Generalized Triangular Transform Coding

Ching-Chih Weng, Chun-Yang Chen, and P. P. Vaidyanathan

Dept. of Electrical Engineering, MC 136-93

California Institute of Technology, Pasadena, CA 91125, USA

E-mail: {cweng@caltech.edu, cych@caltech.edu, and ppvnath@systems.caltech.edu}

Abstract—A general family of optimal transform coders (TC) is introduced here based on the generalized triangular decomposition (GTD) developed by Jiang, et al. This family includes the Karhunen-Loeve transform (KLT), and the prediction-based lower triangular transform (PLT) introduced by Phoong and Lin, as special cases. The coding gain of the entire family, with optimal bit allocation, is equal to those of the KLT and the PLT. Even though the PLT is not applicable for vectors which are not blocked versions of scalar wide sense stationary (WSS) processes, the GTD based family includes members which are natural extensions of the PLT, and therefore also enjoy the so-called MINLAB structure of the PLT which has the unit noise-gain property. Other special cases of the GTD-TC are the GMD (geometric mean decomposition) and the BID (bidiagonal transform). The GMD in particular has the property that the optimal bit allocation (which is required for achieving the maximum coding gain) is a uniform allocation, thereby eliminating the need for bit allocation.¹

I. INTRODUCTION

In transform coder (TC) theory, the Karhunen-Loeve transform (KLT) is known for its optimality properties [1], [6], [17]. For example it provides maximum coding gain when high bit rate scalar quantizers are used in the transform domain. The KLT essentially diagonalizes the autocorrelation matrix of the input vector $\mathbf{x}$ before quantization. The decorrelated components are typically quantized by independent scalar quantizers.²

If the vector $\mathbf{x}$ being transformed is a blocked version of a scalar wide sense stationary (WSS) process $x(n)$, then the coding gain of the KLT can also be achieved by using a different kind of transform called the prediction-based lower triangular transform or PLT, which was introduced into the signal processing literature by Phoong and Lin [13]. The PLT is based on the theory of linear prediction for the scalar WSS process $x(n)$. PLT has smaller design cost because fast algorithms such as the Levinson algorithm can be used instead of matrix diagonalization. The implementation complexity for the PLT is 50% smaller than that of the KLT [13]. However, the PLT as introduced in [13] is not applicable for vectors $\mathbf{x}$ which are *not* blocked versions of scalar WSS processes.

This paper introduces a general family for transform coding based on the generalized triangular decomposition (GTD) introduced by Jiang, et al., in the context of optimal transceiver design in digital communications [8]. We will show that the GTD-TC family has the following features:

- 1) It includes the KLT and PLT as special cases.
- 2) The coding gain for any member of the family is equal to that of the KLT.
- 3) Unlike the PLT, the input vector $\mathbf{x}$ is not required to be a blocked version of a WSS process. One of the attractive features of the PLT is the existence of a structure with

unit noise gain, called the MINLAB structure [13]. The GTD based family includes a PLT-like special case which also enjoys the MINLAB structure. In this sense it extends some of the features of the PLT for the case where $\mathbf{x}$ is not a blocked version of a scalar process.

- 4) Like the KLT and the PLT the GTD family also produces a decorrelated set of components at the inputs of the scalar quantizers. The GTD offers a great deal of freedom in the distribution of the variances of these decorrelated transform domain components.
- 5) Other special cases of the GTD transform coder includes the GMD (geometric mean decomposition) and the BID (bidiagonal transform).
- 6) The GMD in particular has the property that the optimal bit allocation (which is required for achieving the maximum coding gain) is a uniform allocation! *The coding gain that is achieved by the KLT with optimal bit allocation is therefore achieved by the GMD without bit allocation.* Recall here that the closed form formula for optimal bit allocation used by KLT and other transforms [6] often yields non integer values for the bits. The approximation of these with integers would lead to suboptimality of the transform coder. Since the GMD-based method uses identical bits for all the transform domain coefficients without compromising optimality, this disadvantage is not present any more.

The family of GTD coders therefore provides a unified framework for a number of optimal linear transforms for high bit rate coders.

