Joint Source-Channel Coding for Quasi-Static Fading Channels with Quantized Feedback

Farzad Etemadi
Center for Pervasive Comm. and Comp.
Department of EECS
University of California, Irvine, CA 92697, USA
Email: fetemadi@uci.edu

Hamid Jafarkhani
Center for Pervasive Comm. and Comp.
Department of EECS
University of California, Irvine, CA 92697, USA
Email: hamidj@uci.edu

Abstract— THIS PAPER IS ELIGIBLE FOR THE STUDENT PAPER AWARD. We consider the transmission of a Gaussian source over a single-input multiple-output (SIMO) quasi-static fading channel. The goal is to minimize the expected distortion of the reconstructed signal at the receiver. We consider a delay-limited scenario where channel coding is restricted to a single realization of the channel. Channel state information (CSI) is assumed to be known perfectly at the receiver, and a zero-delay, noiseless, fixed-rate feedback link provides a quantized version of the CSI to the transmitter. An upper bound on the performance is derived and it is shown that for practical values of the channel signal to noise ratio (SNR), this bound can be achieved with a very limited knowledge of the channel quality. We show that unlike the rate maximization problems, temporal power adaptation at the transmitter provides significant gains, and the amount of the gain heavily depends on the bandwidth expansion ratio. For asymptotically high SNRs, we derive the distortion exponent of the system, defined as the slope of the expected distortion with respect to the channel SNR. We show that the distortion exponent of limited feedback is equivalent to that of superposition coding without feedback, so long as the number of quantization levels in the feedback scheme is equal to the number of the layers in the superposition coding scheme. For the finite-SNR regime, we propose an optimal and efficient numerical technique to design the feedback scheme. Numerical results for a Rayleigh fading channel are also presented.

I. INTRODUCTION

A Multimedia source, such as an image or a video sequence, usually produces an analog output and the objective in a multimedia communication system is to reproduce the source output at the receiver with minimum distortion. In this context, Shannon's separation theorem states that the source and channel coding tasks can be done separately. The separation theorem is valid for an ergodic channel where the channel codewords extend across a large number of different channel realizations. The ergodicity assumption is easily violated in a slowly varying fading channel with a delay-sensitive application that limits channel coding to a single realization of the channel. In this case, the end-to-end system should be designed using a joint source-channel coding approach.

The joint source-channel coding problem for quasi-static fading channel has been an active research area in recent years [1]–[3]. Most of the existing literature is based on the distortion exponent metric that characterizes the high signal to noise ratio (SNR) behavior of the optimal expected distortion [1].

The distortion exponent provides a useful tool for comparing various transmission strategies, but its asymptotic nature limits its applicability for practical values of the SNR. As a result, a different approach to the joint source-channel coding problem deals directly with the finite-SNR regime [4]–[6].

For slowly fading channels, estimating the channel state information (CSI) at the receiver is relatively simple and incurs a negligible loss in the transmission rate. Assuming a known channel at the receiver only (CSIR), a multiple-input single-output (MISO) or a single-input multiple-output (SIMO) quasi-static fading channel can be modeled as a degraded broadcast channel, and the optimal transmission strategy is multi-layer superposition coding [7]. The joint source-channel coding literature cited in the preceding paragraph falls into the category of CSIR-only problems. On the other hand, if the CSI is also available at the transmitter (CSIT), then the fading channel becomes a memoryless channel and the source-channel separation holds. In this paper, we are considering an intermediate case between the perfect CSIT and no-CSIT scenarios, and assume that the transmitter has access to quantized CSI through a feedback channel.

Transmission over fading channels with partial CSIT has been widely studied in recent years [8]–[10] (and the references therein). The main objective in the existing work is to maximize the transmission rate, and the partial CSIT is usually in the form of quantized or noisy CSI. In particular, Kim *et. al* [8] present a comprehensive study of the rate maximization problem using a noiseless quantized feedback. In this paper, we use the framework of [8] and apply it to the joint source-channel coding problem in the presence of partial CSIT. We derive an upper bound on the system performance and show that for practical values the SNR, a low-rate feedback scheme can achieve this bound. We show that unlike the rate maximization problems, temporal power adaptation at the transmitter provides significant gains, and the amount of the gain heavily depends on the bandwidth expansion ratio. We also derive the distortion exponent of the feedback scheme without power adaptation, and show its equivalence to the distortion exponent of superposition coding without feedback when the number of quantization levels in the feedback scheme is equal to the number of the layers in the superposition coding scheme. For the finite-SNR regime, we

propose an optimal and efficient numerical technique to design the feedback scheme. Finally, we present numerical results for a Rayleigh fading channel. Our discussion will be limited to SIMO channels. The case of multiple transmit antennas is more involved due to the problem of spatial power allocation between the antennas [10], a topic that is considered for future work.

