Recursive Power Allocation in Gaussian Layered Broadcast Coding with Successive Refinement

Chris T. K. Ng*, Deniz Gündüz†, Andrea J. Goldsmith*, and Elza Erkip†

*Dept. of Electrical Engineering, Stanford University, Stanford, CA 94305 USA
†Dept. of Electrical and Computer Engineering, Polytechnic University, Brooklyn, NY 11201 USA
Email: *{ngctk,andrea}@wsl.stanford.edu, †dgundu01@utopia.poly.edu, elza@poly.edu

Abstract— **A transmitter without channel state information wishes to send a delay-limited Gaussian source over a slowly fading channel that has a finite number of discrete fading states. The source is coded in layers, with each layer successively refining the description in the previous one. These coded source layers are then superimposed and simultaneously transmitted to the receiver. The receiver decodes the layers that are supported by the realization of the channel, and combines the descriptions in the decoded layers to reconstruct the source up to a distortion. The expected distortion is minimized by optimally allocating the transmit power among the given number of source layers. For two layers, the allocation is optimal when power is first assigned to the higher layer up to a power ceiling that depends only on the channel fading distribution; all remaining power, if any, is allocated to the lower layer. For multiple layers, the overall expected distortion can be written as a set of recurrence relations, and the minimum expected distortion is found by recursively applying the two-layer optimization procedure at each recurrence step.**

I. Introduction

In an ergodic wireless channel, from the source-channel separation theorem [1], it is optimal to first compress the source and incur the associated distortion at a rate equal to the channel capacity, then send the compressed representation over the channel at capacity with asymptotically small error. However, when delay constraints stipulate that the receiver decodes within a single realization of a slowly fading channel, without channel state information (CSI) at the transmitter, the channel becomes non-ergodic and source-channel separation is not necessarily optimal. In this case it is possible to reduce the end-to-end distortion of the reconstructed source by jointly optimizing the source-coding rate and the transmit power allocation based on the characteristics of the source and the channel. In particular, we consider using the layered broadcast coding approach with successive refinement in the transmission of a Gaussian source over a slowly fading channel, which has a finite number of discrete fading states, in the absence of CSI at the transmitter. The source is coded in layers, with each layer successively refining the description in the previous one. The transmitter simultaneously transmits the codewords of all layers to the receiver by superimposing them with an appropriate power allocation. The receiver successfully decodes the layers supported by the channel realization, and combines the descriptions in the decoded layers to reconstruct the source up to a distortion. In this paper, we are interested in minimizing the expected distortion of the reconstructed source by optimally allocating the transmit power among the layers of codewords. The system model is applicable to communication systems with real-time traffic where it is difficult for the transmitter to learn the channel condition. For example, in a satellite voice system, it is desirable to consider the efficient transmission of the voice streams over uncertain channels that minimize the end-to-end distortion.

The broadcast strategy is proposed in [2] to characterize the set of achievable rates when the channel state is unknown at the transmitter. In the case of a Gaussian channel under Rayleigh fading, [3] describes the layered broadcast coding approach, and derives the optimal power allocation that maximizes expected capacity when the channel has a single-antenna transmitter and receiver. The layered broadcast approach is extended to multiple-antenna channels and the corresponding achievable rates are presented in [4]. In the transmission of a Gaussian source over a Gaussian channel, uncoded transmission is optimal [5] in the special case when the source bandwidth equals the channel bandwidth [6]. For other bandwidth ratios, hybrid digital-analog joint source-channel transmission schemes are studied in [7]–[9]; in these works the codes are designed to be optimal at a target SNR but degrade gracefully should the realized SNR deviate from the target. In particular, [8] conjectures that no code is simultaneously optimal at different SNRs when the source and channel bandwidths are not equal. In this paper, the code considered is not targeted for a specific fading state; we minimize the expected distortion over the fading distribution of the channel.

In [10], the minimum distortion is investigated in the transmission of a source over two independently fading channels in terms of the distortion exponent, which is defined as the exponential decay rate of the expected distortion in the high SNR regime. Upper bounds on the distortion exponent and achievable joint source-channel schemes are presented in [11] for a single-antenna quasi-static Rayleigh fading channel, and later in [12], [13] for multiple-antenna channels. One of the proposed schemes in [12], layered source coding with progressive transmission (LS), is analyzed in terms of expected distortion for a finite number of layers at finite SNR in [14].

This work was supported by the US Army under MURI award W911NF-05-1-0246, the ONR under award N00014-05-1-0168, DARPA under grant 1105741-1-TFIND, a grant from Intel, and the NSF under grant 0430885.

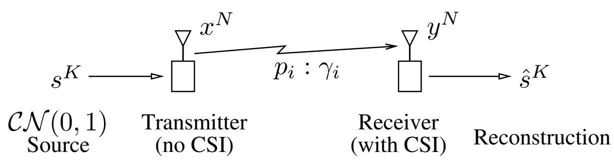

Fig. 1. Source-channel coding without CSI at the transmitter.

The results in [11], [12] show that the broadcast strategy with layered source coding under an appropriate power allocation scheme is optimal for multiple-input single-output (MISO) and single-input multiple-output (SIMO) systems in terms of the distortion exponent. Numerical optimization of the power allocation with constant rate among the layers is examined in [15], while [16] considers the optimization of power and rate allocation and presents approximate solutions in the high SNR regime. Motivated by the optimality of the broadcast strategy in the high SNR regime, in this work we investigate the minimum expected distortion at any arbitrary finite SNR.

