

Delay-limited Throughput Maximization for Fading Channels using Rate and Power Control

Nadeem Ahmed, Mohammad Ali Khojastapour, and Richard G. Baraniuk

Abstract—The fading channels seen in many wireless systems provide a particularly hostile environment for reliable communication. Current metrics for evaluating the performance limits of fading channels have shortcomings. *Ergodic capacity*, representing the ultimate error-free communications limit, only applies to systems with infinite coding delay. Practical systems are delay-limited and must use finite-length codes. For delay-limited systems ϵ -capacity and *delay-limited capacity* are typically used to quantify the communications performance. However, ϵ -capacity is not an estimate of error-free performance while delay-limited capacity tends to be an overly conservative measure. We model practical systems as a single server queue and quantify the communications performance as the average throughput through the queue. Throughput is maximized by optimally selecting the transmission rate and power control strategy. Using this approach we arrive at striking conclusions. First, we show that a throughput very close to ergodic capacity can be achieved with a small coding delay. Second, the optimal transmission rate for some systems can be higher than the ergodic capacity of the channel. Third, we demonstrate the notion that power adaptation does not improve communication performance does not hold for delay-limited systems.

Index Terms—Throughput, capacity, block-fading, rate control, power control, delay, outage.

I. INTRODUCTION

THE communication channels seen in many wireless systems scatter the transmitted signal along its transmission path. The time variation of the channel results in random fluctuations of the received power level, or *fading*. The explosion of wireless communication services has yielded a great need to understand the limits of communication over fading channels.

Transmitters map data to *codewords*, which add redundancy and protect against the effects of the channel. Since fading channels are time-varying each codeword is affected by one or more fading states during transmission, with the specific number affecting the communications performance. The *coding delay*, the time required to encode codewords, is proportional to the codeword length and is quantified by the number of fading states affecting each codeword. The ultimate reliable communications limit over a fading channel is given by its *ergodic capacity* [7]; it is an asymptotic quantity, only achievable with infinite coding delay.

Practical communication systems are *delay-limited* and must use finite-length codewords that are affected by finitely many fading states during transmission, making ergodic capacity unattainable. The need to quantify the communications performance of delay-limited systems gave rise to the concept of *outage* and the notions ϵ -capacity [10] and *delay-limited capacity* [9]. An outage event occurs when the attempted transmission rate is higher than the fading state of the channel

allows, resulting in a decoding error at the receiver. Then, ϵ -capacity is the highest rate that can be supported with outage probability no greater than ϵ . Delay-limited capacity is ϵ -capacity, when outages cannot be tolerated, with $\epsilon = 0$. Both these measures have their short-comings. ϵ -capacity is not an estimate of error-free communications. Delay-limited capacity is an overly conservative estimate of communications performance - it is zero in many common situations [9].

The average *throughput*, or effective data rate between the transmitter and receiver, is a more realistic measure of delay-limited system performance than either ergodic capacity or ϵ -capacity. For example, a delay-limited system transmitting at a rate of 100Kbps and losing 20% of codewords to outages has an 80Kbps throughput. Engineers are interested in maximizing the throughput, rather than the transmission rate, of communications systems. Delay-limited systems use finite-length codewords which can be transmitted multiple times to ensure reliable decoding at the receiver. Using this idea, in [1] the communication system is modelled with a queue and the throughput is related to the transmission rate and the codeword *service time*, the number of transmission attempts required per codeword. The maximum throughput represents the highest *average* reliable data rate that can be transmitted for a particular retransmission scheme.

In [1], systems with knowledge of the channel state only at the receiver are considered. For this system, the average transmit power remains constant, and the throughput is maximized by performing *rate control*. In this paper, we generalize the queueing interpretation to delay-limited systems that have knowledge of the channel state at both transmitter and receiver. The transmit power can be varied based on the condition of the channel and the throughput is maximized by performing both *rate and power control*. Our analysis leads to some striking conclusions. First, for some systems the throughput can approach ergodic capacity with very short codes. Second, the optimal transmission rate that maximizes throughput can be *higher* than ergodic capacity. This is counter to normal practice, where a transmission rate below ergodic capacity is used. Third, the conventional wisdom, that optimal power adaptation negligibly improves communications performance [7], does not hold for delay-limited systems.

