

# Optimality of KLT for High-Rate Transform Coding of Gaussian Vector-Scale Mixtures: Application to Reconstruction, Estimation and Classification\*

Soumya Jana and Pierre Moulin

Beckman Institute, Coordinated Science Lab, and ECE Department  
University of Illinois at Urbana-Champaign  
405 N. Mathews Ave., Urbana, IL 61801  
Email: {jana,moulin}@ifp.uiuc.edu

Submitted: May, 2004. Revised: March, 2006 and June, 2006.

## Abstract

The Karhunen-Loève transform (KLT) is known to be optimal for high-rate transform coding of Gaussian vectors for both fixed-rate and variable-rate encoding. The KLT is also known to be suboptimal for some non-Gaussian models. This paper proves high-rate optimality of the KLT for variable-rate encoding of a broad class of non-Gaussian vectors: Gaussian vector-scale mixtures (GVSM), which extend the Gaussian scale mixture (GSM) model of natural signals. A key concavity property of the scalar GSM (same as the scalar GVSM) is derived to complete the proof. Optimality holds under a broad class of quadratic criteria, which include mean squared error (MSE) as well as generalized  $f$ -divergence loss in estimation and binary classification systems. Finally, the theory is illustrated using two applications: signal estimation in multiplicative noise and joint optimization of classification/reconstruction systems.

**Keywords:** Chernoff distance, classification, estimation,  $f$ -divergence, Gaussian scale mixture, high resolution quantization, Karhunen-Loève transform, mean squared error, multiplicative noise, quadratic criterion.

---

\*This research was supported in part by ARO under contract numbers ARO DAAH-04-95-1-0494 and ARMY WUHT-011398-S1, by DARPA under Contract F49620-98-1-0498, administered by AFOSR, and by NSF under grant CDA 96-24396. This work was presented in part at the DCC Conference, Snowbird, UT, March 2003.

# 1 Introduction

The *Karhunen Loève transform (KLT)* is known to be optimal for orthogonal transform coding of Gaussian sources under the *mean squared error (MSE)* criterion and an assumption of high-rate scalar quantizers. Optimality holds for both fixed-rate [1, 2] as well as variable-rate encoding, where the quantized coefficients are entropy-coded [3]. Conversely, examples have been given where the KLT is suboptimal for compression of non-Gaussian sources also in both fixed-rate and variable-rate frameworks [4]. Now, can we identify nontrivial non-Gaussian sources for which the KLT is optimal? To the best of our knowledge, this remains an open problem in the literature. In this paper, we assume high-resolution scalar quantizers, variable-rate encoding and a general *quadratic criterion*, and show optimality of certain KLT's for encoding a new family of Gaussian mixtures, which we call *Gaussian vector-scale mixture (GVSM)* distributions.

We define GVSM distributions by extending the notion of *Gaussian scale mixture (GSM)*, studied by Wainwright *et al.* [5]. A decorrelated GSM vector  $\mathbf{X}$  is the product  $\mathbf{Z}V$  of a zero-mean Gaussian vector  $\mathbf{Z}$  with identity covariance matrix and an independent scale random vector  $V$ . Extending the above notion, we define a decorrelated GVSM  $\mathbf{X}$  to be the elementwise product  $\mathbf{Z} \odot \mathbf{V}$  of  $\mathbf{Z}$  and an independent scale random vector  $\mathbf{V}$ . In the special case, where components of  $\mathbf{V}$  are identical with probability one, a decorrelated GVSM vector reduces to a decorrelated GSM vector. More generally, a GVSM takes the form  $\mathbf{X} = C^T(\mathbf{Z} \odot \mathbf{V})$ , where  $C$  is unitary. Clearly, conditioned on the scale-vector  $\mathbf{V}$ , the GVSM  $\mathbf{X}$  is Gaussian. The matrix  $C$  governs the correlation and the mixing distribution  $\mu$  of  $\mathbf{V}$  governs the nonlinear dependence among the components of  $\mathbf{X}$ . A GVSM density is a special case of a Gaussian mixture density,  $p(\mathbf{x}) = \int q(\mathbf{x}|\theta) \mu(d\theta)$ , where  $q(\mathbf{x}|\theta)$  is the Gaussian density with parameters denoted by  $\theta = (\mathbf{m}, \Sigma)$  (mean  $\mathbf{m}$  and covariance  $\Sigma$ ) and  $\mu$  is the mixing distribution. Specifically, for  $p(\mathbf{x})$  to be a GVSM density, we require that, under the mixing distribution  $\mu$ , only Gaussian densities  $q(\mathbf{x}|\theta)$  satisfying the following two conditions contribute to the mixture: (1)  $\mathbf{m} = \mathbf{0}$  and (2)  $\Sigma$  has the singular value decomposition of the form  $\Sigma = C^T D C$  for a fixed unitary  $C$ .

The marginal distribution of any GVSM is, clearly, a scalar GVSM, which is the same as a scalar GSM. Scalar GSM distributions, in turn, represent a large collection, where the characteristic function as well as the probability density function (pdf) is characterized by *complete monotonicity* as well as a certain *positive definiteness* property [6, 7]. Examples of scalar GSM densities include densities as varied as the generalized Gaussian family (e.g., Laplacian density), the stable family (e.g., Cauchy density), symmetrized gamma, lognormal

densities, etc. [5]. Obviously, by moving from the scalar to the general GVSM family, one encounters an even richer collection (which is populated by varying the mixing distribution  $\mu$  and the unitary matrix  $C$ ).

Not only does the GVSM family form a large collection, but it is also useful for analysis and modeling of natural signals exhibiting strong nonlinear dependence. In fact, GVSM models subsume two well established and closely related signal models: *spherically invariant random processes (SIRP)* [8] and random cascades of GSM's [5]. SIRP models find application in diverse contexts such as speech modeling [9, 10], video coding [11], modeling of radar clutter [12, 13] and signal modeling in fast fading wireless channels [14]. The GSM cascade models of natural images are also quite powerful [5, 15], and incorporate earlier wavelet models [16, 17, 18]. Note that any SIRP satisfying Kolmogorov's regularity condition is a GSM random process and *vice versa* [19]. In contrast with the extensive statistical characterization of SIRP/GSM models, only limited results relating to optimal compression of such sources are available [20, 21, 22]. Our result fills, to an extent, this gap between statistical signal modeling and compression theory.

Although the objective of most signal coding applications is faithful reconstruction, compressed data are also often used for estimation and classification purposes. Examples include denoising of natural images [23], estimation of camera motion from compressed video [24], automatic target recognition [25, 26, 27] and detection of abnormalities in compressed medical images [28, 29]. Compression techniques can be designed to be optimal in a *minimum mean squared error (MMSE)* sense [30, 31, 32]. Further, a generalized vector quantization method has been suggested for designing combined compression/estimation systems [33]. In detection problems, various *f-divergence* criteria [34], including the *Chernoff distance* measure [35, 36], have been used. Poor [37] generalized the notion of *f-divergence* to also include certain estimation error measures such as the MMSE, and studied the effect of compression on *generalized f-divergences*. In this paper, we optimize the transform coder under a broad class of quadratic criteria, which includes not only the usual MSE, but also the performance loss in MMSE estimation, the loss in Chernoff distance in binary classification, and more generally, the loss in generalized *f-divergence* arising in a variety of decision systems.

There have been some attempts at optimizing transform coders for encoding a given source. Goldberg [38] sets up an optimization problem to design high-rate transform vector quantization systems subject to an entropy constraint. He then finds the optimal wavelet transform for a Gaussian source under a certain transform cost function. Under the MSE distortion criterion, Mallat [3] derived the optimal bit allocation strategy for a general non-Gaussian source using high-resolution variable-rate scalar quantizers and identified the cost

function which the optimal transform must minimize but did not carry out the minimization. The reader can find a lucid account of transform coding in [39]. In the same paper, optimality of KLT for transform coding of Gaussian vectors is dealt with in some detail. In contrast, we shall look into the optimality of KLT for GVSM vectors in the paper. In addition, general results on high-rate quantization can be found [40].

In this paper, we improve upon our earlier results [41] and show that certain KLT's of GVSM vectors, which we call *primary* KLT's, are optimal under high-resolution quantization, variable-rate encoding and a broad variety of quadratic criteria. Specifically, we closely follow Mallat's approach [3] and minimize the cost function given by Mallat for GVSM sources. Deriving this optimality result for GVSM sources presents some difficulty because, unlike the Gaussian case, the differential entropy for GSM does not admit an easily tractable expression. We circumvent this impediment by extending a key concavity property of the differential entropy of scalar Gaussian densities to scalar GSM densities, and applying it to the GVSM transform coding problem. We also derive the optimal quantizers and the optimal bit allocation strategy.

We apply the theory to a few estimation/classification problems, where the generalized  $f$ -divergence measures the performance. For example, we consider signal estimation in multiplicative noise, where GVSM observations are shown to arise due to GVSM noise. Note the special case of multiplicative Gaussian noise arises, for instance, as a consequence of the Doppler effect in radio measurements [42]. Further, our analysis is flexible enough to accommodate competing design criteria. We illustrate this feature in a joint classification/reconstruction design, trading off class discriminability as measured by Chernoff distance against MSE signal fidelity. Similar joint design methodology has proven attractive in recent applications. For example, compression algorithms for medical diagnosis can be designed balancing accuracy of automatic detection of abnormality against visual signal fidelity [28, 29].

The organization of the paper is as follows. Sec. 2 introduces the notation. Poor's result on the effect of quantization on generalized  $f$ -divergence [37], is also mentioned. In Sec. 3, we set up the transform coding problem under the MSE criterion and extend our formulation for a broader class of quadratic criteria. We introduce the GVSM distribution in Sec. 4. In Sec. 5, we establish the optimality of certain primary KLT's for encoding GVSM sources. We apply our theory to the estimation and classification contexts in Sec. 6 and 7, respectively. Sec. 8 concludes the paper. Readers interested only in the results for classical transform coding (under MSE criterion), may skip Secs. 2.2, 3.2, 5.5, 6 and 7.

## 2 Preliminaries

### 2.1 Notation

Let  $\mathbb{R}$  and  $\mathbb{R}_+$  denote the respective sets of real numbers and positive real numbers. In general, lowercase letters (e.g.,  $c$ ) denote scalars, boldface lowercase (e.g.,  $\mathbf{x}$ ) vectors, uppercase (e.g.,  $C$ ,  $X$ ) matrices and random variables, and boldface uppercase (e.g.,  $\mathbf{X}$ ) random vectors. Unless otherwise specified, vectors and random vectors have length  $N$ , and matrices have size  $N \times N$ . The  $k$ -th element of vector  $\mathbf{x}$  is denoted by  $[\mathbf{x}]_k$  or  $x_k$ , and the  $(i, j)$ -th element of matrix  $C$ , by  $[C]_{ij}$  or  $C_{ij}$ . The constants  $\mathbf{0}$  and  $\mathbf{1}$  denote the vectors of all zeros and all ones, respectively. Further, the vector of  $\{x_k\}$  is denoted by  $\mathbf{vect}\{x_k\}$ , whereas  $\text{Diag}\{\mathbf{x}\}$  denotes the diagonal matrix with diagonal entries specified by the vector  $\mathbf{x}$ . By  $I$ , denote the  $N \times N$  identity matrix, and by  $\mathcal{U}$ , the set of  $N \times N$  unitary matrices.

The symbol ‘ $\odot$ ’ denotes the elementwise product between two vectors or two matrices:

$$\begin{aligned} \mathbf{u} \odot \mathbf{v} &= \mathbf{vect}\{u_k v_k\}, \\ [A \odot B]_{ij} &= A_{ij} B_{ij}, \quad 1 \leq i, j \leq N. \end{aligned}$$

In particular, let  $\mathbf{v}^2 = \mathbf{v} \odot \mathbf{v}$ . The symbol ‘ $\sim$ ’ has two different meanings in two unrelated contexts. Firstly, ‘ $\sim$ ’ denotes asymptotic equality. For example, given two sequences  $\{a^{(n)}\}$  and  $\{b^{(n)}\}$  indexed by  $n$ , ‘ $a \sim b$  as  $n \rightarrow \infty$ ’ indicates

$$\lim_{n \rightarrow \infty} \frac{a^{(n)}}{b^{(n)}} = 1.$$

On the other hand, ‘ $\mathbf{V} \sim \mu$ ’ indicates that the random vector  $\mathbf{V}$  follows probability distribution  $\mu$ . Further, the relation ‘ $\mathbf{X} \stackrel{d}{=} \mathbf{Y}$ ’ indicates that the random vectors  $\mathbf{X}$  and  $\mathbf{Y}$  share the same distribution. Define a scalar quantizer by a mapping  $Q : \mathbb{R} \rightarrow \mathcal{X}$ , where the alphabet  $\mathcal{X} \subset \mathbb{R}$  is at most countable. The uniform scalar quantizer with step size  $\Delta > 0$  is denoted by  $Q(\cdot; \Delta)$ . Also denote  $Q(\cdot; \mathbf{\Delta}) = \mathbf{vect}\{Q(\cdot; \Delta_k)\}$ . Moreover, denote the function composition  $f(g(\mathbf{x}))$  by  $f \circ g(\mathbf{x})$ .

The expectation of a function  $f$  of a random vector  $\mathbf{X} \sim \mu$  is given by  $\mathbb{E}[f(\mathbf{X})] = \int \mu(d\mathbf{x}) f(\mathbf{x})$  and is sometimes denoted by  $\mathbb{E}_\mu[f]$ . The differential entropy  $h(\mathbf{X})$  (respectively, the entropy  $H(\mathbf{X})$ ) of a random vector  $\mathbf{X}$  taking values in  $\mathbb{R}^N$  (respectively, a discrete alphabet  $\mathcal{X}$ ) with pdf (respectively, probability mass function (pmf))  $q$  is defined by  $\mathbb{E}[-\log q(\mathbf{X})]$  [43]. Depending on the context, we sometimes denote  $h(\mathbf{X})$  by  $h(q)$  (respectively,  $H(\mathbf{X})$  by  $H(q)$ ). In a binary hypothesis test

$$H_0 : \mathbf{X} \sim \mathcal{P}^0 \quad \text{versus} \quad H_1 : \mathbf{X} \sim \mathcal{P}^1, \quad (2.1)$$

denote  $E_{\mathcal{P}^0}$  by  $E_0$ . Further, the  $f$ -divergence between  $p^0$  and  $p^1$ , the respective Lebesgue densities of  $\mathcal{P}^0$  and  $\mathcal{P}^1$ , is defined by  $E_0[f \circ l]$ , where  $l = p^1/p^0$  is the likelihood ratio and  $f$  is convex [37]. In particular, the  $f$ -divergence with  $f = -\log$  defines Kullback-Leibler divergence  $D(p^0||p^1) := E_0[\log(p^0/p^1)]$  [43].

Denote the zero-mean Gaussian probability distribution with covariance matrix  $\Sigma$  by  $\mathcal{N}(\mathbf{0}, \Sigma)$ , and its pdf by

$$\phi(\mathbf{x}; \Sigma) = \frac{1}{(2\pi)^{\frac{N}{2}} \det^{\frac{1}{2}}(\Sigma)} \exp\left(-\frac{1}{2}\mathbf{x}^T \Sigma^{-1} \mathbf{x}\right). \quad (2.2)$$

In the scalar case ( $N = 1$ ), partially differentiate representation (2.2) of  $\phi(x; \sigma^2)$  twice with respect to  $\sigma^2$ , and denote

$$\phi'(x; \sigma^2) = \frac{\partial \phi(x; \sigma^2)}{\partial(\sigma^2)} = \frac{1}{\sqrt{8\pi\sigma^{10}}} (x^2 - \sigma^2) \exp\left(-\frac{x^2}{2\sigma^2}\right), \quad (2.3)$$

$$\phi''(x; \sigma^2) = \frac{\partial^2 \phi(x; \sigma^2)}{\partial(\sigma^2)^2} = \frac{1}{\sqrt{32\pi\sigma^{18}}} (x^4 - 6\sigma^2 x^2 + 3\sigma^4) \exp\left(-\frac{x^2}{2\sigma^2}\right). \quad (2.4)$$

## 2.2 Generalized $f$ -divergence under Quantization

Now consider Poor's generalization of  $f$ -divergence [37], which measures performance in the estimation/classification systems considered in Sec. 6 and Sec. 7.

**Definition 2.1** [37] *Let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be a continuous convex function,  $l : \mathbb{R}^N \rightarrow \mathbb{R}$  be a measurable function and  $\mathcal{P}$  be a probability measure on  $\mathbb{R}^N$ . The generalized  $f$ -divergence for the triplet  $(f, l, \mathcal{P})$  is defined by*

$$D_f(l, \mathcal{P}) := E_{\mathcal{P}}[f \circ l]. \quad (2.5)$$

The usual  $f$ -divergence  $D_f(l, \mathcal{P}^0) = E_0[f \circ l]$  arising in binary hypothesis testing (2.1) is of course a special case where  $l$  represents the likelihood ratio.

The generalized  $f$ -divergence based on any transformation  $\tilde{\mathbf{X}} = t(\mathbf{X}) \sim \tilde{\mathcal{P}}$  of  $\mathbf{X} \sim \mathcal{P}$ , is given by  $D_f(\tilde{l}, \tilde{\mathcal{P}})$ , where  $\tilde{l}(\tilde{\mathbf{x}}) = E[l(\mathbf{X})|t(\mathbf{X}) = \tilde{\mathbf{x}}]$  gives the divergence upon transformation  $t$ . By Jensen's inequality [37],

$$D_f(\tilde{l}, \tilde{\mathcal{P}}) \leq D_f(l, \mathcal{P})$$

under any  $t$ . Further, denote the incurred loss in divergence by

$$\Delta D_f(t; l, \mathcal{P}) := D_f(l, \mathcal{P}) - D_f(\tilde{l}, \tilde{\mathcal{P}}). \quad (2.6)$$

Of particular interest are transformations of the form

$$t(\mathbf{X}) = Q(\mathbf{s}(\mathbf{X}); \Delta \mathbf{1}), \quad (2.7)$$

i.e., scalar quantization of components of  $\mathbf{s}(\mathbf{X})$  with equal step sizes  $\Delta$ , where  $\mathbf{s} : \mathbb{R}^N \rightarrow \mathbb{R}^N$  is an invertible vector function with differentiable components. Subject to regularity conditions, Poor obtained an asymptotic expression for  $\Delta D_f(t; l, \mathcal{P})$  for such  $t$  as  $\Delta \rightarrow 0$ .