The remainder of the paper is organized as follows. The system model is presented in Section II, and the layered broadcast coding scheme with successive refinement is explained in more detail in Section III. Section IV focuses on the optimal power allocation between two layers, with the analysis being extended to consider recursive power allocation among multiple layers in Section V. Numerical examples of the power allocation algorithm are presented in Section VI, followed by conclusions in Section VII.

II. System Model

Consider the system model illustrated in Fig. 1: A transmitter wishes to send a Gaussian source over a wireless channel to a receiver, at which the source is to be reconstructed up to a distortion. Let the source be denoted by s, which is a sequence of independent identically distributed (iid) zero-mean circularly symmetric complex Gaussian (ZMCSCG) random variables with unit variance: $s \in \mathbb{C} \sim \mathcal{CN}(0,1)$. The transmitter and the receiver each have a single antenna and the channel is described by

$$y = Hx + n, \tag{1}$$

where $x \in \mathbb{C}$ is the transmit signal, $y \in \mathbb{C}$ is the received signal, and $n \in \mathbb{C} \sim \mathcal{CN}(0,1)$ is iid unit-variance ZMCSCG noise.

We assume the random channel gain H has a finite number of discrete fading states. Let $\gamma_i \triangleq |h_i|^2$ denote the channel power gain of a given fading state, where h_i is a realization of H with probability p_i, for $i = 1, \ldots, M$. The receiver has perfect CSI but the transmitter has only channel distribution information (CDI), i.e., the transmitter knows the probability mass function (pmf) of H but not its instantaneous realization. The channel is modeled by a quasi-static block fading process: H is realized iid at the onset of each fading block and remains unchanged over the block duration. We assume decoding at the receiver is *delay-limited*; namely, delay constraints preclude coding across fading blocks but dictate that the receiver decodes at the end of each block. Hence the channel is non-ergodic.

Suppose each fading block spans N channel uses, over which the transmitter describes K of the source symbols. We define the *bandwidth ratio* as $b \triangleq N/K$, which relates the number of channel uses per source symbol. At the transmitter there is a power constraint on the transmit signal $\mathrm{E}\big[|x|^2\big] \leq P$, where the expectation is taken over repeated channel uses. We assume K is large enough to consider the source as ergodic, and N is large enough to design codes that achieve the instantaneous channel capacity of a given fading state with negligible probability of error.

At the receiver, the channel output y is used to reconstruct an estimate $\hat{s}$ of the source. The distortion D is measured by the mean squared error $\mathrm{E}[(s-\hat{s})^2]$ of the estimator, where the expectation is taken over the K-sequence of source symbols and the noise distribution. The instantaneous distortion of the reconstruction depends on the fading realization of the channel; we are interested in minimizing the expected distortion $\mathrm{E}_H[D]$, where the expectation is over the fading distribution.

III. Layered Broadcast Coding with Successive Refinement

To characterize the set of achievable rates when the channel state is unknown at the transmitter, a broadcast strategy is described in [2]. The transmitter designs its codebook by imagining it is communicating with an ensemble of virtual receivers. Each virtual receiver corresponds to a fading state: the realization of the fading state is taken as the channel gain of the virtual receiver. The realized rate at the original receiver is given by the decodable rate of the realized virtual receiver. A fading channel without transmitter CSI, therefore, can be modeled as a broadcast channel (BC). In particular, the capacity region of the BC defines the maximal set of achievable rates among the virtual receivers, which, in terms of the original fading channel, is the maximal set of realized rates among the fading states. In this work, we derive the optimal operating point in the BC capacity region that minimizes the expected distortion $\mathrm{E}_H[D]$ of the reconstructed source.

For fading Gaussian channels, a layered broadcast coding approach is described in [3], [4]. In the layered broadcast approach, the virtual receivers are ordered according to their channel strengths: for single-antenna channels, the channel strength of a virtual receiver is given by the channel power gain of its corresponding fading state. We interpret each codeword intended for a virtual receiver as a *layer* of code, and the transmitter sends the superposition of all layers to the virtual receivers. The capacity region of a single-antenna Gaussian BC is achievable by successive decoding [17], in which each virtual receiver decodes, in additional to its own layer, all the layers below it (the ones with weaker channel strengths). Hence each layer represents the *additional* information over its lower layer that becomes decodable by the original receiver should the layer be realized.

The layered broadcast approach fits particularly well the successive refinability [18], [19] of a Gaussian source. Succes-

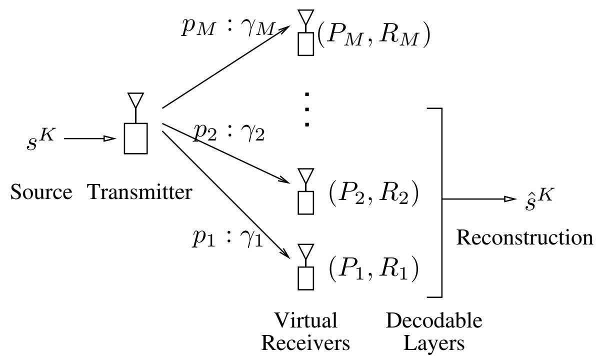

Fig. 2. Layered broadcast coding with successive refinement.

sive refinability states that if a source is first described at rate R_1, then subsequently refined at rate R_2, the overall distortion is the same as if the source were described at rate $R_1 + R_2$ in the first place. As the Gaussian source is successively refinable, naturally, each layer in the broadcast approach can be used to carry refinement information of a lower layer. Concatenation of broadcast channel coding with successive refinement source coding is shown in [11], [12] to be optimal in terms of the distortion exponent for MISO/SIMO systems.