This paper is organized as follows. Section II overviews relevant background information and notation. In Section III, we describe the maximum throughput under different power constraints. Examples of the performance of a system using throughput maximization are shown in Section IV. Finally, we conclude and provide directions for future work in Section V.

II. BACKGROUND AND NOTATION

In many important applications the condition of the fading channel changes on a time scale that is much slower than

Supported in part by NSF, AFOSR, ONR, DARPA, and the Texas Instruments Leadership University Program. Email: [nahmed,amir,richb]@rice.edu, Web: dsp.rice.edu.

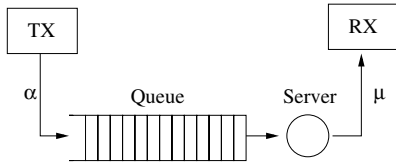


Fig. 1. Queueing representation of communications system. Codewords arrive at rate α and are served by the server at rate μ .

the communications signalling. This motivates modelling the communications channel as a discrete-time block-fading additive white Gaussian noise channel (BF-AWGN) [11]. In this model each “block” of N symbols corresponds to the amount of time the channel remains constant, the channel *coherence time*. The system in the k^{th} block can be written as

$$\underline{y}_k = \underline{x}_k h_k + \underline{w}_k, \quad (1)$$

with $\underline{y}_k, \underline{x}_k \in \mathbb{R}^N$ representing the system output and input. We assume a Gaussian noise process, $\underline{w}_k \sim \mathcal{N}(0, \mathbf{I}^N)$. Scattering by the environment results in reflections of the transmitted signal adding constructively or destructively with the original signal. The multipath interference due to scattering is represented by a random multiplicative gain, $h_k \in \mathbb{R}$, on the transmitted signal. In this work, we assume that $|h_k|^2$ follow a χ_2^2 (chi-squared with 2 degrees of freedom) distribution. This model is commonly used for wireless communication systems without line-of-sight (LOS) between transmitter and receiver.

Codewords span K blocks of the BF-AWGN channel resulting in KN symbols per codeword and a K -block *coding delay*. A system is considered *delay-limited* if $K < \infty$. Codewords contain KN symbols with information encoded at the *transmission rate*, R nats/sec/Hz. The time-variations in the channel are assumed to be i.i.d. from block-to-block. The blocks can physically correspond to slots in time, frequency, or both. The K i.i.d. channel fades affecting each codeword are

$$\underline{h} := [|h_0|^2, |h_1|^2, \dots, |h_{K-1}|^2]. \quad (2)$$

We assume that both transmitter and receiver have perfect (delayless and error free) channel state information (CSI) and know the realization of the K i.i.d. channel states (fades) affecting the current codeword. This model is well-known in literature [5], [8] and applies to wireless multicarrier modulated systems with K parallel subchannels. With CSI, the transmitter can vary the transmit power based on the condition of the channel. Let $\underline{\gamma}$ represent a power control strategy which defines the power allocated to a codeword for any channel realization. Then for the channel, \underline{h} ,

$$\underline{\gamma}(\underline{h}) := [\gamma_0, \gamma_2, \dots, \gamma_{K-1}] \quad (3)$$

represents the power allocation vector for the K blocks the codeword. When performing power control, the transmitter must not violate the specified power constraint. We consider short-term and long-term *average* power constraints. The short-term power constraint

$$\langle \underline{\gamma}(\underline{h}) \rangle := \frac{1}{K} \sum_{k=0}^{K-1} \gamma_k \leq \mathcal{P}_{\text{av}} \quad (4)$$

is a limit on the amount of power allocated within a codeword. In each block of the codeword the power can exceed \mathcal{P}_{av} , while the average over each codeword must be less than \mathcal{P}_{av} . This is equivalent to a total codeword power constraint of $K\mathcal{P}_{\text{av}}$. The long-term constraint

$$E_{\underline{h}} [\langle \underline{\gamma}(\underline{h}) \rangle] \leq \mathcal{P}_{\text{av}} \quad (5)$$

is a more relaxed condition. It allows the average power in any codeword to exceed \mathcal{P}_{av} so long as the average power across all codewords does not.

