Throughput Maximization Over Slowly Fading Channels Using Quantized and Erroneous Feedback

Siavash Ekbatani, Farzad Etemadi, and Hamid Jafarkhani

Abstract

We study rate and power adaptation of data codewords over slowly fading channels, when quantized and erroneous channel state information (CSI) is available to the transmitter. The goal is to maximize the expected data rate using superposition (multi-layer) coding and power control at the transmitter. The proposed CSI quantizer structure resembles channel optimized scalar quantizers (COSQs) with a newly introduced quasi-grey bit-mapping scheme. Our results show that superposition coding provides significant gains when feedback is more erroneous or channel uncertainty at the transmitter is high, whereas power control is more effective with more reliable feedback.

I. INTRODUCTION

In practical communications, delay requirements limit the transmission codewords to a finite number of fading blocks and block fading Gaussian channel models must be adopted [1]. In a slowly fading environment, where each codeword is transmitted over a single fading block, the conventional performance measure is the expected rate versus the outage probability [2]. Acquiring channel state information at the transmitter (CSIT) requires a feedback link and the optimal transmission scheme depends on the available CSIT. At one extreme, when perfect CSIT is available, outage-free transmission is possible at a rate equal to the current realization of the channel mutual information. Moreover, the transmission rate can be further improved if temporal power control is utilized at the transmitter [3]. On the other hand, when no CSIT is available, the best transmission strategy is superposition (multi-layer) coding [4]–[6]. Between these two extremes, one can consider a scenario where only a quantized version of the CSIT is available [8], [10]. Using quantized CSIT with error-free feedback reduces the transmitter’s uncertainty about the channel realization to a single quantization bin. Therefore, superposition coding over each quantization bin is possible [8]. As the number of feedback regions increases, the channel uncertainty decreases and superposition coding becomes less effective. Feedback channels are severely bandwidth and power limited and time sensitive. We intend to reduce the time duration of feedback transmission over the reverse link and also to avoid redundant channel codes and extra resources at the feedback link. In a practical system, all these factors can result in feedback errors. In this letter, we use the framework of [8] together with a noisy feedback link model. With this model, the CSI uncertainty at the transmitter is similar to a no-CSIT scenario, and moreover, we have an additional source of error in the system. Therefore, without a prudent design, ignoring the feedback errors, the system performance may degrade severely [11]. We design a CSI quantizer by joint source-channel coding at the feedback link, using a structure similar to channel optimized scalar quantizers (COSQs) [12], [13]. To facilitate the design process, we also define a particular bit-mapping scheme, named quasi-grey bit-mapping. Our results show that by designing proper COSQs, we can still outperform a no-CSIT system.

This work was supported in part by an ARO Multi-University Research Initiative (MURI) grant # W911NF-04-1-0224.
with any feedback error probability. Moreover, with higher feedback errors, superposition coding becomes more useful, but the power control gain vanishes.

In Section II, we introduce the transmission scheme, the quantizer structure, and the feedback channel model. In Section III, we define the quantizer design problem and present our solution techniques. Section IV provides some numerical results and finally, Section V concludes the paper.

Notations: In the following, \( P(e) \) represents the probability of event \( e \) and \( P(e_1, e_2) \) represents the joint probability of events \( e_1 \) and \( e_2 \). The CDF of a random variable \( X \) is represented by \( F(x) \), where \( x \) denotes the realization of \( X \). Furthermore, all logarithms are natural unless stated otherwise.

II. SYSTEM MODEL AND TRANSMISSION SCHEME

Fig. 1 shows the block diagram of our transmission scheme. We consider a degraded channel \(^1\), such as a single-input multiple-output (SIMO) fading channel, and we express the discrete-time transmit/receive signal model as

\[
y(t) = h x(t) + n(t) : t = 1, \cdots, T
\]  

(1)

where \( t \) represents the index of channel use over a fading block duration. Across a SIMO channel, the transmit signal \( x(t) \) is a scalar and \( h \) and \( y(t) \) are the \( 1 \times M_r \) channel vector and the received signal in the \( t^{th} \) channel use, respectively. The components of the channel vector \( h \) are independent and identically distributed (i.i.d) with some known distribution. Moreover, \( n(t) \) is a circularly symmetric complex additive white Gaussian noise (AWGN) vector with variance 1. Each transmission codeword is limited to a single fading block and the block length \( T \) is large enough, so that transmission codewords with a rate up to the instantaneous mutual information of the channel can be decoded successfully by completely averaging out the noise.

