Joint Optimum Bitwise Decomposition of any Memoryless Source to be Sent over a BSC

Seyed Bahram Zahir Azami[1], Pierre Duhamel[2] and Olivier Rioul[3]

École Nationale Supérieure des Télécommunications
URA CNRS 820
[1,3]Communications and electronics department
[2]Signal and image processing department
46, rue Barrault, 75634 Paris Cedex 13, France
emails: [1]*zahir@com.enst.fr*, [2]*duhamel@sig.enst.fr*, [3]*rioul@com.enst.fr*

Abstract

An arbitrary memoryless source is considered to be sent over a binary symmetric channel. We develop an algorithm, with some similarities to the Lloyd algorithm, capable to optimally make a bitwise decomposition. For a uniform source, this turns to simple quantification with a soft decoding. For other kinds of sources, however, the algorithm has a more important effect. For example, for a Gaussian source, a permutation of indices is imposed. The algorithm is efficient for small number of decomposing bits, say below 6 bits per sample and for the generalized Gaussian sources with a large decay rate ($\theta \geq 2$).

1 Introduction

To transmit a continuous source over a binary channel, a quantizer is needed somewhere in the transmission chain. A **scalar quantizer** (SQ) is considered in this chapter which can be considered as providing a bitwise decomposition of the source. We have already studied the situation with a uniform source in a decomposition of the bits in [5].

As proved in 1969 by Crimmins *et al.* [1] and confirmed by McLaughlin *et al.* in 1995 [4], for a uniform source, the optimum scalar quantization is obtained by a **natural binary coding** (NBC). However, this is not necessarily the case for other kinds of sources. In this paper, this problem is treated for an arbitrary memoryless source.

The first hypothesis to be considered here is to have a memoryless source. However, this is not a very restricting hypothesis. In fact, the considered source is the output of a transformation which can theoretically produce a memoryless output.

Kurtenbach and Wintz have proposed an algorithm to design an optimum encoder/decoder pair with respect to the mean Euclidean distance between the source and the reconstructed signal, in 1969 [3]. In their proposed algorithm, as well as in channel optimized scalar quantization (COSQ) proposed by Farvardin and Vaishampayan [2], the decision thresholds can take any value. This hypothesis is different from that assumed here, *i.e.*, the decision thresholds are chosen such that the quantization intervals are equiprobable, and consequently the decomposed bits become independent.

Hence, despite the recommendation of Kurtenbach and Wintz [3], and that of Farvardin and Vaishampayan [2], in this paper, our second hypothesis is to consider **equiprobable quantization intervals**.

Moreover, it is imposed for the total distortion to be a weighted sum of the distortions due to the different decomposed bits. Having a simple mathematical expression that gives the quantization levels in function of the channel outputs, this condition is easily satisfied.

In this paper, we propose an algorithm which is as a special type of Lloyd's algorithm. The first element of the algorithm consists of updating the quantizer, assuming a fixed dequantizer and the second element is vice versa.

This paper is organized as follows: first we introduce the model, the notations and the hypothesis. Then, we propose our algorithm with its optimality proof. Finally, we give some simulation results.

2 General formalization

In this section we present the global model considered in this paper and introduce its important parts. Also, we give the constraints that we consider in each part and express why these constraints are taken into account.

A source with arbitrary known PDF is considered. Yet, for simplicity, it is assumed that the source has zero mean value. Remark that this does not affect the generality of the presented proofs or algorithms. Figure 1 reflects the configuration considered in this paper.

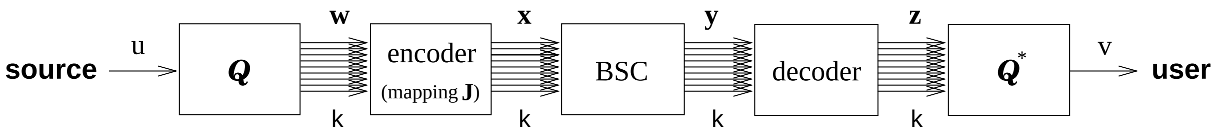

Figure 1: Transmission scheme considered in this paper.

