# Dual Problems in Power Control

Ashutosh Sabharwal, Dinesh Rajan and Behnaam Aazhang

Department of Electrical and Computer Engineering
Rice University, Houston, TX 77005.
{ashu,dinesh,aaz}@rice.edu

**Abstract**

In this paper, we study power control to minimize packet loss for delay-bounded transmission of bursty sources over Gaussian channels and transmission of constant rate sources over time-varying block fading channels. In the above two problems, we consider the effect of queuing delay on the achievable packet loss rates. First, we show that small additional delay helps substantially reduce the packet loss probability in fading channels. Next, we show that for transmission of bursty sources over Gaussian channels, a small additional queuing delay also leads to reduction in packet loss probability. Finally, the duality between the two solutions is highlighted for a specific delay case, suggesting that the duality holds for all delays. A special case of our results is the traditional channel based power control, which typically assumes that each packet is served as soon as it arrives without any queuing delay.

## 1 Introduction

The goal of future wireless systems is to provide a multitude of services with varying requirements (like delays and error rates) efficiently using the available resources of limited power and bandwidth. Techniques to overcome the challenges posed by the wireless channel are employed at the physical layer (*e.g.*, channel coding, multiple transmit and receive antennas), and at the network layer (*e.g.*, ARQ, scheduling). In this paper, we study the interplay between queuing delay (a network layer metric) and transmission power (physical layer metric).

The importance of incorporating network parameters in designing the physical layer of wireless systems has been recognized [1, 2]. The effect of queueing packets in multiple access channel is investigated in [3]. Schedulers that minimize transmission power in fading channels is proposed in [4] under average delay constraints. The interaction between source fidelity and communication networks is investigated in [5].

In this paper, we explicitly construct power control policies that reduce packet outage probabilities[1] in fading channels at the cost of increased delays. We also show an analogous result when transmitting bursty sources through time invariant channels. We find optimal power control policies for transmitting a time varying source through an AWGN channel without any queuing delay. We proceed to construct policies which reduce packet outage probabilities with increasing delays. To demonstrate these analogous results to the fading channel based power control schemes, we construct a Rayleigh source which

[1] defined explicitly in the sequel.

produces time varying information rate. We also show the duality between power control based on channel and source statistics for instantaneous transmission of data which further suggests the duality holds for finite delay transmission schemes too.

Throughout the paper, two scenarios are considered. In the first case, the source produces one packet in each time slot and is transmitted over a time-varying block fading channel. In the second case, a source produces packets with variable information content and is transmitted over a time-invariant additive white Gaussian noise channel. We consider a time-slotted system, where all transmissions occur at periodic time-interval. We assume that the time-slots are long enough to allow transmission close to the mutual information bound, thereby allowing us to use optimal Gaussian codes for each transmission.

The remainder of the paper is organized as follows. We introduce the system model including buffer, fading and source models in Section 2. Effect of delay on outage probabilities in fading channels is investigated in Section 3 and for bursty sources in Section 4. We briefly discuss the duality between the power control for the fading channels and bursty sources with conclusions in Section 5.

# 2 System Model

In this section, we discuss the system model. Since our focus is to study effect of delay on outage performance of algorithms, we discuss the assumptions regarding the buffer model and queuing disciplines in Section 2.1. In Section 2.2, the Rayleigh fading channel model is presented with a brief discussion of relevant results from literature. In Section 2.3, we motivate a *block source* which produces variable number of information bits per time block, and derive power control to minimize outage with stringent delay constraints.

## 2.1 Buffer Model

The system is assumed to be time-slotted. All information is assumed to arrive in packets at the beginning of a time-slot, and is available for transmission only at the beginning of the subsequent time-slot. Thus, each packet of information suffers a delay of at least one time-slot. To simplify the development of major ideas, the transmission systems are subject to one of the following two constraints:

1. *Packet integrity constraint* : Under the packet integrity constraint, the transmission policies are constrained to maintain the integrity of the packets and no fragmentation is allowed. Also, the service discipline is constrained to be first in first out (FIFO).[2]

2. *Packet fragmentation* : Under the packet fragmentation constraint, packets can be fragmented upto arbitrary precision (in essence, assuming a fluid source). The remaining portion of the partially serviced packet can be either stored in the queue or dropped as determined by the power control policy. When the fragmented packets are queued, they are serviced before the other packets (FIFO).

