Research Article

Exploiting Transmit Buffer Information at the Receiver in Block-Fading Channels

Dinesh Rajan

Department of Electrical Engineering, Southern Methodist University, Dallas, TX 77205, USA

Correspondence should be addressed to Dinesh Rajan, rajand@engr.smu.edu

It is well known that channel state information at the transmitter (CSIT) leads to higher throughput in fading channels. We motivate the use of transmit buffer information at receiver (TBIR). The thesis of this paper is that having partial or complete instantaneous TBIR leads to a lower packet loss rate in block-fading channels assuming the availability of partial CSIT. We provide a framework for the joint design and analysis of feedback (FB) and feed-forward (FF) information in fading channels. We then introduce two forms of TBIR—statistical and instantaneous—and show the gains of each form of TBIR using a heuristic scheme. For a Rayleigh fading channel, we show that in certain cases the packet error rate reduces by nearly an order of magnitude with just one bit of feed-forward information of TBIR.

1. Introduction

It is well known that the Shannon capacity of a discrete memoryless channel (DMC) does not increase with feedback from the receiver [1]. However, the capacity of fading channels increases with channel knowledge at the transmitter, and the capacity gain has been quantified for both single-antenna [2] and multiple-antenna [3] systems. Capacity with channel state information (CSI) at both transmitter and receiver has been studied [4], and the effect of errors on the channel knowledge has been quantified [5, 6]. Also, see [7] for a comprehensive review of communication through fading channels. The importance of incorporating traffic conditions in physical layer design has been well recognized [8–10], and cross-layer optimization has been an area of active recent research [11–16]. The use of queue information to optimize physical layer design has been investigated in many settings [17–20].

In this paper, we consider delay-bounded transmission of variable bitrate (VBR) traffic through a block-fading channel. We propose to use transmit buffer information at the receiver (TBIR), and discuss an exemplary implementation system. The novelty in the proposed system is twofold. (i) A new transmission system architecture with a feed-forward channel that transmits queue state information from the transmitter to the receiver is introduced, and (ii) the CSI that is sent to the transmitter via a feedback channel is chosen adaptively based on instantaneous or statistical knowledge of TBIR (the terms buffer and queue are used interchangeably in this paper). The design objective is to minimize overall packet loss for a given buffer size under a long-term power constraint.

The main contributions of this paper may be succinctly summarized as follows.

(i) We provide a framework for the joint design and analysis of feedback (FB) and feed-forward (FF) information over a block-fading channel. We consider two specific forms of TBIR, namely, statistical and instantaneous. The statistical or instantaneous FF information is used to adapt the mechanism that generates the FB information. In particular, a scalar quantizer of the channel fading gain is considered to generate the FB information, and this quantizer is computed based on the available TBIR.

(ii) The performance gain resulting from using the statistical FF information to adapt the channel quantizer at the receiver and generate the CSI is quantified. Further, the additional gain of having instantaneous TBIR over statistical TBIR is shown and quantified in

some simple cases. It turns out that having just one additional bit of FF information can provide about 1 dB saving in power.

(iii) The performance gain of using FF information is also quantified using a simple practical adaptive QAM-based multirate transmission scheme.

The use of FF information does not provide any gain in the two extreme cases of full CSIT and no CSIT. When complete CSIT is available, the number of packets to transmit and the transmission power are determined jointly on the channel and buffer conditions. Also, when no CSIT is available, the transmission rate and power cannot be adapted based on channel conditions, and the use of FF information does not provide any performance benefits. However, for finite (nonzero) FB rates, FF information can lead to reduction in average packet loss. At the receiver, the channel state information (typically, fading amplitude) is measured and quantized to a finite number of bits. If transmit buffer information is not available at the receiver, the quantization thresholds are fixed. However, when buffer information is available at the receiver, the quantization thresholds are adapted based on the TBIR available.

The TBIR is applicable in any point-to-point communication system. In this paper, we consider only a frequency division duplex (FDD) system (in time division duplex (TDD) systems, the channel information can be obtained from data received in prior time slots without requiring explicit FB from the receiver to the transmitter, and such systems are not considered here). The proposed design can be implemented very easily in an 802.11-based WLAN system, where a handshaking mechanism (exchanging RTS and CTS packets) is used prior to actual data transmission. There are also schemes which transmit quantized buffer occupancy information to the receiver in multiuser scenarios; the goal in such situations is to provide fairness or throughput guarantees. In this paper, buffer information is sent to the receiver, even in single-user scenarios, to reduce packet error rate by making more efficient use of the channel state information at the receiver. In a multiuser scenario, information on the various users' transmit buffers can be used both for outage reduction (at the physical layer) and to implement fairness (at MAC layer). For uplink transmissions in cellular systems, even though FF information has to be transmitted from the mobile handset, which could have limited resources, the proposed method can be used to additionally ensure fairness among flows. For downlink transmission, since feedback from receiver to transmitter is limited, feed-forward information can be used to reduce packet loss.

For simplicity of analysis, a *memoryless* source with an *i.i.d.* packet arrival distribution is considered. The analysis directly extends to Markovian source arrivals. Although more sophisticated source models exist, it turns out that the analysis is nontrivial even with these simplified models. Hence, we restrict ourselves to such simple sources in this paper. To demonstrate the applicability of the results, a block-fading channel with Rayleigh fading statistics is used. However, the proposed methods are applicable in general for any block-fading channel, with other fading statistics. We consider a system with finite buffer length, which also results in an upper bound on the average packet delays. Further, with finite buffer length, there is a finite probability of buffer overflow. The overall design objective is to minimize the total packet loss rate resulting from buffer overflows and errors in the transmission over the channel.

The remainder of this paper is organized as follows. In Section 2, we present the basic system under consideration. Sections 3, 4, and 5 focus, respectively, on packet loss rate analysis and optimization with no TBIR, statistical TBIR, and instantaneous TBIR. Numerical results are presented in Section 6. Finally, we conclude in Section 7.

2. System Model and Problem Formulation

Consider a time-slotted system in which a_n fixed-size packets arrive at the transmitter during time slot n and are stored in a buffer of L-size packets before transmission. Let q_n and u_n denote, respectively, the number of packets in the buffer and the number of packets transmitted during time slot n. The buffer update is given by $q_{n+1} = \min(q_n + a_n - u_n, L)$. For simplicity of exposition, we consider a memoryless source arrival model with the distribution of packet arrivals given by $\Pr(a_n = l) = c_l$, $l = 0, \ldots, M$, where M is the maximum number of packet arrivals in one time slot. Clearly, for a valid distribution, $c_l \geq 0$ and $\sum_{l=0}^{M} c_l = 1$. The analysis and results in this paper can be easily extended to other traffic models. Using Little's law [21], the finite buffer length imposes an upper bound on the average delay experienced by the traffic. Further, if a first-come first-serve (FCFS) ordering of the packets in the buffer is assumed, along with a work-conserving scheduler, then the finite-length buffer also implies an upper bound on the absolute delay experienced by the packets.

We consider transmission over a block-fading channel, and assume that the length of one time slot equals T_c, the number of symbols in the coherence interval of the channel. The transmit signal $\mathbf{x}_n$ depends on the number of packets transmitted in each time slot and the coding and modulation schemes. The complex received signal $\mathbf{y}_n$ is given by $\mathbf{y}_n = h_n \mathbf{x}_n + z_n$, where z_n is the additive noise which is modeled as being circularly symmetric Gaussian with zero mean and covariance $\sigma^2 I_{T_c}$, and h_n is the channel gain in time slot n. The real and imaginary parts of h_n are assumed to be independent zero-mean Gaussian, each with variance 1/2. The transmit signal $\mathbf{x}_n$, the received signal $\mathbf{y}_n$, and the noise at the receiver $\mathbf{z}_n$ are T_c-dimensional complex vectors, where T_c is assumed to be a positive integer.

