Joint Source-Channel Coding at the Application Layer

O. Y. Bursalioglu¹, M. Fresia², G. Caire¹, and H. V. Poor²
¹ Ming Hsieh Dept. of Electrical Engineering, University of Southern California, Los Angeles, CA.
² Dept. of Electrical Engineering, Princeton University, Princeton, NJ.

Abstract

The multicasting of an independent and identically distributed Gaussian source over a binary erasure broadcast channel is considered. This model applies to a one-to-many transmission scenario in which some mechanism at the physical layer delivers information packets with losses represented by erasures, and users are subject to different erasure probabilities. The reconstruction signal-to-noise ratio (SNR) region achieved by concatenating a multiresolution source code with a broadcast channel code is characterized and four convex optimization problems corresponding to different performance criteria are solved. Each problem defines a particular operating point on the dominant face of the SNR region. Layered joint source-channel codes are constructed based on the concatenation of embedded scalar quantizers with binary raptor encoders. The proposed schemes are shown to operate very close to the theoretical optimum.

I. INTRODUCTION

Multimedia streaming to wired and wireless/mobile users is one of the most important and rapidly growing applications in the modern Internet. A conventional approach consists of establishing individual streaming sessions from the server to each user. This may be very inefficient when many users wish to receive the same content at the same time (e.g., in mobile television applications). In contrast, conventional analog broadcasting systems send simultaneously the same signal to a potentially unlimited number of receivers, with possibly different reconstruction quality that depends on the channel conditions.

In order to achieve such scalability and “graceful degradation” in a digital network, in this paper we consider a system in which the server generates coded information packets that are multicast to an arbitrarily large number of users. We disregard the details of the physical, link and network layers of the underlying heterogeneous network and model the system, seen at the application layer, as an erasure broadcast channel [1]. To cope with erasures at the application layer, Forward Error Correcting (FEC) codes are considered [2].

We consider an idealized information-theoretic model for the above scenario (see Section II), we investigate its fundamental limits, and provide explicit code constructions that perform remarkably close to the limits with low complexity. In previous work [3], [4], we proposed a class of joint source-channel codes (JSCCs) for the end-to-end quadratic distortion criterion based on the concatenation of transform coding, scalar quantization, and linear channel coding. The latter maps directly the redundant quantization bits onto channel-encoded symbols, eliminating the need for an entropy coding stage, which is typically very fragile to post-decoding residual errors. At the receiver, a Belief-Propagation (BP) iterative decoding algorithm combines the a-priori source probability distribution

The research was supported in part by the National Science Foundation under Grants CNS-06-25637 and CNS-07-22073.
with the observed channel output and produces soft log-likelihood ratios (LLRs) for the quantization bits, which are then used for robust soft-bit reconstruction.

Here we extend our JSCC scheme to the case of multicasting over the broadcast erasure channel. We use multi-resolution (embedded) quantization and layered JSCC, in order to serve simultaneously several users at different channel capacities (i.e., different erasure probabilities). The allocation of source layers to different users can be optimized according to different criteria, as discussed in Section III. Performing close to the theoretical limits requires coding rate adaptation with very fine granularity. In the proposed scheme (see Section IV), this is accomplished by using raptor encoders [5], which are able to produce any arbitrary number of coded symbols “on demand” with a single, low complexity, encoding machine.

Recently, raptor codes [5] have been standardized as application layer FEC coding for Multimedia Broadcast/Multicast Services (MBMS) within 3GPP [6]. We hasten to say that our application is very different from this standard. In 3GPP [6], the “static broadcasting” [7] of a common content file is considered, where each user gathers enough channel observations until it has received enough non-erased symbols such that the whole file can be perfectly decoded. Users may have different decoding delays depending on their erasure probabilities. In contrast, in this work we consider a truly “real-time” multicasting where the decoding delay is the same for all users, but each user reconstructs the source at a possibly different distortion level, depending on its own channel capacity. The numerical results of Section IV show that the proposed coding scheme can achieve end-to-end distortion performance very close to the theoretical limits with finite block length and low encoding/decoding complexity.