Paper outline. The paper is organized as follows. Section II briefly reviews the KLT and the PLT. In Section III we discuss the proposed GTD-TC. Several examples of the GTD-TC, such as the GMD-TC and BID-TC are given here. Section IV provides numerical simulations related to the topic discussed in the paper. Section V concludes this paper.³

Assumptions. All signals and transforms discussed in this paper are assumed to be real-valued. We assume that the $M \times 1$ input $\mathbf{x}(n)$ is a zero mean real-valued wide-sense stationary vector process, with positive definite covariance matrix $\mathbf{R}_x$. The time argument n is dropped when redundant.

II. PRELIMINARIES AND REVIEWS

The transform coder is shown in Fig. 1. The signal $\mathbf{x}$ is first multiplied by an $M \times M$ matrix $\mathbf{T}$ so that $\mathbf{y} = [y_1 \ y_2 \ \cdots \ y_M]^T = \mathbf{T}\mathbf{x}$. The quantizers are scalar quantizers, and are modeled as an additive noise sources so that $\hat{y}_i =$

¹This work is supported in parts by the ONR grant N00014-08-1-0709, and the TMS scholarship 94-2-A-018 of the National Science Council of Republic of China, Taiwan.

²Depending on the objective function to be optimized and the statistical assumptions involved, it can also be argued that the KLT is suboptimal in some sense [2]; we do not get into this aspect here.

³**Notations.** Boldface upper-case letters denote matrices, boldface lower-case letters denote column vectors, and italics denote scalars. The superscripts $(\cdot)^T$ and $(\cdot)^\dagger$ denote the transpose and conjugate transpose operations. $\mathbf{A}_{ij}$ denotes the (i, j) th element of the matrix $\mathbf{A}$. By $\mathbf{A} \succeq \mathbf{B}$, we mean that $\mathbf{A} - \mathbf{B}$ is positive semi-definite. For vector $\mathbf{x}$, the notation $\text{diag}(\mathbf{x})$ denotes the diagonal matrix with diagonal terms equal to the elements in the vector $\mathbf{x}$. For matrix $\mathbf{X}$, the notation $\text{diag}(\mathbf{X})$ denotes the column vector whose elements are the diagonal terms of $\mathbf{X}$. The notation $\mathbf{a} \succ_+ \mathbf{b}$ means that the vector $\mathbf{a}$ majorizes $\mathbf{b}$ additively [12], [10]. Similarly $\mathbf{a} \succ_\times \mathbf{b}$ means that the vector $\mathbf{a}$ majorizes $\mathbf{b}$ multiplicatively [8], [10].

$y_i + q_i$. Suppose the i th quantizer Q_i has b_i bits, then the variance of the quantization error q_i satisfies

$$\sigma_{q_i}^2 = c2^{-2b_i}\sigma_{y_i}^2, \quad (1)$$

where $\sigma_{y_i}^2$ is the variance of the signal input to the i th quantizer. This result generally holds under the high bit rate assumption [6], [11], [17]. The constant c depends on the type of the quantizer and the statistics of y_i . It is assumed that all the scalar quantizers have the same c . The signal is reconstructed at the decoder by multiplying with $\mathbf{T}^{-1}$.

A. Transform Coders and the KLT

The problem of minimizing the arithmetic mean of MSE (AM-MSE) under the average bit rate constraint is solved by the KLT [17]. The KLT uses $\mathbf{T} = \mathbf{U}^T$, where $\mathbf{U}$ is any $M \times M$ orthonormal matrix such that $\mathbf{R}_x = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^T$, where $\mathbf{\Sigma}$ is the diagonal matrix of the eigenvalues $\{\sigma_1^2, \dots, \sigma_M^2\}$ of $\mathbf{R}_x$ (assumed to be in non-increasing order).

Under the high bit rate assumption (1), the optimal bit allocation is given by the bit-loading formula [6], [17]

$$b_i = b + \frac{1}{2} \log_2 \frac{\sigma_i^2}{\det(\mathbf{R}_x)^{\frac{1}{M}}}, \quad (2)$$

where the average bit rate is constrained to be b bits per data stream. The resulting AM-MSE is

$$\mathcal{E}_{KLT} = c2^{-2b} \det(\mathbf{R}_x)^{\frac{1}{M}}. \quad (3)$$

It was shown in [16] that under the high bit rate assumption, it is not a loss of generality to assume that the transform is orthonormal. It should be noted that the KLT decorrelates the signal, so the components of $\mathbf{y}$ are statistically independent (under the Gaussian assumption). This justifies the use of scalar quantizers.

