Stylianou, Evagoras; Charalambous, Charalambos D.; Charalambous, Themistoklis

Joint Rate Distortion Function of a Tuple of Correlated Multivariate Gaussian Sources with Individual Fidelity Criteria

Published in:
Proceedings of IEEE International Symposium on Information Theory, ISIT 2021

DOI:
10.1109/ISIT45174.2021.9518279

Published: 01/09/2021

Document Version
Peer-reviewed accepted author manuscript, also known as Final accepted manuscript or Post-print

Please cite the original version:
Stylianou, E., Charalambous, C. D., & Charalambous, T. (2021). Joint Rate Distortion Function of a Tuple of Correlated Multivariate Gaussian Sources with Individual Fidelity Criteria. In *Proceedings of IEEE International Symposium on Information Theory, ISIT 2021* (pp. 2167-2172). IEEE.
https://doi.org/10.1109/ISIT45174.2021.9518279

Joint Rate Distortion Function of a Tuple of Correlated Multivariate Gaussian Sources with Individual Fidelity Criteria

Evagoras Stylianou*, Charalambos D. Charalambous†, and Themistoklis Charalambous‡
*Department of Electrical and Computer Engineering, Technical University of Munich,
†Department of Electrical and Computer Engineering, University of Cyprus
‡Department of Electrical Engineering and Automation, School of Electrical Engineering, Aalto University
Emails: evagoras.stylianou@tum.de, chadcha@ucy.ac.cy, themistoklis.charalambous@aalto.fi

***Abstract*—In this paper we analyze the joint rate distortion function (RDF), for a tuple of correlated sources taking values in abstract alphabet spaces (i.e., continuous) subject to two individual distortion criteria. First, we derive structural properties of the realizations of the reproduction Random Variables (RVs), which induce the corresponding optimal test channel distributions of the joint RDF. Second, we consider a tuple of correlated multivariate jointly Gaussian RVs, $X_1 : \Omega \to \mathbb{R}^{p_1}, X_2 : \Omega \to \mathbb{R}^{p_2}$ with two square-error fidelity criteria, and we derive additional structural properties of the optimal realizations, and use these to characterize the RDF as a convex optimization problem with respect to the parameters of the realizations. We show that the computation of the joint RDF can be performed by semidefinite programming. Further, we derive closed-form expressions of the joint RDF, such that Gray's [1] lower bounds hold with equality, and verify their consistency with the semidefinite programming computations.**

I. Literature Review, Problem Formulation, and Main Contributions

A. Literature Review

Gray [1, Theorem 3.1, Corollary 3.1] derived lower bounds on the joint rate distortion functions (RDFs), of a tuple of Random Variables (RVs) taking values in arbitrary, abstract spaces, $X_1 : \Omega \to \mathbb{X}_1$, $X_2 : \Omega \to \mathbb{X}_2$, with a weighted distortion, expressed in terms of conditional RDFs, and marginal RDFs. Gray and Wyner in [2], characterized the rate distortion region of a tuple of correlated RVs, using the joint, conditional and marginal RDFs. Xiao and Luo [3, Theorem 6] derived the closed-form expression of the joint RDF for a tuple of scalar-valued correlated Gaussian RVs, with two square-error distortion criteria, while Lapidoth and Tinguely [4] re-derived Xiao's and Luo's joint RDF using an alternative method. Xu, Liu and Chen [5] and Viswanatha, Akyol and Rose [6], generalized Wyner's common information [7] to its lossy counterpart, as the minimum common message rate on the Gray and Wyner rate region with sum rate equal to the joint RDF with two individual distortion functions. The analysis in [5], [6], includes the application of a tuple of scalar-valued, jointly Gaussian RVs. More recent work on rates that lie on the Gray and Wyner rate region are found in [8].

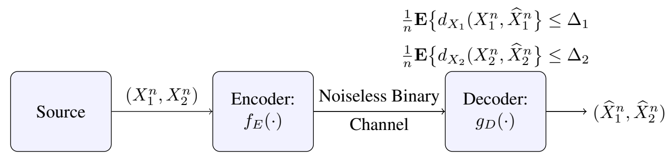

Fig. I.1. Lossy Compression of correlated sources with individual distortion criteria.

B. Problem Formulation

1) The Joint RDF with Individual Distortion Functions: This paper is concerned with the joint RDF of a tuple of RVs taking values in abstract spaces (i.e., continuous-valued RVs), $X_1 : \Omega \to \mathbb{X}_1, X_2 : \Omega \to \mathbb{X}_2$ of reconstructing X_i by $\widehat{X}_i : \Omega \to \widehat{\mathbb{X}}_i$, for $i = 1,2$, subject to two distortion functions $d_{X_i} : \mathbb{X}_i \times \widehat{\mathbb{X}}_i \to [0,\infty), i = 1,2$, defined by

$$R_{X_1,X_2}(\Delta_1,\Delta_2) = \inf_{\mathcal{M}(\Delta_1,\Delta_2)} I(X_1,X_2;\widehat{X}_1,\widehat{X}_2) \qquad \text{(I.1)}$$

where $I(X_1,X_2;\widehat{X}_1,\widehat{X}_2)$ is the mutual information of RVs (X_1,X_2) and $(\widehat{X}_1,\widehat{X}_2)$, the set $\mathcal{M}(\Delta_1,\Delta_2)$ is specified by

$$\mathcal{M}(\Delta_1,\Delta_2) = \Big\{ \widehat{X}_1 : \Omega \to \widehat{\mathbb{X}}_1,\ \widehat{X}_2 : \Omega \to \widehat{\mathbb{X}}_2 \Big| \mathbf{P}_{X_1,X_2,\widehat{X}_1,\widehat{X}_2} \text{ has } (X_1,X_2)\text{-marginal } \mathbf{P}_{X_1,X_2}, \mathbf{E}\big\{d_{X_i}(X_i,\widehat{X}_i)\big\} \leq \Delta_i,\ i = 1,2 \Big\} \qquad \text{(I.2)}$$

and the level of distortions are $\Delta_i \in [0,\infty), i = 1,2$. The joint RDF characterizes the infimum of all achievable rates of a sequence of rate distortion codes, (f_E, g_D), as depicted in Figure I.1, of reconstructing $(X_1^n, X_2^n) \triangleq \{(X_{1,t}, X_{2,t}) : t = 1,2,\ldots,n\}$, by $(\widehat{X}_1^n, \widehat{X}_2^n) \triangleq \{(\widehat{X}_{1,t}, \widehat{X}_{2,t}) : t = 1,2,\ldots,n\}$, where $\widehat{X}_{i,t} : \Omega \to \widehat{\mathbb{X}}_i$ $i = 1,2$, $t = 1,2,\ldots,n$ and $\mathbf{P}_{X_{1,t},X_{2,t}} = \mathbf{P}_{X_1,X_2}$, $\forall t$, with distortion $\frac{1}{n}\mathbf{E}\{d_{X_i}(X_i^n, \widehat{X}_i^n)\} \leq \Delta_i, i = 1,2$, for sufficiently large n. The computation of $R_{X_1,X_2}(\Delta_1,\Delta_2)$ is indispensable in the characterization of the Gray and Wyner rate region, and in the above mentioned applications.

