OPTIMAL POWER/DELAY TRADE-OFFS IN WIRELESS COMMUNICATIONS – SMALL DELAY ASYMPTOTICS

Randall A. Berry
Dept. of ECE, Northwestern Univ.
2145 Sheridan Rd.
Evanston, IL 60208 USA
email: rberry@ece.northwestern.edu

ABSTRACT
In this paper we consider the optimal trade-off between average transmission power and average queueing delay for a single user transmitting data over a wireless fading channel. In particular we study the behavior of this trade-off in the regime of asymptotically large power and small delay. Our focus is on channels which require infinite power to minimize the average delay. For such channels, it is shown that the average delay decreases no faster than $e^{-\alpha P}$ as the average power grows, where α is a parameter that depends on the fading distribution and the arrival rate. We also give an sequence of policies that nearly mathc this lower bound. Several other sub-optimal policies are also discussed.

KEY WORDS
wireless resource allocation, power/rate control, wireless scheduling

1 Introduction

For mobile wireless communication devices, energy efficiency is a key issue that has attracted much interest. A basic technique for improving the energy efficiency of a wireless device is through transmission power control. With data traffic, in addition to adjusting the transmission power used to send each packet of data, energy efficiency can be further improved by adjusting the transmission rate or equivalently the transmission time per packet. In this context, various transmission scheduling approaches have been studied including [1–7]. In these approaches transmission rate and power are adjusted over time based in part on the offered traffic as well as any channel state information. A theme underlying each of these approaches is managing the fundamental trade-off between packet delay and transmission power or energy. Specifically, packet delay can be reduced by transmitting at a higher rate, but this requires more energy per bit. In fading channels, reducing packet delay also prohibits users from optimally allocating their power over time in response to channel variations.

In this paper, we revisit the basic model for transmission scheduling over a fading channel studied in [2, 3]. In this model, data randomly arrives from some higher layer application and is placed into a transmission buffer as

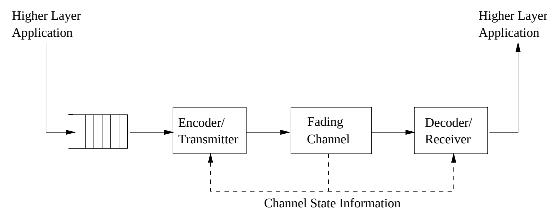

Figure 1. System Model.

shown in Figure 1. Data is periodically removed from the buffer, encoded and transmitted over the fading channel. We focus on the case where each codeword is sent over a fixed number of channel uses, but different codewords may be of different rates. After the codeword is received, it is decoded and sent on to the corresponding higher layer application at the receiver. The transmitter can vary the transmission power and rate based on both the channel state and the buffer occupancy. In [2], the optimal trade-off between the average delay incurred by the arriving data and the long-term average power was studied in a Markov decision framework. Furthermore, the behavior of this trade-off was characterized in the asymptotic regime of large delays (low power). In this regime, it was shown that the rate at which the required power decreases as the average delay, D, increases is[1] $\Theta\left(\frac{1}{D^2}\right)$. This rate can be achieved via a sequence of policies where the only dependence on the buffer occupancy is via a simple threshold rule. Moreover, some dependence on the buffer occupancy is required to achieve this optimal convergence rate.

In this paper, we focus on the behavior of the power/delay trade-off in the asymptotic regime of small delays (high power). In this regime we show that the optimal power delay trade-off behaves quite differently from large delay regime. Specifically, we show that as the average power increases to infinity, the average delay can decrease no faster than $\Omega(e^{-\alpha P})$, for channels whose fading distri-

[1]To characterize the asymptotic behavior of $g(x)$ as $x \to x^*$, we use the following notation $g(x) = O(f(x))$ if $\limsup_{x\to x^*} \frac{|g(x)|}{|f(x)|} < \infty$, $g(x) = \Omega(f(x))$ if $\limsup_{x\to x^*} \frac{|f(x)|}{|g(x)|} < \infty$ and $g(x) = \Theta(f(x))$ if $g(x) = O(f(x))$, and $g(x) = \Omega(f(x))$.