The packet integrity constraint is used in channel based power control, discussed in Sections 2.2 and 3. The packet fragmentation constraint is applied to source based power

[2]FIFO is essential only in power control policies providing absolute delay guarantees. To achieve an average delay bound, FIFO is not critical.

control in Sections 2.3 and 4. The average delay constraint is equivalent to a constraint on average buffer length ($\mathbb{E}(q_t)$) using Little's result [6].

## 2.2 Rayleigh Fading Channel

In the block fading channel, the complex fading coefficient $h_t$ is assumed to be constant over $T_c$ consecutive symbols. The complex received signal $y_t$ is given by,

$$y_t = h_t s_t + \epsilon_t. \tag{1}$$

The transmitted signal $s_t$, the received signal $y_t$ and the additive noise $\epsilon_t$ are $T_c$ dimensional complex row vectors. The real and imaginary parts of $h_t$ are assumed to be independent zero mean Gaussian, each with variance 1/2. Also, the additive noise $\epsilon_t$ is assumed to be circularly symmetric Gaussian with zero mean and covariance $\sigma^2 \mathbf{I}_{T_c}$ where $\mathbf{I}_{T_c}$ is the identity matrix of size $T_c$.

In [7], it was shown that for large $T_c$, the conditional mutual information between $s_t$ and $y_t$, $I(y_t; s_t|h_t)$, is a good indicator of the performance of practical codes. For the case with perfect channel state information at receiver (CSIR), the mutual information $I(y_t; s_t|h_t)$ is given by

$$I(y_t; s_t|h_t) = T_c \log\left(1 + P_t \frac{|h_t|^2}{\sigma^2}\right), \tag{2}$$

where $P_t$ is the average transmitted power over $T_c$ consecutive symbols. Thus, the instantaneous channel capacity is time-varying and depends on the current channel state. Further, in [7], the concept of capacity versus outage was defined for delay constrained transmission. A block of data transmitted at $R$ nats/second/Hz is considered in outage if the instantaneous mutual information $I(y_t; s_t|h_t)$ is less than $R$. If both transmitter and receiver can perfectly track $h_t$, then the transmitter can perform the following power control to minimize outage [8],

$$P_{1,t} = \begin{cases} \frac{e^R - 1}{\gamma_t} & \gamma_t \geq \gamma_o \\ 0 & \gamma_t < \gamma_o \end{cases}, \tag{3}$$

where $\gamma_t = |h_t|^2/\sigma^2$ and $\gamma_o$ is chosen to satisfy the average power constraint, $\int_\gamma P_{1,t} p(\gamma) d\gamma \leq \overline{P}$. The minimum outage probability is computed as $\Pi^*_{1,c} = \text{Prob}(\gamma < \gamma_o) = 1 - e^{-\gamma_o}$. Since each packet experiences a delay of one time-slot before being serviced (either dropped or transmitted), a buffer of size one is sufficient which also explains the subscript 1 in $P_{1,t}$ and $\Pi^*_{1,c}$.

The ergodic capacity, which is the average long-term rate, can be computed as $C_e = \mathbb{E}_{h_t}\left\{1 + P_{\infty,t}\gamma_t\right\}$ using the following power control [9]

$$P_{\infty,t} = \begin{cases} \frac{1}{\gamma_o'} - \frac{1}{\gamma_t} & \gamma_t \geq \gamma_o' \\ 0 & \gamma < \gamma_o' \end{cases}. \tag{4}$$

where $\gamma_o'$ is chosen to satisfy the power constraint, $\int_\gamma P_{\infty,t} p(\gamma) d\gamma \leq \overline{P}$. Note that the outage, $\Pi^*_{\infty,c}$, is zero if $R < C_e$, while some packets can suffer arbitrarily large delays, which explains the subscript $\infty$ in $P_{\infty,t}$ and $\Pi^*_{\infty,c}$.

## 2.3 Rayleigh Variable Rate Source

In the fading channels, the time-variations of the channel leads to a system with a time-varying "supply" of available capacity. Motivated by the fading channel models, we define a block source which has a time-varying "demand." Much like a channel with time-varying mutual information, the dual source produces time-varying bit rate. As an example of such a source, consider the following block Gaussian source.