Our first objective is to identify *structural properties* of realizations of the tuple of RVs $(\widehat{X}_1,\widehat{X}_2)$ in the set $\mathcal{M}(\Delta_1,\Delta_2)$, and structural properties of corresponding induced forward test channel distributions $\mathbf{P}_{\widehat{X}_1,\widehat{X}_2|X_1,X_2}$ or backward test channel distributions $\mathbf{P}_{X_1,X_2|\widehat{X}_1,\widehat{X}_2}$, such that $\mathbf{E}\big\{d_{X_i}(X_i,\widehat{X}_i)\big\} \leq \Delta_i,\ i = 1,2$, i.e., to characterize $R_{X_1,X_2}(\Delta_1,\Delta_2)$.

2) The Joint RDF of a Tuple of Multivariate Gaussian Sources: Our second objective is to compute the joint RDF $R_{X_1,X_2}(\Delta_1,\Delta_2)$, of a tuple of jointly independent and identically distributed multivariate Gaussian RVs, $(X_1^n,X_2^n) \stackrel{\triangle}{=} \{(X_{1,t},X_{2,t}) : t=1,2,\ldots,n\}$, where $X_{i,t}:\Omega\rightarrow\mathbb{R}^{p_i}$, $i=1,2$, $t=1,2,\ldots,n$, i.e., $\mathbf{P}_{X_{1,t},X_{2,t}}=\mathbf{P}_{X_1,X_2}$, $\forall t$ is a multivariate jointly Gaussian distribution and denoted by $(X_1,X_2)\in G(0,Q_{(X_1,X_2)})$, subject to two square-error distortion functions, all defined by

$$Q_{(X_{1,t},X_{2,t})}=\mathbf{E}\Big\{\begin{pmatrix}X_{1,t}\\X_{2,t}\end{pmatrix}\begin{pmatrix}X_{1,t}\\X_{2,t}\end{pmatrix}^{\mathrm{T}}\Big\}=\begin{pmatrix}Q_{X_1} & Q_{X_1,X_2}\\ Q_{X_1,X_2}^{\mathrm{T}} & Q_{X_2}\end{pmatrix} \tag{I.3}$$

$$X_{1,t}\in G(0,Q_{X_1}),\qquad X_{2,t}\in G(0,Q_{X_2}),\quad \forall t, \tag{I.4}$$

$$\widehat{X}_{1,t}:\Omega\rightarrow\widehat{\mathbb{X}}_1\stackrel{\triangle}{=}\mathbb{R}^{p_1},\quad \widehat{X}_{2,t}:\Omega\rightarrow\widehat{\mathbb{X}}_2\stackrel{\triangle}{=}\mathbb{R}^{p_2}\quad\forall t, \tag{I.5}$$

$$d_{X_i}(x_i^n,\widehat{x}_i^n)=\frac{1}{n}\sum_{t=1}^{n}||x_{i,t}-\widehat{x}_{i,t}||^2_{\mathbb{R}^{p_i}},\quad i=1,2, \tag{I.6}$$

where p_i are positive integers for $i=1,2$. Here $X\in G(0,Q_X)$ means X is a Gaussian RV, with zero mean and symmetric nonnegative definite covariance matrix $Q_X\succeq 0$.

C. Main Contributions

1) The derivation of structural properties of test channel distributions $\mathbf{P}_{\widehat{X}_1,\widehat{X}_2|X_1,X_2}$, and corresponding realizations of the reproduction RVs $(\widehat{X}_1,\widehat{X}_2)$ which induce these distributions, and characterize $R_{X_1,X_2}(\Delta_1,\Delta_2)$.
2) The characterization of $R_{X_1,X_2}(\Delta_1,\Delta_2)$ for jointly Gaussian multivariate sources, $X_1:\Omega\rightarrow\mathbb{R}^{p_1},X_2:\Omega\rightarrow\mathbb{R}^{p_2}$, with square-error distortion criteria, (I.3)-(I.6), parametrization of reproduction RVs $(\widehat{X}_1,\widehat{X}_2)$ and corresponding test channels, and calculation of $R_{X_1,X_2}(\Delta_1,\Delta_2)$ using convex numerical algorithms. Further, derivation of closed-form expressions for $R_{X_1,X_2}(\Delta_1,\Delta_2)$, to verify the numerical algorithms. This includes the distortion region $\mathscr{D}_{(X_1,X_2)}$, such that Gray's lower bound [1] holds with equality,

$$R_{X_1,X_2}(\Delta_1,\Delta_2)=R_{X_1}(\Delta_1)+R_{X_2}(\Delta_2)-I(X_1;X_2). \tag{I.7}$$

For case $p_1=p_2=1$, our results reproduce the value of the RDF derived by Xiao and Luo [3]. The tools used in this paper have been used to derive structural properties of distributed RDFs [9] and the nonanticipative RDF of multivariate Gaussian Markov [10] and autoregressive [11] processes.

II. Properties of Realizations of Test Channels

Let $\mathbb{Z}$ and $\mathbb{Z}_+$ be the set of integers and positive integers, respectively. Let $\mathbb{R}$ be the set of real numbers. The expression $\mathbb{R}^{n\times m}$ denotes the set of n by m matrices with elements the real numbers, for $n,\ m\in\mathbb{Z}_+$. For the symmetric matrix $Q\in\mathbb{R}^{n\times n}$, inequality $Q\succ 0$ (resp. $Q\succeq 0$) means the matrix is positive definite (resp. semi-definite). The notation $Q_2\succeq Q_1$ means that $Q_2-Q_1\succeq 0$. For any matrix $A\in\mathbb{R}^{p\times m},(p,m)\in\mathbb{Z}_+\times\mathbb{Z}_+$, we denote its transpose by A^{T}, and for $m=p$, we denote its trace and its determinant by $\mathrm{tr}(A)$ and $\det\big(A\big)$, respectively. The n by n identity (resp. zero) matrix is represented by I_n (resp. 0_n). For matrix $A\in\mathbb{R}^{p\times p}$, $\mathrm{diag}(A)$ is the matrix with diagonal entries those of A and zero elsewhere. Block-diag(A,B) is a square diagonal matrix in which the diagonal elements are square matrices $A\in\mathbb{R}^{p_1\times p_1}$ and $B\in\mathbb{R}^{p_2\times p_2}$, and the off-diagonal elements are zero. Given a triple of real-valued RVs $X_i:\Omega\rightarrow\mathbb{X}_i, i=1,2,3$, we say that RVs (X_2,X_3) are conditional independent given RV X_1 if $\mathbf{P}_{X_2,X_3|X_1}=\mathbf{P}_{X_2|X_1}\mathbf{P}_{X_3|X_1}-$a.s (almost surely); the specification a.s is often omitted. The mutual information between RV X and RV Y is denoted by $I(X;Y)$. The conditional covariance of the two-component vector RV $X=(X_1^{\mathrm{T}},X_2^{\mathrm{T}})^{\mathrm{T}}$, $X_i:\Omega\rightarrow\mathbb{R}^{p_i}, i=1,2$, conditioned on the two-component vector $\widehat{X}=(\widehat{X}_1^{\mathrm{T}},\widehat{X}_2^{\mathrm{T}})^{\mathrm{T}}$, $\widehat{X}_i:\Omega\rightarrow\mathbb{R}^{p_i}, i=1,2$, is denoted by $Q_{(X_1,X_2)|\widehat{X}}\stackrel{\triangle}{=}\mathrm{cov}\Big(X,X\Big|\widehat{X}\Big)\succeq 0$, where