bution is non-zero at zero (such as a Rayleigh fading channel). Where α is a parameter that depends on the fading distribution and the arrival statistics. This bound holds for any transmission policy. We also shown that a sequence of simple "channel threshold" policies can nearly achieve this bound. For comparison, we also consider two other simple suboptimal policies which correspond to a fixed water-filling allocation and a fixed power policy, and show that these have a suboptimal convergence rate in the small delay regime. Note that the rate of $e^{-\alpha P}$ is much a much faster decrease in power than the $\frac{1}{D}$ decrease in delay in the large delay regime. This implies that the savings in power gained by relaxing the delay constraint are much more significant when the delay constraint is stringent. We also consider a class of channels whose distribution is zero at zero. For this case we show that with a constant arrival rate the average delay can decrease no faster than $\log\left(\frac{P^*(1)}{P}\right)$, where $P^*(1)$ is the minimum power required to transmit every arrival independent of the channel state. Before discussing these results we first give a more precise description of the basic model to be considered.

2 Problem Formulation

The fading channel is modeled a discrete time, block fading channel with additive Gaussian noise [8]. In such a channel the transmitted signal is multiplied by a time-varying gain which models the fading. Over each block of N consecutive channel uses, the gain stays fixed. Let $\sqrt{H_m}$ denote the magnitude of complex (base-band) channel gain during the mth block and Θ_m denote the phase. Let $\mathbf{X}_m = (X_{m,1}, \ldots, X_{m,N})$ and $\mathbf{Y}_m = (Y_{m,1}, \ldots, Y_{m,N})$ be vectors in $\mathbb{C}^N$ which denote, respectively, the channel inputs and outputs over the mth block. These are related by:

$$\mathbf{Y}_m = \sqrt{H_m} e^{-j\Theta_m} \mathbf{X}_m + \mathbf{Z}_m, \tag{1}$$

where the additive noise $\mathbf{Z}_m$ is a complex, circularly symmetric Gaussian random vector with zero mean and covariance matrix $\sigma^2 I$. Furthermore, the sequence $\{\mathbf{Z}_m\}$ is i.i.d. Here we assume that the sequence of channel gains, $\{H_n\}$, is a sequence of i.i.d. random variables taking values in a set $\mathcal{H} \subset \mathbb{R}^+$ with a probability density f_H over this set. We assume that both the transmitter and receiver have perfect CSI, *i.e.*, during the mth block, both the transmitter and receiver know the value of H_m and Θ_m. [2]

To model the buffer, we consider a discrete-time "fluid" buffer model in which the time between samples corresponds to a block of N channel uses. Let A_n be the number of bits that arrive between time n and $n-1$, and let S_n be the buffer size at the start of the nth block. Assume that at the start of each block U_n bits are removed from the buffer, encoded and transmitted over the fading channel during the next block. Thus the buffer dynamics are given by:[3]

$$S_{n+1} = \max\{S_n + A_{n+1} - U_n, A_{n+1}\}. \tag{2}$$

Here we assume that the buffer size is infinite and that no packets are lost. We also assume that the arrival process $\{A_n\}$ is a sequence of i.i.d. random variables taking values in a set $\mathcal{A} = [a_{min}, a_{max}]$ with probability distribution $F_A(a)$; here a_{min} and a_{max} are respectively upper and lower bounds on the arrival process. This process is assumed to be independent of the channel fading and noise processes. The expected arrival rate is denoted by $\bar{A}$.

Let $P(h,u)$ be the transmission power required during a block when the channel gain is h and the transmitter chooses to transmit u bits. In the following, we assume that

$$P(h,u) = \frac{\sigma^2}{h}(2^{u/N} - 1). \tag{3}$$

This is the minimum power required so that the mutual information rate over the N channel uses is equal to u/N. Provided that N is large enough, this choice will give a reasonable indication of the power needed to reliably transmit at rate u/N using optimal coding. Most of the following results can be generalized to other functions $P(h,u)$ which indicate the required power needed to transmit at rate u/N bits per channel use in different settings.

