Anytime Capacity of the AWGN+Erasure Channel with Feedback

by

Qing Xu

B.S. (Beijing University) 1997
M.S. (University of California at Berkeley) 2000

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Master of Science

in

Engineering-Electrical Engineering and Computer Science

in the

GRADUATE DIVISION
of the
UNIVERSITY of CALIFORNIA at BERKELEY

Committee in charge:

Professor Anant Sahai, Chair
Professor Raja Sengupta

Fall 2005

The dissertation of Qing Xu is approved:

Chair	Date

Date

University of California at Berkeley

Fall 2005

Anytime Capacity of the AWGN+Erasure Channel with Feedback

by

Qing Xu

Abstract

Anytime Capacity of the AWGN+Erasure Channel with Feedback

by

Qing Xu

Master of Science in Engineering-Electrical Engineering and Computer Science

University of California at Berkeley

Professor Anant Sahai, Chair

We study the feedback anytime reliability of a discrete-time channel with additive white Gaussian noise where the channel output is also subject to i.i.d. erasures. The encoder has noiseless access to the past channel outputs, which includes perfect information regarding which transmissions were erased. There is an average power constraint on the channel input. The channel is an idealized model for a multiple access system where the channel noise is modelled as an additive white Gaussian process and collisions between packets are modelled as erasures. The anytime reliability of a channel is an important property in the design of delay sensitive application such as feedback control over the channel.

As an intermediate step we first study the anytime reliability of packet erasure channels where the erasures are i.i.d. and known to both the transmitter and receiver. The packets are allowed to carry a variable number of bits within them although only one packet may be transmitted at any given time step. These results provide insight into the analysis of AWGN+erasure channel, and are of importance in their own right. For any moment constraint on the size of the packets, we show that the anytime reliability is constant at all rates up to the Shannon capacity of the channel and that this constant is essentially the logarithm of the probability of erasure. For cases where there is both a first moment constraint and a constraint on the maximum number of bits that a packet can carry, we show that the optimum anytime reliability is determined primarily by the peak-packet size, but drops abruptly to zero above the Shannon capacity. We then proceed to show that the feedback anytime reliability of the AWGN+erasure channel is constant at all rates up to the Shannon capacity of the channel and that this constant is essentially the logarithm of the probability of erasure. In order to show achievability, we give a construction consisting of a FIFO queue where the server can adjust its service rate based on the number of bits awaiting transmission. In the AWGN+erasure case the schemes involve a hybrid control system, where the server of the queue takes the data bits and uses them to drive a scalar linear control system with continuous state where the dynamics can switch between fast and slow based on the service rate.

Professor Anant Sahai
Dissertation Committee Chair

Contents

List of Figures

Chapter 1

Introduction

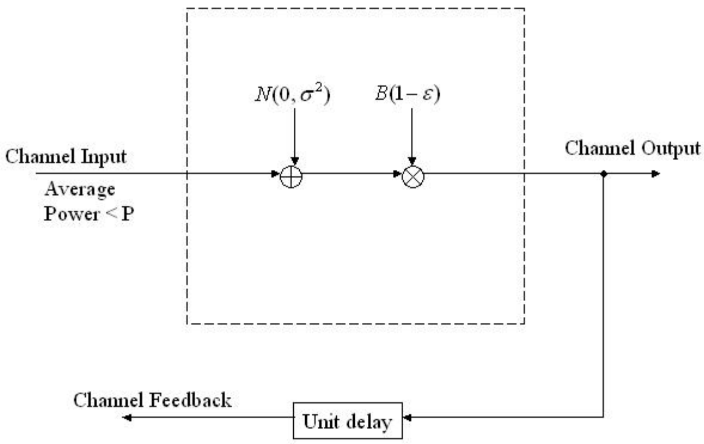

Figure 1.1: The AWGN+erasure channel with feedback

In this thesis we study the anytime reliability of a discrete-time power-constrained AWGN + erasure channel with noiseless feedback. The channel is shown in Figure 1.1. First a white Gaussian noise $N(0, \sigma^2)$ is added to the real-valued channel input, then this output is either erased with probability ε or conveyed to the channel output with no additional error. The erasures across time steps are independent. The average input power of the channel is constrained to be lower than P. We allow the encoder to have access to noiseless feedback of channel output with delay of one unit of discrete

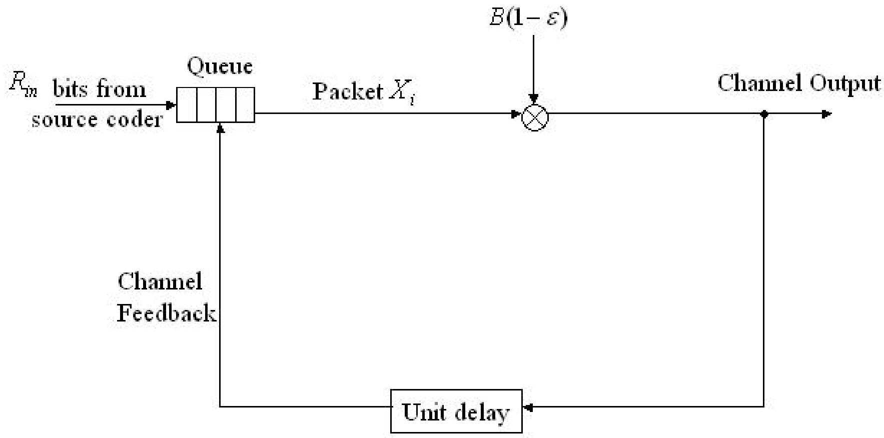

Figure 1.2: Packet erasure channel with feedback, fed by a queue

time to avoid any causality issues.

We study this problem for the following two reasons:

1. The AWGN+erasure channel is an idealized model for a wide range of channels in wireless communication systems.

2. The anytime reliability property of a channel is critical in the design of delay sensitive applications over the channel.

The discrete-time power constrained additive white Gaussian noise channel is one of the most useful idealizations in communication theory. It models communication over a band-limited wireless channel. Real wireless channels are not so simple. They can be subjected to fading, wherein the transmitted signal is attenuated more substantially at some times, possibly due to the interference from other transmitters. One of the simplest models of fading has independent fades from time to time, where the fade is either 1 or 0. To further simplify the model, we will assume knowledge of the fade at the receiver. This model can be thought of as the AWGN+erasure channel in that we can consider the output to be "erased" whenever the fade is 0.

Packet erasures emerge naturally in both wired and wireless communication systems where packet-oriented designs abound, often from an implicit desire to achieve statistical multiplexing among many different users. Erasures generally model two types of events: an unfortunate noise sequence that the underlying error correcting code could not correct or "collisions" at either an intermediate node in a network (where it leads to a packet drop) or over the shared communication medium [14][3] (where it leads to an undecodeable reception of the packet). Generally, interference is reduced by Medium Access Control (MAC) protocols (e.g. the 802.11 MAC protocol [1]), which controls the transmitters so that they are not likely to transmit at the same time. However, in our simplified model we can view a denial by the MAC layer as an erasure since the packet was unable to be sent at the requested time. Knowledge of the erasure can come back to the transmitter through either an acknowledgment packet or by the transmitter observing the packet getting mangled over the link.

For the simplified AWGN+erasure model, rather than considering the erasures as side information known at the receiver, we further simplify the model by viewing the erasures as occurring after the additive noise. Since the channel noise in the model is continuous, the fade is immediately apparent whenever the output is 0.

Since AWGN+erasure channel is a useful abstraction of the environment faced by wireless communications, it is natural to consider using such a channel within control systems. The topic of "control over communication channel" or equivalently the design of "networked control systems" has become a hot topic in both control and communication communities. The high level prospective overview of the problem can be found in [7], and some of the most recent progresses appear in a special issue of IEEE Transactions on Automatic Control. [2] In such a integrated communication and control system, [8], [9], and [10] tell us that we require enough anytime capacity to be able to meet our performance objectives. Essentially, to hold the η-moment of the state of an unstable plant finite, it is necessary and sufficient for the feedback channel's anytime capacity evaluated corresponding to anytime-reliability $\alpha = \eta \log_2 \lambda$ to be greater than $\log_2 \lambda$ where λ is the unstable eigenvalue of the plant. As such, understanding the anytime capacity under various models is essential for us to evaluate engineering questions like whether or not we can have a control system share the communication channel with an existing packet system or whether we need to have a reserved channel for just the control application. Alternatively, sometimes it is possible to schedule the transmissions from nodes in a multiple access network such that interfering nodes share the channel and transmit in turn — e.g. in the wireless token ring protocol [6] — but unless we can evaluate the anytime capacities, it is hard to evaluate whether this additional complexity is justified.

The anytime capacity of a channel relates the bit error with the delay in a system where we require every bit to get through eventually. We review the definition of anytime capacity: [9]

Definition 1.1. $C_{anytime}(\alpha)$, the α-anytime capacity, is the maximum rate at which the channel can be used to communicate with a bit error probability that drops with delay exponentially at a rate of α.

$$C_{anytime}(\alpha) = sup\{R|\exists\mathcal{E}^{\mathcal{R}}, K > 0, \forall N, \exists\mathcal{D}^{\mathcal{R}}_{\mathcal{N}}, P_{error}(\mathcal{E}^{\mathcal{R}}, \mathcal{D}^{\mathcal{R}}_{\mathcal{N}}) < K2^{-\alpha N}\} \qquad (1.1)$$

In above definition $\mathcal{E}^{\mathcal{R}}$ is the anytime encoder, $\mathcal{D}^{\mathcal{R}}_{\mathcal{N}}$ is the decoder, and N is the delay that a bit experiences in units of channel uses. The parameter α is called the *anytime reliability.*

Fundamentally, what we have is a region of achievable (α, R) pairs — the region between the α axis and the anytime capacity curve. Whether we choose to look at maximizing R as a function of α or as maximizing α as a function of R is a matter of convenience for the problem at hand. Once we know the anytime capacity we know the anytime achievable region, and vice versa.

In comparison let's look at the definition of the classical Shannon capacity.

Definition 1.2. The Shannon classical capacity is the maximal rate at which the channel can be used to transmit the data with an arbitrarily small probability of bit error.

$C = sup\{R|\forall\epsilon\exists N, \mathcal{E}^{\mathcal{R}}_{\mathcal{N}}, \mathcal{D}^{\mathcal{R}}_{\mathcal{N}}, P_{error}(\mathcal{E}^{\mathcal{R}}_{\mathcal{N}}, \mathcal{D}^{\mathcal{R}}_{\mathcal{N}}) < \epsilon\}$

Since the probability of error on every single bit goes to zero with increasing delay, it is clear that the anytime capacity is always less than or equal to the classical Shannon capacity.

The rest of the thesis is structured as follows. In the remaining part of this chapter, we review the relevant previous work in 1.1, and summarize the main result of the thesis in 1.2. In Chapter 2 we analyze the feedback anytime reliability of packet erasure channels with different constraints on the packet size. In Chapter 3 we analyze the AWGN+erasure channel. We conclude the thesis in Chapter 4.

1.1 Previous work

The concept of anytime capacity was first introduced by Sahai in his PhD thesis. [9] The anytime capacity of the AWGN channel and binary erasure channel are separately obtained in Chapter 7 of [9]. Both these results assume noiseless channel feedback. In Chapter 6 of the thesis the lower bound of the anytime capacity of AWGN and binary erasure channels without feedback are derived. But so far there is no exact solution of the anytime capacity of these two channels without feedback. We list relevant results of [9] below for completeness.

Theorem 1.3. *For the binary erasure channel with encoders having access to noiseless feedback*

$$C_{anytime}(\alpha) = \frac{\alpha}{\alpha + \log_2\left(\frac{1-\varepsilon}{1-2^{\alpha}\varepsilon}\right)} \tag{1.2}$$

where ε is the erasure probability.[1]

Theorem 1.4. *For the AWGN channel with power constraint P, noise variance σ^2 and encoders having access to noiseless feedback*

$$C_{anytime}(\alpha) = \frac{1}{2}\log_2(1 + \frac{P}{\sigma^2})$$

for any $\alpha > 0$.

Further work on anytime capacity is presented by Sahai and Şimşek in [4] and [11]. They study the anytime channel coding problem for discrete-time channel with noiseless feedback. An upper bound for the anytime exponent is derived, and a family of "time sharing" codes to approach this bound is proposed. It is shown that for binary erasure channel, the bound is achieved by (1.2). They also show that the anytime capacity of the binary symmetric channel is increased with feedback.

In recent years, there are also many works on the control over wireless communication problem. Seiler and Sengupta study the feedback control of an LTI system with the sensor and controller connected by a real erasure channel, i.e. in each time step the feedback controller either receives the system output with no error, or does

[1] The form for the anytime capacity of binary erasure channel comes from solving explicitly the parametric form given in [9].

not receive anything. They derive the existence condition for a linear feedback controller which stabilizes the system in the mean square sense, and the relevant $\mathcal{H}_\infty$ results. [12] [13] They do not consider channel noise and the coding/decoding problem of the channel. Elia studies the control of LTI system when the control command is passed through fading channel to the plant. [5] Like in [13], the objective of controller design is also mean square stability. They concentrate on the controller design and do not consider the coding/decoding of the channel. The fading channel models studied include an FIR channel with Gaussian coefficients and an erasure channel with delay.

To show achievability of the optimum anytime reliability, we use longer packets when the queue is long and shorter packets when the queue is short. Similar in spirit, Tse et.al. study the statistical multiplexing of multiple time-scale Markov streams. [15] They model each stream with a singularly perturbed Markov-modulated process with some state transitions occurring less frequently than others. They estimate the buffer overflow probabilities in various asymptotic regimes in the buffer size, rare transition probabilities, and the number of streams.

1.2 Main results of the thesis

The main results of this thesis are summarized in the following theorems. Theorem 1.5 is the main results for the anytime reliability of the packet erasure channel with constraints on various moment of packet size. Theorem 1.6 is the main result when both the average and maximum size of the packet are constrained. The main result for AWGN+erasure channel is summarized in Theorem 1.7.

Theorem 1.5. *Let X_i be the packet at time i, and $S(X_i)$ be the size of the packet. When the j-th moment of packet size is constrained $E\{[S(X_i)]^j\} \leq \bar{S}_j$, the anytime capacity and reliability of the packet erasure channel depends on the value of j as follows:*

1. $j \geq 1$

   The anytime capacity of the packet erasure channel is

$$C_{anytime}(\alpha) = \begin{cases} (1-\varepsilon)\sqrt[j]{\bar{S}_j} & \textit{if } 0 \leq \alpha \leq -\log_2 \varepsilon \\ 0 & \textit{otherwise} \end{cases} \tag{1.3}$$

Or when viewed in terms of anytime reliability:

$$\alpha^*(R) = \begin{cases} -\log_2 \varepsilon & \textit{if } R < (1-\varepsilon)\sqrt[j]{\bar{S}_j} \\ 0 & \textit{otherwise} \end{cases}$$

2. $j < 1$

The anytime capacity of the packet erasure channel is

$$C_{anytime}(\alpha) = \begin{cases} \infty & \textit{if } 0 \leq \alpha \leq -\log_2 \varepsilon \\ 0 & \textit{otherwise} \end{cases} \tag{1.4}$$

or equivalently for anytime reliability:

$$\alpha^*(R) = -\log_2 \varepsilon \tag{1.5}$$

Theorem 1.6. *Let X_i be the packet at time i, and $S(X_i)$ be the size of the packet. The anytime capacity of the packet erasure channel, when both the average packet size and the peak packet size are constrained, i.e., $E\{S(X_i)\} \leq \bar{S}$ and $\max(S(X_i)) < S_{max}$, is the following:*

$C_{anytime}(\alpha)$

$$= \begin{cases} \min\left((1-\varepsilon)\bar{S}, \frac{\alpha}{\alpha + \log_2\left(\frac{1-\varepsilon}{1-2^{\alpha}\varepsilon}\right)} S_{max}\right) & \textit{if } 0 \leq \alpha \leq -\log_2 \varepsilon \\ 0 & \textit{otherwise} \end{cases} \tag{1.6}$$

Expressed in terms of anytime reliability,

$$\alpha^* = \eta - \log_2(1 + \varepsilon(2^\eta - 1))$$

whenever $R = S_{max}\left(1 - \frac{1}{\eta}\log_2(1 + \varepsilon(2^\eta - 1))\right) < (1-\varepsilon)\bar{S}$, and 0 otherwise.