$$Q_{(X_1,X_2)|\widehat{X}}=\begin{pmatrix}Q_{X_1|\widehat{X}} & Q_{X_1,X_2|\widehat{X}}\\ Q_{X_1,X_2|\widehat{X}}^{\mathrm{T}} & Q_{X_2|\widehat{X}}\end{pmatrix}\in\mathbb{R}^{(p_1+p_2)\times(p_1+p_2)},$$

$$\begin{aligned}Q_{X_1,X_2|\widehat{X}}&\stackrel{\triangle}{=}\mathrm{cov}\Big(X_1,X_2\Big|\widehat{X}\Big).\\ &\stackrel{(1)}{=}\mathbf{E}\Big\{\Big(X_1-\mathbf{E}\Big\{X_1\Big|\widehat{X}\Big\}\Big)\Big(X_2-\mathbf{E}\Big\{X_2\Big|\widehat{X}\Big\}\Big)^{\mathrm{T}}\Big\}\\ &=\mathbf{E}\Big\{E_1E_2^{\mathrm{T}}\Big\},\quad E_i\stackrel{\triangle}{=}X_i-\mathbf{E}\Big\{X_i\Big|\widehat{X}\Big\},\quad i=1,2\end{aligned}$$

and where (1) holds if $(X_1,X_2,\widehat{X}_1,\widehat{X}_2)$ is jointly Gaussian. Similarly for $Q_{X_i|\widehat{X}}, i=1,2$. Consequently, for jointly Gaussian RVs $(X_1,X_2,\widehat{X}_1,\widehat{X}_2)$, and the two-component vector RV $E\stackrel{\triangle}{=}(E_1^{\mathrm{T}},E_2^{\mathrm{T}})^{\mathrm{T}}$, we have $Q_{(X_1,X_2)|\widehat{X}}=\Sigma_{(E_1,E_2)}$ (unconditional).

In Theorem II.1 we identify a structural property of the tuple $(\widehat{X}_1,\widehat{X}_2)$ to achieve a lower bound on $I(X_1,X_2;\widehat{X}_1,\widehat{X}_2)$, for any tuple of RVs (X_1,X_2) with arbitrary distribution $\mathbf{P}_{X_1,X_2}$.

Theorem II.1. *Let $(X_1,X_2,\widehat{X}_1,\widehat{X}_2)$ be arbitrary RVs taking values in the abstract spaces $\mathbb{X}_1\times\mathbb{X}_2\times\widehat{\mathbb{X}}_1\times\widehat{\mathbb{X}}_2$, with arbitrary joint distribution $\mathbf{P}_{X_1,X_2,\widehat{X}_1,\widehat{X}_2}$, and $\mathbb{X}_1\times\mathbb{X}_2-$joint marginal the fixed distribution $\mathbf{P}_{X_1,X_2}$ of (X_1,X_2).*
(a) Define

$$\overline{X}_i^{\mathrm{cm}}=g_i^{\mathrm{cm}}(\widehat{X}_1,\widehat{X}_2)\stackrel{\triangle}{=}\mathbf{E}\Big\{X_i\Big|\widehat{X}\Big\},\quad i=1,2, \tag{II.8}$$

$g_i^{\mathrm{cm}}:\widehat{\mathbb{X}}_1\times\widehat{\mathbb{X}}_2\rightarrow\widehat{\mathbb{X}}_i,\quad g_i^{\mathrm{cm}}(\cdot)$ *are measurable functions,* $i=1,2$.

Then, the following inequality holds:

$$I(X_1,X_2;\widehat{X}_1,\widehat{X}_2)\geq I\big(X_1,X_2;g_1^{\mathrm{cm}}(\widehat{X}_1,\widehat{X}_2),g_2^{\mathrm{cm}}(\widehat{X}_1,\widehat{X}_2)\big). \tag{II.9}$$

Moreover, if there exist RVs $(\widehat{X}_1,\widehat{X}_2)$ such that the functions $g_i^{\mathrm{cm}}(\cdot,\cdot)$ satisfy $g_i^{\mathrm{cm}}(\widehat{X}_1,\widehat{X}_2)=\widehat{X}_i-a.s$ for $i=1,2$, then the inequality in (II.9) holds with equality.
(b) Let $\mathbb{X}_1\times\mathbb{X}_2\times\widehat{\mathbb{X}}_1\times\widehat{\mathbb{X}}_2=\mathbb{R}^{p_1}\times\mathbb{R}^{p_2}\times\mathbb{R}^{p_1}\times\mathbb{R}^{p_2}$, $p_1,p_2\in\mathbb{Z}_+$. For all measurable functions $g_i(\widehat{X}_1,\widehat{X}_2)$, $i=1,2$ then

$$\mathbf{E}\Big\{||X_i-g_i(\widehat{X}_1,\widehat{X}_2)||^2_{\mathbb{R}^{p_i}}\Big\}\geq\mathbf{E}\Big\{||X_i-\mathbf{E}\Big\{X_i\Big|\widehat{X}\Big\}||^2_{\mathbb{R}^{p_i}}\Big\}, i=1,2.$$

(c) If $\mathbb{X}_1\times\mathbb{X}_2\times\widehat{\mathbb{X}}_1\times\widehat{\mathbb{X}}_2=\mathbb{R}^{p_1}\times\mathbb{R}^{p_2}\times\mathbb{R}^{p_1}\times\mathbb{R}^{p_2}$, $p_1,p_2\in\mathbb{Z}_+$, $d_{X_i}(x_i,\widehat{x}_i)=||x_i-\widehat{x}_i||^2_{\mathbb{R}^{p_i}}$, $i=1,2$, $g_i^{\mathrm{cm}}(\widehat{X}_1,\widehat{X}_2)=\widehat{X}_i-a.s$, $i=1,2$, then the joint RDF of (I.1) is characterized by

$$R_{X_1,X_2}(\Delta_1,\Delta_2)=\inf_{\mathcal{M}^{\mathrm{cm}}(\Delta_1,\Delta_2)}I(X_1,X_2;\widehat{X}_1,\widehat{X}_2) \tag{II.10}$$

where $\mathcal{M}^{\mathrm{cm}}(\Delta_1,\Delta_2)$ is specified by the subset of $\mathcal{M}(\Delta_1,\Delta_2)$, with the additional restriction $\widehat{X}_i=\mathbf{E}\Big\{X_i\Big|\widehat{X}\Big\}, i=1,2$.