The average packet loss (Π) depends on the packet loss due to buffer overflows Π_b and the frame error rate of the actual coding scheme. In this paper, we use the probability of outage (2) to bound the frame error rate. In [22], it is shown that for large T_c, the conditional mutual information $I(\mathbf{y}_n, \mathbf{x}_n | h_n)$, between $\mathbf{x}_n$ and $\mathbf{y}_n$, is a good indicator of the performance of practical codes. This mutual information is

given by

$$I(\mathbf{y}_n; \mathbf{x}_n \mid h_n) = T_c \log \left(1 + \frac{P_n |h_n|^2}{\sigma^2} \right)$$
$$= T_c \log \left(1 + P_n \gamma_n \right), \tag{1}$$

where $\gamma_n = |h_n|^2/\sigma^2$ is the normalized instantaneous channel gain and P_n is the transmit power during time slot n. Without loss of generality, we let $\sigma^2 = 1$ and hence γ_n has an exponential distribution. Thus, its density function is given by $f_\gamma(x) = e^{-x}$, $0 < x$, where for simplicity the average value of the exponential distribution is assumed to be unity. The probability of outage in the channel Γ during time slot n (which is a good indicator of the frame error rate in practical systems [22]) is given by

$$\Gamma(u_n, \gamma_n) = \Pr\{I(\mathbf{y}_n; \mathbf{x}_n \mid \gamma_n) < R u_n\}, \tag{2}$$

where u_n is the number of packets of size R transmitted in time slot n. Note that in (2) all u_n packets are encoded jointly and transmitted in time slot n. When an outage occurs, u_n packets are lost. Hence, the average packet loss due to outages in the channel equals $\mathbf{E}_{u,\gamma}[u\Gamma(u,\gamma)]$. By using the information theoretically defined outage probability Γ, we abstract away the actual coding scheme used.

In time slot n, a buffer overflow occurs if $q_n + a_n - u_n > L$. Equivalently, buffer overflow occurs if $a_n > L - q_n + u_n$. The probability of buffer overflow is given by $\sum_{(m,l)} \Pr(q_n = m, u_n = l) \sum_{x=L-m+l+1}^{M} c_x$. However, for given q_n and u_n, different values of a_n result in a different amount of lost packets. Thus, the average packet loss due to buffer overflows is given by

$$\Pi_b = \sum_{(m,l)} \Pr(q_n = m, u_n = l) \sum_{x=L-m+l+1}^{M} (x - L + m - l) c_x. \tag{3}$$

In this sequel, we assume that packets that are lost (due to buffer overflows or loss in the channel) are retransmitted as necessary by higher-layer protocols like TCP. Real-time video/audio traffic can tolerate certain amount of lost packets without serious degradation in performance, and in such cases lost packets may not be retransmitted.

The average packet loss (Π) is given by

$$\Pi = \sum_{m=1}^{L} \int_\gamma m \Pr(u_n = m|\gamma) \Pr(\gamma) \Gamma(m, \gamma) d\gamma$$
$$+ \sum_{m=0}^{L} \sum_{l=0}^{m} \Pr(q_n = m) \Pr(u_n = l|q_n = m) \tag{4}$$
$$\times \sum_{x=L-m+l+1}^{M} (x - L + m - l) \Pr(a_n = x).$$

In the first term above, since γ is a continuous variable, we indicate the average operation using an integral. Also, $\Pr(u_n = m/\gamma)$ represents the conditional probability that m packets are transmitted in time slot n when the channel

gain γ_n equals γ. In subsequent analysis, only quantized information on γ is assumed to be available at the transmitter, and the integral is replaced by a summation. The proposed adaptive transmission schemes choose both the instantaneous transmission rate u_n and power P_n. In an ideal system, the transmit power and rate are determined based jointly on knowledge of instantaneous channel fading and buffer state. Traditional approaches to this problem assume a feedback channel of capacity of, say, N_b bits to transmit the channel state to the transmitter. In this paper, we propose a novel architecture in which partial information about the transmit queue is sent to the receiver using a feed-forward channel of capacity of N_f bits. As will become clear from the numerical results, the proposed use of the N_f bits significantly reduces the average packet loss.

The proposed framework for power and rate control is characterized by the three functions f, g, and e. The transmission rate and power are determined by function f, as $(u_n, P_n) = f(q_n, \hat{\gamma}_n)$. (In this paper, we assume that $\hat{\gamma}_n$ is an error-free quantized version of γ_n. The channel estimation error and errors in the feedback channel are ignored.) With some abuse in notation, we will use $f(q_n, \hat{\gamma}_n)$ to represent both the rate and power. The estimate of γ_n at the transmitter is given by $\hat{\gamma}_n = g(\gamma_n, \hat{q}_n)$, where $\hat{q}_n$ is the information about q_n that is sent as feed-forward (FF) information to the receiver, that is, $\hat{q}_n = e(q_n)$. The schematic of the system is given in Figure 1.

In this sequel, we consider a particular class of functions f, g, e which are described in detail in Section 3. The specific form of these functions can be used to calculate the average transmission power and also to evaluate the average packet loss Π.

The optimization problem of interest can be formally stated as follows:

$$\min_{\{f,g,e\}} \Pi$$
$$\text{s.t. } \mathbb{E}[P_n] \leq P_0, \tag{5}$$

where P_0 is the long-term power constraint. The optimization problem is solved for desired values of N_b and N_f, which will result in appropriate constraints on the functions $\{f,g,e\}$. We now discuss a few special cases.

(i) No CSIT. In this case, $N_b = 0$ and $(u_n, P_n) = f(q_n)$ since $\hat{\gamma}_n$ is a constant independent of the actual channel realization (it is possible that $(u_n, P_n) = f(q_n, \mathbb{E}[\gamma_n])$ if statistical channel knowledge is available at the transmitter; as discussed in Section 4, a system with statistical CSIT is similar to a system with statistical TBIR). The analysis of packet loss probability versus delay in this special case is provided in [18, 23].

(ii) Full CSIT. In this case, which is mainly of theoretical interest, $\hat{\gamma}_n = g(\gamma_n, \hat{q}_n) = \gamma_n$ for all $\hat{q}_n$. Thus, $(u_n, P_n) = f(q_n, \gamma_n)$; the outage performance in this case is studied in [24, 25].

As noted earlier, in both of these cases, FF information is not required and does not decrease the packet loss rate.
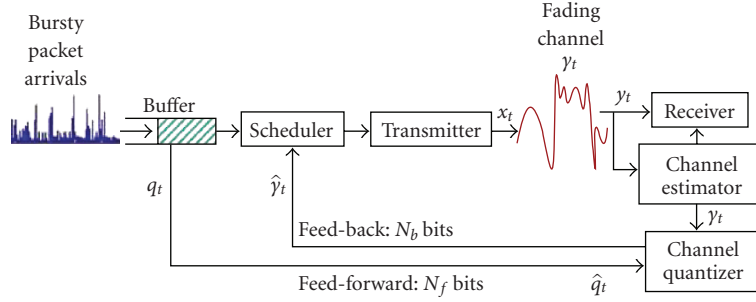
FIGURE 1: Schematic of proposed system incorporating both feedback and feed-forward mechanisms.

(iii) Partial CSIT. This scenario is the main focus of this paper. The analysis and design of FB information are further subdivided into three scenarios: (i) partial CSIT with no TBIR, (ii) partial CSIT with statistical TBIR, and (iii) partial CSIT with instantaneous TBIR. The following sections discuss each of these cases in detail.

3. Partial CSIT: no TBIR

In this section, we derive the performance of a rate and power adaptation scheme in which the feedback information is generated without any knowledge of the packet arrivals, transmit buffer, or delay requirements. The goal is to derive a heuristic approach to solve (5). We first provide details on the channel quantizer design, and then focus on the analysis of the queue at the transmitter. We discuss the computation of the average power, and finally formulate the optimization problem of interest. The results of this section are also useful in formulating the optimization problem in the presence of transmit buffer information at the receiver.