Theorem 1.7. *The anytime capacity of AWGN + erasure channel with average power constraint P, erasure probability ϵ noise variance σ^2, and encoder having access to noiseless feedback, is*

$$C_{anytime}(\alpha) = \begin{cases} \frac{(1-\varepsilon)}{2} \log_2 \left(1 + \frac{P}{\sigma^2}\right) & if \;\; 0 \leq \alpha < -\log_2 \varepsilon \\ 0 & otherwise \end{cases} \tag{1.7}$$

$$\alpha^*(R) = \begin{cases} -\log_2 \varepsilon & if \;\; 0 < R < \frac{1-\varepsilon}{2} \log_2 \left(1 + \frac{P}{\sigma_2}\right) \\ 0 & otherwise \end{cases} \tag{1.8}$$

As shown in Figure 1.3, basically the α-anytime capacity of the AWGN+erasure channel is the same as the Shannon capacity of this channel for all $0 < \alpha < -\log_2 \varepsilon$, and zero otherwise.

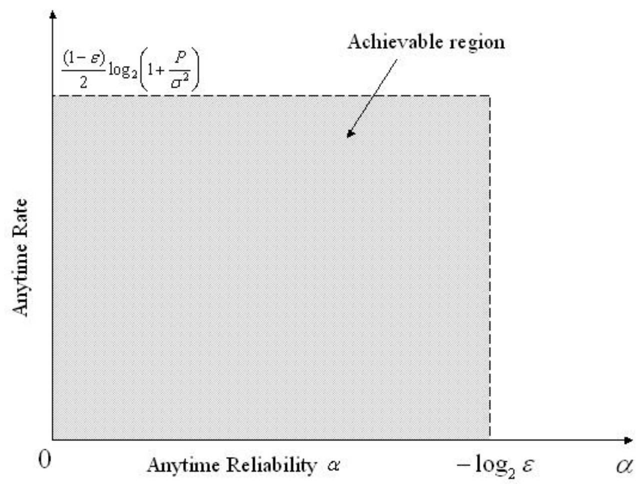

Figure 1.3: The anytime capacity of the AWGN+erasure channel

We compare the anytime capacity of the AWGN channel with feedback, binary erasure channel with feedback, and AWGN+erasure channel with feedback in Figure 1.4. Clearly, erasures reduce the anytime capacity. The reduction is proportional to the erasure rate. The erasures also limit the anytime reliability to be within $-\log_2 \varepsilon$. The seemingly surprising result obtained in this thesis is that when the AWGN is conjoined with erasures, the anytime capacity of the channel does not decay with anytime reliability α, as it does for the binary erasure channel. For comparison we also plot the anytime capacity of an AWGN channel *scheduled* to transmit only 80%

of time (e.g. with the wireless token ring protocol [6]). We can see that the anytime capacity is the same as an AWGM followed by erasures with probability 0.2. But the erasures limit the anytime reliability, whereas scheduling does not. Hence scheduling provides better anytime reliability performance than the uncontrolled probabilistic erasures at the cost of more complex implementation.

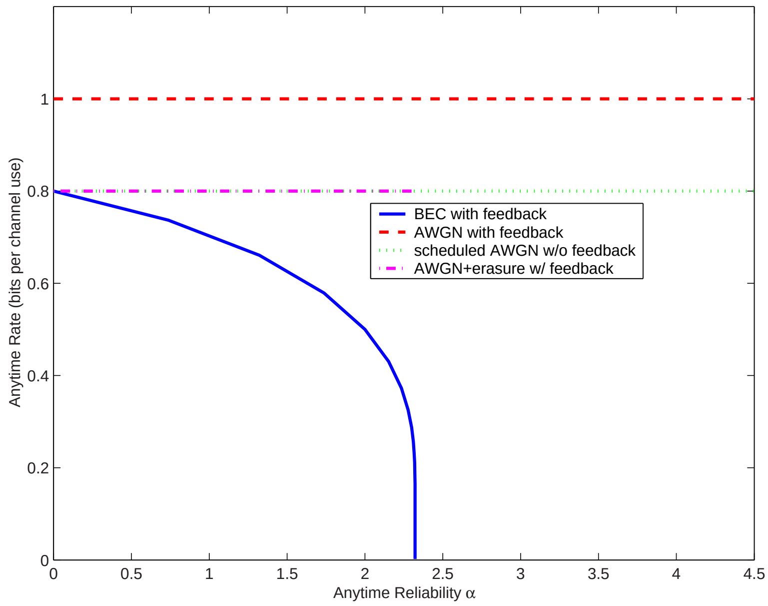

Figure 1.4: Comparison of the capacities: $P = \sqrt{3}$, $\sigma = 1$, and $\varepsilon = 0.2$

Chapter 2

Analysis of the Packet Erasure Channel

In this chapter, we study the anytime reliability of packet erasure channels. We hope to gain insight in the study the AWGN+erasure channel from this intermediate step. At the same time the packet erasure channel is an important model itself as has been discussed in the introduction.

We analyze the anytime capacity of the packet erasure channel with various constraints on the number of bits encoded in a packet, while keeping the transmission of each packet to be one discrete time step. Variable size packets can occur in cases where the transmitter has a choice of modulation to use in transmitting the packet. For example, as shown in Figure 2.1, we can encode 1, 2, or 4 bits in one channel symbol by using different constellations. The case where the packet is simply made longer in time is not covered by the analysis given in this thesis since we assume that the time between allowed packet transmissions is always constant.

The cases of unconstrained packets and fixed-size packets were discussed in [9]. We first review the results for the case of unconstrained packet sizes in 2.1. There, it is obvious that the optimal transmission strategy is just to transmit all the bits awaiting transmission. Next, we review the case of fixed-size packets in 2.2. For this, the optimal strategy is a first-in first-out (FIFO) queue with bits being removed from the queue when a packet is successfully received. With these results in hand, we consider what happens if we impose a moment constraint on the packet size. We consider a general (possibly fractional) moment but show that there is something fundamentally different about moments higher than 1 and those lower than 1. To show achievability of the optimum anytime reliability, we use longer packets when the queue is long. Then in 2.4 we consider what happens when we have a peak packet-size constraint in addition to a first moment constraint. We summarize the chapter in 2.5.

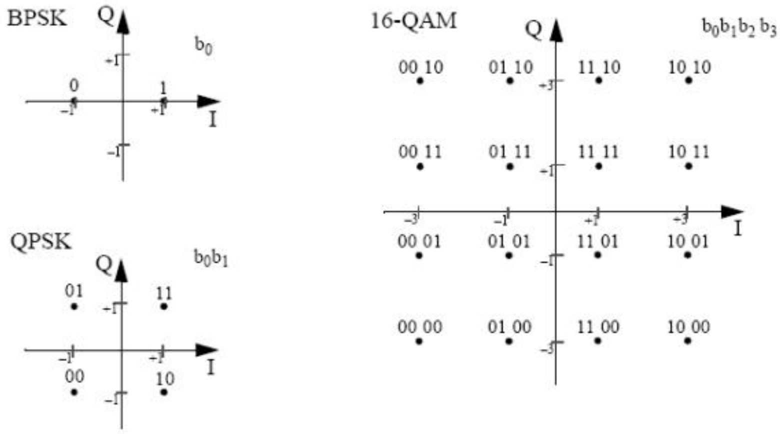

Figure 2.1: BPSK, QPSK, and 16-QAM constellation bit encoding, each carrying 1, 2, or 4 bits per symbol respectively

In all of our analyses, we use a FIFO queue to achieve the anytime capacity. Bits enter the queue at a steady rate R_{in}. Whenever there is a non-erasure in the channel in time step i, a packet X_i is transmitted with no error. By feedback, the queue knows whether an erasure happened in the previous time step. The size of the packet X_i is denoted as $S(X_i)$.

2.1 Unconstrained packets

In this section we analyze the anytime capacity of packet erasure channels with no constraint on the packet size. We assume that the channel noise is zero or the transmit power is unbounded so we can encode arbitrarily large numbers of bits in one packet. When the packet is received, the bit error is negligible.

When there is no constraint on the packet size, the anytime capacity is easy to obtain. At every time step, the channel attempts to transmit all the bits in the queue. This channel is discussed as the real erasure channel in [9] and it is clear that the encoding strategy is optimal.

Theorem 2.1. *The anytime capacity of a packet erasure channel when the packet size is unconstrained is*

$$C_{anytime}(\alpha) = \begin{cases} \infty & \text{if } 0 \leq \alpha \leq -\log_2 \varepsilon \\ 0 & \text{otherwise} \end{cases} \tag{2.1}$$

or expressed in terms of the anytime reliability, we have for all rates R:

$$\alpha^*(R) = -\log_2 \varepsilon \tag{2.2}$$

Proof. There is a bit error only when there is a sequence of consecutive erasures from time t through $t+d$, and the bit error probability is $\varepsilon^d = 2^{d\log_2 \varepsilon}$. Therefore the anytime reliability α must be less than or equal to $-\log_2 \varepsilon$, but the incoming rate can be as high as we would like. □

Figure 2.2 shows the anytime achievable region of the packet erasure channel when the packet size is not constrained. Figure 2.3 depicts the Markov chain of the queue when the packet size is not constrained. Any non-erasure clears the queue. We simulate the packet size as a function of time. The result is shown in Figure 2.4.

2.2 Fixed packet size

When $S(X_i)$ is fixed to be R_{out}, the channel is an R_{out}-bit erasure channel with feedback. The feedback anytime capacity of the binary erasure channel was derived in [8] and [9]. By a simple change of units argument, it is clear that the anytime capacity of the R_{out}-bit channel is the anytime capacity of the binary symmetric channel (1 bit) scaled by R_{out}.

Theorem 2.2. *The anytime capacity of a packet erasure channel with packet size constrained to be R_{out} bits, and encoder having access to noiseless feedback, is given*

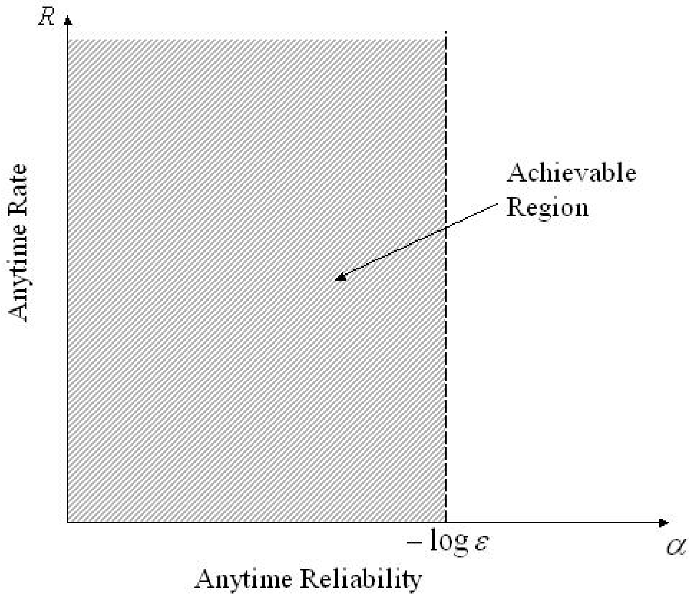

Figure 2.2: The anytime achievable region when the packet size is unconstrained

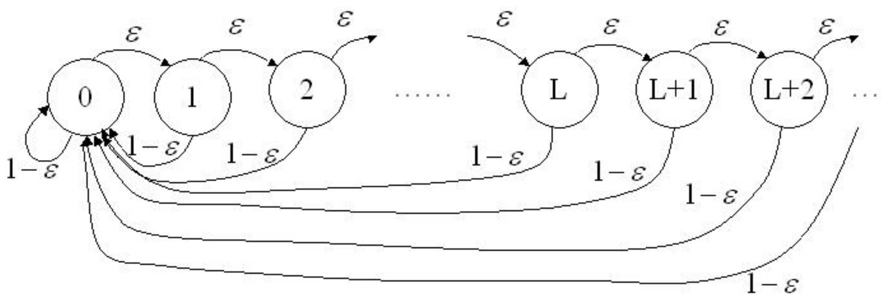

Figure 2.3: Markov chain of queue length when the packet size is unconstrained: R_{in}=1 bit

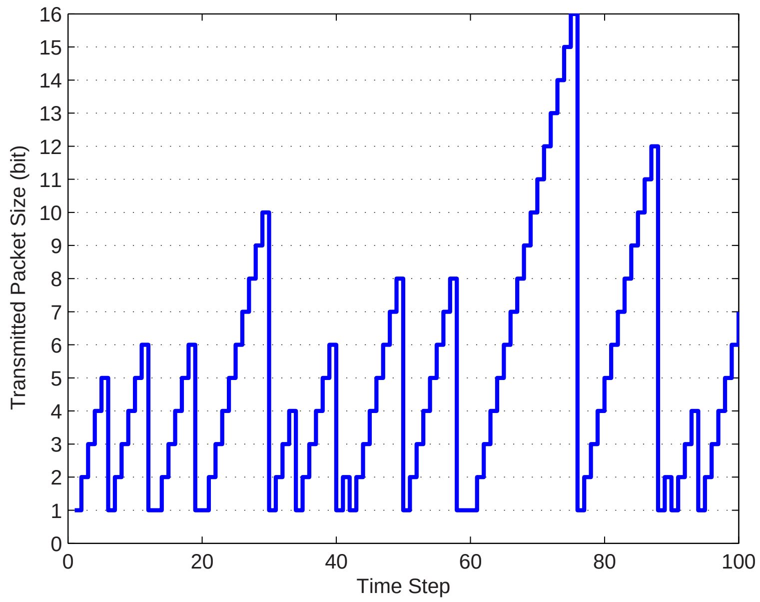

Figure 2.4: Simulation result of packet size: unconstrained packets, $R_{in} = 1$ bit, $\varepsilon = 0.8$

by:

$$C_{anytime}(\alpha) = R_{out} \frac{\alpha}{\alpha + \log_2\left(\frac{1-\varepsilon}{1-2^{\alpha}\varepsilon}\right)} \tag{2.3}$$

where ε *is the erasure probability. Alternatively, the anytime reliability* $\alpha^* = \eta - \log_2(1 + \varepsilon(2^\eta - 1))$ *corresponds to the rate* $R_{out}\left(1 - \frac{1}{\eta}\log_2(1 + \varepsilon(2^\eta - 1))\right)$*, where* η *is a positive parameter.*

The anytime achievable region of the fixed-size packet erasure channel given by equation (2.3) is shown in Figure 2.5,

Although the optimality proof for (2.3) involved a control system [9], it should be clear that the optimal anytime encoder is just a FIFO queue with a server that tries to make R_{out} sized packets and send them out. See Figure 2.6 for an example of the resulting Markov chain.