II. GAUSSIAN SOURCE OVER BINARY ERASURE BROADCAST CHANNEL

The Binary Erasure Broadcast Channel (BEBC) has input alphabet \( \{0, 1\} \), output alphabet \( \{0, 1, e\} \), where “e” denotes an erasure, and is defined by \( L \) channel transition probabilities \( P^{(l)}(y|x) = 1 - \epsilon_l \) for \( y = x \), \( P^{(l)}(e|x) = \epsilon_l \) for \( y = e \) and \( P^{(l)}(y|x) = 0 \) for \( y \neq x, e \), where \( L \) is the number of the different binary erasure channels (BEC) considered.

Without loss of generality, we let \( \epsilon_1 \geq \ldots \geq \epsilon_L \) and denote by \( C_l = 1 - \epsilon_l \) the capacity of the \( l \)-th binary erasure channel (BEC). This channel serves as a simple model for a multicast network where an arbitrarily large number of users are grouped into \( L \) classes, with different channel qualities characterized by their erasure probabilities. For simplicity, we shall refer to each class as a “user” since, in a multicast scenario, all users belonging to the same class are indistinguishable and achieve the same performance.

We consider the transmission of an i.i.d. Gaussian source vector \( S \in \mathbb{R}^k \), with components \( S_l \sim \mathcal{N}(0, 1) \). The encoder maps \( S \) into \( X = f^{(k)}(S) \), where \( f^{(k)} : \mathbb{R}^k \to \{0, 1\}^n \) is a suitable encoding function. At each \( l \)-th user decoder, the received channel output \( Y_l \) is mapped into a reconstructed source vector \( \hat{S}_l = g_l^{(k)}(Y_l) \), where \( g_l^{(k)} : \{0, 1, e\}^n \to \mathbb{R}^k \) is a suitable decoding function. We define the average quadratic distortion \( D_l^{(k)} = \frac{1}{k} \mathbb{E}[\|S - \hat{S}_l\|^2] \), where the expectation is with respect to the joint \( k \)-dimensional probability distribution of \((S, \hat{S}_l)\) induced by the source, by the channel erasures and by the coding scheme. The ratio \( b = n/k \) (channel uses per source symbol) is commonly referred to as the “bandwidth expansion” factor of the system. The reconstruction signal-to-noise ratio (SNR) is defined as \( \text{SNR}_l^{(k)} = -10 \log_{10} D_l^{(k)} \), where \( D_l^{(k)} \) is the distortion for the
A multiresolution source encoder creates 3 layers with \( \text{SNR}_1 < \text{SNR}_2 < \text{SNR}_3 \) for three classes of users with capacities \( C_1 \leq C_2 \leq C_3 \). The multicast scheme is designed such that each user receives an extra layer with respect to previous user. The channel encoder-decoder pair at rate \( C \) is shown as blocks \( CE(C) \) and \( CD(C) \), respectively, in the figure.

For the \( l^\text{th} \) layer. For given \( b \), a \( \text{SNR} \) \( L \)-tuple \((\text{SNR}_1, \ldots, \text{SNR}_L)\) is achievable if there exists a sequence of coding schemes \((f^{(k)}, g_1^{(k)}, \ldots, g_L^{(k)})\) such that \( n = bk \) and, as \( k \to \infty \), \( \text{SNR}_l^{(k)} \to \text{SNR}_l \) for all \( l = 1, \ldots, L \). The achievable \( \text{SNR} \) region of the system is the closure of the convex hull of all achievable \( \text{SNR} \) \( L \)-tuples.