3.1. Channel Quantizer Specification. In this case, since no information about the transmit buffer or traffic arrivals is available at the receiver, the design of the channel quantizer depends only on the channel statistics. Further, the quantization thresholds are chosen to generate a "good" representation of the channel gain. The quantizer thresholds are denoted as β_j, $j = 0, 1, 2, \ldots, 2^{N_b}$. For notational convenience, we let $\beta_0 = 0$ and $\beta_{2^{N_b}} = \infty$. With no TBIR, the β_j coefficients are computed numerically to minimize the mean squared error (MSE) representation of γ_n using the Lloyd-Max algorithm [26]. (For certain source distributions and optimization metrics, the quantizer thresholds β_j can be fully characterized analytically.) The quantized value of the channel state (or gain) is represented as $\hat{\gamma}_n$. In this paper, we use the terms of channel state and channel gain interchangeably. However, in other systems, the transmitter adaptation could be based on the channel phase rather than on amplitude information. The rate and power adaptation are now characterized by the number of packets transmitted for different values of γ_n. Let Y_{ij} denote the number of packets transmitted when there are i packets in the transmit buffer, and the channel gain lies in the jth state, that is,

$q_n = i$ and $\beta_j \leq \gamma_n < \beta_{j+1}$. There is also a natural constraint imposed on the thresholds Y_{ij}, namely, $Y_{ij} \leq Y_{ik}$ for all $k > j$; that is, more packets are transmitted when the instantaneous channel gain γ_n is higher.

For simplicity of exposition and analysis, we map the β_j and Y_{ij} variables into the $\gamma_{k,l}$ variables for $1 \leq l \leq k \leq L$ such that l packets are transmitted during time slot n if $q_n = k$ and $\gamma_{k,l} \leq \gamma_n < \gamma_{k,l+1}$, that is, if buffer has k packets and channel gain lies between certain thresholds. No packets are transmitted if buffer state $q_n = k$ and channel gain $\gamma_n < \gamma_{k,1}$. For notational simplicity, we let $\gamma_{k,0} = 0$ and $\gamma_{k,k+1} = \infty$ for all k. The constraint that Y_{ij} is a nondecreasing function of j implies the following constraint on $\gamma_{k,l}$, namely, $\gamma_{k,l} \leq \gamma_{k,m}$ if $l \leq m$. The thresholding scheme is illustrated in Figure 2. The mapping between $\{Y_{ij}, \beta_i\}$ and $\gamma_{i,j}$ is as follows:

$$\gamma_{i,j} = \beta_k, \quad \text{where } k = \min_{Y_{il}=j} l. \quad (6)$$

In (6), if $k = \phi$ for a given (i, j), then $\gamma_{i,m} = \infty$ for all $m \geq j$. In other words, in buffer state i, the transmission rate never equals or exceeds j packets/time slot.

3.2. Queueing Formulation and Steady-State Analysis. Since we consider stationary models for the traffic arrivals, the channel fading, and the packet transmission policies, the queue state q_n forms a time-homogeneous Markov chain with $(L + 1)$ states, and the steady-state probabilities can be calculated. The transition probabilities p_{ji} between the different queue states are given by $p_{ji} = \Pr\{q_{n+1} = j \mid q_n = i\}$. The transition probabilities are computed as

$$p_{ji} = \begin{cases} \sum_{l=\max[j-i,0]}^{M} c_l \Pr\{u_n \leq i - j + l \mid i\} & \text{if } j = L, \\ \sum_{l=\max[j-i,0]}^{\min[M,j]} c_l \Pr\{u_n = i - j + l \mid i\} & \text{if } j \neq L, \end{cases} \quad (7)$$

where $\Pr(u_n = k \mid i)$, $k = 0, 1, \ldots, L$, is the probability of transmitting k packets in buffer state i and can be computed from the $\gamma_{i,j}$ thresholds. The constraint on the lower bound of l used in the summation in (7) arises from the requirement that to transition from buffer state i to buffer state j, with $j > i$, a minimum of $j - i$ packets must arrive in that time slot. Similarly, the upper bound on l arises from the

0 — β_1	β_1 — β_2	β_2 — β_3	β_3 — ∞	q_n				
$Y_{10}=0$	$Y_{11}=1$	$Y_{12}=1$	$Y_{13}=1$	$q_n=1$	$\gamma_{1,1}=\beta_1$			
$Y_{20}=0$	$Y_{21}=1$	$Y_{22}=1$	$Y_{23}=2$	$q_n=2$	$\gamma_{2,1}=\beta_1$	$\gamma_{2,2}=\beta_3$		
$Y_{30}=0$	$Y_{31}=1$	$Y_{32}=2$	$Y_{33}=2$	$q_n=3$	$\gamma_{3,1}=\beta_1$	$\gamma_{3,2}=\beta_2$	$\gamma_{3,3}=\infty$	
$Y_{40}=0$	$Y_{41}=1$	$Y_{42}=2$	$Y_{43}=3$	$q_n=4$	$\gamma_{4,1}=\beta_1$	$\gamma_{4,2}=\beta_2$	$\gamma_{4,3}=\beta_3$	$\gamma_{4,4}=\infty$
$Y_{50}=1$	$Y_{51}=2$	$Y_{52}=3$	$Y_{53}=4$	$q_n=5$	$\gamma_{5,1}=0$	$\gamma_{5,2}=\beta_1$	$\gamma_{5,3}=\beta_2$	$\gamma_{5,4}=\beta_3,\ \gamma_{5,5}=\infty$

Channel gain γ_n

FIGURE 2: Examples of functions $e(q_n)$, $g(\gamma_n,\hat{q}_n)$, and $f(q_n,\hat{\gamma}_n)$ used when no TBIR or statistical TBIR is available. The corresponding γ_{ij} values are also indicated. Number of feedback bits is $N_b = 2$ and buffer length is $L = 5$.

requirement that when $i > j$ a maximum of j arrivals is allowed. Consequently, we can evaluate p_{ji} as

$$
p_{ji} = \begin{cases} \displaystyle\sum_{l=\max[j-i,0]}^{M} c_l\big(1 - e^{-\gamma_{i,i-j+l+1}}\big) & \text{if } j = L, \\[2ex] \displaystyle\sum_{l=\max[j-i,0]}^{\min[M,j]} c_l\big(e^{-\gamma_{i,i-j+l}} - e^{-\gamma_{i,i-j+l+1}}\big) & \text{if } j \neq L. \end{cases} \tag{8}
$$

The stationary probability of being in buffer state $q_n = i$, denoted by s_i (which is also the invariant distribution of the Markov chain), is then given by

$$
\mathbf{Cs} = \mathbf{s}, \tag{9}
$$

where $\mathbf{s} = \begin{bmatrix} s_0 & s_1 & s_2 & \cdots & s_L \end{bmatrix}'$ and $\mathbf{C}$ is an $(L+1) \times (L+1)$ matrix whose ith row and jth column are p_{ij}.

Thus, the average packet loss Π, which depends on power and rate control policy through the choice of thresholds $\gamma_{i,j}$, is given by

$$
\begin{aligned}
\Pi = {} & \sum_{m=1}^{L} \sum_{\hat{\gamma}} m \Pr(u_n = m, \hat{\gamma}) \Pi_o(m,\hat{\gamma}) \\
& + \sum_{m=0}^{L} \sum_{l=0}^{m} s_m \Pr(u_n = l \mid m) \\
& \times \sum_{x=L+l-m+1}^{M} (x - L + m - l)\Pr(a_n = x)
\end{aligned} \tag{10}
$$

which upon simplification results in

$$
\begin{aligned}
\Pi = {} & \sum_{k=1}^{L} \sum_{i=1}^{2^{N_b}} Y_{ki} s_k \big(e^{-\beta_{i-1}} - e^{-\beta_i}\big) \Pi_o(Y_{ki},\beta_i) \\
& + \sum_{m=0}^{L} \sum_{l=0}^{m} s_m \big(e^{-\gamma_{m,l}} - e^{-\gamma_{m,l+1}}\big) \\
& \times \left(\sum_{x=L+l-m+1}^{M} (x - L + m - l) c_x \right).
\end{aligned} \tag{11}
$$