2.3 Constraint on the j-th moment of packet size

The unconstrained packet size case in section 2.1 imposes no constraint while in many applications, the packet-size constraint of section 2.2 is too rigid and does

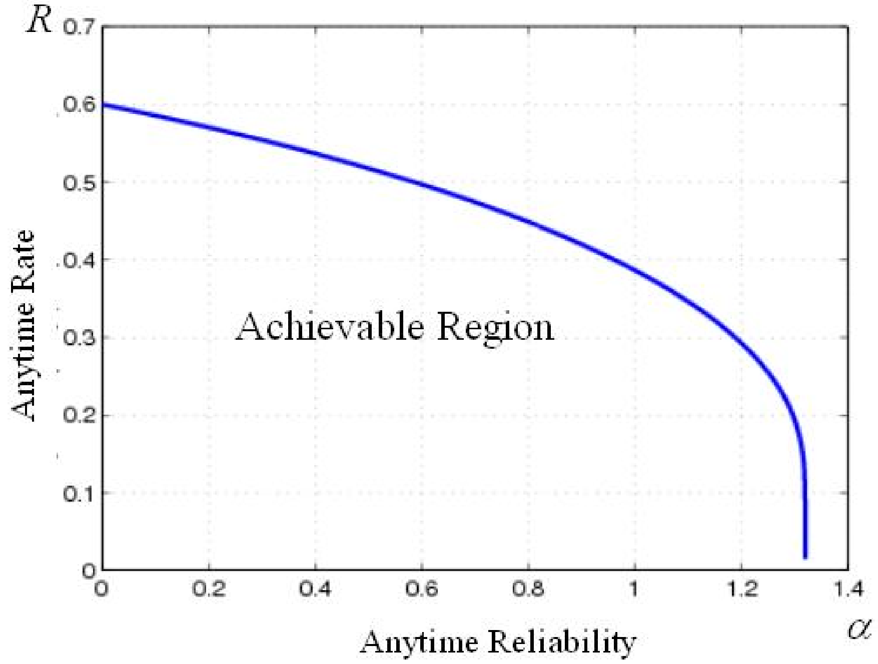

Figure 2.5: The anytime achievable region when packet size is fixed: $R_{out} = 1$ bit, $\varepsilon = 0.4$

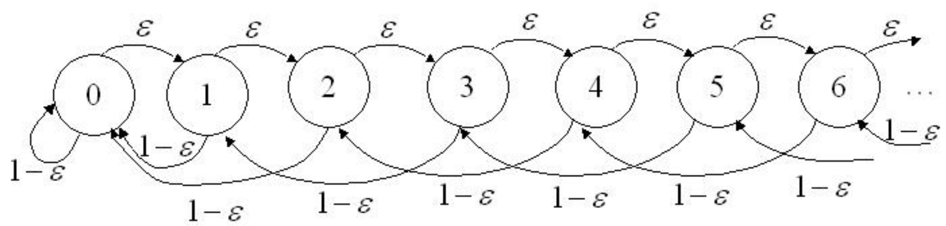

Figure 2.6: Markov chain of queue length when the packet size is fixed: R_{in}=1 bit, Packet Size = 3 bits

not adequately capture the flexibility that might exist. One way of constraining the communication in a flexible way is to impose an average constraint on the j-th moment of the packet size. This ensures that large packets are unlikely and by increasing j, we increase the relative cost of larger packets as compared to smaller ones. $j = 1$ corresponds to the most natural constraint on the average packet size.

To prove achievability, we will use the following type of policy that changes the packet sizes at a certain critical queue length L_c.

- When the queue length l is smaller than or equal to a critical length L_c, the system attempts to transmit $R_{out} = R_1$ bits in each packet.
- When the queue length l is larger than L_c, the system transmits a giant packet to reduce the queue length back down to L_c upon successful reception. Note we are using potentially unboundedly large packets.

To study the anytime capacity of this system we need the following lemma.

Lemma 2.3. *Let the input rate be R_{in} and the system have this queuing rule:*

- Short queue mode*: When the queue length is smaller than L_c, the system transmits packet of size $R_1 > \frac{R_{in}}{(1-\varepsilon)}$*
- Long queue mode*: When the queue length is larger than L_c, the system transmits packets of size larger than R_1.*

Then the probability of having a queue of length $l > d$ bits is upper-bounded by

$$P(l > d) \leq T_1 2^{-\alpha_1^* \left\lfloor \frac{d}{R_{in}} \right\rfloor}$$

where α_1^* is the feedback anytime reliability of the R_1-sized packet erasure channel corresponding to a rate of R_{in} from Theorem 2.2 and T_1 is some positive constant.

Proof. The key is to realize that for all possible realizations of erasures, the queue length can only be shorter than the queue length if we used packets of fixed size R_1. So we can bound the probability of a large queue length by the corresponding probability for the fixed-size packet system. Whenever an erasure happens, the symbol is resent until it is successfully received. The only possible bit error event is that the bit is still in the queue and has not been sent. The queue must therefore contain at least d bits and the anytime reliability of this channel bounds the tail distribution of the queue. How to calculate α_1^* graphically from the binary erasure channel anytime curve is illustrated in Figure 2.7. The constant T_1 is playing the role of the constant K in (1.1). □

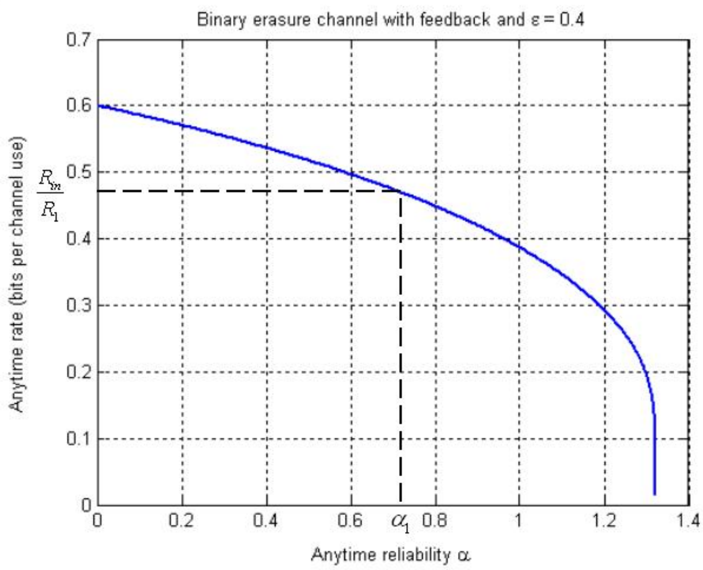

Figure 2.7: Finding α_1^* for given R_1 and R_{in} using the anytime capacity curve of binary erasure channel

With Lemma 2.3, we can prove the following theorems for the anytime capacity of the constrained packet erasure channel. It turns out that the anytime capacity for the case when $j \geq 1$ and $j < 1$ are quite different. We discuss them separately.

Theorem 2.4. *Let X_i be the packet at time i, and $S(X_i)$ be the size of the packet. When the j-th ($j \geq 1$) moment of packet size is constrained $E\{[S(X_i)]^j\} \leq \bar{S}_j$, the anytime capacity of the packet erasure channel is*

$$C_{anytime}(\alpha) = \begin{cases} (1-\varepsilon)\sqrt[j]{\bar{S}_j} & \textit{if } 0 \leq \alpha \leq -\log_2 \varepsilon \\ 0 & \textit{otherwise} \end{cases} \tag{2.4}$$

Or when viewed in terms of anytime reliability:

$$\alpha^*(R) = \begin{cases} -\log_2 \varepsilon & \textit{if } R < (1-\varepsilon)\sqrt[j]{\bar{S}_j} \\ 0 & \textit{otherwise} \end{cases}$$

Proof. Clearly we can only do worse for this channel than the unconstrained packet-erasure channel and hence the upper bound on α is an immediate corollary to Theorem 2.1. Moreover, the anytime capacity can never exceed the Shannon capacity and $(1-\varepsilon)\sqrt[j]{\bar{S}_j}$ is clearly the Shannon capacity of the constrained channel since $\sqrt[j]{\bar{S}_j}$ is the induced constraint on the average number of bits per packet that comes from the moment constraint and the factor of $(1-\varepsilon)$ represents the independent probability that a packet gets through.

For achievability, we use the following queuing rule illustrated in Figure 2.8.

- When the queue length l, in bits, is smaller than or equal to L_c bits, the system takes out $R_{out} = R_1$ bits for every non-erasure. We require $R_1 = \sqrt[j]{\bar{S}_j} - \epsilon_1$, with ϵ_1 being an arbitrary small positive number. Notice to make the system stable we need $R_{in} < (1-\varepsilon)R_1$.

- When the queue length l is larger than L_c, the system transmits a giant packet to reduce the queue length to L_c.

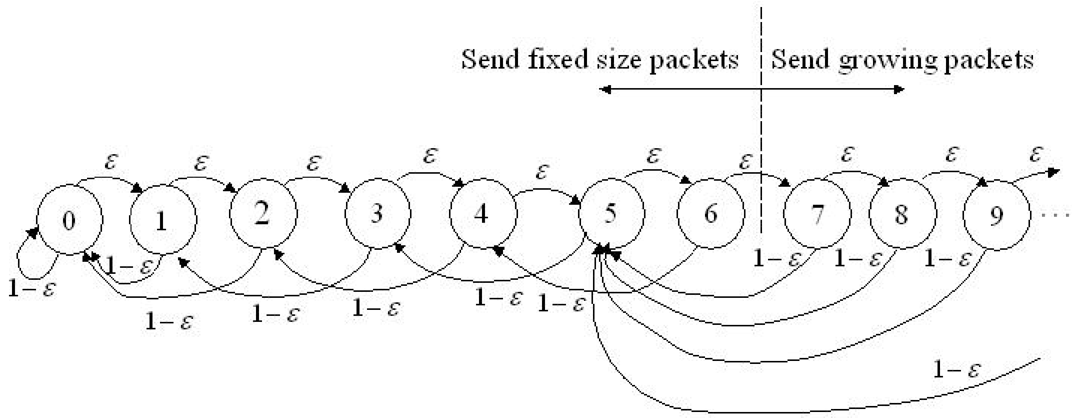

Figure 2.8: Markov chain of queue length when $E\{S^j\} < \bar{S}_j$ and $j \geq 1$: L_c=6, $R_{in} = 1$ bit and $R_{out} = 3$ bits when $l < L_c$

We now bound the probability of error. A bit error implies the bit is not transmitted. If the delay is larger than $\frac{L_c}{R_1}$ and the bit is still not transmitted, the queue must be longer than L_c bits. In such situations, any non-erasure will reduce the queue length to L_c bits and so there must have been a string of consecutive erasures since the instant the queue length grew longer than L_c bits. Thus $P_{error} < \varepsilon^{d-L_c} = \varepsilon^{-L_c} 2^{-d\log_2 \varepsilon}$ as delay d goes large.

We only need to show that the constraint on the j-th moment of queue length is met. From Lemma 2.3 we know that the queue length distribution satisfies $P(l > L_c) \le T_1 2^{-\alpha_1^* \lfloor \frac{L_c}{R_{in}} \rfloor}$, with T_1 and α_1 being positive constants. Hence we have:

$$\begin{aligned} E\{S^j\} &= \sum_{i=0}^{\infty} R_{out}^j(i) P(l = i) \\ &= P(l \le L_c) R_1^j + P(l > L_c) \sum_{i=L_c+1}^{\infty} (i - L_c)^j \cdot P(l = i | l > L_c) \\ &\le R_1^j + T_1 2^{-\alpha_1 \lfloor \frac{L_c}{R_{in}} \rfloor} \sum_{i=L_c+1}^{\infty} (i - L_c)^j \cdot \varepsilon^{i-L_c} \\ &= R_1^j + T_1 2^{-\alpha_1 \lfloor \frac{L_c}{R_{in}} \rfloor} \sum_{i=1}^{\infty} i^j \varepsilon^i \end{aligned} \tag{2.5}$$

Since $\sum_{i=1}^{\infty} i^j \varepsilon^i$ converges and is independent of L_c, the second term is an exponentially decreasing function of L_c. Furthermore, R_1^j is selected to be smaller than $\bar{S}_j$. Therefore we can always find L_c large enough such that the constraint on the j-th moment of the packet size is met. □

In effect, having a constraint on the high moments of the packet size just imposes a bound on the anytime capacity, not the anytime reliability.

Figure 2.9 shows the anytime achievable region of the packet erasure channel when the j-moment ($j \ge 1$) of the packet size is constrained.

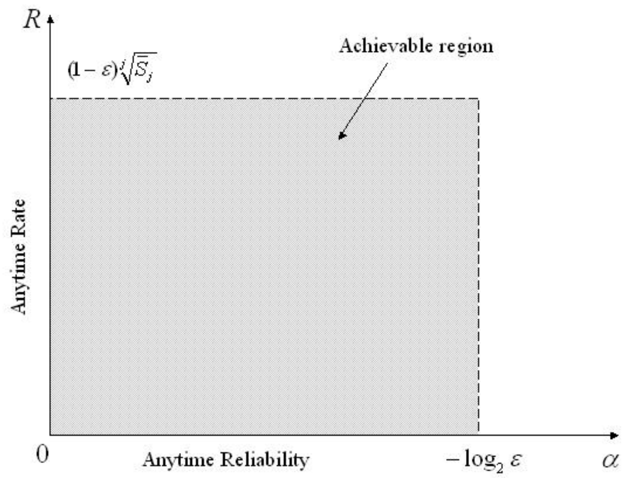

Figure 2.9: The anytime achievable region when the $j-$moment of packet size is constrained ($j \ge 1$)

Figure 2.10 shows the simulated packet size changing with time with different erasure realizations.

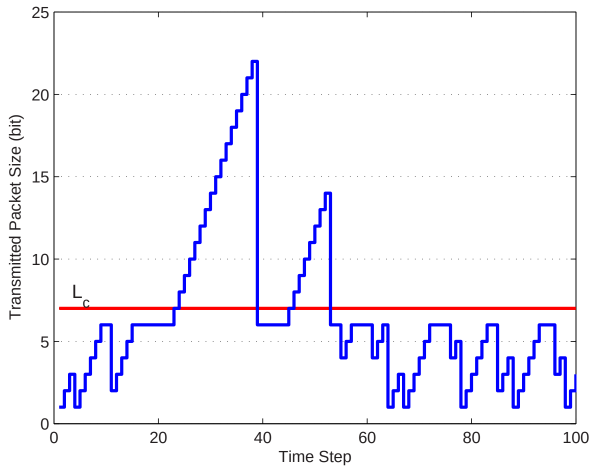

Figure 2.10: Simulation result of packet size: constraint on the $j-$th moment of packet size ($j \geq 1$), $R_{in} = 1$ bit, $L_c = 7$, Packet Size = 6 bits when $l \leq L_c$, $\varepsilon = 0.8$

Theorem 2.5. *Let X_i be the packet at time i, and $S(X_i)$ be the size of the packet. When the j-th ($j < 1$) moment of packet size is constrained by $E\{[S(X_i)]^j\} \leq \bar{S}_j$, the anytime capacity of the packet erasure channel is*

$$C_{anytime}(\alpha) = \begin{cases} \infty & \text{if } 0 \leq \alpha \leq -\log_2 \varepsilon \\ 0 & \text{otherwise} \end{cases} \tag{2.6}$$

or equivalently for anytime reliability:

$$\alpha^*(R) = -\log_2 \varepsilon \tag{2.7}$$

Proof. We now change the policy to be as illustrated in Figure 2.11:

- When the queue is shorter than a critical length L_c bits, no bits are sent: the packet size is zero.
- When the queue is longer than L_c bits, empty the queue with every non-erasure.

As before, a bit error can happen when bit is held in the queue. When the delay is d (large enough), the probability of bit error is the probability of the queue being longer than d, which is upper-bounded as before by $P(l > d) < \varepsilon^{d-L_c}$.

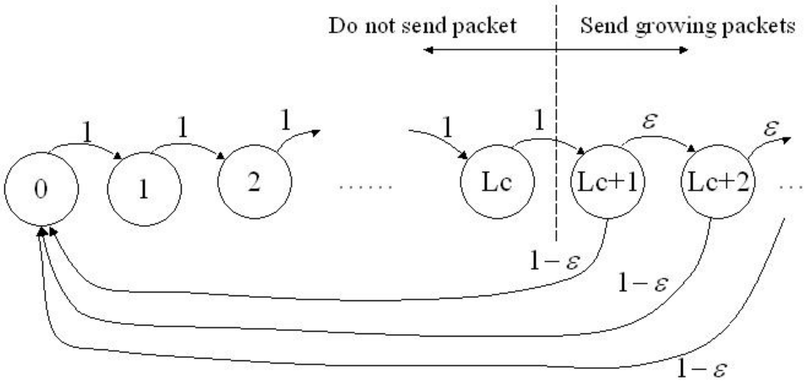

Figure 2.11: Markov chain of queue length when $E\{S^j\} < \bar{S}_j$ and $j < 1$, labels in units of R_{in}

To verify the queue length constraint, we look at stationary distribution of the Markov chain in Figure 2.11. Obviously, $\pi_0 = \pi_1 = \cdots = \pi_{L_c}$, and $\pi_0 = (1-\varepsilon)\sum_{i=1}^{\infty}\pi_{L_c+i}$. Therefore $\pi_0 = \pi_1 = \cdots = \pi_{L_c} = \frac{1}{L_c+1+(1-\varepsilon)} < \frac{1}{L_c}$, and $P(l > L_c) = \sum_{i=1}^{\infty}\pi_{L_c+i} = \frac{\pi_0}{(1-\varepsilon)} < \frac{1}{L_c(1-\varepsilon)}$.