A. Quantizer Structure, Feedback Channel Model, and Bit-mapping

We assume that the receiver has perfect knowledge of the channel vector \( h \). The magnitude of \( h \), defined as its Frobenius norm square, \( \gamma \triangleq ||h||_F^2 \) defines the reliable communication rate over this channel and different channel realizations can be ordered by their magnitudes in terms of the achievable rate. In our system, the variable \( \gamma \) is quantized using a \( K \)-level scalar quantizer defined by the encoder and decoder mappings \( \Phi_e : \mathbb{R} \to \mathcal{J} \) and \( \Phi_d : \mathcal{I} \to \{\gamma_i\} \), respectively. Here, \( \mathbb{R} \) is the field of real numbers and \( \mathcal{J} = \mathcal{I} = \{0, 1, \cdots, K-1\} \) represent the encoder and decoder index sets in the quantizer structure. At the receiver side, the Voronoi regions of the quantizer’s encoder are specified by the partition boundaries \( \gamma^{b}_{0} = 0 < \gamma^{b}_{1} \leq \cdots \leq \gamma^{b}_{K-1} < \gamma^{b}_{K} = \infty \), and the quantizer mapping is defined as \( \Phi_e(\gamma) = j \), if \( \gamma \in [\gamma^{b}_{j}, \gamma^{b}_{j+1}) \). The encoder index \( j \in \mathcal{J} \) is then conveyed back to the transmitter through the feedback link. Unlike the original model of [8], we allow noise in the feedback channel and as a result, the index \( i \in \mathcal{I} \) could be different from \( j \).

Let us define the index transition probabilities, \( p(i|j) \), as the probability of receiving index \( i \) at the transmitter, given index \( j \) was sent from the receiver. In this work, we model the noisy feedback link as a discrete memoryless channel (DMC) based on \( \log_2 K \) uses of binary symmetric channels (BSCs) carrying every feedback bit, with the bit error (cross-over) probability \( \rho_f \). Moreover, we assume that \( \rho_f \) is a constant, known a priori at the design step. A more elaborate finite-state feedback channel model is studied

\(^1\)Superposition is not proven to be optimal over non-degraded channels, such as MIMO channels [6]
in [11], along with the effects of mismatches between the design assumptions and real feedback channel conditions. For a given cross-over probability $\rho_f$, the index transition probabilities are given as $p(i|j) = \rho_f^{d_{i,j}} (1 - \rho_f)^{\log_2(K) - d_{i,j}}, \forall i, j$, where $d_{i,j}$ is the Hamming distance between the binary representations of $i$ and $j$.

To simplify the design process, we propose the following bit-mapping scheme to be used to map feedback indices to binary bits. We discuss the reasonings behind this choice in lemma 1 of the next section.

**Proposition 1**: Assume a specific class of bit-mappings where the index transition probability matrix of the DMC has the following properties:

$$
\forall \ell \neq j : \sum_{k=j}^{K-1} p(k|j) \geq \sum_{k=j}^{K-1} p(k|\ell) \quad (2)
$$

$$
\forall \ell > j, \forall m \leq K - 1 : \sum_{k=m}^{K-1} p(k|\ell) \geq \sum_{k=m}^{K-1} p(k|j) \quad (3)
$$

We call a mapping scheme with properties (2)-(3), a *quasi-grey bit-mapping*. Some examples of this bit-mapping with different quantizer sizes, $K$, are provided in [16].