In this paper, we choose $\Gamma(Y_{ki},\beta_i) = 0$ for all $k = 1,\ldots,L-1$, which is the probability of outage when Y_{ki} packets are transmitted in buffer state k and channel state i. (If we set $\Gamma(m,\beta_i) = \epsilon > 0$, then power P_n can be selected appropriately as $P_n = (e^{lR} - 1)/\hat{\beta}$, where $\int_{\gamma_{k,l}}^{\hat{\beta}} e^{-\gamma} d\gamma = \epsilon$; one such scheme is illustrated in Section 6 using a practical multirate system.) Further, we set $\Gamma(Y_{Li},i) = 0$ for all $i = 2, 3, \ldots, 2^{N_b}$. We consider transmission schemes in which outage occurs in the channel, only when $q_n = L$ and $0 < \gamma_n \leq \beta_1$, that is, $\Gamma(Y_{L1},1) \neq 0$. For all other buffer states and channel gains, packet loss could occur only due to buffer overflows. Qualitatively, the chosen heuristics imply that the only time during which we take a chance on the channel is when the buffer is about to overflow. (Clearly, a more generalized strategy would be to consider more aggressive scheduling for other buffer values also. Such schemes should be considered in future work.) Zero outage in the channel can be guaranteed by transmitting with sufficient power to ensure that the instantaneous mutual information is greater than R (see (12). Note that with no CSIT, zero outage in the channel cannot be guaranteed for all channel fading statistics.

3.3. Average Power Analysis.

Recall that Y_{ij} packets are transmitted in buffer state $q_t = i$ when the channel gain γ_n satisfies the condition $\beta_j \leq \gamma_n < \beta_{j+1}$. The corresponding transmit power that ensures zero outage in the channel is given by

$$
P_n = \frac{e^{Y_{ij}R} - 1}{\beta_j}. \tag{12}
$$

This particular formula for the transmit power is just a restatement of the Gaussian capacity formula [27]. In case $Y_{L1} \neq 0$, then the transmit power when $q_n = L$ and $0 < \gamma_n < \beta_1$ is chosen as

$$
P_n = \frac{e^{Y_{L1}R} - 1}{\hat{\beta}}, \tag{13}
$$

where $0 < \hat{\beta} < \beta_1$ is chosen at the transmitter to satisfy

the power constraint. Clearly, using this transmission power, zero outage in the channel cannot be guaranteed for all $0 < \gamma_n < \beta_1$. Zero outage is only guaranteed for $\hat{\beta} \leq \gamma_n < \beta_1$. The average packet loss can now be rewritten as

$$\Pi = Y_{L1}s_L\left(1 - e^{-\hat{\beta}}\right) + \sum_{m=0}^{L}\sum_{l=0}^{m}s_m\left(e^{-\gamma_{m,l}} - e^{-\gamma_{m,l+1}}\right)$$
$$\times \left(\sum_{x=L+l-m+1}^{M}(x - L + m - l)c_x\right). \tag{14}$$

The average transmit power equals

$$\mathbb{E}[P_n] = \sum_{k=1}^{L}\sum_{l=2}^{2^{N_b}}s_k\left(e^{-\beta_{l-1}} - e^{-\beta_l}\right)\frac{e^{Y_{kl}R} - 1}{\beta_{l-1}}$$
$$+ s_L\left(1 - e^{-\beta_1}\right)\frac{e^{Y_{L1}R} - 1}{\hat{\beta}}. \tag{15}$$

For given Y_{ij} and power constraint P_0, $\hat{\beta}$ can be computed by equating the RHS of (15) to P_0, which is the long-term power constraint. Thus,

$$\hat{\beta} = \frac{s_L\left(1 - e^{-\beta_1}\right)\left(e^{Y_{L1}R} - 1\right)}{P_0 - \sum_{k=1}^{L}\sum_{l=2}^{2^{N_b}}s_k\left(e^{-\beta_{l-1}} - e^{-\beta_l}\right)\left(\left(e^{Y_{kl}R} - 1\right)/\beta_{l-1}\right)}. \tag{16}$$

If RHS of (16) is lesser than 0, then that particular choice of $\{Y_{ij}\}$ cannot be supported with the given buffer constraints. If RHS of (16) is greater than β_1, then it implies that transmitting $Y_{L1} \neq 0$ packets only results in increasing power without any decrease in average packet loss Π for that choice of $\{Y_{ij}, \beta_i\}$.

3.4. Problem Formulation and Solution Methodology. The optimization problem of interest can now be restated as follows:

$$\min_{\{Y_{ij}\}} \Pi$$
$$\text{s.t. } \mathbb{E}[P_n] \leq P_0. \tag{17}$$

Recognize that (17) is a discrete optimization problem, and hence an optimum solution exists and can be computed. (It should be mentioned that one could optimize (17) over all β_i using any other appropriate metric. Since the receiver has no knowledge of buffer, (14)-(15) cannot be used in this particular instance. In the following section, we will optimize over β_i assuming that the receiver has statistical knowledge of buffer.) In this paper, we consider small values of N_b, N_f, and L to illustrate a new concept. Hence, the complexity of solving this optimization problem is not huge. Finding good heuristic solutions to (17) for large system parameters must be considered in future work. The numerical results of the optimization are discussed in Section 6.

The main steps involved in finding the optimal solution may be summarized as follows.

(1) Assume that the number of feedback bits N_b, the β_j coefficients, the power constraint P_0, and the channel fading statistics are given. Create an ordered (lexicographic) set $\mathbb{Y}$ of all feasible combinations of Y_{ij} such that $Y_{ij} \leq Y_{ik}, j < k$, and $Y_{i1} = 0$, $i = 1, \ldots, L - 1$.

(2) Set counter $m = 1$. Consider the mth element of $\mathbb{Y}$.

(3) For that particular combination of Y_{ij}, compute $\hat{\beta}$ that satisfies (15); if no such $\hat{\beta}$ exists, then set the loss probability for this combination equal to 1 and go to step (5).

(4) Compute the total packet loss for the chosen Y_{ij} and the $\hat{\beta}$ computed using (15). Recall that (6) is used to convert between the $Y_{i,j}$ and β_k coefficients.

(5) Set $m = m + 1$. If $m > |\mathbb{Y}|$, then go to step (6); else go to step (3).

(6) Find the minimum value of the total packet loss and the corresponding Y_{ij}.

4. Partial CSIT: Statistical TBIR

In this section, we assume that the receiver has statistical knowledge of the transmit buffer or traffic arrivals. Specifically, we assume that the receiver has knowledge of the packet arrival distribution **c**. Thus, we modify the FB information that is transmitted to better reflect the available knowledge. In particular, we design the channel quantizer in such a way that the overall packet loss is reduced.

It is assumed that the proposed optimization is carried out at the receiver and then the optimal thresholds $\{\beta_i\}$, along with the power and rate adaptive function $f(\cdot)$, are conveyed to the transmitter. Equivalently, one could consider a system where the transmitter has statistical knowledge of the channel statistics. In the latter case, the optimization is performed at the transmitter, and the results are then conveyed to the receiver. Yet another approach might be to have both the transmitter and receiver do the same optimization if they have access to the relevant statistics.

The qualitative reason for the benefit in optimizing the channel quantizer is as follows. In optimal quantizer design with typical metrics like MSE, the objective is to compute the quantizer boundaries and representations' points in each bin to optimize the metric of interest. In the system under consideration, the representation point within each bin is not utilized at the transmitter for adaptation. The power is adapted based on the quantizer boundaries (except at 0). Thus, regular quantizers are not expected to perform well in this context, and this intuition is strengthened by the numerical results in Section 6.

The analysis of average power and average packet loss is similar to that of the no TBIR case. An noted earlier, the main difference is that the β_j thresholds are chosen to optimize system performance rather than to minimize the MSE of γ_n.