Fig. 1. Schematic of a transform coder with scalar quantizers.

B. Prediction-Based Lower Triangular Transform (PLT)

The PLT, proposed in [13], is a signal dependent non-orthonormal transform, which utilizes linear prediction theory [6], [18]. It has the same decorrelation property as the KLT, and is shown to have the same MMSE performance if the so-called “minimum noise structure” and optimal bit allocation are used [13]. In the original article of Phoong and Lin [13], the PLT is used for the vector $\mathbf{x}$ obtained by blocking a scalar WSS $x(n)$. In the following we briefly review the idea of the PLT. We also show that the PLT can actually be used for a vector process which need not to be a blocked version of a scalar process. The development of [13] which was based on linear prediction theory does not apply in this case, but some of the main conclusions continue to be true as we shall elaborate next.

Consider the LDU decomposition [5] of $\mathbf{R}_x$ given by

$$\mathbf{R}_x = \mathbf{L}\mathbf{D}\mathbf{L}^T, \quad (4)$$

Fig. 2. A direct implementation of the PLT.

Here $\mathbf{L}$ is lower triangular with diagonal elements equal to unity, and $\mathbf{D}$ is a diagonal matrix with positive diagonal elements. We can rewrite this as

$$(\mathbf{L}^{-1}\mathbf{R}_x^{\frac{1}{2}})(\mathbf{L}^{-1}\mathbf{R}_x^{\frac{1}{2}})^T = \mathbf{D}.$$

That is, $\mathbf{L}^{-1}\mathbf{x}$ has the diagonal covariance matrix $\mathbf{D}$. So pre-multiplying $\mathbf{x}$ with $\mathbf{L}^{-1}$ results in decorrelation. The transform coder with $\mathbf{T} = \mathbf{L}^{-1}$ will be referred to as the PLT here, and its implementation is shown in Fig. 2. The multipliers s_{km} in the figure are the coefficients in the matrix $\mathbf{L}$. In this implementation the quantizer noise is amplified by $\mathbf{T}^{-1} = \mathbf{L}$. A different implementation, called the minimum noise structure I (MINLAB(I)) [14] is shown in Fig. 3. This structure is shown to have the unity noise gain property [13].⁴ It minimizes the AM-MSE if the bit loading for each quantizer follows the bit loading formula:

$$b_i = b + \frac{1}{2} \log_2 \frac{\mathbf{D}_{ii}}{\det(\mathbf{R}_x)^{\frac{1}{M}}}. \quad (5)$$

The resulting AM-MSE will be

$$\mathcal{E}_{PLT} = c2^{-2b} \det(\mathbf{R}_x)^{\frac{1}{M}}, \quad (6)$$

which is the same as what the KLT can achieve when the optimal bit loading is applied. The reason for the name PLT is that the multipliers s_{km} are related to optimal linear predictor coefficients [13] when $\mathbf{x}(n)$ is the blocked version of a scalar WSS process $x(n)$. For simplicity we shall continue to use the term PLT even when this is not the case.

The PLT achieves the same optimal performance as the KLT but with less computational complexity in the implementation. Other attractive features are mentioned in [13].

Fig. 3. The PLT implemented using MINLAB(I) structure.

III. GENERALIZED TRIANGULAR DECOMPOSITION TRANSFORM CODER

In this section we will show how to construct the GTD-TC from a given covariance matrix. We will also show that actually both the KLT and the PLT are special cases of the

⁴It can be shown [13] that the components of $\mathbf{x}$ and $\hat{\mathbf{x}}$ in Fig. 3 are related as $\hat{x}_k - x_k = q_k$, where q_k is the error introduced at the k th quantizer. So the reconstruction error equals the quantization error, yielding the unit noise gain property [13].

GTD-TC. Several other interesting instances of GTD-TC, i.e., GMD-TC, BID-TC, and the combination of GMD-TC with progressive transmission, will be discussed. Before going into the GTD theory, let us first review the notion of multiplicative majorization [10], [4].