Proof. (a) By properties of mutual information, we have

$$\begin{aligned} I(X_1,X_2;\widehat{X}_1,\widehat{X}_2) &\overset{(1)}{=} I(X_1,X_2;\widehat{X}_1,\widehat{X}_2,\overline{X}_1^{\rm cm},\overline{X}_2^{\rm cm}) \\ &\overset{(2)}{=} I(X_1,X_2;\widehat{X}_1,\widehat{X}_2|\overline{X}_1^{\rm cm},\overline{X}_2^{\rm cm}) \\ &\quad + I(X_1,X_2;\overline{X}_1^{\rm cm},\overline{X}_2^{\rm cm}) \\ &\overset{(3)}{\geq} I(X_1,X_2;\overline{X}_1^{\rm cm},\overline{X}_2^{\rm cm}), \end{aligned} \tag{II.11}$$

where (1) is due to $\overline{X}_i^{\rm cm}, i=1,2$, are functions of $(\widehat{X}_1,\widehat{X}_2)$, (2) is due to the chain rule of mutual information, and (3) is due to $I(X_1,X_2;\widehat{X}_1,\widehat{X}_2|\overline{X}_1^{\rm cm},\overline{X}_2^{\rm cm}) \geq 0$. Thus, (II.9) is obtained. If $g_i^{\rm cm}(\widehat{X}_1,\widehat{X}_2) = X_i - {\rm a.s}, \quad i=1,2$, hold, then $I(X_1,X_2;\widehat{X}_1,\widehat{X}_2|\overline{X}_1^{\rm cm},\overline{X}_2^{\rm cm}) = 0$, and hence the inequality (II.11) become equality. (b) The inequality is well-known, due to the orthogonal projection theorem. (c) This is due to (a), (b). □

III. Structural Properties of Test Channels and Characterization of Joint RDF for Multivariate Jointly Gaussian Sources

This section makes use of Theorem II.1 to derive additional structural properties of test channels for the joint RDF $R_{X_1,X_2}(\Delta_1,\Delta_2)$ of jointly Gaussian sources with square-error distortions, defined by (I.3)-(I.6).

Theorem III.1 (Sufficient conditions for the lower bounds of Theorem II.1 to be achieved)**.** *Consider the quadruple of zero mean RVs $(X_1,X_2,\widehat{X}_1,\widehat{X}_2)$ taking values in $\mathbb{R}^{p_1}\times\mathbb{R}^{p_2}\times\mathbb{R}^{p_1}\times\mathbb{R}^{p_2}$, $p_1,p_2\in\mathbb{Z}_+$, with jointly Gaussian distribution i.e, $\mathbf{P}_{X_1,X_2,\widehat{X}_1,\widehat{X}_2} = \mathbf{P}^G_{X_1,X_2,\widehat{X}_1,\widehat{X}_2}$ and $\mathbb{X}_1\times\mathbb{X}_2-$joint marginal the fixed distribution $\mathbf{P}_{X_1,X_2}$ of (X_1,X_2). Define the vectors,*

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \widehat{X} = \begin{pmatrix} \widehat{X}_1 \\ \widehat{X}_2 \end{pmatrix}, \overline{X}^{\rm cm} \triangleq \mathbf{E}\Big\{ \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \Big| \widehat{X} \Big\} = \begin{pmatrix} \overline{X}_1^{\rm cm} \\ \overline{X}_2^{\rm cm} \end{pmatrix}.$$

(a) If the vector of conditional means satisfy,

$$\overline{X}^{\rm cm} = \mathbf{E}\big\{X\big\} + {\rm cov}\big(X,\widehat{X}\big)\big\{{\rm cov}\big(\widehat{X},\widehat{X}\big)\big\}^{\dagger}\Big(\widehat{X}-\mathbf{E}\big\{\widehat{X}\big\}\Big) = \widehat{X}$$

where † denotes pseudoinvesrse, then the equalities hold:

$$\overline{X}_1^{\rm cm} \triangleq \mathbf{E}\Big\{X_1\Big|\widehat{X}\Big\} = \widehat{X}_1, \ \overline{X}_2^{\rm cm} \triangleq \mathbf{E}\Big\{X_2\Big|\widehat{X}\Big\} = \widehat{X}_2. \tag{III.12}$$

(b) If the inverse of ${\rm cov}\big(\widehat{X},\widehat{X}\big)$ exists and $\mathbf{E}\big\{X\big\} = \mathbf{E}\big\{\widehat{X}\big\} = 0$, then (III.12) holds if Condition 1 holds:

$$\text{Condition 1.} \ {\rm cov}\big(X,\widehat{X}\big)\big\{{\rm cov}\big(\widehat{X},\widehat{X}\big)\big\}^{-1} = I_{p_1+p_2}. \tag{III.13}$$

(c) The lower bounds of Theorem II.1 are achieved, if there exist $(\widehat{X}_1,\widehat{X}_2)$ such that $\overline{X}^{\rm cm} = \widehat{X}$, or the statement of (b) holds.

Proof. Follows by properties of jointly Gaussian RVs. □

In the next lemma, we apply Theorem II.1 and Theorem III.1 to find a parametric jointly Gaussian realization of $(\widehat{X}_1,\widehat{X}_2)$, that induces the set of test channels of the joint RDF $R_{X_1,X_2}(\Delta_1,\Delta_2)$ for (I.3)-(I.6).

Lemma III.1 (Preliminary parametrization of test channel)**.** *Consider the joint RDF $R_{X_1,X_2}(\Delta_1,\Delta_2)$ for (I.3)-(I.6). The following hold.*

(a) A jointly Gaussian distribution $\mathbf{P}_{X_1,X_2,\widehat{X}_1,\widehat{X}_2}$ minimizes $I(X_1,X_2;\widehat{X}_1,\widehat{X}_2)$, subject to two average distortions.
(b) The test channel distribution $\mathbf{P}_{\widehat{X}_1,\widehat{X}_2|X_1,X_2}$ of the joint RDF $R_{X_1,X_2}(\Delta_1,\Delta_2)$ is induced by the parametric Gaussian realization of $(\widehat{X}_1,\widehat{X}_2)$, in terms of the matrices (H,Q_V), as

$$\widehat{X} = HX + V \tag{III.14}$$

$$H \in \mathbb{R}^{(p_1+p_2)\times(p_1+p_2)}, \quad V:\Omega\rightarrow\mathbb{R}^{(p_1+p_2)}, \tag{III.15}$$

$$V \in G(0,Q_{(V_1,V_2)}),\ Q_{(V_1,V_2)} \succeq 0,\ V \text{ and } X \text{ indep.}, \tag{III.16}$$

(c) Consider part (b) and suppose there exist matrices $(H,Q_{(V_1,V_2)})$ such that Theorem III.1.(a) holds, i.e., $\overline{X}^{\rm cm} = \widehat{X}$-a.s., or in the special case Condition 1 holds. Then the infimum in $R_{X_1,X_2}(\Delta_1,\Delta_2)$ is taken over the subset $\mathcal{M}^{{\rm cm},G}(\Delta_1,\Delta_2)\subseteq\mathcal{M}^{\rm cm}(\Delta_1,\Delta_2)$,

$$\mathcal{M}^{{\rm cm},G}(\Delta_1,\Delta_2) \triangleq \Big\{\widehat{X}:\Omega\rightarrow\mathbb{R}^{(p_1+p_2)}\Big| \text{(III.14)} - \text{(III.16)} \textit{ hold}, \ \overline{X}_i^{\rm cm} = \widehat{X}_i,\ \mathbf{E}\big\{||X_i-\widehat{X}_i||^2_{\mathbb{R}^{p_i}}\big\} \leq \Delta_i, i=1,2\Big\} \tag{III.17}$$