A block complex Gaussian source outputs $T_s$ consecutive source symbols which are independent and follow a zero mean Gaussian distribution with variance $\alpha_t^2$. The rate-distortion bound for compressing a source to a distortion level of $D$ requires a rate of

$$x_t = \begin{cases} \log\left(\frac{\sigma_t^2}{D}\right) & \text{if } \sigma_t^2 > D \\ 0 & \text{if } \sigma_t^2 \leq D \end{cases} \tag{5}$$

Thus, the source outputs packets of variable size in each time-slot. An example of variable rate source is video, which can be considered as a concatenation of "scenes," each $K$ video frames long. Each scene has a different information content, captured by $\sigma_t^2$; good video encoders (like MPEG4 encoders) encode scenes with different number of bits with relatively constant distortion level across scenes. Note that $R_t = 0$ for $\sigma_t^2 \leq D$. To maintain a dual relationship with fading channel power control in Section 2.2, we assume $\sigma_t^2 = z_t^2 + D$ where $z_t^2$ is an i.i.d. random process with Rayleigh distribution; hence we label $x_t$ as a *Rayleigh* source.

In this case the complex received signal $y_t$ is given by,$y_t = s_t + \epsilon_t$.. The objective is to transmit the packets with minimal outage with a power and delay constraint using the packet fragmentation policy of Section 2.1. The probability of outage is defined as the ratio of the total information rate dropped to the total information rate produced by the source.[3] Parallel to the development of channel based power control, we derive the outage minimizing power control policy which achieve packet delay of either one time-slot or the delay is unbounded like that achieved by $P_{\infty,t}$.

If all packets are serviced (completely transmitted, partially dropped or completely dropped) in the subsequent time-slot, then the delay of that power control policy is one time-slot. The probability of outage for an arbitrary power control policy, $Q_t$ is given by

$$\Pi_{1,s}(Q) = \int_0^\infty \left(x_t - \log\left(1 + \frac{Q_t}{\sigma^2}\right)\right) p(x_t)dx_t, \tag{6}$$

where $p(x_t)$ is the distribution of the instantaneous arrival rate. The average power used by policy $Q_t$ is $\int_0^\infty Q_t\, p(x_t)dx_t$. Note that power control policy is a function of the buffer state. Since all packets are serviced in the next time-slot, the power control action in time-slot $t$ depends only on rate $x_{t-1}$. The optimal power control policy for delay one, denoted $Q_{1,t}^*$ is given by the following proposition.

**Proposition 1** *For a source producing time varying information rate $x_t$ with distribution $p(x_t)$ transmitted over an AWGN channel with a power constraint $\overline{P}$. The power control policy $Q_{1,t+1}$ that minimizes probability of outage is*

$$Q_{1,t+1} = \begin{cases} \sigma^2\left(e^{x_t} - 1\right), & x_t \leq r_1 \\ \sigma^2\left(e^{r_1} - 1\right), & r_1 \leq x_t \end{cases} \tag{7}$$

[3]Our definition assumes that packets can be fragmented infinitely. An alternate definition of packet outage is to define it as the fraction of packets transmitted succesfully. However the later definition penalizes the loss of two packets with different amounts of information equally in terms of outage.

*where the threshold $r_1$ depends on the available average power $\overline{P}$ and distribution $p(x_t)$. The resulting outage is given by*

$$\Pi^*_{1,s} = \int_{r_1}^{\infty} (x_t - r_1)\, p(x_t) dx_t. \tag{8}$$

**Proof**: See Appendix A □

Note that we cannot transmit more information than what is available, thus, *i.e.*, $0 \leq \log(1+Q_{1,t}/\sigma^2) \leq x_t$. In the absence of a delay constraint, the transmitter will transmit at a constant rate, which is equal to the ergodic rate of the source, $R_e = \mathbb{E}\left\{\log\left(1 + \frac{z_t^2}{D}\right)\right\}$. If the ergodic rate is less than the Gaussian channel capacity, $R_e < C_g$, then the constant transmission server achieves zero outage. The corresponding power control is simply $Q_{\infty,t} = \sigma^2(e^{R_e} - 1)$ in each time-slot.