Let $\mathcal{S} = [0,\infty)$ denote the buffer state space, and let $\mu : \mathcal{S} \times \mathcal{H} \mapsto \mathbb{R}^+$ denote a transmission policy which indicates U_n at any time n as a function of S_n and H_n. Under a given policy, we denote the average delay by $\bar{D}^\mu$ and the long-term average transmission power by $\bar{P}^\mu$. For a given channel and arrival process, in [2], the optimal power/delay curve, $P^*(D)$ is defined as

$$P^*(D) = \inf\{\bar{P}^\mu : \mu \text{ such that } \bar{D}^\mu \le D\}.$$

This curve is decreasing and convex. Asymptotically, as $D \to \infty$, $P^*(D)$ is shown to converge to $\mathcal{P}(\bar{A})$ at a rate of $\Theta(\frac{1}{D^2})$ [2]. The asymptotic value, $\mathcal{P}(\bar{A})$ corresponds to the minimum power required to send at average rate $\bar{A}$, ignoring any delay constraints, i.e. the minimum power so that the channel has an *ergodic capacity* of $\bar{A}/N$ bits per channel use. An example power/delay trade-off curve is shown in Fig. 2.

In the following it will also be useful to define the optimal delay/power curve, $D^*(P)$ as

$$D^*(\bar{P}) = \inf\{\bar{D}^\mu : \mu \text{ such that } \bar{P}^\mu \le \bar{P}\}.$$

Clearly, if $P^*(D)$ is strictly decreasing, then $D^*(\bar{P})$ will simply be its inverse. Given the above model of the buffer dynamics, each packet must spend at least one time unit in the buffer, hence $D^*(\bar{P}) \ge 1$ for all P. The only way that $D^*(\bar{P}) = 1$ is if the transmitter used a policy such that

[2] Since both the transmitter and receiver know Θ_m, we can ignore it in the following.

[3] One reason for choosing this model of the buffer is that is provides an upper bound on the delay incurred by any packet in the underlying continuous time system.

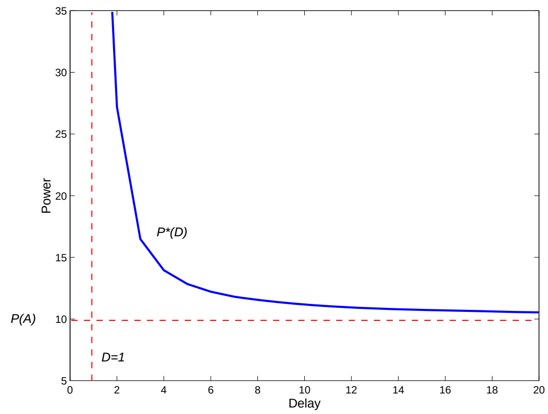

Figure 2. Example power/delay trade-off.

$\mu(S_n, H_n) \geq A_n$ for all n. The minimum power required by such a policy is given by

$$P^*(1) = \mathbb{E}_{A,H} P(H, A),$$

For many practical fading models, such as a Rayleigh fading channel, this quantity will be infinite. In the following, we initially restrict ourselves to such settings. Specifically, we make the following assumption:

Assumption A: Assume that the channel gain H_n, has a continuous density, f_H on $\mathbb{R}^+$ and that $f_H(0) > 0$.

For example, for a Rayleigh fading channel, H has an exponential distribution with $f_H(0) = \frac{1}{\mathbb{E}H}$. This assumption has several consequences that will be used in the following. First, a direct consequence is that $f_H(h) > 0$ for all h within some sufficiently small neighborhood of zero. From this it follows that

$$\Pr(|H_n|^2 < h) = \Theta(h)$$

as h goes to zero. Also, for any such distribution $\mathbb{E}_H\left(\frac{1}{H}\right) = \infty$ and so, $P^*(1) = \infty$. Hence, $D^*(\bar{P})$ will be strictly greater than 1 for any finite $\bar{P}$ but will approach 1 as $\bar{P} \to \infty$. Finally, for any such distribution with finite mean, we have the following lemma.

Lemma 2.1 *Let $f_H(h)$ be a fading density that satisfies Assumption A and has a finite mean. Given any $h_t > 0$, then there exists a constant K_2 such that for all $h < h_t$,*

$$\int_h^\infty \frac{1}{h} f_H(h)\, dh \geq K_1 \ln\left(\frac{1}{h}\right) + K_2,$$

where $K_1 = \inf\{f_H(h) | h \leq h_t\}$.