Proof. (a) This is similar to the classical RDF $R_X(\Delta)$ of a Gaussian RV $X\in G(0,Q_X)$ with square-error distortion. (b) By part (a), the test channel distribution $\mathbf{P}_{\widehat{X}_1,\widehat{X}_2|X_1,X_2}$ is conditionally Gaussian with linear conditional mean $\mathbf{E}\big\{X|\widehat{X}\big\}$ and non-random covariance ${\rm cov}(X,\widehat{X}|X)$. Such a distribution is induced by the realizations (III.14)-(III.16). (c) Follows from Theorem III.1.(c). □

Next, we construct $(H,Q_{(V_1,V_2)})$ such that $\overline{X}_i^{\rm cm} = \mathbf{E}\big\{X_i|\widehat{X}\big\} = \widehat{X}_i - a.s$ for $i=1,2$, and characterize $R_{X_1,X_2}(\Delta_1,\Delta_2)$.

Theorem III.2 (Realization of optimal test channels and characterization of joint RDF)**.** *Consider the joint RDF $R_{X_1,X_2}(\Delta_1,\Delta_2)$ for (I.3)-(I.6).*
(a) The test channel distribution $\mathbf{P}_{\widehat{X}_1,\widehat{X}_2|X_1,X_2}$ of the RDF $R_{X_1,X_2}(\Delta_1,\Delta_2)$ is induced by the parametric realization (III.14)-(III.16), where the matrices, (H,Q_V) satisfy,

$$HQ_{(X_1,X_2)} = Q_{(X_1,X_2)} - \Sigma_{(E_1,E_2)} = Q_{(X_1,X_2)}H^T \succeq 0, \tag{III.18}$$

$$Q_{(V_1,V_2)} = HQ_{(X_1,X_2)} - HQ_{(X_1,X_2)}H^T \succeq 0. \tag{III.19}$$

Moreover, $R_{X_1,X_2}(\Delta_1,\Delta_2)$ is characterized by,

$$R_{X_1,X_2}(\Delta_1,\Delta_2) = \inf_{\mathcal{Q}^{\dagger}(\Delta_1,\Delta_2)} \frac{1}{2}\log\Big\{\frac{\det\big(Q_{(X_1,X_2)}\big)}{\det\big(\Sigma_{(E_1,E_2)}\big)}\Big\}, \tag{III.20}$$

$$\mathcal{Q}^{\dagger}(\Delta_1,\Delta_2) \triangleq \Big\{\Sigma_{(E_1,E_2)} : (H,Q_{(V_1,V_2)}) \textit{ satisfy (III.18), (III.19)}, \ {\rm tr}\big(\Sigma_{E_1}\big)\leq\Delta_1, \ \ {\rm tr}\big(\Sigma_{E_2}\big)\leq\Delta_2\Big\}. \tag{III.21}$$

(b) Suppose $Q_{(X_1,X_2)}\succ 0$. If $R_{X_1,X_2}(\Delta_1,\Delta_2)<\infty$, then the matrices, $(H,Q_{(V_1,V_2)})$, of part (a) reduce to,

$$H = I_{p_1+p_2} - \Sigma_{(E_1,E_2)}Q^{-1}_{(X_1,X_2)}, \tag{III.22}$$

$$Q_{(V_1,V_2)} = \Sigma_{(E_1,E_2)} - \Sigma_{(E_1,E_2)}Q^{-1}_{(X_1,X_2)}\Sigma_{(E_1,E_2)} \succeq 0, \tag{III.23}$$

$$Q_{(X_1,X_2)} - \Sigma_{(E_1,E_2)} \succeq 0, \quad \Longleftrightarrow \tag{III.24}$$

$$\Sigma_{(E_1,E_2)} - \Sigma_{(E_1,E_2)}Q^{-1}_{(X_1,X_2)}\Sigma_{(E_1,E_2)} \succeq 0. \tag{III.25}$$

and $Q^{\dagger}(\Delta_1,\Delta_2)$ *in (III.20) is replaced by* $\overset{\circ}{\mathcal{Q}}(\Delta_1,\Delta_2)$*, given by*

$$\overset{\circ}{\mathcal{Q}}(\Delta_1,\Delta_2) \triangleq \Big\{\Sigma_{(E_1,E_2)}: \ Q_{(X_1,X_2)} \succeq \Sigma_{(E_1,E_2)} \succeq 0, \quad \mathrm{tr}\left(\Sigma_{E_1}\right) \le \Delta_1, \mathrm{tr}\left(\Sigma_{E_2}\right) \le \Delta_2\Big\}. \tag{III.26}$$

Proof. See Appendix VI. □

Lemma III.2. *Consider* $R_{X_1,X_2}(\Delta_1,\Delta_2)$ *of Theorem III.2, defined by (III.20) and assume* $Q_{(X_1,X_2)} \succ 0$*, and* $R_{X_1,X_2}(\Delta_1,\Delta_2) < +\infty$*. The Lagrange functional is,*

$$\mathscr{L} \triangleq \frac{1}{2}\log\left\{\frac{\det\left(Q_{(X_1,X_2)}\right)}{\det\left(\Sigma_{(E_1,E_2)}\right)}\right\} + \mathrm{tr}\left(\Theta\left(\Sigma_{(E_1,E_2)} - Q_{(X_1,X_2)}\right)\right) + \lambda_1\left(\mathrm{tr}\left(\Sigma_{E_1}\right) - \Delta_1\right) + \lambda_2\left(\mathrm{tr}\left(\Sigma_{E_2}\right) - \Delta_2\right) - \mathrm{tr}\left(V\Sigma_{(E_1,E_2)}\right)$$

where $\Theta \succeq 0$*,* $V \succeq 0$*,* $\lambda_i \in [0,\infty), i = 1,2$*. The optimal* $\Sigma_{(E_1,E_2)} \in \overset{\circ}{\mathcal{Q}}(\Delta_1,\Delta_2)$ *for* $R_{X_1,X_2}(\Delta_1,\Delta_2)$ *is found as follows.*
(i) Stationarity:

$$-\frac{1}{2}\Sigma^{-1}_{(E_1,E_2)} + \begin{bmatrix}\lambda_1 I_{p_1} & 0\\ 0 & \lambda_2 I_{p_2}\end{bmatrix} + \Theta + V = 0. \tag{III.27}$$

(ii) Complementary Slackness:

$$\lambda_1\left(\mathrm{tr}\left(\Sigma_{E_1}\right) - \Delta_1\right) = 0, \ \ \lambda_2\left(\mathrm{tr}\left(\Sigma_{E_2}\right) - \Delta_2\right) = 0, \tag{III.28}$$

$$\mathrm{tr}\left(V\Sigma_{(E_1,E_2)}\right) = 0, \ \mathrm{tr}\left(\Theta\left(\Sigma_{(E_1,E_2)} - Q_{(X_1,X_2)}\right)\right) = 0. \tag{III.29}$$