### B. Transmission Scheme, Expected Rate, and Expected Power with Superposition Coding

For superposition coding with $L$ code layers, the quantizer’s decoding operation is performed at the transmitter as a one-to-many mapping $\Phi_d(i) = \{\gamma_{i,\ell}, \mathcal{P}_{i,\ell}\}, \ell = \{0, \cdots, L - 1\}$. Upon receiving index $i$, $0 \leq i \leq K - 1$, the transmitter sends the sum of $L$ Gaussian codewords with rates [4], [8]

$$
R_{i,\ell} = \log \left( 1 + \frac{\gamma_{i,\ell} \mathcal{P}_{i,\ell}}{1 + \gamma_{i,\ell} \sum_{k=\ell+1}^{L-1} \mathcal{P}_{i,k}} \right) : 0 \leq \ell \leq L - 1
$$

nats per channel use, where $\{\gamma_{i,\ell}\}$ represents the channel reconstruction points and $\mathcal{P}_{i,\ell}$ denotes the allocated power to the $\ell$th codeword. The total transmit power in quantization bin $i$ is $\mathcal{P}_i = \sum_{\ell=0}^{L-1} \mathcal{P}_{i,\ell}$. The fundamental assumption in applying this coding scheme is $\gamma_{i,\ell} \geq \gamma_{i,\ell-1}$. Then different code layers can be decoded successively from the highest rate to the lowest and subtracted from the received signal.

With error-free (noiseless) CSIT, the uncertainty associated with index $i$ is confined to $[\gamma_i^b, \gamma_{i+1}^b]$ and consequently, it is further assumed that $\gamma_{i,\ell} \in [\gamma_i^b, \gamma_{i+1}^b]$ for $\ell = \{0, \cdots, L - 2\}$ and $\gamma_{i+1}^b \geq \gamma_{i,(L-1)}$. Note, however, that with noisy CSIT, the latter constraints cannot be valid and the expanded regions of uncertainty lead us to relax these additional structural constraints. Therefore, the reconstruction points of the lower quantization bins are allowed to overlap with the upper bins, e.g., $\gamma_{i,\ell} > \gamma_{i+1}^b$. Therefore, the pertinent quantizer design should be a generalization of the quantizer introduced in [8]. Note that this generalization can significantly increase the design complexity.

We assume that the index $i$ can be reliably communicated to the receiver, like a “*genie*”, with a negligible rate loss of $\lim_{T \to \infty} \frac{\log_2(K)}{T}$ bits per channel use. In theory, the forward link can take infinitely large number of channel uses. Therefore, communicating a limited number of bits to introduce the index $i$ to the receiver is always possible. Conceptually, there are two ways to do that: i) orthogonal or time-multiplexed control/data symbol transmission [14] and ii) nonorthogonal or superimposed control/data symbols [15]. As a result,
the receiver can be aware of the data rate of the transmitted code layers and successive decoding of code layers \( \ell = L - 1 \) to \( \ell = 0 \) can be performed by sequentially subtracting each codeword from the received signal if \( \gamma \geq \gamma_{i\ell} \). Considering the negligible rate loss for communicating \( i \) to the receiver, we note that our results provide a tight upper bound on the actual reliable expected transmission rate.

Under a \textit{long-term (average) power constraint}, we allow the transmitter to choose a transmission power \( P_i = \sum_{\ell=0}^{L-1} P_{i\ell} \) for index \( i \), while the average transmission power is constrained to \( P \). The expected transmission power can be written as

\[
E_p = \sum_{i=0}^{K-1} P(i) P_i = \sum_{i=0}^{K-1} \sum_{j=0}^{K-1} p(i|j) P(j) P_i
\]

\[
= \sum_{i=0}^{K-1} \sum_{j=0}^{K-1} p(i|j) \left[ F(\gamma_{j+1}^b) - F(\gamma_j^b) \right] \sum_{\ell=0}^{L-1} P_{i\ell}
\]

where \( F(\gamma) \) shows the CDF of the random variable \( \gamma \). Therefore, the long-term power constraint can be expressed as \( E_p \leq P \).