Likewise, these exists a constant $\tilde{K}_2$, such for all $h < h_t$,

$$\int_h^\infty \frac{1}{h} f_H(h)\, dh \leq \tilde{K}_1 \ln\left(\frac{1}{h}\right) + \tilde{K}_2,$$

where $\tilde{K}_1 = \sup\{f_H(h) | h \leq h_t\}$.

This implies that the integral $\int_h^\infty \frac{1}{h} f_H(h)$ grows like $\ln\left(\frac{1}{h}\right)$ as $h \to 0$. The lower bound follows because

$$\begin{aligned}
&\int_h^\infty \frac{1}{h} f_H(h)\, dh \\
&\geq \int_h^{h_t} \frac{1}{h} K_1\, dh + \int_{h_t}^\infty \frac{1}{h} f_H(h)\, dh \\
&= K_1(\ln(h_t) - \ln(h)) + \int_{h_t}^\infty \frac{1}{h} f_H(h)\, dh \\
&= K_1 \ln\left(\frac{1}{h}\right) + K_2.
\end{aligned}$$

Note that since H has a finite mean, $f_H(h)$ must go to zero faster than $\frac{1}{H}$ as $H \to \infty$. Hence, $\int_{h_t}^\infty f_H(h)\, dh$ is finite for all $h_t > 0$. The upper bound can be derived in a similar manner.

3 Asymptotic Analysis

3.1 A Lower bound on the convergence rate

In this section we consider a lower bound on the rate at which $D^*(\bar{P})$ approaches 1.

Proposition 3.1 *Assuming the fading distribution satisfies Assumption A, as $P \to \infty$, $D^*(\bar{P}) - 1 = \Omega\left(e^{-\alpha\bar{P}}\right)$, for any $\alpha > \left(\mathbb{E}_A P(1, A) f_H(0)\right)^{-1}$*

In other words, the buffer delay decreases to 1 no faster than $e^{-\alpha P}$ under any sequence of policies.

PROOF. To prove this proposition, we consider a fictitious system where every packet leaves after waiting for 2 time-units without requiring any power (recall each packet must wait at least one time-unit). The maximum delay in the fictitious system will be no more than 2 time-units. For a given average power $\bar{P}$, let $\hat{D}(\bar{P})$ be the minimum average delay experienced by a packet in this fictitious system. Clearly we must have $\hat{D}(\bar{P}) \leq D^*(\bar{P})$ for all $\bar{P}$. We derive our lower bound on the rate of convergence for $D^*(\bar{P})$ by calculating the rate at which $\hat{D}(\bar{P}) - 1$ goes to zero as $\bar{P}$ increases.

Under the assumption that every packet leaves after 2 time-slots, the buffer dynamics can be re-written as

$$S_{n+1} = \max(A_n - U_n + A_{n+1}, A_{n+1}).$$

Also, at each time n an optimal policy will set $U_n \leq A_n$, since any other packets waiting in the buffer will leave the system anyway without requiring any power. Therefore, an optimal policy for this system can be expressed as function of the current channel state, H_n and the number of arrivals A_n. It follows that $\hat{D}(\bar{P})$ is the solution to the following optimization problem:

$$\begin{aligned} \underset{U:\mathcal{H}\times\mathcal{A}\mapsto\mathbb{R}^+}{\text{minimize}} \; & 1+\frac{1}{\bar{A}}\mathbb{E}_{H,A}\left\{(A-U(H,A))^+\right\} \\ \text{subject to: } & \mathbb{E}_{H,A}P(H,U(H,A))\leq \bar{P} \\ & U(h,a)\geq 0, \quad \forall h\in\mathcal{H}, a\in\mathcal{A} \end{aligned}$$

Here the objective function correspond to the expected number of packets in the system under the policy $U(H,A)$ divided by the average arrival rate, which by Little's law is equal to the average delay. Note, we have also used the fact that the arrivals are i.i.d.