Hence we have

$$E\{S^j\} < \frac{1}{L_c(1-\varepsilon)}\sum_{i=1}^{\infty}(L_c+i)^j\varepsilon^i \tag{2.8}$$

As is illustrated in Figure 2.12, due to the sub-linearity of the function $f(x) = x^j$ when $j < 1$ and $x > 0$, we have

$$\begin{aligned} f(x) &< L_c^j + \left(\frac{df(x)}{dx}\right)_{x=L_c} \cdot (x - L_c) \\ &= L_c^j + jL_c^{j-1}(x - L_c) \end{aligned}$$

Therefore continuing from equation (2.8) we have

$$\begin{aligned} E\{S^j\} &< \frac{1}{L_c(1-\varepsilon)}\sum_{i=1}^{\infty}(L_c^j + jL_c^{j-1}i)\varepsilon^i \\ &= \frac{L_c^j\varepsilon}{L_c(1-\varepsilon)^2} + \frac{jL_c^{j-1}}{L_c(1-\varepsilon)^3} \\ &= \frac{\varepsilon}{L_c^{1-j}(1-\varepsilon)^2} + \frac{j}{L_c^{2-j}(1-\varepsilon)^3} \end{aligned} \tag{2.9}$$

Since $j < 1$, both terms in (2.9) decay to zero with L_c. Thus for every R_{in}, we can choose a large enough L_c to satisfy the constraint on the j-th moment of the packet size. □

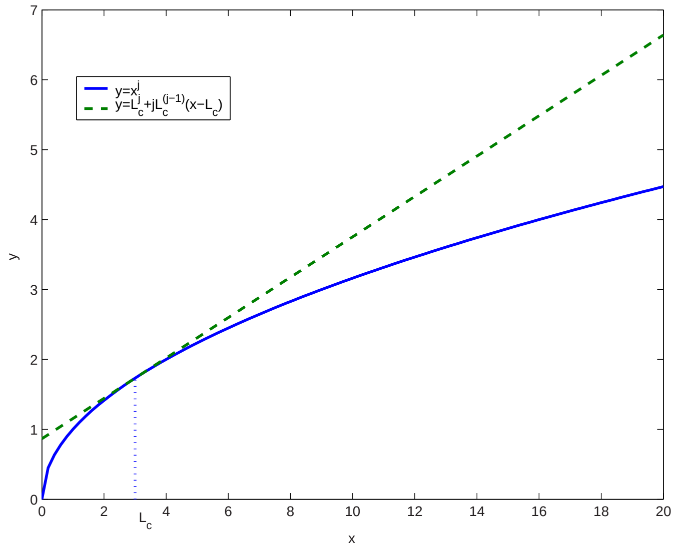

Figure 2.12: Sub-linearity of function $y = x^j$ when $j = 0.5$, $L_c = 3$

So, having a constraint on a moment less than the first moment is as good as having no constraint at all, at least when it comes to the anytime reliability!

Figure 2.13 shows the anytime achievable region of the packet erasure channel when the $j-$moment $(j < 1)$ of the packet size is constrained.

The simulation results for the packet size when the $j-$moment of packet size is constrained with $j < 1$ is shown in Figure 2.14.

2.4 Constraints on both average and peak packet sizes

Usually, it will be unrealistic to allow unboundedly large packet sizes. Seeing as how having a higher moment constraint seems to only effect the anytime capacity and reliability through the induced constraint on the first moment, it is natural to consider the case where not only is $E\{S\} \leq \bar{S}$, but $S < S_{max}$ as well where $S_{max} > \bar{S}$ to avoid triviality. To study the anytime capacity we need the following lemma which refines Lemma 2.3 further:

Lemma 2.6. *Let the input rate be R_{in} and the system have this queuing rule:*

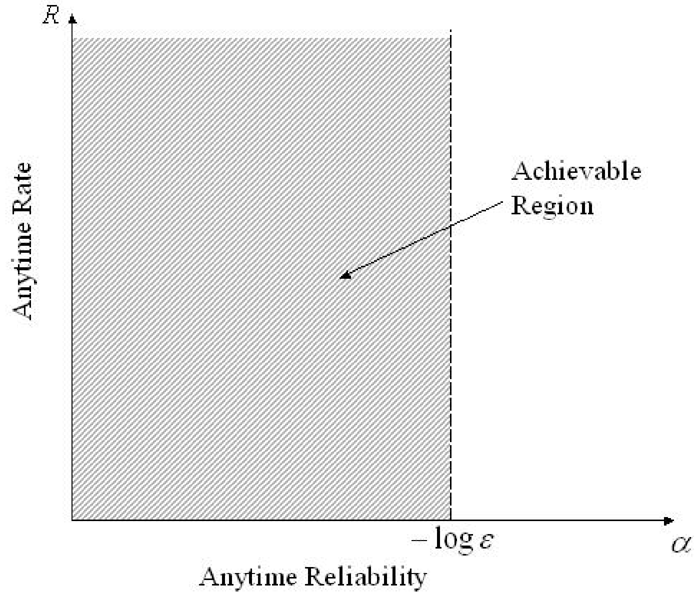

Figure 2.13: The anytime achievable region when the $j-$moment of packet size is constrained ($j < 1$)

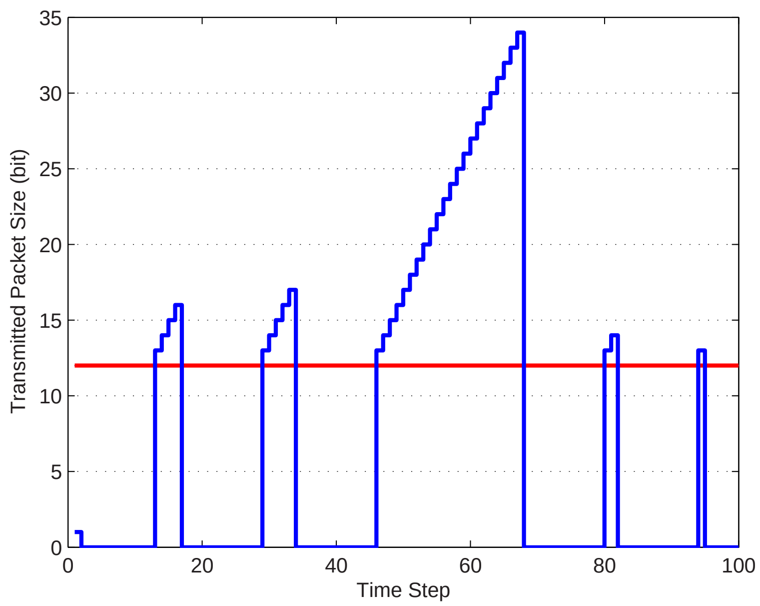

Figure 2.14: Simulation result of packet size: constraint on the $j-$th moment of packet size ($j < 1$), $R_{in} = 1$ bit, $L_c = 12$ bits, $\varepsilon = 0.8$

- Short queue mode*: When the queue length is not larger than L_c, the system transmits packet of fixed size $R_1 > \frac{R_{in}}{(1-\varepsilon)}$ bits.*

- Long queue mode*: When the queue length is larger than L_c, the system transmits packet of the fixed size $R_2 > R_1$ bits.*

Then when queue length $l > L_c$ bits, the probability of having a queue of length $l > d$ is upper-bounded as

$$P\left(l > d\right) \leq T_2' 2^{-\alpha_2^* \left\lfloor \frac{d}{R_{in}} \right\rfloor}$$

where α_2^* is the feedback anytime reliability corresponding to a rate of R_{in} of the R_2-size packet erasure channel from Theorem 2.2 and T_2' is some positive constant.

Proof. This is similar to the proof of Lemma 2.3. Since $R_2 > R_1$, we have $\alpha_2^* > \alpha_1^*$. When $l > L_c$, we consider the part of queue that is over L_c. In this part corresponding to R_2 we have anytime exponent α_2^*. Let $i = l - L_c$, we can bound $P(i \geq d - L_c)$ by $T_2 2^{-\alpha_2^* \left\lfloor \frac{d-L_c}{R_{in}} \right\rfloor}$.

$$\begin{aligned}
P(l > d) &= P(l > L_c)P(l - L_c > d - L_c | l > L_c) \\
&\leq T_1 2^{-\alpha_1^* \left\lfloor \frac{L_c}{R_{in}} \right\rfloor} T_2 2^{-\alpha_2^* \left\lfloor \frac{d-L_c}{R_{in}} \right\rfloor} \\
&\leq T_1 2^{-\alpha_1^* \left(\frac{L_c}{R_{in}} + 1 \right)} T_2 2^{-\alpha_2^* \left(\frac{d-L_c}{R_{in}} + 1 \right)} \\
&= 2^{\alpha_1^* + \alpha_2^*} T_1 T_2 2^{(\alpha_2^* - \alpha_1^*) \frac{L_c}{R_{in}}} 2^{-\alpha_2^* \frac{d}{R_{in}}} \\
&\leq T_2' 2^{-\alpha_2^* \left\lfloor \frac{d}{R_{in}} \right\rfloor}
\end{aligned}$$

since we can bound the queue length of this queuing system by L_c plus the length of one where only R_2 sized packets are used. □

Using Lemmas 2.3 and 2.6 we can prove the following theorem on the anytime capacity of the packet erasure channel when both the average packet size and the peak packet size are constrained.

Theorem 2.7. *Let X_i be the packet at time i, and $S(X_i)$ be the size of the packet. The anytime capacity of the packet erasure channel, when both the average packet size and the peak packet size are constrained, i.e., $E\{S(X_i)\} \leq \bar{S}$ and $\max(S(X_i)) < S_{max}$, is the following:*

$$C_{anytime}(\alpha) = \begin{cases} \min\left((1-\varepsilon)\bar{S}, \frac{\alpha}{\alpha+\log_2\left(\frac{1-\varepsilon}{1-2^{\alpha}\varepsilon}\right)} S_{max}\right) & \text{if } 0 \le \alpha \le -\log_2 \varepsilon \\ 0 & \text{otherwise} \end{cases} \quad (2.10)$$

Expressed in terms of anytime reliability,

$$\alpha^* = \eta - \log_2(1 + \varepsilon(2^\eta - 1))$$

whenever $R = S_{max}\left(1 - \frac{1}{\eta}\log_2(1+\varepsilon(2^\eta-1))\right) < (1-\varepsilon)\bar{S}$, *and 0 otherwise.*

The anytime capacity presented in equation (2.10) is illustrated in Figure 2.15.

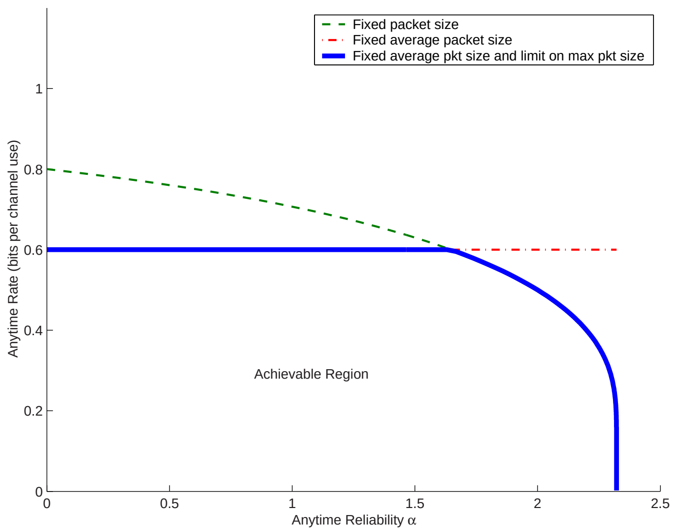

Figure 2.15: Anytime achievable region of packet erasure channel with both average and peak packet size constraints

Proof. The anytime region of this system should be contained in the regions with only one of the constraints: peak or average. Therefore the region lies in their intersection. We achieve every (α, R) point in this intersection by using:

- When the queue length l is smaller than or equal to L_c, the system uses packets of size $R_{out} = R_1$ bits. We require $R_1 = \bar{S} - \epsilon_1$, with ϵ_1 being an arbitrary small positive number. Notice to make the system stable we need $R_{in} < (1-\varepsilon)R_1$.

- When the queue length l is larger than L_c, the system transmits packets with size S_{max}, until the queue is shorter than L_c again.

This queuing rule is illustrated in Figure 2.16.

Similar to the proof of Theorem 2.4, a bit error implies the bit is still awaiting transmission. If the delay is larger than $\left\lfloor \frac{L_c}{R_{in}} \right\rfloor$ and the bit is still not transmitted, the queue must be longer than L_c. From Lemma 2.6 we know that the queue distribution, and therefore the probability of bit error, is bounded by $T_2' 2^{-\alpha_2^*(R_{in})d}$, where $\alpha_2^*(R_{in})$ is the feedback anytime reliability for the S_{max}-size erasure channel corresponding to anytime rate R_{in}. Since we can make ϵ_1 arbitrarily small, we have the anytime reliability as stated in the theorem. To check the average packet size constraint:

$$\begin{aligned} E\{S\} &= \sum_{i=0}^{\infty} R_{out}(i) P(l = i) \\ &= R_1 P(l \leq L_c) + S_{max} P(l > L_c) \\ &\leq R_1 + T_2' S_{max} 2^{-\alpha_2^* \left\lfloor \frac{L_c}{R_{in}} \right\rfloor} \end{aligned} \tag{2.11}$$

Since $R_1 < \bar{S}$ and the second term is an exponential function of L_c, we can always select large enough L_c such that $E\{S\} \leq \bar{S}$. □

The simulated packet size of this case is shown in Figure 2.17.

2.5 Summary

By using variable sized packets, we can substantially increase the anytime reliability of a packet erasure channel. This can be achieved by using bigger packets when we have many bits awaiting transmission. When the constraint on the packet size takes the form of a moment constraint, then the anytime reliability is effectively the same as that of the unconstrained channel for all rates up to capacity. When there is a peak-size constraint as well, then the peak-size constraint dominates the anytime reliability at all rates up to the Shannon capacity of the channel.

In the next chapter, we extend this style of analysis to the AWGN channel with erasures and noiseless feedback. There the doubly-exponential vanishing of the probabilities of error with delay lets us effectively conceptualize the channel as a noiseless packet erasure channel, though there are many technical steps along the way. One of the differences in the two system comes from the power constraint of the AWGN channel. In the analysis in section 2.3 we increase the packet size to arbitrarily large when the queue is long, and we change the critical queue length L_c to meet the expected packet size constraint. At first it seems the AWGN+erasure channel is similar to this case, with the average power constraint analogous to the j-th moment of packet size constraint. However the average power constraint is essentially a constraint on $E\{e^{S(X_i)}\}$, which is more demanding than the packet size j-th moment constraint $E\{S^j(X_i)\}$. Hence the above techniques are not directly applicable.