We apply the layered broadcast approach and successive refinability to perform source-channel coding as outlined in Fig. 2. The fading distribution has M states: the channel power gain realization is γ_i with probability p_i, for $i = 1, \ldots, M$. Accordingly there are M virtual receivers and the transmitter sends the sum of M layers of codewords. Let layer i denote the layer of codeword intended for virtual receiver i, and we order the layers as $\gamma_M > \cdots > \gamma_1 \geq 0$. We refer to layer M as the highest layer and layer 1 as the lowest layer. Each layer successively refines the description of the source s from the layer below it, and the codewords in different layers are independent. Let P_i be the transmit power allocated to layer i, then the transmit symbol x can be written as

$$x = \sqrt{P_1}\, x_1 + \sqrt{P_2}\, x_2 + \cdots + \sqrt{P_M}\, x_M, \tag{2}$$

where $x_1, \ldots, x_M$ are iid ZMCSCG random variables with unit variance.

With successive decoding, each virtual receiver first decodes and cancels the lower layers before decoding its own layer; the undecodable higher layers are treated as noise. Thus the rate R_i intended for virtual receiver i is

$$R_i = \log\left(1 + \frac{\gamma_i P_i}{1 + \gamma_i \sum_{j=i+1}^{M} P_j}\right), \tag{3}$$

where the term $\gamma_i \sum_{j=i+1}^{M} P_j$ represents the interference power from the higher layers. Suppose γ_k is the realized channel power gain, then the original receiver can decode layer k and all the layers below it. Hence the realized rate $R_{\text{rlz}}(k)$ at the original receiver is $R_1 + \cdots + R_k$.

From the rate distortion function of a complex Gaussian source [17], the mean squared distortion is 2^{-bR} when the source is described at a rate of bR per symbol. Thus the realized distortion $D_{\text{rlz}}(k)$ of the reconstructed source $\hat{s}$ is

$$D_{\text{rlz}}(k) = 2^{-bR_{\text{rlz}}(k)} = 2^{-b(R_1 + \cdots + R_k)}, \tag{4}$$

where the last equality follows from successive refinability. The expected distortion $\mathrm{E}_H[D]$ is obtained by averaging over the fading distribution pmf:

$$\mathrm{E}_H[D] = \sum_{i=1}^{M} p_i D_{\text{rlz}}(i) = \sum_{i=1}^{M} p_i 2^{-b(\sum_{j=1}^{i} R_j)}. \tag{5}$$

In this work, we derive the optimal power allocation $P_1^*, \ldots, P_M^*$ among the layers to find the minimum expected distortion $\mathrm{E}_H[D]^*$. If a layer has an expected power gain of zero (i.e., $p_i \gamma_i = 0$), the layer is allocated zero power; hence in the derivation we assume $p_i \gamma_i \neq 0$, for $i = 1, \ldots, M$. Note that the expected distortion is monotonically decreasing in the transmit power P, hence the power constraint can be taken as an equality $\sum_{i=1}^{M} P_i = P$, and the optimization formulated as:

$$\begin{aligned} \mathrm{E}_H[D]^* = &\min_{P_1, \ldots, P_M} \mathrm{E}_H[D] \\ &\text{subject to } P_i \geq 0,\ \textstyle\sum P_i = P,\ \forall i = 1, \ldots, M. \end{aligned} \tag{6}$$

We first consider the power allocation between two layers in the next section, then the analysis is extended to consider more than two layers in Section V. The layered source coding broadcast scheme can be straightforwardly extended to MISO/SIMO systems. The equivalent single-antenna channel distribution is found by using isotropic inputs at the transmitter for MISO systems, and performing maximal-ratio combining at the receiver for SIMO systems.

IV. Two-Layer Optimal Power Allocation

Suppose the channel fading distribution has only two states: the channel power gain realization is either α or β, with $\beta > \alpha \geq 0$. The transmitter then sends two layers of codewords as shown in Fig. 3. Let T_i denote the total transmit power constraint, and T_{i+1} denote the power allocated to layer $i+1$; the remaining power $T_i - T_{i+1}$ is allocated to layer i. The subscript i eases the extension of the analysis when we consider M layers in Section V; in this section, we consider two layers and $i = 1$. The decodable rates for the virtual receivers are denoted by R_i, R_{i+1}; with successive decoding, they are given as follows:

$$R_{i+1} = \log(1 + \beta T_{i+1}) \tag{7}$$

$$R_i = \log\left(1 + \frac{\alpha(T_i - T_{i+1})}{1 + \alpha T_{i+1}}\right). \tag{8}$$

Suppose we generalize slightly and consider the weighted distortion:

$$D_i = u 2^{-bR_i} + w 2^{-b(R_i + R_{i+1})}, \tag{9}$$

where the weights $\{u, w\}$ are non-negative. Note that the weighted distortion D_i is the expected distortion $\mathrm{E}_H[D]$ when the weights $\{u, w\}$ are the probabilities of the fading realizations.

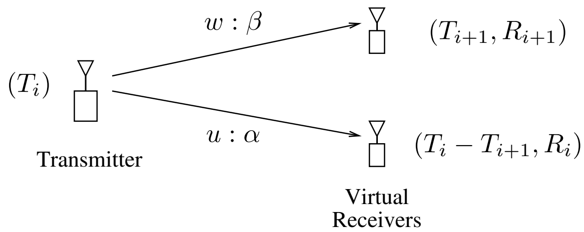

Fig. 3. Power allocation between two layers.

