FEEDBACK CODING SCHEMES FOR CONTROL OVER GAUSSIAN NETWORKS

A Dissertation

Submitted to the Graduate School
of the University of Notre Dame
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy

by

Utsaw Kumar

J. N. Laneman, Co-Director

V. Gupta, Co-Director

Graduate Program in Electrical Engineering
Notre Dame, Indiana

September 2012

FEEDBACK CODING SCHEMES FOR CONTROL OVER GAUSSIAN NETWORKS

Abstract

by

Utsaw Kumar

The presence of inexpensive and powerful sensing and communication devices has made it possible to deploy large scale distributed systems for a variety of applications. Interactions among different components of such a system include communication of information and controlling dynamical processes, among others. Thus, it is important to blend ideas from information theory and control theory to address problems at the core of such distributed systems.

This dissertation looks at communication scenarios, in which there are *feedback channels* available for enhancing communication and control performance over noisy forward links. Such feedback can considerably increase the reliability or reduce the complexity of coding schemes that approach capacity. Moreover, transmission schemes with feedback can be used to stabilize unstable plants over communication channels, where a sensor transmits the plant state information to a remotely placed controller. We develop transmission schemes for a certain class of Gaussian networks with feedback and apply these schemes to obtain sufficient and, in some cases, necessary conditions for stabilizing a plant via a remotely placed controller.

Specifically, we develop coding schemes for a Gaussian relay channel and a Gaussian product channel with feedback. The coding scheme for the Gaussian relay channel with feedback is based on *distributed stochastic approximation algorithms.* We consider two topologies for the relay channel: (i) a cascade of two Gaussian point to point channels, and (ii) a Gaussian relay channel. Further, we use the proposed coding scheme to mean square stabilize an unstable discrete-time linear time invariant (LTI) system over a Gaussian relay channel and derive sufficient conditions for stabilizability. For the Gaussian product channel, we concentrate on the feedback stabilization problem. It is known that linear coding schemes may lead to overly restrictive stabilizability conditions in such scenarios. We develop a *non-linear quantization-based coding scheme* and present the resulting stabilizability conditions. When these conditions are satisfied with equality, the proposed coding scheme transmits data across the product channel at a rate equal to the capacity of the channel; thus, the conditions are necessary as well. We combine coding schemes for the relay channel and the product channel to consider more general network examples, whose achievability and stability results can be obtained using the simpler relay and product channel results derived earlier.

To my parents and brother my ultimate source of inspiration

CONTENTS

FIGURES

ACKNOWLEDGMENTS

Foremost, I would like to express my sincere gratitude to my advisor Dr. J. Nicholas Laneman for his continuous support, patience, motivation, enthusiasm, and immense knowledge. His guidance helped me a lot in research and writing of this dissertation. I have beneted a lot from numerous discussions and interactions with him. With the independence he gives his students in research, research life becomes a really enjoyable experience. A lot of credit for this dissertation also goes to my co-advisor Dr. Vijay Gupta. His insights to distributed estimation have been tremendously important during the course of my research. He was always accessible and willing to help.

I would like to thank my defense committee members, Dr. Yih-Fang Huang, Dr. Tom Fuja, and Dr. Paolo Minero, for taking the time to serve on my committee and numerous discussions that helped me during the course of my Phd.

I have been lucky to have a lot of friends and colleagues making my stay at Notre Dame really enjoyable. In particular, I would like to thank Matthieu Bloch, Brian Dunn, Michael Dickens, Glenn Bradford, Ebrahim MolavianJazi, Zhanwei Sun, Peyman Hesami and Mostafa Khoshnevisan for the numerous fruitful technical discussions. This dissertation would not have been possible without the support and company of my friends: Aditi, Punit, Ankit, Shashank, Sunil, Krishnan, Vaibhav, Sundeep, Seemit, Sachi, Jai, Surya, Gaurav to just name a few. My deepest gratitude goes to my parents for their love, support and encouragement throughout my life. This dissertation would be simply impossible without them. A special mention has to go to the most special gifts of God to me, my brother Ujwal Kumar, who has been a constant source of encouragement and care to me.

CHAPTER 1

INTRODUCTION

Large scale deployments of distributed systems have been made possible over the last decade by advancements in wireless technologies and cost reductions in sensing and computation. Such systems consist of a multitude of sensors, dynamic processes, communication channels, controllers, and so forth. Examples of large scale distributed systems are power grids, traffic regulatory systems and water distribution systems. Different components of such large scale systems interact and cooperate with each other to achieve a common goal, e.g., in power grids, the goal can be distributing the power among different components such that at least a minimum level of service is maintained.

A distributed system architecture allows for greater flexibility in system design and provides more robustness to individual component failures. Although a distributed system architecture simplifies the design of individual components, it also presents challenges regarding compromises on the overall system performance. Collecting information from a multitude of sensors, reliably communicating it to different components and controlling dynamical processes over wireless channels are just a subset of problems at the core of such distributed systems.

Traditionally, different communities have focused on independently approaching specific aspects of such distributed systems. For example, the information theory community has looked at the problem of reliably communicating a message from one point to another over noisy channels through mostly one-way schemes. Furthermore, most results are asymptotic and delay is not an issue. Working independently, the control theory community has looked into the problem of controlling dynamical processes with an objective of system stability, while satisfying some cost constraint. For control problems, delay is important and feedback forms an integral part of the system. However, to tackle large distributed systems, it is important to work at the intersection of information theory and control theory. Studying such distributed systems is inherently a difficult problem because the traditional assumptions of communications and control are no longer valid. Specifically, there is a need to revisit the communication problems under a delay constraint and study the effect of communication losses on control system design.

This dissertation partially addresses such problems by investigating two aspects of network communication with feedback:

- multihop relaying
- parallel communication paths

We develop coding schemes for the relay channel and product channel with feedback and, we apply these schemes to study control problems with communication constraints. We can combine the results for communication and control over these simple channels to study more general network examples. In particular, we look at one network example that is a combination of relays and product channels.

1.1 Communication over Channels with Feedback

A system diagram for a communication link is given in Fig. 1.1. In information theory, stating informally a rate R is said to be achievable if the probability of

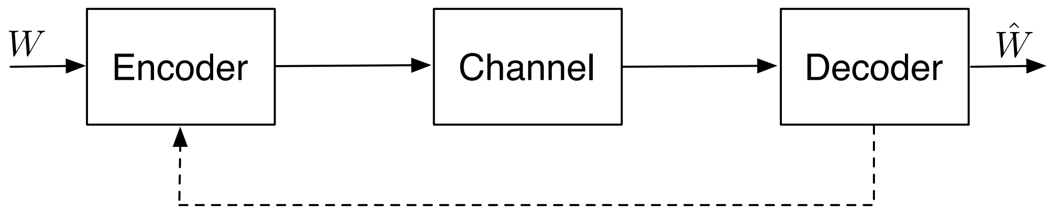

Figure 1.1. Block diagram model for a communication link model.

decoding error converges to zero as the number of channel uses goes to infinity. Over roughly the past sixty years, information theory and communication system engineering have made great advances in understanding the fundamental limits and practical techniques for approaching them. Generally speaking, information theory suggests forward error correction over a very large blocklength, and delay is not important as long as the message is communicated reliably.

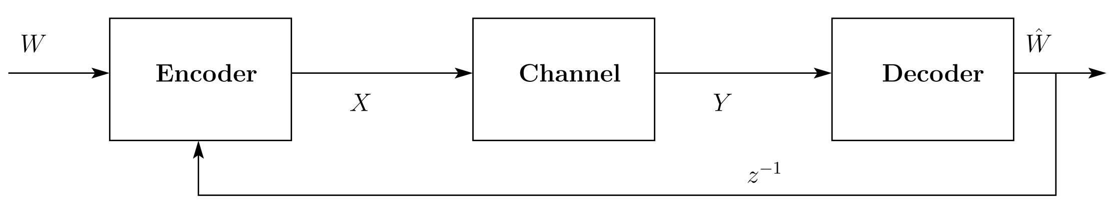

Figure 1.2. Block diagram model of a communication channel with feedback.

In certain communication scenarios, we might have the possibility of using feedback to improve communication over a noisy forward link. Consider a communication channel with perfect feedback as shown in Fig. 1.2. The source needs to communicate a message W to the destination over a memoryless noisy channel. The decoder calculates an estimate $\hat{W}$ of the message. The estimate of the message at the decoder is known at the encoder with a unit time step of delay to preserve causality. Shannon showed that the channel capacity of a memoryless noisy point-to-point channel is, somewhat surprisingly, not increased by noiseless feedback [41]. However, feedback has been shown to be useful in a number of ways. Specifically, feedback helps in decreasing the complexity of capacity-achieving schemes. Also, the rate of decay of the probability of error as a function of the number of channel uses increases with feedback [49].

To illustrate some of the advantages of feedback, consider an example of a memoryless binary erasure channel as shown in Fig. 1.3. The capacity of the erasure channel with or without feedback is $C = 1 - \alpha$ bits per channel use. The capacity expression is intuitive in the sense that, since the proportion α of the bits are lost, one can recover at most the proportion $1 - \alpha$ of the bits. However, a specific coding scheme that achieves this rate is not obvious. If the transmitter obtains perfect feedback of the channel output from the receiver, then it is easy to see how we can achieve capacity. If a bit is erased, retransmit it until it is successfully received. Since the probability of successful transmission is $1 - \alpha$, we achieve a rate of $1 - \alpha$ bits per channel use. Thus, we see that feedback helps in developing a low-complexity encoding and decoding scheme that achieves capacity.

Similar advantages can be seen for an Additive White Gaussian Noise (AWGN) channel with perfect feedback. Schalkwijk and Kailath showed that the delay

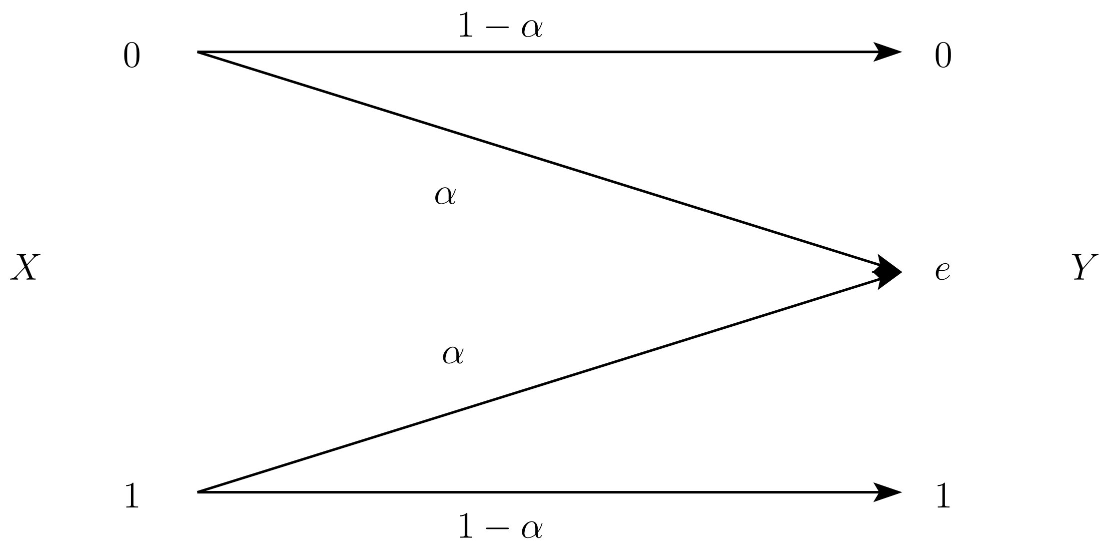

Figure 1.3. Binary Erasure Channel.

and complexity of capacity-approaching schemes for the AWGN channel can be dramatically reduced with noiseless feedback [38]. They proposed a linear iterative coding scheme (popularly known as the Schalkwijk-Kailath (SK) scheme) based on the Robbins-Munro stochastic approximation technique [35]. The encoder has causal access to the estimates $\hat{W}$. They showed that for their scheme, the error probability decays doubly exponentially with time compared to just exponentially for one-way schemes. A good example of such a scenario is communication with a space satellite, where the power from the earth station to the satellite is much larger than the satellite to ground direction, *a fortiori* the former channel can be considered to be noiseless. Ozarow extended the SK scheme to the two-user multiple access channel (MAC) and showed that the capacity region of a MAC is enhanced by the use of feedback [34]. However, many of the available results are

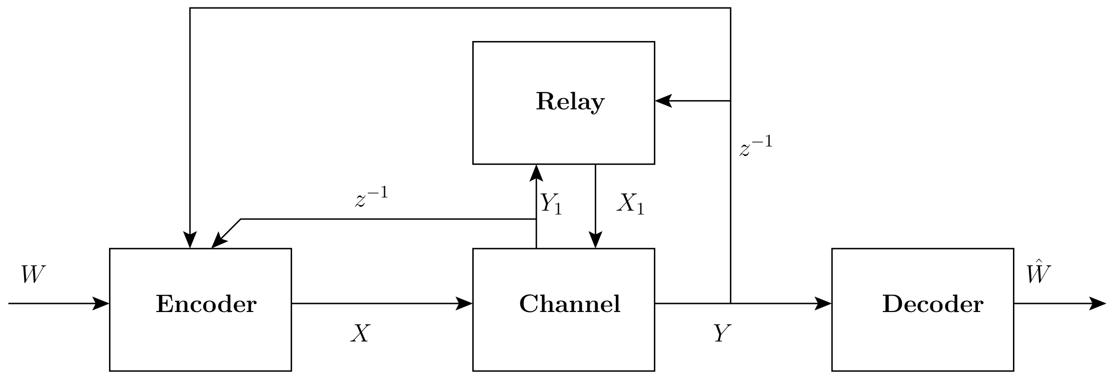

Figure 1.4. Block diagram model of a relay channel with feedback.

for simple channels and networks and are under different simplifying assumptions like noiseless, orthogonal feedback. Thus, there is a need to explore such feedback communication schemes for more general network scenarios, which is the focus of this dissertation.

For example, consider the relay channel with feedback as shown in Fig. 1.4. In [11], Cover and El Gamal give the capacity expression for the relay channel with feedback. They propose coding schemes based on the Block-Markov superposition encoding, that achieve capacity. The relay here functions in full-duplex mode and feedback is over orthogonal channels. However, Block-Markov coding schemes are highly complex, and do not scale well with the number of nodes. In this work, we present feedback coding schemes for two different configurations - a cascade of two point-to-point Gaussian channels and a Gaussian relay channel. The coding schemes are extensions of the aforementioned SK scheme for point-to-point channels.

The imperfect feedback problem has been looked at in [10, 21] and references therein. Though a lot of work needs to be done to fully understand the advantages of imperfect feedback, it is not the focus of this dissertation.

1.2 Contol over Communication Channels

Control theory deals with the problem of controlling a dynamical process with the objective of maintaining system stability while minimizing some cost function. A system block diagram is given in Fig. 1.5. The state of the process is observed by a sensor that generates observations, which can be noisy. Based on an estimate of these observations, the controller generates a control signal that is then applied to the process through an actuator. Feedback is an integral part of control systems and a particular piece of information from the plant might lose its significance if communicated to the controller with delay. The advancements in control theory are extensive and have helped us to solve a large variety of problems. Generally speaking, it has been assumed that an unlimited amount of information can be communicated and processed by the different system components. Recently, there has been a need to understand tradeoffs that result from limited communication in a control loop. The internet, wireless networks, and so forth, have made it possible to have plants linked to remotely placed controllers through communication channels, a problem area commonly referred to as networked control systems.

Networked control systems are now an active area of research (e.g., [1, 9] and the references therein). The performance of such systems is adversely affected by the detrimental effects such as random delays, data loss, data corruption, and so on introduced by the underlying communication network. The major challenge in such problems is to characterize the amount of information transmission needed to achieve a given level of performance. In this dissertation, we study the problem

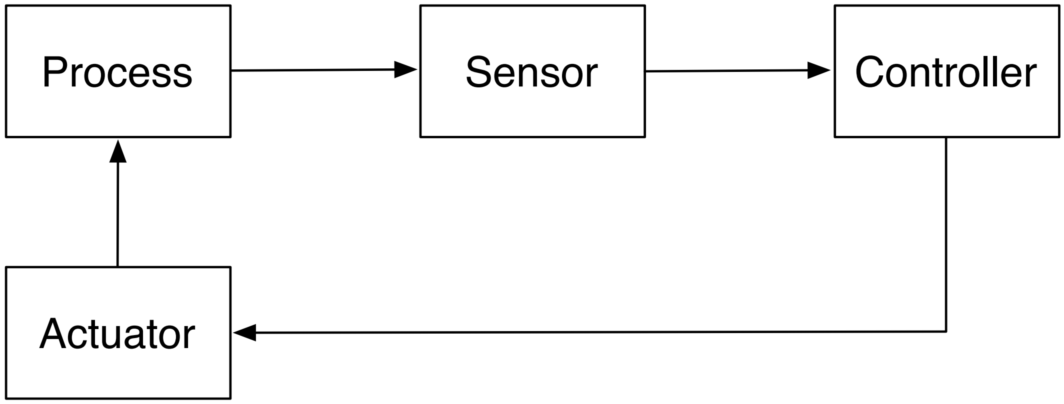

Figure 1.5. Block diagram for a control system model.

of stabilization when communication from the sensor to the controller in Fig. 1.5 is noisy. All the other communication links are assumed to be noiseless.

Consider a plant being stabilized by a remotely placed controller as shown in Fig. 1.6. In control theory, a *data rate theorem* poses a constraint on the minimum information transmission rate required to remotely stabilize a plant. If the process is scalar, it states that a discrete linear time invariant plant with unstable eigenvalue $\lambda > 1$ can be stabilized if and only if the information transmission rate R satisfies $R > \ln \lambda$ nats per second, where $\ln \lambda$ represents how the uncertainty in the process scales with time, also called the *intrinsic entropy* of the process [30]. This result in control theory reminds us of the source coding theorem in information theory [40]. Stated informally, the amount of uncertainty in the source forces a constraint on the minimum communication rate required to make the decoding error arbitrarily small. The analogous result in control theory states that in order to stabilize an unstable process the intrinsic entropy rate $\ln \lambda$ of an un-

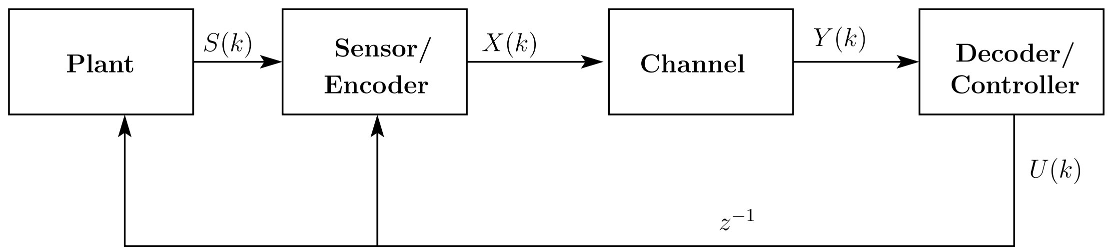

Figure 1.6. Block Diagram of a Plant being Stabilized over a Communication Channel.

stable process should be strictly less than the information transmission rate over the channel. The information transmission rate in turn depends on the power at the transmitter, for example, the maximum information transmission rate over a point-to-point AWGN channel is given by $R = \frac{1}{2} \ln \left(1 + \frac{P}{2\sigma^2}\right)$ nats per channel use, where P is the average transmission power and σ^2 is the noise variance. This lower bounds the amount of power P needed to stabilize a plant with unstable eigenvalue λ.

There is another interpretation to this result. For a fixed average power P and a given channel, the above result specifies the *stability condition*, i.e., the range of values of λ for which the plant can be stabilized. Such stability results are known for the a variety of channels [7, 30, 32]. However, most results are for point-to-point channel configurations and study of stabilization over network configurations, e.g., cascade of point-to-point channels, relay channel, and so forth is just beginning. In this dissertation, we present stability results for a variety of Gaussian channels and discuss how these results can be combined to treat some more general network problems.

1.3 Contributions of the Dissertation

The main contributions of the dissertation can be summarized as follows:

- We develop coding schemes based on distributed stochastic approximation algorithms for a cascade of two point-to-point Gaussian channels with noiseless, orthogonal feedback. The schemes that we propose are capacity achieving and provide a decay in error that is doubly exponential in blocklength.

- We develop a scheme for a Gaussian relay channel with noiseless, orthogonal destination-relay and destination-transmitter feedback. The scheme is also based on distributed stochastic approximation and is shown to provide a doubly exponential error decay for the case of perfect feedback, similar to the cascade case. Our scheme provides a nontrivial, yet simple way to extend the SK schemes to the relay channel case and achieves rates close to capacity, though it is not capacity achieving. We provide numerical results showing how close our achievable rates are to capacity by considering a geometric path-loss model in which the relay is placed on the line joining the transmitter and receiver. We find achievable rates for different positions of the relay and under a joint power constraint P (i.e., the sum of the powers used by the source and relay is equal to P).

- We apply these achievability schemes to study stability conditions for control over a relay channel with feedback.It is known that a relay helps in enhancing the capacity of a point-to-point channel. An obvious question to ask is whether the use of a relay for transmission enhances the stability region even if there is a total power constraint on the sensor and the relay? As earlier, we consider two channel models - a cascade of two point-to-point Gaussian channels and a Gaussian relay channel. We propose schemes to stabilize an unstable discrete-time linear time invariant (LTI) process over these channels and show that in both the cases, the stability region is enhanced as compared to a point-to-point channel. We thus show that it can be useful to use a relay for the control application as well.

- In order to consider network scenarios, it is also important to investigate what coding schemes to use if there are multiple paths between the source and destination. We consider a discrete-time LTI process and a remotely placed controller. The communication from the sensor to the controller is through a Gaussian product channel. We present a non-linear coding scheme based on quantization and derive stabilizability conditions. The stabilizability conditions correspond to the capacity of the product channel. The parallel result in information theory is that in order to achieve the capacity of a Gaussian product channel, you need to solve an optimization problem, that decides what amount of power to allocate to each parallel path. The channel with the least variance of additive noise gets the most power and so on. Our solution for the stabilization problem over the Gaussian product channel has a similar interpretation. The solution requires the transmitter to send the most significant bits over the best channel (corresponding to the least noise or the best Signal-to-Noise (SNR) ratio), the next most significant bits over the next best channel and so on, and send the quantization error over the last channel.

- Using tools from information theory, we show that the sufficiency conditions derived are necessary as well.

- We also consider network examples with two non-interfering relays lying on two independent paths between the source and destination. Using our coding schemes for the Gaussian relay channel with feedback and Gaussian product channel with feedback, we propose and derive achievable rate regions for different configurations. The sufficient conditions for stabilization over such channels follow immediately using the achievability results.

1.4 Outline of the Dissertation

In Chapter 2, a survey of existing literature and techniques used in later chapters is presented. We revisit the classical SK scheme [38] and show that it achieves the capacity of a point-to-point AWGN channel. We show how the SK scheme can be used to stabilize an unstable process over an AWGN link [13]. We review the stochastic approximation algorithms that we use in the later chapters to develop achievability/stability schemes for the relay channel with feedback.

In Chapter 3, we develop and analyze feedback coding schemes for two different configurations of a three node network - a cascade of two Gaussian point-to-point channels and a Gaussian relay channel. A Gaussian relay channel with destination-source feedback was considered in [8]. By contrast, in this work we consider destination-source and destination-relay feedback. Our schemes are based on the nodes running a *distributed stochastic approximation algorithm.*

In Chapter 4, we use the feedback coding schemes developed for a relay channel in Chapter 3 and apply them to stabilize a linear time invariant (LTI) plant over a Gaussian relay channel. We assume that the control action is available at the relay, and that there are average power constraints at both the transmitter and the relay nodes. We present sufficient conditions for the stabilizability of the plant through such schemes. The analysis suggests that it is useful to provide a relay node assisting the plant, even if the total transmission power remains the same.

In Chapter 5, we present necessary and sufficient conditions for stabilizing a discrete-time LTI plant in the mean squared sense when a sensor that observes the state perfectly transmits the plant state information to a remotely placed controller across a Gaussian product channel. It is known that linear coding schemes may lead to overly restrictive stabilizability conditions in such scenarios. We present a non-linear coding scheme based on ideas from distributed source-channel coding and present the resulting stabilizability conditions. If these conditions are satisfied with equality, the proposed coding scheme transmits data across the product channel at a rate equal to the capacity of the channel.

In Chapter 6, we combine the results from earlier chapters to obtain achievability results for a more general network problem. Specifically, we consider a network with two parallel cascade channels between the source and destination. A relay node lies on each path to help the source in sending the information to the destination. We look at two specific configurations of such a network, similar to the variations in Chapter 3.

In Chapter 7, we present concluding remarks and some directions for future research.

CHAPTER 2

BACKGROUND

This chapter provides a survey of relevant literature and techniques used in later chapters. We describe a scheme proposed in [38], which uses stochastic approximation algorithms to communicate over a noisy channel with perfect feedback. We will see how the SK scheme can be used to achieve capacity for point-to-point channels. We also illustrate how the SK scheme can be used to stabilize an unstable plant over a point-to-point channel. We then give a brief overview of stochastic approximation algorithms.

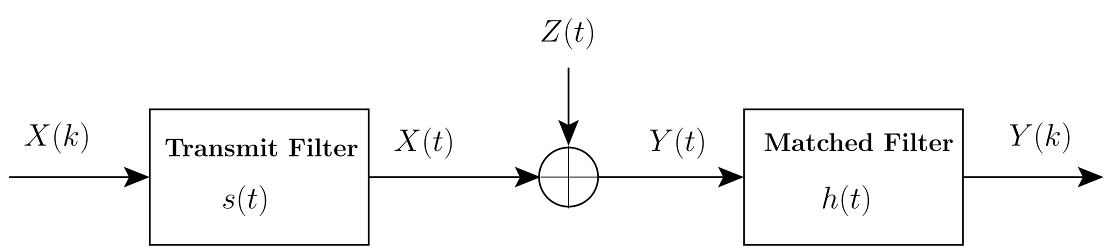

Figure 2.1. Model for an additive noise channel.

2.1 Equivalent Discrete-Time Model

As in [21, 38], to apply stochastic approximation algorithms to communication channels, we need to obtain a discrete-time equivalent of the additive white Gaussian noise (AWGN) channel shown in Fig. 2.1. We assume that information is transmitted by modulating the amplitude of a known waveform, $s(t)$, which is of unit energy and is orthogonal for integer shifts of $\Delta > 0$, i.e.,

$$\int s(t - k\Delta)s(t - j\Delta)dt = \delta_{kj}.$$

Assume that the sequence of numbers $\{X(k)\}$ needs to be transmitted over the channel. The transmitted signal will then be

$$X(t) = \sum_k X(k)s(t - k\Delta), \qquad k = 0, 1, \ldots$$

and the transmitted energy is equal to $\sum_k X^2(k)$. A matched filter, $h(t) = s(-t)$, is used at the receiver to obtain a sequence $\{Y(k) = X(k) + Z(k)\}$, where

$$Z(k) = \int Z(t)s(t - k\Delta)\, dt.$$

If $Z(t)$ is zero-mean, Gaussian and white with power spectral density σ^2, $\{Z(k)\}$ will also be zero-mean, Gaussian and white with

$$\mathbb{E}\left[Z(k)Z(j)\right] = \sigma^2\delta_{kj}.$$

The output of a matched filter for Gaussian noise channels is a sufficient statistic and therefore preserves all the information that is relevant to decision making.

Hence, for the Gaussian channel, it can be seen that the discrete-time channel thus obtained, where a sequence $X(k)$ is transmitted and the sequence $Y(k) = X(k) + Z(k)$ is received, is equivalent to the original continuous-time channel.

2.2 Schalkwijk Kailath Scheme and Extensions

The SK scheme [38] can be employed for communication over a point-to-point channel with perfect feedback, as shown in Fig. 2.2. The forward channel from the source to destination is corrupted by AWGN, whereas the feedback link is modeled as noiseless. The signals transmitted on the forward link have an average power constraint P, i.e., $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{E}[X^2(k)] \leq P$. The input and output of the channel is denoted by X and Y respectively. The noise corrupting the channel is denoted by Z. The output of the channel at time k is given by

$$Y(k) = gX(k) + Z(k),$$

where g is attenuation due to path loss. The noise $Z(k)$ is modeled by a zero-mean Additive White Gaussian Noise (AWGN) process with mean zero and variance σ^2. The encoder maps the message to one of M disjoint, equal-length subintervals of the interval $[0, 1]$ on the real line. The message point to be transmitted corresponds to the midpoint W of a particular message interval. For sufficiently large number of messages, the variance of the message point is $\sigma_W^2 = \frac{1}{12}$. The decoder calculates and stores an estimate $\hat{W}(k)$. Note that the variance of the estimation error $\hat{W}(k) - W$ is denoted by $\alpha(k)$.