Notations: In the sequel, $\| \cdot \|$ is the Euclidean norm of a complex vector. $P(e)$ represents the probability of event e. $F(x)$ and $f(x)$ represent the cumulative distribution function (CDF) and the probability density function (PDF) of a given random variable. log and ln represent base 2 and natural logarithms, respectively. $\mathbb{R}$ is the set of real numbers. $f'(\cdot)$ denotes the derivative of $f(\cdot)$. $f(x) \doteq g(x)$ is the exponential equality, that is $\lim_{x \to \infty} \frac{f(x)}{g(x)} = 1$.

II. SYSTEM MODEL

We consider a discrete-time, complex Gaussian source with independent real and imaginary components of variance 0.5, and the mean square error (MSE) distortion measure. In this case, the distortion-rate (D-R) function of the source is given as $D(R) = 2^{-R}$, where R is the source coding rate in bits per symbol. We further assume that the length of the encoded sequence of source symbols is large enough so that the source can be considered ergodic. The design of the source coder is not considered in our work, and we assume the availability of a source coder capable of achieving the distortion-rate bound. The number of channel symbols transmitted per source symbol is denoted by the bandwidth expansion factor b. The bandwidth expansion factor quantifies the compression or expansion of the source with respect to the available transmission bandwidth.

The discrete-time, baseband model of a slowly fading, single-input multiple-output channel with M receive antennas is given as $\mathbf{r}_t = \mathbf{h} s_t + \mathbf{n}_t$, $t = 1, \cdots, T$, where $\mathbf{h}$, $\mathbf{n}_t$ and $\mathbf{r}_t$ are $M \times 1$ complex vectors and t is the index of the channel use. The channel coefficient vector $\mathbf{h}$ is assumed to be random but constant over a sequence of T channel uses. The elements of $\mathbf{h}$ are independent and identically distributed (i.i.d) according to some known distribution and $\gamma \triangleq \|\mathbf{h}\|^2$. s_t and $\mathbf{r}_t$ represent the transmitted symbol and the received vector, respectively, in the tth channel use and $\mathbf{n}_t$ is a complex Gaussian noise vector with independent real and imaginary parts of zero mean and variance 0.5. We assume that each codeword is limited to a single fading block, and the block length T is large enough such that the ergodic capacity of the forward channel can be achieved when $\mathbf{h}$ is known. For a transmit power constraint of $\mathcal{P}$, the ergodic capacity in the case of perfect CSIT is given as [11]:

$$C = \log(1 + \gamma \mathcal{P}) \quad \text{bits/channel use} \tag{1}$$

Throughout the paper, we assume that the receiver has perfect knowledge of $\mathbf{h}$ and employs maximum ratio combining. Outage-free transmission is possible if γ is known at the transmitter. Thus, the feedback link is solely used to transmit information about γ.