Given T_i, the total power available to the two layers, we optimize over T_{i+1} to minimize the weighted distortion:

$$D_i^* = \min_{0 \le T_{i+1} \le T_i} D_i \tag{10}$$

$$= \min_{0 \le T_{i+1} \le T_i} \Big(\frac{1+\alpha T_i}{1+\alpha T_{i+1}}\Big)^{-b}\Big[u + (1+\beta T_{i+1})^{-b} w\Big]. \tag{11}$$

The minimization can be solved by the Lagrange method. We form the Lagrangian:

$$L(T_{i+1}, \lambda_1, \lambda_2) = D_i + \lambda_1 (T_{i+1} - T_i) - \lambda_2 T_{i+1}. \tag{12}$$

Applying the Karush-Kuhn-Tucker (KKT) conditions, the gradient of the Lagrangian vanishes at the optimal power allocation T_{i+1}^*. Specifically, the KKT conditions stipulate that at T_{i+1}^*, either one of the inequality constraints is active, or $dD_i/dT_{i+1} = 0$, which lead to the solution:

$$T_{i+1}^* = \min(U_{i+1}, T_i) = \begin{cases} U_{i+1} & \text{if } U_{i+1} \le T_i \quad (13a) \\ T_i & \text{else,} \quad (13b) \end{cases}$$

where

$$U_{i+1} \triangleq \begin{cases} 0 & \text{if } \beta/\alpha \le 1 + u/w \\ \dfrac{1}{\beta}\bigg(\Big[\dfrac{w}{u}\Big(\dfrac{\beta}{\alpha} - 1\Big)\Big]^{\frac{1}{1+b}} - 1\bigg) & \text{else.} \end{cases} \tag{14}$$

Interestingly, U_{i+1} depends only on the layer parameters w, β, u, α (which are derived from the channel fading distribution) and the bandwidth ratio b, but not on the total power T_i. In other words, the higher layer is allocated a *fixed* amount of power as long as there is sufficient power available. The optimal power allocation, therefore, adopts a simple policy: first assign power to the higher layer up to a ceiling of U_{i+1}, then assign all remaining power, if any, to the lower layer.

Under optimal power allocation T_{i+1}^*, the minimum weighted distortion as a function of the total power T_i is given by

$$D_i^* = \begin{cases} (1+\alpha T_i)^{-b} W_i & \text{if } U_{i+1} \le T_i \quad (15a) \\ u + (1+\beta T_i)^{-b} w & \text{else,} \quad (15b) \end{cases}$$

where

$$W_i \triangleq (1+\alpha U_{i+1})^b \big[u + (1+\beta U_{i+1})^{-b} w\big]. \tag{16}$$

Note that when the total power constraint $T_{i+1} \le T_i$ is not active (15a), the consequent minimum weighted distortion is analogous to that of a *single* layer with channel power gain α and an equivalent weight W_i. On the other hand, when the total power constraint is active (15b), it is equivalent to one with channel power gain β and an equivalent weight w (with an additive constant u in the distortion). Hence under optimal power allocation, with respect to the minimum expected distortion, the two layers can be represented by a single *aggregate* layer; this idea is explored further when we consider multiple layers in Section V.

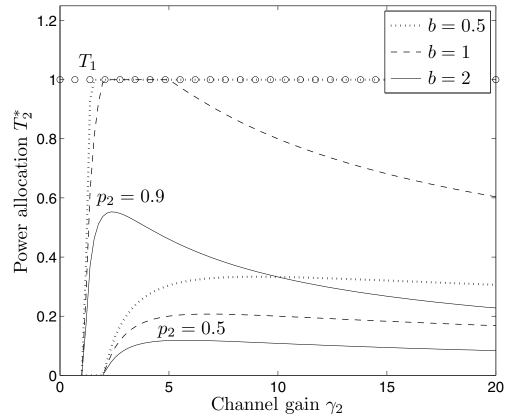

Fig. 4. Optimal power allocation between two layers ($T_1 = 0$ dB).

The optimal power allocation and the minimum expected distortion for a two-layer transmission scheme (i.e., the channel has two discrete fading states) is shown in Fig. 4 and Fig. 5, where T_2^*, the power assigned to the top layer, is computed under the parameters:

$$\begin{aligned} w &= p_2, & \beta &= \gamma_2, \\ u &= 1 - p_2, & \alpha &= 1. \end{aligned} \tag{17}$$

Fig. 4 shows that when the channel power gain γ_2 is small, an increase in γ_2 leads to a larger power allocation T_2^* at the top layer, up to the total available power T_1. However, as γ_2 further increases, T_2^* begins to fall. This is because the transmission power in the higher layer is in effect interference to the lower layer. When the top layer has a strong channel, the overall expected distortion is dominated by the bottom layer; therefore, in such case it is more beneficial distributing the power to minimize the interference. Fig. 5 plots the two-layer minimum expected distortion on a logarithmic scale, and it shows that the relative power gain of the channels has only a marginal impact on the expected distortion, as the overall distortion is in general dominated by the weaker channel.

V. Recursive Power Allocation

In this section we consider the case when the fading distribution has M fading states as depicted in Fig. 2, where M is finite and $M \ge 2$. For notational convenience, we write the power assignment as a cumulative sum starting from the

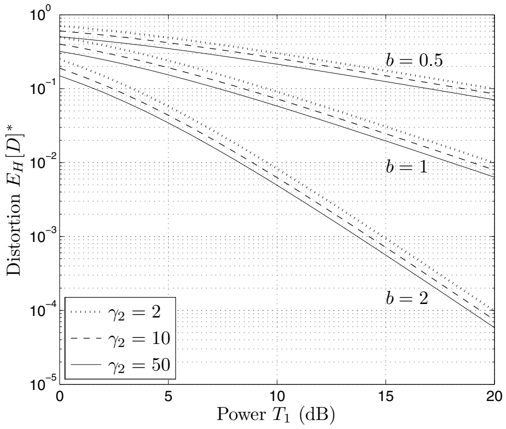

Fig. 5. Two-layer minimum expected distortion ($p_2 = 0.5$).