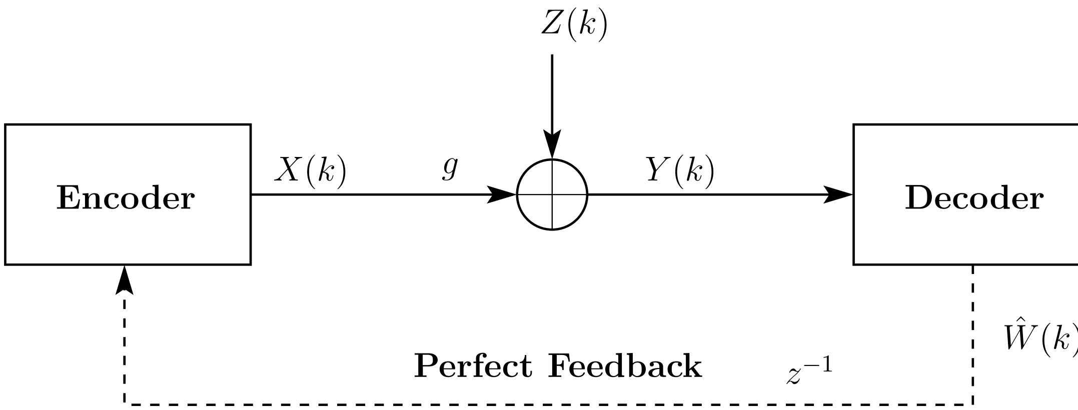

Figure 2.2. Communication over a AWGN channel with perfect feedback.

2.2.1 Coding Scheme

The scheme works as follows. At each time step, the encoder sends the estimation error scaled appropriately to satisfy the power constraints. Note that the encoder knows the estimate at the decoder through the feedback link. The decoder refines its estimate of the message point as it receivers more transmissions from the encoder. Now we describe the scheme in detail.

- *Initialization:* At time step $k = 0$, the encoder transmits the input $X(0)$ given by

$$X(0) = \sqrt{\frac{P}{\sigma_W^2}} W. \tag{2.1}$$

  The decoder receives $Y(0) = gX(0) + Z(0)$ and forms an estimate of W by scaling as follows:

$$\hat{W}(0) = \frac{1}{g}\sqrt{\frac{\sigma_W^2}{P}} Y(0).$$

The estimation error $\epsilon(0)$ is given by

$$\epsilon(0) = \frac{1}{g}\sqrt{\frac{\sigma_W^2}{P}}Z(0).$$

Clearly, $\epsilon(0)$ is zero-mean Gaussian with variance $\alpha(0)$, given by

$$\alpha(0) = \frac{\sigma_W^2 \sigma^2}{g^2 P}. \tag{2.2}$$

The decoder sends the estimate $\hat{W}(0)$ to the encoder.

- *Update:* At each time step $k \geq 1$, the encoder transmits

$$\begin{aligned} X(k) &= \sqrt{\frac{P}{\alpha(k-1)}}\left(\hat{W}(k-1) - W\right) \\ &= \sqrt{\frac{P}{\alpha(k-1)}}\epsilon(k-1). \end{aligned} \tag{2.3}$$

The decoder updates its estimate as follows. At time $k \geq 1$, the decoder calculates the linear minimum mean squared error (MMSE) estimate of W given $Y(k)$ and $\hat{W}(k-1)$ as

$$\hat{W}(k) = \hat{W}(k-1) - \frac{\mathbb{E}[Y(k)\epsilon(k-1)]}{\mathbb{E}[Y^2(k)]}Y(k). \tag{2.4}$$

The decoder sends the estimate $\hat{W}(k)$ to the encoder. Note that the input $X(k)$ satisfies the respective power constraint and that subsequent transmissions are orthogonal to each other (since the MMSE estimation error $\epsilon(k)$ is orthogonal to all observations).

It can be seen that the estimation error $\epsilon(k)$ is Gaussian with zero mean and variance $\alpha(k)$. We now proceed to evaluate a recursive expression for

$\alpha(k)$ as used in the coding scheme presented above. Since $\epsilon(k)$ is defined as $\hat{W}(k) - W$, from (2.4) we obtain

$$\epsilon(k) = \epsilon(k-1) - \frac{\mathbb{E}[Y(k)\epsilon(k-1)]}{\mathbb{E}[Y^2(k)]} Y(k). \tag{2.5}$$

The variance of $\epsilon(k)$ can be obtained as

$$\begin{aligned} \alpha(k) &= \mathbb{E}[\epsilon^2(k)] \\ &= \alpha(k-1) - \frac{\mathbb{E}^2[Y(k)\epsilon(k-1)]}{\mathbb{E}[Y^2(k)]}, \end{aligned} \tag{2.6}$$

with the initial condition in (2.2). The terms in (2.6) can be further evaluated to be

$$\mathbb{E}[Y^2(k)] = g^2 P + \sigma^2, \tag{2.7}$$

and

$$\mathbb{E}[Y(k)\epsilon(k-1)] = g\sqrt{P\alpha(k-1)}. \tag{2.8}$$

Using (2.7) and (2.8) in (2.6), we obtain

$$\alpha(k) = \alpha(k-1)r, \tag{2.9}$$

where $r = \left(\frac{\sigma^2}{g^2 P + \sigma^2}\right)$. Thus,

$$\alpha(k) = \frac{\sigma_W^2 \sigma^2}{g^2 P} \left(\frac{\sigma^2}{g^2 P + \sigma^2}\right)^k. \tag{2.10}$$

Now, we see how this coding scheme can be used for achievability and stability over the AWGN channel.

2.2.2 Achievability for Point-to-Point Channels

The number of messages M can be expressed in terms of the rate R as $M = e^{nR}$. At time $k = n$, the decoder's estimate of the message point W is given by $\hat{W}(n) = W + \epsilon(n)$. The decoder decodes the message as the message point closest to $\hat{W}(n)$. A decoding error occurs if there is a message point $W' \neq W$ closer to $\hat{W}(n)$ than W. The decoding error probability $P_e(n)$ is upper bounded by the probability that the magnitude of $\epsilon(n)$ is greater than half the distance between adjacent points.

$$\begin{aligned} P_e(n) &\leq Pr\left[|\epsilon(n)| > \frac{1}{2(M-1)}\right] \\ &\leq 2Q\left(\frac{1}{2M\sqrt{\alpha(n)}}\right) \\ &= 2Q\left(\frac{e^{n\left(\frac{1}{2}\ln\left(1+\frac{g^2P}{\sigma^2}\right)-R\right)}}{2\sqrt{\alpha(0)}}\right), \end{aligned} \tag{2.11}$$

where $Q(x) := \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp(-\frac{y^2}{2})\,\mathrm{d}y$. The following theorem is the main result for achievability over point-to-point channels [38].

Theorem 2.2.1 *For the point-to-point channel in Figure 2.2, the coding scheme presented in Section 2.2.1 achieves a rate*

$$R < \frac{1}{2}\ln\left(1 + \frac{g^2P}{\sigma^2}\right). \tag{2.12}$$

Proof: From (2.11), we can see that $P_e \to 0$ as $n \to \infty$ if the condition in (2.12) is satisfied. Since $Q(x) \sim \exp(\frac{-x^2}{2})$ for large x, for all rates satisfying (2.12), the error decay is doubly exponential in n.

2.2.3 Extensions of the Achievability Results

Motivated by the advantages of feedback, a large body of work has looked at extending the Schalkwijk-Kailath (SK) scheme to more general cases. Ozarow extended the SK scheme to the two-user multiple access channel (MAC) [34] and the broadcast channel [33]. Infact, [34] established that capacity region for the two-user memoryless MAC with feedback, unlike for the point-to-point channel, is strictly larger than the one without feedback. A common theme of these extensions is to encode by mapping the message to a point on the real line as for the point-to-point AWGN case and then to update the receiver's linear minimum mean squared estimate at each channel use. The transmitter knows the channel outputs via the feedback channel and sends power-normalized versions of the estimation errors. The coding scheme for MAC also uses a phase modulation technique. One of the transmitters operates without any modulation, while the other transmitter sends its power normalized estimation errors multiplied by -1 every other time steps. The purpose of this phase modulation is to align the phases of the inputs so as to increase the mutual information between the channel input and output symbol sequences. The capacity region of the n-user MAC is not known in general. The maximum sum-rate achieved by linear-feedback codes under symmetric power constraints for the n-user MAC was characterized in [3]. Ehsan *et al.* extended Ozarow's scheme to the k-receiver AWGN broadcast channel with feedback using tools from the theory of linear quadratic Gaussian optimal control [2]. Achievability over Gaussian cascade and relay channel for the infinite bandwidth constraint was presented in [21]. However, extensions for the general discrete cascade and relay channel leading to results and coding schemes for the finite bandwidth constraint are largely missing and is the focus of one of the later chapters (Chapter 3).

Kramer extended Ozarow's scheme to interference networks [20]. Merhav and Weissman [29] have extended the "dirty-paper" coding scheme to the case of feedback. More general channels than an AWGN channel have also been considered. Shayevitz and Feder [42, 43] recently provided an interpretation of the SK scheme that allows construction of related schemes for arbitrary channels. In particular, it unifies the classic scheme of Horstein [17] for binary symmetric channels and the SK scheme for AWGN channels.

There has also been a lot of work on studying the impact of noise on the performance of SK-like schemes. For the feedback link corrupted by AWGN, [38] discusses two possible approaches for scenarios in which the feedback link is noisy. In the first strategy, the decoder transmits its estimate $\hat{W}$ on the feedback link. Although the probability of error at the decoder decreases to zero, the rate relative to the capacity of the channel approaches zero simultaneously. In the second strategy, the decoder transmits the output $Y(k)$ of the forward channel on the feedback link. In this case, while the relative rate approaches unity, the mean square error does not reduce to zero. In fact, as [19] shows, as long as the encoding scheme is linear, feeding back the output of the channel cannot result in positive achievable rates with a vanishing probability of error. The recent work in [28] extends the SK scheme to handle feedback noise of finite support, (e.g., quantization noise) by utilizing some ideas from the Witsenhausen counterexample in distributed control [46]. In particular, their modification of the Schalkwijk-Kailath scheme provides a positive rate and a vanishing probability of error, in spite of the noise in the feedback link. Concatenated coding with the feedback code as the inner code and LDPC as the outer code has also been investigated [10].

2.2.4 Stability over Point-to-Point Channels

Consider the set-up shown in Fig. 2.3. The plant is described by an open loop unstable scalar linear time invariant process evolving as

$$S(k+1) = aS(k) + U(k), \tag{2.13}$$

where $S(k) \in \mathbb{R}$ is the state and $U(k) \in \mathbb{R}$ is the control value. We assume that the initial condition $S(0)$ is a random variable with an arbitrary distribution but a finite variance $\sigma^2_{S(0)}$. For ease of exposition, and without loss of generality, we assume that at every time step a sensor observes the state of the process $S(k)$ and transmits information across the communication channel to the controller. The controller calculates an estimate $\hat{S}(k)$ of the initial condition $S(0)$, calculates a control input $U(k)$ and applies it to the process in (2.13). The communication channel from the plant to the controller is modeled as an additive noise channel, while the communication from the controller to the process is assumed to be perfect.

The problem we are interested in this section is to design a transmission scheme and controller $U(k)$ so that the process (2.13) is mean square stabilized.

Definition 2.2.2 *A system is said to be stabilized in the mean squared sense if and only if irrespective of the initial state $S(0)$, the following conditions are satisfied:*

$$\begin{aligned} \mathbb{E}[S(k)] &= 0, \\ \lim_{k\to\infty} \mathbb{E}[S^2(k)] &= 0. \end{aligned} \tag{2.14}$$

The decoder calculates and stores an estimate $\hat{S}(k)$ of the initial condition

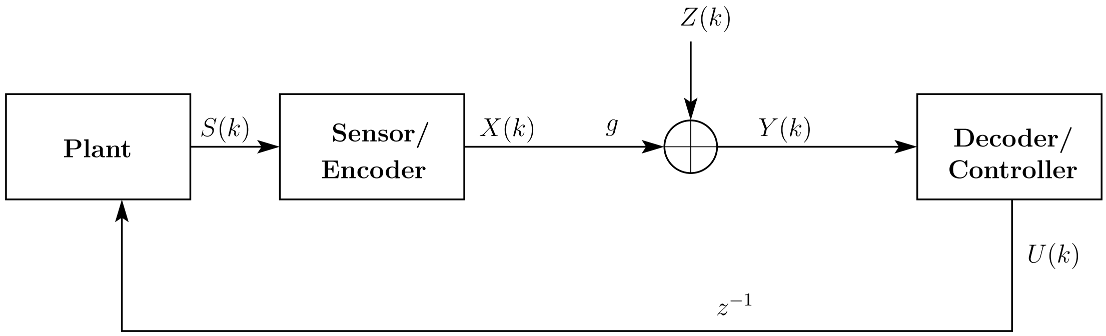

Figure 2.3. Problem setup for an unstable plant being controlled across an AWGN point-to-point channel.

$S(0)$. We define the estimation error as

$$\epsilon(k) := \hat{S}(k) - S(0).$$

The crucial property that needs to be satisfied for mean square stability is that the estimate $\hat{S}(k)$ converges to $S(0)$ at a rapid enough rate as shown by the following result [7, 45].

Lemma 2.2.3 *The LTI system in (2.13) can be mean square stabilized over a communication channel if the following conditions are satisfied:*

$$\begin{aligned} \mathbb{E}[\epsilon(k)] &= 0, \\ \lim_{k\to\infty} a^{2k}\mathbb{E}[\epsilon^2(k)]] &= 0. \end{aligned} \tag{2.15}$$

Proof: Since the controller does not the know the state value $S(0)$ exactly, the controller takes actions using the estimates $\hat{S}(k)$. The controller actions are defined as

$$U(k) := \begin{cases} -a\hat{S}(0), & k = 0 \\ -a^{k+1}(\hat{S}(k) - \hat{S}(k-1)), & k \geq 1. \end{cases} \tag{2.16}$$

Thus, the state $S(k)$ evolves as

$$S(k+1) = -a^{k+1}\left(\hat{S}(k) - S(0)\right) = -a^{k+1}\epsilon(k). \tag{2.17}$$

The mean value and mean squared value of $S(k+1)$ is given by

$$\mathbb{E}[S(k+1)] = -a^{k+1}\mathbb{E}[\epsilon(k)],$$
$$\mathbb{E}[S^2(k+1)] = a^{2(k+1)}\mathbb{E}[\epsilon^2(k)].$$

Thus, if the conditions in (2.15) are satisfied, the process is mean square stabilized. The above result also presents the controller design. We can use the coding scheme in Section 2.2.1 to send $W = S(0)$. Since we represent the estimate of the state as $\hat{S}(k)$ here, $\hat{S}(k) = \hat{W}(k)$. We now present the stability conditions if the coding scheme described above is used to stabilize the process (2.13) [7, 45].

Theorem 2.2.4 *Consider the problem formulation presented in Section 2.2.4 with the coding scheme presented above in Section 2.2.1. The process (2.13) is mean square stabilized over the point-to-point channel if*

$$\ln(a) < \frac{1}{2}\ln\left(1 + \frac{g^2P}{\sigma^2}\right). \tag{2.18}$$

Proof: It is easy to see that $\mathbb{E}[\epsilon(0)] = 0$. It is known that the linear minimum mean squared error is an unbiased estimator. Thus, $\mathbb{E}[\epsilon(k)] = 0$ for all $k \geq 0$, which is the first condition in (2.15). For the second condition, from (2.10) we can see that

$$a^{2k}\alpha(k) = a^{2k}\alpha(0)\left(\frac{\sigma^2}{g^2P+\sigma^2}\right)^k.$$

If the condition in (2.18) is satisfied, then $a^{2k}\alpha(n) \to 0$ and mean square stability is obtained.

Note that the right hand side of the condition in (2.18) is also the maximum rate at which information can be transmitted over a Gaussian point-to-point channel. This is the *data rate theorem* for the point to point channel, that tells what is the minimum power that should be used to stabilize a unstable plant over a point-to-point channel.

2.2.5 Extensions of the Stability Results

The presence of various communication channel models in the control loop has been considered, including channels that introduce data loss (e.g., [16]), delay (e.g., [26]), digital noiseless channels (e.g., [30, 32]), and additive white Gaussian noise channels (e.g., [7]). The works [27, 36] have also considered stabilizability conditions and performance limitations induced by arbitrary communication channels. Stability conditions in the presence of one additive white Gaussian noise (AWGN) channel are available (e.g., [7, 27, 45]). Stabilizing the plant across a broadcast and multiple access channel [51] have also been considered. Interesting parallels of the problem with schemes achieving the capacity of a Gaussian channel with feedback through the Schalkwijk-Kailath (SK) scheme [37, 38] are known [13]. However, to study stabilizability over general Gaussian networks, it is important to study extensions of the SK scheme for multihop scenarios and parallel channels, which is largely missing. We use schemes from statistics called stochastic approximation algorithms to develop coding schemes for the cascade and relay channel in Chapter 3. We provide a brief summary of stochastic approximation in the next section.

2.3 Stochastic Approximation

Stochastic approximation algorithms have a long history [6, 25]. This section reviews stochastic approximation, which are algorithms used to estimate unknown parameters recursively based on the observation data. All these parameter estimation problems can be transformed to the problem of root-finding of an unknown function. Let $Y(k)$ be the observation data and $\hat{T}(k)$ be the estimate of the parameter for the k-th iteration. It can be assumed that the parameter under estimation denoted by W is the root of an unknown function $f(\cdot)$. This is not restrictive since a simple function like $x - W$ can be used.

Robbins and Munro proposed the following algorithm [35]. An oracle (or a black box) provides the value of the function $f(\hat{T}(k))$ evaluated at the point $\hat{T}(k)$, possibly corrupted by a noise component $Z(k)$. The algorithm consists of an iteration of the form

$$\hat{T}(k) = \hat{T}(k-1) - \beta(k)\left(f(\hat{T}(k-1)) + Z(k)\right), \tag{2.19}$$
$$\hat{T}(0) = 0,$$

where $\beta(k)$ is the step size applied at time k satisfying the following set of conditions:

$$\beta(k) > 0, \quad \beta(k) \to 0, \quad \sum_k \beta(k) = \infty. \tag{2.20}$$

It can be proven that $\hat{T}(k)$ converges almost surely to the zero of the function f

as $k \to \infty$, provided the function f is non-decreasing. $Z(k)$'s are independent, zero mean and finite variance satisfies the following conditions [25, 38]:

1. $f(x) \gtreqless 0$ according to $x \gtreqless W$.
2. $\inf\{|f(x)| : \epsilon < |x - W| < 1/\epsilon\} > 0$ for all $0 < \epsilon < 1$.
3. $|f(x)| \leq K|x - W| + K_2$ where K and K_2 are positive constants.
4. $\sup_{k>0} \mathbb{E}\left[Z^2(k)\right] < \infty$.

The conditions on $\beta(k)$ presented above are very restrictive. However, the conditions on $\beta(k)$ can be relaxed based on different applications. For example, the SK scheme described in Section 2.2 maps directly to a stochastic approximation algorithm with $f(x) = \alpha(x - W)$ and the step size $\beta(k)$ corresponding to the coefficients for linear MMSE estimation.

2.4 Distributed Stochastic Approximation (DSA)

The coding scheme we propose for relay and cascade channels is based on distributed stochastic approximation. Distributed stochastic approximation [6] is a distributed method for calculating the zeros of a vector function $\mathbf{f} : \mathbb{R}^m \to \mathbb{R}^m$ from noisy observations of the function. Let $f_i : \mathbb{R}^m \to \mathbb{R}$ denote the i-th component of the function and $\hat{\mathbf{T}}(k)$ denote the current estimate of the zero. The method works as follows. Noisy observations of the form $\tilde{f}_i(\hat{\mathbf{T}}(k)) = f_i(\hat{\mathbf{T}}(k)) + N_i(k)$ are obtained and used to update each component of the estimate $\hat{\mathbf{T}}(k) = [\hat{T}_1(k)\hat{T}_2(k)\ldots\hat{T}_m(k)]$ as

$$\hat{T}_i(k) = \hat{T}_i(k-1) - \beta_i(k)\tilde{f}_i(\hat{\mathbf{T}}(k)),\ i = 1, 2, \ldots, m, \tag{2.21}$$

$$\hat{T}_i(0) = 0,$$

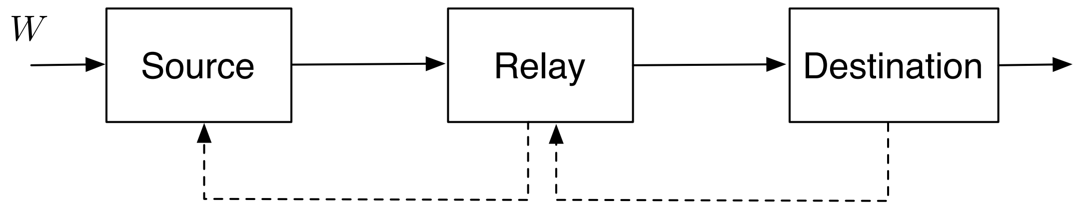

Figure 2.4. A cascade of two point-to-point channels.

$\beta_i(k)$ is the step size as before.

We use this distributed approach to stochastic approximation to extend the SK-scheme to a cascade of point-to-point channels and a relay channel with feedback in Chapter 3. Based on the network topology, we can define a vector function, such that each variable $\hat{T}_i(k)$ of the function corresponds to the estimate at a particular node i in the network. Each dimension j of the function corresponds to an observation at the j-th node in the network. For example, the function for the cascade of point-to-point channels in Fig. 2.4 can be defined as follows:

$$f_i\left(T_1, T_2, T_3\right) = \begin{bmatrix} 0 \\ T_2 - T_1 \\ T_3 - T_2 \end{bmatrix},$$

where variables T_1, T_2 and T_3 correspond to the source, relay and destination respectively. Now if $T_1 = W$, then it can be seen that the zero of this function is $T_1 = T_2 = T_3 = W$. Thus, this vector function gives insights on how to extend the SK scheme to the cascade case. The relay should see a noisy version of the second dimension of the function and destination should see a noisy version of the third dimension of the function. This also gives insights on what inputs each channel should have. More details can be found in the Chapter 3.

2.5 Summary

This chapter provided a survey of relevant tools and techniques used in the dissertation. We described the SK scheme, that uses stochastic approximation algorithms to communicate over a AWGN channel with perfect feedback. We provided a brief overview of stochastic approximation algorithms and show how they the SK scheme maps to a stochastic approximation scheme. We also briefly discuss the distributed stochastic approximation (DSA) algorithm, that is used to extend the SK scheme to a cascade and relay channel with feedback in Chapter 3. We use the coding schemes with feedback developed using DSA to study stabilizability over a relay channel in Chapter 4. We use a quantization-based scheme to extend the SK scheme to develop a coding scheme for the Gaussian product channel with feedback in Chapter 5.

CHAPTER 3

ACHIEVABILITY OVER GAUSSIAN CASCADE AND RELAY CHANNEL

In this chapter, we present and analyze feedback coding schemes for two different configurations of a three node Gaussian networks - a cascade of two point-to-point channels and a relay channel. A Gaussian relay channel with destination-source feedback is considered in [8]. By contrast, in this work we consider destination-source and destination-relay feedback. A similar scheme was presented in [23] where the bandwidth was assumed to be infinite. On the contrary, we present achievable schemes for Gaussian discrete-time channels.

Our schemes are based on mapping the data to be transmitted by the source to a real number and allowing the nodes to run a *distributed stochastic approximation (DSA)* algorithm [6]. As the algorithm evolves, both the relay and destination converge to the message point. In addition to being simple, the schemes provide a doubly exponential error decay in the case of perfect feedback. We achieve rates equal to the capacity for a cascade channel and rates close to capacity for the relay channel with feedback.

The remainder of the chapter is organized as follows. In Section 3.1 we present our system model. In Section 3.2, we briefly discuss the function we use for distributed stochastic approximation algorithms. In Sections 3.3 and 3.4, we present our communication schemes for the cascade and relay channel. we briefly discuss how we can use the schemes proposed in this chapter to obtain achievable rates for the half-duplex relay channel in Section 3.5. We present some concluding remarks in Section 3.6.

3.1 System Model

Consider a source that needs to transmit data to a destination with the assistance of a relay as shown in Fig. 3.1. We adopt the following notation:

- The source, relay, and the destination are denoted by the nodes 1, 2, and 3 respectively
- The input (resp. observation) at node i is denoted by X_i (resp. Y_i)
- The noise corrupting the reception at the relay (resp. destination) is denoted by Z_2 (resp. Z_3)

We assume that the forward and feedback communications take place over orthogonal channels so that they do not interfere with each other.

We consider the following scenario:

- The forward links are modeled by zero-mean AWGN channels. We consider the noises on the links to be mutually independent and white.
- The signal from node i to node j is scaled by a distance-dependent attenuation factor $g_{i,j}$ that depends on the distance $d_{i,j}$ between nodes i and j as

$$g_{i,j} = \begin{cases} bd_{i,j}^{-\eta/2}, & i \neq j, \\ 0, & i = j, \end{cases} \tag{3.1}$$

where η is the path-loss exponent (typically taking values between 2 and 4), and b is a constant. To simplify our analysis, we will assume $b = 1$. A

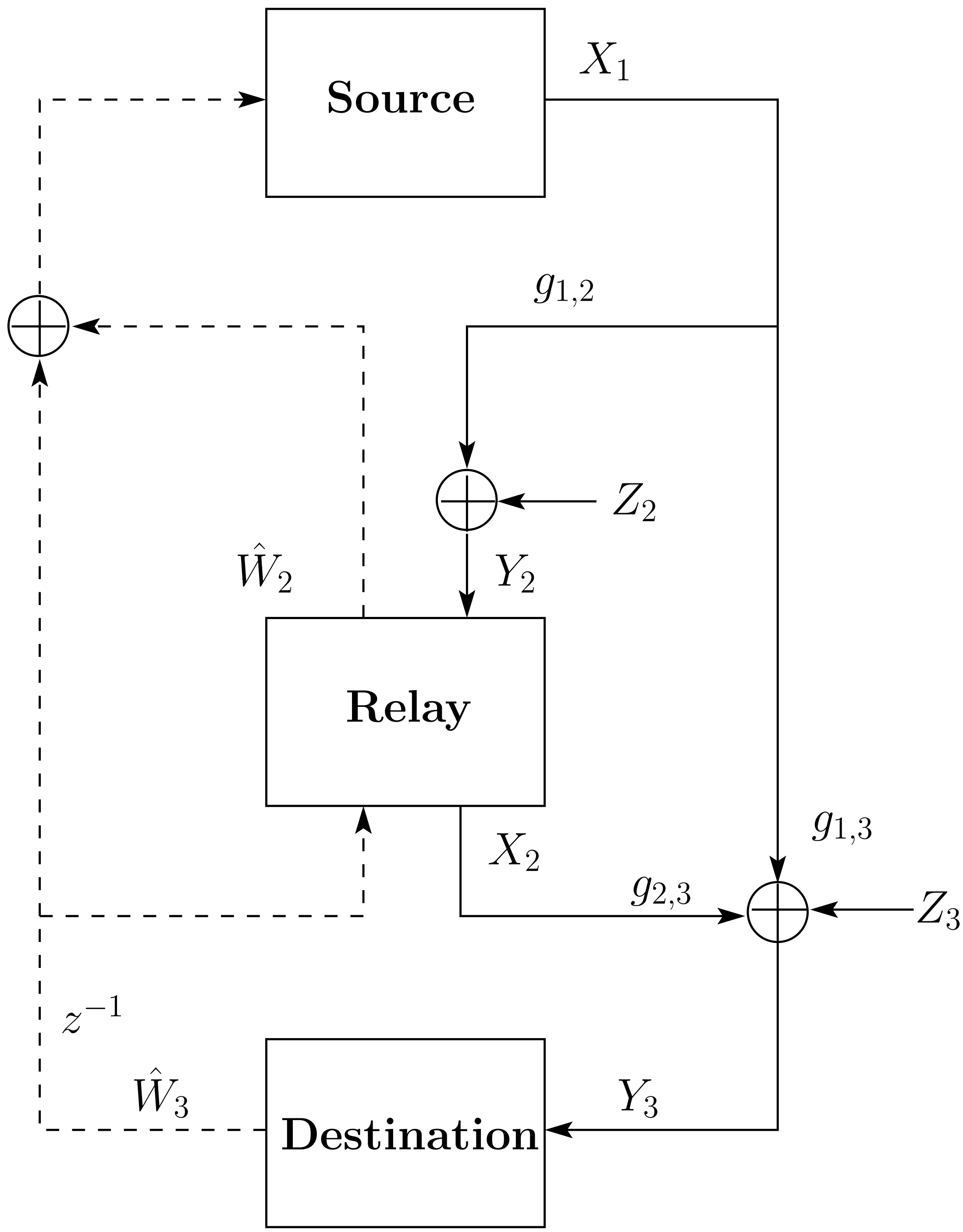

Figure 3.1. Block diagram for a network of 3 nodes.

different choice of b will only result in scaling of the received power identically for all schemes and hence would not affect the performance comparison among them. Every node knows the distance of the other two nodes from itself and hence knows the attenuation factors. If there is no link from node i to node j, then $g_{i,j} = 0$. The observation vector for the forward communication can be written as

- The signals transmitted on the forward links by the source and relay must satisfy average power constraints P_1 and P_2 respectively, i.e.,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{E}[X_i^2(k)] \leq P_i, i = 1, 2$$

.

- The source maps the message to one of M disjoint, equal-length subintervals of the interval $[0, 1]$ on the real line. The message to be transmitted corresponds to the midpoint W of a particular message interval.