**Definition 2.2** (POOR [37]) *The triplet  $(f, l, \mathcal{P})$  is said to satisfy regularity condition  $\mathcal{R}(f, l, \mathcal{P})$  if the following hold. The function  $f$  is convex, its second derivative  $f''$  exists and is continuous on the range of  $l$ . The measure  $\nu$  on  $\mathbb{R}^N$ , defined by  $\nu(d\mathbf{x}) = f'' \circ l(\mathbf{x}) \mathcal{P}(d\mathbf{x})$ , has Lebesgue density  $n$ . Further,*

- a)  $n$  and  $l$  are continuously differentiable;
- b)  $E_{\mathcal{P}}[f \circ l] < \infty$ ;
- c)  $\int d\nu l^2 < \infty$  and  $\int d\nu \|\nabla l\|^2 < \infty$ ;
- d)  $\int_{\mathbb{R}^N} \nu(d\mathbf{x}) \sup_{\{\mathbf{y}: \|\mathbf{x}-\mathbf{y}\| \leq \epsilon\}} \left| \|\nabla l(\mathbf{x})\|^2 - \|\nabla l(\mathbf{y})\|^2 \right| < \infty$  for some  $\epsilon > 0$ ;
- e)  $\sup_{\{\mathbf{x}, \mathbf{y} \in \mathbb{R}^N: \|\mathbf{x}-\mathbf{y}\| < \delta\}} \frac{f'' \circ l(\mathbf{x})}{f'' \circ l(\mathbf{y})} < \infty$  a.s.- $[\nu]$  for some  $\delta > 0$ .

**Theorem 2.3** (POOR, Theorem 3 of [37]) *Subject to regularity condition  $\mathcal{R}(f, l \circ \mathbf{s}, \mathcal{P} \circ \mathbf{s})$ ,*

$$\Delta D_f(t; l, \mathcal{P}) \sim \frac{1}{24} \Delta^2 E_{\mathcal{P}} [\|J^{-1} \nabla l\|^2 f'' \circ l] \quad \text{as } \Delta \rightarrow 0, \quad (2.8)$$

where  $t(\mathbf{x}) = Q(\mathbf{s}(\mathbf{x}); \Delta \mathbf{1})$  and  $J$  is the Jacobian of  $\mathbf{s}$ .

In Sec. 6, we apply Theorem 2.3 to transform encoding (3.11), which takes the form  $t(\mathbf{x}) = Q(\mathbf{s}(\mathbf{x}); \Delta \mathbf{1})$  with invertible  $\mathbf{s}$ .

### 3 Transform Coding Problem

Now we turn our attention to the main theme of the paper, and formulate the transform coding problem analytically. Specifically, consider the orthogonal transform coder depicted in Fig. 1. A unitary matrix  $U \in \mathcal{U}$  transforms source  $\mathbf{X}$  taking values in  $\mathbb{R}^N$  into the vector

$$\bar{\mathbf{X}} = U\mathbf{X}. \quad (3.1)$$

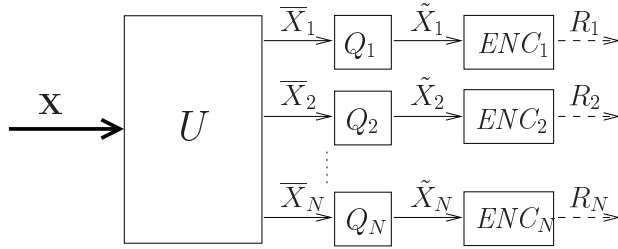


Figure 1: Transform coding: Unitary transform  $U$  and scalar quantizers  $\{Q_k\}$ .

The  $k$ -th ( $1 \leq k \leq N$ ) transformed component  $\bar{X}_k$  is quantized to  $\tilde{X}_k = Q_k(\bar{X}_k)$ . By  $p$ ,  $\bar{p}$  and  $\tilde{p}$ , denote the pdf's of  $\mathbf{X}$  and  $\bar{\mathbf{X}}$ , and the pmf of the quantized vector  $\tilde{\mathbf{X}}$ , respectively. Also, the  $k$ -th marginals of  $\mathbf{X}$ ,  $\bar{\mathbf{X}}$  and  $\tilde{\mathbf{X}}$  are, respectively, denoted by  $p_k$ ,  $\bar{p}_k$  and  $\tilde{p}_k$ . Each quantized coefficient  $\tilde{X}_k$  is independently entropy coded. Denote by  $l_k(\tilde{x}_k)$  the length of the codeword assigned to  $\tilde{X}_k = \tilde{x}_k$ . Then the expected number of bits required to code  $\tilde{X}_k$  is  $R_k(U, Q_k) = \mathbb{E} [l_k(\tilde{X}_k)]$ . Further, by source coding theory [43], the codebook can be chosen such that

$$H(\tilde{X}_k) \leq R_k(U, Q_k) < H(\tilde{X}_k) + 1, \quad (3.2)$$

provided  $H(\tilde{X}_k)$  is finite. Next consider the classical reconstruction at the decoder. More versatile decoders performing estimation/classification are considered in Sec. 6 and Sec. 7.

### 3.1 Minimum MSE Reconstruction

As shown in Fig. 2, the decoder decodes the quantized vector  $\tilde{\mathbf{X}}$  losslessly and reconstructs  $\mathbf{X}$  to

$$\hat{\mathbf{X}} = U^T \tilde{\mathbf{X}}. \quad (3.3)$$

This reconstruction minimizes the *mean-squared error (MSE)*

$$D(U, \{Q_k\}) := \mathbb{E} \left[ \|\mathbf{X} - \hat{\mathbf{X}}\|^2 \right] = \mathbb{E} \left[ \left\| U^T (\bar{\mathbf{X}} - \tilde{\mathbf{X}}) \right\|^2 \right] = \mathbb{E} \left[ \|\bar{\mathbf{X}} - \tilde{\mathbf{X}}\|^2 \right] \quad (3.4)$$

(the second equality follows from (3.1) and (3.3), and the third equality holds because  $U$  is unitary) over all possible reconstructions based on  $\tilde{\mathbf{X}}$  [3]. Note that  $D(U, \{Q_k\}) = \sum_{k=1}^N d_k(U, Q_k)$  is additive over transformed components, where component MSE's are given by

$$d_k(U, Q_k) := \mathbb{E} \left[ \left( \bar{X}_k - \tilde{X}_k \right)^2 \right], \quad 1 \leq k \leq N. \quad (3.5)$$

The goal is to minimize the average bit rate  $R(U, \{Q_k\}) = \frac{1}{N} \sum_{k=1}^N R_k(U, Q_k)$  required to achieve a certain distortion  $D(U, \{Q_k\}) = D$ . We now make the following assumption [46]:

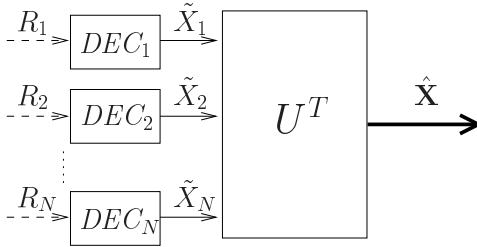


Figure 2: Minimum MSE reconstruction at the decoder.

**Assumption 3.1** *The pdf  $p(\mathbf{x})$  of  $\mathbf{X}$  is continuous. Also, there exists  $U \in \mathcal{U}$  such that  $h(\bar{X}_k) < \infty$  for all  $1 \leq k \leq N$ .*

Note that continuity of  $p(\mathbf{x})$  implies continuity of  $\bar{p}(\bar{\mathbf{x}})$  as well as of each  $\bar{p}_k(\bar{x}_k)$  ( $1 \leq k \leq N$ ) for any  $U \in \mathcal{U}$ . This in turn implies each  $h(\bar{X}_k) > -\infty$ . Hence, by assumption, each  $h(\bar{X}_k)$  is finite for some  $U \in \mathcal{U}$ .

**Assumption 3.2 (HIGH RESOLUTION QUANTIZATION)** *For each  $k$ ,  $\{Q_k^{(n)}\}$  is a sequence of high-resolution quantizers such that*

$$\sup_{\bar{x}_k \in \mathbb{R}} \left[ \bar{x}_k - Q_k^{(n)}(\bar{x}_k) \right] \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Denote  $\tilde{X}_k^{(n)} = Q_k^{(n)}(\bar{X}_k)$ . Clearly, each  $H(\tilde{X}_k^{(n)}) \rightarrow \infty$  as  $n \rightarrow \infty$ . Therefore, by (3.2), we have

$$R_k(U, Q_k) \sim H(\tilde{X}_k), \quad 1 \leq k \leq N. \quad (3.6)$$

Here and henceforth, asymptotic relations hold as  $n \rightarrow \infty$  unless otherwise stated. We drop index  $n$  for convenience.

**Fact 3.3** *Under Assumptions 3.1 and 3.2,*

$$\inf_{Q_k} d_k(U, Q_k) \sim \frac{1}{12} 2^{2h(\bar{X}_k)} 2^{-2R_k}, \quad (3.7)$$

where the infimum is subject to the entropy constraint  $H(\tilde{X}_k) \leq R_k$ . Further, there also exists a sequence of entropy-constrained uniform quantizers  $Q_k$  such that

$$d_k(U, Q_k) \sim \frac{1}{12} 2^{2h(\bar{X}_k)} 2^{-2R_k}. \quad (3.8)$$

Fact 3.3 is mentioned in [40], and follows from the findings reported in [44] and [45]. The same result is also rigorously derived in [46]. Note that Assumption 3.1 (alongwith Assumption 3.2) does not give the most general condition but is sufficient for Fact 3.3 to hold. However, Assumption 3.1 can potentially be generalized in view of more sophisticated results such as [47].

In view of (3.7) and (3.8), uniform quantizers are asymptotically optimal. Hence, we choose uniform quantizers

$$Q_k(\bar{x}_k) \sim Q(\bar{x}_k; \Delta_k), \quad 1 \leq k \leq N, \quad (3.9)$$

such that

$$\mathbf{\Delta} = \mathbf{vect}\{\Delta_k\} = \Delta \boldsymbol{\alpha} \quad (3.10)$$

for some constant  $\boldsymbol{\alpha} \in \mathbb{R}_+^N$  as  $\Delta \rightarrow 0$ . Therefore, the transform encoding  $\{Q_k(\bar{\mathbf{X}})\}$  now takes the form

$$\tilde{\mathbf{X}} = Q(U\mathbf{X}; \mathbf{\Delta}) = Q(U\mathbf{X}; \Delta \boldsymbol{\alpha}). \quad (3.11)$$

In view of (3.9), component MSE (3.5) is given by [46]

$$d_k(\Delta_k) \sim \frac{\Delta_k^2}{12}, \quad 1 \leq k \leq N, \quad (3.12)$$

and the overall MSE (3.4) by

$$D(\mathbf{\Delta}) = \sum_{k=1}^N d_k(\Delta_k) \sim \frac{1}{12} \sum_{k=1}^N \Delta_k^2. \quad (3.13)$$

Using (3.12) and noting equality in (3.7), the component bit rate  $R_k(U, \Delta_k)$  in (3.6) is asymptotically given by

$$R_k(U, \Delta_k) \sim H(\tilde{X}_k) \sim h(\bar{X}_k) - \log \Delta_k, \quad 1 \leq k \leq N, \quad (3.14)$$

which amounts to the average bit-rate

$$R(U, \mathbf{\Delta}) = \frac{1}{N} \sum_{k=1}^N R_k(U, \Delta_k) \sim \frac{1}{N} \sum_{k=1}^N [h(\bar{X}_k) - \log \Delta_k]. \quad (3.15)$$

Subject to a distortion constraint  $D(\mathbf{\Delta}) = D$ , we minimize  $R(U, \mathbf{\Delta})$  with respect to the encoder  $(U, \mathbf{\Delta})$ . For any given  $U$ , optimal step sizes  $\{\Delta_k\}$  are equal, and given by the following well known result, see e.g., [3].

**Result 3.4** *Subject to a distortion constraint  $D$ , the average bit rate  $R(U, \mathbf{\Delta})$  is minimized if the quantizer step sizes are equal and given by*

$$\Delta_k^* = \sqrt{\frac{12D}{N}}, \quad 1 \leq k \leq N, \quad (3.16)$$

*independent of unitary transform  $U$ .*

Note that the optimal  $\mathbf{\Delta}^*$  is also independent of the source statistics. Using (3.16) in (3.15), minimization of  $R(U, \mathbf{\Delta})$  subject to the distortion constraint  $D(\mathbf{\Delta}) = D$  is now equivalent to the unconstrained minimization

$$\min_{U \in \mathcal{U}} \sum_{k=1}^N h(\bar{X}_k). \quad (3.17)$$

The optimization problem in (3.17) is also formulated in [3]. Unlike the optimal  $\mathbf{\Delta}$ , the optimal  $U$  depends on the source statistics.

At this point, let us look into the well known solution of (3.17) for Gaussian  $\mathbf{X}$  (also, see [48]). Specifically, consider  $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{diag}\{\sigma^2\})$ . Hence, under transformation  $\bar{\mathbf{X}} = U\mathbf{X}$ , we have  $\bar{X}_k \sim \mathcal{N}(0, \bar{\sigma}_k^2)$ , where  $\bar{\sigma}_k^2 = \sum_{j=1}^N U_{kj}^2 \sigma_j^2$ ,  $1 \leq k \leq N$ . Hence, using the fact that, for  $X \sim \mathcal{N}(0, \sigma^2)$ , the differential entropy  $h(X) = \frac{1}{2} \log 2\pi e \sigma^2$  is strictly concave in  $\sigma^2$ , and noting  $\sum_{j=1}^N U_{kj}^2 = 1$ , we obtain

$$\sum_{k=1}^N h(\bar{X}_k) = \sum_{k=1}^N \frac{1}{2} \log 2\pi e \bar{\sigma}_k^2 \geq \sum_{k=1}^N \sum_{j=1}^N U_{kj}^2 \frac{1}{2} \log 2\pi e \sigma_j^2 = \sum_{j=1}^N \frac{1}{2} \log 2\pi e \sigma_j^2 = \sum_{j=1}^N h(X_j),$$

where the inequality follows due to Jensen's inequality [43]. The equality in the above inequality holds if and only if  $\{\bar{\sigma}_k^2\}$  is a permutation of  $\{\sigma_j^2\}$ , i.e.,  $U$  is one of the signed permutation matrices, which form the equivalent class of KLT's. Note that the strict concavity of the differential entropy in the variance has been the key in the above optimization.

In this paper, we consider GVSM vectors and identify a concavity property (Property 5.5), similar to the above, in order to derive Theorem 5.7, which states that certain KLT's are the optimal  $U$  if  $\mathbf{X}$  is GVSM. However, first we shall generalize the MSE distortion measure (3.13) to a broader quadratic framework. The optimal transform problem will then be solved in Sec. 5.

### 3.2 Transform Coding under Quadratic Criterion

Consider a design criterion of the form

$$D(U, \mathbf{\Delta}) \sim q(U, \mathbf{\Delta}; \Omega) := \sum_{k=1}^N [U\Omega U^T]_{kk} \Delta_k^2, \quad (3.18)$$

where the weights are constant. Similar quadratic criteria can be found also in [2]. More general formulations, including the cases, where the weights depend on the input or the quantized source, are studied as well; see, e.g., in [49]. At this point, we make the following assumption in order to pose the quadratic problem.

**Assumption 3.5** *The matrix  $\Omega$  is diagonal, positive definite and independent of the encoder  $(U, \mathbf{\Delta})$ .*

Of course, for the choice  $\Omega = \frac{1}{12}I$ ,  $D(U, \mathbf{\Delta})$  is independent of  $U$  and reduces to the MSE criterion (3.13). Observe that the design cost  $D(U, \mathbf{\Delta}) = \sum_{k=1}^N d_k(U, \Delta_k)$  is additive over component costs

$$d_k(U, \Delta_k) = [U\Omega U^T]_{kk} \Delta_k^2, \quad 1 \leq k \leq N. \quad (3.19)$$

The function  $q(U, \mathbf{\Delta}; \Omega)$  in (3.18) is quadratic in  $U$  as well as in each  $\Delta_k$ ,  $1 \leq k \leq N$ . Further, linearity of  $q(U, \mathbf{\Delta}; \Omega)$  in  $\Omega$  accommodates a linear combination of competing quadratic distortion criteria (see Secs. 6.2, 6.3 and 7.3).

Under Assumption 3.1, we minimize the average bit rate  $R(U, \mathbf{\Delta})$  with respect to the pair  $(U, \mathbf{\Delta})$  subject to a distortion constraint  $D(U, \mathbf{\Delta}) = D$ . Thus, by (3.15) and (3.18), we solve the Lagrangian problem

$$\min_{U, \mathbf{\Delta}} [R(U, \mathbf{\Delta}) + \rho D(U, \mathbf{\Delta})] = \min_{U, \mathbf{\Delta}} \left[ \frac{1}{N} \sum_{k=1}^N [h(\bar{X}_k) - \log \Delta_k] + \rho \sum_{k=1}^N [U\Omega U^T]_{kk} \Delta_k^2 \right], \quad (3.20)$$

where the Lagrange multiplier  $\rho$  is chosen to satisfy the design constraint

$$\sum_{k=1}^N [U\Omega U^T]_{kk} \Delta_k^2 = D. \quad (3.21)$$

Towards solving problem (3.20), we first optimize  $\mathbf{\Delta}$  keeping  $U$  constant.