The solution, $U^*(H,A)$, to this problem is similar to the classic "water-filling" power allocation [9]. In particular it can be characterized as follows:

$$P(h,U^*(h,a)) = \begin{cases} 0 & \text{if } h\leq h_L \\ \frac{1}{\lambda}-\frac{\sigma^2}{h} & \text{if } h_L\leq h\leq h_U(a) \\ \frac{\sigma^2}{h}\left(2^{a/N}-1\right) & \text{if } h\geq h_U(a) \end{cases} \tag{4}$$

Here, $\lambda>0$ is a constant chosen to satisfy the average power constraint. The parameters $h_L<h_U(a)$ are channel gain thresholds that satisfy $h_L=\lambda\sigma^2$ and $h_U(a)=\lambda\sigma^2 2^{a/N}$. Notice that the upper threshold depends on the value of A_n. When $h<h_U(a)$, (4) corresponds to the usual water-filling power allocation, where $\frac{1}{\lambda}$ is the "water level" and h_L represents the threshold below which no power is used. When $h>h_U(a)$, the transmitter inverts the channel to transmit at the constant rate a; this is due to fact that $U^*(H,A)\leq A$, as noted above. Also notice that as $\bar{P}$ is increased, λ will decrease and thus so will h_L and $h_U(a)$.

From the above, it follows that

$$\begin{aligned} \hat{D}(P)-1 &= \frac{1}{\bar{A}}\mathbb{E}_{H,A}(A-U^*(H,A))^+ \\ &= \frac{1}{\bar{A}}\Bigg(\Pr(H<h_L)\bar{A}+\Pr(H\geq h_L) \\ &\quad \times \mathbb{E}_{H,A}\left((A-U^*(H,A))^+\middle|H\geq h_L\right)\Bigg) \\ &\geq \Pr(H<h_L) \end{aligned}$$

Here, we have used that $(A-U^*(H,A))^+$ is always non-negative and is equal to A for $H<h_L$. Using Assumption A, we then have, as $h_L\to 0$,

$$\hat{D}(P)-1=\Omega(h_L). \tag{5}$$

Next we bound the average power as follows:

$$\begin{aligned} \bar{P} &= \int_{a_{min}}^{a_{max}}\left(\int_{h_L}^{h_U(a)}\left(\frac{1}{\lambda}-\frac{1}{h}\right)f_H(h)\,dh \right.\\ &\quad \left. +\int_{h_H}^{\infty}\frac{\sigma^2}{h}\left(2^{a/N}-1\right)f_H(h)\,dh\right)dF_A(a) \\ &\geq \int_{a_{min}}^{a_{max}}\int_{h_U(a)}^{\infty}\frac{\sigma^2}{h}\left(2^{a/N}-1\right)f_H(h)\,dh\,dF_A(a) \end{aligned}$$

Pick some constant h_t such that $\inf\{f_H(h)|h\leq h_t\}=K_1>0$. Then, from Lemma 2.1, we have that for all $h<h_t$

$$\int_h^\infty \frac{1}{h}f_H(h)\,dh\geq K_1\ln\left(\frac{1}{h_H}\right)+K_2.$$

For a large enough average power, $h_U(a)$ will be less than h_t for all $a\in\mathcal{A}$. In this case,

$$\bar{P}\geq\int_{a_{min}}^{a_{max}}\left(K_1\ln\left(\frac{1}{h_U(a)}\right)+K_2\right)P(1,a)\,dF_A(a).$$

Note that $h_U(a)=h_L 2^{a/N}$. Hence,

$$\bar{P}\geq K_1\ln\left(\frac{1}{h_L}\right)(\mathbb{E}_A P(1,A))+K_3,$$

where K_3 is a constant independent of h_L.