It is well-known that, under mild conditions on the erasure statistics\(^1\) the capacity region of the BEBC is given by

\[
C = \{(R_1, \ldots, R_L) : R_l \geq 0, \sum_{l=1}^L \frac{R_l}{C_l} \leq 1, l \in \{1, \ldots, L\}\}. \tag{1}
\]

It is seen immediately that this region is achieved by time-sharing between the individual capacities of each user. In fact, the vertices of the region’s dominant face, defined by the hyperplane \( \sum_{l=1}^L \frac{R_l}{C_l} = 1 \) are the points \( c_l = (0, \ldots, C_l, \ldots, 0) \). The BEBC belongs to the class of \textit{stochastically degraded broadcast channels} [8]. With degradation order defined by \( \epsilon_1 \geq \ldots \geq \epsilon_L \), this implies that any message to user \( l \) can be also decoded by all users \( j > l \) (with better channels). Building on [1], we can give an explicit characterization of the \( \text{SNR} \) region for the BEBC achieved by the layered architecture shown in Fig. 1.

An ideal (rate-distortion achieving) multiresolution source code produces \( L \) layers at rates \( R'_1, \ldots, R'_L \) bits per source symbol (See Fig. 1). The \( L \) layers are encoded by a broadcast code for the BEBC at rates \((R_1, \ldots, R_L) \in C\). By definition of \( b \), the channel coding rate and the source coding rate of layer \( l \) are related by \( R'_l = b R_l \). Since the Gaussian source is successively refinable [10], these correspond to distortion levels \( D_l = 2^{-2 \sum_{j=1}^L R'_j} \), or, equivalently, to \( \text{SNRs} \)

\[
\text{SNR}_l = \tilde{b} \sum_{j=1}^l R'_j \tag{2}
\]

(where we define the constant \( \tilde{b} = (20 \log_{10} 2)b \)). Since the BEBC capacity region can be achieved by time-sharing, this reduces to successive transmission of the layers such that layer 1 is encoded in the first segment of \( n_1 \) channel uses, layer 2 in the next \( n_2 \)

\(^1\) As a matter of fact, the capacity region of a broadcast channel depends only on the marginal transition probabilities [8], and hence erasures can be arbitrarily correlated. Furthermore, erasures need not be i.i.d. in time; any erasure process such that the fraction of erased symbols converges almost surely to \( \epsilon_l \) for each \( l = 1, \ldots, L \) yields the same capacity region [9].
channel uses, and so on. Assuming that each channel code operate at the capacity of the corresponding user, we have segments of length $n_t = k b R_l / C_l$.

Let $P(b, C_1, \ldots, C_L)$ denote the region of SNRs achievable with the layered scheme given above. Since $C$ is a polytope and (2) is a linear transformation, also $P(b, C_1, \ldots, C_L)$ is a polytope. This is defined by the inequalities $0 \leq \text{SNR}_1 \leq \cdots \leq \text{SNR}_L$ and by the inequality that determines the dominant face. The vertices $\{v_l\}$ of the dominant face of $P(b, C_1, \ldots, C_L)$ are obtained by mapping via (2) the vertices $\{c_i\}$ of the dominant face of $C$, and are explicitly given by

$$v_1 = (\tilde{b} C_1, \ldots, \tilde{b} C_L), \quad v_2 = (0, \tilde{b} C_2, \ldots, \tilde{b} C_L), \ldots \quad \text{and} \quad v_L = (0, \ldots, 0, \tilde{b} C_L). \quad (3)$$

Any point $x$ on the dominant face of $P(b, C_1, \ldots, C_L)$ must satisfy the hyperplane equation $u^T x = a$ for a certain constant $a$. Imposing this condition for all $x = v_l$ given in (3) and eliminating the constant $a$, we can solve for the vector $u$ perpendicular to the dominant face hyperplane, and obtain the inequality that defines the dominant face as

$$\text{SNR}_1 \left( \frac{1}{C_1} - \frac{1}{C_2} \right) + \text{SNR}_2 \left( \frac{1}{C_2} - \frac{1}{C_3} \right) + \cdots + \text{SNR}_L \left( \frac{1}{C_L} \right) \leq b. \quad (4)$$

It should be noted that the vertex $v_l$ corresponds to transmitting a non-layered source-channel coding scheme at the capacity $C_l$ of user $l$. Therefore, all users $l, l+1, \ldots, L$ receive at the rate-distortion bound of user $l$, and all users $1, \ldots, l-1$, with smaller individual capacity, achieve SNR of 0 dB (i.e., they reconstruct the source as the constant zero vector). As expected, operating at the vertices of the SNR achievable region requires no multiresolution/layered coding scheme. It should also be observed that the $P(b, C_1, \ldots, C_L)$ is achievable by using the concatenation of multiresolution source coding and time-sharing (also referred to as “progressive transmission”) on any degraded broadcast channel with individual single-user capacities $C_1, \ldots, C_L$.

