156) are defined by the constraints in (151)-(152). Equation (156) follows from the fact that $f_m(\{D_{\ell m}\}_{\ell=1}^L)$ depends only on $\{D_{\ell m}\}_{\ell=1}^L$ but not on $\{D_{\ell j}\}_{\ell=1}^L$, $j \neq m$.

Next, we note that each minimization

$$\min_{\{D_{\ell m}\}_{\ell=1}^L} f_m(\{D_{\ell m}\}_{\ell=1}^L) \qquad (158)$$

is identical to the optimization problem we encounter for the scalar Gaussian model in Section III-A, and hence, the minimum is attained by those $\{D_{\ell m}\}_{\ell=1}^L$ that satisfy the constraint in (151) with equality. This implies that the outer bound in Theorem 12 is attainable; completing the proof of Theorem 6.

C. Proof of (146)

We first note that

$$p(x_m, \{y_{\ell m}, u_{\ell m}\}_{\ell=1}^L, w_m) = p(x_m)p(w_m)\left(\prod_{\ell=1}^{L} p(y_{\ell m}|x_m)\right)$$
$$p(\{u_{\ell m}\}_{\ell=1}^L|x_m, w_m, \{y_{\ell m}\}_{\ell=1}^L) \qquad (159)$$

where we use the fact that $(X_m, \{Y_{\ell m}\}_{\ell=1}^L)$ and $W_m = WX^{m-1}\{Y_{\ell,m+1}^M\}_{\ell=1}^L$ are independent, which is a consequence of the joint distribution in (120). Next, we consider the following term

$$p(\{u_{\ell m}\}_{\ell=1}^L|x_m, w_m, \{y_{\ell m}\}_{\ell=1}^L)$$
$$= \sum_{\forall\{y_\ell^{m-1}\}_{\ell=1}^L} p(\{u_{\ell m}\}_{\ell=1}^L, \{y_\ell^{m-1}\}_{\ell=1}^L|x_m, w_m, \{y_{\ell m}\}_{\ell=1}^L) \qquad (160)$$
$$= \sum_{\forall\{y_\ell^{m-1}\}_{\ell=1}^L} p(\{y_\ell^{m-1}\}_{\ell=1}^L|x_m, w_m, \{y_{\ell m}\}_{\ell=1}^L)$$
$$p(\{u_{\ell m}\}_{\ell=1}^L|x_m, w_m, \{y_\ell^m\}_{\ell=1}^L) \qquad (161)$$

where the first term in the summation is

$$p(\{y_\ell^{m-1}\}_{\ell=1}^L|x_m, w_m, \{y_{\ell m}\}_{\ell=1}^L)$$
$$= \prod_{\ell=1}^{L} p(y_\ell^{m-1}|x_m, w_m, \{y_{\ell m}\}_{\ell=1}^L, \{y_j^{m-1}\}_{j=1}^{\ell-1}) \qquad (162)$$
$$= \prod_{\ell=1}^{L} p(y_\ell^{m-1}|w_m) \qquad (163)$$
$$= \prod_{\ell=1}^{L} p(y_\ell^{m-1}|w_m, y_{\ell m}) \qquad (164)$$

where (163)-(164) come from the following Markov chain

$$Y_\ell^{m-1} \to W_m \to X_m, \{Y_{\ell m}\}_{\ell=1}^L, \{Y_j^{m-1}\}_{j=1}^{\ell-1} \qquad (165)$$

which is a consequence of the definition of W_m and the joint distribution in (120).

Next, we consider the second term in the summation given by (161) as follows

$$p(\{u_{\ell m}\}_{\ell=1}^L|x_m, w_m, \{y_j^m\}_{j=1}^L)$$
$$= \prod_{\ell=1}^{L} p(u_{\ell m}|x_m, w_m, \{y_j^m\}_{j=1}^L, \{u_{jm}\}_{j=1}^{\ell-1}) \qquad (166)$$
$$= \prod_{\ell=1}^{L} p(u_{\ell m}|w_m, y_\ell^m) \qquad (167)$$

where (167) comes from the following Markov chain

$$U_{\ell m} \to W_m, Y_\ell^m \to X_m, \{Y_j^m\}_{j=1,j\neq\ell}^L, \{U_{jm}\}_{j=1}^{\ell-1} \qquad (168)$$

which is also a consequence of the definition of W_m and the joint distribution in (120).

Using (164) and (167) in (161), we get

$$p(\{u_{\ell m}\}_{\ell=1}^L | x_m, w_m, \{y_{\ell m}\}_{\ell=1}^L)$$

$$= \sum_{\forall \{y_\ell^{m-1}\}_{\ell=1}^L} \prod_{\ell=1}^L p(y_\ell^{m-1} | w_m, y_{\ell m}) p(u_{\ell m} | w_m, y_\ell^m) \quad (169)$$

$$= \sum_{\forall \{y_\ell^{m-1}\}_{\ell=1}^L} \prod_{\ell=1}^L p(y_\ell^{m-1}, u_{\ell m} | w_m, y_{\ell m}) \quad (170)$$

$$= \prod_{\ell=1}^L p(u_{\ell m} | w_m, y_{\ell m}) \quad (171)$$

using which in (159), we get

$$p(x_m, \{y_{\ell m}, u_{\ell m}\}_{\ell=1}^L, w_m) = p(x_m) p(w_m)$$
$$\prod_{\ell=1}^L p(y_{\ell m} | x_m) p(u_{\ell m} | w_m, y_{\ell m})$$
$$(172)$$

which is the desired result in (146); completing the proof.

APPENDIX III
PROOF OF COROLLARY 2

From the analysis in Appendix II-B, when $\mu_1 \geq \mu_2 \geq 0$, we have

$$\min_{(R_1,R_2)\in\mathcal{R}^p(D_1,D_2)} \mu_1 R_1 + \mu_2 R_2$$

$$= \sum_{m=1}^2 \min_{(D_{1m},D_{2m})\in\mathcal{D}_m} f_m(D_{1m}, D_{2m}) \quad (173)$$

where the function $f_m(D_{1m}, D_{2m})$ is given by

$$f_m(D_{1m}, D_{2m}) = \sum_{\ell=1}^2 \frac{\mu_\ell}{2} \log \frac{\sigma_{\ell m}^2}{D_{\ell m}} + \frac{\mu_2}{2} \log \frac{\sigma_m^2}{D_m}$$
$$+ \frac{\mu_1-\mu_2}{2} \log \frac{1}{D_m} \left(\frac{1}{\sigma_m^2} + \frac{\sigma_{2m}^2 - D_{2m}}{\sigma_{2m}^4} \right)^{-1}$$
$$(174)$$