The feedback transmission scheme operates as follows. The channel realization γ is quantized using a K-level scaler quantizer defined by the encoder and decoder mappings $\Gamma_e : \mathbb{R} \to \mathcal{I}$ and $\Gamma_d : \mathcal{I} \to \{\gamma_i\}$, respectively. $\mathcal{I} = \{0, \cdots, K - 1\}$ represents the index set of the quantizer. The encoding operation is performed at the receiver. The Voronoi regions of the encoder are specified by the partition boundaries $\gamma_0^b = 0 < \gamma_1^b \leq \cdots \leq \gamma_{K-1}^b < \gamma_K^b = \infty$, and the encoder mapping is defined as $\Gamma_e(\gamma) = i$, if $\gamma \in [\gamma_i^b, \gamma_{i+1}^b)$. The encoder index $i \in \mathcal{I}$ is sent to the transmitter through the zero-delay, noiseless feedback link. The decoding operation is performed at the transmitter as $\Gamma_d(i) = \gamma_i$, where $\{\gamma_i\}$ represents the quantizer codebook with $\gamma_i \in [\gamma_i^b, \gamma_{i+1}^b)$. Upon receiving index i, the transmitter allocates power $\mathcal{P}_i$ to a capacity achieving codeword and transmits at rate $\mathcal{R}_i = \log(1 + \gamma_i \mathcal{P}_i)$. The transmission is then successful if $\gamma \geq \gamma_i$, otherwise outage occurs. The source coding rate at the transmitter is $b\mathcal{R}_i$ and the corresponding distortion at the receiver is $D(b\mathcal{R}_i)$, if the transmission is successful. In the case of outage, the resulting distortion is $D_0 = D(0) = 1$. Note that unlike the layered transmission schemes [2]–[6], we do not require the source to be successively refinable.

In [8], it is shown that when the goal is to maximize the expected rate, the optimal quantizer satisfies $\gamma_i^b = \gamma_i$ for $1 \leq i \leq K - 1$. The same logic can be applied to the distortion minimization problem and as a result, the expected distortion cost function is given as

$$\mathcal{E}_D = \sum_{i=0}^{K} [F(\gamma_i) - F(\gamma_{i-1})] \; D(b\mathcal{R}_{i-1}) \tag{2}$$

where $\gamma_{-1} = R_{-1} = 0$ and $\gamma_K = \infty$.

We consider two types of transmission power control strategies. Under the *short-term power constraint*, every codeword is allocated a power budget of $\mathcal{P}_i = \mathcal{P}$, regardless of the transmitter feedback index i. The more relaxed *long-term power constraint* allows the transmitter to choose a transmit power $\mathcal{P}_i$ for index i, while limiting the average transmit power to $\mathcal{P}$. The long-term power constraint is given as

$$\mathcal{E}_P = F(\gamma_1)\mathcal{P}_0 + \sum_{i=1}^{K-1} [F(\gamma_{i+1}) - F(\gamma_i)] \mathcal{P}_i \leq \mathcal{P} \tag{3}$$

The long-term power constraint may result in practically unacceptable large values of $\mathcal{P}_i$ for certain indices. In this case, one may also consider a maximum power constraint in the form of $\mathcal{P}_i \leq \mathcal{P}_m$.

III. PERFORMANCE BOUND

In this section, we derive a performance bound for our proposed transmission scheme. Assuming that perfect CSIT is available, separation theorem applies and outage-free transmission at a rate equal to the ergodic capacity is optimal. For a given power allocation $\mathcal{P}(\gamma)$, the expected distortion is then

$$\mathcal{E}_D = \int_0^\infty D\left(b \log\left(1 + \gamma \mathcal{P}(\gamma)\right)\right) \; f(\gamma) \, d\gamma \tag{4}$$

subject to the constraints

$$\mathcal{E}_P = \int_0^\infty \mathcal{P}(\gamma)\, f(\gamma)\, d\gamma \leq \mathcal{P} \qquad \text{and} \qquad \mathcal{P}(\gamma) \geq 0 \quad (5)$$

The expected distortion lower bound, $\bar{\mathcal{E}}_D$, is achieved for some power allocation function $\bar{\mathcal{P}}(\gamma)$. Using the Lagrange multiplier technique, $\bar{\mathcal{P}}(\gamma)$ can be written as the solution to the following variational problem:

$$\bar{\mathcal{P}}(\gamma) = \arg\min_{\mathcal{P}(\gamma)} \int_0^\infty \mathcal{L}(\gamma, \mathcal{P}(\gamma))\ d\gamma \qquad (6)$$

where $\mathcal{L}(\gamma, \mathcal{P}(\gamma)) = \{D(b\log(1+\gamma\mathcal{P}(\gamma))) + \lambda P(\gamma)\}f(\gamma)$, and $\lambda \geq 0$ is the Lagrange multiplier. Since $D(\cdot)$ is a convex and decreasing function and $\log(\cdot)$ is concave, it is easy to see that the functional $\mathcal{L}$ is convex. The Euler-Lagrange optimality condition [12] requires $-\frac{d}{d\gamma}\frac{\partial\mathcal{L}}{\partial\mathcal{P}'} + \frac{\partial\mathcal{L}}{\partial\mathcal{P}}$ to vanish at the optimal solution, or

$$\frac{b\gamma\, D'\left(b\log\left(1+\gamma\bar{\mathcal{P}}(\gamma)\right)\right)}{\ln(2)\,\left(1+\gamma\bar{\mathcal{P}}(\gamma)\right)} + \lambda = 0 \qquad (7)$$