Notice the similarities between the power control to minimize outage for the two cases (3) and (7). In (3), only good channel are used and packets are dropped in poor channel conditions. Analogously, in (7), high rate packets are partially dropped while low-rate packets are transmitted.

# 3 Delay Bounded Power Control for Fading Channels

In this section, we construct a scheduling policy which reduces packet outage by increasing packet delay. Our main objective is to demonstrate the outage-delay tradeoff, and hence we emphasize on algorithm simplicity and not on its optimality.

Consider a constant arrival process of one fixed size packet every time slot. When all the packets are transmitted in the next slot after their arrival, the achievable mean delay is one time-slot and the resulting packet outage is given by $\Pi^*_{1,c}$. The main idea of the proposed scheduling algorithm can be appreciated by the following special case, where each packet can be delayed by a maximum of two time-slots.

Two-state scheduling policy : In this example, the buffer size is assumed to be two packets. Thus, the buffer can have either one packet or two packets, leading to the two states ($q_t = 1, 2$) in Figure 1. When there is only one queued packet, the scheduler can either transmit it if the channel conditions are good or delay the decision by one time-slot. When there are two queued packets at the beginning of a time-slot, then the transmitter can either transmit one or two packets or drop the head of the line (HOL) packet. Thus, we assume that outage occurs only in state $q_t = 2$. In each state, the decisions are made based only on the current channel conditions, given by the instantaneous channel SNR $\gamma$. The scheduler decisions are summarized in Table 1, where $q_t$ denotes the number of queued packets at the beginning of time-slot $t$. Based on the scheduler actions given in Table 1, the state transition probabilities can be computed as follows:

$$p_{11} = \int_{\gamma_{11}}^{\infty} e^{-\gamma} d\gamma = e^{-\gamma_{11}}, \qquad p_{12} = 1 - p_{11}, \quad p_{21} = \int_{\gamma_{22}}^{\infty} e^{-\gamma} d\gamma = e^{-\gamma_{22}}, \qquad p_{22} = 1 - p_{21}.$$

Denote the probability of being in state $i$ by $s_i$. Then $s_1 = \frac{p_{12}}{p_{12}+p_{21}}$ and $s_2 = \frac{p_{21}}{p_{12}+p_{21}}$. Whenever the scheduler chooses to transmit packets, enough power is used to ensure

Table 1: Policy with buffer of size of 2.

| $q_t$ | $\gamma_t$ | Action |
|---|---|---|
| 1 | $\gamma_t > \gamma_{11}$ | Transmit 1 packet |
| 1 | $\gamma_t < \gamma_{11}$ | Delay 1 packet |
| 2 | $\gamma_t > \gamma_{22}$ | Transmit 2 packets |
| 2 | $\gamma_{21} < \gamma_t < \gamma_{22}$ | Transmit HOL packet, Delay other packet |
| 2 | $\gamma_t < \gamma_{21}$ | Drop HOL packet, Delay other packet |

reliable reception of the packet. The resulting average power is then given by

$$P_{avg} = s_1(e^S - 1)Ei(\gamma_{11}) + s_2(e^S - 1)(Ei(\gamma_{21}) - Ei(\gamma_{22})) + s_2(e^{2S} - 1)Ei(\gamma_{22}), \quad (9)$$

where $Ei(x) = \int_x^\infty \frac{e^{-\gamma}}{\gamma} d\gamma$ and $S$ is the size of the packet. The outage of the resulting two-state scheduler is given by $s_2 \int_0^{\gamma_{21}} e^{-\gamma} d\gamma = s_2(1 - e^{-\gamma_{21}})$. Finding the minimum outage parameters is posed as the following optimization problem,

$$\Pi_{2,c} = \min_{\substack{\gamma_{11},\gamma_{21},\gamma_{22} \\ P_{avg} < \overline{P}}} s_2(1 - e^{-\gamma_{21}}). \quad (10)$$

We state without proof that $\Pi_{2,c} \leq \Pi^*_{1,c}$, thus additional delay reduces the packet outage.

<u>$K$-state scheduling policy</u> : When the buffer size is larger than 2, we can achieve lower outages than with buffer of length 2 at the cost of additional delay. The scheduling decisions of the above two-state scheduling policy are extended as follows:

1. Label the $K$ possible buffer states $q_t$ as $1, 2, \ldots, K$. Packets can be dropped only in buffer state $q_t = K$. Atmost one packet is dropped in any time slot.