III. MULTICAST SYSTEM OPTIMIZATION

Any point on the dominant face of $P(b, C_1, \ldots, C_L)$ is a Pareto-optimal point that corresponds to the maximization of a certain objective function. The choice of the objective function depends on the scenario considered, which determines an appropriate performance criterion. In this work we shall consider the following four objective functions and solve the corresponding optimization problems:

**Max Weighted Total SNR** (MWTS): maximize $\sum_{l=1}^{L} w_l \text{SNR}_l$, where $\{w_l\}$ are given non-negative weights, for a given $b$.

**Min Weighted Total Distortion** (MWTD): minimize $\sum_{l=1}^{L} w_l D_l$, where $\{w_l\}$ are given non-negative weights, for a given $b$.

**Min-Max SNR penalty** (MMS): minimize $\max_{l=1, \ldots, L} \{ \text{SNR}_l^{\text{opt}} - \text{SNR}_l \}$ where $\text{SNR}_l^{\text{opt}} = b C_l$ is the individual rate-distortion bound for user $l$, for a given $b$.

**Min Bandwidth** (MB): minimize $b$, for given SNR constraints, i.e., subject to $\gamma \in P(b, C_1, \ldots, C_L)$, where $\gamma \geq 0$ is a vector of target SNRs.

Because of space limitation, here we will give only a sketch of the solutions of the above problems. Since $P(b, C_1, \ldots, C_L)$ is a polytope, the MWTS problem is immediately given as a linear program and, for all weighting coefficients $\{w_l\}$ the solution is a vertex of $P(b, C_1, \ldots, C_L)$. In particular, it is immediate to see that the solution is obtained by choosing the vertex $v_{l^*}$ such that $l^* = \arg \max_{l \in L} \left\{ C_l \sum_{j=1}^{L} w_j \right\}$. Since the vertices $v_l$
can be achieved by standard single-layer coding, this problem does not lead to interesting
code design and shall not be considered in the following.

The MMS problem can be equivalently formulated as

\[
\begin{align*}
\text{minimize} & \quad t, \\
\text{subject to} & \quad \tilde{b} C_l - \text{SNR}_l \leq t, \quad \forall \ l \in \mathcal{L} \\
\{\text{SNR}_l\} & \in \mathcal{P}(b, C_1, \ldots, C_L).
\end{align*}
\]

(5)

This is a linear program that can be solved by standard numerical methods.

The MB problem is meaningful only if the target SNR vector \( \gamma \) is compatible with
the degradedness of the BEBC, i.e., if it satisfies the inequalities \( 0 \leq \gamma_1 \leq \cdots \leq \gamma_L \). In
this case, we notice that the minimum \( b \) is obtained by imposing the condition that \( \gamma \)
belongs to the dominant face of \( \mathcal{P}(b, C_1, \ldots, C_L) \). This is obtained explicitly from (4) as

\[
\tilde{b} = \gamma_1 \left( \frac{1}{C_1} - \frac{1}{C_2} \right) + \gamma_2 \left( \frac{1}{C_2} - \frac{1}{C_3} \right) + \cdots + \gamma_L \frac{1}{C_L}.
\]

The solution of the MWTD problem is more involved and shall be discussed in the
rest of this section. For analytical convenience, we reparameterize the problem and define
\( r_l = -\log D_l = \log_{10} \frac{\text{SNR}_l}{10} \), \( b' = \log_{10} \tilde{b} \), \( r_0 \Delta = 0 \) and \( \Delta_l = \frac{1}{C_l} - \frac{1}{C_{l+1}} \) with \( C_{L+1} \Delta = \infty \). Then, the problem is given as

\[
\begin{align*}
\text{minimize} & \quad \sum_{l=1}^{L} w_l \exp(-r_l), \\
\text{subject to} & \quad 0 \leq r_1 \leq \cdots \leq r_L \\
\sum_{l=1}^{L} \Delta_l r_l & = b'.
\end{align*}
\]