The *instantaneous capacity* (spectral efficiency actually, but we use the terms interchangeably) is the highest reliable data rate for the K channel fades, \underline{h} , affecting a codeword. It is found by maximizing the mutual information over the K channel fades and is given by

$$C_K(\underline{h}, \underline{\gamma}(\underline{h})) := \frac{1}{K} \sum_{k=0}^{K-1} \log(1 + |h_k|^2 \gamma_k). \quad (6)$$

With finite coding delay, $K < \infty$, an *outage* event occurs when the attempted transmission rate is higher than the instantaneous capacity, $R > C_K(\underline{h}, \underline{\gamma}(\underline{h}))$. The *outage probability* is

$$\begin{aligned} P_{\text{out}}(R, \underline{\gamma}, K) &:= \text{Prob}[R > C_K(\underline{h}, \underline{\gamma}(\underline{h}))] \\ &= E_{\underline{h}}[\mathcal{I}_F(R > C_K(\underline{h}, \underline{\gamma}(\underline{h})))] \end{aligned} \quad (7)$$

where $\mathcal{I}_F(\cdot)$ is the indicator function, which is 1 if the condition inside is true, and 0 otherwise. Then, ϵ -capacity

$$C_\epsilon := \sup_R \sup_{\underline{\gamma}} \{R : P_{\text{out}}(R, \underline{\gamma}(\underline{h}), K) \leq \epsilon, \nu(\underline{\gamma}) \leq \mathcal{P}_{\text{av}}\}, \quad (8)$$

represents the highest rate that can be supported with outage probability less than ϵ . The supremum is taken over all power allocation policies, $\underline{\gamma}$, and $\nu(\underline{\gamma})$ represents either the short-term (4) or long-term (5) power constraint.

When $K = \infty$ it is possible to transmit reliably at any rate less than ergodic capacity, the highest sustained rate over all channel states with arbitrarily small error probability. With CSI at the transmitter and receiver it is given by

$$C_{\text{erg-pc}} := \sup_{\underline{\gamma}} E_{|h|^2} [\log(1 + |h|^2 \underline{\gamma})], \quad (9)$$

where the capacity-achieving power allocation is $\gamma^*(|h|^2) = \left[\frac{1}{\lambda^*} - \frac{1}{|h|^2} \right]_+$ and λ^* is chosen such that $\int_{\lambda^*}^{\infty} \left(\frac{1}{\lambda^*} - \frac{1}{|h|^2} \right) dF(|h|^2) = \mathcal{P}_{\text{av}}$, with $F(|h|^2)$ the cumulative distribution function (CDF) of the channel gains. To achieve $C_{\text{erg-pc}}$, codewords are drawn from an infinite-length codebook with i.i.d. symbols $\sim \mathcal{N}(0, 1)$. Prior to transmission, the N symbols in each block are scaled by $\sqrt{\gamma^*(|h|^2)}$ [5]. When only the receiver has CSI, the transmitter always transmits with $\gamma = \mathcal{P}_{\text{av}}$ and the ergodic capacity is given by [7]

$$C_{\text{erg-const}} := E_{|h|^2} [\log(1 + |h|^2 \mathcal{P}_{\text{av}})]. \quad (10)$$

It is well-known that capacity with power adaptation, $C_{\text{erg-pc}}$, “yields a negligible capacity gain” over capacity with constant power, $C_{\text{erg-const}}$. This is especially apparent for large \mathcal{P}_{av} [7].

III. MAXIMUM ZERO-OUTAGE THROUGHPUT

The conventional measures of communication performance described in the previous section all attempt to quantify the amount of information that can be reliably transmitted if each codeword is transmitted *only once*. However, since delay-limited systems use finite-length codewords they can transmit each codeword multiple times to ensure reliable delivery. Communication systems should exploit this to maximize communications performance. Measures of communication performance should also factor multiple transmission attempts into the analysis as it more accurately models the performance of practical delay-limited systems in use today.

