KTH Electrical Engineering

Sensing and Control of Dynamical Systems Over Networks

ALI ABBAS ZAIDI

Doctoral Thesis in Telecommunications
Stockholm, Sweden 2013

KTH, School of Electrical Engineering
Communication Theory Laboratory
SE-100 44 Stockholm
SWEDEN

ISSN 1653-5146
ISBN 978-91-7501-904-8

Akademisk avhandling som med tillstånd av Kungl Tekniska högskolan framlägges till offentlig granskning för avläggande av teknologie doktorsexamen fredagen den 22 November 2013 klockan 13.15 i hörsal Q1, Kungl Tekniska högskolan, Osquldas väg 6, Stockholm.

Tryck: Universitetsservice US AB

To my mother

Deep thinking will present the clearest picture of every problem. — Imam Ali

Abstract

Rapid advances in sensing, computing, and wireless technologies have led to significant interest in the understanding and development of wireless networked control systems. Networked control systems consist of spatially distributed agents such as plants (dynamical systems), sensors, and controllers, that interact to achieve desired objectives. The sensors monitor the plants and communicate measurements to remotely situated controllers. The controllers apply actions to stabilize and control the plants. Such systems have diverse applications in security, surveillance, industrial production, health monitoring, remote surgery, environment management, space missions, and intelligent transport systems. The objective of the thesis is to understand the fundamental limits and principles involved in the design of sensing and control strategies for dynamical systems controlled over communication networks.

The thesis has three parts. Part I and Part III consider the design of sensing and control strategies for mean-square stabilization of linear dynamical systems over fundamental communication channels such as point-to-point, relay, multiple-access, broadcast, and interference channels. The sensors and other nodes within the communication network are assumed to have average power transmit constraints. Moreover, the communication links between all agents (plants, sensors, controllers) are modeled as Gaussian channels. Necessary as well as sufficient conditions for mean-square stabilization over various network topologies are derived. The necessary conditions are arrived at using information theoretic arguments such as properties of mutual information, directed information, and differential entropy. The sufficient conditions are obtained using delay-free sensing and control policies. These conditions quantify the effect of communication network parameters such as transmit powers, channel noise, and channel interference on the stability of the plant(s). Different settings where linear policies are optimal, asymptotically optimal (in certain parameters of the system) and suboptimal have also been identified. Part II considers the design of real-time sensing and control strategies for minimization of a quadratic cost function of the state process of a system over Gaussian networks. Two fundamental Gaussian networks are considered: i) cascade network and ii) parallel network. For each network, non-linear sensing and control schemes are proposed and sub-optimality of linear strategies is discussed.

The results reveal fundamental limits on the performance of linear systems controlled over Gaussian networks. The methods used to derive these results reveal a close interplay between information theory and control theory.

Acknowledgments

First and foremost, I would like to express my deepest gratitude to Prof. Mikael Skoglund. He has been the most important person during the last seven years of my academic life. Mikael taught me excellent courses in the master program, acted as a supervisor on my master thesis, and then gave me an excellent opportunity to pursue PhD program. Mikael introduced me to very interesting research topics, guided me throughout, and generously provided great opportunities to excel. Mikael is an extremely nice person. Not only I substantially benefitted from his vast technical knowledge, but I also learnt how a gentleman should behave when he is in authority. I am proud to be his student.

I am grateful to Prof. Tobias Josef Oechtering for being the co-advisor on my PhD thesis. His sound knowledge, enthusiasm, and the sense of responsibility makes him a great supervisor. Tobias has a major role in the completion of my PhD studies. Although there are many great aspects of working with Tobias, one thing that I liked most is that he is always available. His welcoming behavior has always made me feel that I can go to him anytime and discuss any matter, irrespective of how long it takes. Tobias has invested a lot of time in supervising me. I hope to make him feel proud by successfully completing this thesis.

Special thanks to Prof. Serdar Yüksel for first hosting me at Queen's University and then taking a supervisory role throughout my PhD. Serdar has been a constant source of encouragement and tremendous support. Despite his busy schedule, he has always generously spent time with me. We had numerous meetings in person as well as on Skype. I have learnt a lot from his scientific knowledge, vision, and work ethics. His views have had a significant impact on transforming my approach towards scientific work. It has been a privilege working closely with Serdar. I will always remember his strong support.

I want to thank all my teachers, in particular, Dr. Magnus Jansson, Dr. Mats Bengtsson, Dr. Ben Slimane, Dr. Ather Gattami, Dr. Erik G. Larsson, Dr. Ragnar Thobaben, Dr. Carlo Fischione, Dr. Tamas Linder, and Dr. Peter Händel, for teaching excellent courses. Moreover, I am deeply grateful to my teachers from Pakistan, Dr. Jamil Ahmad, Prof. Javed Hussain, Brig. Mohsan Shah, Prof. Saeed Murtaza, Prof. Muhammad Asghar, Prof. Lal Badshah and my seniors Azam Mehboob, Talha Hassan, Amir Masood, Irtiza Ali, and Salman Anwar, who have been the great sources of inspiration.