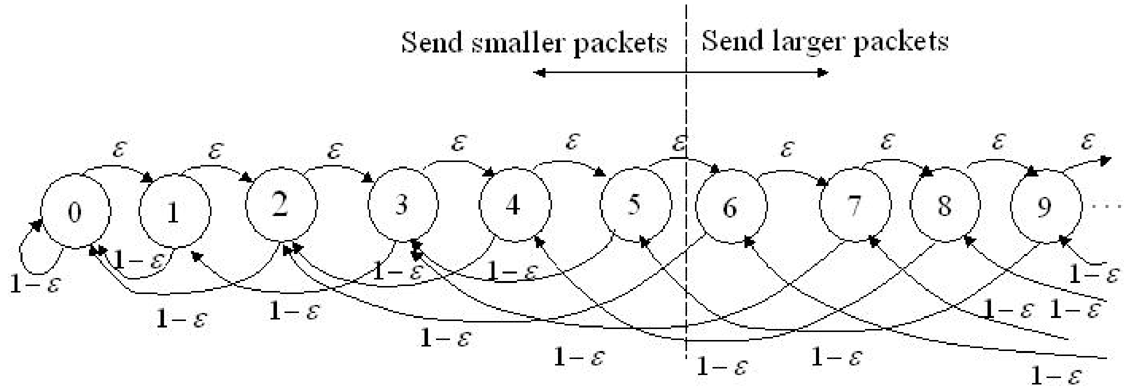

Figure 2.16: An example of Markov chain of queue length when both average packet size and peak packet size are constrained: $L_c = 5$, $R_{in} = 1$, $R_{out} = 3$ when $l \leq L_c$, and $R_{out} = S_{max} = 4$ when $l > L_c$

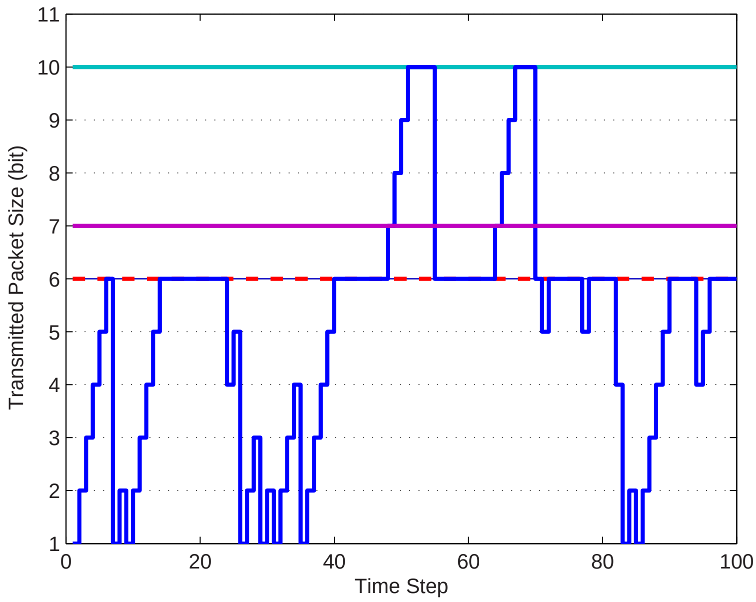

Figure 2.17: Simulation result of packet size: constraints on both the peak and average packet size, $R_{in} = 1$ bit, packet size $= 6$ bits when $l \leq L_c$,peak packet size $= 10$ bits, $L_c = 7$ bits, $\varepsilon = 0.8$

Chapter 3

Analysis of the AWGN+Erasure Channel

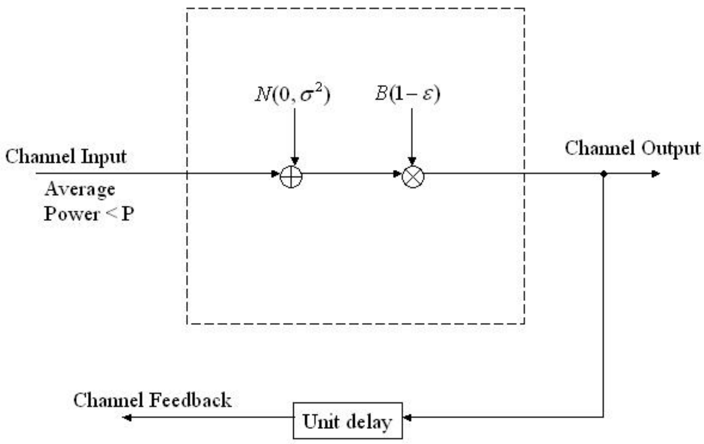

Figure 3.1: The AWGN+erasure channel with feedback

In this chapter we study the anytime reliability of a discrete-time power-constrained AWGN + erasure channel with noiseless feedback. The channel was shown in Figure 1.1, which we repeat in Figure 3.1 for presentation clarity. First a white Gaussian noise $N(0, \sigma^2)$ is added to the real-valued channel input, then this output is either erased with probability ε or conveyed to the channel output with no additional error. The erasures across time steps are independent. The average input power of the channel is constrained to be lower than P. We allow the encoder to have access to noiseless feedback of channel output with delay of one unit of discrete time to avoid any causality issues. The main result was summarized in Theorem 1.7 in Section 1.2. Essentially the anytime capacity of the AWGN + erasure channel is the same as its Shannon capacity, up to the fundamental limit on the anytime reliability posed by the erasures.

To prove achivability, we combine the approach used in Chapter 2 with that taken in Chapter 6 of [9]. The overall idea is to embed the data bits to be transmitted into the disturbance driving a hybrid control system. The continuous state of the control system will represent the uncertainty that the decoder has regarding the data bits and will be used to generate the channel inputs. Because of the perfect feedback, the encoder can keep track of this uncertainty at its end. The bits failed to be transmitted are queued at the encoder. Similar as in Chapter 2, the discrete state of the hybrid control system is the queue length compared to a critical length to be designed. The channel outputs will be used by the decoder to construct its best estimate of the bits. This will be accomplished by running a copy of the same hybrid system that the encoder is running, only without access to the disturbance. Our design must ensure the average transmission power constraint is not violated.

We first describe the detailed approaches we undertake to attack the problem in 3.1. Then in 3.2 we present in detail the construction of the system we use to analyze the anytime capacity and reliability. The analysis of the system is in 3.3. And finally in 3.4 we prove the achievability of the anytime region shown in Figure 1.3.

3.1 Approaches

3.1.1 Use feedback control system to code

We use the same technique as in Chapter 7 of [9]. A simulated feedback control system is used as the encoder, where the feedback is the last channel output. The decoder of the channel is the re-scaled uncontrolled simulated system. We use different encoders and decoders in the long-queue and short-queue mode. We design the system

Table 3.1: Glossary

R_{in} = Constant input bit rate from source

P = Average power constraint of the system

ε = Erasure probability

n_t = AWGN noise

y_t = System output and channel input at simulation time t

z_t = Channel output at simulation time t

x_t = State of the linear system at simulation time t

a_1 = Linear system dynamics in the short queue mode

C_1 = Observer in the short queue mode

D_1 = Estimator in the short queue mode

a_2 = Linear system dynamics in the long queue mode

C_2 = Observer in the long queue mode

n = "speeding up factor" for the encoder to take in the bits when queue is long

l_t = Queue length at simulation time t

L_c = The critical queue length dividing long queue region and short queue region

R_1 = Transmission bit rate when $l_t \leq L_c$

R_2 = Transmission bit rate when $l_t > L_c$

P_1 = Target power when $l_t \leq L_c$

P_2 = Target power when $l_t > L_c$

such that the encoders and decoders in the two modes produce the same mean square error. The simulated system is advanced only when the channel output is not erased, and is paused otherwise.

3.1.2 Real time, virtual time, and simulation time

One advantage of using simulated systems as encoder and decoder is the flexibility to advance or pause the system. In the system we distinguish three time indices.

- Real time t_r

  The real time starts from 0 and always advances in step of size one, regardless of the channel realization. In each real time step, R_{in} bits are generated by the source. At real time t_r there are $R_{in}t_r$ bits that have been generated by the source so far.

- Simulation time t

  Simulation time is the index of the state of the simulated feedback control systems in encoder and decoder. It advances by one if and only if the channel output is not erased. If no erasure has happened since the start of time, the simulation time and the real time are the same.

- Virtual time t_v

  The virtual time indicates the position of the last bit received and decoded by the receiver. It is defined by $\left\lceil \frac{\text{Number of bits received}}{R_{in}} \right\rceil$. It starts from 0, and advances by $\left\lceil \frac{R_{out_i}}{R_{in}} \right\rceil$ with every reception not erased, where $i \in \{1, 2\}$ and R_{out_i} is the time-varying output rate depending on the queue length. Without erasure the virtual time and real time should be the same. But erasures cause bits being stored in the queue and the lag between the virtual time and real time. This lag is the indication of how many bits are delayed due to erasures.

Hence the system works as follows. In each new real-time step, the encoder takes out R_{out_i} bits from the queue, where $i \in \{1, 2\}$ depends on the queue length. The control system in the encoder “tentatively” evolves, with the simulation time “tentatively” advancing by one. If the channel output is not erased, the tentative advance of simulation time is “confirmed” in the control systems in both the encoder and the decoder, where the encoder uses the channel feedback in the next real time step. The virtual time advances by $\lceil \frac{R_{out_i}}{R_{in}} \rceil$ in this case. If on the other hand, the channel output is erased, the simulation time will not advance, and the simulated control systems in both encoder and decoder are “paused”. The virtual time will not advance, with its lag to real time increasing by R_{in}. The detail of the system construction is given in the next section.

The goal of our system design is to reduce the lag between the virtual time and the real time caused by the erasures and maintain small probability of error caused by the white Gaussian noise, under the power constraint. The reader should notice the contrast between our virtual/simulated control system and the idea of just using a standard control system whose state evolves in real time like those in networked control systems. [5] [13] We are using this hybrid queing/control system for ease of analysis since we can use queuing arguments for the queue part and control arguments for the continuous part.

3.2 System construction

To achieve arbitrary points within the achievability region shown in Figure 1.3, we use the approach illustrated in Figure 3.2. Bits arrive into a FIFO queue. If the queue is short (i.e. no bits have waiting for too long), then a certain number of bits corresponding to a rate $R_{out_1} = R_1$ are tentatively taken out of the queue. If the queue is very long (i.e. many old bits are still awaiting a chance at transmission), then a larger number of bits $R_{out_2} = R_2 = nR_1$ are tentatively taken out of the queue. This adjustment of service rate is shown in Figure 3.3. The bits that are tentatively taken out of the queue are used to drive a special simulated source connected to a joint source/channel encoder. If the feedback comes back and shows that the transmission was erased, the bits tentatively taken out are put back into the head of the queue and the evolution of the simulated source is backed out as though discrete time had not advanced at all. The reason that $R_2 = nR_1$ is so we can think of the long-queue behavior as attempting to make the simulated time run n times faster.

Below in all the equations the time indices t and $t + 1$ stand for the simulation time. Thus in real time, $t+1$ and t are two neighboring instants when the channel output is not erased.

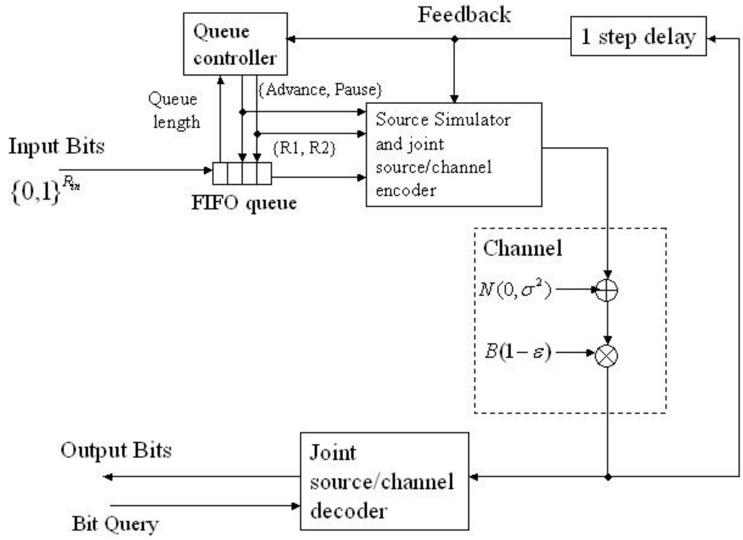

Figure 3.2: Encoding and decoding system overview

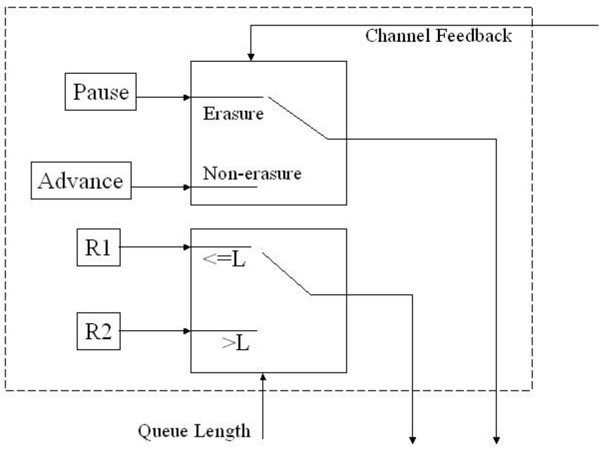

Figure 3.3: Queue controller: adjusting service rate based on queue length

3.2.1 Encoder

The state of the encoding system consists of two parts: The discrete valued state of the queue and the continuous valued state within the simulated source. The goal of the continuous valued state is to provide appropriate continuous valued inputs into the channel and taken together, to try and realize the packet-level abstraction shown in Figure 3.4. The analysis of that abstraction alone has been given in Chapter 2.

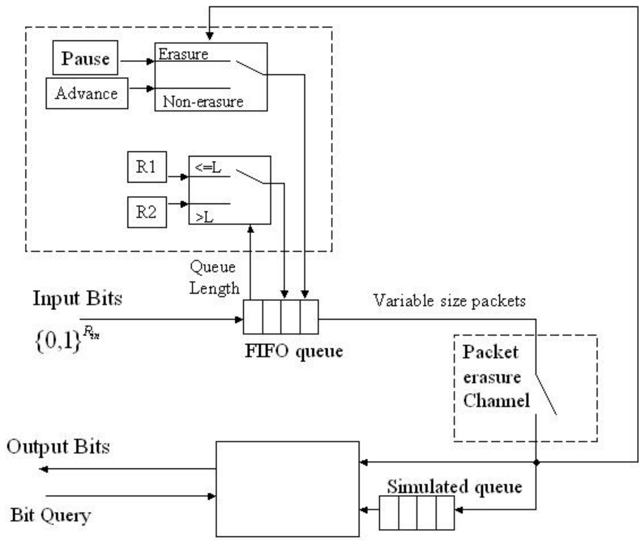

Figure 3.4: Queue level abstraction: bits are removed from queue when a non-erasure happens and n times larger packets are used when queue is long

Figure 3.5 shows the inside of the part that deal with the continuous valued state, and Figures 3.6 and 3.7 show the open-loop dynamics and actions of the controllers respectively. The open-loop dynamics are unstable with $a_2 = a_1^n$ to effectively let it take n time-steps all at once. The controllers apply the control designed to stabilize the system over the noisy feedback link when time should advance. When there is an erasure, the state is left unchanged.[1]

The following are the governing equations of each component.

When $l_t \leq L_c$: $\mathcal{E}_1$

- Linear system $\mathcal{L}_1$

$$
\begin{aligned}
x_{t+1} &= 2^{R_1+\epsilon_1'}(x_t + U_t) + W_{t+1}^{R_1} \\
&= (2+\epsilon_1)^{R_1}(x_t + U_t) + W_{t+1}^{R_1} \\
&= a_1(x_t + U_t) + W_{t+1}^{R_1}
\end{aligned}
$$

[1]In contrast to the "control over communication channel" systems studied in [13] and [5], here we can pause and advance the evolution of the (simulated) control system based on channel feedback.