(6)

This is a convex optimization problem. By differentiating the associated Lagrangian
function \( \Lambda \), we obtain the following Karush-Kuhn-Tucker (KKT) conditions:

\[
\begin{align*}
\forall l \in \mathcal{L}, \quad \frac{\partial \Lambda}{\partial r_l} & = -w_l \exp(-r_l) + \lambda \Delta_l - \mu_l + \mu_{l+1} = 0, \\
\sum_{l=1}^{L} \Delta_l r_l & = b', \\
\mu_l(r_{l-1} - r_l) & = 0, \\
\lambda \left( \sum_{l=1}^{L} \Delta_l r_l - b' \right) & = 0, \quad \mu_l \geq 0, \quad \lambda \geq 0, \quad 0 \leq r_1 \leq \cdots \leq r_L.
\end{align*}
\]

The difficulty here is given by the fact that some groups of variables \( r_l \) with adjacent
indices may be equal. Consider the \( L \) inequalities \( 0 \leq r_1 \leq \cdots \leq r_L \). Some of them
are strict and some hold with equality. There are precisely \( 2^L - 1 \) such configurations,
excluding the configuration of all equalities, which yields the case \( r_l = 0 \) for all \( l \in \mathcal{L} \)
that is clearly not the solution. Although it is not possible to obtain a closed-form solution
of the KKT equations for the general case, we notice that for each configuration of the
inequalities \( 0 \leq r_1 \leq \cdots \leq r_L \) we can find a point for which the gradient of \( \Lambda \) is
zero. Therefore, the solution can be found by searching over all \( 2^L - 1 \) configurations,
and checking the consistency of the KKT conditions for each corresponding tentative
solution. Since the problem is convex, as soon as a consistent solution is found this is
the sought optimal point.

Next, we will show how to obtain a tentative solution that annihilates the gradient of
\( \Lambda \) assuming a certain configuration of the inequalities \( 0 \leq r_1 \leq \cdots \leq r_L \). Suppose that
0 \leq r_1 \leq r_2 \leq \ldots \leq r_{k-1} < r_k = r_{k+1} = \ldots = r_{k+N-1} < r_{k+N} \leq \ldots \leq r_L$, for some $N$ and $k$. From the KKT conditions it follows that $\mu_{k+i} \geq 0$ for $1 \leq i \leq N-1$, and $\mu_{k+i} = 0$ for $i = 0, N$. On letting $\rho_k = \exp(-r_k)$ and equating the partial derivatives of $\Lambda$ to zero, we obtain the system of equations

$$\begin{align*}
-\mu_{k+1} + w_k \rho_k &= \lambda \Delta_k \\
\mu_{k+1} - \mu_{k+2} + w_{k+1} \rho_k &= \lambda \Delta_{k+1} \\
&\vdots \\
\mu_{k+N-1} + w_{k+N-1} \rho_k &= \lambda \Delta_{k+N-1}.
\end{align*}$$

(7)

Summing all equations we find

$$\rho_k = \frac{\lambda \sum_{j=k}^{k+N-1} \Delta_j}{\sum_{j=k}^{k+N-1} w_j}. \quad (8)$$