2.1 Quantizer ($\mathcal{Q}_\mathbf{J}$)

Restriction 1: it is assumed that the intervals are equiprobable

$$p(\mathbf{w}) = (\frac{1}{2})^k \qquad \forall \mathbf{w}$$

the quantization rule is defined as follows

$$\mathrm{u} \in \mathcal{I}_\mathrm{w} \qquad \Longleftrightarrow \qquad \mathbf{w} = [\mathrm{w}_k \ \ldots \ \mathrm{w}_2 \ \mathrm{w}_1]$$

where $\mathrm{w}_i \in \{0, 1\}$ and $[\mathrm{w}_k \ \ldots \ \mathrm{w}_2 \ \mathrm{w}_1]$ is the binary representation of $\mathbf{w}$. The assumption of the equiprobability hypothesis results in

$$p(\mathrm{w}_i) = \frac{1}{2} \quad \forall i \quad \Rightarrow \quad p(\mathbf{w}) = \prod_{i=1}^{k} p(\mathrm{w}_i) = (\frac{1}{2})^k$$

In other words, the different decomposed bits are **independent** from each other. Note that they are also independent in time since the source is assumed to be memoryless.

2.1.1 Interval centroids

Without any constraint on the reconstruction levels and for a noise-free channel, the optimum reconstruction levels can be found by the centroid principle

$$\phi_\mathrm{i} = \mathbf{E}(\mathrm{U}|\mathrm{U} \in \mathcal{I}_\mathrm{i})$$

In vector form, Φ is introduced as the set of all centroids, in an increasing order, or simply the **centroid vector**

$$\Phi = [\phi_0 \ \phi_1 \ \phi_2 \ \ldots \ \phi_{2^k-1}]^t$$

where the t superscript denotes the transpose.

2.1.2 Permutation

In matrix form, the **permutation matrix**, $\mathbf{J}$ is introduced as a row-permuted version of the unity matrix, $\mathbf{I}_{2^k \times 2^k}$

$$\mathbf{J}_{\mathbf{w},\mathbf{i}} = \left\{ \begin{array}{lll} 1 & : & \mathcal{Q}(\mathrm{u}) = \mathbf{i} \quad \& \quad \mathcal{Q}_\mathbf{J}(\mathrm{u}) = \mathbf{w} \\ 0 & : & \text{otherwise} \end{array} \right.$$

Where $\mathcal{Q}(\mathrm{u})$ and $\mathcal{Q}_\mathbf{J}(\mathrm{u})$ represent respectively the non permuted and the permuted version of the quantized source sample, u, by the permutation matrix $\mathbf{J}$. Then $\Phi_\mathbf{J}$ can be written as a permuted version of Φ

$$\Phi_\mathbf{J} = \mathbf{J}.\Phi$$

Restriction 2: In order to keep a simple mathematical expression between the channel outputs and the reconstruction levels, a constraint is imposed on them as in the following equation

$$\mathrm{v} = \psi = \sum_{i=1}^{k} \alpha_i \mathrm{z}_i$$

In matrix form, the expression becomes

$$\begin{array}{rcl} \mathrm{v} = \psi & = & \mathbf{z}^t.\boldsymbol{\alpha} = \boldsymbol{\alpha}^t.\mathbf{z} \\ \Psi & = & \mathbf{B}^t.\boldsymbol{\alpha} \end{array} \qquad (1)$$

where Ψ is the reconstructed centroid vector, $\boldsymbol{\alpha} = [\alpha_k \ \ldots \ \alpha_2 \ \alpha_1]^t$, $\mathbf{z} = [\mathrm{z}_k \ \ldots \ \mathrm{z}_2 \ \mathrm{z}_1]^t$ and $\mathbf{B}$ is the natural binary encoding matrix (A matrix whose successive rows are the binary representations of the natural numbers $0, 1, \cdots, 2^k - 1$).

This constraint gives the conditions to prove the additivity of distortion:

$$D = \sum_{i=1}^{N} \alpha_i^2 \mathbf{E}\left((\mathrm{u}_i - \mathrm{v}_i)^2\right) = \sum_{i=1}^{N} \alpha_i^2 D_i$$

where $D_i = \mathbf{E}\left((\mathrm{u}_i - \mathrm{v}_i)^2\right)$ is the MSE corresponding to the i-th bit, and depends on the channel transition probability, p.

2.2 Channel model

Here we consider ±1 values as in figure 2, in order to simplify our calculations.

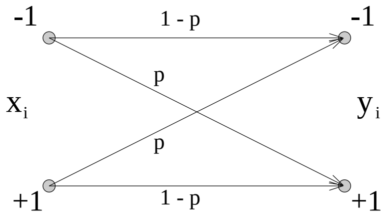

Figure 2: Binary symmetric channel model with ±1 outputs.

It is further assumed that the z_i parameters, *i.e.*, the decoder reconstruction outputs, have **soft values**. This assumption is considered because we aim to minimize the MSE distortion measure and not the error probability.