and the set $\mathcal{D}_m$ consists of (D_{1m}, D_{2m}) pairs satisfying

$$\frac{1}{\sigma_m^2} + \sum_{\ell=1}^2 \frac{\sigma_{\ell m}^2 - D_{\ell m}}{\sigma_{\ell m}^4} = \frac{1}{D_m} \quad (175)$$

$$0 \leq D_{\ell m} \leq \sigma_{\ell m}^2, \quad \ell = 1, 2 \quad (176)$$

Next, we note that

$$f_2(D_{12}, D_{22}) \geq \tilde{f}_2(D_{12}, D_{22}), \quad \forall (D_{12}, D_{22}) \in \mathcal{D}_2 \quad (177)$$

where the function $\tilde{f}_2(D_{12}, D_{22})$ is defined as

$$\tilde{f}_2(D_{12}, D_{22}) = \sum_{\ell=1}^2 \frac{\mu_\ell}{2} \log \frac{\sigma_{\ell 2}^2}{D_{\ell 2}} + \frac{\mu_2}{2} \log \frac{\sigma_2^2}{D_2} \quad (178)$$

This function can be minimized over $(D_{12}, D_{22}) \in \mathcal{D}_2$ to get

$$\min_{(D_{12},D_{22})\in\tilde{\mathcal{D}}_2} \tilde{f}_2(D_{12}, D_{22}) = \frac{\mu_2}{2} \log \frac{\sigma_2^2}{D_2}$$
$$+ \frac{\mu_2}{2} \log \frac{1}{\sigma_{22}^2} \left(\frac{1}{\sigma_2^2} + \frac{1}{\sigma_{22}^2} - \frac{1}{D_2} \right)^{-1} \quad (179)$$

Next, we note that by setting $(D_{12}^* = \sigma_{12}^2, D_{22}^* = \sigma_{22}^4(1/\sigma_2^2 + 1/\sigma_{22}^2 - 1/D_2)) \in \mathcal{D}_2$, we get

$$f_2(D_{12}^*, D_{22}^*) = \min_{(D_{12},D_{22})\in\tilde{\mathcal{D}}_2} \tilde{f}_2(D_{12}, D_{22}) \quad (180)$$

using which, and (179) in (173), we get

$$\min_{(R_1,R_2)\in\mathcal{R}^p(D_1,D_2)} \mu_1 R_1 + \mu_2 R_2 = \min_{(D_{11},D_{21})\in\mathcal{D}_1} f_1(D_{11}, D_{21})$$
$$+ \frac{\mu_2}{2} \log \frac{\sigma_2^2}{D_2} + \frac{\mu_2}{2} \log \frac{1}{\sigma_{22}^2} \left(\frac{1}{\sigma_2^2} + \frac{1}{\sigma_{22}^2} - \frac{1}{D_2} \right)^{-1} \quad (181)$$

which is the desired result in Corollary 2; completing the proof.

APPENDIX IV
PROOF OF COROLLARY 3

Using Theorem 3 for $L = 2$, our outer bound for the parallel Gaussian model can be expressed as follows.

$$\mathrm{T}^+ = \min_{(R_1,R_2)\in\mathcal{R}^o(D_1,D_2)} \mu_1 R_1 + \mu_2 R_2 \quad (182)$$

$$= \min_{(\mathbf{D}_1,\mathbf{D}_2,\mathbf{D})} \sum_{\ell=1}^2 \frac{\mu_\ell}{2} \log \frac{|\boldsymbol{\Sigma}_\ell|}{|\mathbf{D}_\ell|} + \frac{\mu_2}{2} \log \frac{|\mathbf{K}_X|}{|\mathbf{D}|}$$
$$+ \frac{\mu_1 - \mu_2}{2} \log^+ \frac{\left| \left(\mathbf{K}_X^{-1} + \boldsymbol{\Sigma}_2^{-1} - \boldsymbol{\Sigma}_2^{-1}\mathbf{D}_2\boldsymbol{\Sigma}_2^{-1} \right)^{-1} \right|}{|\mathbf{D}|}$$
$$(183)$$

where $(\mathbf{D}_1, \mathbf{D}_2, \mathbf{D})$ are subject to the following constraints

$$\left(\mathbf{K}_X^{-1} + \sum_{\ell=1}^2 \boldsymbol{\Sigma}_\ell^{-1} - \sum_{\ell=1}^2 \boldsymbol{\Sigma}_\ell^{-1}\mathbf{D}_\ell\boldsymbol{\Sigma}_\ell^{-1} \right)^{-1} \preceq \mathbf{D} \quad (184)$$

$$\mathbf{0} \preceq \mathbf{D}_\ell \preceq \boldsymbol{\Sigma}_\ell, \quad \ell = 1, 2 \quad (185)$$

$$\mathbf{D}_{mm} \leq D_m, \quad m = 1, 2 \quad (186)$$

where $\mathbf{D}_{mm}$ denotes the mth diagonal element of $\mathbf{D}$. By restricting $(\mathbf{D}_1, \mathbf{D}_2, \mathbf{D})$ to be diagonal, we have

$$\mathrm{T}^+ \leq \min_{\{D_{\ell m}\}_{\forall \ell, \forall m}} \sum_{m=1}^2 \frac{\mu_1}{2} \log \frac{\sigma_{1m}^2}{D_{1m}} + \frac{\mu_2}{2} \log \frac{\sigma_{2m}^2}{D_{2m}}$$
$$+ \frac{\mu_2}{2} \log \frac{\sigma_m^2}{D_m} + \frac{\mu_1 - \mu_2}{2} \log^+$$
$$\prod_{m=1}^2 \frac{1}{D_m} \left(\frac{1}{\sigma_m^2} + \frac{1}{\sigma_{2m}^2} - \frac{D_{2m}}{\sigma_{2m}^4} \right)^{-1} \quad (187)$$

where $\{D_{\ell m}\}_{\forall \ell, \forall m}$ are subject to the following constraints

$$\left(\frac{1}{\sigma_m^2} + \sum_{\ell=1}^2 \frac{1}{\sigma_{\ell m}^2} - \frac{D_{\ell m}}{\sigma_{\ell m}^4} \right)^{-1} \leq D_m, \quad m = 1, 2 \quad (188)$$

$$0 \leq D_{\ell m} \leq \sigma_{\ell m}^2, \quad \ell = 1, 2 \ m = 1, 2 \quad (189)$$