**Lemma 3.6** *For a given  $U \in \mathcal{U}$ , the optimal quantizer step sizes in problem (3.20) are given by*

$$\Delta_k^*(U) = \sqrt{\frac{D}{N [U\Omega U^T]_{kk}}}, \quad 1 \leq k \leq N. \quad (3.22)$$

**Proof:** Abbreviate  $\bar{\omega}_k = [U\Omega U^T]_{kk}$ ,  $1 \leq k \leq N$ . The Lagrangian in (3.20) is of the form  $\sum_{k=1}^N \theta_k(\Delta_k; U)$ , where

$$\theta_k(\Delta_k; U) = \frac{1}{N} [h(\bar{p}_k) - \log(\Delta_k)] + \rho \bar{\omega}_k \Delta_k^2 \quad (1 \leq k \leq N) \quad (3.23)$$

is strictly convex in  $\Delta_k > 0$ . Indeed,  $\theta_k''(\Delta_k; U) = \frac{1}{N\Delta_k^2} + \rho \bar{\omega}_k > 0$ . Hence the unique minimizer of  $\theta_k(\Delta_k; U)$ , denoted  $\Delta_k^*(U)$ , solves

$$\theta_k'(\Delta_k^*(U); U) = -\frac{1}{N\Delta_k^*(U)} + 2\rho \bar{\omega}_k \Delta_k^*(U) = 0, \quad (3.24)$$

whence

$$\Delta_k^*(U) = \sqrt{\frac{1}{2N\rho\bar{\omega}_k}}, \quad 1 \leq k \leq N. \quad (3.25)$$

Now, using (3.25) in (3.21), we have

$$\sum_{k=1}^N \bar{\omega}_k \Delta_k^{*2}(U) = \frac{1}{2\rho} = D. \quad (3.26)$$

Finally, using (3.26) in (3.25), the result follows.  $\square$

Note that, unlike in classical reconstruction (where  $\Omega \propto I$ ), optimal quantizer step sizes may, in general, be different. However, for a given  $U$ , using Lemma 3.6 in (3.19), the  $k$ -th ( $1 \leq k \leq N$ ) component contributes

$$[U\Omega U^T]_{kk} \Delta_k^{*2}(U) = \frac{D}{N}$$

to the overall distortion  $D$ , i.e., the distortion is equally distributed among the components.

Next we optimize the transform  $U$ . Specifically, using (3.22) in problem (3.20), we pose the unconstrained problem:

$$\min_{U \in \mathcal{U}} \left[ \sum_{k=1}^N h(\bar{X}_k) + \frac{1}{2} \sum_{k=1}^N \log [U\Omega U^T]_{kk} \right]. \quad (3.27)$$

In the special case of classical reconstruction ( $\Omega = \frac{1}{12}I$ ), (3.27) reduces to problem (3.17).

Since  $\Omega$  is diagonal by assumption, any signed permutation matrix  $U$  is an eigenvector matrix of  $\Omega$ . Hence, by Hadamard's inequality [53], such  $U$  minimizes  $\sum_{k=1}^N \log [U\Omega U^T]_{kk}$ . Further, if any such  $U$  also minimizes  $\sum_{k=1}^N h(\bar{X}_k)$ , then it solves problem (3.27) as well. For GVSM sources, we show that such minimizing  $U$ 's indeed exist and are certain KLT's. The next section formally introduces GVSM distributions.

## 4 Gaussian Vector-Scale Mixtures

### 4.1 Definition

**Definition 4.1** A random vector  $\mathbf{X}$  taking values in  $\mathbb{R}^N$  is called GVSM if

$$\mathbf{X} \stackrel{d}{=} C^T (\mathbf{Z} \odot \mathbf{V}), \quad (4.1)$$

where  $C$  is unitary, the random vector  $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, I)$ , and  $\mathbf{V} \sim \mu$  is a random vector independent of  $\mathbf{Z}$  and taking values in  $\mathbb{R}_+^N$ . The distribution of  $\mathbf{X}$  is denoted by  $\mathcal{G}(C, \mu)$ .

In (4.1), conditioned on  $\mathbf{V} = \mathbf{v}$ , the GVSM vector  $\mathbf{X}$  is Gaussian. In fact, any Gaussian vector can be written as  $\mathbf{X} \stackrel{d}{=} C^T (\mathbf{Z} \odot \mathbf{v})$  for a suitable  $C$  and  $\mathbf{v}$ , and, therefore, is a GVSM.

**Property 4.2 (PROBABILITY DENSITY)** A GVSM vector  $\mathbf{X} \sim \mathcal{G}(C, \mu)$  has pdf

$$g(\mathbf{x}; C, \mu) := \int_{\mathbb{R}_+^N} \mu(d\mathbf{v}) \phi(\mathbf{x}; C^T \text{Diag}\{\mathbf{v}^2\} C), \quad \mathbf{x} \in \mathbb{R}^N. \quad (4.2)$$

**Proof:** Given  $\mathbf{V} = \mathbf{v}$  in (4.1),  $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, C^T \text{Diag}\{\mathbf{v}^2\} C)$ . The expression (4.2) follows by removing the conditioning on  $\mathbf{V}$ .  $\square$

In view of (4.2), we call  $\mu$  the *mixing measure* of the GVSM. Clearly, a GVSM pdf can not decay faster than any Gaussian pdf. In fact GVSM pdf's can exhibit heavy tailed behavior (see [5] for details). Next we present a bivariate ( $N = 2$ ) example.

**Example:** Consider  $\mathbf{V} \sim \mu$  with mixture pdf

$$p_{\mathbf{V}}(\mathbf{v}) = \frac{\mu(d\mathbf{v})}{d\mathbf{v}} = 4\lambda v_1 v_2 e^{-(v_1^2 + v_2^2)} + 16(1 - \lambda) v_1 v_2 e^{-2(v_1^2 + v_2^2)}, \quad \mathbf{v} \in \mathbb{R}_+^2, \quad (4.3)$$

parameterized by mixing factor  $\lambda \in [0, 1]$ . Note that the components of  $\mathbf{V}$  are independent for  $\lambda = 0$  and 1. Using (4.3) in (4.2), the GVSM vector  $\mathbf{X} \stackrel{d}{=} \mathbf{Z} \odot \mathbf{V}$  (set  $C = I$  in (4.1)) has pdf

$$p(\mathbf{x}) = \int_{\mathbb{R}_+^2} d\mathbf{v} p_{\mathbf{V}}(\mathbf{v}) \phi(\mathbf{x}; \text{Diag}\{\mathbf{v}^2\}) \quad (4.4)$$

$$= 2\lambda e^{-\sqrt{2}(|x_1| + |x_2|)} + 4(1 - \lambda) e^{-2(|x_1| + |x_2|)}, \quad \mathbf{x} \in \mathbb{R}^2. \quad (4.5)$$

See Appendix A for the derivation of (4.5). Note that  $\mathbf{X}$  is a mixture of two Laplacian vectors with independent components for  $\lambda \in (0, 1)$ .

## 4.2 Properties of Scalar GSM

A GSM vector  $\mathbf{X}$ , defined by [5]

$$\mathbf{X} \stackrel{d}{=} \mathbf{N}V \quad (4.6)$$

(where  $\mathbf{N}$  is a zero-mean  $N$ -variate Gaussian vector,  $V$  is a random variable independent of  $\mathbf{N}$  and taking values in  $\mathbb{R}_+$ ), is a GVSM. To see this, consider the special case

$$\mathbf{V} \stackrel{d}{=} V\boldsymbol{\alpha} \quad (4.7)$$

in (4.1), where  $\boldsymbol{\alpha} \in \mathbb{R}_+^N$  is a constant. Clearly, (4.1) takes the form (4.6), where

$$\mathbf{N} = C^T(\mathbf{Z} \odot \boldsymbol{\alpha}) \sim \mathcal{N}(0, C^T \text{Diag}\{\boldsymbol{\alpha}^2\}C). \quad (4.8)$$

Finally, note that arbitrary  $\mathbf{N}$  can be written as (4.8) for suitable  $\boldsymbol{\alpha}$  and  $C$ . In the scalar case ( $N = 1$ ), GVSM and GSM variables are, of course, identical. For simplicity, denote the scalar GSM pdf by  $g_1(x; \mu) = g(x; 1, \mu)$ .

Our optimality analysis crucially depends on a fundamental property of  $g_1(x; \mu)$ , which is stated below and proven in Appendix B.

**Property 4.3** *Suppose  $g_1(x; \mu)$  is a non-Gaussian GSM pdf. Then, for each pair  $(x_0, x_1)$ ,  $x_1 > x_0 > 0$ , there exists a pair  $(c, \beta)$ ,  $c > 0$ ,  $\beta > 0$ , such that  $g_1(x; \mu) = c\phi(x; \beta^2)$  at  $x \in \{\pm x_0, \pm x_1\}$ . For such  $(c, \beta)$ ,  $g_1(x) > c\phi(x; \beta^2)$ , if  $|x| < x_0$  or  $|x| > x_1$ , and  $g_1(x) < c\phi(x; \beta^2)$ , if  $x_0 < |x| < x_1$ .*

Note that if  $g_1(x; \mu) = \phi(x, \sigma^2)$  were Gaussian, we would have  $c = 1$  and  $\beta = \sigma$ , i.e.,  $c\phi(x; \beta^2)$  would coincide with  $g_1(x; \mu)$ . Further, Property 4.3 leads to the following information theoretic inequality, whose proof appears in Appendix C.

**Property 4.4** *For any  $\gamma > 0$ ,*

$$\int_{\mathbb{R}} dx \phi''(x; \gamma^2) \log g_1(x; \mu) \geq 0; \quad (4.9)$$

*equality holds if and only if  $g_1(x; \mu)$  is Gaussian.*

Property 4.4 gives rise to a key concavity property of GSM differential entropy stated in Property 5.5, which, in turn, plays a central role in the solution of the optimal transform problems (3.17) and (3.27) for a GVSM source.

## 5 Optimal Transform for GVSM Sources

We assume an arbitrary GVSM source  $\mathbf{X} \sim \mathcal{G}(C, \mu)$ . Towards obtaining the optimal transform, first we study the decorrelation properties of  $\mathbf{X}$ .

### 5.1 Primary KLT's

**Property 5.1** (*SECOND ORDER STATISTICS*) *If  $\mathbf{V} \sim \mu$  has finite expected energy  $E[\mathbf{V}^T \mathbf{V}]$ , then  $C$  decorrelates  $\mathbf{X} \sim \mathcal{G}(C, \mu)$  and  $E[\mathbf{X}^T \mathbf{X}] = E[\mathbf{V}^T \mathbf{V}]$ .*

**Proof:** By Definition 4.1,  $\mathbf{X} \stackrel{d}{=} C^T(\mathbf{Z} \odot \mathbf{V})$  where  $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, I)$  is independent of  $\mathbf{V}$ . Hence we obtain the correlation matrix of  $C\mathbf{X}$  as

$$E[C\mathbf{X}\mathbf{X}^T C^T] = E[\mathbf{Z}\mathbf{Z}^T] \odot E[\mathbf{V}\mathbf{V}^T] = I \odot E[\mathbf{V}\mathbf{V}^T] = \text{Diag}\{E[\mathbf{V}^2]\}, \quad (5.1)$$

which shows that  $C$  decorrelates  $\mathbf{X}$ . In (5.1), the first equality holds because  $C$  is unitary, and  $\mathbf{Z}$  and  $\mathbf{V}$  are independent. The second equality holds because  $E[\mathbf{Z}\mathbf{Z}^T] = I$ . Finally, the third equality holds because  $|E[\mathbf{V}\mathbf{V}^T]_{kj}| \leq E[\mathbf{V}^T \mathbf{V}] < \infty$  for any  $1 \leq k, j \leq N$ . Taking the trace in (5.1), we obtain  $E[\mathbf{X}^T \mathbf{X}] = E[\mathbf{V}^T \mathbf{V}]$ .  $\square$

By Property 5.1,  $C$  is the KLT of  $\mathbf{X} \sim \mathcal{G}(C, \mu)$ . In general, the components of the decorrelated vector  $C\mathbf{X}$  are not independent. However, the components of  $C\mathbf{X}$  are independent if the components of  $\mathbf{V} \sim \mu$  too are independent. Moreover, all decorrelating transforms are not necessarily equivalent for encoding. To see this, consider  $\mathbf{X} \sim \mathcal{G}(I, \mu)$  such that the components  $\{X_k\}$  are mutually independent, *not* identically distributed, and have equal second order moments  $E[X_k^2] = 1$ ,  $1 \leq k \leq N$ . Any  $U \in \mathcal{U}$  is a KLT because  $E[\mathbf{X}\mathbf{X}^T] = I$ . However,  $U\mathbf{X}$  has independent components only if  $U$  is a signed permutation matrix. This leads to the notion that not all KLT's of GVSM's are necessarily equivalent and to the following concept of primary KLT's of  $\mathbf{X} \sim \mathcal{G}(C, \mu)$ , a set of unitary transforms that are equivalent to  $C$  for transform coding.

**Definition 5.2** *A primary KLT of  $\mathbf{X} \sim \mathcal{G}(I, \mu)$  is any transformation  $U \in \mathcal{U}$  such that the set of marginal densities of  $U\mathbf{X}$  is, upto a permutation, identical with the set of marginal densities of  $\mathbf{X}$ . A primary KLT of  $\mathbf{X}_C \in \mathcal{G}(C, \mu)$  ( $C \neq I$ ) is any matrix  $UC$ , where  $U$  is a primary KLT of  $\mathbf{X} \sim \mathcal{G}(I, \mu)$ .*

Clearly, any signed permutation matrix is a primary KLT of  $\mathbf{X} \in \mathcal{G}(I, \mu)$  irrespective of  $\mu$ . Also, for any primary KLT  $U$ , we have  $\sum_{k=1}^N h(\bar{X}_k) = \sum_{k=1}^N h(X_k)$ .

Without loss of generality, we assume  $C = I$  and derive the transform  $U^*$  that solves problem (3.17) for a decorrelated random vector  $\mathbf{X} \sim \mathcal{G}(I, \mu)$ . Otherwise, if  $\mathbf{X} \sim \mathcal{G}(C, \mu)$ ,  $C \neq I$ , find  $\check{U}^*$  that solves problem (3.17) for the decorrelated source  $C\mathbf{X} \sim \mathcal{G}(I, \mu)$ , then obtain  $U^* = \check{U}^*C$ .

## 5.2 Finite Energy and Continuity of GVSM Density Function

At this point, let us revisit Assumption 3.1. Subject to a finite energy constraint  $E[\mathbf{V}^T\mathbf{V}] < \infty$  on  $\mathbf{V}$ , the requirement that each  $h(\bar{X}_k) < \infty$ ,  $1 \leq k \leq N$ , is met for any  $U \in \mathcal{U}$ . To see this, note, by Property 5.1, that  $E[\mathbf{X}^T\mathbf{X}] = E[\mathbf{V}^T\mathbf{V}] < \infty$ . Further, recall that subject to a finite variance constraint  $E[Y^2] = \sigma^2$  on a random variable  $Y$ ,  $h(Y)$  attains the maximum  $\frac{1}{2} \log 2\pi e \sigma^2$  [43]. Therefore, we have

$$\sum_{k=1}^N h(\bar{X}_k) \leq \sum_{k=1}^N \frac{1}{2} \log 2\pi e E[\bar{X}_k^2] \leq \frac{N}{2} \log 2\pi e E[\bar{\mathbf{X}}^T\bar{\mathbf{X}}] = \frac{N}{2} \log 2\pi e E[\mathbf{X}^T\mathbf{X}] < \infty,$$

implying each  $h(\bar{X}_k) < \infty$ .

Under Assumption 3.1, we also require the continuity of  $p(\mathbf{x}) = g(\mathbf{x}; C, \mu)$ , a necessary and sufficient condition for which is stated below and proven in Appendix D.

**Property 5.3** (*CONTINUITY OF GVSM PDF*) *A GVSM pdf  $g(\mathbf{x}; C, \mu)$  is continuous if and only if  $\mathbf{V} \sim \mu$  is such that*

$$g(\mathbf{0}; C, \mu) = (2\pi)^{-N/2} E\left[1/\prod_{k=1}^N V_k\right] < \infty.$$

In summary, the following conditions are sufficient for Assumption 3.1 to hold.

**Assumption 5.4** *The source  $\mathbf{X}$  is distributed as  $\mathcal{G}(I, \mu)$  where  $\mathbf{V} \sim \mu$  satisfies*

$$E[\mathbf{V}^T\mathbf{V}] < \infty, \tag{5.2}$$

$$E\left[1/\prod_{k=1}^N V_k\right] < \infty. \tag{5.3}$$

Clearly,  $\bar{\mathbf{X}} = U\mathbf{X} \sim \mathcal{G}(U^T, \mu)$ . Hence, by Property 4.2,

$$\bar{p}(\bar{\mathbf{x}}) = g(\bar{\mathbf{x}}; U^T, \mu) = \int_{\mathbb{R}_+^N} \mu(d\mathbf{v}) \phi(\bar{\mathbf{x}}; U \text{Diag}\{\mathbf{v}^2\} U^T). \quad (5.4)$$

Of course, setting  $U = I$  in (5.4), we obtain  $p(\mathbf{x})$ . Now, marginalizing  $\bar{p}(\bar{\mathbf{x}})$  in (5.4) to  $\bar{x}_k$  ( $1 \leq k \leq N$ ), we obtain

$$\bar{p}_k(\bar{x}_k) = \int_{\mathbb{R}_+^N} \mu(d\mathbf{v}) \phi(\bar{x}_k; \bar{v}_k^2) = \int_{\mathbb{R}_+^N} \bar{\mu}(d\bar{\mathbf{v}}) \phi(\bar{x}_k; \bar{v}_k^2) = \int_{\mathbb{R}_+} \bar{\mu}_k(d\bar{v}_k) \phi(\bar{x}_k; \bar{v}_k^2), \quad (5.5)$$

where  $\bar{\mathbf{v}} = \sqrt{(U \odot U)\mathbf{v}^2}$ , the measure  $\bar{\mu}$  is specified by  $\bar{\mu}(d\bar{\mathbf{v}}) = \mu(d\mathbf{v})$ , and  $\bar{\mu}_k$  denotes the  $k$ -th marginal of  $\bar{\mu}$ . In representation (5.5) of  $\bar{p}_k$ , note that  $\bar{\mu}_k$  depends on  $U$  in a complicated manner. To facilitate our analysis, we now make a slight modification to the GVSM representation (4.2) such that the modified representation of  $\bar{p}_k$  makes the dependence on  $U$  explicit.