For a given λ, (7) can be numerically solved to obtain $\bar{\mathcal{P}}(\gamma)$. A search over λ is then needed to find a $\bar{\mathcal{P}}(\gamma)$ that satisfies (5). For a complex Gaussian source, (7) can be solved analytically:

$$\bar{\mathcal{P}}(\gamma) = \left[\frac{1}{\gamma}\left(\frac{b\gamma}{\lambda}\right)^{\frac{1}{b+1}} - \frac{1}{\gamma}\right]^+ \qquad (8)$$

where $[x]^+ = \max(x, 0)$. The power allocation (8) is strongly influenced by the bandwidth expansion ratio, b. As $b \to 0$, the first term inside the brackets becomes independent of γ and as γ is increased, the power allocation becomes approximately uniform. In this regime, the behavior is similar to the classical water-filling solution for rate maximization problems where temporal power adaptation does not introduce much gain [8], [13]. From a joint source-channel coding point of view, a small b indicates that a small bandwidth is available per source symbol. Thus, the power adaptation policy allocates the power generously, and equally, to all reasonably large values of γ in order to cancel out the effect of the small bandwidth. On the other extreme, as $b \to \infty$, $\bar{\mathcal{P}}(\gamma)$ approaches the Dirac delta function $\delta(\gamma)$. In this regime, the large value of b translates the transmission rates supported by even small values of γ into a good source coding rate. As a result, the power is mostly allocated to poor channel realizations.

In what follows, we present a numerical example of the performance bound. In the case of a Rayleigh SIMO fading channel with M receive antennas, $\gamma = \frac{\chi}{2}$ where χ is a Chi-square random variable with $2M$ degrees of freedom. The PDF and CDF of γ are given as $f(x) = \frac{x^{M-1}e^{-x}}{(M-1)!}$ and $\mathrm{F}(x) = \frac{\Gamma_I(x,M)}{(M-1)!}$, respectively, where $\Gamma_I(x,M) = \int_0^x t^{M-1}e^{-t}dt$ is the incomplete Gamma function. In the plots, SNR represents the average signal to noise ratio at each receive antenna, which is equivalent to the power constraint $\mathcal{P}$.

The effect of the bandwidth expansion factor on the optimal power allocation for a SISO channel is shown in Figure 1. Figure 1(a) shows the absolute gain, defined as the reduction in
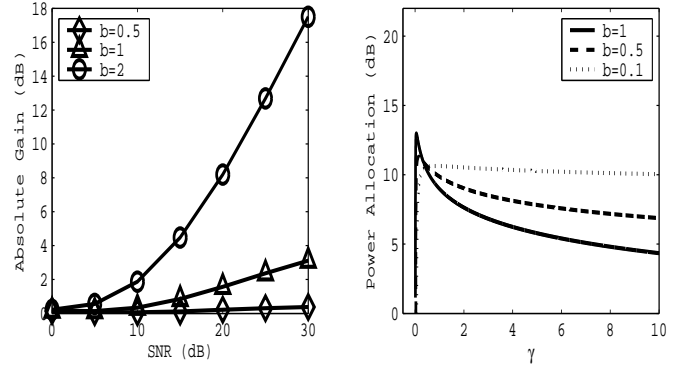

Fig. 1. Power adaptation results for a SISO channel with perfect CSIT (a) Absolute gain (b) Power allocation at $\mathcal{P} = 10$ dB

the expected distortion when temporal power adaptation (long-term) is used compared to when there is no such adaptation (short-term). We note that temporal power adaptation provides significant gains (up to 18 dB), and the amount of gain grows quickly with b. Fig. 1(b) shows the optimal power allocation, $\bar{\mathcal{P}}(\gamma)$, at $\mathcal{P} = 10$ dB. It is seen that the optimal power allocation becomes more non-uniform as b is increased. The results of this section shows that the effect of temporal power adaptation is more pronounced in a joint source-channel coding context when compared to the rate maximization problems.