The optimization problem is now stated as

$$
\min_{\{Y_{ij}, \beta_k\}} \Pi
$$
$$
\text{s.t. } \mathbb{E}[P_n] \le P_0 \tag{18}
$$

Recognize that (18) is a mixed optimization problem and solving it has potentially high complexity. However, for small values of N_b, N_f, and L, the problem is tractable and the main steps in the process are summarized as follows.

(1) Consider set $\mathbb{Y}$ as defined in Section 3. Set counter $m = 1$. Consider the mth element of $\mathbb{Y}$.

(2) For that particular combination of Y_{ij}, compute $\{\beta_i\}$, $\hat{\beta}$ that minimizes (18) (due to the closed-loop nature of the system, we have been unable to find analytical solutions to (18)). This conditional optimization over β_i is easily solved using numerical solvers. Unlike in Section 3, in this case the flexibility in the choice of $\{\beta_i\}$ allows us to increase β_i and $\hat{\beta}$ as high as necessary to satisfy the power constraint.

(3) Compute the total packet loss for the chosen Y_{ij} and the $\{\beta_i\}$, $\hat{\beta}$ parameters computed in step (2).

(4) Set $m = m + 1$. If $m > |\mathbb{Y}|$, then go to step (5); else go to step (2).

(5) Find the minimum value of the total packet loss and the corresponding Y_{ij}.

It should be noted that a similar optimization problem is considered in [25]. The main difference between the analysis in this section and that in [25] is the choice of the heuristic functions f and g. The analysis in [25] is restrictive in that packet losses do not occur in the channel. The results shown in this section generalize and improve the results in [25]. Numerical results of the total packet loss using such statistical TBIR are given in Section 6.

5. Partial CSIT: Instantaneous TBIR

In this section, we consider a communication system as depicted in Figure 1, where the receiver has partial instantaneous knowledge of the transmit buffer conditions. We consider that the receiver has N_f bits of information about the number of packets in the transmit buffer during each time slot. These N_f bits are used to adapt the FB information that is sent to the transmitter in each time slot.

An algorithm depicting the entire process in the system is given in Figure 4. The actions to be taken at the transmitter are represented within the square blocks, while the actions to be taken at the receiver are represented within circles. As discussed in Section 6, the CSI can be calculated at the receiver in two different ways, and hence there is a link indicated in Figure 4 between FF transmission block and CSI computational block. A temporal representation of the entire process is also given in Figure 4.

The gains due to this adaptation can be qualitatively explained as follows. When there are very few packets in the transmit buffer, the probability of buffer overflow is small. Thus, one can delay the packets and wait for good channel conditions to transmit. Consequently, the thresholds γ_{ki} for transmitting i packets are set to high values. On the other hand, when the buffer is nearly full, the probability of buffer overflowing is high. Hence, the thresholds γ_{mi} for transmitting i packets are set to small values, and one may not be able to wait for "good" channel conditions to transmit the packets. In other words, one should take a chance on the channel only when the buffer conditions are "desperate." The numerical values of the optimal thresholds, given in Section 6, confirm this behavior.

The analysis of average packet loss and average power proceeds along similar lines to the earlier cases. The main difference now is that there are multiple sets of β_j coefficients; one set of $\{\beta_j\}$ coefficients is used for each value of $\hat{q}$. These coefficients are represented as $\beta_j(i)$, $i = 1, \ldots, 2^{N_f}$, where N_f is the number of FF bits. The FF information which is generated from the buffer length q_n using the function $\hat{q}_n = e(q_n)$ is assumed to take on values $1, \ldots, 2^{N_f}$. An example of the different functions, $\{f, g, w\}$, is shown in Figure 3. In these figures, the value of $e(q_n)$ is represented in binary digits.

As in the earlier case, the γ_{ij} coefficients can be calculated from the $\{\beta_k(j), Y_{il}\}$ parameters as

$$
\gamma_{i,j} = \beta_k(e(i)), \quad \text{where } k = \min_{Y_{il} = j} l. \tag{19}
$$

As before, if $k = \phi$ for a given (i, j), then $\gamma_{i,m} = \infty$ for all $m \ge j$. The transition probabilities p_{ji} and stationary probabilities s_i are computed using (8) and (9) with the new values of thresholds $\{\beta_k(j), Y_{il}\}$. As in the case of statistical TBIR, it is assumed that packet loss in the channel occurs only in buffer state L when channel gain $\gamma_n < \beta_1(e(L))$. Consequently, the total loss is given by

$$
\begin{aligned}
\Pi_{\text{inst}} = & \; Y_{L1} s_L \left(1 - e^{-\hat{\beta}(e(L))} \right) \\
& + \sum_{m=0}^{L} \sum_{l=1}^{2^{N_b}} s_m \left(e^{-\beta_{l-1}(e(m))} - e^{-\beta_l(e(m))} \right) \\
& \times \left(\sum_{x=L+Y_{ml}-m+1}^{M} (x - L + m - l) c_x \right),
\end{aligned} \tag{20}
$$

where $\hat{\beta}(e(L))$ is used to select the transmit power when $q_n = L$ and $\gamma_n < \beta_1(e(L))$ as $(e^{Y_{L1}R} - 1)/\hat{\beta}(e(L))$. The average transmission power can now be derived as

$$
\begin{aligned}
\overline{P}_{\text{inst}} = & \; \mathbb{E}[P_n] \\
= & \sum_{k=1}^{L} \sum_{l=2}^{2^{N_b}} s_k \left(e^{-\beta_{l-1}(e(k))} - e^{-\beta_l(e(k))} \right) \frac{e^{Y_{kl}R} - 1}{\beta_{l-1}(e(k))} \\
& + s_L (1 - e^{-\beta_1}) \frac{e^{Y_{L1}R} - 1}{\hat{\beta}(e(L))}.
\end{aligned} \tag{21}
$$

The optimization problem is now posed in a manner similar

Example of function $g(\gamma_n, \hat{q}_n)$

Channel gain
with quantization
thresholds shown

$\hat{\gamma} = 00$ ┆ $\hat{\gamma} = 01$ ┆ $\hat{\gamma} = 10$ ┆ $\hat{\gamma} = 11$
─── $\hat{q}_n = 2$
$\beta_1(2)$ $\beta_2(2)$ ... $\beta_3(2)$

$\hat{\gamma} = 00$ ┆ $\hat{\gamma} = 01$ ┆ $\hat{\gamma} = 10$ ┆ $\hat{\gamma} = 11$
─── $\hat{q}_n = 1$
$\beta_1(1)$ $\beta_2(1)$ $\beta_3(1)$

$Y_{10} = 0$ ┆ $Y_{11} = 1$ ┆ $Y_{12} = 1$ ┆ $Y_{13} = 1$ $q_n = 1$
$Y_{20} = 0$ ┆ $Y_{21} = 1$ ┆ $Y_{22} = 1$ ┆ $Y_{23} = 2$ $q_n = 2$ } $\hat{q}_n = 1$
$Y_{30} = 0$ ┆ $Y_{31} = 1$ ┆ $Y_{32} = 2$ ┆ $Y_{33} = 2$ $q_n = 3$

0        $\beta_1(1)$ $\beta_2(1)$ $\beta_3(1)$ ∞

$Y_{40} = 0$ ┆ $Y_{41} = 1$ ┆ $Y_{42} = 2$ ┆ $Y_{43} = 3$ $q_n = 4$ } $\hat{q}_n = 2$
$Y_{50} = 1$ ┆ $Y_{51} = 2$ ┆ $Y_{52} = 3$ ┆ $Y_{53} = 4$ $q_n = 5$