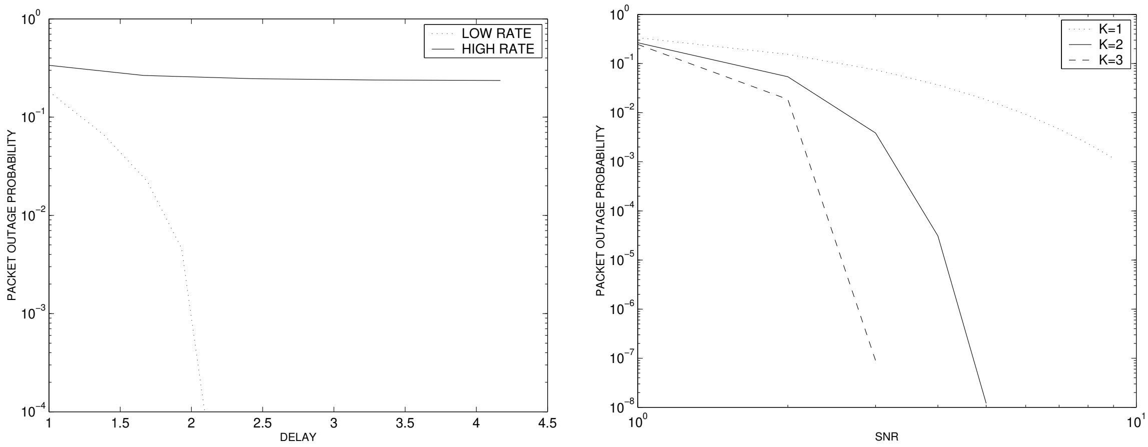
2. In each buffer state $q_t = k$, we can transmit any number of packets from 0 to $k$ depending on channel conditions. Thus, we find $k$ corresponding thresholds, $\gamma_{k1} \leq \gamma_{k2} \ldots \leq \gamma_{kk}$. For simplicity, we assume $\gamma_{k0} = 0$ and $\gamma_{k(k+1)} = \infty$.

3. We transmit $i$ packets in buffer state $q_t = k$ if $\gamma_{ki} < h_t < \gamma_{k(i+1)}$.

The average power and outage expressions can be easily obtained using the same procedure as in the above two-state scheduler example. The following numerical results show that $\Pi_{K,c} \leq \Pi_{K-1,c}$ for flat fading Rayleigh channels.

<u>Numerical Results</u>: The plot of the outage versus buffer length $K$ is given in Figure 2 for two different values of arrival packet sizes at a given average power limit. We fix $\overline{P}$ such that the ergodic capacity of the channel equals 0.7 nats/sec/Hz. For arrival rate of 0.6 nats/sec/Hz, which is lower than the ergodic capacity of the channel, the outage decreases rapidly with increasing buffer length and approaches zero for large buffer lengths. For an average arrival rate of 0.9 nats/sec/Hz, which is higher than the ergodic capacity, the packet outage probability does not approach zero with increasing buffer lengths. The packet outage probability asymptotically approaches $1 - \frac{C_e}{R_{avg}}$. These results are clearly evident from Figure 2(a) which shows the outage variation with delay for arrival rates of 0.6 and 0.9 nats/sec/Hz. The outage probabilities shown are not the lowest outages achievable for a given buffer length and delay, but are an upper bound.

Figure 1: Schematic of 2 state buffer.




(a) (b)

Figure 2: Variation of outage with (a) delay and (b) SNR.

The variation of outage probability with SNR for different buffer sizes is given in Figure 2b. Clearly, with increasing SNR, the outage probability reduces for all buffer sizes. The reduction in outage as the buffer length increases from 1 to 3 is substantial.

# 4 Delay Bounded Power Control for Rayleigh Sources

In Section 2.3, it was shown that analogous to time-varying channels, source time-variations also require power control to reduce outage. In Section 2.3 we presented methods which achieved either a delay of one time-slot or arbitrarily large delays. In this section, we present a scheduler to reduce the outage for time-varying sources at the expense of increased delay, analogous to the development in Section 3. To isolate the effects of time-variations of the source, we assume that the channel is time invariant Gaussian.[4] For the following development, we assume that the transmitter has an infinite size buffer.