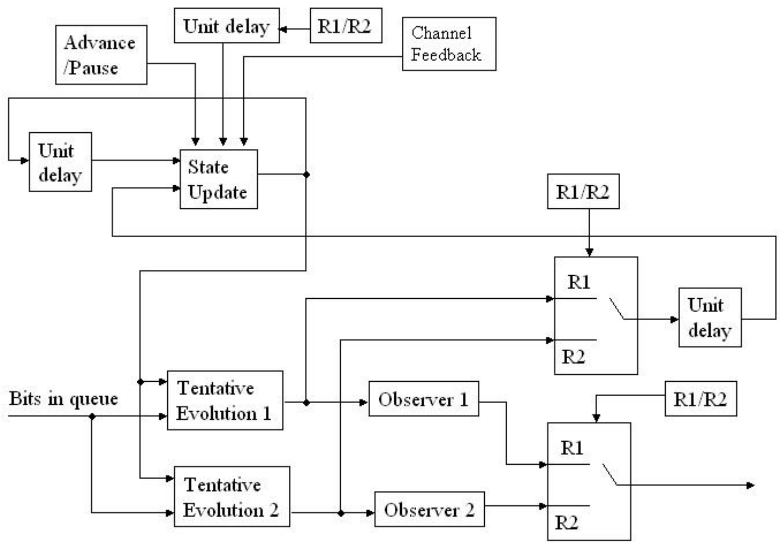

Figure 3.5: Source simulator and joint source/channel encoder

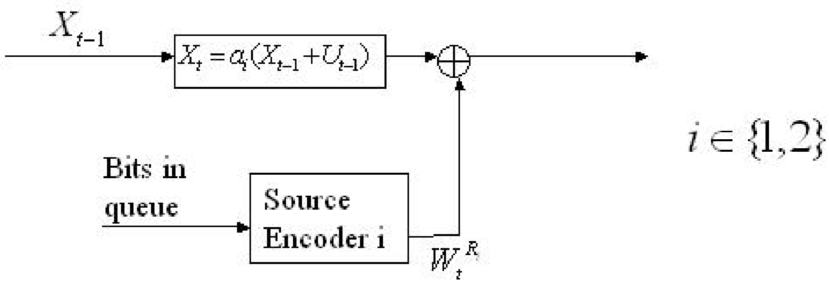

Figure 3.6: Tentative evolution: unstable open-loop dynamics

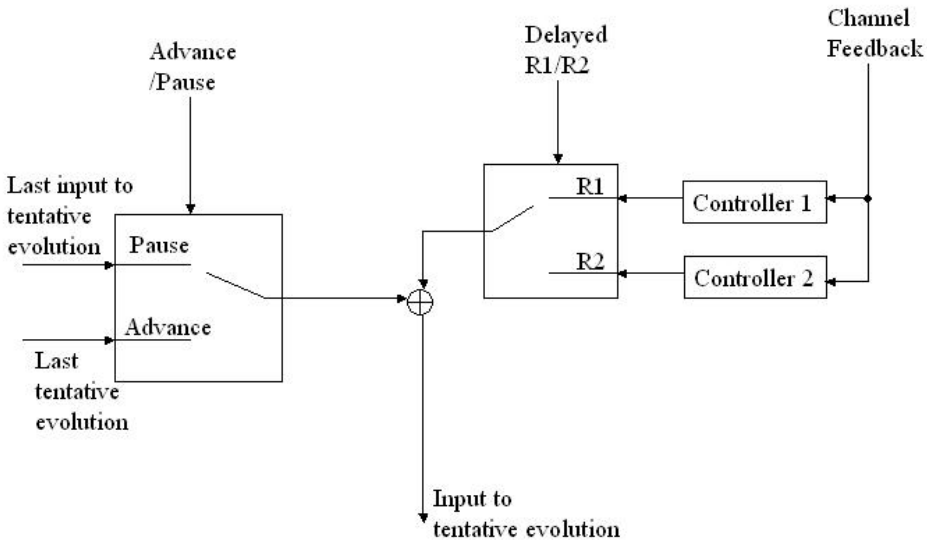

Figure 3.7: State update: advancing time and applying controls

where $a_1 = (2+\epsilon_1)^{R_1} = 2^{R_1+\epsilon_1'}$, $R_1 > \frac{R_{in}}{1-\varepsilon}$ to make the queue stable. $W_t^{R_1}$ is the input to the linear system, which is generated by the source encoder from the bits to be transmitted. The control system attempts to track $W_t^{R_1}$. $W_t^{R_1}$ is required to be bounded in $[-1,1]$ when the queue is short, e.g. if $R_{in} = 1$, $W_t = 1$ when the bit to be transmitted is 1, and $W_t = -1$ when the bit is 0. Obviously, the second moment of the disturbance is upper-bounded by 1.

- Observer and input to the channel

$$y_t = C_1 x_t$$

- Channel Output when non-erasure

$$z_t = y_t + n_t$$

where $n_t \sim \mathcal{N}(0,1)$ and process $\{n_t\}_{t\geq 0}$ is white.

When $l_t > L_c$: $\mathcal{E}_2$

- Linear system $\mathcal{L}_2$

$$\begin{aligned} x_{t+1} &= 2^{R_2+\epsilon_2'}(x_t + U_t) + W_{t+1}^{R_2} \\ &= (2+\epsilon_1)^{R_2}(x_t + U_t) + W_{t+1}^{R_2} \\ &= a_2(x_t + U_t) + W_{t+1}^{R_2} \end{aligned}$$

where $R_2 = nR_1$ and $a_2 = a_1^n$, with $n \in \mathbb{Z}^+$ and $n > 1$. $W_t^{R_2}$ is required to be bounded in $[-2^{n(R_1+\epsilon_1')}, 2^{n(R_1+\epsilon_1')}] = [-a_2, a_2]$, and its second moment is therefore upper-bounded by a_2^2. For more discussions on the generation of W_t see section 6.3 of [9].

- Observer and input to the channel

$$y_t = C_2 x_t$$

- Channel Output when non-erasure

$$z_t = y_t + n_t$$

where $n_t \sim \mathcal{N}(0, 1)$ and process $\{n_t\}_{t\geq 0}$ is white.

3.2.2 Decoder

The decoder shown in Figure 3.8 does not have access to the exact input bits nor the exact continuous state of the encoder. The discrete state (queue length) depends only on the sequence of erasures so far and so the decoder does have access to that. In response to a query asking the value of any particular bit, the decoder checks to see if it is still waiting in the queue. If not, it gives its best estimate of that bit's value.

To extract an estimate of the bit from the received channel outputs so far, the decoder maintains an internal state corresponding to how the encoder's state would have evolved if there have been no inputs W_t. By the linearity of the continuous state evolution in the encoder, the sum of the response of the system to only the controls to the response of the system to only the inputs would be the actual state of the encoder. Although both of those terms individually represent the outputs of unstable processes, their sum is stable. Thus the response to the controls alone has to track closely the response to the inputs alone. The unstable state can be thought to consist of an integer whose binary expansion is the desired data bits. [9] [10]

When $l_t \leq L_c$: $\mathcal{D}_1$

- Estimator

$$\hat{x}_t = D_1 z_t$$

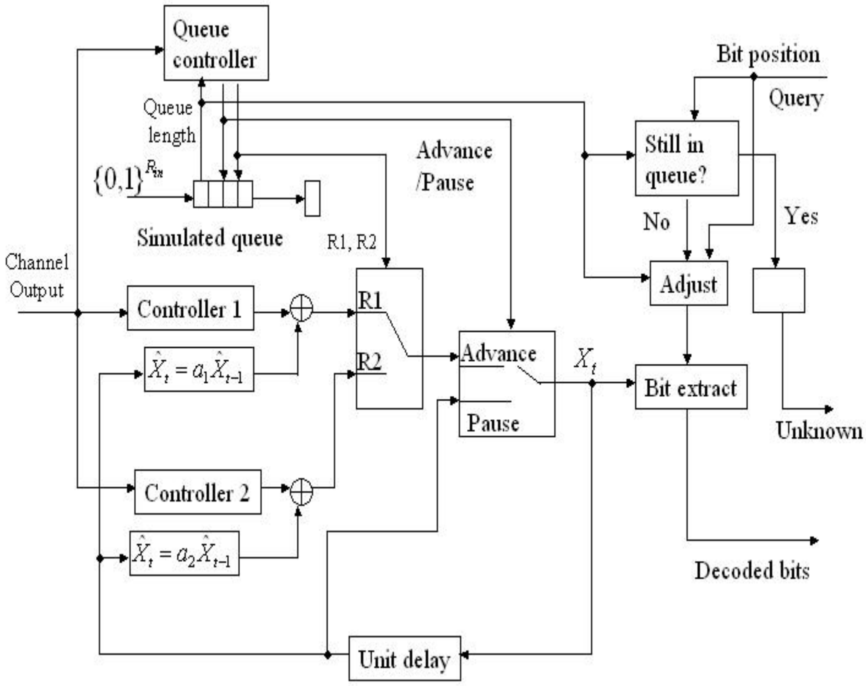

Figure 3.8: Decoder

Motivated by the MMSE estimator, we let $C_1D_1 = \frac{P_1}{1+P_1}$, i.e. $D_1 = \frac{P_1}{C_1(1+P_1)}$

- Control input

$$U_t = -\hat{x}_t$$

When $l_t > L_c$**:** $\mathcal{D}_2$

- Estimator

$$\hat{x}_t = D_2 z_t$$

Motivated by the ML estimator, we let $C_2D_2 = 1$, i.e. $D_2 = \frac{1}{C_2}$.

- Control input

$$U_t = -\hat{x}_t$$

3.2.3 Target powers

In the short queue mode $E\{y_t^2\} \leq P_1$, and in long queue mode $E\{y_t^2\} \leq P_2$, such that overall $E\{y_t^2\} \leq P$. Notice we only constrain the average transmission power, while allowing instantaneous power to be as large as needed.

3.2.4 The order of system evolution

$x_{t-1} \to y_{t-1} \to z_{t-1} \to U_{t-1} \to a_i(x_{t-1} + U_{t-1}) + W_t^{R_i} \to x_t$, where $i \in \{1, 2\}$.

In real time, the system first makes a tentative evolution with $\mathcal{L}_1$ and $\mathcal{L}_2$ in parallel. Only when the previous output is not erased will the control be applied and the newly observed state be transmitted. If the output is erased, the last tentative state will be preserved, and the system goes through a null dynamics, to advance in real time. However the simulation time is not advanced when the channel erases the transmission.

This system works because once a bit is out of the queue, the decoder's estimate of it will converge doubly exponentially in the number of subsequent non-erased channel outputs. As such, the dominant source of errors is the erasures. The average power constraint is met by making the "long queue" behavior rare enough by choosing a suitably large threshold L_c between the two behaviors.

3.3 Analysis of the system

In all the following discussion, time t is the simulation time, i.e., time steps advance only when non-erasure happens in the channel. For presentation simplicity, we assume the variance of the AWGN $\sigma = 1$. When $\sigma \neq 1$, we can scale the power obtained by σ^2, and all other results hold.

3.3.1 State evolution in the short queue mode

Proposition 3.1. *As long as $a_1 < 1 + P_1$, the state of the short queue mode is stable. In such case, the last state of the system before the last long-short transition acts as the initial condition in the short queue mode.*

Proof. Suppose at simulation time $t = \tau$, the queue length is $l_\tau > L_c$ bits, then we have

$$x_\tau = -\frac{a_2}{C_2} n_{\tau-1} + W_\tau^{R_2} \tag{3.1}$$

and

$$U_\tau = -\hat{x}_\tau = -\frac{1}{C_2}(C_2 x_\tau + n_\tau) = -\left(x_\tau + \frac{1}{C_2}n_\tau\right)$$

Suppose at simulation time $\tau+1$, the queue length $l_{\tau+1} \leq L_c$, then

$$\begin{aligned} x_{\tau+1} &= a_1(x_\tau + U_\tau) + W_{\tau+1}^{R_1} \\ &= W_{\tau+1}^{R_1} - \frac{a_1}{C_2}n_\tau \end{aligned}$$

For later simulation time in the short queue mode, we have

$$x_{t+1} = a_1(1 - C_1 D_1)x_t + (W_{t+1}^{R_1} - a_1 D_1 n_t) \tag{3.2}$$

To make the system stable we require

$$a_1(1 - C_1 D_1) < 1 \Leftrightarrow a_1\left(1 - \frac{P_1}{1+P_1}\right) < 1 \Leftrightarrow a_1 < 1 + P_1$$

In such a situation $x_{\tau+t+1}$ is:

$$\begin{aligned} x_{\tau+t+1} &= (a_1(1-C_1D_1))^t x_{\tau+1} + \sum_{j=1}^{t-1} W_{\tau+j+1}^{R_1}(a_1(1-C_1D_1))^{t-j} \\ &\quad - \sum_{j=1}^{t-1} a_1 D_1 n_{\tau+j}(a_1(1-C_1D_1))^{t-j} \\ &= (a_1(1-C_1D_1))^t \frac{a_2}{c_2} n_\tau + \sum_{j=0}^{t-1} W_{\tau+j+1}^{R_1}(a_1(1-C_1D_1))^{t-j} \\ &\quad - \sum_{j=1}^{t-1} a_1 D_1 n_{\tau+j}(a_1(1-C_1D_1))^{t-j} \end{aligned} \tag{3.3}$$

The long queue mode provides the short queue mode with a initial state.

□

Remark 3.2. The power of the channel input during the mode transition is upper-bounded by

$$P_1^{trans} \triangleq C_1^2 E\{x_{\tau+1}^2\} \leq C_1^2\left(1 + \frac{a_1^2}{C_2^2}\right) \tag{3.4}$$

We will discuss the relation between P_1 and P_1^{trans} later in 3.3.3, when all the affecting parameters have been defined.

Proposition 3.3. *When the system is stable, as long as* $a_1 < \sqrt{1+P_1} < 1+P_1$, *there is a unique*

$$C_1 = \sqrt{\frac{P_1(1+P_1-a_1)(1+P_1-a_1^2)}{(1+P_1)(1+P_1+a_1)}}$$

so that the average transmission power does not exceed P_1 *after the first simulation time step following the long-short transition.*

Proof. When the system is stable, (3.3) is a bounded zero mean disturbance plus a zero mean Gaussian noise.

Since $C_1 D_1 = \frac{P_1}{1+P_1}$, the contribution of the bounded disturbance is also bounded as follows

$$M \triangleq \sum_{j=0}^{t}(a_1(1-C_1D_1))^j \quad < \quad \sum_{j=0}^{\infty}(a_1(1-C_1D_1))^j = \frac{P_1+1}{P_1+1-a_1}$$

The second moment of the bounded noise is upper-bounded by $M^2 = \left(\frac{P_1+1}{P_1+1-a_1}\right)^2$.

Since the channel noise is independent of the data to be transmitted, the power of channel input in the short queue mode therefore is upper-bounded by

$$\begin{aligned} P_1 &= C_1^2 M^2 + C_1^2 \sum_{j=0}^{\infty} a_1^2 D_1^2 (a_1^2(1-C_1D_1)^2)^j \\ &= \frac{C_1^2(1+P_1)^2(1+P_1+a_1) + a_1^2 P_1^2 (1+P_1-a_1)}{(1+P_1-a_1)^2(1+P_1+a_1)} \end{aligned}$$

As long as $a_1 < \sqrt{1+P_1} < 1+P_1$, there is a unique solution of C_1:

$$C_1 = \sqrt{\frac{P_1(1+P_1-a_1)(1+P_1-a_1^2)}{(1+P_1)(1+P_1+a_1)}} \tag{3.5}$$

□

Proposition 3.4. *In the short queue mode, the tail of the distribution* $P\left(\left|X_t - \hat{X}_t\right| \geq x\right)$ *dies at least as fast as*

$$f_1(x) \triangleq K e^{-\frac{x}{2\sigma_{a_1}^2}} \tag{3.6}$$

where $K > 0$ and $x > 0$, and

$$\sigma_{a_1}^2 = \frac{P_1(1+P_1)}{(1+P_1-a_1)^2(1+P_1-a_1^2)}$$

Proof. The upper-bound of the variance of the Gaussian noise is

$$\sigma_{a_1}^2 = \sum_{j=0}^{\infty} D_1^2(a_1^2(1-C_1D_1)^2)^j = \frac{P_1(1+P_1)}{(1+P_1-a_1)^2(1+P_1-a_1^2)} \tag{3.7}$$

The tail of the total noise, $P(X \geq x)$, dies as follows

$$P(X \geq x) < \begin{cases} 1 & \text{if } 0 < x \leq M \\ \frac{1}{2}e^{-\frac{(x-M)^2}{2\sigma_{a_1}^2}} & \text{if } x > M \end{cases}$$

We discuss in two cases

1. $x > M$

   In this case,

   $$P(X \geq x) < \frac{1}{2}e^{-\frac{(x-M)^2}{2\sigma_{a_1}^2}} \leq \frac{1}{2}e^{-\frac{x-(M+1)}{2\sigma_{a_1}^2}}$$

   for any $x > M$.

   Therefore $K = e^{\frac{M+1}{2\sigma_{a_1}^2}}$ will meet our need.

2. $0 < x \leq M$

   For K to achieve $1 \leq Ke^{-\frac{x}{2\sigma_{a_1}^2}}, \forall |x| \leq M$, we need $Ke^{-\frac{M}{2\sigma_{a_1}^2}} \geq 1$, i.e., $K \geq e^{\frac{M}{2\sigma_{a_1}^2}}$.