If the transmitter has CSI, then it knows prior to transmission if the K current channel states are sufficient to support reliable communications. If an outage is inevitable, transmission can be delayed until a more favorable channel arises. We will use this *delayed transmission scheme* and quantify the best case communications performance of such a system. The communication system can be thought of as a single server queue (see Figure 1) with the server ensuring reliable transmission by delaying transmission of codewords until the channel condition allows it. The codeword *service time*, S , represents the number of transmission attempts and is a random quantity due to the random nature of the fading channel. Since we know that outage events are independent of one another, the probability that a codeword requires s attempts for successful transmission is

$$\text{Prob}(S = s) = [P_{\text{out}}(R, \underline{\gamma}, K)]^{s-1} [1 - P_{\text{out}}(R, \underline{\gamma}, K)], \quad (11)$$

which implies that the service time distribution is geometric on the positive integers, with parameter $1 - P_{\text{out}}(R, \underline{\gamma}, K)$. The mean is then $E[S] = \frac{1}{1 - P_{\text{out}}(R, \underline{\gamma}, K)}$.

The average amount of data passing through the queue with each transmission attempt is $\frac{R}{E[S]}$ nats/sec/Hz, corresponding to the amount of data in each codeword divided by the number of attempts required to send it. We define the *maximum zero-outage throughput* (MZT) as

$$\text{MZT} := \sup_R \sup_{\underline{\gamma}} \{R[1 - P_{\text{out}}(R, \underline{\gamma}, K)] : \nu(\underline{\gamma}) \leq \mathcal{P}_{\text{av}}\} \quad (12)$$

where the supremum is taken over all power allocations and transmission rates subject to the power constraint.

Maximizing the throughput amounts to maximizing the amount of data passing through the queue and requires optimizing the rate that codeword are encoded. Intuitively, as $R \rightarrow 0$, throughput approaches 0. Similarly, as $R \rightarrow \infty$ frequent outages result in a throughput that also approaches 0.

Multiple transmission attempts per codeword allows for zero-outage (error-free) measures of communication performance for any coding delay. This is not possible if only a single transmission attempt is allowed per codeword. Delay-limited capacity, C_ϵ for $\epsilon = 0$, is 0 in many common situations [9]. However, the improved communications throughput achieved with multiple transmission attempts does have a cost associated with it. The random nature of the service time for any codeword can cause queues to build and queueing delays that must be considered. Using concepts from queueing theory it is possible to quantify this delay. For example,

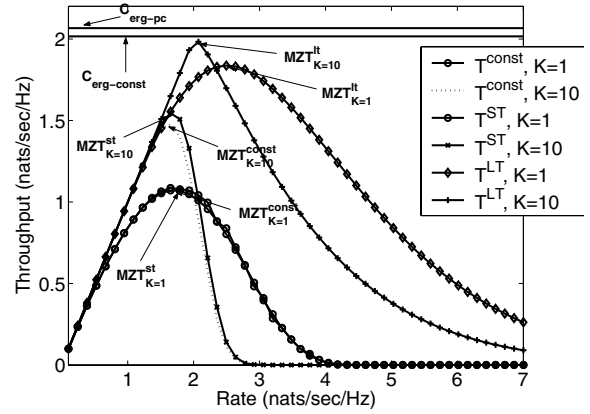


Fig. 2. Throughput vs. rate for constant, short-term and long-term power control policies and power constraint $\nu(\underline{\gamma}) = 10\text{dB}$. For each policy and for each value of K , the MZT is the maximum of each curve. Ergodic capacity with and without power control are included for reference.

if the codewords follow a Poisson arrival process, then the communication system can be modelled with an M/G/1 queue. The delay for each codeword, queueing delay plus the service time, can be determined by the Pollaczek-Khinchin (P-K) theorem [4],

$$D = K \left(\frac{\beta E[S^2]}{2(1 - \beta E[S])} + E[S] \right), \quad (13)$$

where β is the average arrival rate, and $E[S]$ and $E[S^2]$ are the first and second moments of the service time. Scaling by K changes the units to blocks/codeword.

In the following sections, we solve (12) and find MZT under both the short-term and long-term power constraints.

A. MZT under the Short-Term Power Constraint

Under the short-term power constraint, the power in each block of any codeword can deviate from \mathcal{P}_{av} but the total power for the entire code must be less than $K\mathcal{P}_{\text{av}}$. With this power constraint MZT is given by the following theorem.