top layer:

$$T_i \triangleq \sum_{j=i}^{M} P_j, \quad \text{for } i = 1, \ldots, M. \tag{18}$$

The original power assignments $\{P_1, \ldots, P_M\}$ can then be recovered from $\{T_1, \ldots, T_M\}$ by taking their differences. By definition, $T_1 = P$ is given; hence the optimization is over the variables $T_2, \ldots, T_M$:

$$\begin{aligned} \mathrm{E}_H[D]^* = \min_{T_2,\ldots,T_M} \mathrm{E}_H[D] \\ \text{subject to } 0 \le T_M \le \cdots \le T_2 \le P. \end{aligned} \tag{19}$$

A. Expected Distortion Recurrence Relations

In terms of the cumulative power variables $T_1, \ldots, T_M$, the expected distortion in (5) can be written as

$$\mathrm{E}_H[D] = \sum_{i=1}^{M} p_i \Big(\prod_{j=1}^{i} \frac{1 + \gamma_j T_j}{1 + \gamma_j T_{j+1}} \Big)^{-b}, \tag{20}$$

where $T_{M+1} \triangleq 0$. We factor the sum of cumulative products in (20) and rewrite the expected distortion as a set of recurrence relations:

$$D_M \triangleq \big(1 + \gamma_M T_M\big)^{-b} p_M \tag{21}$$

$$D_i = \Big(\frac{1 + \gamma_i T_i}{1 + \gamma_i T_{i+1}} \Big)^{-b} \big(p_i + D_{i+1}\big), \tag{22}$$

where i runs from $M - 1$ down to 1. The term D_i can be interpreted as the cumulative distortion from layers i and above, with $D_1 = \mathrm{E}_H[D]$. Note that D_i depends on only two adjacent power allocation variables T_i and T_{i+1}; therefore, in each recurrence step i, we solve for the optimal T^*_{i+1} in terms of T_i:

$$D^*_M \triangleq D_M \tag{23}$$

$$D^*_i = \min_{0 \le T_{i+1} \le T_i} \Big(\frac{1 + \gamma_i T_i}{1 + \gamma_i T_{i+1}} \Big)^{-b} \big(p_i + D^*_{i+1}\big). \tag{24}$$

In the last recurrence step ($i = 1$), the minimum expected distortion $\mathrm{E}_H[D]^*$ is then given by D^*_1.

B. Reduction through Optimal Power Allocation

We consider the layers from top to bottom. In each recurrence step, the minimum distortion D^*_i in (24) can be found by optimally allocating power between two adjacent layers as described in Section IV. In the first recurrence step ($i = M - 1$), we consider the power allocation between the topmost two layers. The minimal distortion D^*_M is found by setting the parameters in (11) to be:

$$\begin{aligned} w_{M-1} &= p_M, & \beta_{M-1} &= \gamma_M, \\ u_{M-1} &= p_{M-1}, & \alpha_{M-1} &= \gamma_{M-1}, \end{aligned} \tag{25}$$

where the subscripts on the layer parameters w, β, u, α designate the recurrence step. In general, in recurrence step i, the power allocation between layer i and layer $i+1$ can be found by the optimization:

$$D^*_i = \min_{0 \le T_{i+1} \le T_i} \Big(\frac{1 + \alpha_i T_i}{1 + \alpha_i T_{i+1}} \Big)^{-b} \Big[u_i + (1 + \beta_i T_{i+1})^{-b} w_i \Big], \tag{26}$$

the solution of which is given in (15):

$$D^*_i = \begin{cases} (1 + \alpha_i T_i)^{-b} W_i & \text{if } U_{i+1} \le T_i \quad \text{(27a)} \\ u_i + (1 + \beta_i T_i)^{-b} w_i & \text{else.} \quad \text{(27b)} \end{cases}$$

There are two cases to the solution of D^*_i. In the first case, the power allocation is not constrained by the available power T_i, and we substitute (27a) in the recurrence relation (24) to find the minimum distortion in the next recurrence step $i - 1$:

$$\begin{aligned} D^*_{i-1} = \min_{0 \le T_i \le T_{i-1}} & \Big(\frac{1 + \gamma_{i-1} T_{i-1}}{1 + \gamma_{i-1} T_i} \Big)^{-b} \\ & \cdot \Big[p_{i-1} + (1 + \alpha_i T_i)^{-b} W_i \Big]. \end{aligned} \tag{28}$$

The minimization in (28) has the same form as the one in (26), but with the following parameters:

$$\begin{aligned} w_{i-1} &= W_i, & \beta_{i-1} &= \alpha_i, \\ u_{i-1} &= p_{i-1}, & \alpha_{i-1} &= \gamma_{i-1}. \end{aligned} \tag{29}$$

Hence the minimization can be solved the same way as in the last recurrence step. In the second case, the power allocation is constrained by the available power T_i, thus we instead substitute (27b) in (24) in the next recurrence step $i - 1$:

$$\begin{aligned} D^*_{i-1} = \min_{0 \le T_i \le T_{i-1}} & \Big(\frac{1 + \gamma_{i-1} T_{i-1}}{1 + \gamma_{i-1} T_i} \Big)^{-b} \\ & \cdot \Big[p_{i-1} + u_i + (1 + \beta T_i)^{-b} w_i \Big], \end{aligned} \tag{30}$$

which again has the same form as in (26), with the following parameters:

$$\begin{aligned} w_{i-1} &= w_i, & \beta_{i-1} &= \beta_i, \\ u_{i-1} &= p_{i-1} + u_i, & \alpha_{i-1} &= \gamma_{i-1}. \end{aligned} \tag{31}$$

Therefore, in each recurrence step, the two-layer optimization procedure described in Section IV can be used to find the minimum distortion and the optimal power allocation between the current layer and the aggregate higher layer.