The atmosphere in the Communication Theory lab. at KTH is very stimulating, which has been instrumental in my progressive learning. I would like to thank all the colleagues. Our group members are always happy to share knowledge. My understanding in various topics substantially improved due to the in-depth discussions with Hieu Do, Majid Khormuji, Sha Yao, Hamed Farhadi, and Li Minyue. I have also benefitted from various discussions with Nicolas Schrammar, Mattias Andersson, Tai Do, Johannes Kron, Niklas Wernersson, Iqbal Hussain, Amirpasha Shirazinia, Maksym Girnyk, Kittipong Kittichokechai, Ricardo Serrano, Dennis Sundman, Majid Gerami, Ibrahim Bilal, Ji Mingyue, Zhongwei Si, Ahmed Zaki, Chithrupa Ramesh, and John-Olof Nilsson. It has been very productive to collaborate in various projects with Majid Khormuji, Sha Yao, Ibrahim Bilal, Hamed Farhadi, Mattias Andersson, Niklas Wernersson, Lei Bao, Amirpasha Shirazinia and Majid Gerami. I am grateful to Hamed, Amirpasha, Maksym, Kittipong, Majid, and Tai, for proof-reading different parts of this thesis. I am also thankful to Cecilia Forssman, Annika Augustsson, Irene Kindblom, and Raine Tiivel, for the administrative support.

I am indebted to Prof. Anders Ranzter, Prof. Bo Bernhardsson, and Prof. Giacomo Como for their generosity in hosting me at the prestigious Automatic Control lab. in Lund University. The focus period on “Information and Control in Networks” was very useful. It provided me with an excellent opportunity to interact with leading researchers in the field and listen to their great ideas. I also want to thank Prof. Anders Robertsson for a very interesting tour of the robotics lab. It was also my pleasure to meet Naci Saldi, Vinith Misra, Alireza Farhadi, and Saam Coogan.

I am very thankful to Dr. Mikael Gidlund for offering me an internship at ABB Corporate Research. The project was very interesting; I could learn many new things in a short period. For this, the main credit goes to my skillful team members: Thomas Fuglsang, Anders Eskildsen, Majid Gerami, Mikaela Ahlén, and our energetic supervisor, Dr. Johan Sjöberg. It is also my great pleasure to know Dr. Jonas Neander, Dr. Elina Vartiainen, Dr. Shiva Tavallaey, Kan Yu, Peter Nordstrom, and Hamidreza Shariatmadari.

I wish to thank Prof. Nuno Martins for traveling from the US to act as the opponent on my thesis. I am grateful to Prof. Jan Østergaard, Prof. Subhrakanti Dey, and Prof. Dimos Dimarogonas for being the committee members. I am also thankful to Prof. Ming Xiao for the quality review of this thesis.

I had a wonderful time in beautiful Stockholm with very good friends. I want to thank all my friends, in particular I want to mention, Shabbir Maken, Altamash Khan, Hieu Do, Iqbal Hussain, Hamed Farhadi, Maryam Oryani, Maksym Girnyk, Karina, Amirpasha Shirazinia, Nafiseh Shariati, Faisal Altaf, Khurram Yousaf, Tai Do, Thuong Vo, Nicolas Schrammar, Aline Schrammar, Rizwan Raza, Majid Khormuji, Majid Gerami, Ibrahim Bilal, Ahmed Zaki, Kittipong Kittichokechai, Nasir Mehdi, Mahir Hussain, Qaisar Bangash, Li Minyue, Sha Yao, Ji Mingye, Azadeh Karimian, Farshad Naghibi, Mattias Andersson, Zuxing Li, Johannes Kron, Abdullah Nauroze, Rashid Ali, Zhao Wang, Ricardo Serrano, Nan Li, Hammad Kazmi, Samar Abbas, Haopeng Li, Jam Kausar, Chao Wang, Saikat Chatterjee, Jinghong

Yang, Abdullah Niazi, and Dennis Sundman. I will always remember our lunches and dinners, long walks and runs, various sport activities and trips, and amusing conversations. I want to deeply thank Riaz Shah and Jafar Moghal for their never ending hospitalities. I sincerely appreciate all the people involved in organizing cultural and religious activities at Imam Ali Islamic Center and Nordic Islamic Culture Center in Stockholm. I have always enjoyed being a member of KTH Hallen and Tennis Stadion; thanks to Eva and Mikael for the amazing training sessions.