Definition 1: Multiplicative Majorization [10], [4]: Given two vectors $\mathbf{a} = [a_1, a_2, \dots, a_n]$ and $\mathbf{b} = [b_1, b_2, \dots, b_n]$ where a_i and b_i are all positive, we say $\mathbf{a}$ is multiplicatively majorized by $\mathbf{b}$, or $\mathbf{b}$ multiplicatively majorizes $\mathbf{a}$, and we write $\mathbf{a} \prec_{\times} \mathbf{b}$ if

$$\prod_{i=1}^k a_{[i]} \leq \prod_{i=1}^k b_{[i]}, \text{ whenever } 1 \leq k \leq n,$$

and equality holds when $k = n$. Here “[i]” denotes the component of the vector with i -th largest magnitude.

Multiplicative majorization property plays an important role in GTD theory as shown by the following result proved in [8].

Theorem 1: The generalized triangular decomposition (GTD): Let $\mathbf{H} \in \mathbb{C}^{m \times n}$ be a rank- K matrix with singular values $\sigma_{h,1}, \sigma_{h,2}, \dots, \sigma_{h,K}$ in descending order. Let $\mathbf{r} = [r_1, r_2, \dots, r_K]$ be any vector which satisfies

$$\mathbf{a} \prec_{\times} \mathbf{h}, \quad (7)$$

where $\mathbf{a} = [|r_1|, |r_2|, \dots, |r_K|]$ and $\mathbf{h} = [\sigma_{h,1}, \sigma_{h,2}, \dots, \sigma_{h,K}]$. Then there exist matrices $\mathbf{R}$, $\mathbf{Q}$, and $\mathbf{P}$ such that

$$\mathbf{H} = \mathbf{Q}\mathbf{R}\mathbf{P}^\dagger, \quad (8)$$

where $\mathbf{R}$ is a $K \times K$ upper triangular matrix with diagonal terms equal to r_k , and $\mathbf{Q} \in \mathbb{C}^{m \times K}$ and $\mathbf{P} \in \mathbb{C}^{n \times K}$ both have orthonormal columns.

According to the GTD factorization algorithm described in [8], if $\mathbf{H}$ and $\mathbf{r}$ are real valued, then the matrices $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{P}$ can be taken to be real valued. There are many standard decompositions which can be regarded as special instances of the GTD. These are listed below. The first five can be found in standard texts [3], [5].

- 1) The singular value decomposition (SVD).
- 2) The Schur decomposition.
- 3) The QR decomposition.
- 4) The complete orthogonal decomposition.
- 5) The bi-diagonal decomposition.
- 6) The geometric mean decomposition (GMD) [7].

Now consider the transform coding problem again. Suppose the LDU decomposition of $\mathbf{R}_x$ is $\mathbf{R}_x = \mathbf{L}\mathbf{D}\mathbf{L}^T$, as in eq. (4). Decompose $\mathbf{D}^{\frac{1}{2}}\mathbf{L}^T$ using the GTD, i.e.,

$$\mathbf{D}^{\frac{1}{2}}\mathbf{L}^T = \mathbf{Q}\mathbf{R}\mathbf{P}^T. \quad (9)$$

Then we can express $\mathbf{R}_x$ as

$$\begin{aligned} \mathbf{R}_x &= \mathbf{P}\mathbf{R}^T\mathbf{Q}^T\mathbf{Q}\mathbf{R}\mathbf{P}^T \\ &= \mathbf{P}\mathbf{L}_1\text{diag}([\mathbf{R}_{11}^2, \mathbf{R}_{22}^2, \dots, \mathbf{R}_{MM}^2])\mathbf{L}_1^T\mathbf{P}^T, \end{aligned}$$

where $\mathbf{L}_1$ is a unit-diagonal lower triangular matrix which satisfies

$$\mathbf{L}_1\text{diag}([\mathbf{R}_{11}, \mathbf{R}_{22}, \dots, \mathbf{R}_{MM}]) = \mathbf{R}^T.$$

Note that because of the GTD theory, the multiplicative majorization property

$$[\mathbf{R}_{11}^2, \mathbf{R}_{22}^2, \dots, \mathbf{R}_{MM}^2] \prec_{\times} [\sigma_1^2, \sigma_2^2, \dots, \sigma_M^2],$$

holds, where $[\sigma_1^2, \sigma_2^2, \dots, \sigma_M^2]$ are the eigenvalues of $\mathbf{R}_x$ with non-increasing order.