- Each node i updates and stores an estimate $\hat{W}_i(k)$ of W recursively and feeds back its estimate $\hat{W}_i(k)$ on the feedback link. The observation vector for the feedback channel can be written as

$$\begin{bmatrix} Y_{3,2}(k) \\ Y_1(k) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} X_3(k) \\ X_{2,1}(k) \end{bmatrix} .+ \begin{bmatrix} Z_{3,2}(k) \\ Z_1(k) \end{bmatrix}. \tag{3.2}$$

3.2 Distributed Stochastic Approximation (DSA)

The coding scheme we propose is based on distributed stochastic approximation. In our design, every node i updates and stores an estimate $\hat{W}_i(k)$ of the message point W. Since node 1 has perfect knowledge of the message point W, the estimate $\hat{W}_1(k) = W$ identically. We define the following estimation errors:

$$\epsilon_2(k) := \hat{W}_2(k) - W,$$
$$\epsilon_3(k) := \hat{W}_3(k) - W,$$
$$\epsilon_{3,2}(k) := \hat{W}_3(k) - \hat{W}_2(k).$$

The basic idea of the coding scheme is as follows. The source, relay, and the destination implement a distributed stochastic approximation scheme to calculate the zeros of the function $\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^3$, the i-th component of which is given by

$$f_i\left(T_1, T_2, T_3\right) = \begin{cases} 0 & \text{for } i=1 \\ \sum_{j:\ g_{j,i} \neq 0} \mu_i f_{i,j}(T_i, T_j) & \text{for } i=2,3, \end{cases} \tag{3.3}$$

where $f_{i,j}(T_i, T_j) := (T_i - T_j)$, the variable T_i is associated with the i-th node, $T_1 = W$, and μ_i is a term suitably designed to satisfy the transmission power constraints. The estimates $\hat{W}_2(k)$ and $\hat{W}_3(k)$ of the variables T_2 and T_3 are updated as in (2.21), while the estimate $\hat{W}_1(k)$ of the variable T_1 is fixed at the message W, i.e., $\hat{W}_1(k) = W$ identically. Thus, the zero of the function $\mathbf{f}$ is attained by all estimates $\hat{W}_i$'s converging to W. In other words, the distributed stochastic approximation algorithm would eventually yield the estimate at the destination being equal to the initial state W.

3.3 A Cascade of Point-to-Point Channels

Consider a cascade of two point-to-point channels with feedback, i.e., $g_{1,3} = 0$ and relay-source and destination-relay feedback. We know that the capacity of a memoryless noisy channel is not increased by noiseless feedback [41]. Thus, the capacity $C_{i,i+1}$ of a forward link from node i to $i+1$ corrupted by AWGN is given by

$$C_{i,i+1} = \frac{1}{2} \ln \left(1 + \frac{g_{i,i+1}^2 P_i}{\sigma_{i,i+1}^2} \right),$$

where P_i is the average power constraint at node i. The capacity C of the cascade is given by the max-flow min-cut bound [12],

$$C = \min\{C_{1,2}, C_{2,3}\}.$$

3.3.1 Coding Scheme

We now proceed to explain the coding scheme used at the source and the relay. The main idea is that at time step k both the source and the relay send innovations with respect to the estimate at the relay and destination respectively. The code works as follows.

- *Initialization:*
  - At time step $k = 0$, the source node 1 sends the input $X_1(0)$ given by

$$X_1(0) = \sqrt{\frac{P_1}{\sigma_W^2}} W.$$

  - The relay node 2 sends nothing, while it receives

$$Y_2(0) = g_{1,2} X_1(0) + Z_2(0).$$

It then forms an estimate of W by scaling as follows:

$$\hat{W}_2(0) = \frac{1}{g_{1,2}}\sqrt{\frac{\sigma_W^2}{P_1}}Y_2(0).$$

The estimation error $\epsilon_2(0)$ is given by

$$\epsilon_2(0) = \frac{1}{g_{1,2}}\sqrt{\frac{\sigma_W^2}{P_1}}Z_2(0).$$

Clearly, $\epsilon_2(0)$ is zero-mean Gaussian with variance $\alpha_2(0)$, given by

$$\alpha_2(0) = \frac{\sigma_W^2\sigma_2^2}{g_{1,2}^2P_1}. \tag{3.4}$$

The relay sends its estimate $\hat{W}_2(0)$ to the source node.

– At time step $k = 1$, the relay node 1 sends the input $X_1(1)$ given by

$$X_2(1) = \sqrt{\frac{P_2}{\alpha_2(0)}}\hat{W}_2(0).$$

– The destination node 2 receives

$$Y_3(1) = g_{2,3}X_2(1) + Z_3(1).$$

It then forms an estimate of W by scaling as follows:

$$\hat{W}_3(1) = \frac{1}{g_{2,3}}\sqrt{\frac{\alpha_2(0)}{P_2}}Y_3(1).$$

The estimation error $\epsilon_3(1)$ is given by

$$\epsilon_3(1) = \epsilon_2(0) + \frac{1}{g_{2,3}}\sqrt{\frac{\alpha_2(0)}{P_2}}Z_3(1).$$

Clearly, $\epsilon_3(1)$ is zero-mean Gaussian with variance $\alpha_3(1)$, given by

$$\alpha_3(1) = \alpha_2(0)\frac{g_{2,3}^2 P_2 + \sigma_3^2}{g_{2,3}^2 P_2}. \tag{3.5}$$

The destination node sends its estimate $\hat{W}_3(0)$ to the relay node.

- *Update:* At each time step $k \geq 1$, each transmitting node j transmits $f_{i,j}(\hat{W}_i, \hat{W}_j)$ from (3.3) on the forward link.

  – Using the feedback from the relay (resp. destination) at time $k-1$, the source (resp. relay) knows $\hat{W}_2(k-1)$ (resp. $\hat{W}_3(k-1)$). Moreover, the source can calculate $\epsilon_2(k)$, while the relay can calculate $\epsilon_{3,2}(k)$. Note that by the construction that follows, $\epsilon_3(k)$ and $\epsilon_{3,2}(k)$ are jointly Gaussian with zero means, respective variances $\alpha_3(k)$ and $\alpha_{3,2}(k)$ and correlation coefficient $\rho(k)$ given by

$$\rho(k) = \frac{\mathbb{E}[\epsilon_3(k)\epsilon_{3,2}(k)]}{\sqrt{\alpha_3(k)\alpha_{3,2}(k)}}. \tag{3.6}$$

  The quantities $\alpha_2(k)$, $\alpha_3(k), \alpha_{3,2}(k)$ and $\rho(k)$ are deterministic and can be calculated prior to transmission by all the nodes. In particular,

$$\begin{aligned}\alpha_{3,2}(1) &= \alpha_3(1) + \alpha_2(1) - 2\mathbb{E}[\epsilon_3(1)\epsilon_2(1)] \\ &= \alpha_3(1) - \alpha_2(1), \\ \rho(1) &= \frac{\alpha_3(1) - \alpha_2(1)}{\sqrt{(\alpha_3(1) - \alpha_2(1))\alpha_3(1)}}.\end{aligned}$$

We present recursive expressions for these quantities in Section 3.3.2, and assume for the rest of the presentation that these quantities can be calculated by all the nodes.

– At $k \geq 1$, the source transmits

$$\begin{aligned}X_1(k) &= \sqrt{\frac{P_1}{\alpha_2(k-1)}}\left(\hat{W}_2(k-1) - W\right) \\ &= \sqrt{\frac{P_1}{\alpha_2(k-1)}}\epsilon_2(k-1). \end{aligned} \tag{3.7}$$

– At $k \geq 2$, the relay transmits

$$\begin{aligned}X_2(k) &= \sqrt{\frac{P_2}{\alpha_{3,2}(k-1)}}\left(\hat{W}_3(k-1) - \hat{W}_2(k-1)\right) \\ &= \sqrt{\frac{P_2}{\alpha_{3,2}(k-1)}}\epsilon_{3,2}(k-1). \end{aligned} \tag{3.8}$$

Moreover, it updates its estimate as follows. At time $k \geq 1$, the relay calculates the linear minimum mean squared error estimate of W given $Y_2(k)$ and $\hat{W}_2(k-1)$ as

$$\hat{W}_2(k) = \hat{W}_2(k-1) - \frac{\mathbb{E}[Y_2(k)\epsilon_2(k-1)]}{\mathbb{E}[{Y_2}^2(k)]}Y_2(k). \tag{3.9}$$

Notice that the step size $\beta_2(k)$ used by the relay in the distributed stochastic approximation iteration (2.21) is given by $\beta_2(k) = \frac{\mathbb{E}[Y_2(k)\epsilon_2(k-1)]}{\mathbb{E}[{Y_2}^2(k)]}$.

– At time $k \geq 2$, the destination calculates the linear minimum mean squared error estimate of W given $Y_3(k)$ and $\hat{W}_3(k-1)$ as

$$\hat{W}_3(k)=\hat{W}_3(k-1)-\frac{\mathbb{E}[Y_3(k)\epsilon_3(k-1)]}{\mathbb{E}[{Y_3}^2(k)]}Y_3(k). \tag{3.10}$$

The step size $\beta_3(k)$ used by the destination is thus $\beta_3(k) = \frac{\mathbb{E}[Y_3(k)\epsilon_3(k-1)]}{\mathbb{E}[{Y_3}^2(k)]}$. The destination then sends back its estimate $\hat{W}_3(k)$ on the feedback link.

Note that the inputs $X_i(k)$ to the various channels satisfy their respective power constraints.

3.3.2 Calculation of the Variances

We now proceed to evaluate the recursive expressions of $\alpha_2(k)$, $\alpha_3(k)$, $\alpha_{3,2}(k)$, and $\rho(k)$ as used in the coding scheme presented above. The following recursions do not depend on the data, and can be executed by any node.

- *Variance $\alpha_2(k)$ of the error at the relay node:* Since $\epsilon_2(k)$ is defined as $\hat{W}_2(k) - W$, from (3.9) we obtain

$$\epsilon_2(k) = \epsilon_2(k-1) - \frac{\mathbb{E}[Y_2(k)\epsilon_2(k-1)]}{\mathbb{E}[{Y_2}^2(k)]}Y_2(k). \tag{3.11}$$

The variance of $\epsilon_2(k)$ can be obtained as

$$\begin{aligned}\alpha_2(k) &= \mathbb{E}[\epsilon_2^2(k)] \\ &= \alpha_2(k-1) - \frac{\mathbb{E}^2[Y_2(k)\epsilon_2(k-1)]}{\mathbb{E}[{Y_2}^2(k)]},\end{aligned} \tag{3.12}$$

with the initial condition in (5.7). The terms in (3.12) can be further evaluated to be

$$\mathbb{E}[{Y_2}^2(k)] = g_{1,2}^2 P_1 + \sigma_2^2, \tag{3.13}$$

and

$$\mathbb{E}[Y_2(k)\epsilon_2(k-1)] = g_{1,2}\sqrt{P_1\alpha_2(k-1)}. \tag{3.14}$$

Substituting (3.13) and (3.14) into (3.12), we get

$$\alpha_2(k) = \alpha_2(k-1)r, \tag{3.15}$$

where $r = \left(\frac{\sigma_2^2}{g_{1,2}^2 P_1 + \sigma_2^2}\right)$.

- *Variance $\alpha_3(k)$ of the error at the destination node:* Since $\epsilon_3(k)$ is defined as $\hat{W}_3(k) - W$, from (3.10) we obtain

$$\epsilon_3(k) = \epsilon_3(k-1) - \frac{\mathbb{E}[Y_3(k)\epsilon_3(k-1)]}{\mathbb{E}[{Y_3}^2(k)]} Y_3(k). \tag{3.16}$$

The variance of $\epsilon_3(k)$ can be obtained as

$$\begin{aligned}\alpha_3(k) &= \mathbb{E}[\epsilon_3^2(k)] \\ &= \alpha_3(k-1) - \frac{\mathbb{E}^2[Y_3(k)\epsilon_3(k-1)]}{\mathbb{E}[{Y_3}^2(k)]},\end{aligned} \tag{3.17}$$

with the initial condition in Equation (3.5). The terms in (3.17) can be simplified as

$$\mathbb{E}[{Y_3}^2(k)] = g_{2,3}^2 P_2 + \sigma_3^2, \tag{3.18}$$

and

$$\begin{aligned}
&\mathbb{E}[Y_3(k)\epsilon_3(k-1)] \\
&= g_{2,3}\sqrt{\frac{P_2}{\alpha_{3,2}(k-1)}}\mathbb{E}[\epsilon_3(k-1)\epsilon_{3,2}(k-1)] \\
&= g_{2,3}\sqrt{\alpha_3(k-1)}\sqrt{P_2}\rho(k-1).
\end{aligned} \tag{3.19}$$

Substituting (3.18) and (3.19) into (3.17), we get

$$\alpha_3(k) = \alpha_3(k-1)q(k-1), \tag{3.20}$$

where

$$q(k-1) = \frac{g_{2,3}^2 P_2\,(1-\rho^2(k-1))+\sigma_3^2}{g_{2,3}^2 P_2+\sigma_3^2}. \tag{3.21}$$

- *Variance $\alpha_{3,2}(k)$ and the correlation coefficient $\rho(k)$:* To begin with, define

$$\gamma(k) := \mathbb{E}\left[\epsilon_3(k)\epsilon_2(k)\right],$$

so that

$$\mathbb{E}\left[\epsilon_{3,2}(k)\epsilon_2(k)\right] = \gamma(k) - \alpha_2(k).$$

We will now find a recursive relation for $\gamma(k)$. From (3.11) and (3.16), we can write

$$\begin{aligned}\gamma(k) &= \mathbb{E}\Big[(\epsilon_2(k-1) - \beta_2(k)Y_2(k))(\epsilon_3(k-1) - \beta_3(k)Y_3(k))\Big] \\ &= \gamma(k-1) - \beta_2(k)\mathbb{E}[\epsilon_3(k-1)Y_2(k)] - \beta_3(k)\mathbb{E}[\epsilon_2(k-1)Y_3(k)] \\ &\qquad + \beta_3(k)\beta_2(k)\mathbb{E}[Y_2(k)Y_3(k)]. \quad (3.22)\end{aligned}$$

We need to compute the three terms $\mathbb{E}[\epsilon_3(k-1)Y_2(k)]$, $\mathbb{E}[\epsilon_2(k-1)Y_3(k)]$ and $\mathbb{E}[Y_2(k)Y_3(k)]$ to evaluate the above expression. Using the expressions from the coding scheme, we obtain

$$\begin{aligned}\mathbb{E}[\epsilon_3(k-1)Y_2(k)] &= g_{1,2}\sqrt{\frac{P_1}{\alpha_2(k-1)}}\gamma(k-1) \\ \mathbb{E}[\epsilon_2(k-1)Y_3(k)] &= g_{2,3}\sqrt{\frac{P_2}{\alpha_{3,2}(k-1)}}(\gamma(k-1) - \alpha_2(k-1)), \\ \mathbb{E}[Y_2(k)Y_3(k)] &= g_{1,2}g_{2,3}\sqrt{\frac{P_1P_2}{\alpha_2(k-1)\alpha_{3,2}(k-1)}}(\gamma(k-1) - \alpha_2(k-1)).\end{aligned}$$

Using the expectations above and substituting the value of $\beta_2(k)$ from (3.9), we can simplify the sum of first and second terms of (3.22):

$$\gamma(k-1) - \beta_2(k)\mathbb{E}[\epsilon_3(k-1)Y_2'(k)] = r\gamma(k-1). \quad (3.23)$$

Similarly, the third and fourth terms of (3.22) can be simplified to

$$\begin{aligned}&-\beta_3(k)\mathbb{E}[\epsilon_2(k-1)Y_3(k)] + \beta_3(k)\beta_2(k)\mathbb{E}[Y_2'(k)Y_3(k)] \\ &= -\beta_3(k)rg_{2,3}\sqrt{\frac{P_2}{\alpha_{3,2}(k-1)}}\left(\gamma(k-1) - \alpha_2(k-1)\right). \quad (3.24)\end{aligned}$$

Substituting (3.23) and (3.24) in (3.22), we get a recursive relation for $\gamma(k)$

$$\gamma(k)=\frac{r}{\alpha_{3,2}(k-1)\left(g_{2,3}^2P_2+\sigma_3^2\right)}\Big(g_{2,3}^2P_2\left(\alpha_2(k-1)\alpha_3(k-1)-\gamma^2(k-1)\right)$$
$$+\sigma_3^2\alpha_{3,2}(k-1)\gamma(k-1)\Big) \quad (3.25)$$

with the initial value $\gamma(1) = \alpha_2(1)$. Now, we will prove that $\gamma(k) = \alpha_2(k)$ by induction. Lets assume that $\gamma(k-1) = \alpha_2(k-1)$. Thus, $\alpha_{3,2}(k-1) = \alpha_3(k-1) - \alpha_2(k-1)$. Using this and $\gamma(k-1) = \alpha_2(k-1)$ in (3.25), we get

$$\gamma(k) = r\alpha_2(k-1) = \alpha_2(k). \quad (3.26)$$

Now define $\bar{P}_1 := \frac{g_{1,2}^2P_1}{\sigma_2^2}$ and $\bar{P}_2 := \frac{g_{2,3}^2P_2}{\sigma_3^2}$. Given $\alpha_2(k), \alpha_3(k)$, and $\gamma(k)$, we can calculate $\alpha_{3,2}(k)$ and $\rho(k)$ using

$$\begin{aligned}
\alpha_{3,2}(k) &= \alpha_3(k)+\alpha_2(k)-2\gamma(k)\\
&= \alpha_3(k)-\alpha_2(k), && (3.27)\\
\rho(k) &= \frac{\alpha_3(k)-\gamma(k)}{\sqrt{\alpha_3(k)\alpha_{3,2}(k)}}\\
\Rightarrow \rho^2(k) &\overset{(a)}{=} 1-\frac{\alpha_2(k)}{\alpha_3(k)} && (3.28)\\
&\overset{(b)}{=} 1-\frac{\alpha_2(k-1)\frac{1}{\bar{P}_1+1}}{\alpha_3(k-1)\frac{\bar{P}_2(1-\rho^2(k-1))+1}{\bar{P}_2+1}}\\
&\overset{(c)}{=} \frac{\bar{P}_1\bar{P}_2(1-\rho^2(k-1)) + \bar{P}_1+\rho^2(k-1)}{\left(\bar{P}_1+1\right)\left(\bar{P}_2(1-\rho^2(k-1))+1\right)}, && (3.29)
\end{aligned}$$

where (a) follows from (3.26) and (3.27), (b) follows from (3.15) and (3.20) and (c) follows from using the value of $\frac{\alpha_2(k-1)}{\alpha_3(k-1)}$ from (3.28).

We now prove the convergence of the sequence $\rho^2(k)$. To calculate the fixed point, we equate $\rho(k) = \rho(k-1) = \rho$ in (3.29) to get a quadratic equation in ρ^2, the roots of which are 1 and $\frac{\bar{P}_1(\bar{P}_2+1)}{\bar{P}_2(\bar{P}_1+1)}$. Now, define

$$\begin{aligned} \rho^{*2} &:= \min\left\{1, \frac{\bar{P}_1(\bar{P}_2+1)}{\bar{P}_2(\bar{P}_1+1)}\right\} \\ &= \begin{cases} 1 & \text{if } \bar{P}_1 \geq \bar{P}_2 \\ \frac{\bar{P}_1(\bar{P}_2+1)}{\bar{P}_2(\bar{P}_1+1)} & \text{if } \bar{P}_1 < \bar{P}_2 \end{cases} \end{aligned} \tag{3.30}$$

The following lemma proves the convergence of $\rho^2(k)$ to ρ^{*2}.

Lemma 3.3.1 *Consider the function* $f : x \to \frac{\bar{P}_1\bar{P}_2(1-x)+\bar{P}_1+x}{(\bar{P}_1+1)(\bar{P}_2(1-x)+1)}$ *defined on the closed interval* $[0,1]$*, with* $\bar{P}_1, \bar{P}_2 \geq 0$*. For the function* f *and starting point* $x(1) \in [0, \rho^{*2}]$*, where* ρ^{*2} *is defined in* (3.30)*, the infinite sequence* $x(1), x(2) = f(x(1)), x(3) = f(x(2)), \ldots$ *converges to the fixed point* $x^* = \rho^{*2}$*.*

Proof: Note that either $x(1) = f(x(1)) = x^*$ or $0 \leq x(1) < x^*$. For the first case, the lemma follows directly. We need to prove the lemma for the second case. Note that

$$x(k+1) - x(k) = \frac{\bar{P}_2(\bar{P}_1+1)x^2(k) - (\bar{P}_1+\bar{P}_2+2\bar{P}_1\bar{P}_2)x(k) + \bar{P}_1(\bar{P}_2+1)}{(\bar{P}_1+1)(\bar{P}_2(1-x(k))+1)}$$

It can be seen that $x(k+1) - x(k) \geq 0$ for $x(k) \in [0, x^*]$. Thus, the sequence $x(k)$ is monotonically increasing. Also, the function $f(\cdot)$ is continuous and strictly increasing on the interval $[0,1]$. Since, $0 \leq x(1) < x^*$, it follows that $x(2) = f(x(1)) < f(x^*) = x^*$. Similarly, $f(x(k)) < x^*$ for all k. Thus, the sequence is monotonically increasing and upper bounded by x^*, which is also a fixed point of the function f in the interval $[0, x^*]$. From this, it follows that the sequence $x(k)$ converges to x^*. Note that the initial value $\rho^2(1) = \frac{\bar{P}_1}{\bar{P}_1+1} < \rho^{*2}$, thus applying

Lemma 5.2.2 to the sequence of correlation coefficients $\rho^2(k)$, it follows that the sequence converges to ρ^{*2}.

3.3.3 Rate and Error Performance

The number of messages M can be expressed in terms of the rate R as $M = e^{nR}$. Also $\alpha_3(n) = \alpha_3(1) \prod_{k=2}^{n} q(k-1)$. Now $q(n)$ converges to q^* as $\rho^2(k)$ converges to ρ^{*2}.

$$q* := \lim_{n\to\infty} q(n) = \frac{1}{1+\min\{\bar{P}_1, \bar{P}_2\}}. \tag{3.31}$$

At time $k = n$, the decoder's estimate of the message point W is given by $\hat{W}_3(n) = W + \epsilon_3(n)$. The decoder decodes the message as the message point closest to $\hat{W}_3(n)$. A decoding error occurs if there is a message point $W' \neq W$ closer to $\hat{W}_3(n)$ than W. The decoding error probability P_e is upper bounded by the probability that the magnitude of $\epsilon_3(n)$ is greater than half the distance between adjacent points.

$$\begin{aligned} P_e &\leq Pr\left[|\epsilon_3(n)| > \frac{1}{2(M-1)}\right] \\ &\leq 2Q\left(\frac{1}{2M\sqrt{\alpha_3(n)}}\right) \\ &= 2Q\left(\frac{e^{\frac{1}{2}\sum_{k=2}^{n}\ln\frac{1}{q(k-1)} - nR}}{2\sqrt{\alpha_3(1)}}\right), \end{aligned} \tag{3.32}$$

where $Q(x) := \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp(-\frac{y^2}{2})\, \mathrm{d}y$.

We have the following theorem for achievable rates over the cascade channel.

Theorem 3.3.2 *For the cascade channel presented in Section 3.3, the coding scheme presented in Section 3.3.1 achieves a rate* $R < \min\{C_{1,2}, C_{2,3}\}$.

Proof: From (3.32), we can see that $P_e \to 0$ as $n \to \infty$ if

$$R < \lim_{n\to\infty} \frac{1}{2n} \sum_{k=2}^{n} \ln \frac{1}{q(k-1)}.$$

Since $q(k) \to q^*$, where q^* is defined in (3.31), all rates $R < \min\{C_{1,2}, C_{2,3}\}$ can be achieved. Since $Q(x) \sim \exp(\frac{-x^2}{2})$ for large x, for all rates satisfying $R < \min\{C_{1,2}, C_{2,3}\}$, the error decays doubly exponentially in n.

3.4 Relay Channel

Consider a relay channel with destination-relay and destination-source feedback, but no feedback from the relay to the source. The capacity of a discrete memoryless relay channel with such feedback is [11]

$$C = \sup_{P(x_1,x_2)} \min\{\mathbb{I}(X_1; Y_2, Y_3|X_2), \mathbb{I}(X_1, X_2; Y_3)\}. \tag{3.33}$$

For the Gaussian relay channel with feedback, the capacity can be expressed in terms of the channel parameters and the power constraints as $C = \max_{0\leq\kappa\leq 1} C(\kappa)$, where

$$C(\kappa) = \min\left\{\frac{1}{2}\ln\left(1+(1-\kappa^2)\left(\frac{g_{1,2}^2}{\sigma_2^2}+\frac{g_{1,3}^2}{\sigma_3^2}\right)P_1\right),\right.$$
$$\left.\frac{1}{2}\ln\left(1+\frac{g_{1,3}^2}{\sigma_3^2}P_1+\frac{g_{2,3}^2}{\sigma_3^2}P_2+2\kappa\sqrt{\frac{g_{1,3}^2}{\sigma_3^2}\frac{g_{2,3}^2}{\sigma_3^2}P_1P_2}\right)\right\} \tag{3.34}$$

The parameter κ represents the correlation factor between the source input X_1 and relay signal X_2. Decreasing κ increases the mutual information in the broadcast cut channel, but decreases the mutual information in the multiple access cut.

Therefore, depending on the different channel parameters and power constraints, different values of the parameter κ maximizes the channel capacity.

3.4.1 Coding Scheme

We now proceed to explain the coding scheme used at the source and the relay. The code works as follows.

- *Initialization:* At time step $k = 0$,
  - The source node 1 observes W and sends the input $X_1(0)$ given by

$$X_1(0) = \sqrt{\frac{P_1}{\sigma_W^2}} W.$$

  - The relay node 2 sends nothing, while it receives

$$Y_2(0) = g_{1,2} X_1(0) + Z_2(0).$$

    It then forms an estimate of W by scaling as follows:

$$\hat{W}_2(0) = \frac{1}{g_{1,2}} \sqrt{\frac{\sigma_W^2}{P_1}} Y_2(0).$$

    The estimation error $\epsilon_2(0)$ is given by

$$\epsilon_2(0) = \frac{1}{g_{1,2}} \sqrt{\frac{\sigma_W^2}{P_1}} Z_2(0).$$

    Clearly, $\epsilon_2(0)$ is zero-mean Gaussian with variance $\alpha_2(0)$, given by

$$\alpha_2(0) = \frac{\sigma_W^2 \sigma_2^2}{g_{1,2}^2 P_1}. \tag{3.35}$$

– The destination node 3 receives

$$Y_3(0) = g_{1,3}X_1(0) + Z_3(0)$$

and forms an estimate of W given by

$$\hat{W}_3(0) = \frac{1}{g_{1,3}}\sqrt{\frac{\sigma_W^2}{P_1}}Y_3(0).$$

The estimation error $\epsilon_3(0)$ is again zero-mean Gaussian with variance $\alpha_3(0)$, given by

$$\alpha_3(0) = \frac{\sigma_W^2\sigma_3^2}{g_{1,3}^2P_1}. \tag{3.36}$$

- *Update:* At every time step $k \geq 1$,

  – Using the message transmitted by the destination at time $k-1$, the source and the relay both know $\hat{W}_3(k-1)$. Moreover, the source can calculate $\epsilon_3(k)$, while the relay can calculate $\epsilon_{3,2}(k)$. Note that by the construction that follows, $\epsilon_3(k)$ and $\epsilon_{3,2}(k)$ are jointly Gaussian with zero means, respective variances $\alpha_3(k)$ and $\alpha_{3,2}(k)$ and correlation coefficient $\rho(k)$ given by

$$\rho(k) = \frac{\mathbb{E}[\epsilon_3(k)\epsilon_{3,2}(k)]}{\sqrt{\alpha_3(k)\alpha_{3,2}(k)}}. \tag{3.37}$$

  The quantities $\alpha_2(k)$, $\alpha_3(k), \alpha_{3,2}(k)$ and $\rho(k)$ are deterministic and can be calculated prior to transmission by all the nodes. In particular,

$$\alpha_{3,2}(0) = \alpha_3(0) + \alpha_2(0)$$
$$\rho(0) = \frac{\alpha_3(0)}{\sqrt{(\alpha_3(0) + \alpha_2(0))\alpha_3(0)}}.$$

We present recursive expressions for these quantities in Section 3.4.2, and assume for the rest of the presentation that these quantities can be calculated by all the nodes.

– The source transmits

$$\begin{aligned} X_1(k) &= \sqrt{\frac{P_1}{\alpha_3(k-1)}} \left(\hat{W}_3(k-1) - W\right) \\ &= \sqrt{\frac{P_1}{\alpha_3(k-1)}} \epsilon_3(k-1). \end{aligned} \tag{3.38}$$

– The relay transmits

$$\begin{aligned} X_2(k) &= \sqrt{\frac{P_2}{\alpha_{3,2}(k-1)}} \left(\hat{W}_3(k-1) - \hat{W}_2(k-1)\right) \\ &= \sqrt{\frac{P_2}{\alpha_{3,2}(k-1)}} \epsilon_{3,2}(k-1). \end{aligned} \tag{3.39}$$

Moreover, it updates its estimate as follows. Using the feedback from the destination and the received signal from the source, it first calculates

$$
\begin{aligned}
&Y_2'(k)\\
&= Y_2(k) - g_{1,2}\sqrt{\frac{P_1}{\alpha_3(k-1)}}\Big(\hat{W}_3(k-1) - \hat{W}_2(k-1)\Big) && (3.40)\\
&= g_{1,2}\sqrt{\frac{P_1}{\alpha_3(k-1)}}\Big(\hat{W}_2(k-1) - W\Big) + Z_2(k). && (3.41)
\end{aligned}
$$

Using this quantity, the relay calculates the linear minimum mean squared error estimate of W given $Y_2'(k)$ and $\hat{W}_2(k-1)$ as

$$\hat{W}_2(k) = \hat{W}_2(k-1) - \frac{\mathbb{E}[Y_2'(k)\epsilon_2(k-1)]}{\mathbb{E}[Y_2'^2(k)]} Y_2'(k). \quad (3.42)$$

Notice that the step size $\beta_2(k)$ used by the relay in the distributed stochastic approximation iteration (2.21) is given by $\beta_2(k) = \frac{\mathbb{E}[Y_2'(k)\epsilon_2(k-1)]}{\mathbb{E}[Y_2'^2(k)]}$.