### 5.3 Modified GVSM Representation

Suppose

$$\mathbf{V} \stackrel{d}{=} \boldsymbol{\sigma}(\mathbf{W}) \sim \mu \quad (5.6)$$

for some measurable function  $\boldsymbol{\sigma} : \mathbb{R}_+^M \rightarrow \mathbb{R}_+^N$  and some random vector  $\mathbf{W} \sim \nu$  taking values in  $\mathbb{R}_+^M$ . Using (5.6), rewrite the GVSM pdf  $g(\mathbf{x}; C, \mu) = \int_{\mathbb{R}_+^N} \mu(d\mathbf{v}) \phi(\mathbf{x}; C^T \text{Diag}\{\mathbf{v}^2\} C)$  given in (4.2) in the modified representation

$$g(\mathbf{x}; C, \nu, \boldsymbol{\sigma}^2) := \int_{\mathbb{R}_+^M} \nu(d\mathbf{w}) \phi(\bar{\mathbf{x}}; C^T \Sigma(\mathbf{w}) C) \quad (5.7)$$

(which really is a slight abuse of notation), where  $\Sigma(\mathbf{w}) = \text{Diag}\{\boldsymbol{\sigma}^2(\mathbf{w})\}$ . Clearly, any pdf takes the form  $g(\mathbf{x}; C, \nu, \boldsymbol{\sigma}^2)$  if and only if it is GVSM (although for a given  $\nu$  and  $C$ , multiple  $\boldsymbol{\sigma}^2$ 's may give the same pdf  $g(\mathbf{x}; C, \nu, \boldsymbol{\sigma}^2)$ ). In view of (5.7), we call  $\nu$  the *modified mixing measure* and  $\boldsymbol{\sigma}^2$  the *variance function*.

Now, supposing (5.6), write  $\bar{p}(\bar{\mathbf{x}})$  in (5.4) in the modified form

$$\bar{p}(\bar{\mathbf{x}}) = g(\bar{\mathbf{x}}; U^T, \nu, \boldsymbol{\sigma}^2) := \int_{\mathbb{R}_+^M} \nu(d\mathbf{w}) \phi(\bar{\mathbf{x}}; U \Sigma(\mathbf{w}) U^T). \quad (5.8)$$

Note

$$[U \Sigma(\mathbf{w}) U^T]_{kk} = \sum_{j=1}^N U_{kj}^2 \sigma_j^2(\mathbf{w}), \quad 1 \leq k \leq N.$$

Hence, marginalizing (5.8) to  $\bar{x}_k$  ( $1 \leq k \leq N$ ), and denoting  $g_1(\cdot; \nu, \sigma^2) = g(\cdot; 1, \nu, \sigma^2)$ , we obtain

$$\bar{p}_k(\bar{x}_k) = g_1(\bar{x}_k; \nu, \bar{\sigma}_k^2), \quad (5.9)$$

where

$$\bar{\sigma}_k^2 = \sum_{j=1}^N U_{kj}^2 \sigma_j^2. \quad (5.10)$$

Given  $\sigma^2$ , note that the variance function  $\bar{\sigma}_k^2$  is a simple function of  $U$ , and the modified mixing measure  $\nu$  does not depend on  $U$ .

## 5.4 Optimality under MSE Criterion

In this section, we present the optimal transform  $U$  solving problem (3.17). First we need a key concavity property of the differential entropy  $h(g_1(\cdot; \nu, \sigma^2))$ , which is stated below and proven in Appendix E.

**Property 5.5** (*STRICT CONCAVITY IN VARIANCE FUNCTION*) *For a given modified mixing measure  $\nu$ , suppose  $\mathcal{V}(\nu)$  is a convex collection of scalar variance functions such that,*

(i) *for any  $\sigma^2 \in \mathcal{V}(\nu)$ ,  $g_1(x; \nu, \sigma^2)$  is continuous and  $-\infty < h(g_1(\cdot; \nu, \sigma^2)) < \infty$ ;*

(ii) *for any pair  $(\sigma'^2, \sigma^2) \in \mathcal{V}^2(\nu)$ ,  $D(g_1(\cdot; \nu, \sigma'^2) \| g_1(\cdot; \nu, \sigma^2)) < \infty$ .*

*Then  $h(g_1(\cdot; \nu, \sigma^2))$  is strictly concave in  $\sigma^2 \in \mathcal{V}(\nu)$ . If there exists a convex equivalence class of some  $\sigma^2 \in \mathcal{V}(\nu)$  uniquely determining the pdf  $g_1(x; \nu, \sigma^2)$ , then the strict concavity at such  $\sigma^2$  holds only upto the equivalence class.*

We do not know whether nontrivial equivalence classes of this type exist. However, their potential existence has no impact on our analysis.

In further analysis, use the identity variance function  $\sigma^2(\mathbf{w}) = \mathbf{w}^2$ , i.e.,  $\nu = \mu$  in (5.9). First, generate the collection  $\mathcal{Q}(\mu)$  of all possible pdf's  $\bar{p}_k$  by varying  $U$  through all unitary matrices. Note that  $\mathcal{Q}(\mu)$  is independent of  $k$ . By Assumption 5.4,  $h(q)$  is finite for any  $q \in \mathcal{Q}(\mu)$ . In order to apply Property 5.5 to problem (3.17), we also make the following technical assumption.

**Assumption 5.6** *For any pair of pdf's  $(q', q) \in \mathcal{Q}^2(\mu)$ ,  $D(q' \| q) < \infty$ .*

**Theorem 5.7** *Under Assumptions 5.4 and 5.6, and any transformation  $U \in \mathcal{U}$ ,*

$$\sum_{k=1}^N h(\bar{X}_k) \geq \sum_{k=1}^N h(X_k); \quad (5.11)$$

*equality holds if and only if  $U$  is a primary KLT.*

**Proof:** Setting  $\nu = \mu$ , and varying  $U$  through unitary matrices in (5.10), generate the collection  $\mathcal{V}(\mu)$  of  $\bar{\sigma}_k^2 = \sum_{j=1}^N U_{kj}^2 \sigma_j^2$ . Since  $\sum_{j=1}^N U_{kj}^2 = 1$ ,  $\mathcal{V}(\mu)$  is convex. Further, for such  $\mathcal{V}(\mu)$ , the conditions (i) and (ii) in Property 5.5 hold by assumption. Further, by (5.9), we obtain

$$\sum_{k=1}^N h(\bar{p}_k) = \sum_{k=1}^N h \left( g_1 \left( \cdot; \mu, \sum_{j=1}^N U_{kj}^2 \sigma_j^2 \right) \right) \quad (5.12)$$

$$\geq \sum_{k=1}^N \sum_{j=1}^N U_{kj}^2 h(g_1(\cdot; \mu, \sigma_j^2)), \quad (5.13)$$

where the inequality follows by noting  $\sum_{j=1}^N U_{kj}^2 = 1$  and strict concavity of  $h(g_1(\cdot; \mu, \sigma^2))$  in  $\sigma^2$  (Property 5.5), and applying Jensen's inequality [43]. Further, noting such strict concavity in  $\sigma^2$  holds upto the equivalence class uniquely determining  $g_1(\cdot; \mu, \sigma^2)$ , the inequality in (5.13) is an equality if and only if the pdf collection  $\{\bar{p}_k = g_1(\cdot; \mu, \sum_{j=1}^N U_{kj}^2 \sigma_j^2)\}$  is a permutation of the pdf collection  $\{p_j = g_1(\cdot; \mu, \sigma_j^2)\}$ , i.e.,  $U$  is a primary KLT. Finally, noting  $\sum_{k=1}^N U_{kj}^2 = 1$  in (5.13), the result follows.  $\square$

Theorem 5.7 can equivalently be stated as: Any primary KLT  $U^*$  of a GVSM source  $\mathbf{X} \sim \mathcal{G}(I, \mu)$  solves the optimal transform problem (3.17) under the MSE criterion. In the special case when  $\mathbf{X}$  is Gaussian, Assumptions 5.4 and 5.6 hold automatically, and any KLT is a primary KLT. Thus Theorem 5.7 extends the transform coding theory of Gaussian sources to the broader class of GVSM sources. Next, we solve optimal transform problem (3.27) under the broader class of quadratic criteria.

## 5.5 Optimality under Quadratic Criteria

In this section, we shall assume that Assumptions 5.4 and 5.6 hold. In the last paragraph of Sec. 3, we noted that any signed permutation matrix  $U$  is an eigenvector matrix of the diagonal matrix  $\Omega$ , and, by Hadamard's inequality [53], minimizes  $\sum_{k=1}^N \log [U\Omega U^T]_{kk}$ . Referring to Definition 5.2, note that any such  $U$  is also a primary KLT of  $\mathbf{X}$ , and, by Theorem 5.7, minimizes  $\sum_{k=1}^N h(\bar{X}_k)$ . Thus we obtain:

**Lemma 5.8** Any eigenvector matrix  $U$  of  $\Omega$  which is a primary KLT of source  $\mathbf{X} \sim \mathcal{G}(\mu, I)$  solves problem (3.27). The set of solutions includes all signed permutation matrices.

Without loss of generality, choose  $U^* = I$  as the optimal transform. Setting  $U = I$  in (3.6), we obtain optimal quantizer step sizes  $\{\Delta_k^*\}$ . We also obtain optimal bit allocation  $\{R_k^*\}$  and optimal average rate  $R^*$  by setting  $U = I$  in (3.14) and (3.15), respectively.

**Theorem 5.9** Subject to a design constraint  $D(U, \Delta) = D$ , the optimal transform coding problem (3.20) is solved by the pair  $(I, \Delta^*)$ , where optimal quantizer step sizes  $\{\Delta_k^*\}$  are given by

$$\Delta_k^* = \left[ \frac{D}{N\Omega_{kk}} \right]^{\frac{1}{2}}, \quad 1 \leq k \leq N. \quad (5.14)$$

The corresponding optimal bit allocation strategy  $\{R_k^* = R_k(I, \Delta_k^*)\}$  and the optimal average bit rate  $R^* = R(I, \Delta^*)$  are, respectively, given by

$$R_k^* = \left( R^* - \frac{1}{N} \sum_{j=1}^N \left[ h(X_j) + \frac{1}{2} \log \Omega_{jj} \right] \right) + h(X_k) + \frac{1}{2} \log \Omega_{kk}, \quad 1 \leq k \leq N, \quad (5.15)$$

$$R^* = \frac{1}{N} \sum_{k=1}^N \left[ h(X_k) - \frac{1}{2} \log \frac{D}{N\Omega_{kk}} \right]. \quad (5.16)$$

Theorem 5.9 extends the Gaussian transform coding theory in two ways:

1. The source generalizes from Gaussian to GVSM;
2. The performance measure generalizes from the MSE to a broader class of quadratic criteria.

In fact, in the special case of classical reconstruction ( $\Omega = \frac{1}{12}I$ ) of a Gaussian vector  $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \text{Diag}\{\sigma^2\})$  (recall  $h(p_k) = \frac{1}{2} \log 2\pi e \sigma_k^2$ ,  $1 \leq k \leq N$  [43]), the expressions (5.15) and (5.16), respectively, reduce to

$$R_k^* = R^* + \frac{1}{2} \log \frac{\sigma_k^2}{\left[ \prod_{j=1}^N \sigma_j^2 \right]^{\frac{1}{N}}}, \quad 1 \leq k \leq N, \quad (5.17)$$

$$R^* = \frac{1}{2} \log \frac{\pi e N}{6D} + \frac{1}{2N} \log \prod_{k=1}^N \sigma_k^2, \quad (5.18)$$

which gives the classical bit allocation strategy [2].

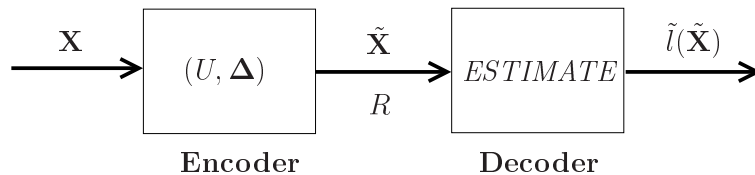


Figure 3: Estimation at the decoder.

At this point, we relax the requirement in Assumption 3.5 that  $\Omega$  be positive definite, and allow nonnegative definite  $\Omega$ 's. We only require  $\Omega$  to be diagonal. Henceforth, we consider this relaxed version instead of the original Assumption 3.5. Denote by  $\mathcal{K} = \{k : \Omega_{kk} \neq 0, 1 \leq k \leq N\}$ , the set of indices of relevant components. Note the dimensionality of the problem reduces from  $N$  to the number of relevant components  $|\mathcal{K}|$ . For optimal encoding of  $\mathbf{X}$ , choose the transformation  $U^* = I$ , discard the irrelevant components  $\bar{X}_k$  (i.e., set  $R_k^* = 0$ ),  $k \notin \mathcal{K}$ , and for each  $k \in \mathcal{K}$ , uniformly quantize  $\tilde{X}_k = Q(\bar{X}_k; \Delta_k^*)$  using the optimal step size  $\Delta_k^* = \left[ \frac{D}{|\mathcal{K}| \Omega_{kk}} \right]^{\frac{1}{2}}$ . Accordingly, compute  $R^*$  and  $R_k^*$ ,  $k \in \mathcal{K}$ , replacing  $N$  by  $|\mathcal{K}|$  and taking summations over  $k \in \mathcal{K}$  only instead of  $1 \leq k \leq N$  in the expressions (5.16) and (5.15).

In the rest of our paper, we present estimation/classification applications of the theory. Specifically, we identify sufficient conditions for our result to apply. We also illustrate salient aspects of the theory with examples.

## 6 Estimation from Rate-Constrained Data

Consider the estimation of a random variable  $Y$  from data  $\mathbf{X} \sim \mathcal{P}$ . Denote the corresponding estimate by  $l(\mathbf{X})$ . However, as shown in Fig. 3, suppose  $\mathbf{X}$  is transform coded to  $\tilde{\mathbf{X}} = t(\mathbf{X}) = Q(U\mathbf{X}; \Delta) \sim \tilde{\mathcal{P}}$  as in (3.11), and the estimator does not have access to  $\mathbf{X}$ . In a slight abuse of notation, denote  $t = (U, \Delta)$ . In absence of  $\mathbf{X}$ , the decoder constructs an estimator  $\tilde{l}(\tilde{\mathbf{X}})$  of  $Y$  based on  $\tilde{\mathbf{X}}$ . Our goal is to design  $(U, \Delta)$  minimizing average bit rate  $R(U, \Delta)$  required for entropy coding of  $\tilde{\mathbf{X}}$  under an estimation performance constraint  $D(U, \Delta) = D$ . We demonstrate the applicability of our result to this design.

### 6.1 Generalized $f$ -Divergence Measure

We consider a broad class of estimation scenarios where generalized  $f$ -divergence measures estimation performance. Take the example of MMSE estimation. Recall that the MMSE

estimator of random variable  $Y$  based on  $\mathbf{X} \sim \mathcal{P}$  is given by the conditional expectation [52]

$$l(\mathbf{X}) = \mathbb{E}[Y|\mathbf{X}]$$

and achieves

$$\text{MMSE} = \mathbb{E}[Y^2] - \mathbb{E}[l^2(\mathbf{X})]. \quad (6.1)$$

Referring to Definition 2.1, here  $\mathbb{E}[l^2(\mathbf{X})] = D_f(l, \mathcal{P})$  is the generalized  $f$ -divergence with convex  $f \circ l = l^2$ . The MMSE estimator of  $Y$  based on compressed observation  $\tilde{\mathbf{X}} = t(\mathbf{X})$  is given by

$$\tilde{l}(\tilde{\mathbf{X}}) = \mathbb{E}[Y|\tilde{\mathbf{X}}] = \mathbb{E}[\mathbb{E}[Y|\mathbf{X}]|t(\mathbf{X})] = \mathbb{E}[l(\mathbf{X})|t(\mathbf{X})],$$

and the corresponding MMSE is given by (6.1) with  $l(\mathbf{X})$  replaced by  $\tilde{l}(\tilde{\mathbf{X}})$ . Therefore, the MMSE increase due to transform encoding  $t = (U, \Delta)$  is given by the divergence loss

$$\Delta D_f((U, \Delta); l, \mathcal{P}) = \mathbb{E}[l^2(\mathbf{X})] - \mathbb{E}[\tilde{l}^2(\tilde{\mathbf{X}})].$$

More generally, the average estimation error incurred by any estimator  $l(\mathbf{X})$  of  $Y$  based on the data  $\mathbf{X}$  is given by  $\mathbb{E}[d(Y, l(\mathbf{X}))]$ , where  $d$  is a distortion metric. Similarly, the corresponding estimator  $\tilde{l}(\tilde{\mathbf{X}}) = \mathbb{E}[l(\mathbf{X})|\tilde{\mathbf{X}}]$  based on the compressed data  $\tilde{\mathbf{X}}$  incurs an average estimation error  $\mathbb{E}[d(Y, \tilde{l}(\tilde{\mathbf{X}}))]$ . We use as our design criterion the increase in estimation error due to transform coding

$$D(U, \Delta) = \mathbb{E}[d(Y, \tilde{l}(\tilde{\mathbf{X}}))] - \mathbb{E}[d(Y, l(\mathbf{X}))],$$

which is given by the divergence loss  $\Delta D_f((U, \Delta); l, \mathcal{P})$  for problems satisfying the following assumption.