(iii) Primal Feasibility: Defined by $\overset{\circ}{\mathcal{Q}}(\Delta_1,\Delta_2)$*.*
(iv) Dual Feasibility: $\lambda_1 \ge 0$*,* $\lambda_2 \ge 0$*,* $\Theta \succeq 0$*,* $V \succeq 0$*.*
Moreover, the following hold.
(a) $V = 0$*, and*

$$\Sigma_{(E_1,E_2)} = \frac{1}{2}\left(\begin{bmatrix}\lambda_1 I_{p_1} & 0\\ 0 & \lambda_2 I_{p_2}\end{bmatrix} + \Theta\right)^{-1} \succ 0. \tag{III.30}$$

(b) If $Q_{(X_1,X_2)} - \Sigma_{(E_1,E_2)} \succ 0$ *then* $\Theta = 0$*, and*

$$\Sigma_{(E_1,E_2)} = \frac{1}{2}\left(\begin{bmatrix}\lambda_1 I_{p_1} & 0\\ 0 & \lambda_2 I_{p_2}\end{bmatrix}\right)^{-1} \succ 0. \tag{III.31}$$

Proof. The derivation is standard hence it is omitted. □

The next two theorems are obtained from Lemma III.2.

Theorem III.3 (Joint RDF for a positive surface)**.** *Consider the characterization of joint RDF* $R_{X_1,X_2}(\Delta_1,\Delta_2)$ *of Theorem III.2, defined by (III.20), and assume* $Q_{(X_1,X_2)} \succ 0$ *(i.e., this implies* $Q_{X_1} \succ 0, Q_{X_2} \succ 0$*). Define the set*

$$\mathscr{D}_{(X_1,X_2)} = \left\{(\Delta_1,\Delta_2) \in [0,\infty)\times[0,\infty) \middle| Q_{(X_1,X_2)} - \Sigma_{(E_1,E_2)} \succ 0\right\}.$$

The joint RDF $R_{X_1,X_2}(\Delta_1,\Delta_2)$ *for* $(\Delta_1,\Delta_2) \in \mathscr{D}_{(X_1,X_2)}$ *is*

$$R_{X_1,X_2}(\Delta_1,\Delta_2) = \frac{1}{2}\log\left\{\frac{\det\left(Q_{(X_1,X_2)}\right)}{\det\left(\Sigma_{E_1}\right)\det\left(\Sigma_{E_2}\right)}\right\} = (I.7)$$

$$\Sigma_{E_1} = \mathrm{diag}\left(\frac{\Delta_1}{p_1},\ldots,\frac{\Delta_1}{p_1}\right), \ \Sigma_{E_2} = \mathrm{diag}\left(\frac{\Delta_2}{p_2},\ldots,\frac{\Delta_2}{p_2}\right)$$

and this is achieved by the covariance matrix $\Sigma_{(E_1,E_2)}$ *with* $\Sigma_{E_1,E_2} = Q_{X_1,X_2|\hat{X}} = 0$*, and Gray's lower bound (I.7) holds.*

Proof. For any element of the set $\mathscr{D}_{(X_1,X_2)}$ then $Q_{(X_1,X_2)} - \Sigma_{(E_1,E_2)} \succ 0$, and the statements follow from Lemma III.2. □

Remark III.1. *For the scalar-valued RVs, i.e.,* $p_1 = p_2 = 1$*, we have verified that Lemma III.2 produces the closed-form expression of* $R_{X_1,X_2}(\Delta_1,\Delta_2)$ *as derived in [3, Theorem 6]. However, for the multivariate case of Lemma III.2, to obtain the closed-form expression is challenging. To make the problem tractable, in Theorem III.4, we use the canonical variable form of the tuple* (X_1,X_2)*, as described in [8] and [12].*

The algorithm to transform the tuple (X_1,X_2) to the canonical variable form is presented below [12, Algorithm 2.10].

Algorithm III.1. *[12, Algorithm 2.10] Transformation of a variance matrix to its canonical variable form.*
Data : $p_1,p_2 \in Z_+$*,* $Q_{(X_1,X_2)} \in \mathbb{R}^{(p_1+p_2)\times(p_1+p_2)}$*, satisfying* $Q_{(X_1,X_2)} = Q^T_{(X_1,X_2)} \succ 0$*, with decomposition (I.3).*
1) Perform singular value decompositions (SVD), $Q_{X_i} = U_i D_i U_i^T$*,* $i = 1,2$ *with* $U_i \in \mathbb{R}^{p_i\times p_1}$*, orthogonal and* $D_i = \mathrm{diag}(d_{1,1},\ldots,d_{i,p_i}) \in \mathbb{R}^{p_1\times p_1}$*,* $d_{i,1} \ge d_{i,2} \ge \cdots \ge d_{i,p_i} > 0$*.*
2) Perform SVD of $D_1^{-\frac{1}{2}}U_1^T Q_{X_1,X_2} U_2 D_2^{-\frac{1}{2}} = U_3 D_3 U_4^T$ *with* $U_3 \in \mathbb{R}^{p_1\times p_1}$*,* $U_4 \in \mathbb{R}^{p_2\times p_2}$ *orthogonal and*

$$D_3 = \text{Block-diag}\left(I_{p_{11}}, D_4, 0_{p_{13}\times p_{23}}\right) \in \mathbb{R}^{p_1\times p_2},$$
$$D_4 = \mathrm{diag}(d_{4,1},\ldots,d_{4,p_{12}}) \in \mathbb{R}^{p_{12}\times p_{22}}, \ 1 > d_{4,1} \ge \cdots \ge d_{4,p_{12}} > 0,$$
$$p_i = p_{i1} + p_{i2} + p_{i3}, \ \ i = 1,2, \quad p_{11} = p_{21}, \quad p_{12} = p_{22}$$

3) Compute the new variance matrix and the transformation to the canonical variable representation $(X_1 \mapsto S_1 X_1,\ X_2 \mapsto S_2 X_2)$ *according to*

$$Q_{\mathrm{cvf}} = \begin{pmatrix} I_{p_1} & D_3\\ D_3^T & I_{p_2}\end{pmatrix}, \quad S_1 = U_3^T D_1^{-\frac{1}{2}} U_1^T, \ S_2 = U_4^T D_2^{-\frac{1}{2}} U_2^T$$