- At time $k \geq 1$, the destination calculates the linear minimum mean squared error estimate of W given $Y_3(k)$ and $\hat{W}_3(k-1)$ as

$$\hat{W}_3(k) = \hat{W}_3(k-1) - \frac{\mathbb{E}[Y_3(k)\epsilon_3(k-1)]}{\mathbb{E}[{Y_3}^2(k)]} Y_3(k). \quad (3.43)$$

  The step size $\beta_3(k)$ used by the destination is thus $\beta_3(k) = \frac{\mathbb{E}[Y_3(k)\epsilon_3(k-1)]}{\mathbb{E}[{Y_3}^2(k)]}$.

Note that the inputs $X_i(k)$ to the various channels satisfy their respective power constraints.

3.4.2 Calculation of the Variances

We now proceed to evaluate the recursive expressions of $\alpha_2(k)$, $\alpha_3(k)$, $\alpha_{3,2}(k)$, and $\rho(k)$ as used in the coding scheme presented above. The following recursions do not depend on the data, and can be executed by any node.

- *Variance $\alpha_2(k)$ of the error at the relay node:* Since $\epsilon_2(k)$ is defined as $\hat{W}_2(k) - W$, from (3.42) we obtain

$$\epsilon_2(k) = \epsilon_2(k-1) - \frac{\mathbb{E}[Y_2'(k)\epsilon_2(k-1)]}{\mathbb{E}[Y_2'^2(k)]} Y_2'(k). \tag{3.44}$$

The variance of $\epsilon_2(k)$ can be obtained as

$$\begin{aligned}\alpha_2(k) &= \mathbb{E}[\epsilon_2^2(k)] \\ &= \alpha_2(k-1) - \frac{\mathbb{E}^2[Y_2'(k)\epsilon_2(k-1)]}{\mathbb{E}[Y_2'^2(k)]},\end{aligned} \tag{3.45}$$

with the initial condition in Equation (3.35). The terms in (3.45) can be further evaluated to be

$$\mathbb{E}[Y_2'^2(k)] = g_{1,2}^2 \frac{P_1}{\alpha_3(k-1)} \alpha_2(k-1) + \sigma_2^2, \tag{3.46}$$

and

$$\mathbb{E}[Y_2'(k)\epsilon_2(k-1)] = g_{1,2} \sqrt{\frac{P_1}{\alpha_3(k-1)}} \alpha_2(k-1). \tag{3.47}$$

Substituting (3.46) and (3.47) into (3.45), we get

$$\alpha_2(k) = \alpha_2(k-1) r(k-1), \tag{3.48}$$

where $r(k-1) = \left(\frac{\sigma_2^2}{g_{1,2}^2 \frac{P_1}{\alpha_3(k-1)} \alpha_2(k-1) + \sigma_2^2} \right)$.

- *Variance $\alpha_3(k)$ of the error at the destination node:* Since $\epsilon_3(k)$ is defined as $\hat{W}_3(k) - W$, from (3.43) we obtain

$$\epsilon_3(k) = \epsilon_3(k-1) - \frac{\mathbb{E}[Y_3(k)\epsilon_3(k-1)]}{\mathbb{E}[{Y_3}^2(k)]} Y_3(k). \tag{3.49}$$

The variance of $\epsilon_3(k)$ can be obtained as

$$\begin{aligned} \alpha_3(k) &= \mathbb{E}[\epsilon_3^2(k)] \\ &= \alpha_3(k-1) - \frac{\mathbb{E}^2[Y_3(k)\epsilon_3(k-1)]}{\mathbb{E}[{Y_3}^2(k)]}, \end{aligned} \tag{3.50}$$

with the initial condition in Equation (3.36). The terms in (3.50) can be simplified as

$$\mathbb{E}[{Y_3}^2(k)] = g_{1,3}^2 P_1 + g_{2,3}^2 P_2 + 2 g_{1,3} g_{2,3} \sqrt{P_1 P_2} \rho(k-1) + \sigma_3^2, \tag{3.51}$$

and

$$\begin{aligned} \mathbb{E}[Y_3(k)\epsilon_3(k-1)] &= g_{1,3}\sqrt{P_1 \alpha_3(k-1)} + g_{2,3}\sqrt{\frac{P_2}{\alpha_{3,2}(k-1)}} \mathbb{E}[\epsilon_3(k-1)\epsilon_{3,2}(k-1)] \\ &= \sqrt{\alpha_3(k-1)} \left(g_{1,3}\sqrt{P_1} + g_{2,3}\sqrt{P_2}\rho(k-1) \right). \end{aligned} \tag{3.52}$$

Substituting (3.51) and (3.52) into (3.50), we get

$$\alpha_3(k) = \alpha_3(k-1) q(k-1), \tag{3.53}$$

where

$$q(k-1)=\frac{g_{2,3}^2 P_2\,(1-\rho^2(k-1))+\sigma_3^2}{g_{1,3}^2 P_1+g_{2,3}^2 P_2+2g_{1,3}g_{2,3}\sqrt{P_1P_2}\rho(k-1)+\sigma_3^2}. \tag{3.54}$$

- *Variance $\alpha_{3,2}(k)$ and the correlation coefficient $\rho(k)$:* To begin with, define

$$\gamma(k) := \mathbb{E}\left[\epsilon_3(k)\epsilon_2(k)\right],$$

so that

$$\mathbb{E}\left[\epsilon_{3,2}(k)\epsilon_2(k)\right] = \gamma(k) - \alpha_2(k).$$

We will now find a recursive relation for $\gamma(k)$. From (3.44) and (3.49), we can write

$$\begin{aligned}\gamma(k) &= \mathbb{E}\Big[(\epsilon_2(k-1) - \beta_2(k)Y_2'(k))(\epsilon_3(k-1) - \beta_3(k)Y_3(k))\Big]\\ &= \gamma(k-1) - \beta_2(k)\mathbb{E}[\epsilon_3(k-1)Y_2'(k)] - \beta_3(k)\mathbb{E}[\epsilon_2(k-1)Y_3(k)]\\ &\qquad + \beta_3(k)\beta_2(k)\mathbb{E}[Y_2'(k)Y_3(k)].\end{aligned} \tag{3.55}$$

We need to compute the three terms $\mathbb{E}[\epsilon_3(k-1)Y_2'(k)]$, $\mathbb{E}[\epsilon_2(k-1)Y_3(k)]$ and $\mathbb{E}[Y_2'(k)Y_3(k)]$ to evaluate the above expression. Using the expressions from the coding scheme, we obtain

$$\mathbb{E}[\epsilon_3(k-1)Y_2'(k)] = g_{1,2}\sqrt{\frac{P_1}{\alpha_3(k-1)}}\gamma(k-1)$$

$$\mathbb{E}[\epsilon_2(k-1)Y_3(k)] = g_{1,3}\sqrt{\frac{P_1}{\alpha_3(k-1)}}\gamma(k-1) + g_{2,3}\sqrt{\frac{P_2}{\alpha_{3,2}(k-1)}}(\gamma(k-1) - \alpha_2(k-1)),$$

$$\mathbb{E}[Y_2'(k)Y_3(k)] = g_{1,3}g_{1,2}\frac{P_1}{\alpha_3(k-1)}\gamma(k-1) + g_{1,2}g_{2,3}\sqrt{\frac{P_1P_2}{\alpha_3(k-1)\alpha_{3,2}(k-1)}}(\gamma(k-1) - \alpha_2(k-1)).$$

Using the expectations above and substituting the value of $\beta_2(k)$ from (3.42), we can simplify the sum of first and second terms of (3.55):

$$\gamma(k-1) - \beta_2(k)\mathbb{E}[\epsilon_3(k-1)Y_2'(k)] = r(k-1)\gamma(k-1). \tag{3.56}$$

Similarly, the third and fourth terms of (3.55) can be simplified to

$$\begin{aligned} &-\beta_3(k)\mathbb{E}[\epsilon_2(k-1)Y_3(k)] + \beta_3(k)\beta_2(k)\mathbb{E}[Y_2'(k)Y_3(k)] \\ &= -\beta_3(k)r(k-1)\Big[g_{1,3}\sqrt{\frac{P_1}{\alpha_3(k-1)}}\gamma(k-1) \\ &\qquad + g_{2,3}\sqrt{\frac{P_2}{\alpha_{3,2}(k-1)}}(\gamma(k-1) - \alpha_2(k-1))\Big]. \end{aligned} \tag{3.57}$$

Substituting (3.56) and (3.57) in (3.55), we get a recursive relation for $\gamma(k)$, with the initial value $\gamma(0) = 0$. Given $\alpha_2(k), \alpha_3(k)$, and $\gamma(k)$, we can calculate $\alpha_{3,2}(k)$ and $\rho(k)$ using

$$\alpha_{3,2}(k) = \alpha_3(k) + \alpha_2(k) - 2\gamma(k),$$
$$\rho(k) = \frac{\alpha_3(k) - \gamma(k)}{\sqrt{\alpha_3(k)\alpha_{3,2}(k)}}.$$

3.4.3 Rate and Error Performance

The number of messages M can be expressed in terms of the rate R as $M = e^{nR}$. Also $\alpha_3(n) = \alpha_3(0)\prod_{k=0}^{n-1} q(k)$.

At time $k = n$, the decoder's estimate of the message point W is given by $\hat{W}_3(n) = W + \epsilon_3(n)$. The decoder decodes the message as the message point closest to $\hat{W}_3(n)$. As before, a decoding error occurs if there is a message point $W' \neq W$ closer to $\hat{W}_3(n)$ than W. The decoding error probability P_e is upper bounded by the probability that the magnitude of $\epsilon_3(n)$ is greater than half the distance between adjacent points.

$$\begin{aligned} P_e &\leq Pr\left[|\epsilon_3(n)| > \frac{1}{2(M-1)}\right] \\ &\leq 2Q\left(\frac{1}{2M\sqrt{\alpha_3(n)}}\right) \\ &= 2Q\left(\frac{e^{\frac{1}{2}\sum_{k=1}^{n}\ln\frac{1}{q(k-1)} - nR}}{2\sqrt{\alpha_3(0)}}\right), \end{aligned} \tag{3.58}$$

where $Q(x) := \int_x^\infty \frac{1}{\sqrt{2\pi}}\exp(-\frac{y^2}{2})\,\mathrm{d}y$.

We have the following theorem for achievable rates over the relay channel with feedback.

Theorem 3.4.1 *For the relay channel presented in Section 3.4, the coding scheme* *presented in Section 3.4.1 achieves a rate*

$$R < \liminf_{n\to\infty} \frac{1}{2n} \sum_{k=1}^{n} \ln \frac{1}{q(k-1)}, \tag{3.59}$$

where $q(k-1)$ *is given by* (3.54).

Proof: From (3.58), we can see that $P_e \to 0$ as $n \to \infty$ if

$$R < \liminf_{n\to\infty} \frac{1}{2n} \sum_{k=1}^{n} \ln \frac{1}{q(k-1)}.$$

Since $Q(x) \sim \exp(\frac{-x^2}{2})$ for large x, for all rates satisfying (3.59), the error decays doubly exponentially in n.

3.4.4 Numerical Results

Consider a relay channel setting where the destination is placed at a (spatial) distance of 2 units from the source. The relay is placed on the line joining the source and destination. The total power is fixed at $P_1 + P_2 = 1$. The system parameters are considered to be $\sigma_2^2 = \sigma_3^2 = 1$ and $\eta = 2$. All the logarithms are taken to the base e. We plot the limit in (3.59) and the capacity as a function of the power P_1 and distance of the relay from the plant in Fig. 3.2. The system parameters are the same as above. We plot the gap to capacity in Fig. 3.3. As shown, we achieve rates close to capacity.

3.5 Half-Duplex Relay Channel

We can consider half-duplex relays as compared to full-duplex relays considered earlier. The relay would listen at every even time step and transmit at every odd

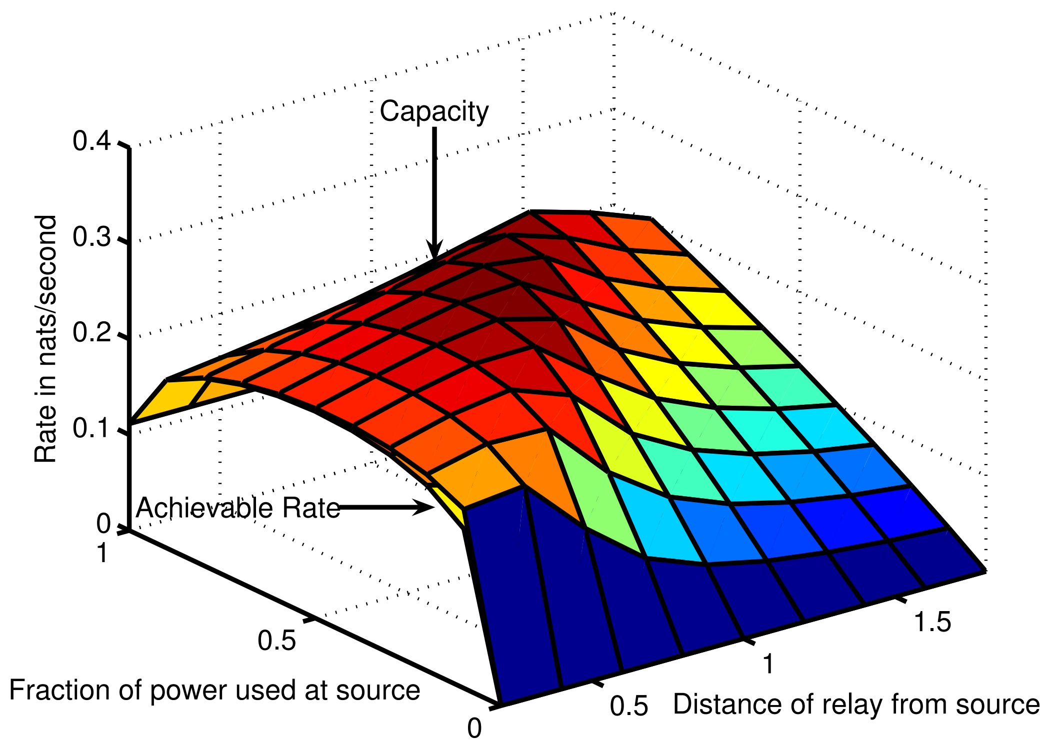

Figure 3.2. Achievable rate as a function of fraction of power used at the source and distance of relay from source with the relay operating in full-duplex.

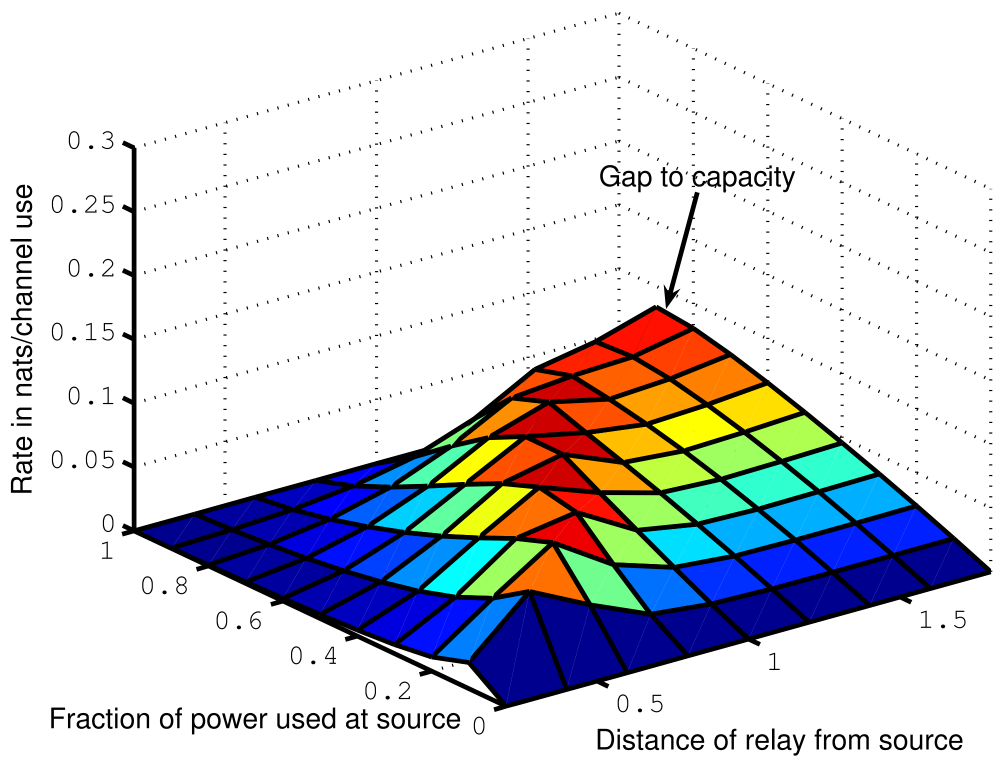

Figure 3.3. Gap to capacity as a function of fraction of power used at the source and distance of relay from source with the relay operating in full-duplex.

time step. In such a scenario, at every even time step, the destination sees a noisy version of the input from only the source, whereas at every odd time step, it sees a noisy version of the sum of inputs from the source and relay node, i.e.,

$$Y_3(k) = \begin{cases} g_{1,3}X_1(k) + g_{2,3}X_2(k) + Z_3(k) & k \text{ odd,} \\ g_{1,3}X_1(k) + Z_3(k) & k \text{ even.} \end{cases}$$

The feedback structure is similar to the relay channel case in 3.4.

We now proceed to explain the coding scheme used at the source and the relay. The code works as follows.

- *Initialization:* The initialization step is same as in Section 3.4.1.
- *Update:* At every time step $k \geq 1$,
  - Using the message transmitted by the destination at time $k-1$, the source and the relay both know $\hat{W}_3(k-1)$. Moreover, the source can calculate $\epsilon_3(k)$, while the relay can calculate $\epsilon_{3,2}(k)$. Note that by the construction that follows, $\epsilon_3(k)$ and $\epsilon_{3,2}(k)$ are jointly Gaussian with zero means, respective variances $\alpha_3(k)$ and $\alpha_{3,2}(k)$ and correlation coefficient $\rho(k)$ given by (3.37)

    We present recursive expressions for these quantities in later in Section 3.4.2, and assume for the rest of the presentation that these quantities can be calculated by all the nodes.
  - The source transmits

$$X_1(k) = \sqrt{\frac{P_1}{\alpha_3(k-1)}}\epsilon_3(k-1).$$

– The relay transmits

$$X_2(k) = \begin{cases} \sqrt{\frac{2P_2}{\alpha_{3,2}(k-1)}}\epsilon_{3,2}(k-1) & k \text{ odd,} \\ 0 & k \text{ even.} \end{cases}$$

Moreover, it updates its estimate at even time steps (k = 2,4, ...) as in (3.41) and (3.42).

– At time $k \geq 1$, the destination calculates the linear minimum mean squared error estimate of W given $Y_2'(k)$ and $\hat{W}_3(k-1)$ as in ((3.43))

Note that the inputs $X_i(k)$ to the various channels satisfy their respective power constraints.

3.5.1 Calculation of the Variances

We can see that recursive expressions of $\alpha_2(k)$, $\alpha_3(k)$, $\alpha_{3,2}(k)$, and $\rho(k)$ as used in the coding scheme presented above at odd time steps is same as in Section 3.4.1, where as the recursive expressions at even time steps can be calculated by substituting $P_2 = 0$ in Section 3.4.2. Thus,

- *Variance $\alpha_2(k)$ of the error at the relay node:*

$$\alpha_2(k) = \begin{cases} \alpha_2(k-1)r(k-1) & k \text{ even,} \\ \alpha_2(k-1) & k \text{ odd.} \end{cases}$$

where $r(k-1) = \left(\frac{\sigma_2^2}{g_{1,2}^2 \frac{P_1}{\alpha_3(k-1)}\alpha_2(k-1)+\sigma_2^2}\right)$. Note that the above expression for even time steps same as in (3.48).

- *Variance $\alpha_3(k)$ of the error at the destination node:*

$$\alpha_3(k)=\alpha_3(k-1)q(k-1), \tag{3.60}$$

where

$$q(k-1) = \begin{cases} \frac{2g_{2,3}^2 P_2\left(1-\rho^2(k-1)\right)+\sigma_3^2}{g_{1,3}^2 P_1+2g_{2,3}^2 P_2+2g_{1,3}g_{2,3}\sqrt{2P_1P_2}\rho(k-1)+\sigma_3^2} & k \text{ odd,} \\ \frac{\sigma_3^2}{g_{1,3}^2 P_1+\sigma_3^2} & k \text{ even.} \end{cases} \tag{3.61}$$

Note that the above expression is same as in (3.54) with P_2 replaced with 0 for even time steps and $2P_2$ for odd time steps.

- *Variance $\alpha_{3,2}(k)$ and the correlation coefficient $\rho(k)$:* Define

$$\gamma(k) := \mathbb{E}\left[\epsilon_3(k)\epsilon_2(k)\right],$$

so that

$$\mathbb{E}\left[\epsilon_{3,2}(k)\epsilon_2(k)\right] = \gamma_h(k) - \alpha_2(k).$$

We can write a recursive relation for $\gamma(k)$ in a similar way as in (3.55) . Similarly, we can calculate $\alpha_{3,2}(k)$ and $\rho(k)$ using

$$\alpha_{3,2}(k) = \alpha_3(k) + \alpha_2(k) - 2\gamma(k),$$
$$\rho(k) = \frac{\alpha_3(k) - \gamma(k)}{\sqrt{\alpha_3(k)\alpha_{3,2}(k)}}.$$

3.5.2 Rate and Error Performance

The number of messages M can be expressed in terms of the rate R as $M = e^{nR}$. Also $\alpha_3(n) = \alpha_3(0)\prod_{k=0}^{n-1} q(k)$.

At time $k = n$, the decoder's estimate of the message point W is given by $\hat{W}_3(n) = W + \epsilon_3(n)$. The decoder decodes the message as the message point closest to $\hat{W}_3(n)$. As before, a decoding error occurs if there is a message point $W' \neq W$ closer to $\hat{W}_3(n)$ than W. The decoding error probability P_e is upper bounded by the probability that the magnitude of $\epsilon_3(n)$ is greater than half the distance between adjacent points.

$$\begin{aligned} P_e &\leq Pr\left[|\epsilon_3(n)| > \frac{1}{2(M-1)}\right] \\ &\leq 2Q\left(\frac{1}{2M\sqrt{\alpha_3(n)}}\right) \\ &= 2Q\left(\frac{e^{\frac{1}{2}\sum_{k=1}^{n} \ln \frac{1}{q(k-1)} - nR}}{2\sqrt{\alpha_3(0)}}\right), \end{aligned} \tag{3.62}$$

where $Q(x) := \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp(-\frac{y^2}{2})\, \mathrm{d}y$.

We have the following theorem for achievable rates over the half-dulpex relay channel.

Theorem 3.5.1 *For the half-duplex relay channel with feedback, the coding scheme presented above achieves a rate*

$$R < \liminf_{n\to\infty} \frac{1}{2n} \sum_{k=1}^{n} \ln \frac{1}{q(k-1)}, \tag{3.63}$$

where $q(k-1)$ *is given by* (3.61).

Proof: The proof is similar to the proof of Theorem 3.4.1.

Since $Q(x) \sim \exp(\frac{-x^2}{2})$ for large x, for all rates satisfying (3.63), the error decays doubly exponentially in n.

3.5.3 Numerical Example

Consider a relay channel setting similar to the one in Section 3.4.4. We plot the achievable rates for the half-duplex relay channel in Fig. 3.4

3.6 Summary

This chapter develops communication strategies based upon distributed stochastic approximation algorithms for a cascade of two point-to-point channels and a Gaussian relay channel with perfect feedback. The strategies have been shown to provide a doubly exponential error decay in the number of iterations or channel uses. We also discuss the case of half-duplex relaying. In the next chapter, we will see how can these coding schemes be applied to obtain stabilizability conditions over a Gaussian relay channel.

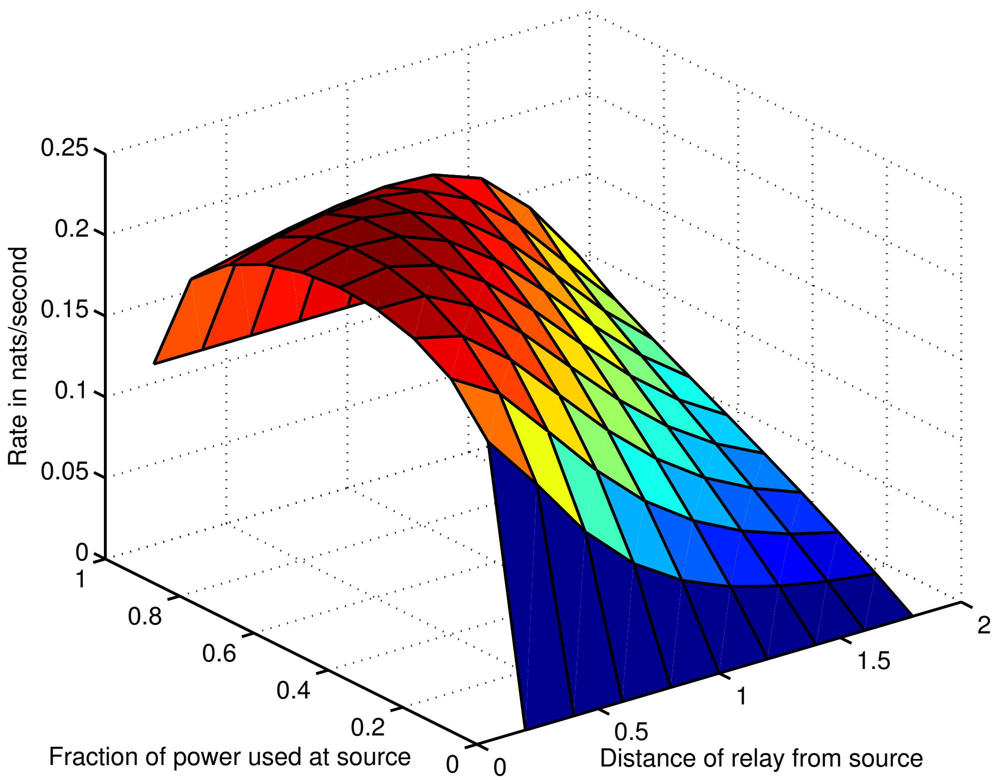

Figure 3.4. Achievable rate as a function of fraction of power used at the source and distance of relay from source with the relay operating in half-duplex.

CHAPTER 4

STABILITY ACROSS A GAUSSIAN CASCADE AND RELAY CHANNEL

In this chapter, we present sufficient conditions for stabilizing a discrete-time LTI plant in the mean square sense by a remotely placed controller over a Additive White Gaussian (AWGN) channel in the presence of a relay node. We consider two arrangements: (i) when the controller has a direct path from the sensor (the relay channel), and (ii) when the controller does not have a direct path from the sensor (the cascade channel). The coding schemes that we utilize are based on coding scheme presented in Section 3.4. A similar scheme was presented in [23] for the relay channel in the context of communication across a relay channel with feedback. We analyze the stability region of the closed loop system with the proposed scheme under average transmission power constraints for both the sensor and the relay node. Interestingly, our analysis suggests that the stability region may be increased by using a relay node even if the total transmission power remains the same.

The remainder of the chapter is organized as follows. We discuss the problem setup in Section 4.1. In Section 4.3.2 the encoder and decoder design at the sensor, relay node and controller are presented. The stability conditions with the proposed coding scheme for a scalar plant are presented in Section 4.3.5. We generalize the stability results to a vector plant in Section 5.2.3. The conditions for a scalar plant are numerically illustrated in Section 4.4 showing that the presence of a relay node improves the stability region even though the total transmission power is not increased. Some concluding remarks are given in Sections 5.4.

4.1 Problem Setup

The problem consists of the following elements.

4.1.1 Process

Consider an open loop unstable linear time invariant process evolving as

$$S(k+1) = AS(k) + BU(k), \tag{4.1}$$

where $S(k) \in \mathbb{R}^m$ is the state and $U(k) \in \mathbb{R}$ is the control value. The initial condition $S(0)$ is assumed to be a random variable with a finite covariance $\sigma^2_{S(0)}$ and an arbitrary probability distribution. For pedagogical ease and without loss of generality, we assume that the state of the process is observable by a sensor.