Next, we set

$$D_{12} = \sigma_{12}^2 \quad (190)$$

$$D_{22} = \frac{\mu_2}{\mu_1} \sigma_{22}^4 \left(\frac{1}{\sigma_2^2} + \frac{1}{\sigma_{22}^2} \right) \quad (191)$$

which are feasible, i.e., satisfy the constraints in (188)-(189), due to the assumptions in (38)-(39). Next, we note the following

$$\prod_{m=1}^{2} \frac{1}{D_m} \left(\frac{1}{\sigma_m^2} + \frac{1}{\sigma_{2m}^2} - \frac{D_{2m}}{\sigma_{2m}^4} \right)^{-1}$$
$$= \frac{1}{D_1} \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_{21}^2} - \frac{D_{21}}{\sigma_{21}^4} \right)^{-1} \frac{1}{D_2} \frac{\mu_1}{\mu_1 - \mu_2} \left(\frac{1}{\sigma_2^2} + \frac{1}{\sigma_{22}^2} \right)^{-1} \tag{192}$$
$$\geq \frac{1}{D_1} \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_{21}^2} \right)^{-1} \frac{1}{D_2} \frac{\mu_1}{\mu_1 - \mu_2} \left(\frac{1}{\sigma_2^2} + \frac{1}{\sigma_{22}^2} \right)^{-1} \tag{193}$$
$$> 1 \tag{194}$$

where, in (192), we use (190)-(191), (193) follows from the fact that $D_{21} \geq 0$, and (194) is due to the assumption in (40). Hence, using (190)-(191) and (193) in (187), we get

$$T^+ \leq \min_{(D_{11}, D_{21}) \in \hat{\mathcal{D}}_1} f_1(D_{11}, D_{21}) + \frac{\mu_2}{2} \log \frac{\sigma_2^2}{D_2}$$
$$+ \frac{\mu_2}{2} \log \frac{\mu_1}{\mu_2} \frac{1}{\sigma_{22}^2} \left(\frac{1}{\sigma_2^2} + \frac{1}{\sigma_{22}^2} \right)^{-1}$$
$$+ \frac{\mu_1 - \mu_2}{2} \log \frac{\mu_1}{\mu_1 - \mu_2} \frac{1}{D_2} \left(\frac{1}{\sigma_2^2} + \frac{1}{\sigma_{22}^2} \right)^{-1} \tag{195}$$

where the set $\hat{\mathcal{D}}_1$ is defined as the union of (D_{11}, D_{21}) pairs satisfying

$$\left(\frac{1}{\sigma_1^2} + \sum_{\ell=1}^{2} \frac{1}{\sigma_{\ell1}^2} - \frac{D_{\ell1}}{\sigma_{\ell1}^4} \right)^{-1} \leq D_1 \tag{196}$$
$$0 \leq D_{\ell1} \leq \sigma_{\ell1}^2 \tag{197}$$

We note that $\mathcal{D}_1 \subseteq \hat{\mathcal{D}}_1$, where $\mathcal{D}_1$ is the region defined in Corollary 3. Hence, using this fact in (195), we get

$$T^+ \leq \min_{(D_{11}, D_{21}) \in \mathcal{D}_1} f_1(D_{11}, D_{21}) + \frac{\mu_2}{2} \log \frac{\sigma_2^2}{D_2}$$
$$+ \frac{\mu_2}{2} \log \frac{\mu_1}{\mu_2} \frac{1}{\sigma_{22}^2} \left(\frac{1}{\sigma_2^2} + \frac{1}{\sigma_{22}^2} \right)^{-1}$$
$$+ \frac{\mu_1 - \mu_2}{2} \log \frac{\mu_1}{\mu_1 - \mu_2} \frac{1}{D_2} \left(\frac{1}{\sigma_2^2} + \frac{1}{\sigma_{22}^2} \right)^{-1} \tag{198}$$

which is the desired result in Corollary 3; completing the proof.

APPENDIX V
PROOF OF LEMMA 4

We first note the following Markov chain

$$(U_j, \mathbf{Y}_j) \rightarrow (W, \mathbf{X}) \rightarrow (U_{[1:L] \setminus j}, \mathbf{Y}_{[1:L] \setminus j}) \tag{199}$$

whose proof is given in Appendix V-A. Next, we note that

$$E \left[\mathbf{S}_{\mathcal{A}^c} | \mathbf{X}, \{U_\ell\}_{\ell \in \mathcal{A}^c}, W \right] = \sum_{\ell \in \mathcal{A}^c} \mathbf{A}_\ell E \left[\mathbf{Y}_\ell | \mathbf{X}, \{U_\ell\}_{\ell \in \mathcal{A}^c}, W \right] \tag{200}$$
$$= \sum_{\ell \in \mathcal{A}^c} \mathbf{A}_\ell E \left[\mathbf{Y}_\ell | \mathbf{X}, U_\ell, W \right] \tag{201}$$

where (201) follows from the Markov chain in (199). Now, we consider $\text{mmse}(\mathbf{S}_{\mathcal{A}^c} | \mathbf{X}, \{U_\ell\}_{\ell \in \mathcal{A}^c}, W)$ as follows

$$\text{mmse}(\mathbf{S}_{\mathcal{A}^c} | \mathbf{X}, \{U_\ell\}_{\ell \in \mathcal{A}^c}, W)$$
$$= E \left[\left(\mathbf{S}_{\mathcal{A}^c} - E[\mathbf{S}_{\mathcal{A}^c} | \mathbf{X}, \{U_\ell\}_{\ell \in \mathcal{A}^c}, W] \right) \right.$$
$$\left. \times \left(\mathbf{S}_{\mathcal{A}^c} - E[\mathbf{S}_{\mathcal{A}^c} | \mathbf{X}, \{U_\ell\}_{\ell \in \mathcal{A}^c}, W] \right)^\top \right] \tag{202}$$
$$= E \left[\left(\sum_{\ell \in \mathcal{A}^c} \mathbf{A}_\ell (\mathbf{Y}_\ell - E[\mathbf{Y}_\ell | \mathbf{X}, U_\ell, W]) \right) \right.$$
$$\left. \times \left(\sum_{\ell \in \mathcal{A}^c} \mathbf{A}_\ell (\mathbf{Y}_\ell - E[\mathbf{Y}_\ell | \mathbf{X}, U_\ell, W]) \right)^\top \right] \tag{203}$$
$$= \sum_{\ell \in \mathcal{A}^c} \mathbf{A}_\ell \text{mmse}(\mathbf{Y}_\ell | \mathbf{X}, U_\ell, W) \mathbf{A}_\ell^\top$$
$$+ \sum_{i \in \mathcal{A}^c} \sum_{\substack{j \in \mathcal{A}^c \\ j \neq i}} \mathbf{A}_i E \left[(\mathbf{Y}_i - E[\mathbf{Y}_i | \mathbf{X}, U_i, W]) \right.$$
$$\left. \times (\mathbf{Y}_j - E[\mathbf{Y}_j | \mathbf{X}, U_j, W])^\top \right] \mathbf{A}_j^\top \tag{204}$$