Theorem III.4. *Consider the statement of Theorem III.2.(b), with* $(X_1,X_2) \in G(0,Q_{(X_1,X_2)}), Q_{(X_1,X_2)} \succ 0$*. Determine the canonical variable form of the tuple* (X_1,X_2)*, according to [12, Definition 2.2] by using algorithm Algorithm III.1, and restrict attention to indices,* $p_{11} = p_{21} = 0$*. Then,* $n = p_{12} = p_{22}$ *and* $p_1 = p_{12} + p_{13}, p_2 = p_{22} + p_{23}$*. Similarly, transform* $(E_1,E_2) \in G(0,\Sigma_{(E_1,E_2)})$ *with* $\overline{p}_{11} = \overline{p}_{21} = 0$ *and* $\overline{n} = \overline{p}_{12} = \overline{p}_{22}$*,* $\overline{p}_1 = \overline{p}_{12} + \overline{p}_{13}, \overline{p}_2 = \overline{p}_{22} + \overline{p}_{23}$*.*
The joint RDF $R_{X_1,X_2}(\Delta_1,\Delta_2)$ *of Theorem III.2.(b), is equivalently characterized by*

$$R_{X_1,X_2}(\Delta_1,\Delta_2) = \inf_{\overset{\circ}{\mathcal{Q}}(\Delta_1,\Delta_2)} \frac{1}{2}\log\left\{\frac{\det\left(D_1\right)\det\left(D_2\right)\det\left(Q_{\mathrm{cvf}}\right)}{\det\left(\overline{D}_1\right)\det\left(\overline{D}_2\right)\det\left(\Sigma_{\mathrm{cvf}}\right)}\right\}$$

where,

$$\overset{\circ}{\mathcal{Q}}(\Delta_1,\Delta_2) \triangleq \Big\{\overline{n} \in \mathbb{Z}_+, \ \overline{d}_{4,i} \in (0,1), \ i = 1,\ldots,\overline{n}, \quad \overline{d}_{1,i} \in (0,\infty), \ i = 1,\ldots,\overline{p}_1, \ \overline{d}_{2,i} \in (0,\infty), \ i = 1,\ldots,\overline{p}_2 : \quad \sum_{i=1}^{\overline{p}_1}\overline{d}_{1,i} \le \Delta_1, \ \sum_{i=1}^{\overline{p}_2}\overline{d}_{2,i} \le \Delta_2, \ Q_{(X_1,X_2)} - \Sigma_{(E_1,E_2)} \succeq 0\Big\}$$

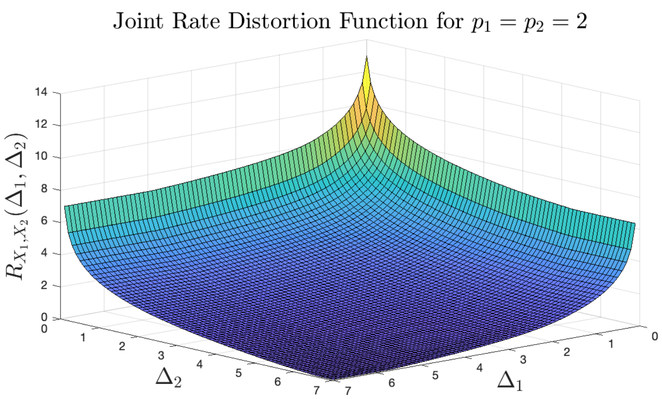

Fig. III.2. Joint RDF $R_{X_1,X_2}(\Delta_1,\Delta_2)$ of source of Section IV, $p_1 = p_2 = 2$.

and

$$\det\left(\Sigma_{\text{cvf}}\right) = \det\left(I_{\overline{p}_1} - \overline{D}_3 \overline{D}_3^{T}\right)$$
$$= \begin{cases} 1, & \text{if} \quad \overline{p}_{13} > 0,\ \overline{p}_{23} > 0,\ \overline{p}_{12} = \overline{p}_{22} = 0, \\ \prod_{i=1}^{\overline{n}} \left(1 - \overline{d}_{4,i}^2\right), & \text{if} \quad \overline{p}_{12} = \overline{p}_{22} = \overline{n},\ \overline{p}_{13} \geq 0,\ \overline{p}_{23} \geq 0. \end{cases}$$

Proof. By Theorem III.2.(b) and applying [12, Definition 2.2] and Algorithm III.1 we obtain the results. ☐

Remark III.2. *$R_{X_1,X_2}(\Delta_1,\Delta_2)$ of Theorem III.4, is much easier to optimize, due to its structure.*

IV. Evaluation of the Joint RDF via SDP

We can express the optimization problem of Theorem III.2 as a semidefinite program (SDP) as follows, define $\Xi_1^{\text{T}} = \text{Block-diag}\big(I_{p_1}\ 0_{p_2}\big)$ and $\Xi_2^{\text{T}} = \text{Block-diag}\big(0_{p_1}\ I_{p_2}\big)$,

$$\begin{aligned} \min_{\Sigma_{(E_1,E_2)}} \quad & \frac{1}{2} \log \left\{ \frac{\det\left(Q_{(X_1,X_2)}\right)}{\det\left(\Sigma_{(E_1,E_2)}\right)} \right\} \\ \text{s.t.} \quad & Q_{(X_1,X_2)} - \Sigma_{(E_1,E_2)} \succeq 0,\ \Sigma_{(E_1,E_2)} \succeq 0, \\ & \text{tr}\left(\Xi_i^{\text{T}} \Sigma_{(E_1,E_2)} \Xi_i\right) \leq \Delta_i,\ i = 1,2 \end{aligned} \tag{IV.32}$$

Then, we can solve the SDP (IV.32) by using the CVX [13]. Below, we calculate the optimal $\Sigma_{(E_1,E_2)}$ for a multivariate example $X_i : \Omega \to \mathbb{R}^2$, $i = 1,2$, with covariance,

$$Q_{(X_1,X_2)} = \left(\begin{array}{cc|cc} 3.929 & -0.11 & 0.642 & 0.976 \\ -0.11 & 2.629 & -0.859 & 0.337 \\ \hline 0.642 & -0.859 & 2.142 & 1.797 \\ 0.976 & 0.337 & 1.797 & 3.495 \end{array}\right).$$

Fig. III.2 depicts $R_{X_1,X_2}(\Delta_1,\Delta_2), (\Delta_1,\Delta_2) \in [0,\infty) \times [0,\infty)$. Below we distinguish two cases.
Case 1. Given distortions $(\Delta_1,\Delta_2) = (0.4, 0.5)$, the solution of (III.20), (III.26) is given by

$$\Sigma_{(E_1,E_2)} = \text{diag}\left(0.2, 0.2, 0.25, 0.25\right),\ Q_{(X_1,X_2)} - \Sigma_{(E_1,E_2)} \succ 0$$

Distortions Δ_1 and Δ_2 are equally divided among the diagonal elements of the first and second 2-by-2 diagonal blocks of $\Sigma_{(E_1,E_2)}$ respectively, and the rest of the values are zero. Hence, $(0.4, 0.5) \in \mathscr{D}_{(X_1,X_2)}$; this re-confirms Theorem III.3.

Case 2. Given distortions $(\Delta_1,\Delta_2) = (1.65, 1.85)$, the optimal error covariance matrix is given by,

$$\Sigma_{(E_1,E_2)} = \left(\begin{array}{cc|cc} 0.849 & -0.0017 & -0.0053 & 0.0036 \\ -0.0017 & 0.801 & -0.144 & 0.0961 \\ \hline -0.0053 & -0.144 & 0.804 & 0.293 \\ 0.0036 & 0.0961 & 0.293 & 1.05 \end{array}\right)$$

and $Q_{(X_1,X_2)} - \Sigma_{(E_1,E_2)} \succeq 0$ but not positive definite. Unlike Case 1, $\Sigma_{(E_1,E_2)}$ is not block-diagonal, i.e., $\Sigma_{E_1,E_2} \neq 0$, as in Theorem III.3, hence $(1.65, 1.85) \notin \mathscr{D}_{(X_1,X_2)}$. This choice of distortions corresponds to Lemma III.2.(b).