4.1.2 Communication channel

In this work, we concentrate on a wireless channel modeled as an additive white Gaussian noise channel. Such a channel has input and output alphabets as $\mathbb{R}$, and the output $y(k)$ at time k is related to the input $x(k)$ as

$$y(k) = gx(k) + w(k), \tag{4.2}$$

where $w(k)$ is white Gaussian noise with mean 0 and covariance $\Sigma_w > 0$. The factor g is an attenuation factor that depends on the distance d between the transmitter and receiver. Typical model for such a factor is

$$g = bd^{-\eta/2},$$

where η is the path loss exponent (typically $2 \leq \eta \leq 4$ for wireless communication), and b is a constant. In addition, there is a power constraint of the form

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{E}[x^2(k)] \leq P. \tag{4.3}$$

4.1.3 Sensor to controller communication

We are interested in analyzing if a relay node can help in better control performance over the baseline case when there is no relay node. In the baseline case, the sensor transmits to the controller that is located a distance d away across a channel of the form (4.2). Thus, at every time k, it transmits a scalar of the form

$$X_1(k) = h_1\left(k, S(0), S(1), \cdots, S(k), U(0), \ldots, U(k-1)\right), \tag{4.4}$$

while satisfying the power constraint (4.3). When the relay node is present, there are two cases that we consider. Let the relay node be placed at a distance d_1 from the sensor and d_2 from the controller. In either case, the sensor transmits a quantity $X_1(k)$ which is a causal function of the information it has access to, while satisfying a power constraint

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{E}[X_1^2(k)] \leq P_1. \tag{4.5}$$

In the first case, shown in Figure 4.1, a noisy version of $X_1(k)$ is received at both the relay and the controller. We refer to this setup as the Gaussian relay channel.

In the second case, shown in Figure 4.2, $X_1(k)$ is not received at the controller. We refer to this setup as the cascade channel. We denote the quantity received at time k at the relay from the sensor by $Y_2(k)$ and the corresponding noise by $Z_2(k)$, i.e.,

$$Y_2(k) = g_{12}X_1(k) + Z_2(k),$$

where $g_{12} = bd_1^{-\eta/2}$. The relay node transmits a quantity $X_2(k)$ that is a causal function of the information it has access to. Moreover the relay satisfies a power constraint

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \mathbb{E}[X_2^2(k)] \le P_2. \tag{4.6}$$

In the relay channel case, this transmission interferes with the quantity transmitted by the sensor. Thus, the controller receives

$$Y_3(k) = g_{13}X_1(k) + g_{23}X_2(k) + Z_3(k),$$

where $g_{13} = bd^{-\eta/2}$ and $g_{23} = bd_2^{-\eta/2}$. In the cascade channel case,

$$Y_3(k) = g_{23}X_2(k) + Z_3(k).$$

We assume that the noises $Z_2(.)$ and $Z_3(.)$ are mutually independent white noises with mean 0 and variances σ_2^2 and σ_3^2 respectively. Note that if $P_1 + P_2 = P$, then the relay and the sensor node share the power. In such a case, it is not clear if the addition of a relay node will help.

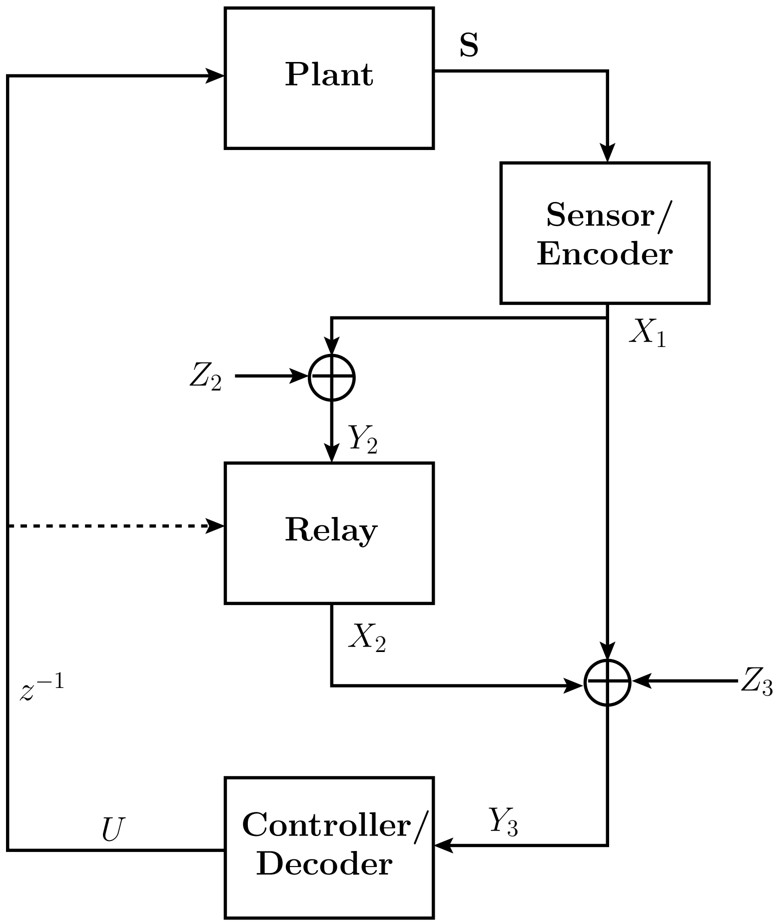

Figure 4.1. Problem setup for an unstable plant being controlled by a controller across an AWGN relay channel

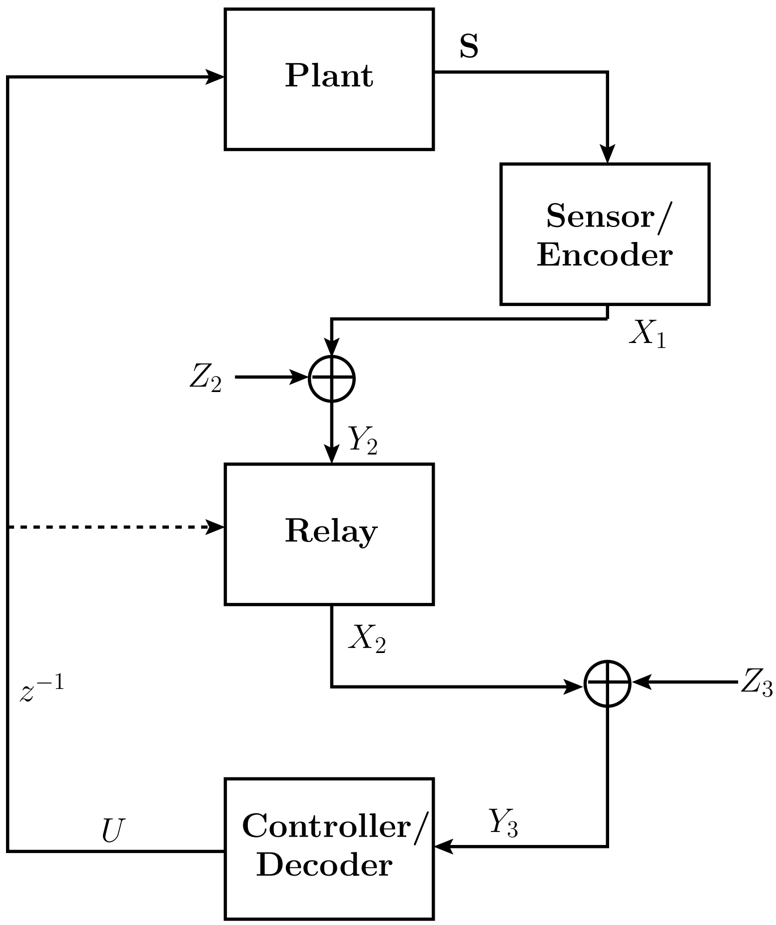

Figure 4.2. Problem setup for an unstable plant being controlled by a controller across an AWGN cascade channel

4.1.4 Controller to actuator communication

The controller calculates a control input

$$U(k) = h_3(k, U(0), \cdots, U(k-1), Y_3(0), \cdots, Y_3(k)) \tag{4.7}$$

and transmits it to the actuator. We assume that the controller is not power limited; thus, the actuator receives $U(k)$ without corruption. In addition, due to the broadcast nature of the communication, the sensor and the relay also receive the control $U(k)$.

4.1.5 Problem statement

We are interested in designing the functions $X_1(.)$ and $X_2(.)$ (that we refer to as the design of the coding scheme), and of $U(.)$ (that we refer to as the design of the controller), such that that the process (4.1) is mean squares stabilized. In particular, we are interested in the question if the stabilizability region of the process can be enhanced using a relay node even if the total transmission power used by the sensor and the relay nodes stays constant under the following constraints on the encoder and controller design:

- *Constraint C_1:* The control action must satisfy a controller cost constraint, $\sum_{k=0}^{\infty} \mathbb{E}[U(k)^T U(k)] < \infty$.
- *Constraint C_2:* The power constraints (4.5) and (4.6) are satisfied.
- *Constraint C_3:* Causality contraints (4.4) and (4.7) are satisfied.

Recall that a system is said to be mean squares stabilized if and only if irrespective of the initial state $S(0)$, the following conditions are satisfied:

$$\begin{aligned} \mathbb{E}[S(k)] &= 0, \\ \lim_{k\to\infty} \mathbb{E}[S(k)S^T(k)] &= 0. \end{aligned} \tag{4.8}$$

Moreover, our solution must satisfy the constraint C_1 and the power constraints (4.5) and (4.6).

4.2 Notation and Preliminary Results

Notation: Define the i-th basis vector by e_i. Thus, $e_i \in \mathbb{R}^m$ has all elements 0, except the i-th one which is unity. By $\ln(x)$ we mean logarithm to the base e. Denote the m eigenvalues of matrix M by $\lambda_j(M)$, $j = 1, \cdots, m$. We refer to the real numbers by $\mathbb{R}$ and positive integers by $\mathbb{Z}_+$.

4.2.1 Baseline Case

For this case we assume that the sensor and controller communicate over a AWGN channel. The following theorem [31] gives the sufficient conditions for the plant to be stabilized over a point-to-point channel.

Theorem 4.2.1 *Consider the problem formulation presented in Section 4.1 with the process given by (4.1) and absence of the relay node. The process (4.1) is mean square stabilized over the AWGN channel if*

$$\sum_{i=1}^{m} \max\{0, \ln(\lambda_i(A))\} < \frac{1}{2}\ln\left(1 + \frac{g_{1,3}^2 P}{\sigma_3^2}\right). \tag{4.9}$$

4.3 Main Results

Our solution consists of two parts. We propose an encoding scheme that ensures that at every time k, the controller can calculate an estimate $\hat{S}_3(k)$ of the state value $S(0)$ such that the error

$$\epsilon_3(k) = \hat{S}_3(k) - S(0) \tag{4.10}$$

has a covariance that decreases geometrically in k. We then show that using $\hat{S}_3(k)$ which satisfies this property, the controller can achieve mean squared stability of the process. We begin by proving the second part.

4.3.1 Controller Design

Theorem 4.3.1 *Consider the problem formulation stated in Section 4.1. Assume that*

$$\mathbb{E}[\epsilon_3(k)] = 0, \tag{4.11}$$

$$\lim_{k\to\infty} A^k \mathbb{E}[\epsilon_3(k)\epsilon_3^T(k)](A^T)^k = 0. \tag{4.12}$$

Then, the process (4.1) can be mean squared stabilized by a suitable choice of the controller.

Proof: Let K be a stabilizing controller, i.e., let $A + BK$ be Schur-stable. Such a K exists since (A, B) is controllable. Now, use the controller

$$U(k) = K\hat{S}(k),$$

where

$$\hat{S}(k) = A^k \hat{S}_3(k) + \sum_{j=1}^{k} A^{k-j} BU(j-1). \tag{4.13}$$

With this controller, the process (4.1) evolves as

$$S(k+1) = (A + BK)S(k) + BK\delta(k), \tag{4.14}$$

where

$$\delta(k) = S(k) - \hat{S}(k) = -A^k \epsilon_3(k).$$

Since $S(0)$ is zero mean, equation (4.11) implies that $\mathbb{E}[\delta(k)] = 0$ and in turn $\mathbb{E}[S(k)] = 0$ at every k. Moreover, if (4.12) is satisfied, $\mathbb{E}[\delta(k)\delta^T(k)] \to 0$ as $k \to \infty$. Thus, since $A+BK$ is stable, (4.14) yields that $\lim_{k\to\infty} \mathbb{E}[S(k)S^T(k)] = 0$. Note that even though we use a certainty equivalence controller, the controller does not generate its estimate using a Kalman filter. This result also provides the controller design.

We now provide a coding scheme that ensures that the relations (4.11-4.12) are satisfied. For pedagogical ease, we begin by describing the scheme for the case $m = 1$. The extension to the vector case is provided later.

4.3.2 Scalar LTI plant

In this section, we consider the process to evolve as

$$S(k+1) = aS(k) + U(k), \tag{4.15}$$

with $a > 1$.

4.3.3 Basic Approach

In our coding scheme, the sensor, relay, and the controller implement a distributed stochastic approximation scheme to calculate the zeros of a function $\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^3$ whose i-th component is given by

$$f_i\left(T_1, T_2, T_3\right) = \begin{cases} 0 & \text{for } i=1 \\ \displaystyle\sum_{j:\ g_{j,i}\neq 0} \mu_i(T_i - T_j) & \text{for } i=2,3, \end{cases} \tag{4.16}$$

where $T_1 = S(0)$, the relay node executes the recursion (2.21) for the variable T_2, the controller node executes the recursion (2.21) for the variable T_3, and μ_i's are suitably designed to satisfy the transmission power constraints (4.5) and (4.6). Since the zeros of the function $\mathbf{f}$ are obtained at $T_1 = T_2 = T_3 = S(0)$, the estimates $\hat{T}_2(k)$ and $\hat{T}_3(k)$ serve as an estimate of the initial state $S(0)$ at the relay and the controller, respectively. By the properties of distributed stochastic approximation, both these quantities converge to $S(0)$ as $k \to \infty$ in the mean squared sense.

To better reflect the usage of T_i's as estimates, we define $\hat{S}_i(k) = \hat{T}_i(k)$, for $i = 2, 3$. Finally, define

$$\epsilon_2(k) := \hat{S}_2(k) - S(0) \tag{4.17}$$

$$\epsilon_{3,2}(k) := \hat{S}_3(k) - \hat{S}_2(k). \tag{4.18}$$

At every time k, the sensor can calculate $\epsilon_3(k)$, while the relay can calculate $\epsilon_{3,2}(k)$. Since we calculate an linear mmse estimate, our coding scheme will ensure that the quantities $\epsilon_3(k)$ and $\epsilon_{3,2}(k)$ are jointly Gaussian with zero means, covariances

$\alpha_3(k)$ and $\alpha_{3,2}(k)$ (respectively), and correlation coefficient

$$\rho(k) = \frac{\mathbb{E}[\epsilon_3(k)\epsilon_{3,2}(k)]}{\sqrt{\alpha_3(k)\alpha_{3,2}(k)}}. \tag{4.19}$$

The quantities $\alpha_2(k)$, $\alpha_3(k), \alpha_{3,2}(k)$ and $\rho(k)$ are deterministic and can be calculated prior to transmission by all the nodes. In particular,

$$\alpha_{3,2}(0) = \alpha_3(0) + \alpha_2(0)$$
$$\rho(0) = \frac{\alpha_3(0)}{\sqrt{(\alpha_3(0) + \alpha_2(0))\alpha_3(0)}}.$$

Recursive expressions to calculate these quantities are provided later.

4.3.4 Coding Scheme

The coding scheme is the same as used in Section 3.4.1 with $W = S(0)$. Detailed analysis can be found in [24] or Section 3.4.1. For future reference, we denote the coding scheme by $\mathcal{S}(S(0), \hat{S}_3(k))$, where $S(0)$ is the initial condition, and $\hat{S}_3(k)$ is the estimate of $S(0)$ at time k at the controller.

4.3.5 Stability Analysis

We now present the stability conditions when the coding scheme described above is used to stabilize the process (4.15).

Theorem 4.3.2 *Consider the problem formulation presented in Section 4.1 with the process given by (4.15) and the coding scheme and controller presented in Section 4.3.4. The process (4.15) is mean square stabilized over the AWGN relay* *channel if*

$$\ln(a) < \liminf_{n\to\infty} \frac{1}{2n} \sum_{k=1}^{n} \ln \frac{1}{q(k-1)}, \tag{4.20}$$

where $q(k-1)$ *is calculated as in Equation (3.54).*

Proof: Since $\epsilon_3(0) = \epsilon_{3,2}(0) = \epsilon_2(0) = 0$ and all updates in the coding scheme are linear, it is straight-forward to see that equation (4.11) is satisfied. Now, (3.53) yields

$$a^{2k}\alpha_3(k) = a^{2k}q(k-1)\alpha_3(k-1).$$

The ratio test then yields that a sufficient condition for equation (4.12) to be satisfied is that the condition in (4.20) be satisfied. Since the coding scheme in Section 4.3.4 uses the controller used in the proof of Theorem 4.3.1, the process (4.15) is mean square stabilized. Further, the bound in Equation (4.20) is non-trivial.

Corollary 4.3.3 *The right hand side of the equation (4.20) is finite.*

Proof: From (3.54), we see that $\rho(k-1)$ minimizes the numerator as well as maximizes the denominator. Thus, $1/q(k-1)$ is maximized by choosing $\rho(k-1) = 1$, or in other words,

$$\frac{1}{q(k-1)} \leq \frac{g_{1,3}^2 P_1 + g_{2,3}^2 P_2 + 2g_{1,3}g_{2,3}\sqrt{P_1P_2} + \sigma_3^2}{\sigma_3^2}.$$

This implies that

$$\liminf_{n\to\infty} \frac{1}{2n} \sum_{k=1}^{n} \ln \frac{1}{q(k-1)} \leq \frac{1}{2} \ln \left(1 + \frac{g_{1,3}^2 P_1 + g_{2,3}^2 P_2 + 2g_{1,3}g_{2,3}\sqrt{P_1P_2}}{\sigma_3^2} \right).$$

Notice that Theorem 4.3.2 has been presented in terms of how unstable the process can be for a given transmission power usage. However, we can also interpret it in terms of the minimum power required to stabilize a given process. While this is a minor point in the context of a scalar process, we will see that such a reinterpretation is quite useful for vector processes.

The special case of the cascade channel can be derived from the above results that have been presented for the relay channel.

Theorem 4.3.4 *The LTI system in (4.15) can be mean square stabilized over the cascade channel if (4.20) is satisfied with* $g_{1,3} = 0$.

Proof: The cascade channel is the same as the relay channel, except that there is no path from the sensor to the controller. The coding scheme and the analysis in Theorem 4.3.2 thus hold for the cascade channel with $g_{1,3} = 0$. There is a slight difference in the initialization step, which is similar to the initialization for the cascade channel in Section 3.3.

4.3.6 Vector LTI plant

In this Section, we consider the process in (4.1) to have dimension m arbitrary. Note that we only present the results for the relay channel in this section. The results for the cascade channel follow by substituting $g_{1,3} = 0$ as used earlier. Without loss of generality, we assume that the matrices A is in the modal form and can be expressed as

$$A = \begin{bmatrix} A_s & 0 \\ 0 & A_u \end{bmatrix},$$

where $A_s \in \mathbb{R}^{(m-l)\times(m-l)}$ and $A_u^{-1} \in \mathbb{R}^{l\times l}$ are Schur stable. Note that $0 \le l \le m$, and we assume that an empty A_s (resp. A_u) corresponds to $l = m$ (resp. $l = 0$).

4.3.7 Coding Scheme

The basic approach of the coding scheme for a vector process is to transmit the last l elements of the initial state $S(0)$ to the controller. To achieve this aim, l coding schemes proposed in Section 4.3.2 are used in parallel for each individual element. Thus, for each $j = 0, 1, \cdots, l-1$, at the sequence of times $kl+j$ ($k \in \mathbb{Z}_+$), the sensor, relay, and controller implement the coding scheme $\mathcal{S}(S^j(0), S_3^j(k))$, with $S^j(k) = e_j^T S(0)$ and $S_3^j(k) = e_j^T S_3(k)$. The controller calculates the control input as follows. It maintains an estimate $\hat{S}_3(k)$ of the initial state $S(0)$. At each time k, such that $j = k \mod l$, it performs the following actions:

- Update $\hat{S}_3(k)$ as

$$\hat{S}_3(k) = \hat{S}_3(k-1) - e_j^T \hat{S}_3(k-1) e_j + S_3^j(k) e_j,$$

  with the initial condition $\hat{S}_3(-1) = 0$.

- Calculate $\hat{S}(k)$ using the relation (4.13).

- Transmit both the control $U(k) = K\hat{S}(k)$ and the estimate $S_3^j(k)$.

4.3.8 Stability Analysis

We have the following stability characterization.

Theorem 4.3.5 *Consider the problem formulation presented in Section 4.1 with the coding scheme presented in Section 4.3.4 for each unstable state. The process (4.1) with l unstable states is mean square stabilized over the AWGN relay* *channel if the following condition is satisfied:*

$$\max_{1\le i\le m}\{0, \ln|\lambda_i(A)|\} < \liminf_{n\to\infty}\frac{1}{2nl}\sum_{k=1}^{n}\ln\frac{1}{q(k)}, \tag{4.21}$$

where

$$q(k-1) = \frac{\alpha_3(k)}{\alpha_3(k-1)}$$

is calculated as in Equation (3.54).

Proof: Since, A_s is stable, we do not need to update the first $m-l$ components of the error vector ϵ_3, which remain constant at 0. The other l components of ϵ_3 are updated every l time steps. Since $\epsilon_3(0) = 0$ and all updates in the coding scheme are linear, it is straightforward to see that equation (4.11) is satisfied. Since A_s is schur stable, A_s^k approaches 0 as $k \to \infty$. Using the result from Theorem 4.3.2, we can see that $\lambda_j^{2k}\mathbb{E}\left[(\epsilon_3^j(k))^2\right] \to 0$, where $m-l+1 \le j \le m$ and $\epsilon_3^j(k) = e_j^T\epsilon_3(k)$, if the following condition is satisfied:

$$\ln|\lambda_j(A)| < \liminf_{n\to\infty}\frac{1}{2nl}\sum_{k=1}^{n}\ln\frac{1}{q(k)},$$

where

$$q(k-1) = \frac{\alpha_3(k)}{\alpha_3(k-1)}$$

is calculated as in Equation (3.54). Note that since the diagonal elements in (4.12) approach 0 as $k \to \infty$, using Cauchy-Schwarz inequality, it follows that the non-diagonal elements in (4.12) also approach 0. The theorem follows by noting that the above condition needs to be satisfied for $\forall 1 \le i \le m$. The only parameter that depends on the choice of design parameters such as the transmission power is $q(k)$. Consequently, we can rewrite the above result as follows.

Corollary 4.3.6 *Consider the problem formulation presented in Section 4.1 with the coding scheme presented in Section 4.3.4 for each unstable state. The minimum powers at the sensor and the relay nodes that are needed for mean suqare stabilizing the process (4.1) with l unstable states must be sufficient to guarantee that*

$$\liminf_{n\to\infty} \frac{1}{2n} \sum_{k=1}^{n} \ln \frac{1}{q(k)} > 2l \max_{1\le j\le m} \{0, \ln|\lambda_j(A)|\}, \tag{4.22}$$

where

$$q(k-1) = \frac{\alpha_3(k)}{\alpha_3(k-1)}$$

is calculated as in Equation (3.54).

Proof: The proof follows directly from Theorem 4.3.5.

4.3.9 Constraint C_1

$\hat{S}_3(k)$ is a linear minimum mean squared error (mmse) estimate of $S(0)$, thus the controller satisfies the constraint C_1.

Proposition 4.3.7 *If $\hat{S}_3(k)$ is a linear minimum mean squared error (mmse) estimate of $S(0)$, and the process (4.1) is mean squared stabilized, then the controller proposed in Theorem 4.3.1 satisfies $\sum_{k=0}^{\infty} \mathbb{E}[U^T(k)U(k)] < \infty$.*

Proof: First note that if $\hat{S}_3(k)$ is a linear minimum mean squared error (mmse) estimate of $S(0)$, then $\hat{S}(k)$ in (4.13) is an MMSE estimate of $S(k)$. This implies that

$$\mathbb{E}[S(k)S^T(k)] = \mathbb{E}[\hat{S}(k)\hat{S}^T(k)] + \mathbb{E}[\delta(k)\delta^T(k)].$$

Thus,

$$\sum_{k=0}^{\infty} \mathbb{E}[U^T(k)U(k)] = \sum_{k=0}^{\infty} \mathbb{E}[\hat{S}^T(k)K^T K\hat{S}(k)]$$
$$= \sum_{k=0}^{\infty} \mathbb{E}[tr(\hat{S}(k)\hat{S}^T(k)K^T K)].$$

The above term is finite if $\sum_{k=0}^{\infty} \mathbb{E}[tr(\hat{S}(k)\hat{S}^T(k))]$ is finite. We have

$$\sum_{k=0}^{\infty} \mathbb{E}[tr(\hat{S}(k)\hat{S}^T(k))] = \sum_{k=0}^{\infty} \mathbb{E}[tr(S(k)S^T(k))] + \sum_{k=0}^{\infty} \mathbb{E}[tr(\delta(k)\delta^T(k))]$$

If the process (4.1) is mean squared stabilized, the first summation is finite. The second summation can be written as

$$\sum_{k=0}^{\infty} \mathbb{E}[tr(\delta(k)\delta^T(k))] = \sum_{k=0}^{\infty} \mathbb{E}[tr(A^k \epsilon_3(k)\epsilon_3^T(k)(A^T)^k)]$$
$$= \sum_{k=0}^{\infty} \sum_{j=1}^{m} \lambda_j^{2k} \mathbb{E}[(\epsilon_3^j(k))^2]$$
$$\leq \sum_{j=1}^{m} \sum_{k=0}^{\infty} \alpha_3(0) n \prod_{i=0}^{k-1} \lambda_j^{2n} q(i),$$

where the last step follows from the fact that a particular estimation error is updated every n time steps. If the condition in (4.21) is satisfied, the above term is finite.

4.4 Numerical Results

Consider the process (4.15) where the controller is placed at a (spatial) distance of 2 units from the sensor. The relay is placed at the midpoint of the line joining the sensor and controller. The system parameters are considered to be $P_1 = P_2 =$

$0.5, \sigma_{S(0)}^2 = \sigma_2^2 = \sigma_3^2 = 1$ and $\eta = 2$. All the logarithms are taken to the base e. Under these constraints, the limit in (4.20) can be evaluated to be 0.1671 for the relay channel. Under the same constraints, the limit for the cascade channel is 0.1995. If we consider that there is no relay node and use the power $P_1 + P_2 = 1$ at the sensor node, the sufficient condition for mean square stability (e.g., [7]) is $\ln(a) < 0.1116$. This shows that using a relay node may be beneficial even if the total power constraint on transmitting nodes remains the same, i.e., even if the power used by the relay node is at the expense of the power used by the sensor node.

To illustrate this we plot the limits in (4.20) as a function of the power P_1 and distance of the relay from the plant for the relay channel and cascade channel in Figures 4.3 and 4.4 respectively. The system parameters are the same as above. The total power is fixed at $P_1 + P_2 = 1$. As shown, for any given position of the relay, we can always find a power distribution for which the stability region is enhanced.

As in Section 3.5, we also plot stability conditions when half-duplex relays are considered in this chapter as compared to the full-duplex relaying. The relay would listen at every even time step and transmit at every odd time step and have variance update equation as in (3.61). Figure 4.5 illustrates the performance of half-duplex relaying.

4.5 Summary

In this chapter, we derived sufficient conditions for mean square stabilizability of a linear time invariant open loop unstable plant over a relay AWGN channel and a cascade of two point-to-point AWGN channels. We proposed a coding scheme,

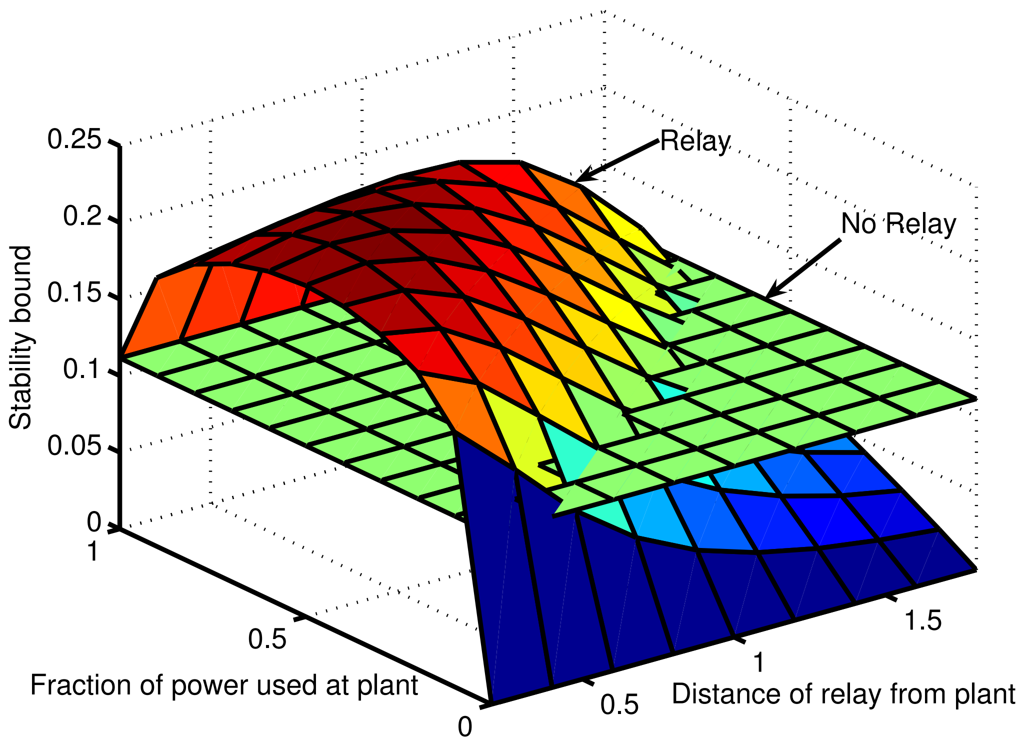

Figure 4.3. Sufficient condition bound as a function of fraction of power used at the plant and distance of relay from plant for a full-duplex relay channel.