Our intuition for the following power control policy comes from the zero outage power control $Q_{\infty,t}$, which uses the same power in every time-slot. The infinite delay policy $Q_{\infty,t}$ completely smoothens the variations of the source, and sends at a constant rate. So, policies having finite queuing delay should attempt to smoothen the source variations via appropriate queuing. The proposed policy can be concisely defined as follows:

1. Based on the buffer state $q_t$, the transmitter flushes the buffer if $q_t < r_l$ else

[4] This parallels our development in Section 3 where we had assumed a constant rate source, and time-varying channel.

transmits rate $r_l$. The power used is given by

$$Q_{k,t} = \sigma^2 \left( e^{\min(q_t, r_l)} - 1 \right) \tag{11}$$

2. From the remaining buffer contents, $g_t = q_t - \log\left(1 + \frac{Q_{k,t}}{\sigma^2}\right)$, the excess rate beyond $r_u$ is dropped. Thus, the buffer update equations are

$$\underbrace{d_t}_{\text{dropped}} = g_t - \min(g_t, r_u) \tag{12}$$

$$q_{t+1} = q_t + x_t - d_t - \log\left(1 + \frac{Q_{k,t}}{\sigma^2}\right) \tag{13}$$

The average delay is controlled by the thresholds $r_l$ and $r_u$; larger $r_u$ implies larger delays and lower outage. An illustrative example of the performance of the proposed power control policy is shown in Figure 3. In Figure 3(a), the outage performance is shown as a function of delay for two different channel SNRs. The ergodic rate of the source $R_e$ equals 2 nats/sec/Hz. At high SNR, when the Gaussian channel capacity equals 2.04 nats/sec/Hz, which is greater than $R_e$, the outage decreases monotonically with increasing mean delay. On the other hand, at low SNR when the Gaussian channel capacity equals 1.84 nats/sec/Hz, which is lower than $R_e$, the outage asymptotically reaches $1 - \frac{C_e}{R_e}$. The variation of outage probability with SNR for two different delays (1 and 15 time-slots) is shown in Figure 3(b). Clearly, with increasing SNR the outage probability reduces for all delays.

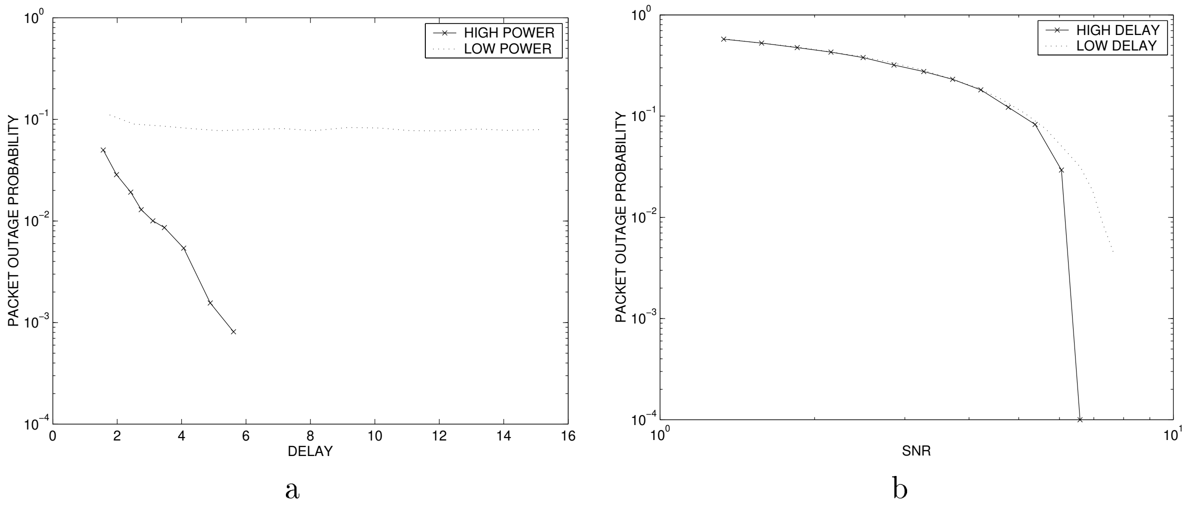
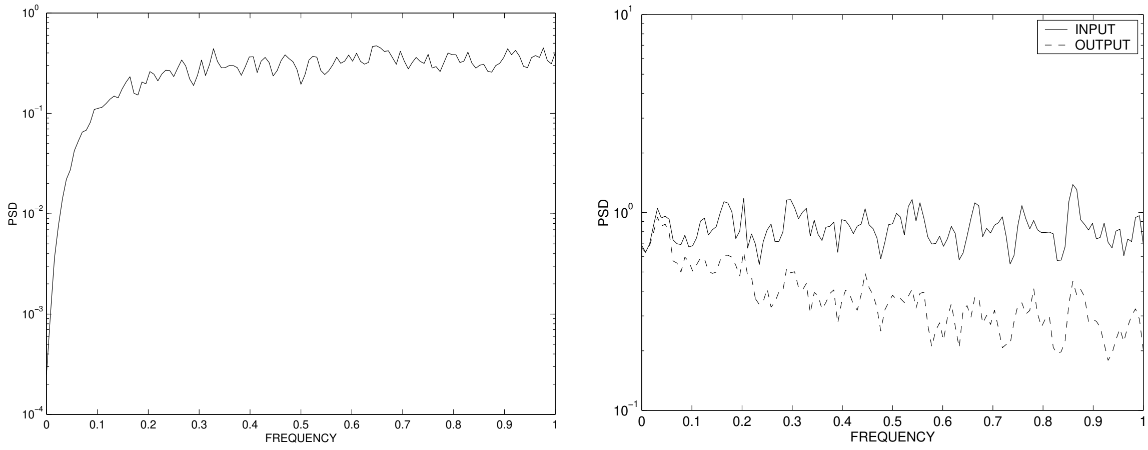