If we pass the signal $\mathbf{x}$ through the orthonormal matrix $\mathbf{P}^T$ to produce $\mathbf{z}$, i.e., $\mathbf{z} = \mathbf{P}^T\mathbf{x}$, the covariance of $\mathbf{z}$ is

$$\mathbf{R}_z = \mathbf{P}^T\mathbf{R}_x\mathbf{P} = \mathbf{L}_1\text{diag}([\mathbf{R}_{11}^2, \mathbf{R}_{22}^2, \dots, \mathbf{R}_{MM}^2])\mathbf{L}_1^T.$$

Therefore, $\mathbf{L}_1$ is the lower triangular matrix of the LDU form of $\mathbf{R}_z$. If now apply the PLT $\mathbf{L}_1^{-1}$ to the signal $\mathbf{z}$, the components of the resulting vector are decorrelated. The system is called GTD-TC, and is demonstrated in Fig. 4 for $M = 4$. Here we have used the MINLAB(I) structure [13]. The multipliers s_{km} are the entries of the matrix $\mathbf{L}_1^{-1}$. The bit loading formula becomes

$$b_i = b + \frac{1}{2} \log_2 \frac{\mathbf{R}_{ii}^2}{\det(\mathbf{R}_z)^{\frac{1}{M}}} = b + \frac{1}{2} \log_2 \frac{\mathbf{R}_{ii}^2}{\det(\mathbf{R}_x)^{\frac{1}{M}}}, \quad (10)$$

where we have used $\det(\mathbf{R}_z) = \det(\mathbf{P}^T\mathbf{R}_x\mathbf{P}) = \det(\mathbf{R}_x)$. The AM-MSE is invariant to the orthonormal matrix $\mathbf{P}$ at the decoder, therefore the AM-MSE is the same as the one for the PLT part for the transform coding of $\mathbf{z}$. As in eq. (6), the MSE is

$$\mathcal{E}_{GTD} = c2^{-2b} \det(\mathbf{R}_z)^{\frac{1}{M}} = c2^{-2b} \det(\mathbf{R}_x)^{\frac{1}{M}}, \quad (11)$$

which is the same as the MSE for KLT and PLT with optimal bit allocation. Note that this result is true because of the minimum noise structure for the PLT (which has unit noise gain).

Fig. 4. The GTD Transform coder implemented using MINLAB(I) structure.

We can regard $\mathbf{P}$ and $\mathbf{P}^T$ as the *precoder* and *postcoder*, and the system in between as the PLT part as indicated in the figure. Since there are infinitely many GTD realizations [8], this framework includes many transform coders that achieve the maximized coding gain. Actually it contains both the KLT and the PLT as special cases:

- 1) Suppose in (9), the GTD $\{\mathbf{Q}, \mathbf{R}, \mathbf{P}\}$ is taken as the SVD of $\mathbf{D}^{\frac{1}{2}}\mathbf{L}^T$:

$$\mathbf{D}^{\frac{1}{2}}\mathbf{L}^T = \mathbf{V}\Sigma^{\frac{1}{2}}\mathbf{U}^T.$$

In this case, we actually have $\mathbf{R}_x = \mathbf{L}\mathbf{D}\mathbf{L}^T = \mathbf{U}\Sigma\mathbf{U}^T$, thus $\mathbf{P} = \mathbf{U}$, which consists of the eigenvectors of the input covariance matrix. We also have $\mathbf{R}_z = \mathbf{U}^T\mathbf{R}_x\mathbf{U} = \Sigma$. In this case, the GTD-TC is reduced to the KLT. The PLT part in Fig. 4 is simply a series of scalar quantizers, and the optimal bit loading is according to the formula (2).