Transmission at rate \( R_{i\ell} \) can be reliably decoded if index \( i \) appears at the transmitter and \( \gamma \geq \gamma_{i\ell} \). Therefore, the achievable expected rate is \( E_R = \sum_{i=0}^{K-1} \sum_{\ell=0}^{L-1} P(\gamma \geq \gamma_{i\ell}, i) R_{i\ell} \) where \( R_{i\ell} \) was defined in Equation (4). The decoding event probability \( P(\gamma \geq \gamma_{i\ell}, i) \) depends on the location of the reconstruction point \( \gamma_{i\ell} \) among quantization bins. This probability can be expressed as

\[
P(\gamma \geq \gamma_{i\ell}, i) = \sum_{j=0}^{K-1} 1_j(\gamma_{i\ell}) \left[ P\left( \gamma_{i\ell} \leq \gamma \leq \gamma_{j+1}^b, i \right) + P\left( \gamma \geq \gamma_{j+1}^b, i \right) \right]
\]

with the indicator function \( 1_j(\gamma_{i\ell}) = 1 \) if \( \gamma_{i\ell} \in [\gamma_j^b, \gamma_{j+1}^b) \) and 0 otherwise. Moreover, in (6), we have \( P(\gamma_{i\ell} \leq \gamma \leq \gamma_{j+1}^b, i) = p(i|j) \left[ F(\gamma_{j+1}^b) - F(\gamma_{i\ell}) \right] \) and \( P(\gamma \geq \gamma_{j+1}^b, i) = \sum_{n=j+1}^{K-1} p(i|n) \left[ F(\gamma_{n+1}^b) - F(\gamma_{n}^b) \right] \). Therefore, the achievable expected rate can be expressed as

\[
E_R = \sum_{i=0}^{K-1} \sum_{\ell=0}^{L-1} \mu_{i\ell} R_{i\ell}
\]

where we can use the definition of the \textit{rate reward} \( \mu_{i\ell} \) [4], [8]:

\[
\mu_{i\ell} = \sum_{j=0}^{K-1} 1_j(\gamma_{i\ell}) \left\{ p(i|j) \left[ F(\gamma_{j+1}^b) - F(\gamma_{i\ell}) \right] + \sum_{n=j+1}^{K-1} p(i|n) \left[ F(\gamma_{n+1}^b) - F(\gamma_{n}^b) \right] \right\}
\]

\textbf{III. QUANTIZER DESIGN}

The quantizer design objective is to find the values of the boundaries, \( \{\gamma_{i}^b\} \), the channel reconstruction points, \( \{\gamma_{i\ell}\} \), and the power allocation factors, \( \{P_{i\ell}\} \) to maximize the expected rate, \( E_R \), conditioned on the power constraint, \( E_p \leq P \). This optimization problem can be expressed as

\[
\max \quad E_R \quad \text{s.t.} \quad \gamma_i^b - \gamma_{i0} \leq 0, \quad \gamma_{i\ell} - \gamma_{i(\ell+1)} \leq 0, \quad E_p - P \leq 0
\]

The structure of the optimal quantizer depends on the mapping scheme, \( j' = \delta(j) \) and \( i = \delta^{-1}(i') \). Our design strategy is to start with a fixed mapping scheme that can simplify the design procedure, i.e., the
quasi-grey bit-mapping introduced in (2) and (3). Also, to simplify the design process, specially while dealing with a long-term power constraint, we use the following proposition:

**Proposition 2:** Suppose that the power levels assigned to different quantization bins, \( P_i = \sum_{\ell=0}^{L-1} P_{i\ell} \) satisfy the following constraint:

\[
P_0 \leq P_1 \leq \cdots \leq P_{K-1}
\]

Namely, the power levels are non-decreasing with the quantization indices. In other words, if power control is allowed in the system, we assign more power to indices representing the stronger channels.

We won’t claim the optimality of the quantizer structure designed based on the above propositions. However, we will show that the proposed design demonstrates desirable properties of the expected rate of the system. Using propositions 1 and 2, the following lemma shows that we can notably reduce the complexity of the quantizer design algorithm.