From this it follows that as $\bar{P}\to\infty$, $h_L=\Omega\left(e^{-\alpha\bar{P}}\right)$, where $\alpha=(K_1(\mathbb{E}_A P(1,A)))^{-1}$. Combining this with (5), we have $\hat{D}-1=\Omega\left(e^{-\alpha\bar{P}}\right)$ and so $D^*(P)-1=\Omega\left(e^{-\alpha\bar{P}}\right)$ either, as desired. Finally, we note that in the above bound, $K_1\geq f_H(0)$ and be made arbitrarily close to this value by choosing h_t small enough. This gives the desired lower bound on α. ■

3.2 A nearly order optimal sequence of policies

In this section, we show that the above bound can nearly be achieved by a simple sequence of *channel threshold* policies in which the transmitter only transmitts when the channel gain H is larger than a threshold h_{th}. Given that the channel is greater than this threshold, we consider a policy that sets $U_n=A_n+\delta$, where $\delta>0$ is a small constant.[4] Clearly, if $s_n/N<A_n+\delta$ then there is not enough information in the buffer to transmit. In this case we can assume the transmitter sends extra "dummy" bits. This is clearly a poor choice from the view of saving power, but is sufficient for our purposes. A sequence of these policies can nearly achieve the optimal convergence rate in the low power regime.

Proposition 3.2 *There exists a sequence of channel threshold policies $\{\mu_K\}$ such that as $k\to\infty$, $\bar{P}^{\mu_k}\to\infty$, and $D^{\mu_k}-1=O(\exp(-\alpha P^{\mu_k}))$, for any $\alpha<(f_H(0)(\mathbb{E}_A P(1,A)))^{-1}$*

PROOF. Let $\{\mu_k\}$ be a sequence of channel threshold policies as in the lemma with a fixed parameter δ for each policy, and such that as $k\to\infty$, h_{th} decreases to 0.

[4] Note that policy does not base the transmission decision on only H_n and S_n, but can be viewed as using the past history of S_n and U_n. Such a dependence is not needed for an optimal policy.

The average power of such a policy is given by

$$\bar{P}^{\mu_k} = \int_{h_{th}}^{\infty} \int_{a_{min}}^{a_{max}} \frac{\sigma^2}{h} \left(2^{a+\delta/N} - 1\right) f_H(h)\, dh\, dF_A(a).$$

By a similar argument as in the proof of Prop. 3.1, it can be shown that as $\bar{P}^{\mu_k}$ increases, $h_{th} = O\left(e^{-\alpha \bar{P}^{\mu_k}}\right)$, where $\alpha = (\mathbb{E}_A P(1, A+\delta)\tilde{K}_1)^{-1}$ and $\tilde{K}_1$ is the constant from Lemma 2.1.

To bound the average delay under such a policy we use the following lemma:

Lemma 3.3 *For any policy where $U_n - A_n$ is an i.i.d. sequence, the average buffer occupancy is bounded by*

$$\mathbb{E}S \leq \frac{\sigma_\Delta^2}{2(\bar{\Delta})} + \bar{A},$$

where $\bar{\Delta}$ and σ_Δ^2 are the mean and variance of $(U_n - A_n)$.

This is essentially the same as Kingman's upper bound on the average delay for a continuous-time G/G/1 queue [10]. Notice that here we are bounding the average buffer occupancy in a discrete-time queue with dynamics given by (2); however, the same bounding techniques can be used.

Using this lemma and applying Little's law, we that the average delay is bounded by

$$\bar{D}^{\mu_k} - 1 \leq \frac{\sigma_\Delta^2}{2(\bar{\Delta})\bar{A}} \tag{6}$$

Evaluating this for a channel threshold policy we have,

$$\bar{D}^{\mu_k} - 1 \leq \frac{(1-q)q(\delta^2 + 2\delta\bar{A}) + q\bar{A}^2 - q^2(\bar{A})^2}{2((1-q)\delta - (q)\bar{A})(\bar{A})}, \tag{7}$$

where $q = \Pr(|H|^2 < h_{th})$. From this it is clear that $\bar{D}^{\mu_k} - 1 = O(q)$ as $q \to 0$. Again using Assumption A, $q = \Theta(h_{th})$, and hence we have $D^{\mu_k} - 1 = O\left(\exp(-\alpha P^{\mu_k})\right)$ as desired. ■

3.3 Suboptimal policies

Next we illustrate two simple policies with suboptimal convergence rates. First we consider a simple *fixed power* policy which does not depend on the buffer state. By this we mean a policy in which the transmitter transmits at a fixed power, $\bar{P}$ in each slot, i.e.,

$$\Psi^{\bar{P}}(h) = N \log\left(1 + \tfrac{h\bar{P}}{\sigma^2}\right) \quad \text{for all } s, h$$