- 2) In (9), suppose $\{\mathbf{Q}, \mathbf{R}, \mathbf{P}\}$ is taken as the QR decomposition of $\mathbf{D}^{\frac{1}{2}}\mathbf{L}^T$. Since $\mathbf{D}^{\frac{1}{2}}\mathbf{L}^T$ is by itself an upper triangular matrix, we actually have $\mathbf{P} = \mathbf{I}$ and $\mathbf{Q} = \mathbf{I}$. In this case, the GTD-TC reduces to the original PLT-TC.

In the following, we will introduce three new transform coder schemes based on GTD theory.

A. Geometric Mean Decomposition – GMD

Geometric mean decomposition is a special case of the GTD. It arises in optimal transceiver design [12], [20]. In the GTD, if the diagonal terms of $\mathbf{R}$ are identical and equal to $(\prod_{i=1}^K \sigma_{h,i})^{\frac{1}{K}}$, then it is called GMD. GMD always exists for any matrix $\mathbf{H}$ since the multiplicative majorization property always holds. Suppose the GMD is used for the transform coder: in (9), $\mathbf{R}$ has all diagonal terms equal to $\bar{\sigma} = (\prod_{i=1}^M \sigma_i)^{\frac{1}{M}}$. The bit loading formula becomes

$$b_i = b + \frac{1}{2} \log_2 \frac{\bar{\sigma}^2}{\det(\mathbf{R}_x)^{\frac{1}{M}}} = b. \quad (12)$$

because $\det(\mathbf{R}_x) = \bar{\sigma}^{2M}$. The preceding equation says that all the quantizers are assigned the same number of bits. Note that since any GTD-TC achieves the same optimal performance, GMD-TC manifests a very good property – achieving the maximized coding gain without the need for bit loading.

B. Bi-Diagonal Transformation – Hessenberg Form

A matrix $\mathbf{B}$ is said to be bidiagonal if it has the form demonstrated below for the 4×4 case.

$$\mathbf{B} = \begin{bmatrix} b_{00} & b_{01} & 0 & 0 \\ 0 & b_{11} & b_{12} & 0 \\ 0 & 0 & b_{22} & b_{23} \\ 0 & 0 & 0 & b_{33} \end{bmatrix}$$

If the GTD form of $\mathbf{D}^{\frac{1}{2}} \mathbf{L}^T$ is $\mathbf{QBP}^T$, where $\mathbf{B}$ is a bi-diagonal matrix, then we call it the bi-diagonal transform coder (BID-TC). It can be seen that

$$\mathbf{R}_x = \mathbf{LDL}^T = \mathbf{PB}^T \mathbf{BP}^T,$$

where $\mathbf{B}^T \mathbf{B}$ is a tri-diagonal matrix.

The advantages of the BID-TC coder lie in its reduced computational complexity. To reduce a symmetric matrix to a tri-diagonal form by orthonormal transformation is computationally much less complex compared to eigenvalue decomposition [3]. The detail of reducing a symmetric matrix to the tri-diagonal form is discussed in [3], and requires only several Householder transformations. The LDU decomposition for a symmetric tri-diagonal matrix is also easy, which requires only $O(M)$ operations now, instead of $O(M^2)$ for general symmetric matrices. Therefore, the design cost for the BID-TC is less than KLT. Also, due to the bi-diagonal structure of $\mathbf{B}$, the implementation cost for the inner PLT part is also reduced, which is only in the order of $O(M)$. This can be seen in Fig. 5, which shows the MINLAB(I) structure for the BID-TC encoder. Signal feedforward paths are only required for the adjacent data streams. The number of signal feedforward paths is much less than for the original PLT.

The detail comparison between the design and implementation costs for various GTD based coders are summarized in Table I.

Fig. 5. The BID Transform coder implemented using MINLAB(I) structure.

IV. SIMULATIONS

In this section we provide the numerical simulations for GTD based coders. The signal $\mathbf{x}$ is generated by a zero mean Gaussian vector process with prescribed covariance matrix $\mathbf{R}_x$. The number of data streams $M = 8$ in the experiments. Uniform roundoff quantizers are assumed. Each quantizer adapts its step size according to the variance of the Gaussian input (pp. 818 of [17]). For each case, we run the Monte Carlo simulations for calculating the AM-MSE. In each trial, we first generate the input covariance matrix by multiplying a fixed diagonal matrix with a randomly generated orthonormal matrix on the left and its transpose on the right. The input vector $\mathbf{x}$ is then generated according to this covariance matrix. In the following we provide simulation comparisons of different transform coders with and without optimal bit allocation.