0       $\beta_1(2)$ $\beta_2(2)$ $\beta_3(2)$ ∞

Example of function $f(q_n, \hat{\gamma}_n)$ Example of
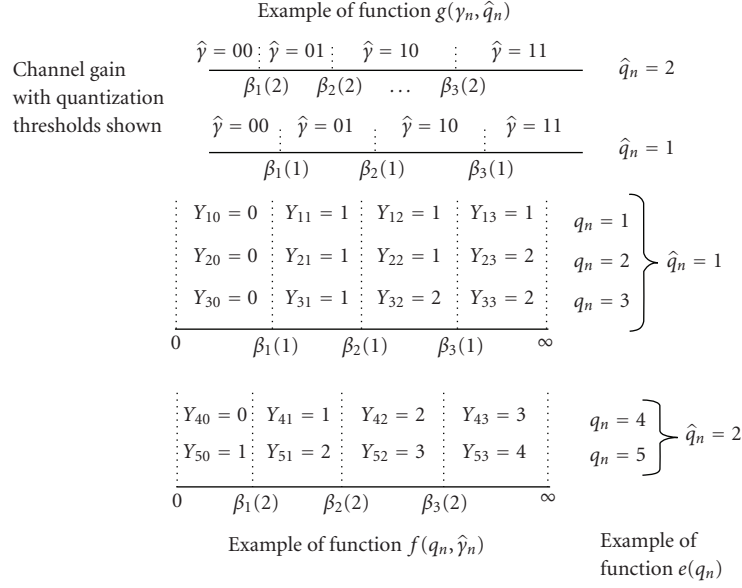   function $e(q_n)$

FIGURE 3: Examples of functions $e(q_n)$, $g(\gamma_n, \hat{q}_n)$, and $f(q_n, \hat{\gamma}_n)$ used when $N_f = 1$ bit of instantaneous TBIR is available. Number of feedback bits is $N_b = 2$ and buffer length is $L = 5$.
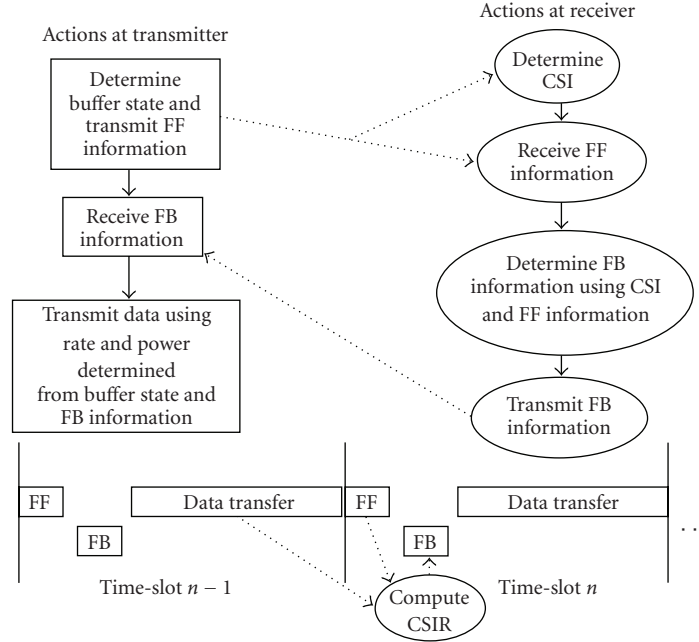

FIGURE 4: Summary of the main steps involved at the transmitter and receiver in implementing the proposed joint FF-FB architecture.

to the earlier cases as

$$\min_{\{Y_{ij}, \beta_k(l)\}} \Pi_{\text{inst}}$$
$$\text{s.t. } P_{\text{inst}} \leq P_0. \tag{22}$$

Recognize that (22) is a mixed optimization problem and solving it has potentially high complexity, like in the case of statistical TBIR. The procedure used to solve (22) is similar to that of the statistical TBIR case and is not repeated here.

Numerical values of the optimal thresholds along with the average packet loss are studied in the following section.

6. Numerical Results and Discussions

In this section, we numerically study the performance of the proposed adaptation strategies with no TBIR, statistical TBIR, and instantaneous TBIR. We also briefly discuss implementation issues and extensions of proposed concepts.
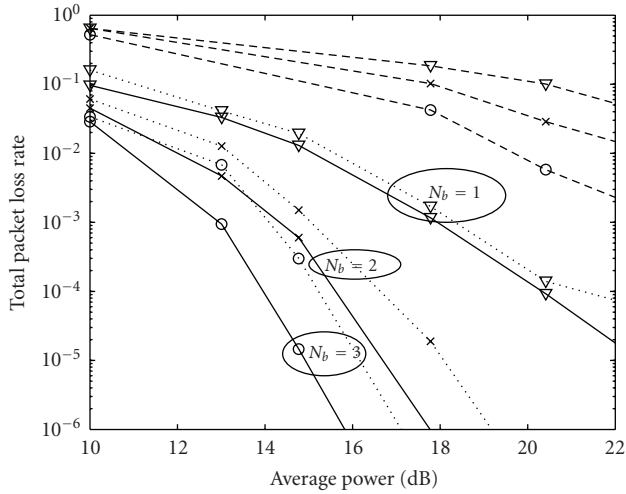
FIGURE 5: Variation of average packet loss with SNR for buffer length $L = 2$. The performance of the scheme with no TBIR (dashed lines), statistical TBIR (dotted lines), and one bit of instantaneous TBIR (solid lines) is shown.

— Instantaneous TBIR, $N_f = 1$ bit
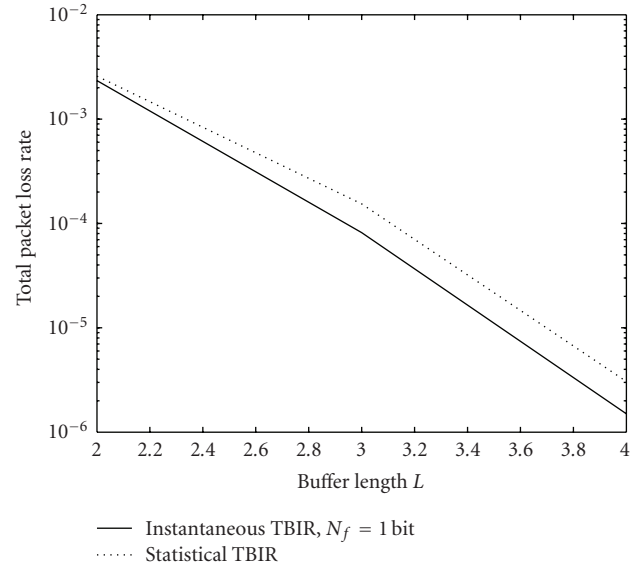········ Statistical TBIR

FIGURE 6: Variation of average packet loss with buffer length for statistical TBIR (dotted lines) and one bit of instantaneous TBIR (solid lines) is shown.

6.1. Numerical Results

Optimal Thresholds: Statistical and Instantaneous TBIR. The result of solving (18) and (22) for the same arrival traffic ($c_l = 0.5$, $l = 0, 1$), buffer length $L = 2$, and one bit feedback is given below. In both cases, the optimal thresholds $Y_{10} = 0$ and $Y_{ij} = 1$ for all $(i, j) \neq (1, 0)$. In the statistical TBIR case, $\beta_1 = 0.104$ and $\hat{\beta} = 2.9 \times 10^{-3}$. In the case of instantaneous TBIR with $N_f = 1$ bit, the optimal functions $e(q_n) = 0$, $q_n = 0, 1$, and $e(q_n) = 1$, $q_n = 2$. The corresponding thresholds $\beta_1(1) = 0.12$, $\beta_1(2) = 0.04$, and $\hat{\beta} = 1.5 \times 10^{-3}$. These optimal thresholds confirm the qualitative behavior explained in Section 5.

Packet Loss Versus SNR. The plot of the average packet loss versus SNR is given in Figure 5 for the three cases of no TBIR, statistical TBIR, and one bit of instantaneous TBIR. Results for three different feedback channel capacities of $N_b = 1$, 2, and 3 bits are shown. In Figure 5, a buffer of $L = 2$ length packets is used to store packets generated by an *on-off* source with arrival distribution of $c_l = 0.5$, $l = 0, 1$. The performance gains of using statistical TBIR over no TBIR are huge; for example, the power saving is about 9 dB to achieve packet error rate of 1% using $N_b = 3$ bits. Thus, showing the importance of adapting the channel quantizer at the receiver is based on statistical buffer conditions.