**Assumption 6.1** *Estimation of random variable  $Y$  is associated with a triplet  $(f, l, \mathcal{P})$  such that the following hold:*

- i)  $\mathbb{E}[d(Y, l(\mathbf{X}))] = c - D_f(l, \mathcal{P})$  ( $\mathbf{X} \sim \mathcal{P}$ ) for some constant  $c$  independent of  $(f, l, \mathcal{P})$ ;*
- ii)  $\mathbb{E}[d(Y, \tilde{l}(\tilde{\mathbf{X}}))] = c - D_f(\tilde{l}, \tilde{\mathcal{P}})$  ( $\tilde{\mathbf{X}} \sim \tilde{\mathcal{P}}$ );*
- iii) Regularity condition  $\mathcal{R}(f, l, \mathcal{P})$  (Definition 2.2) holds.*

Clearly, a limited set of distortion metrics  $d$  is admissible. In the special case of MMSE estimation,  $d$  is the squared error metric and we have  $c = \mathbb{E}[Y^2]$  according to (6.1).

## 6.2 Estimation Regret Matrix

Next we obtain a high resolution approximation of  $D(U, \Delta) = \Delta D_f((U, \Delta); l, \mathcal{P})$  using Poor's Theorem 2.3. The function  $q$  is defined in (3.18).

**Lemma 6.2** *Subject to regularity condition  $\mathcal{R}(f, l, \mathcal{P})$ ,*

$$D(U, \Delta) = \Delta D_f((U, \Delta); l, \mathcal{P}) \sim q(U, \Delta; \Gamma(f, l, \mathcal{P})) = \sum_{k=1}^N [U \Gamma(f, l, \mathcal{P}) U^T]_{kk} \Delta_k^2, \quad (6.2)$$

where

$$\Gamma(f, l, \mathcal{P}) := \frac{1}{24} E_{\mathcal{P}} [(\nabla l \nabla^T l) f'' \circ l] \quad (6.3)$$

is independent of the pair  $(U, \Delta)$ .

The proof appears in Appendix F. We call  $\Gamma(f, l, \mathcal{P})$  the *estimation regret matrix*, because it determines the performance loss in estimating  $Y$  due to transform coding of  $\mathbf{X} \sim \mathcal{P}$ . In the special case, where  $f \circ l = l^2$  and  $l = p$  (Lebesgue density of  $\mathcal{P}$ ),  $\Gamma(f, p, p)$  gives a scaled version of the *Fisher information matrix* for location parameters [52].

The design criterion  $D(U, \Delta)$  given by (6.2) is of the form (3.18). Consequently, we pose the optimal transform coding problem for estimating  $Y$  by setting  $\Omega = \Gamma(f, l, \mathcal{P})$  in problem (3.20). To proceed, we need certain properties of  $\Gamma$ , which are given in the following and proven in Appendix G.

**Lemma 6.3** *The matrix  $\Gamma(f, l, \mathcal{P})$  is nonnegative definite. In addition, if both  $p(\mathbf{x})$  (Lebesgue density of  $\mathcal{P}$ ) and  $l(\mathbf{x})$  are symmetric in each  $x_k$  about  $x_k = 0$ ,  $1 \leq k \leq N$ , then  $\Gamma$  is diagonal.*

**Assumption 6.4** *The data  $\mathbf{X}$  follow GVSM distribution  $\mathcal{P} = \mathcal{G}(I, \mu)$  such that  $\mathbf{V} \sim \mu$  satisfies (5.2) and (5.3), and Assumption 5.6 holds for  $\mathcal{Q}(\mu)$ . Further,  $l(\mathbf{x})$  is symmetric in each  $x_k$  about  $x_k = 0$ ,  $1 \leq k \leq N$ .*

Of course, Assumption 5.4 holds. Noting the GVSM pdf  $p(\mathbf{x})$  is symmetric, by Lemma 6.3, the second condition implies  $\Gamma(f, l, \mathcal{P})$  is diagonal. Thus the relaxed version of Assumption 3.5 holds for  $\Omega = \Gamma(f, l, \mathcal{P})$  (see the penultimate paragraph of Sec. 5.5).

Estimating a random vector  $\mathbf{Y}$  based on data  $\mathbf{X}$  amounts to constructing an estimator  $l_k(\mathbf{X})$  of each component  $Y_k$  ( $1 \leq k \leq N$ ) separately. In other words, there are  $N$  scalar

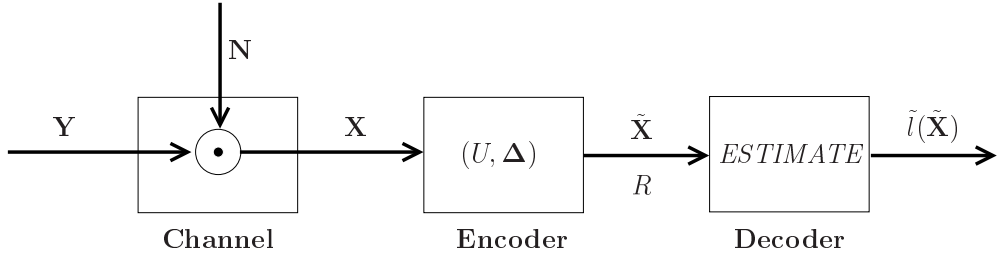


Figure 4: Signal estimation in multiplicative noise.

estimation problems associated with the triplets  $(f, l_k, \mathcal{P})$ ,  $1 \leq k \leq N$ . By (6.2), the overall increase in estimation error is given by

$$D(U, \Delta) = \sum_{k=1}^N \Delta D_f((U, \Delta); l_k, \mathcal{P}) \sim \sum_{k=1}^N q(U, \Delta; \Gamma(f, l_k, \mathcal{P})). \quad (6.4)$$

Referring to (3.18), recall that  $q(U, \Delta; \Omega) = \sum_{k=1}^N [U \Omega U^T]_{kk} \Delta_k^2$  is linear in  $\Omega$ . Hence the optimal transform coding problem can be formulated by setting

$$\Omega = \sum_{k=1}^N \Gamma(f, l_k, \mathcal{P})$$

in (3.20).

Next we present a case study of signal estimation in multiplicative noise, where GVSM vectors are shown to arise naturally.

### 6.3 Multiplicative Noise Channel

Consider the estimation system depicted in Fig. 4, where the signal vector  $\mathbf{Y}$  taking values in  $\mathbb{R}_+^N$  is componentwise multiplied by independent noise vector  $\mathbf{N}$  to generate the vector

$$\mathbf{X} = \mathbf{Y} \odot \mathbf{N} \sim \mathcal{P}. \quad (6.5)$$

Consider MMSE estimation of  $\mathbf{Y}$ . Ideally, an MMSE estimator  $l_k(\mathbf{X}) = \mathbb{E}[Y_k | \mathbf{X}]$  of each  $Y_k$  ( $1 \leq k \leq N$ ) is obtained based on  $\mathbf{X}$ . Instead consider transform encoding  $\tilde{\mathbf{X}} = Q(U\mathbf{X}; \Delta)$  of  $\mathbf{X}$  to an average bit rate  $R(U, \Delta)$ , and obtain the corresponding estimator  $\tilde{l}_k(\tilde{\mathbf{X}}) = \mathbb{E}[Y_k | \tilde{\mathbf{X}}]$  of each  $Y_k$  based on  $\tilde{\mathbf{X}}$ . For each  $k$  ( $1 \leq k \leq N$ ), recall that Assumption 6.1 holds subject to regularity conditions  $\mathcal{R}(f, l_k, \mathcal{P})$ . Also recall that the increase in MMSE in the  $k$ -th component due to transform coding is given by  $\Delta D_f((U, \Delta); l_k, \mathcal{P})$  where  $f \circ l_k = l_k^2$  and, by (6.4), the overall increase in MMSE is given by  $D(U, \Delta) \sim q(U, \Delta; \sum_{k=1}^N \Gamma(f, l_k, \mathcal{P}))$ . Subject

to a design constraint  $D(U, \Delta) = D$ , the optimal transform encoder  $(U, \Delta)$  is designed minimizing  $R(U, \Delta)$  in (3.15).

If the noise vector  $\mathbf{N}$  is a decorrelated GVSM, then the observed vector  $\mathbf{X}$  is also a decorrelated GVSM irrespective of the distribution  $\mu$  of the signal vector  $\mathbf{Y}$ . Thus our design potentially applies to the corresponding transform coding problem. To see that  $\mathbf{X}$  is indeed GVSM, write  $\mathbf{N} \stackrel{d}{=} \mathbf{V} \odot \mathbf{Z}$ , where  $\mathbf{V}$  takes values in  $\mathbb{R}_+^N$  and  $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, I)$  is independent of  $\mathbf{V}$  (Definition 4.1). Hence we can rewrite (6.5) as

$$\mathbf{X} \stackrel{d}{=} \mathbf{Y} \odot (\mathbf{V} \odot \mathbf{Z}) = (\mathbf{Y} \odot \mathbf{V}) \odot \mathbf{Z},$$

where  $\mathbf{Y} \odot \mathbf{V}$  takes values in  $\mathbb{R}_+^N$  and is independent of  $\mathbf{Z}$ .

For the sake of simplicity, suppose the components of the noise vector  $\mathbf{N} \stackrel{d}{=} \mathbf{Z} \sim \mathcal{N}(\mathbf{0}, I)$  are zero-mean *i.i.d.* Gaussian with variance one. Thus the observed vector  $\mathbf{X}$  in (6.5) takes the form

$$\mathbf{X} \stackrel{d}{=} \mathbf{Y} \odot \mathbf{Z} \sim \mathcal{P} = \mathcal{G}(\mu, I), \quad (6.6)$$

where  $\mathbf{Y} \sim \mu$ . Hence, if  $\mathbf{V} \sim \mu$  satisfies (5.2) and (5.3), then the first requirement in Assumption 6.4 is satisfied. Further, from (6.6), the  $k$ -th ( $1 \leq k \leq N$ ) MMSE estimate  $l_k$  can be obtained as

$$l_k(\mathbf{x}) = \mathbb{E}[Y_k | \mathbf{X} = \mathbf{x}] = \frac{\int_{\mathbb{R}_+^N} \mu(d\mathbf{y}) y_k \phi(\mathbf{x}; \text{Diag}\{\mathbf{y}^2\})}{\int_{\mathbb{R}_+^N} \mu(d\mathbf{y}) \phi(\mathbf{x}; \text{Diag}\{\mathbf{y}^2\})}, \quad (6.7)$$

which exhibits symmetry in each  $x_k$  ( $1 \leq k \leq N$ ) about  $x_k = 0$ . This satisfies the second requirement in Assumption 6.4.

Lastly, we give a flavor of the regularity condition  $\mathcal{R}(f, l_k, \mathcal{P})$ . We fix  $k = 1$ . Referring to Definition 2.2, note  $f \circ l_1 = l_1^2$  is convex and  $f'' = 2$  is continuous in the range of  $l_1$ . Hence

$$d\nu = f'' \circ l_1 d\mathcal{P} = 2 d\mathcal{P}. \quad (6.8)$$

Consequently, condition a) requires continuous differentiability of  $p$ . Now, let us revisit the bivariate ( $N = 2$ ) example given in Sec. 4, where  $\mathbf{V}$  played the role of the signal vector  $\mathbf{Y}$ . Reproducing  $p(\mathbf{x})$  from (4.5) for  $\lambda = \frac{1}{2}$ , note that

$$p(\mathbf{x}) = e^{-\sqrt{2}(|x_1|+|x_2|)} + 2e^{-2(|x_1|+|x_2|)}$$

exhibits discontinuity in its gradient at  $\mathbf{x} = \mathbf{0}$ . Hence  $\mathcal{R}(f, l_1, \mathcal{P})$  fails to hold. However, there exists distribution  $\mu$  of  $\mathbf{Y}$  such that the regularity condition  $\mathcal{R}(f, l_1, \mathcal{P})$  holds. Specifically, consider the polynomial (heavy) tailed pdf

$$\frac{\mu(d\mathbf{y})}{d\mathbf{y}} = \frac{25/2}{y_1^6 y_2^6} + \frac{8}{y_1^5 y_2^5}, \quad \mathbf{y} \in [1, \infty)^2. \quad (6.9)$$

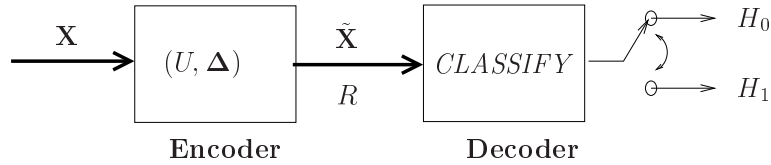


Figure 5: Binary classification at the decoder.

Although the corresponding  $p(\mathbf{x})$  does not admit a closed form expression, verification of  $\mathcal{R}(f, l_1, \mathcal{P})$  is not difficult and is left to the reader.

## 7 Binary Classification from Rate-Constrained Data

Finally, we apply our result to binary classification. Consider a test between binary hypotheses:

$$H_0 : \mathbf{X} \sim \mathcal{P}^0 \text{ versus } H_1 : \mathbf{X} \sim \mathcal{P}^1,$$

in a Bayesian framework with  $\Pr[H_i] = \pi_i$ ,  $i = 0, 1$ . Assume the computational resources in the encoder are insufficient for implementing the optimal likelihood ratio test<sup>1</sup>, but, as depicted in Fig. 5, are sufficient for constructing  $\tilde{\mathbf{X}} = t(\mathbf{X}) = Q(U\mathbf{X}; \Delta)$  (see (3.11)). The decoder, in the absence of  $\mathbf{X}$ , decides between  $H_1$  and  $H_0$  based on  $\tilde{\mathbf{X}}$ . Next we optimize the encoder  $(U, \Delta)$  subject to an average bit rate constraint  $R$ .

### 7.1 Regular $f$ -divergence

The usual  $f$ -divergence  $D_f(l, \mathcal{P}^0)$  ( $l$  indicating the likelihood ratio) arises naturally in binary classification [52]. In particular, the minimum probability of error is achieved by a likelihood ratio test and is given by

$$P_e(l, \mathcal{P}^0) = \mathbb{E}_{\mathcal{P}^0} [\min\{\pi_0, \pi_1 l\}] = 1 - D_f(l, \mathcal{P}^0), \quad (7.1)$$

where

$$f \circ l = 1 - \min\{\pi_0, \pi_1 l\} \quad (7.2)$$

is continuous and convex in  $l$ . Recall that the minimum probability of error based on compressed vector  $\tilde{\mathbf{X}} = t(\mathbf{X})$  is given by

$$P_e(\tilde{l}, \tilde{\mathcal{P}}^0) = 1 - D_f(\tilde{l}, \tilde{\mathcal{P}}^0), \quad (7.3)$$

---

<sup>1</sup>Given adequate resources, the encoder would detect the class index based on  $\mathbf{X}$  and transmit the decision.

where  $\tilde{\mathbf{X}} \sim \tilde{\mathcal{P}}^0$  under  $H_0$  and  $\tilde{l}(\tilde{\mathbf{x}}) = \mathbb{E}[l(\mathbf{X})|t(\mathbf{X}) = \tilde{\mathbf{x}}]$  is the likelihood ratio of  $\tilde{\mathbf{X}}$  at  $\tilde{\mathbf{x}}$ . Subtracting (7.1) from (7.3), the increase in probability error is given by the divergence loss  $\Delta D_f(t; l, \mathcal{P}^0)$  as defined in (2.6).

However, the function  $f$  in (7.2) is not smooth and violates the regularity condition  $\mathcal{R}(f, l, \mathcal{P}^0)$ . As a result, Theorem 2.3, does not apply. Our approach is to use  $f$ -divergences  $D_f(l, \mathcal{P}^0)$  such that  $\mathcal{R}(f, l, \mathcal{P}^0)$  holds, and adopt the divergence loss  $\Delta D_f(t; l, \mathcal{P}^0)$  as the design cost. Fortunately, this regularity requirement is not overly restrictive. For instance, upper as well as lower bounds on  $P_e$  can be obtained based on regular  $f$ -divergences. Specifically, for a certain sequence  $\{f_\alpha\}$ , such a bound can be made arbitrarily tight as  $\alpha \rightarrow \infty$  [51]. Here the regularity condition  $\mathcal{R}(f_\alpha, l, \mathcal{P}^0)$  holds for any finite  $\alpha$ . Also of particular interest is the Chernoff (upper) bound on  $P_e(l, \mathcal{P}^0)$ . This bound is the negative of a regular  $f$ -divergence such that

$$f \circ l = -\pi_0^{1-s} \pi_1^s l^s, \quad (7.4)$$

where  $s \in (0, 1)$  is the Chernoff exponent. Under certain conditions, there exists  $s \in (0, 1)$  such that the Chernoff bound is asymptotically tight as  $N \rightarrow \infty$  [52].