C. Feasibility of Unconstrained Minimizer

When we proceed to the next recurrence step, however, it is necessary to determine which set of parameters in (29), (31) should be applied. Note that in the optimization in (26), if the available power T_i is unlimited (i.e., $T_i = \infty$), then the optimal power allocation is $T^*_{i+1} = U_{i+1}$ as given in (13a); hence U_{i+1} is the *unconstrained* minimizer of D_i. Consequently, we can first assume the minimization in (26) is unconstrained by T_i and its solution is given by (27a). If the unconstrained allocation U_{i+1} is found to be feasible, then it is indeed the optimal allocation. On the other hand, if U_{i+1} is subsequently shown to be infeasible, then we backtrack to the minimization in (26) and adopt the *constrained* solution given by (27b). In this case, $T^*_{i+1} = T_i$ as given in (13b), which implies layer i is inactive since $P^*_i = T_i - T^*_{i+1} = 0$.

We ascertain the feasibility of U_{i+1} by verifying that it does not exceed the inflowing power allocation T_i from the lower layer i, which in turn depends on the power allocation T_{i-1} from the next lower layer $i-1$ and so on. The procedure can be accomplished by the recursive algorithm shown in Algorithm 1. We start by allocating power between the topmost two layers (line 1). In each recursion step, we compute the unconstrained allocation U (line 3). If U does not exceed the total power P, we first assume it is feasible, and proceed in the recursion to find the power allocation T^*_i from the lower layer (line 10). If U turns out to be infeasible, then we repeat the allocation step with the constrained minimization parameters (line 16). The recursion continues until the bottom layer is reached (line 4). In the best case, if the unconstrained allocations for all layers are feasible, the algorithm has complexity $O(M)$. In the worst case, if all unconstrained allocations are infeasible, each recursion step performs two power allocations and the algorithm has complexity $O(2^M)$.

Algorithm 1 Recursive Power Allocation

1: ALLOC($M-1, p_M, \gamma_M, p_{M-1}, \gamma_{M-1}$) ▷ Start from top

2: **procedure** ALLOC(i, w, β, u, α)
3:     Compute U from w, β, u, α
4:     **if** $i = 1$ **then** ▷ Bottom layer
5:         $T^*_2 \leftarrow \min(U, P)$
6:         **return**
7:     **end if**
8:     **if** $U < P$ **then** ▷ Within total power P
9:         Compute W from U, w, β, u, α
10:         ALLOC($i-1, W, \alpha, p_{i-1}, \gamma_{i-1}$) ▷ Unconstrained
11:         **if** $T^*_i \geq U$ **then**
12:             $T^*_{i+1} \leftarrow U$ ▷ U is feasible
13:             **return**
14:         **end if**
15:     **end if**
16:     ALLOC($i-1, w, \beta, p_{i-1}+u, \gamma_{i-1}$) ▷ Constrained
17:     $T^*_{i+1} \leftarrow T^*_i$
18: **end procedure**

VI. Numerical Results

In this section, we present numerical results produced by the recursive power allocation algorithm described in Section V. In the examples, we assume the channel pmf is taken from a discretized Rayleigh fading distribution. Specifically, for a channel under Rayleigh fading with unit power, the channel power gain γ is exponentially distributed with unit mean, and its probability density function (pdf) is given by

$$f(\gamma) = e^{-\gamma}, \quad \text{for } \gamma \geq 0. \tag{32}$$

We truncate the pdf at $\gamma = \Gamma$, quantize γ into M evenly spaced levels:

$$\gamma_i \triangleq (i-1)\frac{\Gamma}{M-1}, \quad \text{for } i = 1, \dots, M, \tag{33}$$

and discretize the probability distribution of γ to the closest lower level γ_i:

$$p_i \triangleq \int_{(i-1)\frac{\Gamma}{M-1}}^{i\frac{\Gamma}{M-1}} f(\gamma)\, d\gamma, \quad \text{for } i = 1, \dots, M-1 \tag{34}$$

$$p_M \triangleq \int_{\Gamma}^{\infty} f(\gamma)\, d\gamma. \tag{35}$$

While it is possible to consider the optimal discretization of a fading distribution that minimizes expected distortion, in this paper we assume the channel pmf is given and do not consider such a step.

The optimal power allocation for the discretized Rayleigh fading pmf is shown in Fig. 6 and Fig. 7. The Rayleigh fading pdf is truncated at $\Gamma = 2$, and discretized into $M = 25$ levels. The truncation is justified by the observation that in the output the highest layers near Γ are not assigned any power. Fig. 6 plots the optimal power allocation P^*_i's for different layers (indexed by the channel power gain γ_i) at SNRs $P = 0$ dB, 5 dB, and 10 dB, with the bandwidth ratio $b = 1$. We observe that the highest layers are inactive ($P^*_i = 0$), and within the range of active layers a lower layer is in general allocated more power than a higher layer, except at the lowest active layer where it is assigned the remaining power. As SNR increases, the power allocations of the higher layers are unaltered, but the range of active layers extends further into the lower layers. On the other hand, Fig. 7 plots the allocation P^*_i's for different bandwidth ratios $b = 0.5, 1, 2$ at the SNR of 0 dB. It can be observed that a higher b (i.e., more channel uses per source symbol) has the effect of spreading the power allocation further across into the lower layers.