3 Lloyd type algorithm

We propose an algorithm which can be viewed as a special type of Lloyd-Max algorithm to optimize the distortion due to source and channel. We aim to minimize the average Euclidean distance:

$$D = \mathbf{E}(\mathrm{U} - \mathrm{V})^2$$

Like other algorithms of Lloyd type, this algorithm consists of two elements: optimizing the quantizer, assuming a known dequantizer and vice versa.

3.1 Optimizing the dequantizer assuming a fixed quantizer

In this part, the distortion is minimized, due to the source and to the channel, assuming a known quantizer, $\mathcal{Q}_{\mathbf{J}}$, by optimizing the parameters of the dequantizer, $\mathcal{Q}^*$, *i.e.*, the $\boldsymbol{\alpha}$ parameters.

$$D = \mathbf{E}\Big((\mathrm{U} - \phi) - (\mathrm{V} - \phi)\Big)^2 =$$

$$\underbrace{\mathbf{E}\Big((\mathrm{U} - \phi)^2\Big)}_{D_q} + \underbrace{\mathbf{E}\Big((\mathrm{V} - \phi)^2\Big)}_{\mathbf{E}\Big((\sum_{i=1}^{k} \alpha_i \mathrm{z}_i - \phi)^2\Big)} \; - \underbrace{2\mathbf{E}\Big((\mathrm{U} - \phi)(\mathrm{V} - \phi)\Big)}_{2A=0}$$

For the first term, let us define $D_q \equiv \mathbf{E}\Big((\mathrm{U} - \phi)^2\Big)$. This is the distortion purely due to the quantization and is independent of the dequantizer. The nullity of the third term can be easily proved.

So we rewrite the above equation, derivate it with respect to $\boldsymbol{\alpha}$ and set it equal to zero

$$\begin{aligned} \frac{\partial}{\partial \boldsymbol{\alpha}^t}\mathbf{E}\Big((\mathrm{V} - \phi)^2\Big) &= \mathbf{E}\Big(\frac{\partial}{\partial \boldsymbol{\alpha}^t}(\mathrm{V} - \phi)^2\Big) \\ = 2\mathbf{E}\Big((\mathrm{V} - \phi)\frac{\partial \mathrm{v}}{\partial \boldsymbol{\alpha}^t}\Big) &= 2\mathbf{E}\Big((\mathrm{V} - \phi)\mathbf{z}\Big) = \mathbf{0} \end{aligned}$$

Using equation (1) and the orthogonality principle yields

$$\underbrace{\mathbf{E}(\mathbf{z}.\mathbf{z}^t)}_{\mathbf{R}}.\boldsymbol{\alpha} = \underbrace{\mathbf{E}(\phi.\mathbf{z})}_{\mathbf{r}} \qquad (2)$$

hence it can be concluded the optimum value for $\boldsymbol{\alpha}$ is

$$\boldsymbol{\alpha}^* = \mathbf{R}^{-1}.\mathbf{r} \qquad (3)$$

3.1.1 Calculation of minimized distortion

With $\mathrm{V}^* = \mathbf{z}^t.\boldsymbol{\alpha}^* = \mathbf{z}^t.\mathbf{R}^{-1}.\mathbf{r} = \mathbf{z}^t.\mathbf{R}^{-1}.\mathbf{E}(\phi.\mathbf{z})$, we can write

$$\begin{aligned} D_{\min} &= D_q + \mathbf{E}\Big((\mathrm{V}^* - \phi)^2\Big) \\ &= D_q + \mathbf{E}\Big((\mathrm{V}^* - \phi).(\mathbf{z}^t.\boldsymbol{\alpha}^* - \phi)\Big) \end{aligned}$$

The term $\mathbf{E}\Big((\mathrm{V}^* - \phi).\mathbf{z}^t.\boldsymbol{\alpha}^*\Big)$ is equal to zero because of the orthogonality principle.

$$\begin{aligned} D_{\min} &= D_q + \mathbf{E}\Big((\phi - \mathrm{V}^*).\phi\Big) \\ &= D_q + \mathbf{E}\Big((1 - \mathbf{z}^t.\mathbf{R}^{-1}.\mathbf{z}).\phi^2\Big) \\ &= D_q + \frac{1}{2^k}\Phi^t.\Phi - \mathbf{r}^t.\mathbf{R}^{-1}.\mathbf{r} \end{aligned}$$