Theorem 1: The maximum zero-outage throughput for the short-term power constraint $\langle \underline{\gamma}(\underline{h}) \rangle \leq \mathcal{P}_{\text{av}}$ is

$$\begin{aligned} \text{MZT}^{\text{st}} = & \sup_R \{ R E_{\underline{h}} \{ \mathcal{I}_F(R < \log(\prod_{i=0}^{\mu_{\underline{h}}-1} |h_i|^2)) \\ & + \mu_{\underline{h}} \log(K\mathcal{P}_{\text{av}} + \sum_{i=0}^{\mu_{\underline{h}}-1} \frac{1}{|h_i|^2}) - \mu_{\underline{h}} \log \mu_{\underline{h}} \} \} \end{aligned} \quad (14)$$

where $\mu_{\underline{h}} \in \{1, 2, \dots, K\}$.

Proof sketch: MZTst is found by minimizing the outage probability for each transmission rate, and then taking the supremum over all rates. Outage probability minimization is equivalent to maximizing the instantaneous capacity [5]. For any rate we substitute the form of the power allocation strategy that maximizes the instantaneous capacity into (12) to arrive at (14). The complete proof can be found in [2]. ■

While (14) is generally difficult to solve analytically, a semi-explicit solution can be found when $K = 1$. If the channel gains follow a χ_2^2 distribution, then MZTst = $\sup_R [R e^{-\frac{eR}{\gamma}}]$. Since $K = 1$, the optimal power control

policy is $\gamma = \mathcal{P}_{\text{av}}$. In this case, the solution to the problem is the same as in [1]. The optimal transmission rate, $R^{\text{st}} = \mathcal{W}(\mathcal{P}_{\text{av}})$, where \mathcal{W} is Lambert's W function. Then,

$$\text{MZT}^{\text{st}} = R^{\text{st}} e^{-\left(\frac{e^{R^{\text{st}}}-1}{\mathcal{P}_{\text{av}}}\right)}. \quad (15)$$

B. MZT under the Long-Term Power Constraint

Under the long-term power constraint, the average power for any particular codeword can exceed \mathcal{P}_{av} , but the long-term average over all codewords cannot. It is less restrictive than the short-term power constraint, and the MZT is given by the following theorem.

Theorem 2: The maximum zero-outage throughput for the long-term power constraint $E_{\underline{h}}[\langle \gamma(\underline{h}) \rangle] \leq \mathcal{P}_{\text{av}}$ is

$$\text{MZT}^{\text{lt}} = \sup_R \{ R E_{\underline{h}} [\mathcal{I}_F(\Gamma(K, R, \underline{h}) < s_R^*) + w_0 \mathcal{I}_F(\Gamma(K, R, \underline{h}) = s_R^*)] \} \quad (16)$$

where $\Gamma(K, R, \underline{h}) = \sum_{i=0}^{\mu_{\underline{h}}-1} \left\{ \left[\frac{e^{KR}}{\prod_{j=0}^{\mu_{\underline{h}}-1} |h_j|^2} \right]^{\frac{1}{\mu_{\underline{h}}}} - \frac{1}{|h_i|^2} \right\}$, $\mu_{\underline{h}} \in \{1, 2, \dots, K\}$, $0 \leq w_0 \leq 1$, and $s_R^* \in \mathbb{R}$.

Proof sketch: MZT^{lt} is found by minimizing the outage probability [5] for any transmission rate and then taking the supremum over all rates. The outage minimizing power control policy, $\tilde{\gamma}$, minimizes the power required to maintain a target rate [5]. Transmission is delayed if the power allocated for a particular channel, $\langle \tilde{\gamma}(\underline{h}) \rangle$, exceeds a threshold, s_R^* . The threshold is chosen such that the long-term power constraint is satisfied. Substituting this solution into (12) results in (16). We denote the optimal transmission rate R^{lt} and the corresponding power cutoff point, $s_{R^{\text{lt}}}^*$. The entire proof can be found in [2]. ■

MZT^{lt} is difficult to express analytically, but when $K = 1$ the optimal solution satisfies several sufficient conditions.