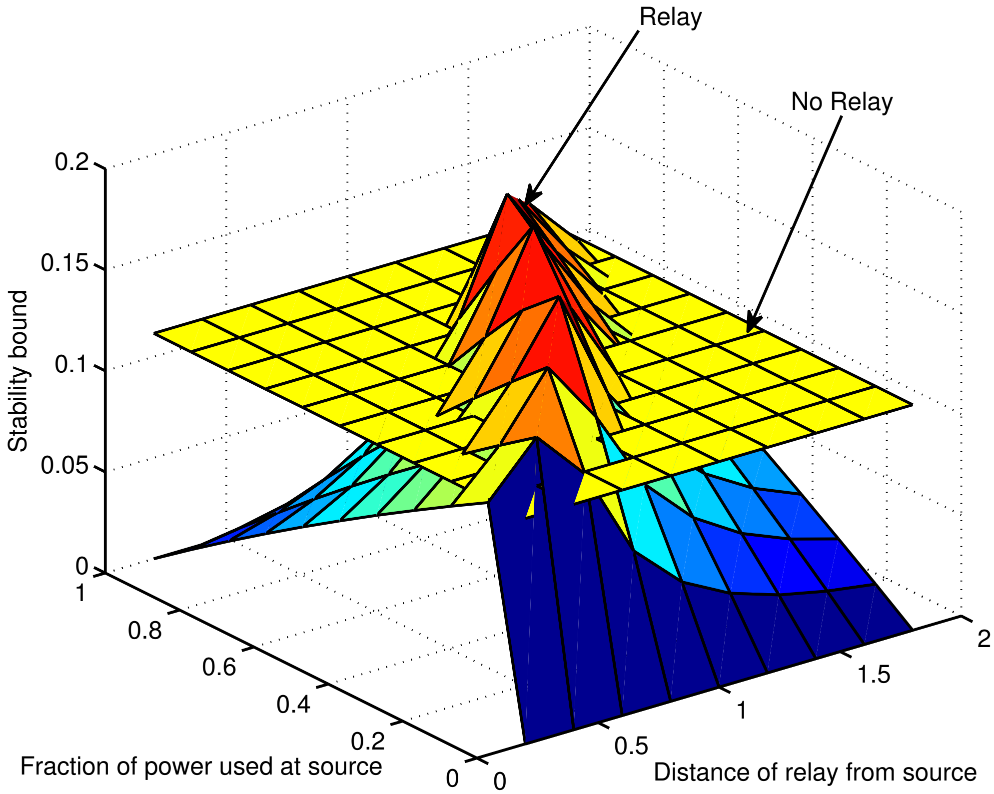

Figure 4.4. Sufficient condition bound as a function of fraction of power used at the plant and distance of relay from plant for a cascade channel.

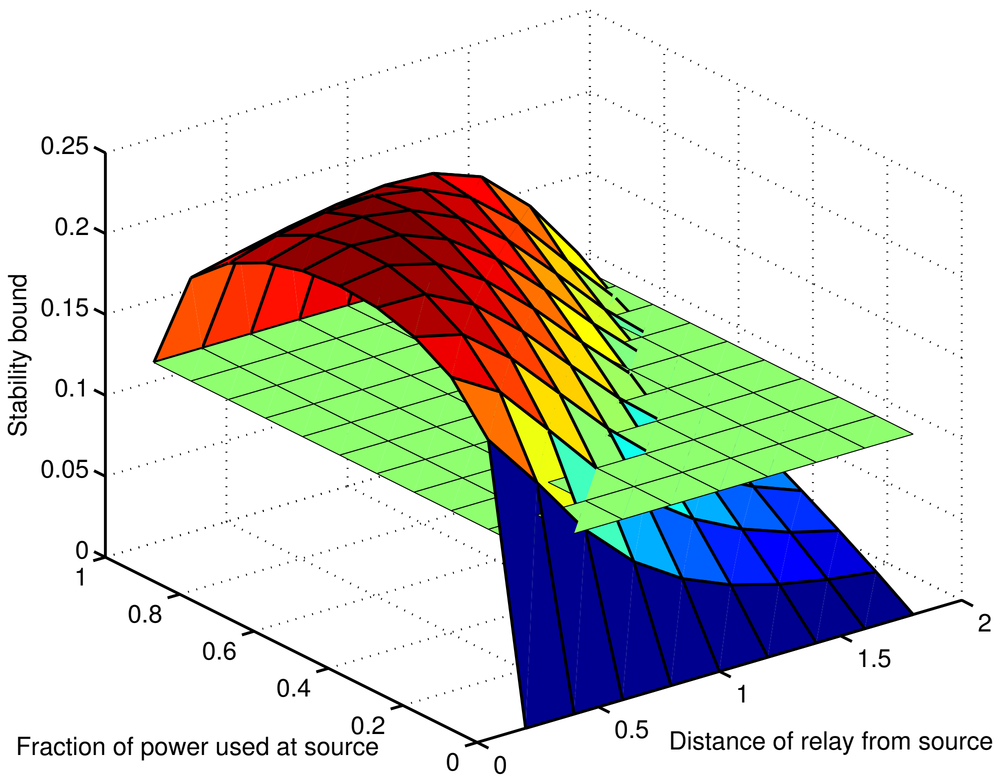

Figure 4.5. Sufficient condition bound as a function of fraction of power used at the plant and distance of relay from plant for a half-duplex relay channel.

which makes use of distributed stochastic approximation algorithms to ensure appropriately fast rate of convergence of the estimate of the initial state at the controller to the correct value. The scheme can also be interpreted as a distributed version of the Schalkwijk-Kailath coding scheme for point-to-point channels with both the sensor and the relay transmitting innovations with respect to the estimate at the controller at every step. We analyzed the stability region of the closed loop system with the proposed scheme under average transmission power constraints for both the sensor and the relay node. Interestingly, the stability region may be increased by using a relay node even if the total transmission power remains the same.

CHAPTER 5

STABILITY ACROSS A GAUSSIAN PRODUCT CHANNEL

In this chapter, we are interested in stabilizability of a scalar unstable linear time invariant (LTI) discrete time system across a Gaussian product channel (also known as parallel Gaussian channels). The Gaussian product channel models a continuous-time waveform Gaussian channel in which the transmitter sends information to the receiver across multiple parallel channels, each parallel channel being individually modeled by an additive white Gaussian noise channel. The parallel channels may represent different frequency bands, time instances, or in general different "degrees of freedom". Control across such a channel is inherently more difficult than control across a single channel. For instance, while it is known that for a single AWGN channel, the optimal encoding policies are linear [4, 7, 27, 45], for the parallel channel case, Yüksel and Taikonda [50] presented a counterexample which shows that in general linear encoding strategies may lead to overly restrictive stabilizability conditions[1]. Since design of optimal non-linear strategies is not trivial, analytical results and coding schemes for this setup are largely lacking. The inadequacy of linear controllers in achieving optimal stabilizability conditions or performance for a parallel Gaussian channel was also noted

[1]Note that while the general set-up in [50] considers a multi-sensor setting, the specific numerical example they provide is identical to the case when one sensor can transmit information across a Gaussian product channel with two parallel channels.

in [44]. However, neither [50] nor [44] proposed a non-linear controller and encoder structure.

By constructing one particular stabilizing non-linear encoder and decoder structure, we present sufficient conditions for stabilizing a scalar discrete-time LTI plant in the mean squared sense when the sensor transmits information to a remotely placed controller across a Gaussian product channel. When the sufficient conditions for stability using our coding scheme are satisfied with equality, data about the initial condition is transmitted at a rate equal to the capacity of a Gaussian product channel. We then derive the necessary conditions for stability using tools from information theory, thus giving a lower bound on the channel capacity required for stabilization in terms of the unstable eigenvalues of the plant.

The remainder of the chapter is organized as follows. We discuss the problem setup in Section 5.1. In Section 5.2 the encoder, decoder and controller design are presented, which gives us the sufficiency results. In Section 5.3, a proof the necessary condition is presented. Some directions for future work are outlined in Section 5.4.

5.1 Problem Setup

Consider the model shown in Fig. 5.1. An open loop unstable linear time invariant process evolves according to

$$S(k+1) = AS(k) + BU(k), \tag{5.1}$$

where $S(k) \in \mathbb{R}^l$ is the state and $U(k) \in \mathbb{R}$ is the control value. We assume that the state of the process is observable by a sensor. We assume that the pair (A, B) is controllable. We assume that each component $S^i(0)$ $(i = 1, 2, \ldots, l)$ of the initial condition $S(0)$ is a random variable uniformly distributed in the interval $[c, d]$ with a finite variance $\sigma^2_{S^i(0)} = \frac{(d-c)^2}{12}$.

For ease of exposition and without loss of generality, we assume that at every time step a sensor observes the state of the process $S(k)$ and transmits suitable information across the communication channel to the controller. The controller calculates a control input $U(k)$ and the actuator applies it to the process in (5.1). The communication channel from the plant to the controller is modeled as a Gaussian product channel, while the communication from the controller to the process is assumed to be noiseless. The Gaussian product channel consists of m channels. The input and output of the i-th channel at time k is denoted by $X_i(k)$ and $Y_i(k)$ respectively. The noise corrupting the i-th channel is denoted by Z_i. The output of the i-th channel at time k is given by

$$Y_i(k) = g_i X_i(k) + Z_i(k),$$

where g_i is an attenuation factor due to path loss. Each noise $Z_i(k)$ is modeled by a zero-mean AWGN process with mean zero and variance σ_i^2. The noises on the various links are assumed to be mutually independent and white. We impose three constraints on the encoder and controller design:

- *Constraint C_1:* The control action must satisfy a controller cost constraint, $\sum_{k=0}^{\infty} \mathbb{E}[U(k)^2] < \infty$.
- *Constraint C_2:* There is an average power constraint imposed on the signals transmitted by the encoders on the different channels and a common power constraint on the total power used. Thus, the encoding schemes must be

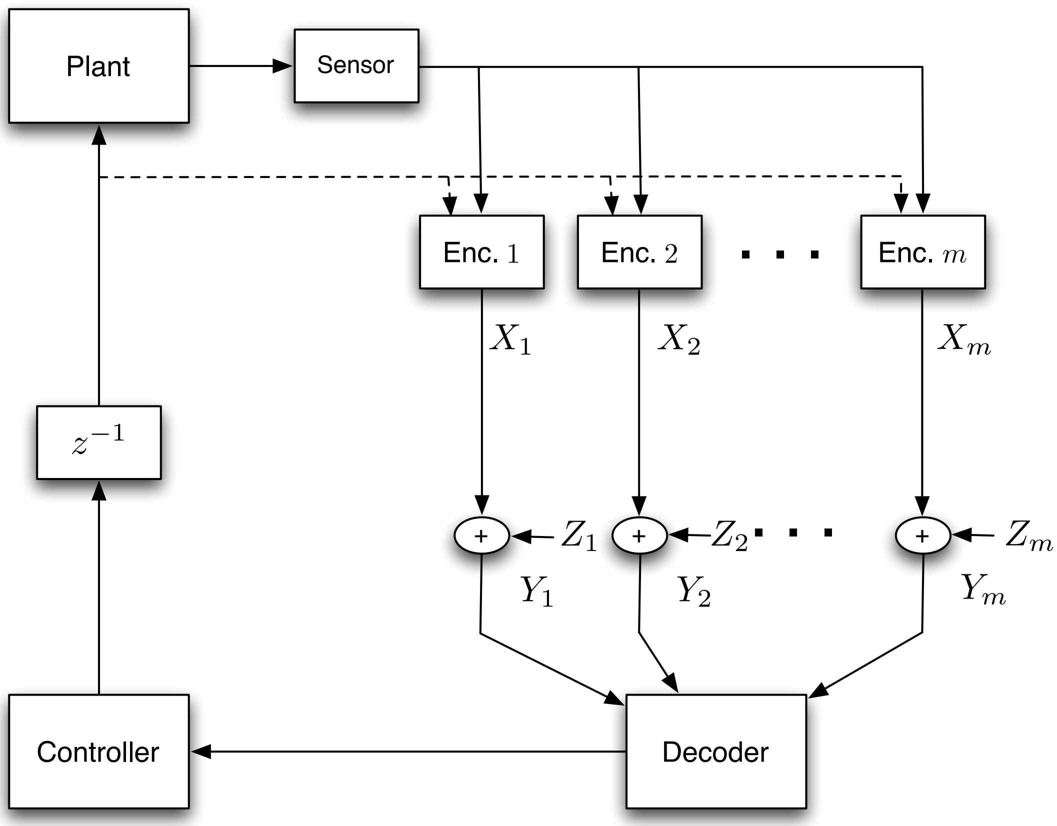

Figure 5.1. Model for an unstable plant being controlled by a controller across a Gaussian product channel

such that the transmitted signals satisfy

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{E}[X_i^2(k)] \le P_i, \; i = 1, 2, \ldots m,$$

$$\sum_{i=1}^{m} P_i \le P.$$

- *Constraint C_3:* The encoders are assumed to be causal, but otherwise unconstrained in terms of computation and memory. The information structure at the encoders is as follows. If h_i is the encoding policy at the encoder for the i-th input, then

$$X_i(k) = h_i(S(0), \ldots, S(k), U(0), \ldots, U(k-1)).$$

Notice that we have assumed access to the control input at every encoder. Such an assumption makes sense since given the state values at time k and $k+1$, the sensor can calculate $U(k)$. In our construction, the controller will transmit additional information besides the control input. We assume that the actuator can isolate the correct component of the transmitted vector and apply it to the process.

5.1.1 Problem Statement

The problem we are interested in this chapter is two fold. First, we want to design the maps h_i's and controller $U(k)$ so that the process (5.1) is mean square stabilized, while satisfying the design constraints C_1, C_2 and C_3, thus giving us sufficient conditions for stability. The design of the encoder map involves designing a scheme to divide the total power amongst the various inputs $X_1, X_2, \ldots X_m$ in an optimal way. Recall that a system is said to be stabilized in the mean squared sense if and only if, irrespective of the initial state $S(0)$, the following conditions are satisfied:

$$\begin{aligned} \mathbb{E}[S(k)] &= 0, \\ \lim_{k\to\infty} \mathbb{E}[S(k)S^T(k)] &= 0. \end{aligned} \tag{5.2}$$

Secondly, we are also interested in characterizing the necessary conditions of stability. We do this by using tools from information theory to lower bound the second moment of the state of the plant.

5.1.2 Main Result:

The following is the main result of the chapter:

Theorem 5.1.1 *Consider the problem formulation presented in Section 5.1. The necessary and sufficient condition for stabilizing the process (5.1) in mean square sense over the Gaussian product channel is:*

$$\sum_{i=1}^{l} \max\{0, \ln|\lambda_i(A)|\} < \max \sum_{i=1}^{m} \frac{1}{2}\ln\left(1+\frac{g_i^2 P_i}{\sigma_i^2}\right), \tag{5.3}$$

where the maximization is over power allocations satisfying $\sum_{i=1}^{m} P_i = P$.

We prove the above theorem in two parts: Section 5.2 presents a design that achieves stability if (5.3) is satisfied and the necessity of (5.3) for stability is proven in Section 5.3.

Remark 5.1.2 *The condition shown to be necessary for stability in Theorem 5.1.1 is seen to be less restrictive than the conditions shown to be necessary in [44, Theorem 6] This difference arises because [44] concentrates on the class of linear en-*

coders, decoders, and controllers. Since we allow for non-linear structures, we obtain less restrictive necessary conditions. Further, the condition in Theorem 5.1.1 is also sufficient for stabilizability. In that sense, this is a tight characterization of the stabilizability region.

5.2 Sufficiency Results

In this section, we prove that if the condition (5.3) holds, then the system (5.1) can be stabilized. Our solution consists of two parts. First, we prove that if the controller has access to an estimate $\hat{S}(k)$ of the initial condition $S(0)$ of (5.1), such that the error

$$\epsilon(k) := \hat{S}(k) - S(0) \tag{5.4}$$

has a covariance matrix that decreases geometrically in k, then the process can be stabilized by a suitable design of the controller. Then, we construct an encoding scheme that ensures that such an estimate is available to the controller. We begin with the first part.

5.2.1 Controller Design

The controller design is same as for the relay channel in Chapter 4 and reader is referred to Section 4.3.1.

We now provide a coding scheme that ensures that the relations (4.11) and (4.12) are satisfied. To set the notation, we begin by describing the scheme for the case of state dimension $l = 1$. The extension to the vector case is provided later.

5.2.2 Scalar LTI Plant

In this section, we consider the process (5.1) to evolve as

$$S(k+1) = aS(k) + U(k), \tag{5.5}$$

with $|a| > 1$.

The encoders distribute the information about $S(0)$ amongst the various channel inputs $X_1(k), X_2(k), \ldots, X_m(k)$. Let $S_i(k)$ be the information about $S(0)$ transmitted through the i-th ($i \in \{1, 2, \ldots, m\}$) channel. $X_i(k)$ is calculated by scaling $S_i(k)$ to satisfy the power constraint C_2. The controller observes the outputs of the channels $Y_1(k), Y_2(k), \ldots, Y_m(k)$. and extracts relevant information $\hat{S}_1(k), \hat{S}_2(k), \ldots, \hat{S}_m(k)$. The estimate $\hat{S}(k)$ of the initial state $S(0)$ can be then calculated at the decoder as a function of the information collected from different links, i.e.,

$$\hat{S}(k) = f(k, \hat{S}_1(k), \hat{S}_2(k), \ldots, \hat{S}_m(k)).$$

We defined the overall estimation error in (5.4). We let the estimation error for information sent through the i-th channel be denoted by

$$\epsilon_i(k) := \hat{S}_i(k) - S_i(0).$$

Finally, let $\alpha(k)$ (resp. $\alpha_i(k)$) represent the variance of the estimation error $\epsilon(k)$ (resp. $\epsilon_i(k)$). The crucial property that needs to be satisfied for mean square stability is that the estimate $\hat{S}(k)$ converges to $S(0)$ at a rapid enough rate as shown by Theorem 4.3.1.

We will now present our coding scheme and show that it satisfies the constraints in Theorem (4.3.1). We start with considering the special case when we have only one channel ($m = 1$), extend to the case when there are two channels ($m = 2$) ,and then generalize it for m channels. Note that the case $m = 1$ is similar to the SK scheme presented in section 2.2.1, but we reproduce the derivation for notational convenience.

5.2.2.1 Special Case m = 1

The coding scheme for the case when $m = 1$ works as follows [7, 45]. Note that since there is only one channel, $S_1(0) = S(0), \hat{S}_1(k) = \hat{S}(k), \epsilon_1(k) = \epsilon(k)$ and $P_1 = P$.

Initialization: At time step $k = 0$, the encoder transmits

$$X_1(0) = \sqrt{\frac{P_1}{\sigma^2_{S_1(0)}}} S_1(0). \tag{5.6}$$

The decoder forms an estimate of $S_1(0)$ as follows:

$$\hat{S}_1(0) = \frac{1}{g_1}\sqrt{\frac{\sigma^2_{S_1(0)}}{P_1}} Y_1(0).$$

The estimation error $\epsilon_1(0)$ is given by

$$\epsilon_1(0) = \frac{1}{g_1}\sqrt{\frac{\sigma^2_{S_1(0)}}{P_1}} Z_1(0).$$

Clearly, $\epsilon_1(0)$ is zero-mean Gaussian with variance $\alpha_1(0)$, given by

$$\alpha(0) = \alpha_1(0) = \frac{\sigma^2_{S_1(0)}\sigma_1^2}{g_1^2 P_1}. \tag{5.7}$$

The controller calculates the control input according to (4.13) and transmits both the control $U(0) = K\bar{S}(0)$ and the estimate $\hat{S}(0)$ to the process.

Update: At each time step $k \geq 1$, the encoder transmits

$$X_1(k) = \sqrt{\frac{P_1}{\alpha_1(k-1)}}\epsilon_1(k-1). \tag{5.8}$$

The decoder updates its estimate as follows. At time $k \geq 1$, the decoder calculates the linear minimum mean squared error (MMSE) estimate of $S_1(0)$ given $Y_1(k)$ and $\hat{S}_1(k-1)$ as

$$\hat{S}_1(k) = \hat{S}_1(k-1) - \frac{\mathbb{E}[Y_1(k)\epsilon_1(k-1)]}{\mathbb{E}[{Y_1}^2(k)]}Y_1(k). \tag{5.9}$$

The controller calculates the control input according to (4.13) and transmits both the control $U(k) = K\bar{S}(k)$ and the estimate $\hat{S}(k)$ to the process. Note that the input $X_1(k)$ satisfies the power constraint C_2.

It can be seen that the estimation error $\epsilon_1(k)$ is Gaussian with zero mean and variance $\alpha_1(k)$. We now proceed to evaluate the recursive expression for $\alpha_1(k)$ as used in the coding scheme presented above. Since $\epsilon_1(k)$ is defined as $\hat{S}_1(k) - S_1(0)$, from (5.9) we obtain

$$\epsilon_1(k) = \epsilon_1(k-1) - \frac{\mathbb{E}[Y_1(k)\epsilon_1(k-1)]}{\mathbb{E}[{Y_1}^2(k)]}Y_1(k). \tag{5.10}$$

The variance of $\epsilon_1(k)$ can be obtained as

$$\alpha_1(k) = \mathbb{E}[\epsilon_1^2(k)] = \alpha_1(k-1) - \frac{\mathbb{E}^2[Y_1(k)\epsilon_1(k-1)]}{\mathbb{E}[{Y_1}^2(k)]}, \tag{5.11}$$

with the initial condition in (5.7). The terms in (5.11) can be further evaluated to be

$$\mathbb{E}[{Y_1}^2(k)] = g_1^2 P_1 + \sigma_1^2, \tag{5.12}$$

and

$$\mathbb{E}[Y_1(k)\epsilon_1(k-1)] = g_1\sqrt{P_1\alpha_1(k-1)}. \tag{5.13}$$

Using (5.12) and (5.13) in (5.11), we obtain

$$\alpha_1(k) = \alpha_1(k-1)r_1, \tag{5.14}$$

where $r_1 = \left(\frac{\sigma_1^2}{g_1^2 P_1 + \sigma_1^2}\right)$. Note that since $\alpha(k) = \alpha_1(k)$ and $P_1 = P$,

$$\alpha(k) = \frac{\sigma_{S(0)}^2 \sigma_1^2}{g_1^2 P}\left(\frac{\sigma_1^2}{g_1^2 P + \sigma_1^2}\right)^k. \tag{5.15}$$

We now present the stability conditions when the coding scheme described above is used to stabilize the process (5.5), same as Theorem 4.3.2.

Corollary 5.2.1 *Consider the problem formulation presented in Section 5.1 with the controller structure and coding scheme presented above for a scalar plant and no of channels $m = 1$. The process (5.5) is mean square stabilized if*

$$\ln(a) < \frac{1}{2}\ln\left(1 + \frac{g_1^2 P}{\sigma_1^2}\right). \tag{5.16}$$

Note that for a scalar system as we are considering, if encoders have access to control law K, and the system matrices A and B, then they can calculate $\hat{S}(k)$ from $U(k)$.

5.2.2.2 Special Case m=2

We consider next the case when $m = 2$ before we present the scheme for an arbitrary m. To develop a coding scheme for the case when more than one channel is present, we revisit a relevant result from information theory [39], and recognized also in [50]. For a distributed source-channel coding to be optimal in the information-theoretic sense, two conditions need to be satisfied:

- The information transmitted on all the channels should be independent.
- Information is being transmitted through all the channels at respective channel capacities (the source and the channel need to be matched).

As explained in [50], one of the reasons linear schemes are not optimal is that it is not possible to make the signals transmitted on different channels independent when linear schemes are used. We develop a non-linear encoding scheme which will ensure that we transmit independent information over the two channels.

For transmission over the Gaussian product channel with $m = 2$, consider the following construction. Recall that $S(0)$ is uniformly distributed over $[c, d]$. Divide the interval $[c, d]$ into M_1 (fix an integer M_1 whose value will be chosen later) disjoint, equal-length message intervals as shown in Fig. 5.2. Let c_j where $j \in \{1, 2, \ldots, M_1 - 1\}$ be the partition points. Also define $c_0 := c$ and $c_{M_1} := d$. A point x is said to be in the j-th interval I_j if $x \in [c_{j-1}, c_j]$. Consider a quantizer $T_1(\cdot)$ which maps each point of the j-th interval ($j \in \{1, 2, \ldots, m\}$) to the midpoint of that particular interval. The message to be sent on the first channel is $T_1(S(0))$. The message to be sent on the second channel corresponds to the quantization

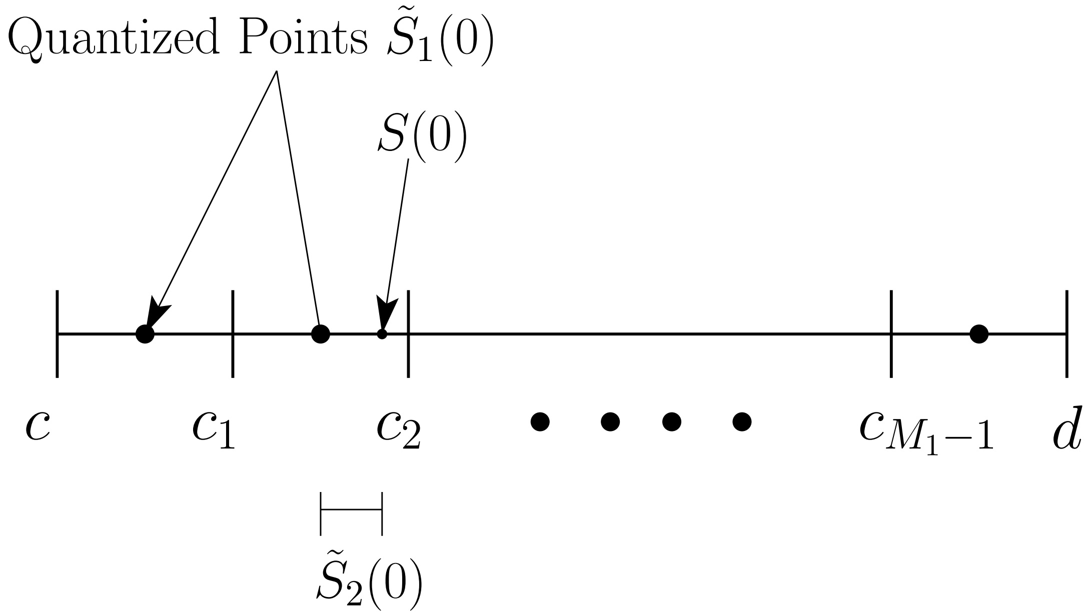

Figure 5.2. Generation of $\tilde{S}_1$ and $\tilde{S}_2$

error $S(0) - T_1(S(0))$. Thus we design the quantizer $T_1(\cdot)$ as follows:

$$\begin{aligned}\tilde{S}_1(0) &= T_1(S(0)) \\ &\stackrel{(a)}{=} c + \left(j - \frac{1}{2}\right)\frac{1}{M_1} \text{ if } S(0) \in I_j, \\ \tilde{S}_2(0) &= S(0) - \tilde{S}_1(0).\end{aligned} \tag{5.17}$$

Lemma 5.2.2 *The random variables $\tilde{S}_1$ and $\tilde{S}_2$ defined in (5.17) are independent.*

Proof: Consider the conditional probability $\Pr(\tilde{S}_2 = \gamma | \tilde{S}_1 \in I_j)$.

$$\begin{aligned}\Pr(\tilde{S}_2 = \gamma | \tilde{S}_1 \in I_j) &= \frac{\Pr(\tilde{S}_2 = \gamma, \tilde{S}_1 \in I_j)}{\Pr(\tilde{S}_1 \in I_j)} \\ &= M_1 \Pr(\tilde{S}_2 = \gamma, \tilde{S}_1 \in I_j) \\ &\overset{(a)}{=} \sum_{j=1}^{M_1} \Pr(\tilde{S}_2 = \gamma, \tilde{S}_1 \in I_j) \\ &= \Pr(\tilde{S}_2 = \gamma),\end{aligned}$$

where (a) follows using the fact that $\Pr(\tilde{S}_2 = \gamma, \tilde{S}_1 \in I_j)$ is same for all j's. Thus, the random variables are independent. Now, we define the information to be sent on the two parallel channels as

$$S_1(0) \triangleq \tilde{S}_1(0), \;\; S_2(0) \triangleq M_1 \tilde{S}_2(0).$$

It can be seen that $S_1(0)$ takes values uniformly from a set with M_1 elements. The number of intervals M_1 is related the information rate R_1 and number of channel uses k over the first channel as

$$M_1 = 2^{kR_1}. \tag{5.18}$$

Note also that $S_2(0)$ is uniformly distributed in the interval $[-\frac{|d-c|}{2}, \frac{|d-c|}{2}]$ and thus has a variance $\sigma^2_{S_2(0)} = \frac{(d-c)^2}{12}$.