Using this policy the average power is clearly equal to $\bar{P}$. The expected rate under a fixed power policy is increasing with $\bar{P}$; the rate at which it is increasing can be shown to satisfy:

$$\mathbb{E}_H\left\{\Psi^{\bar{P}}(H)\right\} = \Theta(\log(\bar{P})).$$

Next we consider the variance of $\Psi^{\bar{P}}(H)$. It can be shown that as $\bar{P}$ increases, the variance is increasing. However, asymptotically the variance is bounded. Specifically, we have

Lemma 3.4 *Under a fixed power policy, $\Psi^{\bar{P}}$, for all $\bar{P}$,*

$$\mathit{var}\left(\Psi^{\bar{P}}(H)\right) \leq \tfrac{N}{2}\mathit{var}\left(\log(H)\right)$$

To see this note that

$$\mathrm{var}\left(\Psi^{\bar{P}}(H)\right) = \frac{1}{2}\mathbb{E}_{H,\tilde{H}}\left(\Psi^{\bar{P}}(H) - \Psi^{\bar{P}}(\tilde{H})\right),$$

where $\tilde{H}$ is another random variable, independent of H and identically distributed. Substituting the expression for $\Psi^{\bar{P}}(h)$, we have

$$\mathrm{var}\left(\Psi^{\bar{P}}(H)\right) = \frac{1}{2}\mathbb{E}_{H,\tilde{H}}\left\{N \log\left(\frac{1 + \frac{H\bar{P}}{\sigma^2}}{1 + \frac{\tilde{H}\bar{P}}{\sigma^2}}\right)\right\}.$$

As $\bar{P} \to \infty$ this converges to

$$\tfrac{N}{2}\mathbb{E}_{H,\tilde{H}}\left\{\log\left(\frac{H}{\tilde{H}}\right)\right\} = \tfrac{N}{2}\mathrm{var}\left(\log(H)\right),$$

as desired.

Using these observations in the bound from lemma 3.3, it follows that the average buffer occupancy using a fixed power policy is upper bounded by

$$\mathbb{E}S \leq \frac{\sigma_A^2 + \frac{N}{2}\mathrm{var}\left(\log(H)\right)}{2\left(\mathbb{E}_H\left\{\Psi^{\bar{P}}(H)\right\} - \bar{A}\right)} + \bar{A}$$

Therefore, from Little's law, the average buffer delay, $\bar{D}$ is upper-bounded by

$$\bar{D} \leq \frac{K}{\mathbb{E}_H\left\{\Psi^{\bar{P}}(H)\right\} - \bar{A}} + 1.$$

where K is a constant, depending on the arrival and channel statistics but not on the average power.

Finally, since $\mathbb{E}_H\left\{\Psi^{\bar{P}}(H)\right\}$ increases at rate $\Theta(\log P)$, it follows that the average delay using a fixed power policy decreases at a rate faster than $O\left(\frac{1}{logP}\right)$.

Next we give a lower bound for the convergence rate of a fixed power policy. Specifically we have

Proposition 3.5 *For any sequence of fixed power policies, $D^{\mu_k} - 1 = \Omega(\frac{1}{P\log P})$ as $P \to \infty$.*

This can be shown by considering the following lower bound, which is analogous to Kingman's lower bound for a G/G/1 queue [10]. Specifically, the average buffer occupancy is lower bounded by

$$\bar{S} - \bar{A} \geq \frac{\mathbb{E}((A-U)^+)^2}{2(\bar{U} - \bar{A})}$$

where $(A-U)^+ = \max(A-U,0)$, and A and U are two random variables with the respective steady state distributions.

As before the denominator will increase at rate $\Theta(\log(P))$. To bound the numerator, note that by Markov's inequality

$$\mathbb{E}((A-U)^+)^2 \geq \Pr(A-U > \epsilon)\epsilon^2$$

and

$$\Pr(A-U \geq \epsilon) = \Pr\left(|H|^2 \leq \frac{2^(A-\epsilon)}{P}\right).$$

By Assumption A, this last term is decreasing at rate $\Theta(\frac{1}{P})$. Therefore combining these results we have that $D^{\mu_k} - 1 = \Omega\left(\frac{1}{P\log P}\right)$ as desired. Note this is much slower than the optimal convergence rate obtained by a channel threshold policy.