where (203) is due to (201). Next, we consider the cross-terms in (204) as follows

$$E \left[(\mathbf{Y}_i - E[\mathbf{Y}_i | \mathbf{X}, U_i, W])(\mathbf{Y}_j - E[\mathbf{Y}_j | \mathbf{X}, U_j, W])^\top \right]$$
$$= E \left[E \left[(\mathbf{Y}_i - E[\mathbf{Y}_i | \mathbf{X}, U_i, W]) \right. \right.$$
$$\left. \left. \times (\mathbf{Y}_j - E[\mathbf{Y}_j | \mathbf{X}, U_j, W])^\top | \mathbf{X}, W \right] \right] \tag{205}$$
$$= E \left[E \left[(\mathbf{Y}_i - E[\mathbf{Y}_i | \mathbf{X}, U_i, W]) | \mathbf{X}, W \right] \right.$$
$$\left. \times E \left[(\mathbf{Y}_j - E[\mathbf{Y}_j | \mathbf{X}, U_j, W])^\top | \mathbf{X}, W \right] \right] \tag{206}$$
$$= \mathbf{0} \tag{207}$$

where (206) is due to the Markov chain in (199). Using (207) in (204), we get

$$\text{mmse}(\mathbf{S}_{\mathcal{A}^c} | \mathbf{X}, \{U_\ell\}_{\ell \in \mathcal{A}^c}, W)$$
$$= \sum_{\ell \in \mathcal{A}^c} \mathbf{A}_\ell \text{mmse}(\mathbf{Y}_\ell | \mathbf{X}, U_\ell, W) \mathbf{A}_\ell^\top \tag{208}$$

which completes the proof of Lemma 4.

A. Proof of (199)

We first consider the joint distribution in (54) as follows

$$p(\mathbf{x}, \{\mathbf{y}_\ell, u_\ell\}_{\ell=1}^L, w) = p(\mathbf{x}) p(w) \left(\prod_{\substack{\ell=1 \\ \ell \neq j}}^{L} p(\mathbf{y}_\ell | \mathbf{x}) p(u_\ell | \mathbf{y}_\ell, w) \right)$$
$$p(\mathbf{y}_j | \mathbf{x}) p(u_j | \mathbf{y}_j, w) \tag{209}$$

which implies

$$U_\ell \rightarrow (\mathbf{Y}_\ell, W) \rightarrow \mathbf{X}, \quad \ell = 1, \dots, L \tag{210}$$

Next, we note that

$$p(\mathbf{y}_j|\mathbf{x})p(u_j|\mathbf{y}_j, w) = p(\mathbf{y}_j|\mathbf{x})p(u_j|\mathbf{y}_j, w, \mathbf{x}) \quad (211)$$

$$= p(\mathbf{y}_j|\mathbf{x})\frac{p(u_j, \mathbf{y}_j, w, \mathbf{x})}{p(\mathbf{y}_j, \mathbf{x})p(w)} \quad (212)$$

$$= \frac{p(u_j, \mathbf{y}_j, w, \mathbf{x})}{p(\mathbf{x})p(w)} \quad (213)$$

$$= \frac{p(u_j, \mathbf{y}_j, w, \mathbf{x})}{p(w, \mathbf{x})} \quad (214)$$

$$= p(u_j, \mathbf{y}_j|w, \mathbf{x}) \quad (215)$$

where (211) comes from the Markov chain in (210), (212) and (214) follow from the fact that $(\mathbf{X}, \mathbf{Y}_j)$ and W are independent which is a consequence of the factorization in (209). Using (215) in (209), we get

$$p(\mathbf{x}, \{\mathbf{y}_\ell, u_\ell\}_{\ell=1}^L, w) = p(\mathbf{x})p(w)\left(\prod_{\substack{\ell=1 \\ \ell \neq j}}^L p(\mathbf{y}_\ell|\mathbf{x})p(u_\ell|\mathbf{y}_\ell, w)\right)$$
$$p(u_j, \mathbf{y}_j|\mathbf{x}, w) \quad (216)$$

which implies the Markov chain in (199); completing the proof.

APPENDIX VI
PROOF OF THEOREM 8

The singular value decomposition of the matrices $\{\mathbf{H}_\ell\}_{\ell=1}^L$ are given by

$$\mathbf{H}_\ell = \mathbf{U}_\ell \mathbf{\Lambda}_\ell \mathbf{V}_\ell^\top, \quad \ell = 1, \ldots, L \quad (217)$$

where $\{\mathbf{U}_\ell\}_{\ell=1}^L$ and $\{\mathbf{V}_\ell\}_{\ell=1}^L$ are orthogonal matrices. Next, we show that without loss of generality, we can assume that $\{\mathbf{H}_\ell\}_{\ell=1}^L$ are square matrices. To this end, we define the following observations

$$\bar{\mathbf{Y}}_\ell = \mathbf{U}_\ell^\top \mathbf{Y}_\ell \quad (218)$$

$$= \mathbf{\Lambda}_\ell \mathbf{V}_\ell \mathbf{X} + \bar{\mathbf{N}}_\ell \quad (219)$$

where $\bar{\mathbf{N}}_\ell$ is again a zero-mean Gaussian random vector with an identity covariance matrix. We note that the rate-distortion region for the observations $\{\bar{\mathbf{Y}}_\ell\}_{\ell=1}^L$ is identical to the rate-distortion region for the observations $\{\mathbf{Y}_\ell\}_{\ell=1}^L$, since we obtain the observations $\{\bar{\mathbf{Y}}_\ell\}_{\ell=1}^L$ from $\{\mathbf{Y}_\ell\}_{\ell=1}^L$ by an invertible transform. Now, we show that there is no loss of generality to assume that the matrices $\{\mathbf{H}_\ell\}_{\ell=1}^L$ are square matrices. Assume that $\mathbf{H}_\ell$ is an $r_\ell \times M$ matrix. Hence, $\mathbf{\Lambda}_\ell$ is also an $r_\ell \times M$ matrix. First, consider $r_\ell > M$. In this case, $r_\ell - M$ entries of $\bar{\mathbf{Y}}_\ell$ consists of only noise. Since the noise $\bar{\mathbf{N}}_\ell$ is i.i.d., we can drop these $r_\ell - M$ entries of the observation $\bar{\mathbf{Y}}_\ell$ without altering the rate-distortion region. Hence, when $r_\ell > M$, there is an equivalent model with the same rate-distortion region and $r_\ell = M$. Next, assume $r_\ell < M$. In this case, we can add $M - r_\ell$ i.i.d. noise entries to the observation $\bar{\mathbf{Y}}_\ell$ without altering the rate-distortion region. Hence, when $r_\ell < M$, there is also an equivalent model with the same rate-distortion region and $r_\ell = M$. Consequently, from now on, we assume that $r_1 = \ldots = r_L = M$.