Intuitively, the higher layers have stronger channels but suffer from larger risks of being in outage, while the lower layers provide higher reliability but at the expense of having to cope with less power-efficient channels. Accordingly the optimal power allocation is concentrated around the middle layers. Furthermore, as SNR increases, the numerical results suggest that, to minimize the expected distortion in a Rayleigh fading channel, it is more favorable taking on the weaker channels with the extra power rather than accepting the larger risks from the higher layers.

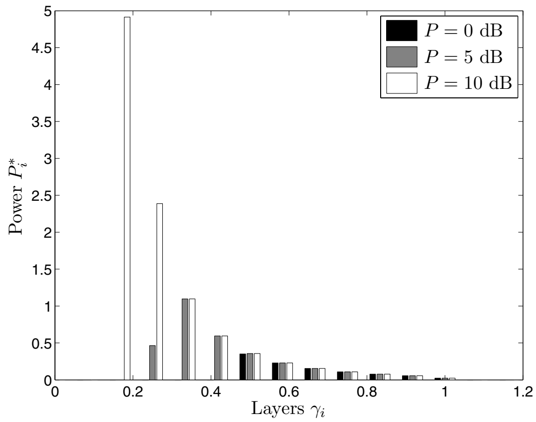

Fig. 6. Optimal power allocation ($b = 1$).

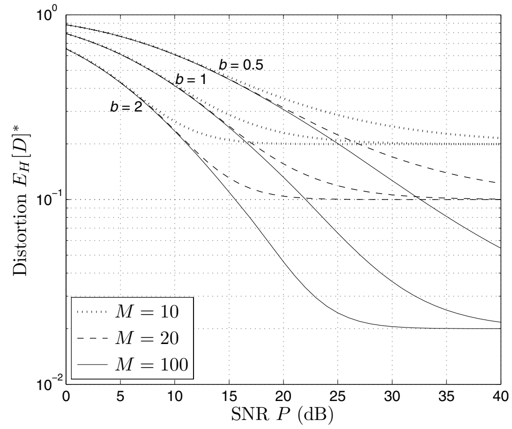

Fig. 8. Minimum expected distortion under optimal power allocation.

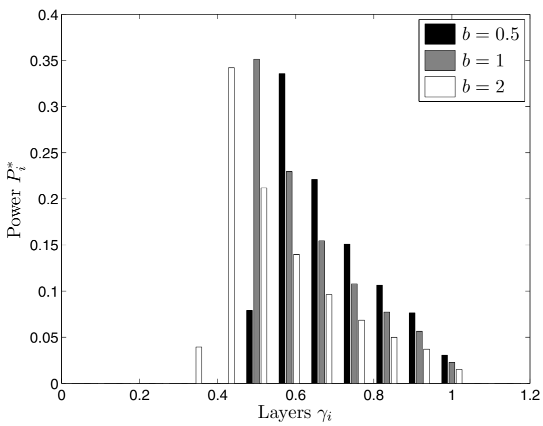

Fig. 7. Optimal power allocation ($P = 0$ dB).

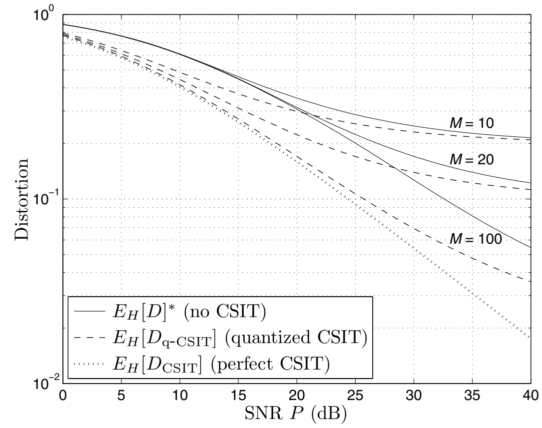

Fig. 9. Expected distortion lower bounds with CSIT ($b = 0.5$).

The minimum expected distortion $\mathrm{E}_H[D]^*$ under optimal power allocation is shown in Fig. 8 on a logarithmic scale. When the bandwidth ratio b is higher, $\mathrm{E}_H[D]^*$ decreases as expected. However, the improvement in $\mathrm{E}_H[D]^*$ from refining the resolution M of the discretization is almost negligible at low SNRs. At high SNRs, on the other hand, the distortion is dominated by the outage probability

$$P_{\text{out}} \triangleq \Pr\{\gamma_1 = 0 \ \text{ is realized}\} \tag{36}$$

$$= \int_0^{\frac{\Gamma}{M-1}} f(\gamma)\, d\gamma, \tag{37}$$

which is decreasing in M. Therefore, when the SNR is sufficiently high, the expected distortion $\mathrm{E}_H[D]^*$ reaches a floor that is dictated by P_{out}. This behavior is due to having evenly-spaced γ_i's; the performance could be improved by allowing γ_i's to take on arbitrary levels.

As a comparison, we consider the expected distortion lower bounds when the system has CSI at the transmitter (CSIT). Under the discretized Rayleigh fading pmf, suppose the realized channel power gain is known to be γ_k, then it is optimal for the transmitter to concentrate all power on layer k to achieve the instantaneous distortion $D_{\text{q-CSIT}} = (1 + \gamma_k P)^{-b}$. Thus with the quantized CSIT, the expected distortion is given by

$$\mathrm{E}_H[D_{\text{q-CSIT}}] = \sum_{k=1}^{M} p_k (1 + \gamma_k P)^{-b}. \tag{38}$$

In terms of the original Rayleigh fading pdf $f(\gamma)$, with perfect CSIT, the expected distortion is similarly given by

$$\mathrm{E}_H[D_{\text{CSIT}}] = \int_0^{\infty} e^{-\gamma} (1 + \gamma P)^{-b}\, d\gamma, \tag{39}$$

where the definite integral can be evaluated numerically. The expected distortions are plotted in Fig. 9 for the cases of no CSIT, quantized and perfect CSIT. It can be observed that at low SNRs, quantized CSIT is nearly as good as perfect CSIT, whereas at high SNRs, quantized CSIT provides only marginal improvement over no CSIT, as the expected distortion is dominated by the probability of outage.