**Lemma 1:** Using a quasi-grey bit-mapping scheme and with nondecreasing power levels in each bin, the optimal solutions of Equation (9) satisfy \( \gamma_{i}^{*} = \gamma_{i0}^{*} \) for \( i = 1, 2, \cdots, K - 1 \). In other words, each quantizer boundary coincides with the smallest reconstruction point of the corresponding bin.

**Proof:**

We use contradiction to prove the statement of the Lemma. Consider the two quantizer structures shown in Fig. 1. Suppose that the optimal quantizer structure resembles Quantizer A, where for an arbitrary index \( j \in \{1, 2, \cdots, K - 1\} \), we have \( \gamma_j^{b} < \gamma_j^{0} \), thus there is a non-empty interval \( \gamma \in \Psi_{j} = [\gamma_j^{b}, \gamma_j^{0}] \), with feedback index \( j \) at the receiver. We compare the performance of this system with the one using Quantizer B that possesses the same structure as Quantizer A, except that \( \gamma_j^{0} = \gamma_{j0}^{*} \).

Using Quantizer A, for the channel realizations in \( \Psi_{j} \), the probability of outage or the probability that the transmission codeword does not contribute to the aggregate rate of the system can be expressed as \( P_{out}^{A} = \sum_{k=0}^{K-1} p(k|j) \). With Quantizer B, the same channel realizations experience the outage probability \( P_{out}^{B} = \sum_{k=0}^{K-1} p(k|j-1) \). According to the property (2), we can confirm that \( P_{out}^{A} \geq P_{out}^{B} \).

Moreover, with Quantizer A, the average transmission power given a channel realization in \( \Psi_{j} \) occurs can be written as \( P_{av}^{A} = \sum_{k=0}^{K-1} p(k|j)P_{k} \). With Quantizer B, the same channel realizations incur the average transmission power equal to \( P_{av}^{B} = \sum_{k=0}^{K-1} p(k|j-1)P_{k} \). Now, assuming proposition 2, i.e. the power codebook \( \{P_{i}\} \) is non-decreasing, (10), and noting the property (3) of the quasi-grey bit-mapping, it is straightforward to conclude that \( P_{av}^{A} \leq P_{av}^{B} \).

The rest of the regions, other than \( \Psi_{j} \) behave similarly with Quantizers A and B. Therefore, Quantizer B provides greater aggregate data rate compared to Quantizer A, with less power consumption. This contradicts the optimality of Quantizer A and the proof is complete.

Using Lemma 1, we can replace the boundary variables \( \{\gamma_{i}^{b}\} \) with the minimum reconstruction points of the corresponding regions, \( \gamma_{i0}^{*} \). As a result, the optimization variables of the expected rate cost function, \( \mathcal{E}_{R} \), can be reduced to \( \{\gamma_{i\ell}^{*}, P_{i\ell}\} \). Then, the long-term power constraint and the expected rate objective function can be respectively expressed as

\[
\mathcal{E}_{P} = \sum_{i=0}^{K-1} \nu_i \sum_{\ell=0}^{L-1} P_{i\ell} \leq \mathcal{P} : \quad \nu_i = p(i|0) \mathcal{F} \left( \gamma_{i0} + \sum_{j=1}^{K-1} p(i|j) \left[ \mathcal{F} \left( \gamma_{(j+1)0} \right) - \mathcal{F} \left( \gamma_{j0} \right) \right] \right)
\]

and

\[
\mathcal{E}_{R} = \sum_{i=0}^{K-1} \sum_{\ell=0}^{L-1} \mu_{i\ell} \mathcal{R}_{i\ell} : \quad \mu_{i\ell} = \sum_{j=0}^{K-1} \mathbf{1}_{j}(\gamma_{i\ell}) \left\{ p(i|j) \left[ \mathcal{F} \left( \gamma_{(j+1)0} \right) - \mathcal{F} \left( \gamma_{i\ell} \right) \right] + \sum_{n=j+1}^{K-1} p(i|n) \left[ \mathcal{F} \left( \gamma_{(n+1)0} \right) - \mathcal{F} \left( \gamma_{n0} \right) \right] \right\}
\]