Now, assume that we have $G+1$ disjoint groups of equal elements $r_l$, where group $g = 0$ contains the elements equal to zero, and groups $g = 1, \ldots, G$ contain non-zero elements. Also, let $r_{kg}$ denote the value of the elements in group $g$, where $k_g, k_g + 1, \ldots, k_g + N_g - 1$ is the support of group $g$ and, obviously, $\sum_{g=0}^{G} N_g = L$. Similarly to (8), letting $\rho_{kg} = \exp(-r_{kg})$, and defining $A_g = \sum_{j=k_g}^{k_g+N_g-1} \Delta_j$ and $W_g = \sum_{j=k_g}^{k_g+N_g-1} w_j$, we obtain $\rho_{kg} = \frac{\lambda A_g}{W_g}$. Replacing this into the dominant-face equation, which must hold with equality (otherwise, some of the distortions could be improved), we find

$$b' = \sum_{l=1}^{L} \Delta_l r_l = \ldots = - \sum_{g=1}^{G} A_g \log \left( \frac{\lambda A_g}{W_g} \right) \Rightarrow -\log(\lambda) = \frac{b' + \sum_{g=1}^{G} A_g \log \frac{A_g}{W_g}}{\sum_{g=1}^{G} A_g}. \quad (9)$$

Eventually, for each configuration of the equality groups we can compute $\lambda$ from (9), $\rho_{kg}$ as $\rho_{kg} = \frac{\lambda A_g}{W_g}$ and $\{\mu_l\}$ using (7), in order to check the KKT conditions consistency. It should not be neglected that a group $g = 0$ of $N_0$ zero elements $0 = r_1 = \ldots = r_{N_0}$ may exist. This yields $\mu_1, \ldots, \mu_{N_0} \geq 0$, $\mu_{N_0+1} = 0$ and $\lambda \Delta_l - \mu_l + \mu_{l+1} = 1$ for $1 \leq l \leq N_0$, so that the values of $\mu_1, \ldots, \mu_{N_0}$ can be also found.

**IV. Joint Source-Channel Coding for the BEBC**

In this section we present a practical coding scheme based on embedded scalar quantization and binary raptor codes [5] that closely approaches the theoretical limits of Section III at finite block length and constant encoding/decoding complexity per source symbol. It is well-known that entropy-coded scalar quantization achieves rate-distortion performance within a fixed gap from the optimal limit, for i.i.d. smooth source statistics and quadratic distortion [11]. Furthermore, embedded uniform scalar quantizers that achieve (roughly) a 6dB improvement in the reconstruction SNR for each additional quantization bit are easily designed [12]. Let $Q : \mathbb{R} \rightarrow \mathbb{F}_2^N$ denote a scalar embedded quantizer, and let $\mathbf{u} = Q(S)$ denote the binary index of length $N$ corresponding to the source symbol $S$. A source vector $S = (S_1, \ldots, S_k)$ is mapped componentwise into a sequence of binary indices $U = [u_1, \ldots, u_k]$, formatted as an $N \times k$ binary array. We shall refer to the $i$-th row of $U$, denoted by $U_{i,:}$, as the $i$-th “bit-plane”. Without loss of generality, we let $U_{1,:}$ denote the sign bit-plane, and $U_{2,:}, \ldots, U_{N,:}$ denote the magnitude bit-planes with
Figure 2. The tetrahedron shown in the figure is the achievable SNR region for a BEBC with individual capacities $C_1 = 0.36, C_2 = 0.81, C_3 = 0.9$ and $b = 9.5$. The solutions of the MWTS, MWTD (for unit weights) and MMS problems are shown as points on the dominant face. The corresponding points are, respectively: $\text{SNR}^\text{MWSD} = [0.46 0.33 0.33], \text{SNR}^\text{MWTD} = [19.228 22.98 22.98], \text{SNR}^\text{MMS} = [7.29 32.77 37.92]$

decreasing order of significance. The quantizer distortion as a function of the bit-planes is given by $D_Q(p) = \alpha 2^{-2p}$ for $p = 1, 2, \ldots, N$, where $\alpha$ is a fixed constant (see [13]). For the class of quantizers used in this work we have $\alpha \approx 1.5$ as shown in Fig. 3 (a).