IV. DISTORTION EXPONENT

The distortion exponent, Δ, is defined as $\Delta = -\lim_{P\to\infty}\frac{\log\mathcal{E}_D}{\log\mathcal{P}}$. The distortion exponent of the proposed feedback scheme in the case of short-term power constraint is derived in the following theorem.

Theorem 1: Assuming a short-term power constraint, the distortion exponent of the proposed scheme is given as

$$\Delta = M\left(1 - \frac{1-\frac{b}{M}}{1-\left(\frac{b}{M}\right)^{K+1}}\right) \qquad (9)$$

that is the same as the distortion exponent of a system with no CSIT that employs K-layer superposition coding.

Proof: We assume that in the high-SNR regime, $R_i = r_i\log\mathcal{P}$, where r_i is the multiplexing gain of layer i. The corresponding outage probability is given as

$$\mathrm{F}(\gamma_i) = \mathrm{P}_{out}(r_i\log\mathcal{P}) \doteq \mathcal{P}^{-d^*(r_i)} \qquad (10)$$

where $d^*(r) = M(1-r)$ is the optimal diversity at the multiplexing gain of r for the SIMO system [14]. Since $r_{i-1} \leq r_i$, $d^*(r_{i-1}) \geq d^*(r_i)$ and $\mathrm{F}(\gamma_i) - \mathrm{F}(\gamma_{i-1}) \doteq \mathrm{F}(\gamma_i)$. Also, $D(bR_{i-1}) = 2^{-br_{i-1}\log\mathcal{P}} = \mathcal{P}^{-br_{i-1}}$, using (2) we have

$$\mathcal{E}_D \doteq \sum_{i=0}^K \mathcal{P}^{-\left\{M(1-r_i)+br_{i-1}\right\}} \qquad (11)$$

To find the optimal distortion exponent, we need to solve the following optimization problem

$$\max_{0 \le r_0 \le \cdots \le r_{K-1}} \min \left\{ M(1-r_0),\ M(1-r_1)+br_0,\ \cdots, \right.$$
$$\left. M(1-r_{K-1})+br_{K-2},\ br_{K-1} \right\} \quad (12)$$

Each term inside the brackets in (12) is non-negative and represents a hyper-plane in (r_{i-1}, r_i). The inter-section of all hyper-planes is then the maximum value that is smaller than all the terms. As a result, the solution is obtained by equating all the terms in (12). The corresponding system of equations is solved in [2], and the resulting distortion exponent is given by (9).[1] $\qquad\square$

Single-layer superposition coding is the same as a system with no feedback ($K = 1$). On the other hand, infinite-layer superposition coding has the same distortion exponent as the feedback scheme with full CSIT [2]. Theorem 1 shows that for any K, K-layer superposition coding and K-level feedback are equivalent in the distortion exponent sense.

V. Quantizer Design

A. Short-Term Power Constraint

Under the short-term power constraint $\mathcal{P}_i = \mathcal{P}$, and the expected distortion minimization problem is stated as

$$\min_{\{\gamma_i,\ 0 \le i \le K-1\}} \mathcal{E}_D \quad (13)$$
$$\text{s.t.} \quad \gamma_{i+1} - \gamma_i \ge 0$$

The rate maximization problem for noiseless feedback channels has been numerically solved in [8] assuming that the Karush-Kuhn-Tucker (KKT) conditions hold. Due to the lack of convexity, the KKT solution may not be a global optimum. In what follows, we propose an efficient and globally optimal dynamic programming solution to (13).