Note also that our quantizer design process involves using contiguous Voronoi regions, which is not proven to be optimal.
Similar to [8], we adopt an iterative solution of the above multi-variable optimization problem using the following algorithm:

1) For a given power allocation set, \( \{P_{i\ell}\} \), find the reconstruction points \( \{\gamma_{i\ell}\} \) for \( 0 \leq i \leq K - 1 \) and \( 0 \leq \ell \leq L - 1 \). Assume that the initial power allocation is uniform \( (P_{i\ell} = P/L) \). The coupling between the quantization bins in (12) makes the optimization challenging. However, for a fixed set of \( \{\gamma_{i0}\}, i = 1, 2, \ldots, K - 1 \), the expected rate expression (12) decouples for \( \gamma_{00} \) and \( \gamma_{i\ell}, i = 0, 1, \ldots, K - 1, \ell = 1, 2, \ldots, L - 1 \). Therefore, use exhaustive search over the set \( \{\gamma_{i0}\}, i = 1, 2, \ldots, K - 1 \). In each iteration of the search, find the maximizers of the objective function separately with respect to \( \gamma_{00} \) and \( \gamma_{i\ell}, i = 0, 1, \ldots, K - 1, \ell = 1, 2, \ldots, L - 1 \), searching over \([\gamma_{i(\ell-1)}, \infty)\) in the \( \ell = 1 \) to \( \ell = L - 1 \) order. Note that as defined in the description of the quantizer structure in Section II-A, the ordering of the variables, \( \gamma_{i0} \geq \gamma_{(i-1)0}, i = 1, 2, \ldots, K - 1 \) can reduce the complexity of the search.

2) For a given set \( \{\gamma_{i\ell}\} \), find the optimal \( \{P_{i\ell}\} \) for \( 0 \leq i \leq K - 1 \) and \( 0 \leq \ell \leq L - 1 \). In this step, a multi-step numerical search should be used [8]. Choose a Lagrange multiplier \( \lambda \) and invoke the solution of optimal power allocation over parallel Gaussian channels from [4] to solve

\[
\max_{\{P_{i\ell}\}} \sum_{i=0}^{K-1} \sum_{\ell=0}^{L-1} \mu_{i\ell} R_{i\ell} \\
\text{s.t.} \quad \sum_{i=0}^{K-1} \nu_{i} \left( \sum_{\ell=0}^{L-1} P_{i\ell} \right) \leq P
\]

Given \( i \) and \( \lambda \), find the optimal \( P_{i} \), sequentially in the order of \( i = K - 1, K - 2, \ldots, 0 \), using

\[
P_{i} = \max \left\{ \left[ \max_{\ell} \left( \frac{\mu_{i\ell}}{\nu_{i}\lambda} - \frac{1}{\gamma_{i\ell}} \right) \right], P_{i+1} \right\}
\]

where \([x]^{+} = \max(x, 0)\). The above solution confines the space of valid quantization bin powers to satisfy (10). Next, for the given \( \lambda \), define the utility function \( U(z) \) as

\[
U(z) = \arg \max_{\ell} \frac{\mu_{i\ell}}{\gamma_{i\ell} + z} - \lambda : z \in [0, P_{i}]
\]

The optimal power allocated to Layer \( \ell \) is the length of the interval defined as [4], [8]

\[
P_{i\ell}^{*} = \max_{z \in [0, P_{i}]} z - \min_{z \in [0, P_{i}]} z \\
\text{s.t.} \quad U(z) = \ell \quad \text{s.t.} \quad U(z) = \ell
\]

In each step of the algorithm, given \( \lambda \), if the consumed power from (11) is greater (less) than the power constraint, increase (decrease) \( \lambda \) and reiterate until the power constraint is satisfied with equality.

3) Iterate between steps 1) and 2) until convergence. In each step, the expected rate either increases or remains unchanged. As a result, the algorithm converges to a locally optimal solution, since the expected rate is bounded above.