Remark: Equation (4) gives a simple way to calculate the distortion as it can be remarked that $\sigma_u^2 = D_q + \frac{1}{2^k}\Phi^t.\Phi$

$$D_{\min} = \sigma_u^2 - \mathbf{r}^t.\mathbf{R}^{-1}.\mathbf{r} \qquad (4)$$

3.1.2 Calculation of $\mathbf{R}, \mathbf{r}$

It can be shown that for these conditions we have:

$$\begin{aligned} \mathbf{R} &= \Big[\mathbf{E}(\mathrm{z}_i.\mathrm{z}_j)\Big]_{i,j} = \mathbf{I}_{(k \times k)} \\ \mathbf{r} &= \Big[\mathbf{E}(\mathrm{z}_i.\phi)\Big]_i = \frac{1}{2^k}\Big(\mathbf{J}.\widehat{\mathbf{B}}\Big)^t.\Phi = \frac{1}{2^k}\widehat{\mathbf{B}}^t.\mathbf{J}^t.\Phi \end{aligned}$$

where $\widehat{\mathbf{B}}_{i,j} = \mathbf{E}_{\text{channel}}(\mathrm{z}|\mathrm{w} = \mathbf{B}_{i,j})$.

$$\begin{aligned} \mathbf{B}_{i,j} &= \begin{cases} +1 \Rightarrow \widehat{\mathbf{B}}_{i,j} = 1(1-p) - 1p = 1 - 2p \\ -1 \Rightarrow \widehat{\mathbf{B}}_{i,j} = -1(1-p) + 1p = -1 + 2p \end{cases} \\ \widehat{\mathbf{B}} &= (1 - 2p)\mathbf{B} \end{aligned}$$

With these values of $\mathbf{R}$ and $\mathbf{r}$, we rewrite the expression for $\boldsymbol{\alpha}$ as

$$\boldsymbol{\alpha} = \mathbf{r} = \frac{1}{2^k}\widehat{\mathbf{B}}^t.\mathbf{J}^t.\Phi$$

3.2 Optimizing the quantizer assuming a fixed dequantizer

The second part of algorithm consists in optimizing the quantizer, $\mathcal{Q}_{\mathbf{J}}$, assuming a known dequantizer, $\mathcal{Q}^*$. Since the equiprobability constraint must be respected, the only flexibility is in the permutation of indices. The following lemma proves that a permutation sorting the reconstruction levels in an increasing order is the optimum permutation.

Lemma 1 *For a known decoder, the optimum encoder is the one which sorts the decoded centroids in an increasing magnitude order.*

4 Complete algorithm

The complete algorithm of bitwise decomposition is described below for a memoryless source $\mathbf{U}$, either knowing its PDF, or having sufficient number of its samples. Both versions of the algorithm are depicted in the block diagram of figure 3. The only difference of these two versions is in their initializations.

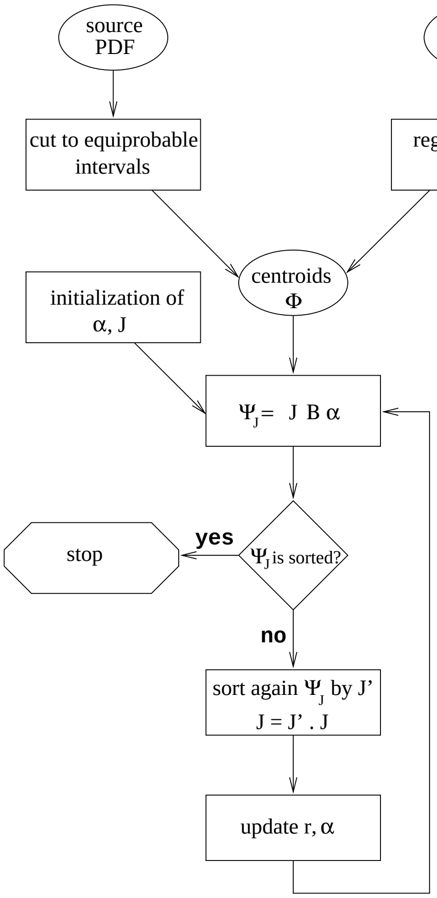

Figure 3: Block diagram of the optimized bitwise decomposition algorithm for either a known PDF or a source with enough known samples.

In practice, to avoid the local minima, the algorithm may be applied several times with different initializations and the best obtained results will be taken. For small values of N this is less important, as usually the global optimum is obtained rapidly. For larger values of N, more initializations is suggested.