Theorem 3: If the channel gains follow a χ_2^2 distribution, then for $K = 1$, $\text{MZT}^{\text{lt}} = R^{\text{lt}} e^{-\left(\frac{e^{R^{\text{lt}}}-1}{s_{R^{\text{lt}}}^*}\right)}$, where R^{lt} and $s_{R^{\text{lt}}}^*$ satisfy

$$e^{R^{\text{lt}}} E_i\left(1, \frac{e^{R^{\text{lt}}}-1}{s_{R^{\text{lt}}}^*}\right) = \mathcal{P}_{\text{av}}, \quad (17)$$

$$(s_{R^{\text{lt}}}^*)^2 - \mathcal{P}_{\text{av}} R^{\text{lt}} e^{R^{\text{lt}}} e^{\left(\frac{e^{R^{\text{lt}}}-1}{s_{R^{\text{lt}}}^*}\right)} = 0 \quad (18)$$

where $E_i(1, x) = \int_1^\infty \frac{e^{-xt}}{t} dt$.

Proof sketch: Condition (17), is a sufficient condition for the optimal s^* , and is obtained by finding the optimal short-term cutoff for a given transmission rate. Condition (18), is a sufficient condition for R^{lt} , and is obtained by finding the optimal rate for the outage minimizing power control policy. The details of the proof can be found in [2]. ■

IV. EXAMPLES AND DISCUSSION

Our simulations show that in order to maximize the communications throughput, the transmission rate needs to be selected very carefully. Figure 2 plots the throughput, for various coding delays, achieved as a function of the transmission rate for constant [1], short-term and long-term power allocation. MZT, the maximum throughput, for each power control strategy corresponds to the peak of each curve. Transmitting at a rate other

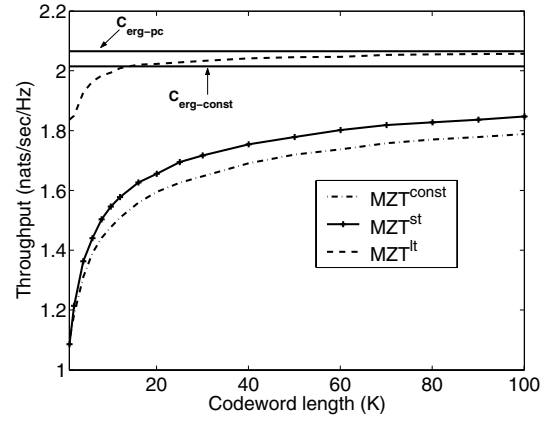


Fig. 3. Throughput vs. K for constant, short-term, and long-term power control policies and power constraint $\nu(\gamma) = 10\text{dB}$. In each case as K increases, the throughput approaches ergodic capacity, with the long-term policy converging most quickly.

than the one that corresponds to MZT can result in a large drop in throughput. For example if $K = 10$ and the transmission rate is optimally selected, then under the constant and short-term power allocation strategies, $\text{MZT}^{\text{st}} = 1.55$ nats/sec/Hz and $\text{MZT}^{\text{const}} = 1.48$ nats/sec/Hz, respectively. However, if we arbitrarily set the transmission rate to $R = 2$ nats/sec/Hz, then the throughput drops to $T^{\text{st}} = 1.18$ nats/sec/Hz and $T^{\text{const}} = 1.04$ nats/sec/Hz, respectively. Arbitrarily selecting the transmission rate can lead to large drops in throughput. The larger the K the more significant the loss in throughput is. This can be seen in Figure 2 by comparing the MZT for a particular power constraint and noticing that for larger K the curve is narrower. Therefore, the optimal rate should be carefully chosen by solving (14) and (16) for either of the short-term and long-term power constraints and by the procedure in [1] for constant power allocation.