Now we transmited the messages $S_1(0)$ and $S_2(0)$ over the two channels recursively and independently of each other using the encoding scheme used for transmitting $S_1(0)$ in the case $m = 1$ in Section 5.2.2.1 (see (5.6) and (5.8)). The decoder forms estimates $\hat{S}_i(k)$ of $S_i(0)$, $i = 1, 2$, the variances of which can be written using (5.15) as

$$\alpha_1(k) = \frac{\sigma^2_{S_1(0)}\sigma_1^2}{g_1^2 P_1}\left(\frac{\sigma_1^2}{g_1^2 P_1 + \sigma_1^2}\right)^k \triangleq \alpha_1(0) r_1^k, \tag{5.19}$$

$$\alpha_2(k) = \frac{\sigma^2_{S_2(0)}\sigma_2^2}{g_2^2 P_2}\left(\frac{\sigma_2^2}{g_2^2 P_2 + \sigma_2^2}\right)^k \triangleq \alpha_2(0) r_2^k. \tag{5.20}$$

Note that the estimation errors do not depend on the control inputs, and hence does not effect the controller design. The control input in this case is calculated as

$$U(k) = \left[\hat{S}_1(k), \hat{S}_2(k), K\bar{S}(k)\right]^T, \tag{5.21}$$

where the third component $K\bar{S}(k)$ is calculated using (4.13). The third component of $U(k)$ as defined above is extracted and applied to the process (5.5) by the actuator, whereas the i-th component ($i = 1, 2$) is used by the encoder i to update the i-th input, as given by equation (5.8). Note that because of the construction described above (5.17), the information sent on the parallel channels $i = 1, 2, \ldots, m$ are mutually independent. Also, except at time step $k = 0$, the inputs to both the channels have a Gaussian distribution and are thus matched to the respective Gaussian channels. We now show that the condition (5.3) is sufficient for stability with this construction.

Corollary 5.2.3 *Consider the problem formulation presented in Section 5.1 with the coding scheme presented above for $m = 2$. The process (5.5) is mean square stabilized over the Gaussian product channel with $m = 2$ if*

$$\ln(a) < \sum_{i=1}^{2} \frac{1}{2} \ln\left(1 + \frac{g_i^2 P_i}{\sigma_i^2}\right), \tag{5.22}$$

where the maximization is over power allocations satisfying $\sum_{i=1}^{2} P_i = P$.

Proof: It is easy to see that $\mathbb{E}[\epsilon(0)] = \mathbb{E}[\epsilon_1(0) + \frac{\epsilon_2(0)}{M_1}] = 0$. It is known that the linear minimum mean squared error is an unbiased estimator. Thus, $\mathbb{E}[\epsilon(k)] = 0$ for all $k \geq 0$, which is the first condition in (4.12). To evaluate the estimation error variance $\alpha(k)$, we write

$$\begin{aligned}\mathbb{E}[\epsilon^2(k)] &\overset{(a)}{=} \Pr(\hat{S}_1(k) \neq S_1(0))\mathbb{E}[\epsilon^2(k)|\hat{S}_1(k) \neq S_1(0)] \\ &\quad + \Pr(\hat{S}_1(k) = S_1(0))\mathbb{E}[\epsilon^2(k)|\hat{S}_1(k) = S_1(0)]. \end{aligned} \tag{5.23}$$

The terms above can be written or bounded as follows.

$$\begin{aligned}\Pr(\hat{S}_1(k) \neq S_1(0)) &\leq \Pr\left[|\epsilon_1(n)| > \frac{1}{2M_1}\right] \\ &= 2Q\left(\frac{1}{2M_1\sqrt{\alpha_1(n)}}\right) \\ &\overset{(a)}{=} 2Q\left(\frac{2^{k\left(R_1 - \frac{1}{2}\ln(1+\frac{g_1^2 P_1}{\sigma_1^2})\right)}}{2\sqrt{\frac{\sigma_{S_1(0)}^2 \sigma_1^2}{g_1^2 P_1}}}\right), \\ \mathbb{E}[\epsilon^2(k)|\hat{S}_1(k) \neq S_1(0)] &\overset{(b)}{\leq} (d-c)^2, \\ \Pr(\hat{S}_1(k) = S_1(0)) &\leq 1, \\ \mathbb{E}[\epsilon^2(k)|\hat{S}_1(k) = S_1(0)] &= \frac{\alpha_2(k)}{M_1^2}, \\ &\overset{(c)}{=} \frac{\sigma_2^2}{12 g_2^2 P_2 2^{2kR_1}}\left(\frac{\sigma_2^2}{g_2^2 P_2 + \sigma_2^2}\right)^k, \end{aligned}$$

where $Q(x) \triangleq \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp(-\frac{y^2}{2})\, \mathrm{d}y$, (a) follows from (5.18) and (5.19), (b) follows using the fact the estimation error is upper bounded by the maximum distance between any two points on $[d, c]$ and (c) follows from (5.18) and (5.20). Thus, we can upper bound $a^{2k}\mathbb{E}[\epsilon^2(k)]$ as

$$a^{2k}\mathbb{E}[\epsilon^2(k)] \leq a^{2k} Q\left(\frac{2^{k\left(R_1 - \frac{1}{2}\ln(1+\frac{g_1^2 P_1}{\sigma_1^2})\right)}}{2\sqrt{\frac{\sigma_{S_1(0)}^2 \sigma_1^2}{g_1^2 P_1}}}\right)(d-c) + \frac{\sigma_2^2}{12 g_2^2 P_2} \frac{a^{2k}}{2^{2kR_1}} \left(\frac{\sigma_2^2}{g_2^2 P_2 + \sigma_2^2}\right)^k. \quad (5.24)$$

Since $Q(x) \sim \exp(\frac{-x^2}{2})$ for large x, the $Q(\cdot)$ term in (5.24) decreases doubly exponentially in k. On the other hand, the term a^{2k} increases exponentially. This implies that if $R_1 < \frac{1}{2}\ln\left(1 + \frac{g_1^2 P_1}{\sigma_1^2}\right)$, the first term in (5.24) goes to zero as $k \to \infty$ irrespective of the value of a. Moreover, if

$$\begin{aligned} \ln a &< R_1 + \frac{1}{2}\ln\left(1 + \frac{g_2^2 P_2}{\sigma_2^2}\right) \\ &< \frac{1}{2}\ln\left(1 + \frac{g_1^2 P_1}{\sigma_1^2}\right) + \frac{1}{2}\ln\left(1 + \frac{g_2^2 P_2}{\sigma_2^2}\right), \end{aligned}$$

then the second term in (5.24) also approaches zero as $k \to \infty$. Now since we are allowed to choose P_1 and P_2 while satisfying the total power constraint in C_2, we can optimize the right hand side of the above equation to increase the stability region. Thus, if the condition in (5.22) is satisfied, then $a^{2k}\alpha(k) \to 0$ and mean square stability is obtained using the controller design. The optimization in (5.22) can be carried out using Lagrange multipliers and is a standard result in information theory [12, Chapter 10]. An interpretation of the optimization is presented in Section 5.2.6, after discussing the coding scheme for the case when $m > 2$ channels are present.

5.2.2.3 General Case: Arbitrary value of m

The encoding scheme presented above can be generalized for transmission over the Gaussian product channel with $m > 2$ channels as follows. Divide the interval $[c, d]$ into M_1 disjoint, equal-length message intervals. Then divide each of these M_1 intervals into a further M_2 subintervals and so on till M_{m-1}. Define I_j^i to be the j-th interval ($j \in \{1, 2, \ldots, M_i\}$) for the i-th ($i \in \{1, 2, \ldots, m-1\}$) level quantizer Q_i that maps the message in each point of the interval I_j^i to the midpoint of that interval. The message to be sent on the i-th channel ($i = 1, 2, \ldots, m-1$) corresponds to the output of the i-th quantizer $Q_i(\cdot)$. The message to be sent on the m-th channel corresponds to the quantization error. Thus we design a set of $m-1$ quantizers as follows:

$$
\begin{aligned}
\tilde{S}_1(0) &= Q_1(S(0)), \\
\tilde{S}_2(0) &= Q_2(S(0) - \tilde{S}_1(0)), \\
&\vdots \\
\tilde{S}_{m-1}(0) &= Q_{m-1}(S(0) - \sum_{i=1}^{m-2} \tilde{S}_i(0)), \\
\tilde{S}_m(0) &= S(0) - \sum_{i=1}^{m-1} \tilde{S}_i(0).
\end{aligned}
$$

The following lemma is a generalization of the lemma 5.2.2 for arbitrary m presented without proof.

Lemma 5.2.4 *The random variables $\tilde{S}_i$, $i = 1, 2, \ldots, m$ defined in (5.17) are mutually independent.*

Now, we define the messages to be sent on the i-th ($i = 1, \ldots, m$) parallel channel as

$$S_i(0) \triangleq \left(\prod_{j=1}^{i-1} M_j \right) \tilde{S}_i(0).$$

Note that $S_i(0)$, $i = 1, 2, \ldots, m-1$ takes values uniformly from a set with M_i elements. As before, M_i is related to the information rate R_i and number of channel uses k over the i-th channel as

$$M_i = 2^{kR_i}, \; i = 1, 2, \ldots, m-1. \tag{5.25}$$

Also, $\tilde{S}_m(0)$ is uniformly distributed in the interval $[-\frac{|d-c|}{2}, \frac{|d-c|}{2}]$ and thus has a variance of $\frac{(d-c)^2}{12}$.

The encoding scheme is as follows. We transmit the messages $S_i(0)$, $i = 1, 2, \ldots, m$ over the m channels recursively in the same way as we transmitted $S_1(0)$ in for the case $m = 1$ (see (5.6) and (5.8)). The decoder forms estimates $\hat{S}_i(k)$ of $S_i(0)$, $i = 1, 2, \ldots, m$, the variances of which can written down using (5.15) as

$$\alpha_i(k) = \frac{\sigma^2_{S_i(0)} \sigma_i^2}{g_i^2 P_i} \left(\frac{\sigma_i^2}{g_i^2 P_i + \sigma_i^2} \right)^k \triangleq \alpha_i(0) r_i^k. \tag{5.26}$$

For future reference, we denote the coding scheme by $\mathcal{S}(S(0), \hat{S}(k), m)$, where $S(0)$ is the initial condition, $\hat{S}(k)$ is the estimate of $S(0)$ at time k at the controller and m is the number of parallel channels.

The controller design is as follow. The controller calculates and transmits the input

$$U(k) = \left[\hat{S}_1(k), \ldots, \hat{S}_m(k), K\bar{S}(k) \right]^T, \tag{5.27}$$

where the last component is calculated using (4.13). The $m+1$-th component of $U(k)$ defined above is extracted and applied to the process (5.5), whereas the i-th component ($1 \leq i \leq m$) is used by the encoder i to update the i-th input. Note that because of the construction described above, the random variables transmitted on the parallel channels $i = 1, 2, \ldots, m$ are mutually independent. Also, except at time step $k = 0$, the inputs to both channels have a Gaussian distribution and are thus matched to the respective Gaussian channels. We can now prove the sufficiency part of Theorem 5.1.1.

Corollary 5.2.5 *Consider the problem formulation presented in Section 5.1 for a scalar plant with the coding scheme presented above for an arbitrary number of channels m. The process (5.5) is mean square stabilized over the Gaussian product channel if*

$$\ln(a) < \max \sum_{i=1}^{m} \frac{1}{2} \ln\left(1 + \frac{g_i^2 P_i}{\sigma_i^2}\right), \tag{5.28}$$

where the maximization is over power allocations satisfying $\sum_{i=1}^{m} P_i = P$.

Proof:

We use the encoder, decoder and controller design outlined above. It is easy to see that $\mathbb{E}[\epsilon(0)] = 0$. It is known that the linear minimum mean squared error is an unbiased estimator. Thus, $\mathbb{E}[\epsilon(k)] = 0$ for all $k \geq 0$, which is the first condition in (4.12). Define the event $E := (\hat{S}_1(k) = S_1(0), \hat{S}_2(k) = S_2(0), \ldots, \hat{S}_{m-1}(k) = S_{m-1}(0))$. We can write the estimation error variance $\alpha(k)$ as

$$\begin{aligned} \mathbb{E}[\epsilon^2(k)] &= \Pr(\bar{E})\mathbb{E}[\epsilon^2(k)|\bar{E}] + \Pr(E)\mathbb{E}[\epsilon^2(k)|E] \\ \Rightarrow a^{2k}\mathbb{E}[\epsilon^2(k)] &= a^{2k}(\Pr(\bar{E})\mathbb{E}[\epsilon^2(k)|\bar{E}] + \Pr(E)\mathbb{E}[\epsilon^2(k)|E]) \end{aligned} \tag{5.29}$$

Using arguments similar to the proof of Corollary 5.2.3, we can prove that the first term in (5.29) goes to zero irrespective of the value of a if the following conditions are satisfied simultaneously.

$$R_i < \frac{1}{2} \ln \left(1 + \frac{g_i^2 P_i}{\sigma_i^2}\right) \forall i = 1, 2, \ldots, m-1. \tag{5.30}$$

To obtain a condition for the second term to approach zero, rewrite

$$\begin{aligned} a^{2k}\Pr(E)\mathbb{E}[\epsilon^2(k)|E] \leq a^{2k}\mathbb{E}[\epsilon^2(k)|E] &= \frac{a^{2k}\alpha_m(k)}{\prod_{i=1}^{m-1} M_i} \\ &= \frac{\sigma_m^2}{12 g_m^2 P_m} \frac{a^{2k}}{\prod_{i=1}^{m-1} 2^{2kR_i}} \left(\frac{\sigma_m^2}{g_m^2 P_m + \sigma_m^2}\right)^k. \end{aligned}$$

Thus, a sufficient condition for the term to approach zero is that

$$\begin{aligned} \ln a &< \sum_{i=1}^{m-1} R_i + \frac{1}{2} \ln \left(1 + \frac{g_m^2 P_m}{\sigma_m^2}\right) \\ &< \sum_{i=1}^{m} \frac{1}{2} \ln \left(1 + \frac{g_i^2 P_i}{\sigma_i^2}\right), \end{aligned}$$

where the last inequality follows using (5.30). Now since we are allowed to choose P_i while satisfying the total power constraint in C_2, we can optimize the right hand side of the above equation to increase the stability region. Thus, if the condition in (5.3) is satisfied, then $a^{2k}\alpha(k) \to 0$ and mean square stability is obtained.

5.2.3 Vector LTI plant

In this section, we consider the process in (5.1) to have dimension an arbitrary l. Without loss of generality, we assume that the matrice A is in the modal form and can be expressed as

$$A = \begin{bmatrix} A_s & 0 \\ 0 & A_u \end{bmatrix}, \tag{5.31}$$

where $A_s \in \mathbb{R}^{(l-n)\times(l-n)}$ and $A_u^{-1} \in \mathbb{R}^{n\times n}$ are Schur stable. Note that $0 \leq n \leq l$, and we assume that an empty A_s (resp. A_u) corresponds to $n = l$ (resp. $n = 0$). Divide the state $S(k)$ into corresponding parts

$$S(k) = \begin{bmatrix} S_s(k) \\ S_u(k) \end{bmatrix},$$

where $S_s(k)$ is formed by the first $l - n$ components of $S(k)$ and $S_u(k)$ by the last n components.

5.2.4 Coding Scheme

The basic approach of the coding scheme for a vector process is to transmit the last n elements of the initial state $S(0)$ to the controller. To achieve this aim, n coding schemes proposed in Section 5.2.2.3 are used in parallel for each individual element. Thus, for each $j = 0, 1, \cdots, n-1$, at the sequence of times $kn + j$ ($k \in \mathbb{Z}_+$), the sensor, relay, and controller implement the coding scheme $\mathcal{S}(S^j(0), \hat{S}^j(k), m)$, with $S^j(k) = e_j^T S(0)$ and $\hat{S}^j(k) = e_j^T \hat{S}(k)$. The controller calculates the control input as follows. It maintains an estimate $\hat{S}(k)$ of the initial state $S(0)$. At each time k, such that $j = k \mod n$, it performs the following actions:

- Update $\hat{S}(k)$ as

$$\hat{S}(k) = \hat{S}(k-1) - e_j^T \hat{S}(k-1) e_j + \hat{S}^j(k) e_j,$$

with the initial condition $\hat{S}_3(-1) = 0$.

- Calculate $\bar{S}(k)$ using the relation (4.13).
- Transmit the control input $U(k)$ and given by.

$$U(k) = \left[\hat{S}_1^j(k), \ldots, \hat{S}_m^j(k), K\bar{S}(k)\right]^T,$$

where T represents a transpose and the last component is calculated using (4.13).

5.2.5 Stability Analysis

We have the following proof for the sufficiency part of Theorem 5.1.1.

Proof for sufficiency of Theorem 2.1 *for vector plants*: We use the encoding scheme outlined above for each unstable state. Since, A_s is stable, we do not need to update the first $l-n$ components of the error vector $\epsilon(k)$, which remain constant at 0. The other n components of $\epsilon(k)$ are updated every n time steps. Since $\epsilon(0) = 0$ and all updates in the coding scheme are linear, it is straightforward to see that equation (4.11) is satisfied. Since A_s is Schur stable, A_s^k approaches 0 as $k \to \infty$. Using the result from scalar case, we can see that $\lambda_j^{2k}\mathbb{E}\left[(\epsilon^j(k))^2\right] \to 0$, where $l - n + 1 \leq j \leq l$ and $\epsilon^j(k) = e_j^T \epsilon(k)$, if the following condition is satisfied:

$$\ln|\lambda_j(A)| < \frac{1}{n} \max \sum_{i=1}^{m} \frac{1}{2} \ln\left(1 + \frac{g_i^2 P_i}{\sigma_i^2}\right),$$

where the maximization is over power allocations satisfying $\sum_{i=1}^m P_i = P$. Note that since the diagonal elements in (4.12) approach 0 as $k \to \infty$, using Cauchy-Schwarz inequality, it follows that the non-diagonal elements in (4.12) also approach 0. The theorem follows by noting that the above condition needs to be satisfied for $\forall 1 \leq i \leq m$. Note that the left hand side of (5.3) is a sum of the logarithms of all the unstable eigenvalues of the plant and right hand side depends on the channel parameters and transmission power. The result suggests that there is a minimum transmission power needed to ensure mean square stability. Consequently, we can rewrite the in an alternative form.

Corollary 5.2.6 *Consider the problem formulation presented in Section 5.1 with the coding scheme presented in Section 5.2.2.3 for each unstable state. The minimum power at the sensor that is needed for mean square stabilizing the process (4.1) of dimension l must be sufficient to guarantee that*

$$\max \sum_{i=1}^{m} \frac{1}{2} \ln\left(1 + \frac{g_i^2 P_i}{\sigma_i^2}\right) > \sum_{i=1}^{l} \max\left\{0, \ln|\lambda_i(A)|\right\}, \tag{5.32}$$

where the maximization is over power allocations satisfying $\sum_{i=1}^{m} P_i = P$.

5.2.6 Water-filling Solution

The optimization problem in (5.22) or (5.28) can be solved using Lagrange multipliers and has a well known interpretation. The solution is given by [12, Chapter 10]

$$P_i = \left(\lambda - \frac{\sigma_i^2}{g_i^2}\right)^+ = \max\left\{\lambda - \frac{\sigma_i^2}{g_i^2}, 0\right\},$$

where the Lagrange multiplier λ is chosen to satisfy

$$\sum_{i=1}^{m} \left(\lambda - \frac{\sigma_i^2}{g_i^2}\right)^+ = P.$$

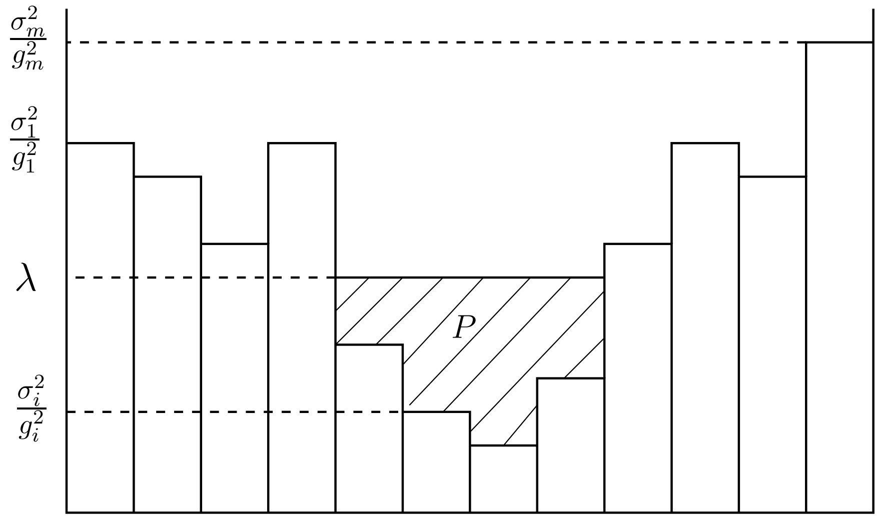

Figure 5.3. Water-filling for the product channel. Area of the shaded region is equal to P.

The optimal solution has the water-filling interpretation as shown in Fig. 5.3. The vertical levels indicate the noise levels in the various channels. As the power P is increased, power is first allotted to the channel with lowest noise, then the next lowest and so on. This power distribution is identical to the way water fills itself in a container, hence the name "water-filling".

Without loss of generality, assume

$$\sigma_1 \leq \sigma_2 \leq \ldots \leq \sigma_m^2.$$

Then, the optimal power allocation can be shown to be

$$P_i^* = \begin{cases} \dfrac{P + \sum_{j=1}^{t} \dfrac{\sigma_j^2}{g_j^2}}{t} - \sigma_i^2 & i \leq t, \\ 0 & i > t \end{cases}$$

where t is the largest value for which power is allocated to that particular channel. The capacity can be evaluated as

$$C = \frac{1}{2} \ln \frac{\left(P + \sum_{j=1}^{t} \frac{\sigma_j^2}{g_j^2}\right)^t}{\prod_{j=1}^{t} \frac{\sigma_j^2}{g_j^2}} \text{ nats per channel use.}$$

Using this capacity result and Corollary 5.2.6, we can find a lower bound on the power needed to stabilize the plant over a product channel.

$$P \geq \sqrt[t]{\prod_{j=1}^{t} \frac{\sigma_j^2}{g_j^2} \prod_{\lambda_i : \lambda_i > 1} |\lambda_i|^2} - \sum_{j=1}^{t} \frac{\sigma_j^2}{g_j^2}$$

5.2.7 Constraint C_1

The constraints C_2 and C_3 are satisfied by construction of the coding scheme. We can also show that the constraint C_1 is satisfied by the proposed design.

Proposition 5.2.7 *If $\hat{S}(k)$ is a linear MMSE estimate of $S(0)$, and the process (5.1) is mean squared stabilized, then the controller proposed in Theorem 4.3.1 satisfies $\sum_{k=0}^{\infty} \mathbb{E}[U^T(k)U(k)] < \infty$.*

Proof:

$$\begin{aligned}\sum_{k=0}^{\infty} \mathbb{E}[U^T(k)U(k)] &= \sum_{k=0}^{\infty} \mathbb{E}[\hat{S}^T(k)K^T K \hat{S}(k)] \\ &= \sum_{k=0}^{\infty} \mathbb{E}[\mathrm{tr}(\hat{S}(k)\hat{S}^T(k)K^T K)].\end{aligned}$$

The above term is finite if $\sum_{k=0}^{\infty} \mathbb{E}[\mathrm{tr}(\hat{S}(k)\hat{S}^T(k))]$ is finite. We have

$$\sum_{k=0}^{\infty} \mathbb{E}[\mathrm{tr}(\hat{S}(k)\hat{S}^T(k))] = \sum_{k=0}^{\infty} \mathbb{E}[\mathrm{tr}(S(k)S^T(k))] + \sum_{k=0}^{\infty} \mathbb{E}[\mathrm{tr}(\delta(k)\delta^T(k))]$$

If the process (5.1) is mean squared stabilized, the first summation is finite. The second summation can be written as

$$\begin{aligned}\sum_{k=0}^{\infty} \mathbb{E}[\mathrm{tr}(\delta(k)\delta^T(k))] &= \sum_{k=0}^{\infty} \mathbb{E}[\mathrm{tr}(A^k \epsilon_3(k)\epsilon_3^T(k)(A^T)^k)] \\ &= \sum_{k=0}^{\infty}\sum_{j=1}^{m} \lambda_j^{2k}\mathbb{E}[(\epsilon_3^j(k))^2].\end{aligned}$$

If the condition in (5.3) is satisfied, the above term is finite.

5.3 Necessity Results

To prove the necessary part of Theorem 5.1.1, we find a lower bound on the second moment of the state of the plant $S(k)$ and show that the condition in (5.3) is a necessary condition for this lower bound to converge to zero. A similar approach to proving necessity has been used for different settings [15, 30, 32]. Let Y_k represent a collection of all observations over the parallel channels at time k: $Y_k = \{Y_1(k), \ldots, Y_m(k)\}$ and let Y^k represent a collection of $Y_j, 0 \le j \le k$: $Y^k = \{Y_0, \ldots, Y_k\}$. Let $N(k)$ be the conditional entropy power of $S(k)$ conditioned on the event $\{Y^{k-1} = y^{k-1}\}$ averaged over all y^{k-1}.

$$N(k) = \frac{1}{2\pi e}\mathbb{E}_{Y^{k-1}}\left[e^{\frac{2}{n}h(S_u(k)|Y^{k-1}=y^{k-1})}\right], \tag{5.33}$$

where S_u are unstable states of the plant under the decomposition (5.31).

The following result establishes a relation between (5.2) and $N(k)$.

Lemma 5.3.1 *A necessary condition for* (5.2) *to hold is that* $\lim_{k\to\infty} N(k) = 0$.

Proof: It is known that conditional entropy power provides a lower bound on the mean square value of $S_u(k)$ [15]. Thus,

$$\mathbb{E}_{Y^{k-1}}\left[||S_u(k)||^2\right] \leq e^{(1-1/n)}\frac{1}{2\pi e}e^{\frac{2}{n}h(S_u(k)|Y^{k-1}=y^{k-1})}.$$

Taking expectation on both the sides, we obtain

$$\mathbb{E}\left[||S_u(k)||^2\right] \leq e^{(1-1/n)}N(k),$$

from which the result follows.

We now proceed with the proof by deriving a necessary condition for $N(k)$ to be bounded. Using Lemma 5.3.1, a necessary condition for mean square stability is provided by a necessary condition for $N(k)$ to be bounded. For this, we find a recursive equation for the evolution of $N(k)$.

$$\begin{aligned}
N(k+1) &= \frac{1}{2\pi e}\mathbb{E}_{Y^k}\left[e^{\frac{2}{n}h(S_u(k+1)|Y^k=y^k)}\right] \\
&= \frac{1}{2\pi e}\mathbb{E}_{Y^k}\left[e^{\frac{2}{n}h(A_uS_u(k)+B_uU(k)|Y^k=y^k)}\right] \\
&\overset{(a)}{=} \frac{1}{2\pi e}\mathbb{E}_{Y^k}\left[e^{\frac{2}{n}h(A_uS_u(k)|Y^k=y^k)}\right] \qquad (5.34)\\
&= \frac{|\det A_u|^{2/n}}{2\pi e}\mathbb{E}_{Y^k}\left[e^{\frac{2}{n}h(S_u(k)|Y^k=y^k)}\right] \\
&= \frac{|\det A_u|^{2/n}}{2\pi e}\mathbb{E}_{Y^{k-1}}\left[\mathbb{E}_{Y^k|Y^{k-1}}\left[e^{\frac{2}{n}h(S_u(k)|Y^k=y^k)}\right]\right] \\
&\overset{(b)}{\geq} \frac{|\det A_u|^{2/n}}{2\pi e}\mathbb{E}_{Y^{k-1}}\left[e^{\frac{2}{n}\mathbb{E}_{Y^k|Y^{k-1}}[h(S_u(k)|Y^k=y^k)]}\right] \\
&= \frac{|\det A_u|^{2/n}}{2\pi e}\mathbb{E}_{Y^{k-1}}\left[e^{\frac{2}{n}\mathbb{E}_{Y^k}[h(S_u(k)|Y^k=y^k)]}\right], \qquad (5.35)
\end{aligned}$$

where (a) follows from the fact that the input $U(k)$ is a function of Y^k and (b) follows using Jensen's inequality [12, Chapter 2]. The expectation of the entropy term in (5.35) can be written as

$$\begin{aligned}
&\mathbb{E}_{Y^k}[h(S_u(k)|Y^k = y^k)] \\
&= \mathbb{E}_{Y^k}[h(S_u(k)|Y^{k-1} = y^{k-1}) - I(S_u(k); Y_k|Y^{k-1} = y^{k-1})] \\
&\geq \mathbb{E}_{Y^{k-1}}[h(S_u(k)|Y^{k-1} = y^{k-1})] - C, \qquad (5.36)
\end{aligned}$$

where C is the capacity of the Gaussian product channel. Using (5.35), (5.36) and (5.33) we obtain

$$N(k+1) \geq |\det A_u|^{2/n} e^{(-2/m)C} N(k)$$

Thus, using the expression for the capacity of a Gaussian product channel (right hand side of (5.3)), we can arrive at the necessary part of the proof in Theorem 5.1.1.

5.4 Summary

We present encoder and decoder designs to achieve mean square stability, which are extensions of the Schalkwijk-Kailath coding scheme for point-to-point channels. When the sufficient conditions for stability using our coding scheme are satisfied with equality, data about the initial condition is being transmitted at a rate equal to the capacity of a Gaussian product channel, which implies that our scheme is optimal. We use tools from information theory to prove the necessary part of the result.

An immediate extension of this work would be to consider the effect of process or (and) sensor noise in (5.1). Such extensions, however, would rely on extensions of the SK scheme to situations with noisy feedback [22] and are likely to be more involved. Another direction of future work can involve other noise distributions to model wireless channels more accurately.