The quantizer output $U$ forms a discrete memoryless source, with entropy $H(u)$ (in units of bits/source symbol) that can be decomposed according to the well-known entropy chain rule. For simplicity of notation, the conditional entropy of bit-plane $i$ given the previous bit-planes will be denoted by $H_i = \frac{1}{k} H(U_{i+1}|U_{1: i}, \ldots, U_{i-1:i})$. The rate-distortion function achieved by embedded scalar quantization followed by entropy coding of the quantization indices, namely, the set of rate-distortion points with coordinates $(\sum_{i=1}^p H_i, D_Q(p))$ for $p = 1, \ldots, N$, is close to the Gaussian rate-distortion function (see Fig. 3(a)). It is also well-known that powerful graph-based codes, such as low

Figure 3. (a): Comparison between the reconstruction SNR-Rate trade-offs for Gaussian source obtained by rate-distortion function (SNR$_{rd}(R')$), by the set of rate-distortion points of the embedded quantizer with coordinates $(\sum_{i=1}^p H_i, D_Q(p))$ for $p = 1, \ldots, N$ (SNR$_{eq}$) and by $D_Q(R') = \alpha 2^{-2R}$ where $\alpha = 1.5$; (b): This figure illustrates the rounding of the SNR solutions of the information-theoretic optimization problems to the actual SNRs achievable by the proposed scheme, with corresponding allocation of the bit-planes.

Notice that the optimization problems discussed in Section III refer to $\alpha = 1$ but they can be straightforwardly adapted to different values of $\alpha$. 

99
density parity check codes (LDPCs) and raptor codes [14], [5], achieve very small bit-error probability at rates very close to the capacity of the BEC, with linear encoding and decoding complexity. This suggests that the theoretical optimal performance can be approached by the following general coding architecture:

1- Let \( \{\text{SNR}_l\} \) denote the desired theoretical operating point on the dominant face of \( P(b, C_1, \ldots, C_L) \), obtained as the solution of the system optimization of Section III.

2- Allocate \( p_l \) bit-planes (from 1 to \( p_l \)) to layer \( l \), such that \( -10 \log_{10} D_Q(p_l) \approx \text{SNR}_l \). Since the quantizer offers only \( N \) possible distortion levels, this step requires some suitable rounding of the target \( \text{SNR}_l \). We set \( p_0 = 0 \) for convenience of notation.

3- For each \( l \), the block of corresponding \( p_l - p_{l-1} \) bit-planes is entropy-coded by any suitable lossless compression algorithm. The corresponding block of \( B_l = k \sum_{i=1+p_{l-1}}^{p_l} H_i \) entropy-coded bits are mapped into a codeword of length \( n_l = B_l/(C_l - \delta_l) \) channel-coded symbols, where \( \delta_l \) denotes the “gap-to-capacity” of the binary code used at layer \( l \).

4- The channel encoded blocks are transmitted in sequence, with total block length \( n = \sum_{l=1}^{L} n_l \).

5- The \( l^{th} \) decoder, sequentially decodes all layer codewords from 1 to \( l \). If the post-decoding error probability is sufficiently low, the source can be reconstructed at the quantizer with reconstruction SNR, \( -10 \log_{10} D_Q(p_l) \), that is close to the target \( \text{SNR}_l \) by construction.

The overall bandwidth expansion factor of the scheme is given by

\[
\hat{b} = \frac{\sum_{l=1}^{L} n_l}{k} = \sum_{l=1}^{L} \frac{\sum_{i=1+p_{l-1}}^{p_l} H_i}{C_l - \delta_l}.
\]

For a good family of codes, achieving small gap-to-capacity, and an appropriate choice of the bit-plane allocation to the layers, the achieved \( \hat{b} \) is only slightly larger than the “design” bandwidth expansion factor \( b \).

As noticed in [3] and [4], the weak link in the above code construction consists of the entropy-coding stage: since the entropy coding is known to be non-robust to channel errors, in order to approach the quantization distortion, each layer must be decoded with extremely low residual bit-error probability. On one hand, we want to perform close to capacity (small gap \( \delta_l \)). On the other hand, we need to achieve essentially zero bit-error probability at the decoder output. In practice, we wish to make use of modern low-density graph-based codes with BP decoding [5],[15]. With these codes it is easy to achieve bit-error probabilities between \( 10^{-2} \) and \( 10^{-3} \) at block lengths between 1000 and 10000. However, these post-decoding bit-error probabilities are completely unsuited for entropy decoding, and this would result in a very large distortion (see [3] for a detailed analysis of the effect of residual channel decoding errors on the end-to-end distortion, when entropy coding is implemented by an arithmetic coding [8] algorithm).