## 7.2 Common KLT under Both Hypotheses

Now consider the divergence loss  $\Delta D_f((U, \mathbf{\Delta}); l, \mathcal{P}^0)$  due to transform encoding  $t = (U, \mathbf{\Delta})$ . Replacing  $(f, l, \mathcal{P})$  by the triplet  $(f, l, \mathcal{P}^0)$  at hand in Lemma 6.2, the high-resolution approximation is given by

$$\Delta D_f((U, \mathbf{\Delta}); l, \mathcal{P}^0) \sim q(U, \mathbf{\Delta}; \Gamma) = \sum_{k=1}^N [U\Gamma U^T]_{kk} \Delta_k^2,$$

where  $\Gamma(f, l, \mathcal{P}^0)$  takes the form (6.3). We call this  $\Gamma$  the *discriminability regret* matrix as it determines the divergence loss between  $p_0$  and  $p_1$  due to transform coding. Take the example of the negative of Chernoff bound. Specifically, using  $f \circ l = -\pi_0^{1-s} \pi_1^s l^s$  in (6.3), we obtain

$$\Gamma = \frac{s(1-s)}{24} \pi_0^{1-s} \pi_1^s \mathbb{E}_0 [(\nabla l(\mathbf{X}) \nabla^T l(\mathbf{X})) l^{s-2}(\mathbf{X})], \quad s \in (0, 1). \quad (7.5)$$

Equivalently, we can consider the Chernoff distance  $-\log \mathbb{E}_0 [l^s(\mathbf{X})]$ , which gives the negative exponent in Chernoff bound. By straightforward computation, the loss in Chernoff distance also takes the quadratic form  $q(U, \mathbf{\Delta}; \Gamma)$  (see (3.18) for an expression of  $q$ ) [37], where

$$\Gamma = \frac{s(1-s)}{24} \frac{\mathbb{E}_0 [(\nabla l(\mathbf{X}) \nabla^T l(\mathbf{X})) l^{s-2}(\mathbf{X})]}{\mathbb{E}_0 [l^s(\mathbf{X})]}. \quad (7.6)$$

Note that the  $\Gamma$  in (7.5) scales to the  $\Gamma$  in (7.6), which of course does not take the form (6.3) exactly. More importantly, the latter  $\Gamma$  still conveys the same discriminability regret information between  $p^0$  and  $p^1$  as the former  $\Gamma$ . In the special case where the conditional distributions are Gaussian,  $\mathcal{P}^i = \mathcal{N}(\mathbf{0}, \Sigma^i)$ , with diagonal covariance matrices  $\Sigma^i = \text{Diag}\{(\boldsymbol{\sigma}^i)^2\}$ ,  $i = 0, 1$ , the matrix  $\Gamma$  in (7.6) is also diagonal with diagonal entries

$$\Gamma_{kk} = \frac{s(1-s)}{24} \frac{\left((\sigma_k^0)^2 - (\sigma_k^1)^2\right)^2}{(\sigma_k^0)^2 (\sigma_k^1)^2 \left(s(\sigma_k^0)^2 + (1-s)(\sigma_k^1)^2\right)}, \quad 1 \leq k \leq N. \quad (7.7)$$

For some  $k$ ,  $\Gamma_{kk} = 0$  if and only if  $\sigma_k^0 = \sigma_k^1$ . In other words, since there is no discriminability at all, there is no scope for further degradation.

In order to optimize the transform encoder  $(U, \boldsymbol{\Delta})$ , set  $\Omega = \Gamma$  in (3.18). We make the following assumption so that Theorem 5.9 gives the optimal design.

**Assumption 7.1** *Under  $H_i$ ,  $i = 0, 1$ ,  $\mathbf{X} \sim \mathcal{P}^i = \mathcal{G}(I, \mu^i)$ . Further, for each  $i$ ,  $\mathbf{V} \sim \mu^i$  satisfies (5.2) and (5.3). Moreover, the mixture  $\mu = \pi^0 \mu^0 + \pi^1 \mu^1$  is such that Assumption 5.6 holds for  $\mathcal{Q}(\mu)$ .*

Removing class conditioning, the source  $\mathbf{X}$  follows the mixture distribution

$$\pi^0 \mathcal{P}^0 + \pi^1 \mathcal{P}^1 = \mathcal{G}(I, \pi^0 \mu^0 + \pi^1 \mu^1) = \mathcal{G}(I, \mu),$$

and  $\mathbf{V} \sim \mu$  satisfies (5.2) and (5.3). Hence Assumption 5.4 holds. Further,  $p^0(\mathbf{x})$  and  $p^1(\mathbf{x})$  are both symmetric, hence so is the likelihood ratio  $l(\mathbf{x}) = p^1(\mathbf{x})/p^0(\mathbf{x})$ . Therefore, by Lemma 6.3, the matrix  $\Gamma(f, l, \mathcal{P}^0)$  is diagonal. This satisfies the relaxed version of Assumption 3.5 (see the penultimate paragraph of Sec. 5.5) for  $\Omega = \Gamma(f, l, \mathcal{P}^0)$ . Finally, by an earlier argument, replacing  $I$  in Assumption 7.1 by any  $C \in \mathcal{U}$  does not increase generality. Thus, in effect, we assume  $\mathbf{X}$  is decorrelated by a common KLT  $C$  under both  $H_0$  and  $H_1$ .

### 7.3 Example: Joint Classification/Reconstruction

As mentioned earlier, our analytical framework allows the flexibility to accommodate competing criteria. We highlight this aspect of our design using a joint classification/reconstruction paradigm where classification accuracy is traded off against reconstruction error.

Consider the usual transform encoding  $\tilde{\mathbf{X}} = Q(U\mathbf{X}, \boldsymbol{\Delta})$  at the encoder. As depicted in Fig. 6, the decoder simultaneously decides between classes  $H_0$  and  $H_1$ , and reconstructs

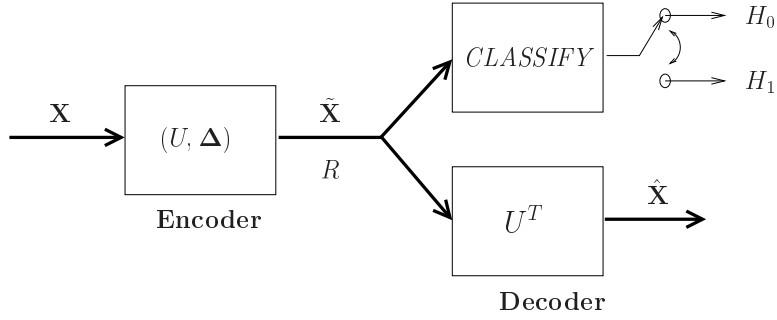


Figure 6: Joint classification/reconstruction at the decoder.

$\mathbf{X}$  to  $\hat{\mathbf{X}}$  based on the encoded vector  $\tilde{\mathbf{X}}$ . We assume Gaussian conditional distributions  $\mathbf{X} \sim \mathcal{P}^i = \mathcal{N}(\mathbf{0}, \Sigma^i)$  under  $H_i$ ,  $i = 0, 1$ . Due to Assumption 7.1, the covariance matrices  $\Sigma^i$ 's are diagonal. We measure classification accuracy by Chernoff distance, loss in which takes the asymptotic expression  $q(U, \mathbf{\Delta}; \Gamma) = \sum_{k=1}^N [U\Gamma U^T]_{kk} \Delta_k^2$ , where  $\Gamma$  is given by (7.7). Moreover, measuring reconstruction fidelity by the MSE  $q(U, \mathbf{\Delta}; \frac{1}{12}I) = \sum_{k=1}^N \frac{1}{12} \Delta_k^2$ , we adopt as the design criterion, the linear combination

$$D(U, \mathbf{\Delta}) = (1 - \lambda) q(U, \mathbf{\Delta}; \Gamma) + \lambda q(U, \mathbf{\Delta}; \frac{1}{12}I) = \sum_{k=1}^N [U\Omega U^T]_{kk} \Delta_k^2,$$

where

$$\Omega = (1 - \lambda)\Gamma + \lambda \frac{1}{12}I$$

and  $\lambda \in [0, 1]$  is a weighting factor. Note  $\lambda = 0$  and  $\lambda = 1$  correspond to the respective cases where only Chernoff loss and only MSE are used for optimization. We vary the mixing factor  $\lambda$  and find a particular value that achieves a desired tradeoff. The bit rate is computed based on the mixture distribution

$$\mathbf{X} \sim \mathcal{P} = \pi_0 \mathcal{P}^0 + \pi_1 \mathcal{P}^1 = \pi_0 \mathcal{N}(\mathbf{0}, \Sigma^0) + \pi_1 \mathcal{N}(\mathbf{0}, \Sigma^1).$$

We minimize  $D(U, \mathbf{\Delta})$  subject to an average bit-rate constraint  $R(U, \mathbf{\Delta}) = R$  in the high-resolution regime. By Theorem 5.9,  $(U, \mathbf{\Delta}) = (I, \mathbf{\Delta}^*)$  gives an optimal transform encoder. Further, (5.15) and (5.16) give the bit allocation strategy and the average bit rate, respectively. The differential entropies  $h(X_k)$ ,  $1 \leq k \leq N$ , are computed using MATLAB's numerical integration routine.

We assume  $\pi_0 = \pi_1 = \frac{1}{2}$ , and choose Chernoff exponent  $s = \frac{1}{2}$ , i.e., the Chernoff distance now coincides with the Bhattacharyya distance [52]. We also choose an average rate  $R = 3$  bits per component, which corresponds to a high-resolution regime. The dimensionality of the vectors is  $N = 64$ . In Fig. 7 (a), the conditional standard deviations  $\{\sigma_k^i\}$ , are plotted

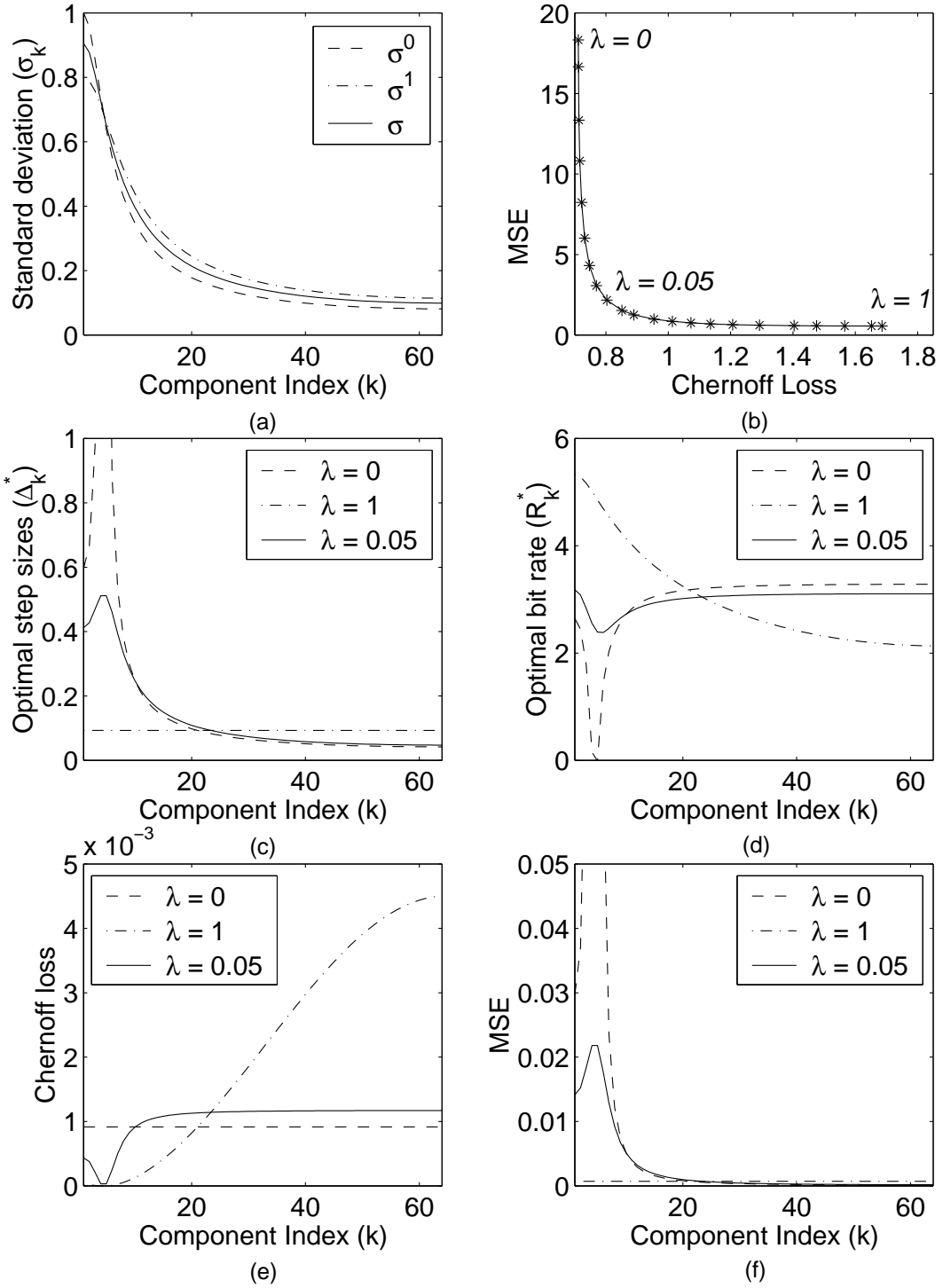


Figure 7: Average bit rate  $R = 3$  bits per component: (a) Component standard deviations, (b) Tradeoff between Chernoff loss and MSE, (c) Optimal quantizer step sizes, (d) Optimal bit allocation, (e) Component Chernoff loss, and (f) Component MSE.

under each hypothesis  $H_i$ ,  $i = 0, 1$ . On the same figure, we also plot the mixture standard deviations  $\sigma_k = \sqrt{\frac{1}{2} \left( (\sigma_k^0)^2 + (\sigma_k^1)^2 \right)}$ ,  $1 \leq k \leq 64$ . In Fig. 7(b), the loss in Chernoff distance is traded off against the MSE. Both the MSE corresponding to  $\lambda = 0$  and the Chernoff loss corresponding to  $\lambda = 1$  are substantial. But attractive tradeoffs are obtained for intermediate values of  $\lambda$ , e.g.,  $\lambda = 0.05$ . In Fig. 7(c)–(f), we plot optimal quantizer step sizes, optimal bit allocation, component-wise loss in Chernoff distance and component-wise MSE, respectively, for  $\lambda = 0, 0.05$  and  $1$ .

Quantizer step sizes and the bit allocation exhibit significant variation when optimized based on Chernoff distance alone ( $\lambda = 0$ ). Corresponding to the components with nearly equal competing conditional variances, the quantizer step sizes tend to be large and the number of allocated bits small<sup>2</sup>. Conversely, when competing variances are dissimilar, the step sizes are small and the number of allocated bits large. Further, the Chernoff loss is equally distributed among the components, whereas MSE (which is  $\frac{1}{12}$  times the square of the corresponding step size) varies significantly across different coefficients.

For a pure MSE criterion ( $\lambda = 1$ ), the optimal step sizes are equal as is MSE across the components. More bits are allocated to the coefficients with larger mixture variance. Further, the Chernoff loss shows significant variation across the components. For  $\lambda = 0.05$ , on the other hand, the amount of variation in different quantities is usually between the two extremes corresponding to  $\lambda = 0$  and  $\lambda = 1$ . Overall, the Chernoff loss is marginally more than that for  $\lambda = 0$  whereas the MSE is marginally higher than that for  $\lambda = 1$ , providing a good tradeoff between these criteria.

## 8 Conclusion

High-rate optimality of KLT is well known for variable-rate transform coding of Gaussian vectors, and follows because the decorrelated components are mutually independent. Motivated by such optimality, early research modeled natural signals as Gaussian (see, e.g., [2]). However, recent empirical studies show that natural signals such as natural images exhibit strong nonlinear dependence and non-Gaussian tails [5]. Such studies also observe that one wavelet transform approximately decorrelates a broad class of natural images, and propose wavelet image models following GSM distribution. In this paper, we have extended GSM models to

---

<sup>2</sup>According to Fig. 7(d), the allocated number of bits might be negative (see component 5), which is physically impossible. For such components, quantizer step sizes are very large, and the high-resolution approximations used in the derivation no longer hold.

GVSM, and showed optimality of certain KLT's. Applied to the natural images, our result implies near optimality of the decorrelating wavelet transform among all linear transforms. The theory applies to classical reconstruction as well as various estimation/classification scenarios where the generalized  $f$ -divergence measures performance. For instance, we have considered MMSE estimation in a multiplicative noise channel, where GVSM data have been shown to arise when the noise is GVSM. Further, closely following empirical observations, a class-independent common decorrelating transform is assumed in binary classification, and the framework has been extended to include joint classification/reconstruction systems. In fact, the flexibility of the analysis to accommodate competing design criteria has been illustrated by trading off Chernoff distance measures of discriminability against MSE signal fidelity.

The optimality analysis assumes high resolution quantizers. Further, under the MSE criterion, the optimal transform minimizes the sum of differential entropies of the components of the transformed vector. When the source is Gaussian, the KLT makes the transformed components independent and minimizes the aforementioned sum. On the contrary, when the source is GVSM, certain KLT's have been shown to minimize the above sum, although the decorrelated components are not necessarily independent. The crux of the proof lies in a key concavity property of the differential entropy of the scalar GSM. The optimality also extends to a broader class of quadratic criteria under certain conditions.

The aforementioned concavity of the differential entropy is a result of a fundamental property of the scalar GSM: Given a GSM pdf  $p(x)$  and two arbitrary points  $x_1 > x_0 > 0$ , there exist a constant  $c > 0$  and a zero-mean Gaussian pdf  $\phi(x)$  such that  $p(x)$  and  $c\phi(x)$  intersect at  $x = x_0$  and  $x_1$ . For such pair  $(c, \phi)$ , the function  $c\phi(x)$  lies below  $p(x)$  in the intervals  $[0, x_0)$  and  $(x_1, \infty)$ , and above in  $(x_0, x_1)$ . This property has potential implications beyond transform coding. Specifically, we conjecture that given an arbitrary pdf  $p(x)$ , the abovementioned property holds for all admissible pairs  $(x_0, x_1)$  if and only if  $p(x)$  is a GSM pdf. If our conjecture is indeed true, then this characterization of the GSM pdf is equivalent to the known characterizations by complete monotonicity and positive definiteness [5, 19, 6, 7].