Optimal bit allocation: Fig. 6 compares the AM-MSE performance of different transform coders with optimal bit allocation, for input covariance matrix with high and low condition numbers, respectively. “Transform-wBL” means we adopt the specified transform with optimal bit loading. For example “KLTwBL” uses the KLT with the bit loading formula (2). “PLTwBL” is the method mentioned in [13], with the optimal bit loading formula (5). “UNCwBL” is the case when we have no transform; we directly quantize the input $\mathbf{x}$ with optimal bit allocation

$$b_i = b + \frac{1}{2} \log_2 \frac{\sigma_{x,i}^2}{(\prod_{k=1}^M \sigma_{x,i}^2)^{\frac{1}{M}}}.$$

Since the input to the quantizers x_i are correlated to each other in general, direct scalar quantization without transformation results in performance loss compared to the GTD-TCs even when the optimal bit loading scheme is applied. “BIDwBL” is the bi-diagonal transform coder discussed in III-B. The bit loading formula is as in (10). “GMDTFC” is the GMD transform coder. Since the signal variance in each data stream is the same, no bit loading is needed. This allows us to build the same scalar quantizers for all data streams. It can be seen from the figure that with optimal bit loading, all GTD-TCs perform about the same. This is consistent with the analysis made in Sec. III. Direct quantization without transforms (UNCwBL) results in about 5 bits per data stream performance loss for Fig. 6.

Fig. 6. Performance of different transform coders with optimal bit allocation. Input covariance matrix has high condition number (10^4).

TABLE I
DESIGN AND IMPLEMENTATION COSTS OF TRANSFORM CODERS

	Design cost	Impl. cost(precoder part)	Impl. cost(PLT part)
KLT	EVD, $O(M^3)$	$O(M^2)$	0
PLT	LDU, $O(M^2)$	0	$O(M^2)$
GMD-TC	EVD and GMD [8], $O(M^3)$	$O(M^2)$	$O(M^2)$
BID-TC	Hessenberg form $O(M^3)$ and easy LDU $O(M)$	$O(M^2)$	$O(M)$
General GTD-TC	EVD and GTD [8], $O(M^3)$	$O(M^2)$	$O(M^2)$

Uniform bit allocation: Fig. 7 compares the AM-MSE performance of different transform coders with uniform bit allocation, for input covariance matrix with high and low condition numbers, respectively. Here “transform-nBL” means we adopt some specific transform with no optimal bit loading, i.e., we allocate the same number of bits to each data stream. However, the step size of each scalar quantizer is adapted according to variance of the Gaussian input (P.818 of [17]). “KLTnBL” uses KLT for the transform. “PLTnBL” is the method mentioned in [13] but with no bit loading. “UNCnBL” is the case when we have no transform but directly quantize the input $\mathbf{x}$. No bit loading is applied either. “BIDwBL” is the bi-diagonal transform coder discussed in III-B with no bit loading. “GMDTFC” is the GMD transform coder. It can be seen from the figure that with no bit loading applied, GMD performs much better than the other methods, since the GMD without bit allocation is theoretically as good as the other methods with optimal bit allocation.

Fig. 7. Performance of different transform coders with uniform bit allocation. Input covariance matrix has high condition number (10^7).

In the simulation results, the reader will notice that for values of b (average number of bits) exceeding three (low condition number case), and exceeding six (for high condition number case), the theoretical predictions are indeed verified to be true.⁵ Namely, with no bit allocation, GMD performs much better than KLT, PLT, and the BID. These later methods with no bit allocation have performance comparable to direct quantization. Furthermore, with optimal bit allocation, all these methods (GMD, KLT, and BID) have identical performances. For small values of b these theoretical predictions (which are based on the high bit rate assumption) are seen to be (understandably) less and less true.

⁵It should be mentioned here that such relatively large values for b are not uncommon in areas such as multispectral image compression [15].