Next, we define

$$\mathbf{H}_{\ell,\alpha} = \mathbf{U}_\ell(\mathbf{\Lambda}_\ell + \alpha\mathbf{I})\mathbf{V}_\ell^\top, \quad \ell = 1, \ldots, L \quad (220)$$

where $\alpha > 0$. We note that $\{\mathbf{H}_{\ell,\alpha}\}_{\ell=1}^L$ are invertible, i.e., $\{\mathbf{H}_{\ell,\alpha}^{-1}\}_{\ell=1}^L$ exist, and in particular, we have

$$\mathbf{H}_{\ell,\alpha}^{-1} = \mathbf{V}_\ell(\mathbf{\Lambda}_\ell + \alpha\mathbf{I})^{-1}\mathbf{U}_\ell^\top, \quad \ell = 1, \ldots, L \quad (221)$$

Using $\{\mathbf{H}_{\ell,\alpha}\}_{\ell=1}^L$, we define an *enhanced model* as follows

$$\mathbf{Y}_{\ell,\alpha} = \mathbf{H}_{\ell,\alpha}\mathbf{X} + \mathbf{N}_\ell, \quad \ell = 1, \ldots, L \quad (222)$$

Using these enhanced observations in (222), we can rewrite the original observations in (98) as follows

$$\mathbf{Y}_\ell = \mathbf{H}_\ell\mathbf{H}_{\ell,\alpha}^{-1}\mathbf{Y}_{\ell,\alpha} + \tilde{\mathbf{N}}_\ell, \quad \ell = 1, \ldots, L \quad (223)$$

where $\tilde{\mathbf{N}}_\ell$ is a zero-mean Gaussian random vector, and independent of $\{\mathbf{Y}_{\ell,\alpha}\}_{\ell=1}^L$ and $\{\tilde{\mathbf{N}}_j\}_{j=1, j\neq\ell}^L$. The decomposition in (223) is possible, since we have

$$(\mathbf{H}_\ell\mathbf{H}_{\ell,\alpha}^{-1})(\mathbf{H}_\ell\mathbf{H}_{\ell,\alpha}^{-1})^\top = \mathbf{U}_\ell\mathbf{\Lambda}_\ell^2(\mathbf{\Lambda}_\ell + \alpha\mathbf{I})^{-2}\mathbf{U}_\ell^\top \preceq \mathbf{I} \quad (224)$$

Moreover, due to the decomposition in (223), we can assume that the following holds

$$p(\mathbf{x}, \{\mathbf{y}_{\ell,\alpha}, \mathbf{y}_\ell\}_{\ell=1}^L) = p(\mathbf{x})\prod_{\ell=1}^L p(\mathbf{y}_{\ell,\alpha}|\mathbf{x})p(\mathbf{y}_\ell|\mathbf{y}_{\ell,\alpha}) \quad (225)$$

which implies that the original observations $\{\mathbf{Y}_\ell\}_{\ell=1}^L$ are degraded versions the enhanced observations $\{\mathbf{Y}_{\ell,\alpha}\}_{\ell=1}^L$. Consequently, we have

$$\mathcal{R}(\mathbf{D}) \subseteq \mathcal{R}_\alpha(\mathbf{D}) \quad (226)$$

where $\mathcal{R}_\alpha(\mathbf{D})$ denotes the rate-distortion region for the enhanced model defined by (222). Next, we note that the enhanced model defined by (222) is equivalent to the following one

$$\bar{\mathbf{Y}}_{\ell,\alpha} = \mathbf{H}_{\ell,\alpha}^{-1}\mathbf{Y}_{\ell,\alpha} = \mathbf{X} + \bar{\mathbf{N}}_{\ell,\alpha}, \quad \ell = 1, \ldots, L \quad (227)$$

where the covariance matrix of $\bar{\mathbf{N}}_{\ell,\alpha}$ is given by

$$\mathbf{\Sigma}_{\ell,\alpha} = \left(\mathbf{H}_{\ell,\alpha}^\top\mathbf{H}_{\ell,\alpha}\right)^{-1}, \quad \ell = 1, \ldots, L \quad (228)$$

Using Theorem 1, we can obtain an outer bound for the rate-distortion region of the model defined by (227), which is equivalent to the enhanced model given by (222). In particular, we have $\mathcal{R}_\alpha(\mathbf{D}) \subseteq \mathcal{R}_\alpha^o(\mathbf{D})$, where $\mathcal{R}_\alpha^o(\mathbf{D})$ is given by the union of rate tuples $(R_1, \ldots, R_L)$ satisfying

$$\sum_{\ell \in \mathcal{A}} R_\ell$$
$$\geq \frac{1}{2}\log^+ \frac{\left|\left(\mathbf{K}_X^{-1} + \sum_{\ell \in \mathcal{A}^c}\mathbf{\Sigma}_{\ell,\alpha}^{-1} - \sum_{\ell \in \mathcal{A}^c}\mathbf{\Sigma}_{\ell,\alpha}^{-1}\tilde{\mathbf{D}}_\ell\mathbf{\Sigma}_{\ell,\alpha}^{-1}\right)^{-1}\right|}{|\mathbf{D}|}$$
$$+ \sum_{\ell \in \mathcal{A}}\frac{1}{2}\log\frac{|\mathbf{\Sigma}_{\ell,\alpha}|}{|\tilde{\mathbf{D}}_\ell|} \quad (229)$$

for all $\mathcal{A} \subseteq \{1, \ldots, L\}$, where the union is over all $\{\tilde{\mathbf{D}}_\ell\}_{\ell=1}^L$ satisfying

$$\mathbf{D} \succeq \left(\mathbf{K}_X^{-1} + \sum_{\ell=1}^L \boldsymbol{\Sigma}_{\ell,\alpha}^{-1} - \sum_{\ell=1}^L \boldsymbol{\Sigma}_{\ell,\alpha}^{-1} \tilde{\mathbf{D}}_\ell \boldsymbol{\Sigma}_{\ell,\alpha}^{-1} \right)^{-1} \quad (230)$$

$$\boldsymbol{\Sigma}_{\ell,\alpha} \succeq \tilde{\mathbf{D}}_\ell \succeq \mathbf{0}, \quad \ell = 1, \ldots, L \quad (231)$$