Combining the above two cases, we can always pick $K = e^{\frac{M+1}{2\sigma_{a_1}^2}}$ such that $P(X \geq x) \leq Ke^{-\frac{x}{2\sigma_{a_1}^2}}, \forall x \in \mathbb{R}^+$. □

3.3.2 State evolution in the long queue mode

Proposition 3.5. *Except for the first step after short-long mode transition in the simulation time, the system in the long queue mode has no memory of history before the transition.*

Proof. Suppose at simulation time $t = \tau$, the queue length is $l_\tau \leq L_c$, we have

$$x_\tau = a_1(1 - C_1 D_1)x_{\tau-1} + (W_\tau^{R_1} - a_1 D_1 n_{\tau-1}) \tag{3.8}$$

and

$$U_\tau = -(D_1 C_1 x_\tau + D_1 n_\tau)$$

Suppose at simulation time $\tau + 1$, the queue length is $l_{\tau+1} > L_c$, and the system is switched to the long queue mode, then we have

$$\begin{aligned} x_{\tau+1} &= a_2(x_\tau + U_\tau) + W_{\tau+1}^{R_2} \\ &= a_2(1 - C_1 D_1)x_\tau - a_1 D_1 n_\tau + W_{\tau+1}^{R_2} \end{aligned}$$

For the later simulation time t in the long queue mode

$$x_t = W_t^{R_2} - \frac{a_2}{C_2} n_{t-1} \tag{3.9}$$

Hence the system has no memory of the short queue mode behavior in the long queue mode, except for the first time step after transition. This enforced forgetting of the past in the long-queue mode is just a technical tool to simplify analysis such that we can ignore speed of convergence issues after the queue transition. □

Proposition 3.6. *The expected power of $x_{\tau+1}$ at the transition can be bounded as follows:*

$$\begin{aligned} P_2^{trans} &\triangleq C_2^2 E\{x_{\tau+1}^2\} \\ &\leq a_2^2 C_2^2 (1 - C_1 D_1)^2 P_1 + a_1^2 C_2^2 D_1^2 + C_2^2 a_2^2 \\ &= a_2^2 C_2^2 \left((1 - C_1 D_1)^2 P_1 + \frac{a_1^2}{a_2^2} D_1^2 + 1 \right) \end{aligned} \tag{3.10}$$

There is no problem of stability in the long queue mode. Suppose that the power constraint is satisfied with equality, then we have:

Proposition 3.7. *When the system is in the long queue mode, and is* not *at the step when the short-long mode transition happens, the expected power of the channel input is upper-bounded by*

$$P_2 = a_2^2(1 + C_2^2) \tag{3.11}$$

Proof.

$$P_2 = C_2^2 E\{x_t^2\} \leq C_2^2 a_2^2 \left(1 + \frac{1}{C_2^2}\right) = a_2^2(1 + C_2^2)$$

□

Proposition 3.8. *In the long queue mode, the tail of the error distribution* $P\left(\left|X_t - \hat{X}_t\right| \geq x\right)$ *dies at least as fast as*

$$f_2(x) \triangleq \frac{1}{2} e^{-\frac{x^2}{2\sigma_{a_2}^2}} \tag{3.12}$$

where $\sigma_{a_2}^2 = \frac{a_2^2}{P_2 - a_2^2}$.

Proof. The error is zero mean Gaussian with variance smaller than

$$\sigma_{a_2}^2 = \frac{1}{C_2^2} = \frac{1}{\frac{P_2}{a_2^2} - 1} = \frac{a_2^2}{P_2 - a_2^2} \tag{3.13}$$

It goes to infinity as a_2 approaches $\sqrt{P_2}$.

□

3.3.3 The relation between P_1 and P_1^{trans}

We hope to bound the power in the short queue mode by P_1, hence we need to study the relation between P_1 and P_1^{trans}.

Proposition 3.9. *The power in the short queue mode is bounded by* P_1*, as long as*

$$C_2^2 \geq \frac{a_1^2 C_1^2}{P_1 - C_1^2} \tag{3.14}$$

Proof. To bound P_1^{trans} by P_1, using (3.4), we have:

$$P_1^{trans} \leq P_1 \Leftrightarrow C_1^2 \left(1 + \frac{a_1^2}{C_2^2}\right) \leq P_1 \Leftrightarrow C_2^2 \geq \frac{a_1^2 C_1^2}{P_1 - C_1^2}$$

Since from (3.5), the RHS of the last equation is positive, we can always select a big enough C_2 such that $P_1^{trans} \leq P_1$.

□

3.3.4 Compatibility Results

From above results we can see the following two points:

1. Without noise n_t, the un-controlled state of the system depends only on the virtual time.

2. The controlled state depends on the virtual time, the data to be transmitted, and the channel noise since the last queue mode transition.

The first point is true because without channel noise and control, the state of the system is only driven by the bounded disturbance. The disturbance is generated by taking bits out of the source coder at proper rate. The rate is determined by the queue length, or equivalently, the virtual time.

The second point is true because in the long queue mode, the only channel noise matters is the one in the immediately past time step.

Below we use these results to study the estimation error of the state, and therefore the bit error probability. We obtain the relation between bit error probability and delay thereafter. First we have the following proposition.

Proposition 3.10. *The probability of bit error with delay d and queue length l bits, denoted as $P_{error}(d, l)$, is bounded by*

$$P_{error}(d, l) \leq f\left(\frac{\epsilon_1}{2 + 2\epsilon_1} 2^{(\log_2 a_1)d - l}\right)$$

where $f(\cdot) \triangleq \max(f_1(\cdot), f_2(\cdot))$, with $f_1(x) \triangleq Ke^{-\frac{x}{2\sigma^2_{a_1}}}$ as defined in proposition 3.4 and $f_2(x) \triangleq \frac{1}{2}e^{-\frac{x^2}{2\sigma^2_{a_2}}}$ as defined in proposition 3.8.

Proof. Suppose we only have the short queue mode. Then from the derivation of Sahai's thesis ([9], pg. 110), for a given queue length l when the bit is transmitted, let d be the time delay, i be the number of bits received, ζ be the error, and ϵ_1 be defined as $a_1 = 2^{R_1 + \epsilon_1'} = (2 + \epsilon_1)^{R_1}$, then we have:

$$\begin{aligned}
P_{error}(d, l) &\leq P\left(|\zeta| > \frac{\epsilon_1}{2 + 2\epsilon_1}(2 + \epsilon_1)^{-i}\right) \\
&= P\left(\left|X_t - \hat{X}_t\right| > \frac{\epsilon_1}{2 + 2\epsilon_1}(2 + \epsilon_1)^{(\log_{2+\epsilon_1} a_1)d - l}\right) \\
&= f_1\left(\frac{\epsilon_1}{2 + 2\epsilon_1} 2^{(\log_2 a_1)d - l}\right)
\end{aligned}$$

Similarly, if we only have the long queue mode, then the bit error probability satisfies the following inequality:

$$P_{error}(d,l) \le f_2\left(\frac{\epsilon_1}{2+2\epsilon_1}2^{(\log_2 a_2)d-l}\right)$$

where $a_2 = 2^{R_2+\epsilon_2'} = (2+\epsilon_1)^{R_2}$.

Below we write $f(\cdot) = \max(f_1(\cdot), f_2(\cdot))$. Function $f(\cdot)$ is a decreasing function.

For the short queue mode, the proposition is obviously true.

For the long queue mode, since

$$\log_2(a_2) > \log_2(a_1)$$

we have

$$\frac{\epsilon_1}{2+2\epsilon_1}2^{(\log_2 a_2)d-l} \ge \frac{\epsilon_1}{2+2\epsilon_1}2^{(\log_2 a_1)d-l} \tag{3.15}$$

Therefore in the long queue mode

$$\begin{aligned} P_{error}(d,l) &\le f\left(\frac{\epsilon_1}{2+2\epsilon_1}2^{(\log_2 a_2)d-l}\right) \\ &\le f\left(\frac{\epsilon_1}{2+2\epsilon_1}2^{(\log_2 a_1)d-l}\right) \end{aligned}$$

□

Let $\Omega' = \frac{\epsilon_1}{2+2\epsilon_1}$. Now we remove the condition on l and get the probability of bit error as the following:

$$\begin{aligned} P_{error}(d) \le &\sum_{j=0}^{L_c} P(l \ge j) f\left(\Omega' 2^{(\log_2 a_1)d-j}\right) + \sum_{k=L_c+1}^{d-1} P(l \ge k) f\left(\Omega' 2^{(\log_2 a_1)d-k}\right) + \\ &P(l \ge d)\cdot 1 \\ &\le L' \sum_{j=0}^{L_c} 2^{-\alpha_1 j} \exp\left(-\frac{\Omega' 2^{R_1 d-j}}{2\sigma_{a_1}^2}\right) + L'' \sum_{k=L_c+1}^{d-1} 2^{-\alpha_2 k} \exp\left(-\frac{\Omega' 2^{R_1 d-k}}{2\sigma_{a_1}^2}\right) \\ &+ L''' 2^{-\alpha_2 d} \end{aligned} \tag{3.16}$$

with L', L'', and L''' being positive constants. We consider the three terms in turn below. The goal is to prove there exist designs of the system such that $P_{error}(d)$ goes to zero with delay d at least exponentially at the desired rate α.

First look at the third term, the simplest one, in (3.16). For any given $\alpha < -\log_2 \varepsilon$, we can pick an $\alpha_2 \in (\alpha, -\log_2 \varepsilon)$ to make the third term decay fast enough with delay.

Lemma 3.11. *The first term in (3.16), denoted as $G_1(d)$, decays to zero at least doubly exponentially with delay d.*

Proof. The first term in equation (3.16) satisfies the following inequalities

$$\begin{aligned} G_1(d) &\triangleq L' \sum_{j=0}^{L_c} 2^{-\alpha_1 j} \exp\left(-\frac{\Omega' 2^{R_1 d - j}}{2\sigma_{a_1}^2}\right) \\ &< L' \left(\frac{1-2^{-\alpha_1 L_c}}{1-2^{-\alpha_1}}\right) \exp\left(-\frac{\Omega'}{2\sigma_{a_1}^2} 2^{-L_c} 2^{R_1 d}\right) \end{aligned}$$

Hence the proposition is true. □

Since a double exponential function of d decays to zero faster than any exponential function of d, for any system design and α, $G_1(d)$ will decay faster than desired.

Lemma 3.12. *The second term in equation (3.16), denoted as $G_2(d)$, can go to zero with d at desired rate $\alpha < -\log_2 \varepsilon$ by selecting the long queue mode rate R_2.*

Proof. Denote the second term in (3.16) as

$$G_2(d) = L'' \sum_{k=L_c+1}^{d-1} g(k)$$

where

$$g(k) \triangleq 2^{-\alpha_2 k} \exp\left(-\frac{\Omega' 2^{R_1 d - k}}{2\sigma_{a_1}^2}\right)$$

We will show below that $g(k)$ can be upper-bounded by an exponential function of d. The value of k can only be integers between $d-1$ and L_c+1, but we assume for now $k \in \mathbb{R}^+$ and $L_c + 1 \leq k \leq d-1$ and differentiate $g(k)$ to find its trend with changing k, , then

$$\frac{dg}{dk} = g(k) \cdot \ln 2 \cdot \left(-\alpha_2 + \frac{\Omega'}{2\sigma_{a_1}^2} 2^{R_1 d - k}\right)$$

Since $g(k) > 0, \forall k \in [L_c + 1, d-1]$, we look at the sign of

$$h(k) \triangleq -\alpha_2 + \frac{\Omega'}{2\sigma_{a_1}^2} 2^{R_1 d - k}$$

Obviously $h(k)$ is a decreasing function of k. When

$$d > \max\left(\frac{L_c + 1 + \log_2\left(\frac{2\sigma_{a_1}^2\alpha_2}{\Omega'}\right)}{R_1}, \frac{-1 + \log_2\left(\frac{2\sigma_{a_1}^2\alpha_2}{\Omega'}\right)}{R_1 - 1}\right) \tag{3.17}$$

$h(k)$ is always positive in the interval $[L_c + 1, d - 1]$, hence $g(k)$ increases with k. $g(k)$ achieves its maximum value when $k = d - 1$. Therefore $G_2(d)$ satisfies the following

$$\begin{aligned} G_2(d) =& L'' \sum_{k=L_c+1}^{d-1} 2^{-\alpha_2 k} \exp\left(-\frac{\Omega' 2^{R_1 d - k}}{2\sigma_{a_1}^2}\right) \\ \leq& L''(d - L_c - 2)2^{-\alpha_2(d-1)} \exp\left(-\frac{\Omega' 2^{(R_1-1)d+1}}{2\sigma_{a_1}^2}\right) \qquad (3.18) \\ =& L'' 2^{\alpha_2} \cdot (d - L_c - 2)2^{-\alpha_2 d} \exp\left(-\frac{\Omega' 2^{(R_1-1)d+1}}{2\sigma_{a_1}^2}\right) \qquad (3.19) \end{aligned}$$

$G_2(d)$ is the product of $(d - L_c - 2)2^{-\alpha_2 d}$ with a doubly exponential function of d. Hence $\forall \alpha < -\log_2 \varepsilon$, there exists α_2 lying between α and $-\log_2 \varepsilon$ such that $G_2(d)$ decays to zero at least exponentially at rate higher than α. □

We have finished analyzing all three terms in (3.16) and proved the following proposition.

Proposition 3.13. *The probability of bit error (3.16) goes to zero exponentially with d at any desired rate $\alpha < -\log_2 \varepsilon$ by properly designing the system.*

3.4 Proof of achievability

In this section we prove Theorem 1.7, the main theorem of the thesis. For clarity we repeat the theorem below.

Theorem 3.14. *The anytime capacity of AWGN + erasure channel with average power constraint P', noise variance σ^2, and encoder having access to noiseless feedback is*

$$C_{anytime}(\alpha) = \begin{cases} \frac{(1-\varepsilon)}{2}\log_2\left(1 + \frac{P'}{\sigma^2}\right) & \text{if } 0 \leq \alpha \leq -\log_2 \varepsilon \\ 0 & \text{otherwise} \end{cases} \tag{3.20}$$

Proof. Now suppose that the power constraint P' is known, then the achievable data rate is in $\big[0, \frac{1-\epsilon}{2}\log_2(1+\frac{P'}{\sigma^2})\big)$ bits per channel use as shown in Figure 1.3. Below for simplicity we denote $P = \frac{P'}{\sigma^2}$, and all power values obtained will have to be scaled by σ^2 to get the real values. For any given point (R_{in}, α) in the achievable region in Figure 1.3, the following is the procedure of system design to achieve it. Without specific order, the parameters of the system to be determined are P_1, R_1, ϵ_1, a_1, C_1, D_1, α_1, α_2, n, R_2, a_2, C_2, D_2, P_2, and L_c. The system design can follow these steps below in the indicated order.

1. P_1

   Select an arbitrary P_1 such that $2^{\frac{2R_{in}}{1-\varepsilon}} - 1 < P_1 < P$. This is always possible since the given R_{in} is in the achievable region.

2. R_1 and ϵ_1

   $R_1 = \frac{R_{in}}{1-\varepsilon} + \delta_1$, where δ_1 is an arbitrary small positive number. Select an arbitrary ϵ_1 which satisfies $(2+\epsilon_1)^{R_1} = (2+\epsilon_1)^{\frac{R_{in}}{1-\varepsilon}+\delta_1} \leq \sqrt{P_1+1}$. These are enabled by the selection of P_1.

3. a_1

   Value of a_1 comes from R_1 and ϵ_1 as $a_1 = (2+\epsilon_1)^{R_1} = (2+\epsilon_1)^{\frac{R_{in}}{1-\varepsilon}+\delta_1}$.

4. C_1 and D_1

   $C_1 = \sqrt{\frac{P_1(1+P_1-a_1)(1+P_1-a_1^2)}{(1+P_1)(1+P_1+a_1)}}$ comes from equation (3.5). $D_1 = \frac{P_1}{C_1(1+P_1)}$, motivated by MMSE estimator.

5. α_1

   α_1 is obtained by looking up the anytime capacity curve of the erasure channel with noiseless feedback using Theorem 2.2 and Figure 2.7. The value of α_1 corresponds to rate $\frac{R_{in}}{R_1}$.

6. α_2

   Choose α_2 such that $\alpha < \alpha_2 < -\log_2\varepsilon$, and (3.18) decreases with d at rate higher than α.

7. n and R_2

   Choose n such that $R_2 = nR_1, n \in \mathbb{Z}^+, n > 1$. And n is selected such that on the anytime capacity curve of the erasure channel with noiseless feedback, the normalized rate $\frac{R_{in}}{R_2}$ corresponds to anytime exponent α_2 or larger.

8. a_2

$$a_2 = a_1^n = (2 + \epsilon_1)^{nR_1}$$

9. C_2 and D_2

   C_2 is selected to satisfy (3.14), and $D_2 = \frac{1}{C_2}$.

10. P_2

    P_2 comes from equation (3.11), i.e. $P_2 = a_2^2(1 + C_2^2)$.

11. L_c

    Select L_c such that the average power constraint is met. Formally, let $P_{2,max} \triangleq \max(P_2, P_2^{trans})$. It is clear from equations (3.10) and (3.11) that $P_{2,max}$ does not depend on L_c. We require

$$\begin{aligned} E\{y_t^2\} &= Prob(l \leq L_c)P_1 + Prob(l > L_c)P_{2,max} \\ &< P_1 + Prob(l > L_c)P_{2,max} \\ &\leq P_1 + 2^{-\alpha_2^* \lfloor \frac{L_c}{R_{in}} \rfloor} P_{2,max} \end{aligned}$$

    where the first inequality comes from lemma 2.3.