IV. NUMERICAL RESULTS

This section presents the properties of the solutions to the rate maximization problem (9), with the discussed simplifications. First in Fig. 2, we plot the reconstruction points of the optimized COSQs designed for a \( 1 \times 1 \) channel, with \( L = 2 \) layer superposition coding and \( K = 2 \) feedback regions. The leftmost plot is the noiseless feedback quantizer and the rightmost one is the no-CSIT solution. As the quality of the feedback link degrades (error probability increases), the resulting quantizer converges to the no-CSIT solution. Moreover, for extremely noisy systems, the reconstruction points associated with different
quantization bins overlap. For example, $\gamma_{10} < \gamma_{01}$ for $\rho_f = 0.25$. Note that had we restricted our system to non-overlapping regions, the possibility of converging to a no-CSIT solution would have been diminished. Converging to no-CSIT solution can be interpreted as follows. As the feedback becomes less reliable, the source coding effect of the COSQ diminishes and the feedback bit is utilized to protect the small amount of channel information that feedback carries.

Fig. 3 shows the performance of the same system with a short-term (instantaneous) power constraint, where $P_i = P, \forall i$. In this figure, the performance of the system with $L = 1$ and $L = 2$ code layers and different feedback error values are shown. The following observations can be made here. First, the performance of the system with noisy feedback falls between those of no-CSIT and error-free feedback systems. Second, the gain associated with superposition coding (layering gain) increases as the residual uncertainty (post feedback) increases, as in noisy feedback and the extreme case of no-feedback. In Fig. 4, we investigate the properties of the solutions in more details. This figure shows the absolute gain of superposition coding versus the transmission power. When a short-term power constraint is used, the absolute gain is the same as the layering gain, i.e., the gain of superposition coding with respect to single-layer coding. When long-term power control is used, however, the absolute gain has an additional component due to temporal power adaptation, i.e., the power gain. The following observations can be made from Fig. 4. First, as expected, the power gain diminishes with increasing the SNR, since the system becomes less power constrained. This can be seen from the convergence of short-term and long-term curves at high-SNR. Second, as the feedback quality degrades, the layering gain increases. This is due to the fact that for higher feedback error probabilities, the ambiguity regarding the real channel realization is increased and layering, as a way of utilizing uncertainty, becomes more effective. Furthermore, at low-SNR, the absolute gain of high-quality feedback channels are greater. This can be explained noting that when power is scarce, power adaptation is more effective, provided that the feedback is reliable. In contrast, power adaptation loses its significance in the high-SNR regime and layering gain, which is more effective for systems with poor feedback, becomes dominant.

V. CONCLUSIONS

We proposed a quantizer structure for superposition coding with partial CSIT that generalized the existing structure in the literature. In the presence of feedback channel errors, the proposed COSQ that is a joint source-channel coding module, can adjust the CSI source coding and feedback channel coding rates for any given feedback channel bit rate and error probability. With low-quality feedback, temporal power control is less effective, but superposition coding provides significant gains and the latter gain increases as the feedback channel quality degrades. These properties result from the properties of the optimized quantizer structure that with highly erroneous feedback converges to a no-CSIT scheme, where temporal power adaptation is impossible, while superposition coding is optimal.

ACKNOWLEDGMENT

The authors would like to thank anonymous reviewers for their constructive comments that have improved the quality of presentation of this paper.

---

4This phenomenon can be observed more clearly, using more feedback bits [16].
REFERENCES

Fig. 2. The reconstruction points of a $1 \times 1$ multi-layer coding system with $K = 2$-level quantizers and $L = 2$-layer codes at SNR = 40(dB) and with different BSC cross-over probabilities. The overlapping of the feedback regions is highlighted.

Fig. 3. Multi-layer coding over a $1 \times 1$ channel with $K = 2$ feedback regions and different (known) BSC cross-over probabilities.
Fig. 4. Absolute gain of a $1 \times 1$ system with $L = 2$-layer superposition coding with short-term (SP) and long-term (LP) power constraints and $K = 2$ feedback regions.