The maximum throughput using codewords affected by only a few fading states can approach ergodic capacity. This is seen in Figure 3, which plots MZT vs. coding delay K for an average transmit power of $\mathcal{P}_{\text{av}} = 10$ dB. The figure shows MZT under the constant, short-term and long-term power allocation strategies. Under the long-term power constraint, $\text{MZT}^{\text{lt}} = 2.00$ nats/sec/Hz when $K = 10$. This is very close to ergodic capacity, $C_{\text{erg-pc}} = 2.07$ nats/sec/Hz that is achievable only when $K = \infty$. This is particularly interesting as it suggests that power control is more important than coding delay (ergodicity) for maximizing throughput. It is also worth noting that the more relaxed the power constraint, the higher the MZT, i.e. $\text{MZT}^{\text{const}} \leq \text{MZT}^{\text{st}} \leq \text{MZT}^{\text{lt}}$. This relation holds for any K , since the constant power allocation is a special case of the short-term power constraint, which in turn is a special case of the long-term power constraint.

The importance of optimally selecting the transmission rate is re-emphasized in Figure 4 which plots the optimal transmission rate, corresponding to MZT, as a function of coding delay, K , for an average transmit power of $\mathcal{P}_{\text{av}} = 10$ dB. The optimal transmission rate, especially for small K , can fluctuate a great deal. In fact a very non-intuitive phenomenon is observed - in some cases, the optimal transmission rate can actually be higher than ergodic capacity. For example when

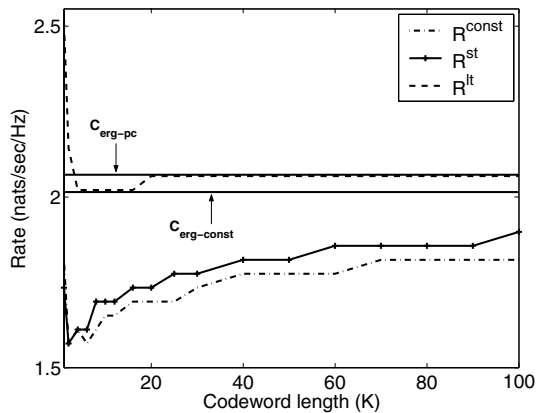


Fig. 4. Optimal transmission rate vs. K for constant, short-term and long-term power control policies and power constraint $\nu(\gamma) = 10\text{dB}$. In each case as K increases, the optimal transmission rate approaches ergodic capacity, with the long-term policy converging most quickly. It is interesting to note that the optimal transmission rate, for some K , is *above* ergodic capacity.

$K = 1$, the optimal transmission rate, under the long-term power constraint, should be $R^{\text{lt}} = 2.51$ nats/sec/Hz, while $C_{\text{erg-pc}} = 2.07$ nats/sec/Hz. Here the optimal transmission rate is more than 21% higher than the capacity of the channel. This is counter to common practice, where a transmission rate lower than capacity is normally used. We emphasize that this is not a violation of the ergodic capacity theorem [7], as the resulting *throughput is always less than ergodic capacity*.

Conventional wisdom tells us that when $K = \infty$, optimal power allocation yields “a negligible [ergodic] capacity gain” over a constant power strategy [7]. It is very interesting to note that this well-known property does not hold for delay-limited systems, when $K < \infty$. This can be seen in Figure 5 by comparing the maximum throughput with constant power, $\text{MZT}^{\text{const}}$, and under the long-term power constraint, MZT^{lt} . When $K = 10$ an average throughput of 1 nat/sec/Hz is achievable with SNRs of 6.4dB and 3.2dB, respectively. Here power control provides a 3.2dB gain over constant power transmission. If the throughput is increased to 1.5 nats/sec/Hz the power control gain remains constant at 3.2dB. This is very different than when $K = \infty$; the power control gain, seen by comparing $C_{\text{erg-pc}}$ and $C_{\text{erg-const}}$, shrinks from 0.85dB at 1 nat/sec/Hz to 0.47dB at 1.5 nats/sec/Hz.

To re-emphasize the importance of the power control we see in Figure 5 that for $K = 10$ the maximum achievable throughput with a long term power constraint, MZT^{lt} , is only about 0.5dB away from ergodic capacity of the channel, $C_{\text{erg-pc}}$, which is achieved with $K = \infty$. More surprisingly, for low SNR, MZT^{lt} is even greater than the ergodic capacity without power control, $C_{\text{erg-const}}$. We achieve a better throughput with $K = 10$ and power control, MZT^{lt} , than with $K = \infty$ and constant power, $C_{\text{erg-const}}$. This shows that power control is more important than large coding delays (ergodicity).