The performance of instantaneous TBIR shows power saving of about 1 dB over statistical TBIR for $N_b = 2, 3$. For $N_b = 1$, the instantaneous TBIR only shows marginal reduction in packet loss rate. The results thus suggest that even 1 bit of FF can be extremely useful in improving overall system performance. Alternately, at a given power constraint, the packet error rate reduces substantially with just 1 bit of FF; for example, at an SNR of 15 dB, the packet error rate is reduced by nearly an order of magnitude for $N_b = 3$ bits.

Further, the packet loss versus SNR curve for ($N_b = 3$, $N_f = 0$) intersects the curve for ($N_b = 2$, $N_f = 1$) at multiple points; this indicates that sometimes increasing N_f by one bit reduces packet loss rates more than increasing N_b by one bit. However, it should be mentioned that the goal here is to improve the system performance using the given FB bits, by adding FF bits. Moreover, we conjecture that for highly bursty sources (source having large variations in packets' arrivals), the gain of 1 bit of FF would be higher than using an additional bit of FB. The question of whether adding an extra bit of FB is better than adding a bit of FF is challenging; the answer depends critically on the traffic arrivals and channel statistics and should be investigated carefully in future work.

Packet Loss Versus Buffer Length L. The variation of the total average packet loss with buffer length L is given in Figure 6. It is clear that using 1 bit of FF can significantly reduce the average packet loss for the same number of FB bits. Note that for a delay of 1 time slot, the use of FF information does not reduce packet losses since packets cannot be delayed and the transmission rate cannot be adapted to channel conditions. This case is loosely analogous to the use of CSIT in discrete memoryless channels, in which CSIT does not increase capacity but could provide simpler methods to achieve capacity. However, for delays greater than 1 time slot, even though the source is a discrete memoryless source, the use of a buffer and greater flexibility in allowed delay introduces "memory" into the buffer state; thus, FF information provides performance gains (lower packet loss).

Packet Loss Versus Number of Feedback Bits N_b. The variation of average packet loss with the number of feedback bits N_b is given in Figure 7 for both the statistical and instantaneous TBIR cases. In this case, the same *on-off* traffic as in the

TABLE 1: The power required to achieve a desired average packet error using a convolutional code with variable QAM.

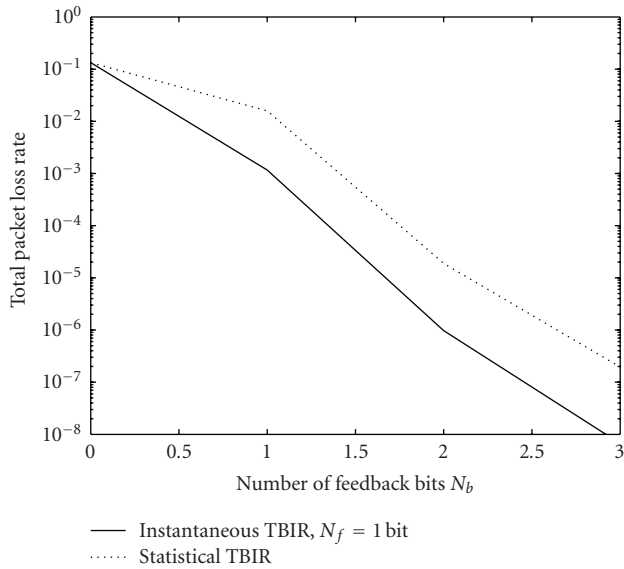| No. of packets/time slot	No. of bits/modulation symbol	Average power (dB)
1	2	5.5
2	4	12
3	6	22.5

FIGURE 7: Variation of average packet loss with number of feedback bits is shown for statistical TBIR (dotted lines) and one bit of instantaneous TBIR (solid lines).

earlier case, with one packet arrival on average, in every other time slot is considered. The loss rate with zero bits of CSIT is computed as follows. With no CSIT, power is transmitted at a constant rate (only depending on source arrivals). For this particular traffic, the transmission power is given by $P_{\mathrm{no\,CSIT}} = (e^R - 1)/\gamma_{\mathrm{no\,CSIT}}$. From the given power constraint of P_0, we can compute $\gamma_{\mathrm{no\,CSIT}}$ as $\gamma_{\mathrm{no\,CSIT}} = (e^R - 1)/2P_0$, where the factor of 2 comes from the fact that a packet is transmitted in only 50% of the time slots. The packet loss rate then equals $\int_0^{\gamma_{\mathrm{no\,CSIT}}} e^{-x}\, dx = 1 - e^{-\gamma_{\mathrm{no\,CSIT}}}$. It can be seen from Figure 7 that even a few bits of FB and one bit of FF can provide significant gains in performance.

6.2. Implementation Strategies. In this paper, we have assumed that the transfer of FB and FF information takes place at the beginning of a time slot before data communication in that time slot. It is also assumed that channel state information is available right at the beginning of the time slot. There are potentially many ways to implement these strategies; a couple of strategies are illustrated below.

(i) It is conceivable that CSI is computed at the receiver based on the reception of the FF information. Given that the FF information is likely to be only a few bits, accurate CSIR may be difficult to obtain. However, if pilot or synchronization bits are sent along with the

FF information, accurate CSI can be obtained from these bits.

(ii) In this paper, we assumed that the fading states in two different time slots are independent of each other. However, many practical communication systems exhibit considerable correlation in the fading process. This correlation can be used to obtain estimates of the CSI from prior time slots.

These strategies are pictorially depicted in Figure 4.

6.3. Practical Multirate System. Thus, the analysis so far in the paper is based on the information-theoretic concept of outage and transmission at rates close to Shannon capacity using finite block-length codes. We now demonstrate the application of FF information using a practical coding and modulation scheme. A similar coding and modulation scheme is used in [18] for multirate transmission over an AWGN channel.

We assume the size of each packet to be 25 bits, and the channel bandwidth and transmit pulse shape are such that 25 symbols can be transmitted in each time slot. The transmitter can choose to transmit 0, 1, 2, or 3 packets in each time slot. The data bits of all the packets in a time slot are jointly encoded, using a convolutional encoder of rate 1/2 with constraint length of 3 and generator matrix [4 7] [28]. The output of the convolutional encoder is modulated using a variable rate QAM depending on u_n according to Table 1. For example, to transmit 2 packets per time slot, the scheme needs to transmit 100 coded bits (2 packets $\times$25 bits/packet $\times$2 coded bits/information bit) using 25 symbols, which implies 4 bits/symbol; hence we choose a simple rectangular 16-QAM constellation in this case. The power required to achieve a packet error rate of 0.02 is given in Table 1 assuming instantaneous $\gamma_n = 1$ (an alternative is to change the coding rate assuming that the modulation (number of constellation points) is fixed, say, 4-QAM; to transmit 1, 2, or 3 packets per time slot, coding rates of 1/6, 1/3, or 1/2 could be used, resp.) For other values of γ_n, the power in Table 1 should be scaled by γ_n.