## A Derivation of (4.5)

Using (4.3) in (4.4), we obtain

$$\begin{aligned}
p(\mathbf{x}) &= \int_{\mathbb{R}_+^2} d\mathbf{v} \phi(\mathbf{x}; \text{Diag}\{\mathbf{v}^2\}) \left[ 4\lambda v_1 v_2 e^{-(v_1^2+v_2^2)} + 16(1-\lambda)v_1 v_2 e^{-2(v_1^2+v_2^2)} \right] \\
&= 4\lambda \frac{1}{2\pi} \int_{\mathbb{R}_+} dv_1 \exp\left(-\frac{x_1^2}{2v_1^2} - v_1^2\right) \int_{\mathbb{R}_+} dv_2 \exp\left(-\frac{x_2^2}{2v_2^2} - v_2^2\right) \\
&\quad + 16(1-\lambda) \frac{1}{2\pi} \int_{\mathbb{R}_+} dv_1 \exp\left(-\frac{x_1^2}{2v_1^2} - 2v_1^2\right) \int_{\mathbb{R}_+} dv_2 \exp\left(-\frac{x_2^2}{2v_2^2} - 2v_2^2\right). \quad (\text{A.1})
\end{aligned}$$

Making use of the known integral (3.325 of [54])

$$\int_{\mathbb{R}_+} dv \exp\left(-\frac{a^2}{v^2} - b^2 v^2\right) = \frac{\sqrt{\pi}}{b} \exp(-2ab), \quad a \geq 0, b > 0,$$

for the pairs

$$(a, b) = \left(\frac{1}{\sqrt{2}}|x_1|, 1\right), \left(\frac{1}{\sqrt{2}}|x_2|, 1\right), \left(\frac{1}{\sqrt{2}}|x_1|, \sqrt{2}\right) \text{ and } \left(\frac{1}{\sqrt{2}}|x_2|, \sqrt{2}\right)$$

in (A.1), we obtain (4.5).

## B Proof of Property 4.3

Solving the simultaneous equations

$$g_1(x_i; \mu) = c\phi(x_i; \beta^2), \quad i = 0, 1, \quad (\text{B.1})$$

using expression (2.2) of  $\phi$ , obtain

$$\beta = \sqrt{\frac{x_1^2 - x_0^2}{2 \log \frac{g_1(x_0; \mu)}{g_1(x_1; \mu)}}}, \quad (\text{B.2})$$

$$c = \sqrt{2\pi} \beta \exp\left(\frac{x_0^2}{2\beta^2}\right) g_1(x_0; \mu). \quad (\text{B.3})$$

Using  $g_1(x_0; \mu) > g_1(x_1; \mu)$  ( $g_1(x; \mu)$  is decreasing in  $(0, \infty)$ ) in (B.2) and (B.3), we obtain  $\beta > 0$  and  $c > 0$ . By symmetry of both  $g_1(x; \mu)$  and  $\phi(x; \beta^2)$  about  $x = 0$ ,  $g_1(x; \mu) = c\phi(x; \beta^2)$  at  $x \in \{\pm x_0, \pm x_1\}$ . This proves the existence of the pair  $(c, \beta)$ .

Now define the function

$$\psi(x) := \exp\left(\frac{x^2}{2\beta^2}\right) [g_1(x; \mu) - c\phi(x; \beta^2)]. \quad (\text{B.4})$$

Clearly,

$$\operatorname{sgn}[\psi(x)] = \operatorname{sgn}[g_1(x; \mu) - c\phi(x; \beta^2)]. \quad (\text{B.5})$$

Using expression (4.2) of  $g_1(x; \mu)$  and (2.2) of  $\phi(x; \beta^2)$  in (B.4), we obtain

$$\begin{aligned} \psi(x) &= \int_{\mathcal{A}^>} \mu(dv) \frac{1}{\sqrt{2\pi v^2}} \exp\left(\frac{x^2}{2} \left(\frac{1}{\beta^2} - \frac{1}{v^2}\right)\right) + \frac{1}{\sqrt{2\pi\beta^2}} [\mu(\mathcal{A}^-) - c] \\ &\quad + \int_{\mathcal{A}^<} \mu(dv) \frac{1}{\sqrt{2\pi v^2}} \exp\left(-\frac{x^2}{2} \left(\frac{1}{v^2} - \frac{1}{\beta^2}\right)\right), \end{aligned} \quad (\text{B.6})$$

where

$$\mathcal{A}^{\text{op}} = \{v \in \mathbb{R}_+ : v \text{ op } \beta\}, \quad \text{op} \in \{>, <, =\}.$$

In (B.6), note if  $\mu(\mathcal{A}^>) > 0$ , the first integral is strictly increasing in  $x > 0$ , and if  $\mu(\mathcal{A}^<) > 0$ , the second integral is strictly decreasing in  $x > 0$ . Hence  $\psi(x)$  has at most four zeros. In fact, in view of (B.5),  $\psi(x)$  inherits the zeros of  $g_1(x; \mu) - c\phi(x; \beta^2)$ , and, therefore, has exactly four zeros given by  $\{\pm x_0, \pm x_1\}$ , which are all distinct. Consequently,  $\psi(x)$  changes sign at each zero. Now, observe that, as  $x \rightarrow \infty$ , the first integral in (B.6) increases without bound whereas the second integral is positive, implying

$$\lim_{x \rightarrow \infty} \psi(x) = \infty. \quad (\text{B.7})$$

Hence, in view of (B.5), the result follows.  $\square$

## C Proof of Property 4.4

**Lemma C.1** *For any  $\gamma > 0$ ,*

$$\begin{aligned} \operatorname{sgn}[\phi''(x; \gamma^2)] &= 0, & \forall |x| = x_0, x_1, \\ &= 1, & \forall 0 \leq |x| < x_0, |x| > x_1, \\ &= -1, & \forall x_0 < |x| < x_1. \end{aligned} \quad (\text{C.1})$$

where  $x_1 = \sqrt{(3 + \sqrt{6})\gamma^2}$ , and  $x_0 = \sqrt{(3 - \sqrt{6})\gamma^2}$ .

**Proof:** By expression (2.4), the polynomial  $f(x) = x^4 - 6\gamma^2 x^2 + 3\gamma^4$  determines the sign of  $\phi''(x, \gamma^2)$ . Since the zeros of  $f(x)$  occur at

$$x = \pm \sqrt{(3 \pm \sqrt{6})\gamma^2},$$

and are all distinct,  $f(x)$  changes sign at each zero crossing. Finally, noting

$$\lim_{x \rightarrow \infty} f(x) = \infty,$$

the result follows.  $\square$

At this point, we need the following technical condition that allows interchanging order of integration and differentiation.

**Fact C.2** *Suppose  $\xi(x)$  is a function such that  $|\xi(x)| \leq |f(x)|$  for some polynomial  $f(x)$  and all  $x \in \mathbb{R}$ . Then*

$$\int_{\mathbb{R}} dx \phi''(x; \gamma^2) \xi(x) = \frac{\partial^2}{\partial(\gamma^2)^2} \int_{\mathbb{R}} dx \phi(x; \gamma^2) \xi(x).$$

**Proof:** By the expressions (2.3) and (2.4) of  $\phi'(x; \gamma^2)$  and  $\phi''(x; \gamma^2)$ , respectively, and the assumption on  $\xi(x)$ , each of  $|\phi'(x; \gamma^2) \xi(x)|$  and  $|\phi''(x; \gamma^2) \xi(x)|$  is dominated by a function of the form  $|\check{f}(x)|\phi(x; \gamma^2)$ , where  $\check{f}(x)$  is a polynomial. Further, due to exponential decay of  $\phi(x; \gamma^2)$  in  $x$ ,  $|\check{f}(x)|\phi(x; \gamma^2)$  integrates to a finite quantity. Hence, applying the dominated convergence theorem [50], the result follows.  $\square$

For instance, setting  $\xi(x) = 1$  in Fact C.2, we obtain

$$\int_{\mathbb{R}} dx \phi''(x; \gamma^2) = \frac{\partial^2}{\partial(\gamma^2)^2} \left[ \int_{\mathbb{R}} dx \phi(x; \gamma^2) \right] = \frac{\partial^2}{\partial(\gamma^2)^2} [1] = 0. \quad (\text{C.2})$$

**Lemma C.3** *For any  $\gamma > 0$  and any  $\beta > 0$ ,*

$$\int_{\mathbb{R}} dx \phi''(x; \gamma^2) \log \phi(x; \beta^2) = 0. \quad (\text{C.3})$$

**Proof:** By expression (2.2), we have

$$\log \phi(x; \beta^2) = -\log \sqrt{2\pi\beta^2} - \frac{x^2}{2\beta^2}. \quad (\text{C.4})$$

Using (C.4) and noting  $\int_{\mathbb{R}} dx \phi(x; \gamma^2) = 1$  and  $\int_{\mathbb{R}} dx x^2 \phi(x; \gamma^2) = \gamma^2$ , we obtain

$$\int_{\mathbb{R}} dx \phi(x; \gamma^2) \log \phi(x; \beta^2) = -\log \sqrt{2\pi\beta^2} - \frac{\gamma^2}{2\beta^2}. \quad (\text{C.5})$$

Differentiating (C.5) twice with respect to  $\gamma^2$ , we obtain

$$\frac{\partial^2}{\partial(\gamma^2)^2} \int_{\mathbb{R}} dx \phi(x; \gamma^2) \log \phi(x; \beta^2) = 0. \quad (\text{C.6})$$

Setting  $\xi(x) = \log \phi(x; \beta^2)$  (which, by (C.4), is a polynomial), and applying Fact C.2 to the left hand side of (C.6), the result follows.  $\square$

**Proof of Property 4.4:** By Property C.3, equality holds in (4.9) for any Gaussian  $g_1(x; \mu)$ . It is, therefore, enough to show that (4.9) holds with a strict inequality for any non-Gaussian  $g_1(x; \mu)$ . Choose  $x_1 = \sqrt{(3 + \sqrt{6})} \gamma^2$ , and  $x_0 = \sqrt{(3 - \sqrt{6})} \gamma^2$ . Given such  $(x_0, x_1)$ , obtain the pair  $(c, \beta)$  from Lemma 4.3. Hence, by Lemmas 4.3 and C.1, we have

$$\text{sgn}[g_1(x; \mu) - c\phi(x; \beta^2)] = \text{sgn}[\phi''(x; \gamma^2)]. \quad (\text{C.7})$$

Noting  $\log x$  is monotone increasing in  $x > 0$ , and both  $g_1(x; \mu)$  and  $c\phi(x; \beta^2)$  take positive values, we have

$$\text{sgn}[g_1(x; \mu) - c\phi(x; \beta^2)] = \text{sgn}[\log g_1(x; \mu) - \log c\phi(x; \beta^2)]. \quad (\text{C.8})$$

Now, multiplying the right hand sides of (C.7) and (C.8), we obtain

$$\phi''(x; \gamma^2) [\log g_1(x; \mu) - \log c\phi(x; \beta^2)] > 0$$

almost everywhere, which, upon integration with respect to  $x$ , yields

$$\int_{\mathbb{R}} dx \phi''(x; \gamma^2) [\log g_1(x; \mu) - \log c\phi(x; \beta^2)] > 0. \quad (\text{C.9})$$

Rearranging, we finally obtain

$$\int_{\mathbb{R}} dx \phi''(x; \gamma^2) \log g_1(x; \mu) > \log c \int_{\mathbb{R}} dx \phi''(x; \gamma^2) + \int_{\mathbb{R}} dx \phi''(x; \gamma^2) \log \phi(x; \beta^2). \quad (\text{C.10})$$

In the right hand side of (C.10), the first term vanishes by (C.2) and the second term vanishes by Property C.3.  $\square$

## D Proof of Property 5.3

**Proof:** Let  $\Lambda = \text{Diag}\{\mathbf{v}^2\}$ . Noting

$$\phi(\mathbf{x}; C^T \Lambda C) \leq \phi(\mathbf{0}; C^T \Lambda C) = \frac{1}{(2\pi)^{N/2} \prod_{k=1}^N v_k}, \quad (\text{D.1})$$

we have

$$g(\mathbf{x}; C, \mu) = \int_{\mathbb{R}_+^N} \mu(d\mathbf{v}) \phi(\mathbf{x}; C^T \Lambda C) \leq \int_{\mathbb{R}_+^N} \mu(d\mathbf{v}) \phi(\mathbf{0}; C^T \Lambda C) = g(\mathbf{0}; C, \mu), \quad (\text{D.2})$$

where

$$g(\mathbf{0}; C, \mu) = \int_{\mathbb{R}_+^N} \mu(d\mathbf{v}) \frac{1}{(2\pi)^{N/2} \prod_{k=1}^N v_k} = \frac{1}{(2\pi)^{N/2}} \mathbb{E} \left[ \frac{1}{\prod_{k=1}^N V_k} \right]. \quad (\text{D.3})$$

(‘if’ part) If  $\mathbb{E} \left[ \frac{1}{\prod_{k=1}^N V_k} \right] < \infty$ , combining (D.2) and (D.3), we obtain

$$g(\mathbf{x}; C, \mu) \leq g(\mathbf{0}; C, \mu) < \infty. \quad (\text{D.4})$$

Since  $g(\mathbf{x}; C, \mu)$  is bounded, by the dominated convergence theorem [50], we obtain

$$\lim_{\|\mathbf{h}\| \rightarrow 0} g(\mathbf{x} + \mathbf{h}; C, \mu) = \int_{\mathbb{R}_+^N} \mu(d\mathbf{v}) \lim_{\|\mathbf{h}\| \rightarrow 0} \phi(\mathbf{x} + \mathbf{h}; C^T \Lambda C) = \int_{\mathbb{R}_+^N} \mu(d\mathbf{v}) \phi(\mathbf{x}; C^T \Lambda C) = g(\mathbf{x}; C, \mu), \quad (\text{D.5})$$

which proves continuity of  $g(\mathbf{x}; C, \mu)$ .

(‘only if’ part) Continuity of  $g(\mathbf{x}; C, \mu)$  implies  $g(\mathbf{0}; C, \mu) = \frac{1}{(2\pi)^{N/2}} \mathbb{E} \left[ \frac{1}{\prod_{k=1}^N V_k} \right] < \infty$ .

□

## E Proof of Property 5.5

Property 4.4 implies

$$\frac{\partial^2}{\partial (\gamma^2)^2} \int_{\mathbb{R}} dx \phi(x; \gamma^2) \log g_1(x) \geq 0 \quad (\text{E.1})$$

for any  $\gamma > 0$  and any continuous  $g_1(x; \nu, \sigma^2)$ . (Here and henceforth, substitute  $g_1(x; \nu, \sigma^2)$  for  $g_1(x; \mu)$  while referring earlier results.) To see this, fix the pair  $(x_0, x_1)$  in Lemma 4.3, and obtain the corresponding pair  $(c, \beta)$ . As  $x \rightarrow \infty$ , we have  $g_1(x; \nu, \sigma^2) > c\phi(x; \beta^2)$ , implying

$$|\log g_1(x; \nu, \sigma^2)| < |\log c\phi(x; \beta^2)| \quad (\text{E.2})$$

(because  $g_1(x; \nu, \sigma^2) \rightarrow 0$ ). In view of representation (2.2), note that  $\log c\phi(x; \beta^2)$  is a polynomial in  $x$ . Hence, in view of (E.2), a polynomial  $f(x)$  can be constructed for any continuous  $g_1(x; \nu, \sigma^2)$  such that

$$|\log g_1(x; \nu, \sigma^2)| \leq |f(x)| \quad \text{for all } x \in \mathbb{R}. \quad (\text{E.3})$$

Therefore, setting  $\xi(x) = \log g_1(x; \nu, \sigma^2)$ , and applying Fact C.2 to the left hand side of (4.9), we obtain (E.1).