V. CONCLUSIONS

The main purpose of the paper has been to provide a general framework for a family of linear transform coders based on the GTD. The GTD has in the past been found to be of great importance in digital transceiver optimization, but has hitherto not been considered for transform coding. The KLT and PLT transforms are special cases belonging to the GTD transform coder family. Some of the new transform coders that have been presented as members of this family include the GMD and the BID coders. The BID has the advantage that the computational complexity of the PLT part is significantly less. The GMD has the property that no bit allocation is needed in order to achieve the optimal coding gain, which the KLT and PLT can only achieve this gain with the help of bit allocation.

REFERENCES

- [1] A. N. Akansu and Y. Liu, “On signal decomposition techniques,” *Optical engr.*, vol. 30, pp.912 - 920, Jul. 1991.
- [2] M. Effros, H. Feng, and K. Zeger, “Suboptimality of Karhunen-Loeve transformation for transform coding,” *IEEE Trans. Info. Theory.*, vol. 50, No. 8, pp.1605 - 1619, Aug. 2004.
- [3] G. H. Golub, and C. F. Van Loan, “Matrix computations,” The Johns Hopkins Univ. Press 1996.
- [4] R. A. Horn, and C. R. Johnson, *Topics in matrix analysis*, Cambridge Univ. Press 1991.
- [5] R. A. Horn, and C. R. Johnson, *Matrix analysis*, Cambridge Univ. Press 1985.
- [6] N. S. Jayant, and P. Noll, *Digital coding of waveforms*, Englewood Cliffs, NJ: Prentice-Hall, 1984.
- [7] Y. Jiang, J. Li, and W. W. Hager, “Joint transceiver design for MIMO communications using geometric mean decomposition,” *IEEE Trans. Sig. Proc.*, pp.3791-3803, Oct. 2005.
- [8] Y. Jiang, W. W. Hager, and J. Li, “Generalized triangular decomposition,” *Mathematics of computation.*, Oct. 2007.
- [9] S. Mallat, and F. Falzon, “Analysis of Low Bit Rate Image Transform Coding,” *IEEE Trans. Sig. Proc.*, vol. 46, No. 4, pp.1027 - 1042, Apr. 1998.
- [10] A. W. Marshall, and I. Olkin, *Inequalities: Theory of majorization and its applications*, Academic Press. 1979.
- [11] D. L. Mary, and D. T. M. Stock, “A theoretical high-rate analysis of causal versus unitary online transform coding,” *IEEE Trans. Sig. Proc.*, vol. 54, No. 4, pp.1472 - 1482, Apr. 2006.
- [12] D. P. Palomar, and Y. Jiang, “MIMO transceiver design via majorization theory,” *Foundations and Trends in Communications and Information Theory*, vol. 3, issue 4, pp. 331-551, Nov. 2006.
- [13] S. M. Phoong, and Y. P. Lin, “Prediction-based lower triangular transform,” *IEEE Trans. Sig. Proc.*, vol. 48, No. 7, pp.1947 - 1955, Jul. 2000.
- [14] S. M. Phoong, and Y. P. Lin, “MINLAB: minimum noise structure for ladder-based biorthogonal filter banks,” *IEEE Trans. Sig. Proc.*, vol. 48, No. 2, pp.465 - 476, Feb. 2000.
- [15] J. A. Saghri, A. G. Tescher, and J. T. Reagan, “Practical transform coding of multispectral imagery,” *IEEE Signal Processing Magazine*, pp.32 - 43, Jan. 1995.
- [16] P. P. Vaidyanathan, “Theory of optimal orthonormal subband coders,” *IEEE Trans. Sig. Proc.*, vol. 46, pp.1528 - 1543, Jun. 1998.
- [17] P. P. Vaidyanathan, *Multirate systems and filter banks*, Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [18] P. P. Vaidyanathan, *The theory of linear prediction*, Morgan & Claypool publishers, 2008.
- [19] C. C. Weng, C. Y. Chen, and P. P. Vaidyanathan, “MIMO Transceivers with Decision Feedback, Bit Loading and Limited Feedback: Theory and Optimization,” *submitted to IEEE Trans. Sig. Proc.*
- [20] J. K. Zhang, A. Kavcic, and K. M. Wong, “Equal-diagonal QR decomposition and its application to precoder design for successive-cancellation detection,” *IEEE Trans. Info. Theory*, pp. 154 - 172, Jan. 2005.