V. CONCLUSIONS

We presented a framework for analyzing the communications performance of practical delay-limited systems over fading channels. Communications engineers are interested in the throughput between the transmitter and receiver. We formulate the problem as a single server queue and equate

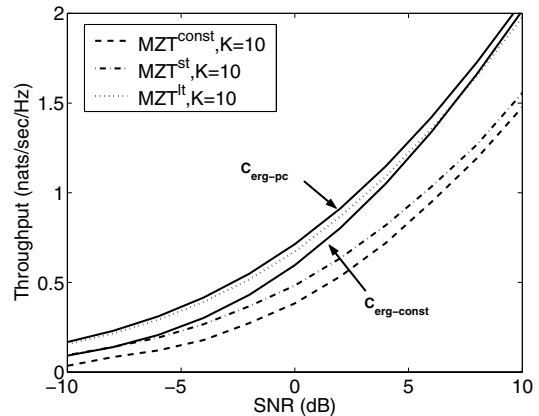


Fig. 5. Throughput vs. SNR for constant, short-term and long-term power control policies for $K = 10$. Ergodic capacity with and without power control are also shown.

the throughput of the system with the amount of information passing through the queue. We assume the transmitter has knowledge of the channel state and can adjust the transmit power level based on this knowledge. Using this model we maximize the throughput of the system optimally choosing the transmission rate and power allocation strategy.

Our analysis led to some striking realizations. First, we show that the throughput can approach ergodic capacity with very short coding delays using optimal rate and power allocation. Second, the optimal transmission rate, that maximizes throughput, for some systems can be higher than ergodic capacity. This is rather non-intuitive and counter to what is done in practice today, where a transmission rate below ergodic capacity is chosen. We note that the resulting throughput is always less than ergodic capacity. Third, we show that the conventional wisdom, for delay unconstrained systems, that optimal power adaptation negligibly improves communication performance does not hold for delay-limited systems.

REFERENCES

- [1] N. Ahmed and R. Baraniuk, “Throughput Measures for Delay-Constrained Communications in Fading Channels,” *41st Annual Allerton Conference on Communication, Control and Computing*, Oct. 2003
- [2] N. Ahmed, M.A. Khojastepour and R.G. Baraniuk, “Optimal Rate and Power Control for Throughput Maximization For Communication over Fading Channels” *In preparation*.
- [3] R. Gallager, “Information Theory and Reliable Communication,” *New York: Wiley*, 1968.
- [4] L. Kleinrock, “Queueing Systems - Volume I: Theory,” *New York: Wiley*, 1975.
- [5] G. Caire, G. Taricco and E. Biglieri, “Optimum Power Control Over Fading Channels,” *IEEE Trans. Info. Theory*, v. 45, no. 5, Jul. 1999.
- [6] R. Berry and R. Gallager, “Communication over Fading Channels with Delay Constraints,” *IEEE Trans. Info. Theory*, v. 48, no. 5, May 2002.
- [7] A. Goldsmith and P. Varaiya, “Capacity of Fading Channels with Channel Side Information,” *IEEE Trans. Info. Theory*, v. 43, no. 6, Nov. 1997.
- [8] E. Biglieri, J. Proakis and S. Shamai, “Fading Channels: Information-Theoretic and Communications Aspects,” *IEEE Trans. Comm.*, v. 44, no. 6, Oct. 1998.
- [9] S. Hanly and D. Tse, “Multiaccess Fading Channels - Part II: Delay-Limited Capacities,” *IEEE Trans. Info. Theory*, v. 44, no. 7, Nov. 1998.
- [10] S. Ozarow, S. Shamai and A. Wyner, “Information Theoretic Considerations for Cellular Mobile Radio,” *IEEE Trans. Vehic. Tech.*, v. 43, no. 2, May 1994.
- [11] R. J. McEliece and W. E. Stark, “Channels with Block Interference,” *IEEE Trans. Info. Theory*, v. 30, no. 1, Jan. 1984.
- [12] E. Teletar and R. G. Gallager, “Combining Queueing Theory with Information Theory for Multiaccess,” *IEEE Jour. Sel. Areas in Comm.*, v.13, no.6, Aug. 1995.