VII. Conclusion

We considered the problem of source-channel coding over a delay-limited fading channel without CSI at the transmitter, and derived the optimal power allocation that minimizes the end-to-end expected distortion in the layered broadcast coding transmission scheme with successive refinement. Between two layers of codewords, the allocation is optimal when power is first assigned to the higher layer up to a power ceiling that depends only on the channel fading distribution; all the remaining power, if any, is then allocated to the lower layer. Next we write the minimum expected distortion as a set of recurrence relations, and in each recurrence step the two-layer optimization procedure solves the power allocation between the current layer and the aggregate higher layer. The parameters of the aggregate layer, however, depend on whether the optimization is constrained by the available power from its lower layer, which is only given in the next recurrence step. Therefore, a recursive algorithm was proposed to first compute the unconstrained power allocation that minimizes the expected distortion, which if feasible is the optimal allocation, otherwise the allocation needs to be recomputed adopting the constrained solution of the optimization.

In addition, we applied the recursive power allocation algorithm to the pmf of a discretized Rayleigh fading distribution. We observed that the optimal power allocation is concentrated around the middle layers, and within this range the lower layers are assigned more power than the higher ones. As the SNR increases, the allocations of the higher layers remain unchanged, and the extra power is allocated to the idle lower layers. In [20] this work is extended to study the limiting process as the discretization resolution tends to infinity. Specifically, this extension derives the optimal power distribution that minimizes the expected distortion when the fading distribution of the channel is given by a continuous probability density function.

References

[1] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," *IRE Nat. Conv. Rec., part 4*, pp. 142–163, 1959.

[2] T. M. Cover, "Broadcast channels," *IEEE Trans. Inform. Theory*, vol. 18, no. 1, pp. 2–14, Jan. 1972.

[3] S. Shamai (Shitz), "A broadcast strategy for the Gaussian slowly fading channel," in *Proc. IEEE Int. Symp. Inform. Theory*, 1997, p. 150.

[4] S. Shamai (Shitz) and A. Steiner, "A broadcast approach for a single-user slowly fading MIMO channel," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2617–2635, Oct. 2003.

[5] T. J. Goblick, Jr., "Theoretical limitations on the transmission of data from analog sources," *IEEE Trans. Inform. Theory*, vol. 11, no. 4, pp. 558–567, Oct. 1965.

[6] M. Gastpar, B. Rimoldi, and M. Vetterli, "To code, or not to code: Lossy source-channel communication revisited," *IEEE Trans. Inform. Theory*, vol. 49, no. 5, pp. 1147–1158, May 2003.

[7] S. Shamai (Shitz), S. Verdú, and R. Zamir, "Systematic lossy source/channel coding," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 564–579, Mar. 1998.

[8] U. Mittal and N. Phamdo, "Hybrid digital-analog (HDA) joint source-channel codes for broadcasting and robust communications," *IEEE Trans. Inform. Theory*, vol. 48, no. 5, pp. 1082–1102, May 2002.

[9] Z. Reznic, M. Feder, and R. Zamir, "Distortion bounds for broadcasting with bandwidth expansion," *IEEE Trans. Inform. Theory*, vol. 52, no. 8, pp. 3778–3788, Aug. 2006.

[10] J. N. Laneman, E. Martinian, G. W. Wornell, and J. G. Apostolopoulos, "Source-channel diversity for parallel channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 10, pp. 3518–3539, Oct. 2005.

[11] D. Gunduz and E. Erkip, "Source and channel coding for quasi-static fading channels," in *Proc. of Asilomar Conf. on Signals, Systems and Computers*, Nov. 2005.

[12] ——, "Joint source-channel codes for MIMO block fading channels," *submitted*, 2006.

[13] G. Caire and K. Narayanan, "On the SNR exponent of hybrid digital-analog space time coding," in *Proc. Allerton Conf. Commun., Contr., Comput.*, 2005.

[14] F. Etemadi and H. Jafarkhani, "Optimal layered transmission over quasi-static fading channels," in *Proc. IEEE Int. Symp. Inform. Theory*, 2006, pp. 1051–1055.

[15] S. Sesia, G. Caire, and G. Vivier, "Lossy transmission over slow-fading AWGN channels: a comparison of progressive, superposition and hybrid approaches," in *Proc. IEEE Int. Symp. Inform. Theory*, 2005, pp. 224–228.

[16] K. E. Zachariadis, M. L. Honig, and A. K. Katsaggelos, "Source fidelity over fading channels: Erasure codes versus scalable codes," in *Proc. IEEE Globecom Conf.*, vol. 5, 2005, pp. 2558–2562.

[17] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley-Interscience, 1991.

[18] W. H. R. Equitz and T. M. Cover, "Successive refinement of information," *IEEE Trans. Inform. Theory*, vol. 37, no. 2, pp. 269–275, Mar. 1991.

[19] B. Rimoldi, "Successive refinement of information: Characterization of the achievable rates," *IEEE Trans. Inform. Theory*, vol. 40, no. 1, pp. 253–259, Jan. 1994.

[20] C. T. K. Ng, D. Gündüz, A. J. Goldsmith, and E. Erkip, "Minimum expected distortion in Gaussian layered broadcast coding with successive refinement," submitted to *IEEE Int. Symp. Inform. Theory*, 2007.