We assume that each γ_i is optimized over the interval $[0, \gamma_{max}]$ with the step size δ, where $\gamma_{max} = N_\gamma \delta$ for some integer N_γ. Thus, $\{j\delta\}$, $0 \le j \le N_\gamma$, defines the set of values that can be assumed by γ_i. Let $\Delta_D^i(n, j)$ denote the contribution of $\{\gamma_{i-1} = n\delta, \gamma_i = j\delta\}$ to the expected distortion and $\Delta_P^i(n, j)$ be the power cost associated with the pair, then

$$\Delta_D^i(n, j) = [\mathrm{F}(j\delta) - \mathrm{F}(n\delta)]\ D\left(b\log\left(1 + n\delta\mathcal{P}_{i-1}\right)\right) \quad (14)$$

for $1 \le i \le K - 1$, and

$$\Delta_D^K(n, j) = [1 - \mathrm{F}(n\delta)]\ D\left(b\log\left(1 + n\delta\mathcal{P}_{K-1}\right)\right) \quad (15)$$

$$\Delta_P^i(n, j) = \begin{cases} \mathrm{F}(j\delta)\,\mathcal{P}_0, & i = 1 \\ [\mathrm{F}(j\delta) - \mathrm{F}(n\delta)]\,\mathcal{P}_i, & 2 \le i \le K - 1 \\ [1 - \mathrm{F}(n\delta)]\,\mathcal{P}_{K-1}, & i = K \end{cases} \quad (16)$$

for $0 \le n, j \le N_\gamma$. For $\gamma_i = j\delta$, let the optimal value of γ_{i-1} be $s_{ij}\delta$. Also for $\gamma_i = j\delta$, let $\mathcal{E}_D^{ij}$ be the minimum expected distortion contributed by $\{\gamma_0, \cdots, \gamma_i\}$. The dynamic programming principle then implies that

$$\mathcal{E}_D^{ij} = \mathcal{E}_D^{(i-1)s_{ij}} + \Delta_D^i(s_{ij}, j) \quad (17)$$

[1]In the superposition coding scheme of [2], the rates are defined incrementally. That is, if $\gamma \ge \gamma_i$, then $\sum_{n=0}^i R_n$ is recovered. However, the resulting system of equations is algebraically equivalent to our formulation.

for $1 \le i \le K$, $0 \le j \le N_\gamma$ and the initial value $\mathcal{E}_D^{0n} = \mathrm{F}(n\delta)D_0$. The accumulated power cost for (17) is

$$\mathcal{E}_P^{ij} = \mathcal{E}_P^{(i-1)s_{ij}} + \Delta_P^i(s_{ij}, j) \quad (18)$$

for $1 \le i \le K$, $0 \le j \le N_\gamma$ and the initial value $\mathcal{E}_P^{0n} = 0$. For $1 \le i \le K$ and $0 \le j \le N_\gamma$, the survivor index, s_{ij}, is

$$s_{ij} = \arg \min_{0 \le n \le j} \left\{ \mathcal{E}_D^{(i-1)n} + \Delta_D^i(n, j) \right\} \quad (19)$$
$$\text{s.t.} \quad \mathcal{E}_P^{(i-1)n} + \Delta_P^i(n, j) \le \mathcal{P} \quad (20)$$

The $\gamma_{i-1} \le \gamma_i$ constraint is implicitly enforced by the search space $0 \le n \le j$ in (19). If for certain i and j, no solution for s_{ij} can be found, then $\gamma_i = j\delta$ is excluded from the search in the next step of the algorithm. In other words, a trellis path that violates the power constraint is pruned. In the last step of the algorithm, $i = K$ and γ_K assumes only one value, that is $\gamma_K = \infty$ and $j = \infty$. The minimum distortion $\mathcal{E}_D^{K\infty}$ is then the optimal value of the cost function, $\mathcal{E}_D^\star$. The optimal solution $\{\gamma_i^\star\}$ is recovered using the survivors $\{s_{ij}\}$ and the backward pass of the Viterbi algorithm.

B. Long-term Power Constraint

Under a long-term power constraint, the distortion minimization problem is formally stated as

$$\min_{\{\gamma_i, \mathcal{P}_i,\ 0 \le i \le K-1\}} \mathcal{E}_D \quad (21)$$
$$\text{s.t.} \quad \gamma_{i+1} - \gamma_i \ge 0$$
$$\mathcal{E}_P \le \mathcal{P}, \quad \mathcal{P}_i \ge 0$$

We solve this constrained optimization problem using the Lagrange multiplier technique. Define the Lagrangian $\mathcal{L} = \mathcal{E}_D + \lambda\mathcal{E}_P$, where $\lambda \ge 0$ is the Lagrange multiplier. Minimization of $\mathcal{L}$ is performed iteratively by alternating between the optimization over $\{\gamma_i\}$ given $\{\mathcal{P}_i\}$, and vice versa. Note that the dynamic programming technique of Section V-A is formulated for any given power allocation, and as a result, it can be used to optimize $\{\gamma_i\}$ when $\{\mathcal{P}_i\}$ is given.