Figure 3: Packet outage probability versus (a) average delay and (b) SNR. Rayleigh source with $R_e = 2$ nats/sec/Hz, AWGN channel.

# 5 Discussion and Conclusions

The proposed power control policies for the fading channels and bursty sources (like the proposed Rayleigh sources) indicate a source-channel duality, in the same spirit as the duality for Gaussian channels and sources. For instance, the power control for fading channels with delay of one time-slot is inversely propotional to instantaneous channel




(a) (b)

Figure 4: Spectrum of the input and output processes using (a) power control for Rayleigh channel, $P_t$ (input has zero spectral content other than at DC) and (b) power control for Rayleigh source, $Q_t$, with average delay of 1.3 time-slots.

SNR, $\gamma_t$ whenever a packet is transmitted. The power control for Rayleigh source is linearly propotional to the instantaneous source variance for the packets which are completely transmitted. Since the packet fragmentation is used in source power control, the two power control policies are different in the constant power transmission regions.

The relationship between the two power control policies is further strengthened by observing the spectrum of the input and output processes in Figure 4. In fading channels, the arrival process is constant rate and hence has frequency content only at zero frequency. Since the power control chooses to wait for some packets and at times, transmit more than one packet, the departure packet process has high frequencies as shown in Figure 4(a). Analogously, the input and output spectrum of the power control policy for Rayleigh source shows that it behaves as a low-pass filter, thereby smoothening the bursty source by queuing.

Via constructive power control policies, we showed the gains due to queuing in fading channels and for bursty sources. The proposed power control policies jointly design the network level operation of scheduling packets and the physical layer operation of power control to achieve the performance gains. The results in this paper will potentially contribute to our understanding of importance of delay in physical layer design.

# References

[1] A. Ephremides and B. Hajek, “Information theory and communication networks: An unconsummated union,” *IEEE Info. Th.*, vol. 44, pp. 2416–2434, Oct. 1998.

[2] R. G. Gallager, “A perspective on multiaccess channels,” *IEEE Trans. Info. Th.*, vol. IT-31, pp. 124–142, Mar. 1985.

[3] I. E. Telatar and R. G. Gallager, “Combining queuing theory with information theory for multiaccess,” *IEEE J. Sel. Ar. Comm.*, vol. 13, pp. 963–969, Aug. 1995.

[4] B. Collins and R. Cruz, “Transmission policies for time varying channels with average delay constraints,” in *Proc. Allerton Int. Conf. on Comm. Control and Comp.*, pp. 709–717, 1999.