CHAPTER 6

GAUSSIAN NETWORKS

In this chapter, we briefly discuss the problem of achievability and stabilizability over some other classes of Gaussian channels, which provides some insights on transmission schemes for Gaussian networks. As discussed earlier, feedback schemes for general network scenarios have not been fully explored. However, using the schemes developed in earlier chapters, we can design feedback schemes for some specific network scenarios. Specifically, Chapters 3, 4 help understand the roles of intermediate relay nodes that have memory and processing abilities. Chapter 5 deals with the problem of communicating if we have multiple paths from the source (or plant) to the destination (or controller). These paths exhibit different levels of rate and reliability and thus need to be exploited by the system designer.

Throughout the chapter we consider the following scenario.

- Forward and feedback communications take place over orthogonal channels so that they do not interfere with each other.
- The forward links are modeled by zero-mean AWGN channels. We consider the noises on different links to be mutually independent and white.

We discuss some specific network problems that can be solved using results from previous chapters in Section 6.1.

6.1 Gaussian Networks

We consider three network examples corresponding to two parallel paths with relays to illustrate how we can use the results of previous chapters to solve some network problems. For all the examples in this section, there are two independent paths between the source and destination, with a relay on each path helping the source in transmitting to the destination. The transmissions from either relay is not seen at the other relay. The rest of the links are as shown for the three different scenarios. We adopt the following notation:

- The source and destination are denoted by the nodes 1 and 4 respectively, whereas the two relays are denoted by nodes $2, 3$
- The inputs to the channel to the two relays are denoted by X_1 and X_2 respectively. The inputs from the relays are denoted by X_3 and X_4 respectively.
- The outputs at the two relays are denoted by Y_1 and Y_2 respectively. The outputs at the destination is denoted by Y_3 and Y_3 respectively.
- All the forward links are corrupted by AWGN. The noises at relay 1, relay 2 are denoted by Z_1 and Z_2 respectively. The inputs X_3 and X_4 are corrupted by noises Z_3 and Z_4 respectively. Noise Z_i has variance σ_i^2.

We consider the following scenario.

- The signal from node i to node j is scaled by a distance-dependent attenuation factor $g_{i,j}$ that depends on the distance $d_{i,j}$ between nodes i and j as

$$g_{i,j} = \begin{cases} d_{i,j}^{-\eta/2}, & i \neq j, \\ 0, & i = j, \end{cases} \tag{6.1}$$

where η is the path-loss exponent.

- The signals transmitted on the forward links by the i-th node, $i \in \{1, 2, 3\}$ must satisfy average power constraints P_i, and P_1 is distributed over the two paths as $P_{1,1}$ and $P_{1,2}$.

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{E}[X_1^2(k) + X_2^2(k)] \leq P_{1,1} + P_{1,2} \leq P_1,$$
$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{E}[X_3^2(k)] \leq P_2,$$
$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{E}[X_4^2(k)] \leq P_3.$$

6.1.1 Example 1

Consider a Gaussian network as shown in Figure 6.1. We can use the results from Sections 3.3 and 5.2.2.2 to propose a coding scheme for the given network. Divide the interval [0,1] into M_1 equal-length message intervals, then divide each of the M_1 intervals into M_2 message intervals. M_i is related to the rate R_i as $M_i = \exp(nR_i)$. Send M_1 and M_2 over the parallel channels as in Section 5.2.2.2. The transmission over the relay channel is as in Section 3.3.

The following theorem gives the achievable rates for the given network:

Theorem 6.1.1 *For the network in Figure 6.1, the coding scheme presented above*

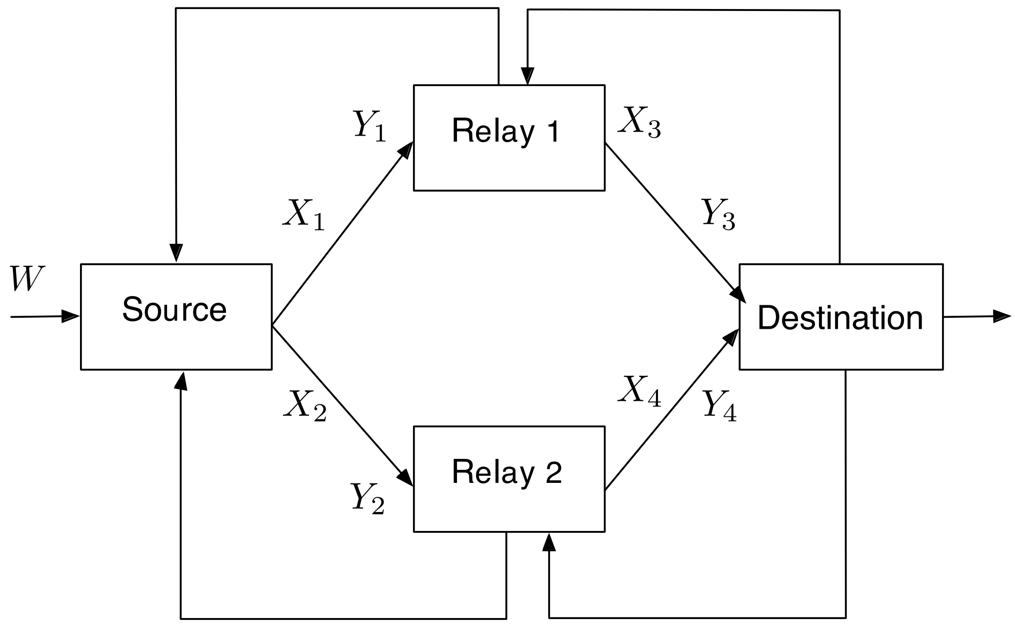

Figure 6.1. Problem setup for an unstable plant being controlled over a network of nodes: Example 1.

achieves a rate

$$R < \max_{P_{1,1}+P_{1,2}=P_1} \Bigg(\overbrace{\min\left(\frac{1}{2}\ln\left(1+\frac{g_{1,2}^2 P_{1,1}}{\sigma_1^2}\right), \frac{1}{2}\ln\left(1+\frac{g_{2,4}^2 P_2}{\sigma_3^2}\right)\right)}^{C_1}$$
$$+ \overbrace{\min\left(\frac{1}{2}\ln\left(1+\frac{g_{1,3}^2 P_{1,2}}{\sigma_2^2}\right), \frac{1}{2}\ln\left(1+\frac{g_{3,4}^2 P_3}{\sigma_4^2}\right)\right)}^{C_2} \Bigg).$$

Proof: Given $P_{1,1}$ and $P_{1,2}$, using Theorem 3.3.2, we can see that the achievable rates for the two paths are

$$R_1 \leq C_1, R_2 \leq C_2.$$

The overall achievable rate is $C_1 + C_2$, that can further be optimized for the distribution of power on both the paths [12, Chapter 10]. Thus, we arrive at the result.

6.1.2 Example 2

We can also consider a network with destination-relay and destination-source feedback as shown in Figure 6.2.

We can use the results from Sections 3.4 and 5.2.2.2 to propose a coding scheme for the given network. Divide the interval [0,1] into M_1 equal-length message intervals, then divide each of the M_1 intervals into M_2 message intervals. M_i is related to the rate R_i as $M_i = \exp(nR_i)$. Send M_1 and M_2 over the parallel channels as in Section 5.2.2.2. The transmission over the relay channel is as in Section 3.4.

The following theorem gives the achievable rates for the given network:

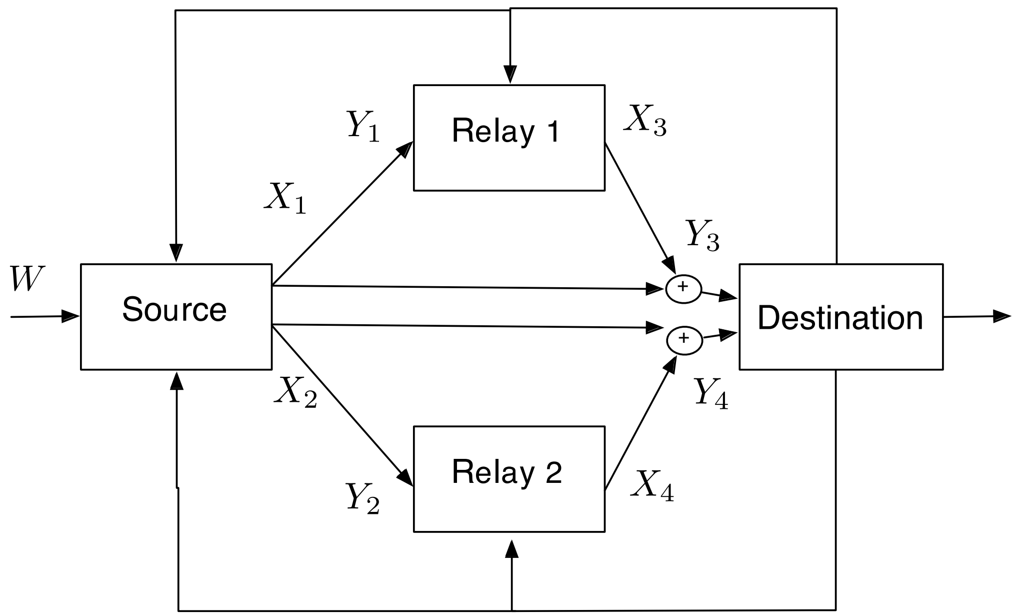

Figure 6.2. Problem setup for an unstable plant being controlled over a network of nodes: Example 2.

Theorem 6.1.2 *For the network in Figure 6.2, the coding scheme presented above achieves a rate*

$$R < \max_{P_{1,1}+P_{1,2}=P_1} \left(\liminf_{n\to\infty} \frac{1}{2n} \sum_{k=1}^{n} \ln \frac{1}{q_1(k-1)} + \liminf_{n\to\infty} \frac{1}{2n} \sum_{k=1}^{n} \ln \frac{1}{q_2(k-1)} \right),$$

where

$$q_1(k-1) = \frac{g_{2,4}^2 P_2 \left(1-\rho_1^2(k-1)\right) + \sigma_3^2}{g_{1,4}^2 P_{1,1} + g_{2,4}^2 P_2 + 2 g_{1,4} g_{2,4} \sqrt{P_{1,1} P_2}\, \rho_1(k-1) + \sigma_3^2},$$

$$q_2(k-1) = \frac{g_{3,4}^2 P_3 \left(1-\rho_2^2(k-1)\right) + \sigma_4^2}{g_{1,4}^2 P_{1,2} + g_{2,4}^2 P_3 + 2 g_{1,4} g_{3,4} \sqrt{P_{1,2} P_3}\, \rho_2(k-1) + \sigma_4^2},$$

where ρ_1 and ρ_2 can be calculated as in Section 3.4.2.

Proof: The proof is similar to the proof of 6.1.1 and has been omitted.

6.1.3 Example 3

Consider a network with only destination-source feedback for the two parallel channels. The following result from [8, Theorem 5] helps us to characterize the result for such a feedback configuration of the relay. Note that this result is only for the first parallel path in Figure 6.3, but can similarly be written for the second parallel path.

Theorem 6.1.3 *For the Gaussian relay channel with only destination-source feedback, the following rate is achievable*

$$R < \frac{1}{2} \ln \left(1 + \frac{g_{1,4}^2 P_{1,1} \left(1 + \sqrt{\dfrac{g_{2,4}^2 P_2}{g_{1,2}^2 P_{1,1} + \sigma_1^2}}\, \rho_1^* \right)^2}{\dfrac{g_{2,4}^2 P_2 \sigma_1^2}{g_{1,2}^2 P_{1,1} + \sigma_1^2} + \sigma_3^2} \right),$$

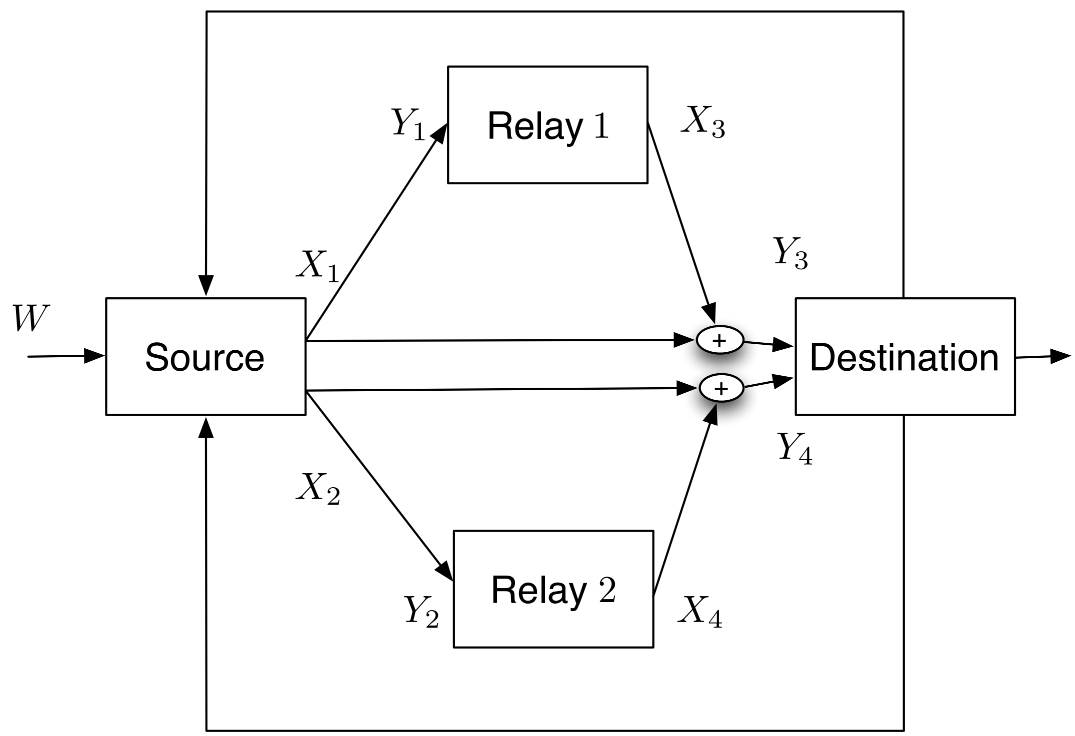

Figure 6.3. Problem setup for an unstable plant being controlled over a network of nodes: Example 2.

where the correlation coefficient ρ^* *is given by the unique solution in* $[0,1]$ *of the following equation*

$$\rho^2\left(g_{1,4}^2P_{1,1}+2g_{2,4}g_{1,2}^2\sqrt{\frac{P_{1,1}^2P_2}{g_{1,2}^2P_{1,1}+\sigma_1^2}}\rho+\frac{g_{2,4}^2P_2g_{1,2}^2P_{1,1}}{g_{1,2}^2P_{1,1}+\sigma_1^2}\rho^2+\frac{\sigma_1^2g_{2,4}^2P_2}{g_{1,2}^2P_{1,1}+\sigma_1^2}+\sigma_3^2\right)$$
$$=g_{2,4}^2P_2\frac{\sigma_1^2}{g_{1,2}^2P_{1,1}+\sigma_1^2}+\sigma_3^2 \quad (6.2)$$

For deriving the above theorem, a SK like coding scheme is proposed in [8] in which the source transmits maximally informative updates (similar to our coding scheme in Chapter 3) and the relay applies a simple amplify-and-forward strategy. The destination node calculates the linear minimum mean squared estimate estimate (mmse) of the message given all its observations. Note that the relay uses an amplify-and-forward strategy because it does not get any feedback from the destination node. The achievability theorem for the example in Figure 6.3 can be specified in a similar way as in Theorems 6.1.1 and 6.1.2, with the capacity on each parallel path being the achievability rate specified by Theorem 6.1.3.

6.2 Summary

We present a few network examples and show how our results from the earlier chapters be used to obtain achievability and stability results. Note that using a similar approach for the different parallel paths, our approach can be extended to the corresponding cases with arbitrary number of parallel paths. If the relay sees feedback from the destination, then it sends corresponding power-normalized estimation errors on the forward link, otherwise it just uses a simple amplify-and-forward strategy. We can also consider other simple network examples. Studying general network examples, is however out of the scope of this dissertation. Such a general network scenario can have a complicated architecture and studying achievability and thus stabilizability results is a very difficult problem.

CHAPTER 7

CONCLUSIONS AND FUTURE WORK

In this dissertation, we investigate coding schemes, that are extensions of the Schalkwijk-Kailath scheme [38] for certain classes of Gaussian channels with feedback. We examine three different channels - a cascade of point-to-point channels, a relay channel and a product channel. Transmission schemes that make use of feedback can be used to stabilize unstable plants over communication channels, where a sensor transmits the plant state information to a remotely placed controller. Using the coding schemes developed, we find corresponding stability conditions for stabilization of dynamical processes over the aforementioned channels. We then use these simplified channel results to characterize achievability and stability results for a few specific network examples.

We present a feedback coding schemes for two different configurations of a three node network - a cascade of two Gaussian point-to-point channels and a Gaussian relay channel in Chapter 3. Our schemes are based on allowing the nodes to run a *distributed stochastic approximation algorithm*. We present results for full-duplex and half-duplex relaying. Though our scheme achieves the capacity for the cascade channel, our scheme is not capacity-achieving for the relay channel case. A corresponding open problem is to develop an iterative coding scheme that achieves the capacity of a generalized relay channel with feedback. A complex Block-Markov superposition scheme that achieves capacity was given in [11].

We use the feedback coding schemes developed for a relay channel in Chapter 3 and apply it to stabilize a linear time invariant (LTI) plant over a Gaussian relay channel in Chapter 4. It is known that a relay helps in enhancing the achievability region even if the total transmission power remains the same. We present the corresponding result for sufficient conditions for the stabilizability of the plant through such schemes. The analysis suggests that it is useful to provide a relay node assisting the plant, even if the total transmission power remains the same. Derivation of tight necessary conditions is still an open problem.

In Chapter 5, we present necessary and sufficient conditions for stabilizing a discrete-time LTI plant in the mean squared sense when a sensor transmits the plant state information to a remotely placed controller across a Gaussian product channel. We present a non-linear coding scheme based on ideas from distributed source-channel coding and present the resulting stabilizability conditions. When these conditions are satisfied with equality, the proposed coding scheme transmits data across the product channel at a rate equal to the capacity of the channel, indicating that the sufficiency results are also necessary. We then alternatively prove that the derived condition is necessary using tools from information theory. However, our scheme needs dedicated feedback links for all the channels. A coding scheme that can satisfy the necessary condition and uses only one feedback link from the controller to the sensor is still an open problem.

Having studied two basic, yet important components of network problems - relay and parallel paths, in Chapter 6, we use the results from previous chapters to obtain achievability results for some specific network problems. Characterizing achievability or stability results for general network scenarios, however, still remains an open problem. Our scheme, though scalable, is very different to analyze as linear minimum mean square estimates (mmse) require computations that increase in complexity as the number of nodes in the network increases. Non-optimal schemes that do not necessarily calculate mmse estimates can be considered. Such coding schemes would be similar to consensus algorithms [5, 14, 18] and can be useful in that regard.

Problems pertaining to computation of reliability functions and achievability results for the case of noisy feedback and peak energy constraints is open in general. Wyner [47] showed that the error probability of the SK scheme degrades to exponential form under a peak energy constraint. Xiang *et al.* [48] show that if the noise power in the feedback link is sufficiently small, the best error exponent for conmmunicating an M-ary message can be strictly larger than the one without feedback. Their proof involves two feedback coding schemes. It remains to be seen if such coding schemes can be extended to relay channels, MAC channels, and so forth.

BIBLIOGRAPHY

1. P. Antsaklis and J. Baillieul. Special Issue on Networked Control Systems. *IEEE Transactions on Automatic Control*, 49, 2004.

2. E. Ardestanizadeh, P. Minero, and M. Franceschetti. LQG Control Approach to Gaussian Broadcast Channels with Feedback. *submitted to IEEE Transactions on Information Theory*, 2011.

3. E. Ardestanizadeh, M. Wigger, Y.-H. Kim, and T. Javidi. Linear-Feedback Sum-Capacity for Gaussian Multiple Access Channels. *IEEE Transactions on Information Theory*, 58(1):224–236, Jan. 2012.

4. R. Bansal and T. Basar. Simultaneous Design of Measurement and Control Strategies for Stochastic Systems with Feedback. *Automatica*, 25:679–694, Sep. 1989.

5. F. Benezit, A. Dimakis, P. Thiran, and M. Vetterli. Gossip along the way: Order-optimal consensus through randomized path averaging. In *Proceedings of the Allerton Conference on Communication, Control, and Computing*, 2007.

6. V. S. Borkar. *Stochastic Approximation: A Dynamical Systems Viewpoint.* Cambridge University Press, 2008.

7. J. H. Braslavsky, R. H. Middleton, and J. S. Freudenberg. Feedback Stabi-

lization over Signal-to-Noise Ratio Constrained Channels. *IEEE Transactions on Automatic Control*, 52:1391–1403, 2007.

8. S. I. Bross and M. A. Wigger. On the Discrete Memoryless Relay Channel in the Presence of Receiver-Transmitter Feedback. *IEEE Trans. on Information Theory*, 55(1):275–291, Jan. 2009.

9. L. Bushnell. Special Issue on Networks and Control. *IEEE Control Systems Magazine*, 21:903–908, Feb. 2001.

10. Z. Chance and D. J. Love. Concatenated Coding for the AWGN Channel with Noisy Feedback. *IEEE Transactions on Information Theory*, 57(10): 6633–6649, Oct. 2011.

11. T. M. Cover and A. El Gamal. Capacity Theorems for the Relay Channel. *IEEE Trans. on Information Theory*, 25(5):572–584, Sep. 1979.

12. T. M. Cover and J. A. Thomas. *Elements of Information Theory.* John Wiley & Sons, Inc., New York, 1991.

13. N. Elia. When Bode meets Shannon: Control-Oriented Feedback Communication Schemes. *IEEE Transaction on Automatic Control*, 49(9):1477–1488, Sep. 2004.

14. L. Fang and P. Antsaklis. On communication requirements for multi-agents consensus seeking. In *Proceedings of Workshop NESC05: University of Notre Dame, Lecture Notes in Control and Information Sciences (LNCIS)*, pages 331:53–68. Springer, 2006.

15. J. S. Freudenberg, R. H. Middleton, and V. Solo. The Minimal Signal-to-Noise Ratio Required to Stabilize over a Noisy Channel. In *American Control Conference*, Minneapolis, Jun. 2006.

16. V. Gupta, A. F. Dana, J. Hespanha, R. M. Murray, and B. Hassibi. Data Transmission over Networks for Estimation and Control. *IEEE Trans. on Automatic Control*, Aug. 2009.

17. M. Horstein. Sequential Transmission using Noiseless Feedback. *IEEE Trans. on Information Theory*, pages 136–143, Jul. 1963.

18. S. Kar and J. Moura. Distributed average consensus in sensor networks with random link failures. In *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, Apr. 2007.

19. Y.-H. Kim, A. Lapidoth, and T. Weissman. The Gaussian Channel with Noisy Feedback. In *Proc. IEEE Int. Symp. Information Theory (ISIT)*, pages 1416–1420, Nice, France, Jun 2007. doi: 10.1109/ISIT.2007.4557421.

20. G. Kramer. Feedback Strategies for White Gaussian Interference Networks. *IEEE Trans. on Information Theory*, IT-48, No. 6:1423–1438, Jun. 2002.

21. U. Kumar. Feedback Communication via Stochastic Approximation. Master's thesis, University of Notre Dame, Notre Dame, IN, Dec. 2009.

22. U. Kumar, J. N. Laneman, and V. Gupta. Cooperative Communications with Feedback via Stochastic Approximation. In *Proc. IEEE Information Theory Workshop*, Taormina, Sicily, Oct. 2009.

23. U. Kumar, J. N. Laneman, and V. Gupta. Cooperative Communications with

Feedback via Stochastic Approximation. In *Proc. IEEE Information Theory Workshop*, Taormina, Sicily, Oct. 2009.

24. U. Kumar, V. Gupta, and J. N. Laneman. Sufficient Conditions for Stabilizability over Gaussian Relay and Cascade Channels. In *Proc. IEEE Conference on Decision and Control*, Atlanta, Georgia, Dec. 2010.

25. H. J. Kushner and G. G. Yin. *Stochastic Approximation and Recursive Algorithms and Applications.* Springer-Verlag, Inc., New York, 2003.

26. R. Luck and A. Ray. An Observer-based Compensator for Distributed Delays. *Automatica*, 26:903–908, Aug. 1990.

27. N. C. Martins and M. A. Dahleh. Feedback Control in the Presence of Noisy Channels: Bode-Like Fundamental Limitations of Performance. *IEEE Transactions on Automatic Control*, 53:1604–1615, Aug. 2008.

28. N. C. Martins and T. Weissman. Coding for Additive White Noise Channels With Feedback Corrupted by Quantization or Bounded Noise. *IEEE Trans. on Information Theory*, IT-54(9):4274–4282, Sep. 2008.

29. N. Merhav and T. Weissman. Coding for the Feedback Gel'Fand-Pinsker Channel and the Feedforward Wyner-Ziv Source. In *Proc. IEEE Int. Symp. Information Theory (ISIT)*, pages 1506–1510, Sep. 2005.

30. P. Minero, M. Franceschetti, S. Dey, and G. N. Nair. Data Rate Theorem for Stabilization Over Time-Varying Feedback Channels. *IEEE Transactions on Automatic Control*, 54:243–255, Feb. 2009.

31. G. N. Nair and R. J. Evans. Mean Square Stabilizability of Stochastic Linear

Systems with Data Rate Constraints. In *Proc. IEEE Conference on Decision and Control*, 2002.

32. G. N. Nair, F. Fagnani, S. Zampieri, and R. J. Evans. Feedback Control under Data Rate Constraints: An Overview. *Proceedings of the IEEE*, 95:3369–3395, Jan. 2007.

33. L. Ozarow and S. Leung-Yan-Cheong. An Achievable Region and Outer Bound for the Gaussian Broadcast Channel with Feedback (Corresp.). *IEEE Trans. on Information Theory*, 30(4):667–671, Jul. 1984.

34. L. H. Ozarow. The Capacity of the White Gaussian Multiple Access Channel with Feedback. *IEEE Trans. on Information Theory*, 30(4):623–629, Jul. 1984.

35. H. Robbins and S. Munro. A Stochastic Approximation Method. *Ann. Math. Stat.*, 22:400–407, Sep. 1951.

36. A. Sahai and S. Mitter. The Necessity and Sufficiency of Anytime Capacity. *IEEE Trans. on Information Theory*, 52:3369–3395, 2006.

37. J. P. M. Schalkwijk. A Coding Scheme for Additive Noise Channels with Feedback - II: Band-Limited Signals. *IEEE Trans. Inform. Theory*, IT-12: 183– 189, Apr. 1966.

38. J. P. M. Schalkwijk and T. Kailath. A Coding Scheme for Additive Noise Channels with Feedback - I: No bandwidth constraint. *IEEE Trans. Inform. Theory*, IT-12:172– 182, Apr. 1966.

39. S. Shamai, S. Verdu, and R. Zamir. Systematic Lossy Source/Channel Coding. *IEEE Trans. Inform. Theory*, 44:564– 579, Mar. 1998.

40. C. Shannon. A Mathematical Theory of Communication. *Bell Systems Technical Journal*, 27:379423, 623656, 1948.

41. C. Shannon. The Zero-Error Capacity of a Noisy Channel. *IRE Trans. on Information Theory*, IT-2:8–19, Sep. 1956.

42. O. Shayevitz and M. Feder. Communication with Feedback via Posterior Matching. In *Proc. IEEE Int. Symp. Information Theory (ISIT)*, pages 391–395, Nice, France, Jun. 24–29, 2007.

43. O. Shayevitz and M. Feder. The Posterior Matching Feedback Scheme: Capacity Achieving and Error Analysis. In *Proc. IEEE Int. Symp. Information Theory (ISIT)*, pages 900–904, Toronto, ON, Canada, Jul. 6–11, 2008.

44. Z. Shu and R. H. Middleton. Stabilization over Power-Constrained Parallel Gaussian Channels. *IEEE Transactions on Automatic Control*, 56(7):1718–1724, Jul. 2011.

45. S. Tatikonda and S. K. Mitter. Control over Noisy Channels. *IEEE Transactions on Automatic Control*, 49:1196 – 1201, Jul. 2004.

46. H. Witsenhausen. A Counterexample in Stochastic Optimum Control. *SIAM Journal on Control*, 6:131, 1968.

47. A. D. Wyner. On the SchalkwijkKailath Coding Scheme with a Peak Energy Constraint. *IEEE Transactions on Information Theory*, 14(1):129–134, Jan. 1968.

48. Y. Xiang and Y.-H. Kim. Gaussian Channel with Noisy Feedback and Peak Energy Constraint. *submitted to IEEE Transactions on Information Theory*, Aug. 2011.

49. H. Yamamoto and K. Itoh. Asymptotic Performance of a Modified Schalkwijk-Barron Scheme for Channels with Noiseless Feedback. *IEEE Trans. on Information Theory*, 25:729–733, Nov. 1979.

50. S. Yüksel and S. Tatikonda. Distributed Sensing and Control over Communication Channels and a Counterexample. In *Proc. IEEE Conference on Decision and Control*, UIUC, Illinois, Sep. 2007.

51. S. A. A. Zaidi, T. Oechtering, and M. Skoglund. Sufficient Conditions for Closed-Loop Control over Multiple-Access and Broadcast Channels. In *Proc. IEEE Conference on Decision and Control*, Atlanta, Georgia, Dec. 2010.