Next consider a sequence of "fixed water-filling" policy. By this we mean a policy that uses a water-filling power (and rate) allocation, once again independent of the buffer state. For such a sequence of policies the same bounds once again apply.

Proposition 3.6 *For any sequence of fixed water-filling policies, $\{\mu_k\}$, with $\bar{P}^{\mu_k} \to \infty$, $barD^{\mu_K} - 1 = O\left(\frac{1}{\log P^{\mu_K}}\right)$. Furthermore assuming the channel gain satisfies assumption A, then $\bar{D}^{\mu_K} - 1 = \Omega\left(\frac{1}{\bar{P}^{\mu_K}\log \bar{P}^{\mu_k}}\right)$*

The proof of this follows similar arguments to the above.

3.4 Finite state channels

In this section we briefly discuss a class of channels that does not satisfy assumption A. Specifically, consider a finite state channel, *i.e.*, a channel where $|\mathcal{H}| < \infty$, and assume that for each $h \in \mathcal{H}$, $h > 0$. In this case, $P^*(1)$ will be finite. Therefore, we consider the rate at which the average delay converges to 1 as a function of how fast $\bar{P}$ is converging to $P^*(1)$. Specifically we have

Proposition 3.7 *For a finite state channel, as $P \to P^*(1)$, $D^*(P) = \Omega\left(\log\left(\frac{P^*(1)}{P}\right)\right)$.*

This can be proved using a similar argument as in Prop. 3.1.

4 Conclusions

In this paper we consider the behavior of the optimal power delay trade-off in the regime of small delay and large power. In this regime we showed that the average delay decreases at the rate of $e^{-\alpha P}$, as the power P increases, assuming that the channel gain's density is strictly positive at zero. This implies that when delay is tightly constrained, a small increase in the delay will result in substantial power saving. We also illustrated that a simple channel threshold policy is order optimal and discussed several other suboptimal policies.

5 Acknowledgments

The author wishes to thank Vijay Subramanian at Motorola for several helpful discussions, and Bruce Hayek for suggesting the proof to Prop. 3.1.

References

[1] B. Collins and R. Cruz, "Transmission Policies for Time Varying Channels with Average Delay Constraints," in *Proc. 1999 Allerton Conf. on Commun. Control, & Comp.*, (Monticello, IL), 1999.

[2] R. Berry and R. Gallager, "Communication over fading channels with delay constraints," *IEEE Trans. on Information Theory*, vol. 48, pp. 1135–1149, May 2002.

[3] R. Berry and R. Gallager, "Buffer control for communication over fading channels," in *Proc. 2000 IEEE International Symposium on Info. Th.*, (Sorrento, Italy), p. 409, June 25-30 2000.

[4] R. Berry, "Some dynamic resource allocation problems in wireless networks," in *Proc. of SPIE* (E. Chong, ed.), vol. 4531, (Denver, CO), pp. 37–48, 2001.

[5] B. Prabhakar, E. Uysal-Biyikoglu, and A. E. Gamal, "Energy-efficient transmission over a wireless link via lazy packet scheduling," in *Proc. IEEE Infocom 2001*, 2001.

[6] Bettesh and S. Shamai, "Optimal power and rate control for fading channels," in *Proc. IEEE Vehicular Technology Conference, Spring 2001*, (Rhodes, Greece), May 6-9 2001.

[7] D. Rajan, A. Sabharwal, and B. Aazhang, "Transmission policies for bursty traffic sources on wireless channels," in *Proc. of 35th Annual CISS*, (Baltimore), Mar 2001.

[8] L. Ozarow, S. Shamai, and A. Wyner, "Information Theoretic Considerations for Cellular Mobil Radio," *IEEE Tranactions on Vehicular Technology, Vol. 43, No. 2*, pp. 359–378, May 1994.

[9] R. Gallager, *Information Theory and Reliable Communication*. New York: John Wiley and Sons, 1968.

[10] L. Kleinrock, *Queueing Systems*, vol. II. John Wiley and Sons, 1976.