For the pair  $(\sigma'^2, \sigma^2) \in \mathcal{V}^2(\nu)$ , define the functional

$$\theta(\nu, \sigma'^2, \sigma^2) := -h(g_1(\cdot; \nu, \sigma'^2)) - D(g_1(\cdot; \nu, \sigma'^2) \| g_1(\cdot; \nu, \sigma^2)) \quad (\text{E.4})$$

$$= \int_{\mathbb{R}} dx g_1(x; \nu, \sigma'^2) \log(g_1(x; \nu, \sigma^2)) \quad (\text{E.5})$$

$$= \int_{\mathbb{R}} dx \int_{\mathbb{R}_+^N} \nu(d\mathbf{w}) \phi(x; \sigma'^2(\mathbf{w})) \log(g_1(x; \nu, \sigma^2)). \quad (\text{E.6})$$

Here (E.5) follows by definitions of  $h(\cdot)$  and  $D(\cdot \| \cdot)$ . By assumption, verify from (E.4) that

$$-\infty < \theta(\nu, \sigma'^2, \sigma^2) < \infty. \quad (\text{E.7})$$

Hence, by Fubini's Theorem [50], the order of integration in (E.6) can be interchanged to obtain

$$\theta(\nu, \sigma'^2, \sigma^2) = \int_{\mathbb{R}_+^N} \nu(d\mathbf{w}) \int_{\mathbb{R}} dx \phi(x; \sigma'^2(\mathbf{w})) \log(g_1(x; \nu, \sigma^2)) \quad (\text{E.8})$$

Now, by (E.1), the inner integral in (E.8) is convex in  $\sigma'^2(\mathbf{w})$ , implying  $\theta(\nu, \sigma'^2, \sigma^2)$  is (essentially) convex in  $\sigma'^2$ . Therefore,

$$-\theta(\nu, \sigma_\lambda^2, \sigma_\lambda^2) \geq -(1 - \lambda)\theta(\nu, \sigma_0^2, \sigma_0^2) - \lambda\theta(\nu, \sigma_1^2, \sigma_1^2) \quad (\text{E.9})$$

for arbitrary pair  $(\sigma_0^2, \sigma_1^2) \in \mathcal{V}^2(\nu)$  and any convex combination  $\sigma_\lambda^2 = (1 - \lambda)\sigma_0^2 + \lambda\sigma_1^2$ , where  $\lambda \in (0, 1)$ . Adding  $(1 - \lambda)\theta(\nu, \sigma_0^2, \sigma_0^2) + \lambda\theta(\nu, \sigma_1^2, \sigma_1^2)$  to both sides of (E.9), and noting

$$\begin{aligned} -\theta(\nu, \sigma^2, \sigma^2) &= h(g_1(\cdot; \nu, \sigma^2)), \\ \theta(\nu, \sigma'^2, \sigma'^2) - \theta(\nu, \sigma'^2, \sigma^2) &= D(g_1(\cdot; \nu, \sigma'^2) \| g_1(\cdot; \nu, \sigma^2)) \end{aligned}$$

from (E.4), we obtain

$$\begin{aligned} &h(g_1(\cdot; \nu, \sigma_\lambda^2)) - (1 - \lambda)h(g_1(\cdot; \nu, \sigma_0^2)) - \lambda h(g_1(\cdot; \nu, \sigma_1^2)) \\ &\geq (1 - \lambda)D(g_1(\cdot; \nu, \sigma_0^2) \| g_1(\cdot; \nu, \sigma_\lambda^2)) + \lambda D(g_1(\cdot; \nu, \sigma_1^2) \| g_1(\cdot; \nu, \sigma_\lambda^2)) \\ &\geq 0. \end{aligned} \quad (\text{E.10})$$

Here (E.10) holds by nonnegativity of  $D(\cdot \| \cdot)$  [43]. This inequality is an equality if and only if

$$g_1(x; \nu, \sigma_0^2) = g_1(x; \nu, \sigma_1^2) = g_1(x; \nu, \sigma_\lambda^2) \quad \text{for all } x \in \mathbb{R}.$$

Hence the result.  $\square$

## F Proof of Lemma 6.2

Write  $\Delta = \Delta\alpha$  as in (3.10), where  $\alpha \in \mathbb{R}^N$ . Choose

$$\mathbf{s}(\mathbf{x}) = \text{Diag}\{1/\alpha\}U\mathbf{x} \quad (\text{F.1})$$

so that  $t(\mathbf{x}) = Q(\mathbf{s}(\mathbf{x}); \Delta\mathbf{1}) = Q(U\mathbf{x}; \Delta)$ . Since  $\mathbf{s}(\mathbf{x})$  is linear, if  $\mathcal{R}(f, l \circ \mathbf{s}, \mathcal{P} \circ \mathbf{s})$  holds, then so does  $\mathcal{R}(f, l, \mathcal{P})$ , and *vice versa*. From (F.1), the Jacobian of  $\mathbf{s}(\mathbf{x})$  is given by

$$J = U^T \text{Diag}\{1/\alpha\}, \quad (\text{F.2})$$

which is independent of  $\mathbf{x}$ . Inverting  $J$  in (F.2), we obtain  $J^{-1} = \text{Diag}\{\alpha\}U$ . Replacing in (2.8), we obtain

$$\begin{aligned} \Delta D_f((U, \Delta); l, \mathcal{P}) &\sim \frac{1}{24} \Delta^2 \mathbb{E}_{\mathcal{P}} [\|\alpha \odot (U\nabla l)\|^2 f'' \circ l] \\ &= \frac{1}{24} \Delta^2 \text{tr} [(\alpha\alpha^T) \odot (U\mathbb{E}_{\mathcal{P}} [(\nabla l \nabla^T l) f'' \circ l] U^T)] \\ &= \frac{1}{24} \sum_{k=1}^N \Delta^2 \alpha_k^2 [U\Gamma U^T]_{kk}, \end{aligned}$$

which gives the desired result.  $\square$

## G Proof of Lemma 6.3

The first statement immediately follows from the definition (6.3) of  $\Gamma$  and the convexity of  $f$ . Now, if  $l(\mathbf{x})$  is symmetric in  $x_k$  about  $x_k = 0$ , then  $\frac{\partial l(\mathbf{x})}{\partial x_k}$  is antisymmetric in  $x_k$  about  $x_k = 0$ . Note, from (6.3), that

$$\Gamma_{jk} = \mathbb{E} \left[ \frac{\partial l(\mathbf{X})}{\partial X_j} \frac{\partial l(\mathbf{X})}{\partial X_k} f'' \circ l(\mathbf{X}) \right], \quad 1 \leq j, k \leq N. \quad (\text{G.1})$$

The quantity inside the square braces in (G.1) is antisymmetric in  $X_k$  about  $X_k = 0$ , if  $j \neq k$ . Hence, by symmetry of  $p(\mathbf{x})$  in  $x_k$ , we obtain  $\Gamma_{jk} = 0$ ,  $j \neq k$ , which proves the second statement.  $\square$

## References

- [1] J.-Y. Huang and P. Schultheiss, "Block quantization of correlated Gaussian random variables," *IEEE Trans. on Commun. Systems*, vol. 11, pp. 289-296, Sep. 1963.
- [2] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwer, Boston MA, 1992.
- [3] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 1999.
- [4] M. Effros, H. Feng, and K. Zeger, "Suboptimality of the Karhunen-Loève transform for transform coding," *IEEE Trans. on Inform. Theory*, vol. 50, no. 8, pp. 1605-1619, Aug. 2004.
- [5] M. J. Wainwright, E. P. Simoncelli, and A. S. Willsky, "Random cascades on wavelet trees and their use in analyzing and modeling natural images," *Applied and Computational Harmonic Analysis*, Vol. 11, pp. 89-123, 2001.
- [6] I. J. Schoenberg, "Metric spaces and completely monotone functions," *The Annals of Mathematics*, Second Series, Vol. 39, Issue 4, pp 811-841, Oct. 1938.
- [7] D. F. Andrews and C. L. Mallows, "Scale mixtures of normal distributions", *J. Roy. Statist. Soc.*, Vol. 36, pp. 99—102, 1974.
- [8] I. F. Blake, and J. B. Thomas, "On a class of processes arising in linear estimation theory," *IEEE Trans. on Inform. Theory*, Vol. IT-14, pp. 12-16, Jan. 1968.
- [9] H. Brehm, and W. Stammer, "Description and generation of spherically invariant speech-model signals," *Signal Processing*, Vol. 12, pp. 119-141, 1987.
- [10] P. L. D. León II, and D. M. Etter, "Experimental results of subband acoustic echo cancelers under spherically invariant random processes," *Proc. ICASSP'96*, pp. 961-964.
- [11] F. Müller, "On the motion compensated prediction error using true motion fields," *Proc. ICIP'94*, Vol. 3., pp. 781-785, Nov. 1994.
- [12] E. Conte, and M. Longo, "Characterization of radar clutter as a spherically invariant random process," *IEE Proc. F, Commun., Radar, Signal Processing*, Vol. 134, pp. 191-197, 1987.

- [13] M. Rangaswamy, "Spherically invariant random processes for modeling non-Gaussian radar clutter," *Conf. Record of the 27th Asilomar Conf. on Signals, Systems and Computers*, Vol. 2, pp. 1106-1110, 1993.
- [14] A. Abdi, H. A. Barger, and M. Kaveh, "Signal modeling in wireless fading channels using spherically invariant processes," *Proc. ICASSP'2000*, Vol. 5, pp. 2997-3000, 2000.
- [15] J. K. Romberg, H. Choi, and R. G. Baraniuk, "Bayesian tree-structured image modeling using wavelet-domain hidden Markov models," *IEEE Trans. on Image Proc.*, Vol. 10, No. 7, pp. 1056-1068, Jul. 2001.
- [16] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. on Pattern Anal. and Machine Intell.*, Vol. 11, No. 7, pp. 674-693, 1989.
- [17] S. M. LoPresto, K. Ramchandran and M. T. Orchard, "Image coding based on mixture modeling of wavelet coefficients and a fast estimation-quantization framework," *Proceedings of Data Compression Conference*, pp. 221-230, 1997.
- [18] J. Liu, and P. Moulin, "Information-theoretic analysis of interscale and intrascale dependencies between image wavelet coefficients," *IEEE Trans. on Image Proc.*, Vol. 10, No. 11, pp. 1647-1658, Nov. 2001.
- [19] K. Yao, "A representation theorem and its applications to spherically-invariant random processes," *IEEE Trans. on Inform. Theory*, Vol. IT-19, No. 5, pp. 600-608, Sep. 1973.
- [20] H. M. Leung, and S. Cambanis, "On the rate distortion functions of spherically invariant vectors and sequences," *IEEE Trans. on Inform. Theory*, Vol. IT-24, pp. 367-373, Jan. 1978.
- [21] M. Herbert, "Lattice quantization of spherically invariant speech-model signals," *Archiv für Elektronik und Übertragungstechnik*, AEÜ-45, pp. 235-244, Jul. 1991.
- [22] F. Müller, and F. Geiszläger, "Vector quantization of spherically invariant random processes," *Proc. IEEE Int'l. Symp. on Inform. Theory*, p. 439, Sept. 1995.
- [23] S. G. Chang, B. Yu and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Trans. on Image Proc.*, Vol. 9, No. 9, pp. 1532-1546, Sep. 2000.

- [24] Y.-P. Tan, D.D. Saur, S.R. Kulkarni and P.J. Ramadge, “Rapid estimation of camera motion from compressed video with application to video annotation,” *IEEE Trans. on Circ. and Syst. for Video Tech.*, Vol. 10, No. 1, Feb. 2000.
- [25] J.-W. Nahm and M.J.T. Smith, “Very low bit rate data compression using a quality measure based on target detection performance,” *Proc. SPIE*, 1995.
- [26] S.R.F. Sims, “Data compression issues in automatic target recognition and the measurement of distortion,” *Opt. Eng.*, Vol. 36, No. 10, pp. 2671-2674, Oct. 1997.
- [27] A. Jain, P. Moulin, M. I. Miller and K. Ramchandran, “Information-theoretic bounds on target recognition performance based on degraded image data,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 9, pp. 1153-1166, Sep. 2002.
- [28] K. L. Oehler and R. M. Gray, “Combining image compression and classification using vector quantization,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 17, No. 5, pp. 461-473, May 1995.
- [29] K.O. Perlmutter, S.M. Perlmutter, R.M. Gray, R. A. Olshen, and K. L. Oehler, “Bayes risk weighted vector quantization with posterior estimation for image compression and classification,” *IEEE Trans. Im. Proc.*, Vol. 5, No. 2, pp. 347—360, Feb. 1996.
- [30] T. Fine, “Optimum mean-squared quantization of a noisy input,” *IEEE Trans. on Inform. Theory*, Vol. 11, No. 2, pp. 293–294, Apr. 1965.
- [31] J. K. Wolf and J. Ziv, “Transmission of noisy information to a noisy receiver with minimum distortion,” *IEEE Trans. on Inform. Theory*, Vol. 16, No. 4, pp. 406–41, Jul. 1970.
- [32] Y. Ephraim and R. M. Gray, “A unified approach for encoding clean and noisy sources by means of waveform and autoregressive model vector quantization,” *IEEE Trans. on Inform. Theory*, Vol. 34, No. 4, pp. 826–834, Jul. 1988.
- [33] A. Rao, D. Miller, K. Rose and A. Gersho, “A generalized VQ method for combined compression and estimation,” *Proc. ICASSP*, pp. 2032-2035, May 1996.
- [34] H. V. Poor and J. B. Thomas, “Applications of Ali-Silvey distance measures in the design of generalized quantizers,” *IEEE Trans. on Communications*, Vol. COM-25, pp. 893-900, Sep. 1977.

- [35] G. R. Benitz and J. A. Bucklew, "Asymptotically optimal quantizers for detection of i.i.d. data," *IEEE Trans. on Inform. Theory*, Vol. 35, No. 2, pp. 316-325, Mar. 1989.
- [36] D. Kazakos, "New error bounds and optimum quantization for multisensor distributed signal detection," *IEEE Trans. on Communications*, Vol. 40, No. 7, pp. 1144-1151, Jul. 1992.
- [37] H. V. Poor, "Fine quantization in signal detection and estimation," *IEEE Trans. on Inform. Theory*, Vol. 34, No. 5, pp. 960—972, Sept. 1988.
- [38] M. Goldberg "Applications of wavelets to quantization and random process representations," Ph. D. dissertation, Stanford University, May, 1993.
- [39] V. K. Goyal, "Theoretical foundations of transform coding," *IEEE Signal Processing Magazine*, vol. 28, No. 5, pp. 9-21, Sep. 2001.
- [40] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Trans. on Inform. Theory*, Vol. 44, pp. 2325-2383, Oct. 1998.
- [41] S. Jana, and P. Moulin, "Optimal transform coding of Gaussian mixtures for joint classification/reconstruction," *Proc. DCC'03*, pp. 313–322.
- [42] G. Zhou and G. B. Giannakis, "Harmonics in Gaussian multiplicative and additive noise: Cramer-Rao bounds," *IEEE Trans. on Sig. Proc.*, Vol. 43, No. 5, pp. 1217-1231, May 1995.
- [43] T. M. Cover, and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, 1991.
- [44] P. L. Zador, "Topics in the asymptotic quantization of continuous random variables," Bell Laboratories Technical Memorandum, 1966.
- [45] V. N. Koshelev, "Quantization with minimal entropy," *Probl. Pered. Inform.*, No. 14, pp. 151-156, 1963.
- [46] H. Gish, and J. Pierce, "Asymptotically efficient quantizing," *IEEE Trans. on Inform. Theory*, Vol. IT-14, pp. 676-683, Sep. 1968.
- [47] R. M. Gray, T. Linder and J. Li, "A Lagrangian formulation of Zador's entropy-constrained quantization theorem," *IEEE Trans. on Inform. Theory*, Vol. 48, No. 3, pp. 695-707, Mar. 2002.

- [48] N. S. Jayant and P. Knoll, *Digital Coding of Analog Waveforms*, Prentice-Hall, 1988.
- [49] T. Linder, R. Zamir and K. Zeger, “High-resolution source coding for non-difference distortion measures: multidimensional companding,” *IEEE Trans. on Inform. Theory*, Vol. 45, No. 2, pp. 548-561, Mar. 1999.
- [50] J. W. Dshalalow, *Real Analysis: An Introduction to the Theory of Real Functions and Integration*, Chapman & Hall/CRC, 2001.
- [51] H. Avi-Itzhak and T. Diep, “Arbitrarily tight upper and lower bounds on the Bayesian probability of error”, *IEEE Trans. on Pattern Anal. and Machine Intel.*, Vol. 18, No. 1, pp. 89-91, Jan. 1996.
- [52] H. L. Van Trees, *Detection, Estimation and Modulation Theory, Part I*, John Wiley & Sons, 1968.
- [53] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, 1979.
- [54] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, Academic Press, 1980.

**Soumya Jana:** Soumya Jana received the B.Tech. degree in Electronics and Electrical Communication Engineering from the Indian Institute of Technology, Kharagpur, in 1995, the M.E. degree in Electrical Communication Engineering from the Indian Institute of Science, Bangalore, in 1997, and the Ph.D. degree in Electrical Engineering from the University of Illinois at Urbana-Champaign in 2005. He worked as a software engineer in Silicon Automation Systems, India, in 1997-1998. He is currently a postdoctoral researcher at the University of Illinois.

**Pierre Moulin:** Pierre Moulin received the degree of Ingénieur civil électricien from the Faculté Polytechnique de Mons, Belgium, in 1984, and the M.Sc. and D.Sc. degrees in Electrical Engineering from Washington University in St. Louis in 1986 and 1990, respectively.

He was a researcher at the Faculté Polytechnique de Mons in 1984–85 and at the Ecole Royale Militaire in Brussels, Belgium, in 1986–87. He was a Research Scientist at Bell Communications Research in Morristown, New Jersey, from 1990 until 1995. In 1996, he joined the University of Illinois at Urbana-Champaign, where he is currently Professor in the Department of Electrical and Computer Engineering, Research Professor at the Beckman Institute and the Coordinated Science Laboratory, and Affiliate Professor in the Department of Statistics. His fields of professional interest are image and video processing, compression, statistical signal processing and modeling, decision theory, information theory, information hiding, and the application of multiresolution signal analysis, optimization theory, and fast algorithms to these areas.

Dr. Moulin has served as an Associate Editor of the IEEE Transactions on Information Theory and the IEEE Transactions on Image Processing, Co-chair of the 1999 IEEE Information Theory Workshop on Detection, Estimation, Classification and Imaging, Chair of the 2002 NSF Workshop on Signal Authentication, and Guest Associate Editor of the IEEE Transactions on Information Theory's 2000 special issue on information-theoretic imaging and of the IEEE Transactions on Signal Processing's 2003 special issue on data hiding. During 1998-2003 he was a member of the IEEE IMDSP Technical Committee. More recently he has been Area Editor of the IEEE Transactions on Image Processing and Guest Editor of the IEEE Transactions on Signal Processing's supplement series on Secure Media. He is currently serving on the Board of Governors of the SP Society and as Editor-in-Chief of the Transactions on Information Forensics and Security.

He received a 1997 Career award from the National Science Foundation and a IEEE Signal Processing Society 1997 Senior Best Paper award. He is also co-author (with Juan Liu) of

a paper that received an IEEE Signal Processing Society 2002 Young Author Best Paper award. He was 2003 Beckman Associate of UIUC's Center for Advanced Study, received UIUC's 2005 Sony Faculty award, and was plenary speaker at ICASSP 2006.