Next, we set $\mathbf{D}_\ell = \mathbf{H}_{\ell,\alpha} \tilde{\mathbf{D}}_\ell \mathbf{H}_{\ell,\alpha}^\top$, $\ell = 1, \ldots, L$, using which in (229)-(231), we can express the outer bound $\mathcal{R}_\alpha^o(\mathbf{D})$ as the union of rate tuples $(R_1, \ldots, R_L)$ satisfying

$$\sum_{\ell \in \mathcal{A}} R_\ell \geq \frac{1}{2} \log^+ \frac{\left| \left(\mathbf{K}_X^{-1} + \sum_{\ell \in \mathcal{A}^c} \mathbf{H}_{\ell,\alpha}^\top (\mathbf{I} - \mathbf{D}_\ell) \mathbf{H}_{\ell,\alpha} \right)^{-1} \right|}{|\mathbf{D}|}$$

$$+ \sum_{\ell \in \mathcal{A}} \frac{1}{2} \log \frac{1}{|\mathbf{D}_\ell|} \quad (232)$$

for all $\mathcal{A} \subseteq \{1, \ldots, L\}$, where the union is over all $\{\mathbf{D}_\ell\}_{\ell=1}^L$ satisfying

$$\left(\mathbf{K}_X^{-1} + \sum_{\ell=1}^L \mathbf{H}_{\ell,\alpha}^\top (\mathbf{I} - \mathbf{D}_\ell) \mathbf{H}_{\ell,\alpha} \right)^{-1} \preceq \mathbf{D} \quad (233)$$

$$\mathbf{0} \preceq \mathbf{D}_\ell \preceq \mathbf{I}, \quad \ell = 1, \ldots, L \quad (234)$$

In view of (226), we have the following

$$\mathcal{R}(\mathbf{D}) \subseteq \mathcal{R}_\alpha^o(\mathbf{D}), \quad \forall \alpha > 0 \quad (235)$$

which implies that

$$\mathcal{R}(\mathbf{D}) \subseteq \lim_{\alpha \to 0} \mathcal{R}_\alpha^o(\mathbf{D}) \quad (236)$$

Hence, to obtain an outer bound for the rate-distortion region of the general model defined by (98), it is sufficient to obtain the limiting region $\lim_{\alpha \to 0} \mathcal{R}_\alpha^o(\mathbf{D})$. To this end, we introduce the following lemma.

Lemma 5: For all $\mathcal{A}^c \subseteq \{1, \ldots, L\}$, we have

$$\lim_{\alpha \to 0} \left(\mathbf{K}_X^{-1} + \sum_{\ell \in \mathcal{A}^c} \mathbf{H}_{\ell,\alpha}^\top (\mathbf{I} - \mathbf{D}_\ell) \mathbf{H}_{\ell,\alpha} \right)^{-1}$$

$$= \left(\mathbf{K}_X^{-1} + \sum_{\ell \in \mathcal{A}^c} \mathbf{H}_\ell^\top (\mathbf{I} - \mathbf{D}_\ell) \mathbf{H}_\ell \right)^{-1} \quad (237)$$

The proof of Lemma 5 is given in Appendix VI-A. Using this lemma in (232)-(234), we obtain the region $\lim_{\alpha \to 0} \mathcal{R}_\alpha(\mathbf{D})$ as the union of rate tuples $(R_1, \ldots, R_L)$ satisfying

$$\sum_{\ell \in \mathcal{A}} R_\ell \geq \frac{1}{2} \log^+ \frac{\left| \left(\mathbf{K}_X^{-1} + \sum_{\ell \in \mathcal{A}^c} \mathbf{H}_\ell^\top (\mathbf{I} - \mathbf{D}_\ell) \mathbf{H}_\ell \right)^{-1} \right|}{|\mathbf{D}|}$$

$$+ \sum_{\ell \in \mathcal{A}} \frac{1}{2} \log \frac{1}{|\mathbf{D}_\ell|} \quad (238)$$

for all $\mathcal{A} \subseteq \{1, \ldots, L\}$, where the union is over all $\{\mathbf{D}_\ell\}_{\ell=1}^L$ satisfying

$$\left(\mathbf{K}_X^{-1} + \sum_{\ell=1}^L \mathbf{H}_\ell^\top (\mathbf{I} - \mathbf{D}_\ell) \mathbf{H}_\ell \right)^{-1} \preceq \mathbf{D} \quad (239)$$

$$\mathbf{0} \preceq \mathbf{D}_\ell \preceq \mathbf{I}, \quad \ell = 1, \ldots, L \quad (240)$$

which is the desired result in Theorem 8; completing the proof.

A. Proof of Lemma 5

In the proof of Lemma 5, we use the following fact.
Lemma 6: Let $\mathbf{A} \succ \mathbf{0}$. Then we have

$$\lim_{\mathbf{B} \to \mathbf{0}} (\mathbf{A} + \mathbf{B})^{-1} = \mathbf{A}^{-1} \quad (241)$$

Next, we note that

$$\sum_{\ell \in \mathcal{A}^c} \mathbf{H}_{\ell,\alpha}^\top (\mathbf{I} - \mathbf{D}_\ell) \mathbf{H}_{\ell,\alpha} = \sum_{\ell \in \mathcal{A}^c} \mathbf{H}_\ell^\top (\mathbf{I} - \mathbf{D}_\ell) \mathbf{H}_\ell + \mathbf{M}(\alpha)$$

$$(242)$$

where $\lim_{\alpha \to 0} \mathbf{M}(\alpha) = \mathbf{0}$. We define

$$\mathbf{M}_{\mathcal{A}^c} = \mathbf{K}_X^{-1} + \sum_{\ell \in \mathcal{A}^c} \mathbf{H}_\ell^\top (\mathbf{I} - \mathbf{D}_\ell) \mathbf{H}_\ell \quad (243)$$

Using (242)-(243) and Lemma 6, we have

$$\left(\mathbf{K}_X^{-1} + \sum_{\ell \in \mathcal{A}^c} \mathbf{H}_{\ell,\alpha}^\top (\mathbf{I} - \mathbf{D}_\ell) \mathbf{H}_{\ell,\alpha} \right)^{-1}$$

$$= \left(\mathbf{M}_{\mathcal{A}^c} + \mathbf{M}(\alpha) \right)^{-1} \quad (244)$$

$$= \mathbf{M}_{\mathcal{A}^c}^{-1} \quad (245)$$

$$= \left(\mathbf{K}_X^{-1} + \sum_{\ell \in \mathcal{A}^c} \mathbf{H}_\ell^\top (\mathbf{I} - \mathbf{D}_\ell) \mathbf{H}_\ell \right)^{-1} \quad (246)$$

which is the desired end result in Lemma 5; completing the proof.