For a given $\{\gamma_i\}$, the optimization over $\{\mathcal{P}_i\}$ is done by setting the derivatives $\frac{\partial\mathcal{L}}{\partial\mathcal{P}_i}$ to zero, that results in

$$\frac{b\omega_i\gamma_i}{\ln(2)(1+\gamma_i\mathcal{P}_i)}D'\left(b\log\left(1+\gamma_i\mathcal{P}_i\right)\right)+\lambda\nu_i = 0, \qquad 0 \le i \le K-1 \quad (22)$$

where $\nu_0 = \mathrm{F}(\gamma_1)$ and $\nu_i = \omega_i = \mathrm{F}(\gamma_{i+1}) - \mathrm{F}(\gamma_i)$ for $i > 0$. For a given λ, (22) can be numerically solved to obtain $\mathcal{P}_i$ and then a search over λ ensures that the power constraint is satisfied. For a Gaussian source, (22) simplifies to

$$\mathcal{P}_i = \left[\frac{1}{\gamma_i}\left(\frac{b\omega_i\gamma_i}{\lambda\nu_i} \right)^{\frac{1}{b+1}} - \frac{1}{\gamma_i} \right]^+ \quad (23)$$

As $K \to \infty$, $\gamma_0 \to 0$ and $\frac{\omega_i}{\nu_i} \to 1$ for all i. In this case, (23) becomes a discrete version of (8).

Finally, when the power allocation is additionally constrained by a maximum power constraint $\mathcal{P}_i \le \mathcal{P}_m$, the water-filling operation is performed up to the maximum power, and the solution is given as [8]
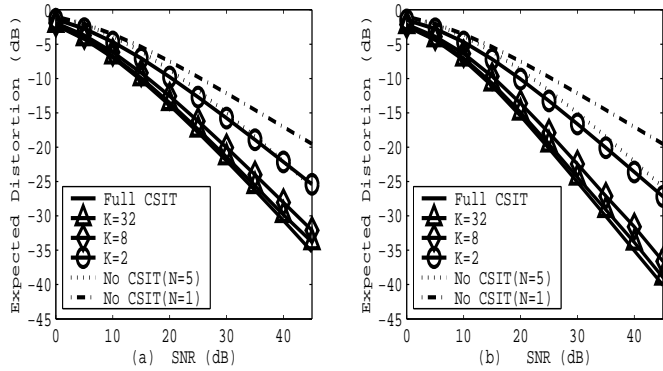
Fig. 2. Expected distortion for a SISO channel with $b = 1$ (a) Short-term power constraint (b) Long-term power constraint

$$\mathcal{P}_i = \left[\min\left(\frac{1}{\gamma_i}\left(\frac{b\omega_i\gamma_i}{\lambda\nu_i}\right)^{\frac{1}{b+1}} - \frac{1}{\gamma_i}, \mathcal{P}_m\right)\right]^+ \qquad (24)$$

VI. NUMERICAL RESULTS

In this section, we provide the numerical results of this paper. Throughout this section, we assume a SISO Rayleigh fading channel. The case of a SIMO channel is similar and does not introduce additional insight. Performance results without feedback are based on superposition coding with N layers [6]. Note that the feedback scheme with $K = 1$ is equivalent to superposition coding with $N = 1$, and there is only one threshold (γ_0) to be optimized in both systems. Beyond 5 layers, the performance improvement offered by additional layers in the superposition scheme is negligible [6]. As a result, in our numerical experiments we consider the no-CSIT scenario with $N = 5$ practically equivalent to a system with an infinite number of layers, or an optimal no-CSIT scheme.