The main difference from the earlier theoretical formulation is that in this case the total packet loss rate is calculated by setting $\Gamma(m, \beta_i)$ to a desired nonzero frame error rate. In Table 1, the packet error rate Π_{m,β_i} is set at 2% for all m, β_i. The performance of the proposed scheme with $(N_b = 2, N_f = 0)$, $(N_b = 2, N_f = 1)$, and $(N_b = 3, N_f = 0)$ is shown in Figure 8. It can be seen that just 1 bit of FF results in approximately 1 dB saving in power, just like in the analysis based on information-theoretic outage probabilities. Further, at high SNR, the addition of 1 bit of FF to a scheme with $N_b = 2$ bits performs nearly as well as
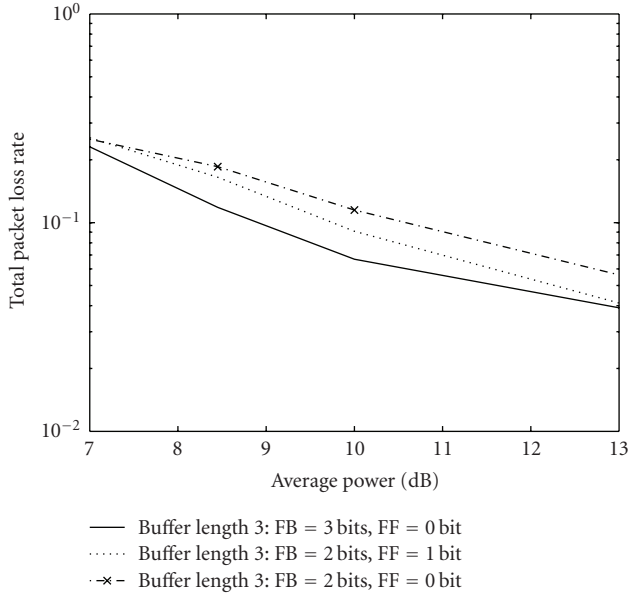
FIGURE 8: Variation of average packet loss with SNR for statistical TBIR and instantaneous TBIR with buffer length $L = 3$.

the scheme with one additional bit of FB, that is, $N_b = 3$ bits. Recognize that the packet loss hits an asymptote around 4% due to the design choice. Smaller values of the asymptote may be achieved by appropriately choosing a scheme with lower transmission packet error rate, that is, $\Pi_{m,\beta_i} \leq 0.02$. However, such schemes require a higher transmit power and could result in performance degradation at small to medium SNRs. The optimal choice of operational frame error rate should be computed based on the SNR of interest. The results in this section are presented to merely indicate the feasibility of using the proposed scheme in a practical multirate system.

6.4. MIMO Systems. The proposed formulation directly extends to systems with MIMO architecture. For instance, in the special case of an MISO or SIMO system, the capacity [3] is given by $0.5 \log(1 + (P/N_t)H^t H/\sigma^2)$, where N_t, N_r are the numbers of transmit and receive antennas and H is an $N_r \times N_t$ matrix of channel gains. The effective channel gain is thus a scalar given by $H^t H$ in the SIMO case (or HH^t in the MISO case). The proposed formulation can be applied to this effective channel gain which has a Chi-squared distribution if the channels between the transmit and receive antennas are modeled as independent Gaussian. In the general MIMO case, the *effective* channel gain would not be a scalar and one needs to optimize the feedback and feed-forward coefficients using a vector quantizer.

7. Conclusions

In this paper, we proposed a new communication system architecture in which information about transmit buffer is sent as feed-forward information to the receiver. This FF information is used in conjunction with the FB of channel state information to reduce average packet loss in

fading channels. Moreover, the proposed design framework provides a mechanism to provide delay guarantees to the arrival traffic.

We are currently working on studying the effects of FF for more realistic traffic arrival models. One limitation of the proposed scheme is the computational complexity of solving the optimization problems. It is likely that in practical systems the number of feedback bits will be limited to a small number (1–4 bits per coherence interval), and thus complexity is determined mainly by buffer length L. Conceptual extensions to frequency-selective fading channel are relatively straightforward using an orthogonal transmission scheme like OFDM; however, the complexity of the vector quantizer optimization that results needs to be investigated in future work. The impact of causal CSI knowledge along with errors in the FB and FF channels should also be considered carefully in future studies.

References

[1] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, NY, USA, 1991.

[2] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Transactions on Information Theory*, vol. 43, no. 6, pp. 1986–1992, 1997.

[3] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[4] M. Medard and A. Goldsmith, "Capacity of time-varying channels with channel side information at the sender and receiver," in *Proceedings of the International Conference on Communications (ICC '99)*, pp. 20–16, Vancouver, Canada, June 1999.

[5] A. Lapidoth and S. Shamai, "Fading channels: how perfect need "perfect side information" be?" *IEEE Transactions on Information Theory*, vol. 48, no. 5, pp. 1118–1134, 2002.

[6] M. Médard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Transactions on Information Theory*, vol. 46, no. 3, pp. 933–946, 2000.

[7] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: information-theoretic and communications aspects," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2619–2692, 1998.

[8] R. Gallager, "A perspective on multiaccess channels," *IEEE Transactions on Information Theory*, vol. 31, no. 2, pp. 124–142, 1985.

[9] A. Ephremides and B. Hajek, "Information theory and communication networks: an unconsummated union," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2416–2434, 1998.

[10] A. Ephremides, "Some wireless networking problems with a theoretical conscience," in *Codes, Curves, and Signals: Common Threads in Communications*, pp. 201–212, Kluwer Academic publishers, Dordrecht, The Netherlands, 1998.

[11] T. ElBatt and A. Ephremides, "Joint scheduling and power control for wireless ad-hoc networks," in *Proceedings of the 21st Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '02)*, vol. 2, pp. 976–984, New York, NY, USA, June 2002.

[12] E. Yeh, "Delay optimal rate allocation in multiaccess communications: a cross layer view," in *Proceedings of the IEEE*

Workshop on Multimedia Signal Processing (MMSP '02), pp. 404–407, St. Thomas, Virgin Islands, USA, December 2002.

[13] L. Tong, V. Naware, and P. Venkitasubramanian, "Signal processing in random access: a cross-layer perspective," *IEEE Signal Processing Magazine*, vol. 21, no. 5, pp. 29–39, 2004.

[14] R. A. Berry and E. M. Yeh, "Cross-layer wireless resource allocation," *IEEE Signal Processing Magazine*, vol. 21, no. 5, pp. 59–68, 2004.

[15] S. Shakkottai, T. S. Rappaport, and P. C. Karlsson, "Cross-layer design for wireless networks," *IEEE Communications Magazine*, vol. 41, no. 10, pp. 74–80, 2003.

[16] E. Shih, S.-H. Cho, N. Ickes, et al., "Physical layer driven protocol and algorithm design for energy-efficient wireless sensor networks," in *Proceedings of the 7th Annual International Conference on Mobile Computing and Networking (MOBICOM '01)*, pp. 272–287, Rome, Italy, July 2001.

[17] R. A. Berry and R. G. Gallager, "Communication over fading channels with delay constraints," *IEEE Transactions on Information Theory*, vol. 48, no. 5, pp. 1135–1149, 2002.

[18] D. Rajan, A. Sabharwal, and B. Aazhang, "Delay-bounded packet scheduling of bursty traffic over wireless channels," *IEEE Transactions on Information Theory*, vol. 50, no. 1, pp. 125–144, 2004.

[19] B. E. Collins and R. L. Cruz, "Transmission policies for time varying channels with average delay constraints," in *Proceedings of the 37th Annual Allerton Conference on Communication, Control, and Computing*, pp. 709–717, Monticello, Ill, USA, September 1999.

[20] R. Negi, *Power adaptation strategies for delay constrained channels*, Ph.D. thesis, Stanford University, Stanford, Calif, USA, September 2000.

[21] D. P. Bertsekas and R. Gallager, *Data Networks*, Prentice-Hall, Upper Saddle River, NJ, USA, 1992.

[22] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Transactions on Vehicular Technology*, vol. 43, no. 2, pp. 359–378, 1994.

[23] D. Rajan, *Power efficient transmission policies for multimedia traffic over wireless channels*, Ph.D. thesis, Rice University, Houston, Tex, USA, April 2002.

[24] A. Sabharwal, D. Rajan, and B. Aazhang, "Dual problems in power control," in *Proceedings of the 39th Annual Allerton Conference on Communication, Control and Computing*, Monticello, Ill, USA, October 2001.

[25] D. Rajan, A. Sabharwal, and B. Aazhang, "Outage behavior with delay and CSIT," in *Proceedings of the IEEE International Conference on Communications (ICC '04)*, vol. 1, pp. 578–582, Paris, France, June 2004.

[26] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic publishers, Dordrecht, The Netherlands, 1992.

[27] C. E. Shannon, "A mathematical theory of communication," *Bell Systems Technical Journal*, vol. 27, pp. 379–423, 1948.

[28] J. G. Proakis, *Digital Communications*, Mc-Graw Hill, Boston, Mass, USA, 4th edition, 2001.