    Since P_1 is selected to be smaller than P, there are always L_c large enough to satisfy (3.21).

Now all the parameters are selected and the system is designed to achieve the given (R_{in}, α).

Since we can always have a string of consecutive erasures between the time the bit was sent and the current time, the bit error probability can never decay faster than ε^d. Thus the anytime capacity of the AWGN+erasure channel is zero for $\alpha > -\log_2 \varepsilon$.

For $0 \leq \alpha < -\log_2 \varepsilon$, we have proved above the anytime capacity of the channel is larger than $\frac{1-\varepsilon}{2}\log_2(1+P)$. But the Shannon capacity of the AWGN+erasure channel with power constraint P, either with or without feedback is $\frac{1-\varepsilon}{2}\log_2(1+P)$, and anytime capacity of a channel cannot be larger than its Shannon capacity. Therefore

$$C_{anytime}(\alpha) = \frac{1-\varepsilon}{2}\log_2(1+P),\ \forall 0 \leq \alpha < -\log_2(\varepsilon)$$

$\square$

3.5 Simulation results and discussions

Figures 3.9-3.14 show the simulation results for one realization of a particular system. The input rate R_{in} is 1 bit. The erasure probability $\varepsilon = 0.4$. The output rate in the short queue mode is $R_{short} = 2$ bits. Notice that the system is stable even when only short queue mode is applied. The bit rate for long queue mode, R_{long}, is 3 bits. The two modes transit at the critical queue length $L_c = 4$ bits. In all these figures, the horizontal axis is the real time.

Figure 3.9 shows an arbitrary realization of the erasure/non-erasure in the channel in real time. Each filled dot means a successful transmission, and each empty dot stands for a erasure. All the following figures in this section come from this same realization of the channel.

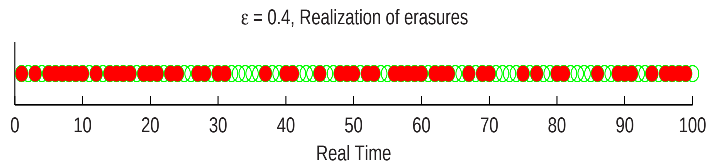

Figure 3.9: Successful transmission and erasure: $\epsilon = 0.4$, filled dots are successes and empty dots are erasures

Figure 3.10 shows the virtual time as a function of real time in this particular channel realization. We compare the virtual times of a system with short-queue mode only and of a system with two modes. As discussed above, erasures make the virtual time lag from the real time. System design should keep the lag small. In Figure 3.10 we can see when two modes are applied, the high data rate in the long queue mode accelerates the reduction of the lag, hence the virtual time in the two-mode case is closer to the real time than when only the short queue mode is applied.

Figure 3.11 explicitly shows the lag between virtual time and real time. We can see that before the first non-erasure after the system enters the long queue mode for the first time, the lags in the two cases are the same. But when a successful transmission occurs when the queue is long, the two-mode system transmits more bits than the one-mode system, and reduces the lag faster. Also notice since the system is stable even only when the short queue mode is applied, the lag of the one-mode system converges to zero eventually, although longer time has elapsed.

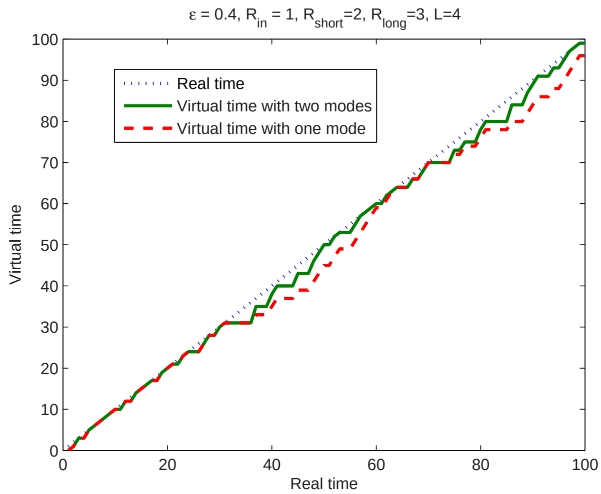

Figure 3.10: A realization of Virtual Time as a function of Real time: $\epsilon = 0.4$, $R_{in} = 1$ bit, $R_{short} = 2$ bits, $R_{long} = 3$ bits, $L_c = 4$

Figure 3.12 is the indicator function of long queue mode, i.e., when the system is in long queue mode, the function returns value 1, otherwise it returns 0.

Figure 3.13 shows the average power corresponding to the channel realization in Figure 3.9. We can see that when the system stays in the short queue mode, the average transmission power increases gradually and converges P_1, and when the system stays in the long queue mode, the average transmission power is $P_2 > P_1$ due to the higher data rate. At the instant the system transits from the long queue to short queue mode the average power is $P_1^{trans} < P_1$ by design. As the system stays in the

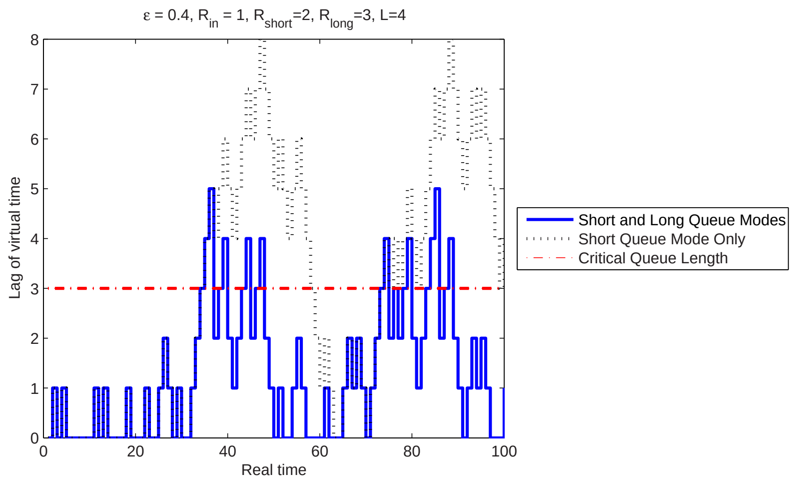

Figure 3.11: Lag between Real Time and Virtual Time with 1 mode and two modes: $\epsilon = 0.4$, $R_{in} = 1$ bit, $R_{short} = 2$ bits, $R_{long} = 3$ bits, $L_c = 4$

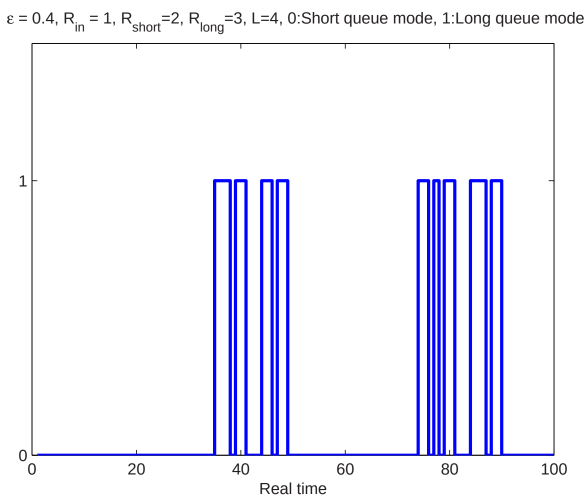

Figure 3.12: Time period in short and long queue mode: $\epsilon = 0.4$, $R_{in} = 1$ bit, $R_{short} = 2$ bits, $R_{long} = 3$ bits, $L_c = 4$, Short queue: 0, Long queue: 1

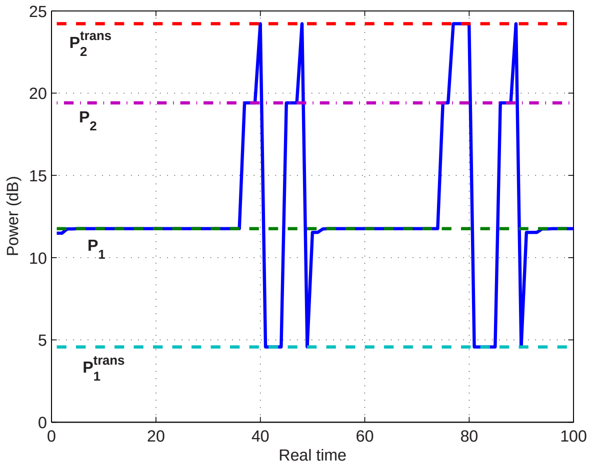

Figure 3.13: Average power: $\epsilon = 0.4$, $R_{in} = 1$ bit, $R_{short} = 2$ bits, $R_{long} = 3$ bits, $L_c = 4$

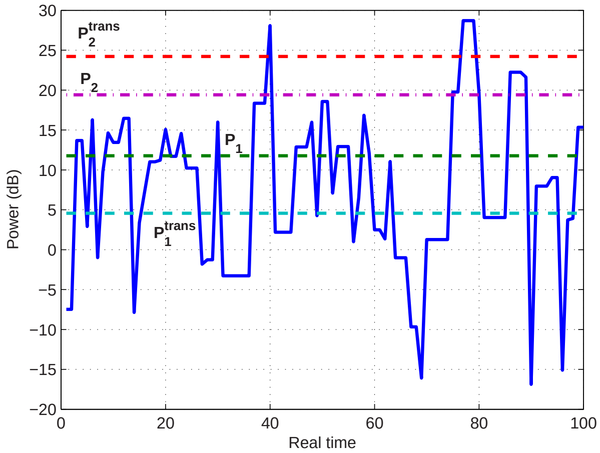

Figure 3.14: A realization of the actual power: $\epsilon = 0.4$, $R_{in} = 1$ bit, $R_{short} = 2$ bits, $R_{long} = 3$ bits, $L_c = 4$

short queue mode, the power increases from P_1^{trans} and converges to P_1. Similarly the average power changes to P_2^{trans} at the instant the system transits to the long queue mode. Then the average power goes to P_2 in the long queue mode until the system transits back to the short queue mode. In contrast to the average power, Figure 3.14 shows the actual power for the realizations of the erasures shown in Figure 3.9, and an arbitrary particular realization of the Gaussian noise. The correspondence between the actual power in Figure 3.14 and average power in Figure 3.13 is obvious. Our design limits the average power, while allows instant power to be high. But as we have observed, it is rare for the actual power to be extremely high.

Chapter 4

Conclusion and Future Work

We have constructively proved the anytime capacity of the AWGN+erasure channel, when the encoder has noiseless channel feedback, is the same as its Shannon capacity. In Sahai's Ph.D. thesis [9], the direct relation between the anytime capacity of a channel with the stability of a control system over the channel is obtained. We copy the relevant corollary here.

Corollary 4.1. *A controlled scalar discrete-time unstable linear Markov process with parameter A driven by a bounded noise can be η-stabilized across a noisy channel with the observer having access to noiseless feedback if there is an $\epsilon > 0$ for which $C_{anytime}(\eta \log_2 A + \epsilon) > \log_2 A$ for the channel with noiseless feedback. In particular, if $C_{anytime}(2 \log_2 A + \epsilon) > \log_2 A$, then we can stabilize it in the mean-squared sense.*

Therefore the anytime capacities we derived in the thesis all have their stability implications in the feedback control over the corresponding channels.

We believe that this style of analysis can be extended to give us the anytime reliabilities of many Gaussian channels with noiseless feedback and channel state side information available at both the transmitter and receiver. We are interested in anytime capacity of AWGN+erasure channel without feedback. Other two interesting cases are shown in Figure 4.1 and 4.2. Figure 4.1 shows a channel in which the order of erasure and AWGN are reversed from the channel we studied. In the channel studied in this thesis (Figure 1.1), when the output and the feedback is 0, we know that an erasure has happened. But in the channel shown in Figure 4.1, it is more difficult to determine the occurrence of erasures. If we have side information from the erasure to both the encoder and decoder, then the channel is the same as the AWGN+erasure channel we have studied. Figure 4.2 shows the case when the fading is governed by more complex processes than Bernoulli process, e.g. a Markovian process. Another interesting extension is the FIR channel with Gaussian parameters which is studied by Elia in [5] in feedback control system.

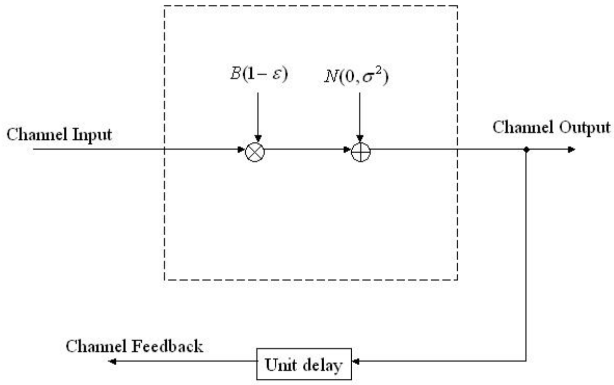

Figure 4.1: Erasure+AWGN channel with feedback

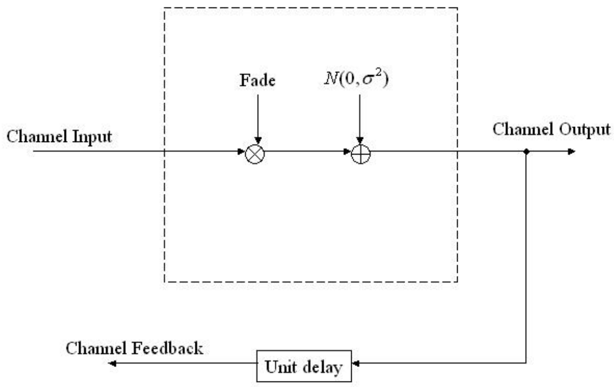

Figure 4.2: Fading+AWGN channel with feedback

Bibliography

[1] Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications. *IEEE Standard 802.11-1997*, 1997.

[2] Special issue on networked control systems. *IEEE Transactions on Automatic Control*, 49(9), September 2004.

[3] N. Abramson. Multiple access in wireless digital networks. *Proceedings of IEEE*, 82:1360–1369, 1994.

[4] T. Şimşek. *Anytime channel coding with feedback*. PhD thesis, University of California at Berkeley, 2004.

[5] N. Elia. Feedback stabilization in the presence of fading channels. *Proceedings of the American Control Conference*, pages 4438–4443, June 2003.

[6] D. Lee, A. Attias, A. Puri, R. Sengupta, S. Tripakis, and P. Varaiya. A wireless token ring protocol for intelligent transportation systems. *IEEE 4th International Conference on Intelligent Transportation Systems, Oakland, CA*, 2001.

[7] R. Murray, K. Astrom, S. Boyd, R. Brockett, and G. Stein. Future directions in control in an information-rich world. *IEEE Control Systems Magazine*, 23(2):20–33, April 2003.

[8] A. Sahai. Evaluating channels for control: Capacity reconsidered. *American Control Conference*, 4:2358–2362, June 2000.

[9] A. Sahai. *Anytime Information Theory*. PhD thesis, Massachusetts Institute of Technology, February 2001.

[10] A. Sahai. The necessity and sufficiency of anytime capacity for scalar control over a noisy communication link. *43rd IEEE Conference on Decision and Control (CDC)*, December 2004.

[11] A. Sahai and T. Şimşek. On the reliability function of discrete memoryless channels with feedback and variable-delay decoding. *IEEE Information Theory Workshop*, October 2004.

[12] P. Seiler. *Coordinated Control of Unmanned Aeiral Vehicles.* PhD thesis, University of California, Berkeley, 2001.

[13] P. Seiler and R. Sengupta. Analysis of communication losses in vehicle control problems. *American Control Conference*, 2:1491–1496, June 2001.

[14] F. Tobagi and L. Kleinrock. Packet switching in radio channels: Part I- carrier sense multiple-access modes and their throughput/delay characteristics. *IEEE Trans. Comm.*, COM-23(12):1400–1416, December 1975.

[15] D. Tse, R. Gallager, and J. Tsitsiklis. Statistical multiplexing of multiple time-scale markov streams. *IEEE Journal on Selected Areas in Communications*, 13(6):1028–1038, August 1995.