V. Conclusion

The joint RDF $R_{X_1,X_2}(\Delta_1,\Delta_2)$, with individual distortion criteria, is analyzed, with emphasis on the structural properties of realizations of the reproduction RVs $(\widehat{X}_1, \widehat{X}_2)$ of (X_1, X_2), and corresponding optimal test channel distribution, $\mathbf{P}_{\widehat{X}_1,\widehat{X}_2|X_1,X_2}$. Closed-form expressions of $R_{X_1,X_2}(\Delta_1,\Delta_2)$ are derived for a strictly positive surface of the distortion region, and a numerical technique is presented, which verifies the closed-form expressions.

VI. Appendices

Proof of Theorem III.2. Consider (III.14)-(III.16). To identify $(H, Q_{(V_1,V_2)})$ such that $\overline{X}_i^{\text{cm}} = \mathbf{E}\Big\{X_i \Big| \widehat{X}\Big\} = \widehat{X}_i, i = 1,2$, we make use of the following preliminary calculations. The covariance of X and $\widehat{X}$ is,

$$Q_{X,\widehat{X}} = \mathbf{E}\Big\{X\Big(HX+V\Big)^{\text{T}}\Big\} = Q_{(X_1,X_2)} H^{\text{T}}. \tag{VI.33}$$

By (III.14)-(III.16), the covariance of $\widehat{X} = HX + V$ is

$$Q_{\widehat{X}} = \mathbf{E}\Big\{\widehat{X}\widehat{X}^{\text{T}}\Big\} = HQ_{(X_1,X_2)}H^{\text{T}} + Q_{(V_1,V_2)}, \tag{VI.34}$$

Consider the special case when *Condition 1,* (III.13) holds:

$$\begin{aligned} & \text{cov}\big(X,\widehat{X}\big)\big\{\text{cov}\big(\widehat{X},\widehat{X}\big)\big\}^{-1} = I_{p_1+p_2} \iff Q_{X,\widehat{X}} Q_{\widehat{X}}^{-1} = I_{p_1+p_2} \\ & \Longrightarrow Q_X H^{\text{T}} = HQ_X H^{\text{T}} + Q_{(V_1,V_2)} \quad \text{by (VI.33), (VI.34)} \\ & \Longrightarrow Q_{(V_1,V_2)} = Q_{(X_1,X_2)}H^{\text{T}} - HQ_{(X_1,X_2)}H^{\text{T}}. \end{aligned} \tag{VI.35}$$

Next, we turn to the identification of H. By the definition of covariance of the errors, then $\Sigma_{(E_1,E_2)} \stackrel{\triangle}{=} \text{cov}(X, X|\widehat{X})$, and

$$\begin{aligned} \Sigma_{(E_1,E_2)} &= \text{cov}(X,X) - \text{cov}(X,\widehat{X})\big\{\text{cov}(\widehat{X},\widehat{X})\big\}^{-1}\text{cov}(X,\widehat{X})^{\text{T}} \\ &= Q_{(X_1,X_2)} - HQ_{(X_1,X_2)}, \ \text{by (III.13),(VI.33)} \\ &\Longrightarrow HQ_{(X_1,X_2)} = Q_{(X_1,X_2)} - \Sigma_{(E_1,E_2)} = Q_{(X_1,X_2)}H^{\text{T}} \quad \text{(VI.36)} \\ &\Longrightarrow H = I_{p_1+p_2} - \Sigma_{(E_1,E_2)} Q_{(X_1,X_2)}^{-1}, \ \text{if } Q_{(X_1,X_2)} \succ 0. \end{aligned}$$

Using (VI.36) into (VI.35) then we have

$$Q_{(V_1,V_2)} = Q_{(X_1,X_2)}H^{\text{T}} - HQ_{(X_1,X_2)}H^{\text{T}} = Q_{(V_1,V_2)}^{\text{T}} \tag{VI.37}$$
$$= Q_{(X_1,X_2)} - \Sigma_{(E_1,E_2)} - HQ_{(X_1,X_2)}H^{\text{T}}. \tag{VI.38}$$

Hence, $(H, Q_{(V_1,V_2)})$ are obtained. The general case is shown by using properties of pseudoinverse. The rest follow.

VII. Acknowledgments

This work was supported in parts by the European Regional Development Fund and the Republic of Cyprus through the Research Promotion Foundation Projects EXCELLENCE/1216/0365 and EXCELLENCE/1216/0296.

REFERENCES

[1] R. Gray, "A new class of lower bounds to information rates of stationary sources via conditional rate-distortion functions," *IEEE Transactions on Information Theory*, vol. 19, no. 4, pp. 480–489, 1973.

[2] R. Gray and A. Wyner, "Source coding for a simple network," *Bell System Technical Journal*, vol. 53, no. 9, pp. 1681–1721, 1974.

[3] J.-J. Xiao and Z.-Q. Luo, "Compression of correlated Gaussian sources under individual distortion criteria," in *43rd Allerton Conference on Communication, Control, and Computing*, 2005, pp. 438–447.

[4] A. Lapidoth and S. Tinguely, "Sending a bivariate Gaussian over a Gaussian MAC," *IEEE Transactions on Information Theory*, vol. 56, no. 6, pp. 2714–2752, 2010.

[5] G. Xu, W. Liu, and B. Chen, "A lossy source coding interpretation of wyner's common information," *IEEE Transactions on Information Theory*, vol. 62, no. 2, pp. 754–768, 2015.

[6] K. B. Viswanatha, E. Akyol, and K. Rose, "The lossy common information of correlated sources," *IEEE Transactions on Information Theory*, vol. 60, no. 6, pp. 3238–3253, 2014.

[7] A. Wyner, "The common information of two dependent random variables," *IEEE Transactions on Information Theory*, vol. 21, no. 2, pp. 163–179, 1975.

[8] C. D. Charalambous and J. H. van Schuppen, "Characterization of conditional independence and weak realizations of multivariate Gaussian random variables: Applications to networks," in *IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2020, pp. 2444–2449.

[9] M. Gkagkos and C. D. Charalambous, "Structural properties of optimal test channels for Gaussian multivariate distributed sources," in *Proceedings of the 2021 IEEE International Symposium on Information Theory*, 2021.

[10] C. D. Charalambous, T. Charalambous, C. Kourtellaris, and J. H. van Schuppen, "Structural properties of nonanticipatory epsilon entropy of multivariate Gaussian sources," in *IEEE International Symposium on Information Theory (ISIT)*, 2020, pp. 2867–2872.

[11] C. D. Charalambous, C. Kourtellaris, T. Charalambous, and J. H. van Schuppen, "Generalizations of nonanticipative rate distortion function to multivariate nonstationary Gaussian autoregressive processes," in *2019 IEEE 58th Conference on Decision and Control (CDC)*, 2019, pp. 8190–8195.

[12] C. D. Charalambous and J. H. van Schuppen, "A new approach to lossy network compression of a tuple of correlated multivariate Gaussian rvs," *arXiv preprint arXiv:1905.12695*, 2019.

[13] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.