[5] D. N. C. Tse, *Variable-rate lossy compression and its effects on communication networks.* PhD thesis, Mass. Inst. Tech., Feb. 1995.

[6] D. Bertsekas and R. Gallager, *Data Networks.* Prentice Hall, 1992.

[7] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Info. Th.*, vol. 43, pp. 359–378, May 1994.

[8] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," tech. rep., AT&T Bell Labs, 1995. (Appeared in *Eur. Trans. Tel.*, vol. 10, no. 6, pp. 585–595, 1999).

[9] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Info. Th.*, vol. 43, pp. 1986–92, Nov. 1997.

# A Proof of Proposition 1

For every power control policy including the optimal policy $Q^*(x_t)$, there exists $U^*(x_t)$ such that $Q^*(x_t) = \sigma^2(e^{U^*(x_t)} - 1)$. Consider two values of rate $r_1$ and $r_2$ and let $r_1 < r_2$ without loss of generality. The proof relies on proving that $U^*(r_1)$ should equal $U^*(r_2)$ if the constraints are satisfied or $U^*(r_1) = r_1$ and is lesser than $U^*(r_2)$. We will prove the following two conditions,

- If $U^*(r_1) = r_1$ then $U^*(r_2) \geq r_1$.
- If $U^*(r_1) < r_1$ then $U^*(r_2) = U^*(r_1)$.

Let $U^*(r_1) = r_1$ and $U^*(r_2) < r_1$. Consider another power control policy

$$u^*(r) = \begin{cases} \log\left(\frac{p_R(r_1)e^{U^*(r_1)} + p_R(r_2)e^{U^*(r_2)}}{p_R(r_1)+p_R(r_2)}\right) & r = r_1, r_2, \\ U^*(r) & \forall r \neq r_1, r_2. \end{cases} \tag{14}$$

with $U^*(r_1) = r_1$ and $U^*(r_2) = r_1 - \delta$ for some $\delta > 0$. It is easy to see that $0 \leq u^*(r_i) \leq r_i$ for $i = 1, 2$. It can also be easily verified that $u^*$ and $U^*$ require the same average power. However $\Pi_Q - \Pi_q = \log\left(\frac{p_R(r_1)e^{r_1} + p_R(r_2)e^{(r_1-\delta)}}{p_R(r_1)+p_R(r_2)}\right)(p_R(r_1) + p_R(r_2)) - (\log(e^{r_1})p_R(r_1) + \log(e^{r_2})p_R(r_2))$ which is strictly positive for $\delta > 0$ due to concavity of the log function. Hence, we find a different policy $u^*$ with same power and lower outage which contradicts the assumption that $U^*$ is optimal. So, If $U^*(r_1) = r_1$ then $U^*(r_2) \geq r_1$.

Now, let $U^*(r_1) < r_1$ and $U^*(r_1) \neq U^*(r_2)$. First, if $U^*(r_1) > U^*(r_2)$, then we can show that there exists policy $u^*$ as specified by (14) having lower outage at the same average power. Second, if $U^*(r_1) < U^*(r_2)$, there are two different possibilities, depending on the sign of $s = \frac{p_R(r_1)e^{U^*(r_1)} + p_R(r_2)e^{U^*(r_2)}}{p_R(r_1)+p_R(r_2)} - r_1$. If $s > 0$, then again we can show that policy $u^*$ as defined in (14) has lower outage at the same power. If $s < 0$, we can find a new policy $\tilde{u}(r)$ as follows,

$$\tilde{u}(r) = \begin{cases} U^*(r) & r \neq r_1, r_2 \\ r_1 & r = r_1 \\ \log\left(e^{U^*(r_2)} + \frac{p_R(r_1)(e^{U^*(r_1)} - e^{r_1})}{p_R(r_2)}\right) & r = r_2 \end{cases} \tag{15}$$

Again, we can show $\tilde{u}$ has lower outage at the same power. Since in all cases we contradict the assumption that $U^*$ is optimal, it follows that $U^*(r_1) = U^*(r_2)$.

Hence we have proved that if $U^*(r_1) = r_1$ then $U^*(r_2) \geq r_1$ and if $U^*(r_1) < r_1$ then $U^*(r_2) = U^*(r_1)$ and consequently (7) follows immediately. □