APPENDIX VII
PROOFS OF THEOREM 2 AND THEOREM 9

We obtain the inner bound for the rate-distortion region $\mathcal{R}(\mathbf{D})$ by evaluating the Berger-Tung achievable scheme with jointly Gaussian auxiliary random vectors. For that purpose, we consider the most general form of the vector Gaussian CEO model defined by the observations in (98). In other words, we first obtain an inner bound for the most general form given by (98), i.e., we prove Theorem 9, and next, show that Theorem 2 follows from Theorem 9. Let $\mathcal{R}^{\mathrm{BT}}(\mathbf{D})$ denote the Berger-Tung inner bound. $\mathcal{R}^{\mathrm{BT}}(\mathbf{D})$ is given by the union of rate tuples $(R_1, \ldots, R_L)$ satisfying [5]

$$\sum_{\ell \in \mathcal{A}} R_\ell \geq I(\mathbf{X}; \{U_\ell\}_{\ell \in \mathcal{A}} | \{U_\ell\}_{\ell \in \mathcal{A}^c}) + \sum_{\ell \in \mathcal{A}} I(\mathbf{Y}_\ell; U_\ell | \mathbf{X})$$

$$(247)$$

for all $\mathcal{A} \subseteq \{1, \ldots, L\}$, where the union is over all $(U_1, \ldots, U_L)$ satisfying the Markov chain

$$U_j \to \mathbf{Y}_j \to \mathbf{X} \to \mathbf{Y}_k \to U_k, \quad j \neq k \quad (248)$$

and the distortion constraint

$$\mathrm{mmse}(\mathbf{X} | U_1, \ldots, U_L) \preceq \mathbf{D} \quad (249)$$

We select the auxiliary random variables $\{U_\ell\}_{\ell=1}^L$ as follows

$$U_\ell = \mathbf{Y}_\ell + \bar{\mathbf{N}}_\ell, \quad \ell = 1, \ldots, L \quad (250)$$

where $\{\bar{\mathbf{N}}_\ell\}_{\ell=1}^L$ are zero-mean independent Gaussian random vectors with covariance matrices $\{\bar{\boldsymbol{\Sigma}}_\ell\}_{\ell=1}^L$, and are independent of $\{\mathbf{Y}_\ell\}_{\ell=1}^L, \mathbf{X}$. We assume that the covariance matrices

$\{\bar{\boldsymbol{\Sigma}}_\ell\}_{\ell=1}^L$ are strictly positive definite, i.e., we have $\bar{\boldsymbol{\Sigma}}_\ell \succ \mathbf{0}$, $\forall \ell \in \{1, \ldots, L\}$. This assumption arises from the fact that if one of these matrices is singular, for example, if $\bar{\boldsymbol{\Sigma}}_\ell$ is singular, then, as we will show soon, the corresponding MMSE matrix $\mathrm{mmse}(\mathbf{Y}_\ell|\mathbf{X}, U_\ell)$ will be singular as well, and consequently, $I(U_\ell; \mathbf{Y}_\ell|\mathbf{X}) \to \infty$. When the auxiliary random variables $\{U_\ell\}_{\ell=1}^L$ are selected to be Gaussian as in (250), the rate bound in (247) becomes

$$\sum_{\ell \in \mathcal{A}} R_\ell \geq \frac{1}{2} \log \frac{|\mathrm{mmse}(\mathbf{X}|\{U_\ell\}_{\ell \in \mathcal{A}^c})|}{|\mathrm{mmse}(\mathbf{X}|\{U_\ell\}_{\ell=1}^L)|}$$
$$+ \sum_{\ell \in \mathcal{A}} \frac{1}{2} \log \frac{1}{|\mathrm{mmse}(\mathbf{Y}_\ell|\mathbf{X}, U_\ell)|} \quad (251)$$

where, as it will become clear soon, all MMSE matrices are strictly positive definite; implying that the rate bounds in (251) are finite.

Next, we evaluate the MMSE terms in (251). Using the definition of auxiliary random variables in (250), we have (see (108) and (111) in Appendix I-A)

$$\mathbf{D}_\ell \triangleq \mathrm{mmse}(\mathbf{Y}_\ell|\mathbf{X}, U_\ell) \quad (252)$$
$$= \left(\mathbf{I} + \bar{\boldsymbol{\Sigma}}_\ell^{-1}\right)^{-1}, \quad \ell = 1, \ldots, L \quad (253)$$

where $\mathbf{D}_\ell$ satisfies the following orders

$$\mathbf{0} \prec \mathbf{D}_\ell \preceq \mathbf{I} \quad (254)$$

where the upper bound on $\mathbf{D}_\ell$ follows from the following fact

$$\mathrm{mmse}(\mathbf{Y}_\ell|\mathbf{X}, U_\ell) = \mathrm{mmse}(\mathbf{N}_\ell|\mathbf{N}_\ell + \bar{\mathbf{N}}_\ell) \preceq \mathbf{I} \quad (255)$$

Using (253), we have

$$\bar{\boldsymbol{\Sigma}}_\ell = \left(\mathbf{D}_\ell^{-1} - \mathbf{I}\right)^{-1}, \quad \ell = 1, \ldots, L \quad (256)$$

Next, we evaluate the MMSE matrices $\mathrm{mmse}(\mathbf{X}|\{U_\ell\}_{\ell \in \mathcal{A}^c})$ as follows (see (108) and (111) in Appendix I-A)

$$\mathrm{mmse}(\mathbf{X}|\{U_\ell\}_{\ell \in \mathcal{A}^c}) = \left(\mathbf{K}_X^{-1} + \sum_{\ell \in \mathcal{A}^c} \mathbf{H}_\ell^\top (\mathbf{I} + \bar{\boldsymbol{\Sigma}}_\ell)^{-1} \mathbf{H}_\ell\right)^{-1} \quad (257)$$
$$= \left(\mathbf{K}_X^{-1} + \sum_{\ell \in \mathcal{A}^c} \mathbf{H}_\ell^\top (\mathbf{I} - \mathbf{D}_\ell) \mathbf{H}_\ell\right)^{-1} \quad (258)$$

where we used the following identity

$$\left(\mathbf{I} + \bar{\boldsymbol{\Sigma}}_\ell\right)^{-1} = \mathbf{I} - \mathbf{D}_\ell \quad (259)$$

which can be shown by using (256). Hence, using (253) and (258) in (251), we obtain the inner bound as the union of rate tuples $(R_1, \ldots, R_L)$ satisfying