5 Practical experiments

In order to test the performance of the bitwise decomposition algorithm, the generalized Gaussian distribution is considered as the source model. This distribution is given below

$$\begin{aligned} p(\mathrm{u}) &= \frac{\theta\eta(\theta,\sigma)}{2\Gamma(\frac{1}{\theta})}\exp\Big(-\Big(\eta(\theta,\sigma)|u|\Big)^{\theta}\Big) \\ \eta(\theta,\sigma) &= \sigma^{-1}\Big(\frac{\Gamma(\frac{3}{\theta})}{\Gamma(\frac{1}{\theta})}\Big)^{\frac{1}{2}} \end{aligned}$$

with $\theta > 0$ describing the exponential rate of decay, σ a positive quantity representing a scale parameter, and $\Gamma(.)$ being the Gamma function. The variance of the associated random variable is given by $\sigma_u^2 = \sigma^2$.

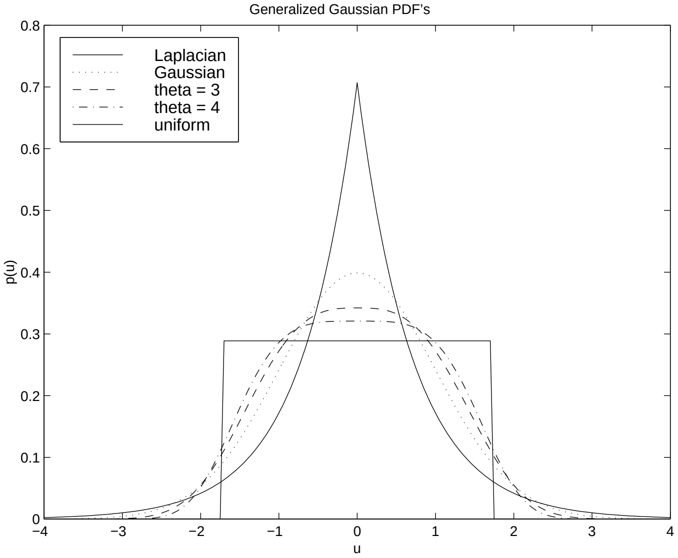

Figure 4: PDF of Generalized Gaussian sources with $\theta = 1, 2, 3, 4, \infty$.

It is known that $\theta = 1, 2$ and ∞ correspond to Laplacian, Gaussian and uniform sources, respectively. See figure 4, where these distributions are illustrated for $\sigma_u^2 = 1$. In the following subsections, some sources are chosen to be treated with the proposed bitwise decomposition algorithm. The performances and results will be presented.

5.1 Uniform source

For a uniform source, the optimization yields the same results obtained analytically by Crimmins *et al.*, in 1969 [1], using Fourier transform on finite Abelian groups. More recently, in 1995, McLaughlin *et al.* [4], proved the same theorem, using a vector for formulating the index assignment and showing that the MSE due to channel errors is determined by the projections of its columns onto the eigenspaces of the multidimensional channel transition matrix. This result is a natural binary coding (no permutation) and $\alpha_i \propto 2^i$.

5.2 Gaussian source

The Gaussian source is studied here, as a second example with $k = 4$ bits per sample and $p = 10^{-2}$. Contrarily to the uniform source case, here, we have not a natural binary permutation. Figure 5 depicts the quantization tree. It is obvious that with these α_is, the optimum permutation is not the natural one.

We obtain $\boldsymbol{\alpha} = [0.6206\ 0.5332\ 0.4351\ 0.2810]^t$ and an SNR equal to 12.05 dB.

5.3 Generalized Gaussian source

In this paper, some generalized Gaussian sources with different values of θ have been studied as examples. Fig-

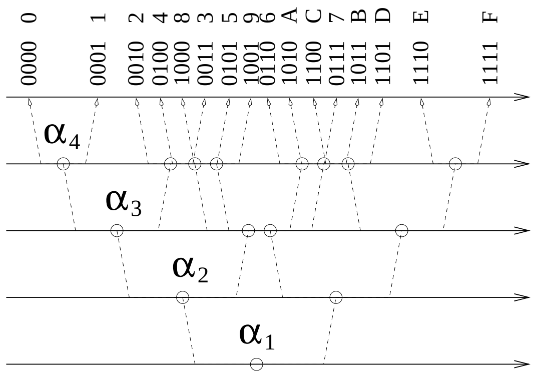

Figure 5: 4 bits Quantization tree for a Gaussian source: permutation is not natural.

ure 6 depicts a global comparison of all of the performance bounds obtained for some generalized Gaussian sources with various values of θ, using the Lagrangian method as explained in [7].