Figure 2 shows the performance results for $b = 1$. The performance bounds with perfect CSIT are also shown. We note that even one bit of feedback provides gains of 5 dB and 7 dB, respectively, with the short-term and long-term power constraints. It can be seen from the figure that 5 bits of feedback ($K = 32$) performs very close to a system with perfect CSIT over a wide range of signal to noise ratios. This is expected because according to Theorem 1, the distortion exponent of a system with 5-bits of feedback is the same as 5-layer superposition coding. The 5-layer superposition coding, in turn, almost achieves the infinite-layer performance that has the same exponent as the full CSIT system. Also, at very large SNRs, a system with 1-bit feedback underperforms 5-layer superposition coding. This is again a direct result of Theorem 1, since the distortion exponent of 1-bit feedback is equivalent to that of two-layer superposition coding. Despite its asymptotic limitations, the results of this experiment nevertheless show the effectiveness of low-rate feedback for practical values of the SNR. Moreover, the results motivate the combination of quantized feedback and superposition coding, originally

suggested in [8], to overcome the asymptotic limitations of quantized feedback.

VII. CONCLUSION

In this paper, the problem of transmitting a Gaussian source over a single-input multiple-output quasi-static fading channel with a delay constraint was considered. An upper bound on the performance was derived assuming perfect channel knowledge at the transmitter, and it was shown that for practical values of the channel signal to noise ratio, this bound can be achieved with a very limited knowledge of the channel quality. The performance bound also indicated that unlike the rate maximization problems, temporal power adaptation at the transmitter is very effective in minimizing the distortion, especially for large bandwidth expansion ratios. The distortion exponent of a system with a K-level feedback and no power adaptation was derived and shown to be the same as the distortion exponent of a system without feedback that employs K-layer superposition coding. An efficient and optimal numerical algorithm for feedback design was presented and the effectiveness of feedback was shown numerically for a Rayleigh fading channel.

REFERENCES

[1] J. N. Laneman, E. Martinian, G. W. Wornell, J. G. Apostolopoulos, "Source-Channel Diversity for Parallel Channels," *IEEE Trans. Info. Th.*, Oct. 2005.
[2] D. Gunduz, E. Erkip, "Distortion Exponent of MIMO Fading Channels," *Info. Th. Work.*, Mar. 2006.
[3] G. Caire, K. Narayanan , "On the SNR Exponent of Hybrid Digital-Analog Space Time Coding," *Allerton Conf. Comm., Cont., and Comp.*, Sep. 2005.
[4] K. E. Zachariadis, M.L. Honig, A. K. Katsaggelos, "Source Fidelity over Fading Channels: Erasure Codes versus Scalable Codes," *IEEE Globecom*, Nov. 2005.
[5] F. Etemadi, H. Jafarkhani, "Optimal Layered Transmission Over Quasi-Static Fading Channels," *IEEE Int. Symp. Info. Th.*, Jul. 2006.
[6] F. Etemadi, H. Jafarkhani, "Optimal Rate and Power Allocation for Layered Transmission with Superposition Coding," *IEEE Data Comp. Conf.*, Mar. 2007.
[7] S. Shamai, A. Steiner, "A Broadcast Approach for a Single-User Slowly Fading MIMO Channel," *IEEE Trans. Info. Th.*, Oct. 2003.
[8] T. T. Kim, M. Skoglund, "On the Expected Rate of Slowly Fading Channels with Quantized Side Information," *Asilomar Conf. Sig., Sys., and Comp.*, Nov. 2005.
[9] S. Ekbatani, F. Etemadi, H. Jafarkhani, "Transmission Over Slowly Fading Channels Using Unreliable Quantized Feedback," *IEEE Data Comp. Conf.*, Mar. 2007.
[10] J. C. Roh, B. D. Rao, "Multiple Antenna Channels with Partial Channel State Information at the Transmitter," *IEEE Trans. Wirel. Comm.*, Mar. 2004.
[11] G. J. Foschini, M. Gans, "On the Limits of Wireless Communication in a Fading Environment When Using Multiple Antennas," *Wirel. Pers. Comm.*, Mar. 1998.
[12] I. M. Gelfand, S. V. Fomin. *Calculus of Variations.* Dover Pub., 2000.
[13] A. J. Goldsmith, P. P. Varaiya, "Capacity of Fading Channels with Channel Side Information," *IEEE Trans. Info. Th.*, Nov. 1997.
[14] L. Zheng, D. N. C. Tse, "Diversity and Multiplexing: A Fundamental Tradeoff in Multiple-Antenna Channels," *IEEE Trans. Info. Th.*, May 2003.