$$\sum_{\ell \in \mathcal{A}} R_\ell \geq \frac{1}{2} \log \frac{\left|\left(\mathbf{K}_X^{-1} + \sum_{\ell \in \mathcal{A}^c} \mathbf{H}_\ell^\top (\mathbf{I} - \mathbf{D}_\ell) \mathbf{H}_\ell\right)^{-1}\right|}{\left|\left(\mathbf{K}_X^{-1} + \sum_{\ell=1}^L \mathbf{H}_\ell^\top (\mathbf{I} - \mathbf{D}_\ell) \mathbf{H}_\ell\right)^{-1}\right|}$$
$$+ \sum_{\ell \in \mathcal{A}} \frac{1}{2} \log \frac{1}{|\mathbf{D}_\ell|} \quad (260)$$

for all $\mathcal{A} \subseteq \{1, \ldots, L\}$, where the union is over all positive semi-definite matrices $\{\mathbf{D}_\ell\}_{\ell=1}^L$ satisfying

$$\left(\mathbf{K}_X^{-1} + \sum_{\ell=1}^L \mathbf{H}_\ell^\top (\mathbf{I} - \mathbf{D}_\ell) \mathbf{H}_\ell\right)^{-1} \preceq \mathbf{D} \quad (261)$$
$$\mathbf{0} \preceq \mathbf{D}_\ell \preceq \mathbf{I}, \quad \ell = 1, \ldots, L \quad (262)$$

where the first constraint in (261) is obtained by using (258) in (249), and the second constraint in (262) comes from (254). Hence, in view of (260)-(262), we obtain the inner bound given in Theorem 9; completing the proof.

Next, we show that Theorem 2 follows from Theorem 9. We note that the observations in (1) are equivalent to the general form of the observations in (98), when one sets $\mathbf{H}_\ell = \boldsymbol{\Sigma}_\ell^{-1/2}$, $\ell = 1, \ldots, L$. Using this observation in (260)-(262) in conjunction with the definition $\mathbf{D}_\ell = \boldsymbol{\Sigma}_\ell^{-1/2} \tilde{\mathbf{D}}_\ell \boldsymbol{\Sigma}_\ell^{-1/2}$, $\ell = 1, \ldots, L$, one can get the inner bound in Theorem 2; completing the proof.

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, "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.

[4] V. Prabhakaran, D. Tse, and K. Ramachandran, "Rate region of the quadratic Gaussian CEO problem," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun./Jul. 2004, p. 119.

[5] S.-Y. Tung, "Multiterminal source coding," Ph.D. dissertation, School Electr. Comput. Eng., Cornell Univ., Ithaca, NY, USA, 1978.

[6] 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.

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

[8] 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.

[9] S. Tavildar and P. Viswanath, "On the sum-rate of the vector Gaussian CEO problem," in *Proc. Conf. Rec. 39th Asilomar Conf. Signals, Syst. Comput.*, Oct./Nov. 2005, pp. 3–7.

[10] J. Chen and J. Wang, "On the vector Gaussian CEO problem," in *Proc. IEEE ISIT*, Aug. 2011, pp. 2050–2054.

[11] 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.

[12] 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.

[13] D. P. Palomar and S. Verdu, "Gradient of mutual information in linear vector Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 141–154, Jan. 2006.

[14] E. Ekrem and S. Ulukus, "An alternative proof for the capacity region of the degraded Gaussian MIMO broadcast channel," *IEEE Trans. Inf. Theory*, vol. 58, no. 4, pp. 2427–2433, Apr. 2012.

[15] E. Ekrem and S. Ulukus, "An outer bound for the vector Gaussian CEO problem," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2012, pp. 576–580.

[16] E. Ekrem and S. Ulukus. (Feb. 2012). "An outer bound for the vector Gaussian CEO problem." [Online]. Available: http://arxiv.org/abs/1202.0536

[17] J. Wang and J. Chen, "On the vector Gaussian L-terminal CEO problem," in *Proc. IEEE ISIT*, Jul. 2012, pp. 571–575.

[18] A. Dembo, "Information inequalities and uncertainty principles," Dept. Statist., Stanford Univ., Stanford, CA, USA, Tech. Rep. 75, 1990.

[19] 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.

[20] H. Weingarten, T. Liu, S. Shamai (Shitz), 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.

Ersen Ekrem received his Ph.D. degree from the Department of Electrical and Computer Engineering at the University of Maryland, College Park in August 2012. Prior to that, he received the B.S. and M.S. degrees in Electrical and Electronics Engineering from Bogazici University, Istanbul, Turkey, in 2006 and 2007, respectively. Currently, he is with Qualcomm, Santa Clara, USA.

He received the Distinguished Dissertation Fellowship from the ECE Department at the University of Maryland, College Park, in 2012. His research interests include information theory and wireless communications.

Sennur Ulukus (S'90–M'98) is a Professor of Electrical and Computer Engineering at the University of Maryland at College Park, where she also holds a joint appointment with the Institute for Systems Research (ISR). Prior to joining UMD, she was a Senior Technical Staff Member at AT&T Labs-Research. She received her Ph.D. degree in Electrical and Computer Engineering from Wireless Information Network Laboratory (WINLAB), Rutgers University, and B.S. and M.S. degrees in Electrical and Electronics Engineering from Bilkent University. Her research interests are in wireless communication theory and networking, network information theory for wireless communications, signal processing for wireless communications, information theoretic physical layer security, and energy harvesting communications.

Dr. Ulukus received the 2003 IEEE Marconi Prize Paper Award in Wireless Communications, an 2005 NSF CAREER Award, the 2010-2011 ISR Outstanding Systems Engineering Faculty Award, and the 2012 George Corcoran Education Award. She served as an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY (2007-2010) and IEEE TRANSACTIONS ON COMMUNICATIONS (2003-2007). She served as a Guest Editor for the *Journal of Communications and Networks* for the special issue on energy harvesting in wireless networks (2012), IEEE TRANSACTIONS ON INFORMATION THEORY for the special issue on interference networks (2011), IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS for the special issue on multiuser detection for advanced communication systems and networks (2008). She served as the TPC co-chair of the Communication Theory Symposium at 2014 IEEE Globecom, Communication Theory Symposium at 2013 IEEE ICC, Physical-Layer Security Workshop at 2011 IEEE Globecom, Physical-Layer Security Workshop at 2011 IEEE ICC, 2011 Communication Theory Workshop (IEEE CTW), Wireless Communications Symposium at 2010 IEEE ICC, Medium Access Control Track at 2008 IEEE WCNC, and Communication Theory Symposium at 2007 IEEE Globecom. She was the Secretary of the IEEE Communication Theory Technical Committee (CTTC) in 2007-2009.