A vast literature has dealt with the issue of making entropy coding “robust” to channel errors (see [16], [17], [18] and references therein). Here we shall follow a more radical approach. As proposed in [3], we bypass explicit entropy coding and map directly the redundant quantizer bit-planes onto channel-encoded symbol using a single linear encoding operation. It is well-known that linear coding achieves the entropy-rate of a discrete source. Also, linear coding achieves the capacity of symmetric channels. Since the concatenation of two linear maps (the first for compression and the second for channel coding) is linear, it follows that the limiting performance of ideal lossless compression
and channel coding can be achieved by a single linear encoding stage. The proposed coding scheme maps each bit-plane $U_i$, for $p_{l-1} < i \leq p_l$ (i.e., associated to layer $l$) to a block of channel-coded symbols $X_i = U_i G_i$, where $G_i$ is a $k \times m_i$ generator matrix of a suitable linear code. This is indicated as the “JSCC encoder” in the block diagram of Fig. 4. The rate of this linear encoder is chosen such that $\frac{m_i}{k}$ is slightly larger than the theoretical limit $\frac{H_i}{C_l}$. The rate margin must be optimized numerically and it depends on the block length $k$, on the capacity $C_l$ and on the actual code construction. The coded block for layer $l$ is obtained by concatenating the codewords $[X_{p_{l-1}+1}, \ldots, X_{p_l}]$ for all bit-planes allocated to layer $l$, so that $n_l = \sum_{i=1}^{p_l} m_i$. At the layer-$l$ receiver, all bit-planes $1 \leq i \leq p_l$ are decoded in sequence by a multi-stage decoder (“JSCC decoder” in the block diagram of Fig. 4). Each component decoder in the multi-stage decoder is based on the BP algorithm. Details are omitted for the sake of space limitation, and the reader is referred to [3] where the scheme is explained in details. The BP decoder for the bit-plane $i$ incorporates the source a-priori information given by the a-priori probability distribution of the symbols of $U_i$, conditioned on the knowledge of the previous bit-planes $U_{1}, \ldots, U_{i-1}$. These are obtained from the decoder output of the previous stages in the multistage decoder. The BP iterative decoding algorithm produces soft outputs, in the form of approximated posterior log-likelihood ratios of the bit-plane symbols. This information can be used to reconstruct the quantizer quantization points according to the posterior-mean estimator principle. This yields the well-known soft-bit reconstruction scheme (see [3] and references therein), which further improves the scheme performance with respect to hard-decoding reconstruction.

Building on our results in [3], we implement the linear encoders operating on each
bit-plane via a *systematic* raptor encoder with degree distribution [5]

\[ \Omega(x) = 0.008x + 0.494x^2 + 0.166x^3 + 0.073x^4 + 0.083x^5 \\
+ 0.056x^6 + 0.037x^7 + 0.056x^8 + 0.025x^9 + 0.003x^{10}. \]

Raptor codes are particularly suited for this application since they have an efficient systematic encoder (linear complexity) and an efficient BP iterative decoding algorithm that easily incorporates the bit-plane a-priori probabilities. The same basic raptor encoder can be used at each bit-plane, since it can produce an arbitrary number of coded symbols and then adapt the coding rate for any desired ratio \( H_i/C_i \).

We report a few simulation result for source block length \( k = 10000 \) in Tabs. I and II. In Tab. I, the result for the MWTD and MMS scenarios are shown. The optimal \( \{\text{SNR}_{ij}\}^* \) values for each user are given both for \( \alpha = 1 \) (rate-distortion limit) and \( \alpha = 1.5 \) (scalar quantizers) cases. The allocation of the bit-planes to the layers (expressed by the vector \( p \)), is obtained by rounding the optimal SNR values to discrete SNR values of the embedded quantizer (see Fig. 3 (b)), as explained before.

### References