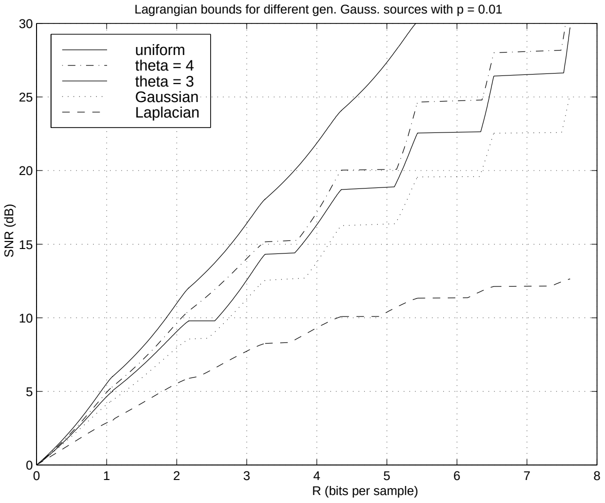

Figure 6: Lagrangian performance bound for some generalized Gaussian PDF's for $p = 10^{-2}$.

In the following section, some concluding remarks are provided about the bitwise decomposition algorithm.

6 Conclusion

The bitwise decomposition algorithm provides an extremely simple system. All the parameters of this system consist of $2^N - 1$ quantization thresholds, a 2^N length mapping table in the encoder and an N size $\boldsymbol{\alpha}$ vector in the decoder.

For all kinds of sources, the system performance is good for relatively small values of N, say for N below 5 or 6 bits. We observed that for example with a Gaussian source and values of N up to 4 bits, the performance of the simple bitwise decomposition system is very close to its optimum value (the Lagrangian bound). The same statement holds for other kinds of sources with this algorithm.

However, the performance begins to saturate for $N > 6$ bits (except for the uniform source). In general, the performance obtained by bitwise decomposition algorithm for the generalized Gaussian sources is better when θ has a big value. For small values of θ, such as Laplacian source ($\theta = 1$), the performance is not very good for $N > 4$ bits per sample.

In fact, the algorithm tries to represent 2^N centroids just by N parameters of the $\boldsymbol{\alpha}$ vector. For small values of N this is possible but for larger values of N, this becomes more difficult (except for the uniform source which is a particular case).

A possible improvement of this system can be obtained by including a channel coding scheme to the system. For example, a hierarchical protection may be implemented. This is the subject of our future research and is a generalization of [6] from a uniform source to an arbitrary generalized Gaussian source.

References

1. T.R. Crimmins, H.M. Horwitz, C.J. Palermo, and R.V. Palermo, *"Minimization of Mean-Squared Error for Data Transmitted Via Group Codes"*, IEEE Transactions on Information Theory **15** (1969), no. 1, 72–78.

2. N. Farvardin and V. Vaishampayan, *"Optimal Quantizer Design for Noisy Channels : An Approach to Combined Source-Channel Coding"*, IEEE Transactions on Information Theory **33** (1987), no. 6, 827–838.

3. A.J. Kurtenbach and P.A. Wintz, *"Quantizing for Noisy Channels"*, IEEE Transactions on Communication Technology **17** (1969), no. 2, 291–302.

4. S. W. McLaughlin, D.L. Neuhoff, and J.J. Ashley, *"Optimal Binary Index Assignment for a Class of Equiprobable Scalar and Vector Quantizers"*, IEEE Transactions on Information Theory **41** (1995), no. 6, 2031–2037.

5. S.B. ZahirAzami, P. Duhamel, and O. Rioul, *"Combined Source-Channel Coding for Binary Symmetric Channels and Uniform Memoryless Sources"*, Proceedings of the *Seizième Colloque sur le traitement du Signal et des Images* (GRETSI) (Grenoble, France), Sep. 1997.

6. ———, *"Joint Source-Channel Coding of Memoryless Sources Over Binary Symmetric Channels"*, Proceedings of the IEEE Global Telecommunications Conference (Sydney, Australia), Nov. 1998, pp. 3614–3619.

7. S.B. ZahirAzami, O. Rioul, and P. Duhamel, *"Performance Bounds for Joint Source-Channel Coding of Uniform Memoryless Sources Using a Binary Decomposition"*, Proceedings of European Workshop on Emerging Techniques for Communication Terminals (COST) (Toulouse, France), July 1997, pp. 259–263.