I have great memories from my trips to Canada, Iran, The Netherlands, and China. I wish to deeply thank my uncle (Javed), aunt (Nafisah), cousins (Yavar, Sahar, Kiran), niece (Samrina) and nephews (Eshan and Ayan) for the unforgettable moments that we spent in Canada. I always miss you. I was also lucky to meet some very nice people in Kingston, Ayça Özçelikkale, Hakan Olgun, Huma Irfan, Prabeen Joshi, and Nervoz Şen. Without them, my life in Kingston could not be so interesting. I have really enjoyed my trips to Tehran and Mashhad; for which I am very thankful to my uncle (Athar), aunt (Zahra), and parents of Amirpasha (Jafar and Zahra), for their great hospitality. In Eindhoven, I had a wonderful time with Jamil Ahmad, who has been a great teacher. It was a pleasure to meet Houyu Si (Zhongwei's younger brother) in Beijing, who guided me to many interesting places.

I am extra-ordinary lucky to be born in a very caring and supportive family. I wish to deeply thank all my uncles, especially, Faiz, Riaz, and Imtiaz, and my aunts Farhana, Rehana, Farzana, Tabassum, Tajdar, and Nafisah, for all the love and support. I have spent wonderful moments of my life with my cousins, especially, Saif, Saad, Tooba, Sabeeka, Sobia, Sadaf, Nadia, Zehra, Fatima, Iqtedar, Shumaila, Fareeha, Kanwal, Fareena, Yavar, Uroos, Tashkeel, Sahar, Zainab, Jafar, and Najaf. My childhood teacher, Sabeeka, deserves a very special acknowledgement for her sincere efforts and love. I have always enjoyed spending time with my niece and nephews, in particular, Masooma and Fatima (tama) have a special place in my heart. Finally, I wish to express deepest gratitude to my dearest parents and my younger sister (Farheen) and brother (Hasan), for their unconditional love and for making a perfect family.

Ali Abbas Zaidi
Stockholm, November 2013

Contents

Notation

$\mathbb{R}^n$	$n \times 1$ dimensional real-valued vector space
$\|X\|$	Euclidean norm of a vector X
$\lvert a \rvert$	Absolute value of a
$X_{[0,k]}$	A sequence $\{X_0, X_1, \ldots, X_k\}$
$f'(x)$	Derivative of function $f(x)$
$p(x)$	Probability density function of a continuous random variable X
$p(x \mid y)$	Conditional probability density function of X given Y
$\mathbb{E}[X]$	Expected value of a random variable X
$h(X)$	Differential entropy of a random variable X
$h(X \mid Y)$	Conditional differential entropy of X given Y
$I(X;Y)$	Mutual Information between X and Y
$I\left(X_{[0,k]} \to Y_{[0,k]}\right)$	Directed Information between $X_{[0,k]}$ and$Y_{[0,k]}$
log	Natural logarithm
sup	Supremum
inf	Infimum
lim sup	Limit superior
lim inf	Limit inferior
A	System matrix
λ_i	i-th eigenvalue of system matrix A
$\det(A)$	Determinant of a matrix A
$\mathrm{Tr}[A]$	Trace of a matrix A
A^T	Transpose of matrix A
$\Lambda \succ 0$	Λ is positive definite matrix
$\mathrm{diag}(a_1, a_2, \ldots, a_n)$	Diagonal matrix with diagonal entries $\{a_1, a_2, \ldots, a_n\}$
$\mathcal{N}(m, \sigma^2)$	Gaussian distribution with mean m and variance σ^2
$Q(x)$	Complementary error function $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{\tau^2}{2}} d\tau$
$\mathrm{sgn}(\cdot)$	Sign function

List of Acronyms

AF	Amplify-and-forward
AWGN	Additive white Gaussian noise
DM	Decision Maker
i.i.d.	Independent and identically distributed
LQG	Linear Quadratic Gaussian
LTI	Linear time invariant
LTV	Linear time variant
MMSE	Minimum mean squared error
pdf	Probability density function
RHS	Right hand side
SK	Schalkwijk-Kailath
SNR	Signal-to-noise ratio
TDMA	Time division multiple-access

List of Figures

Chapter 1

Introduction

In the language of engineering and science, a group of interacting, interrelated, or interdependent elements forming a complex whole is often called a system [FD]. For example, the solar system comprising of the Sun and its planetary system of eight planets, their moons, and other non-stellar objects. A system is dynamical if its behavior changes over time due to external or internal factors, for example physical changes in a human body, climate change on earth, and motion of a vehicle. If the behavior of a dynamical system can be steered in a desired way by applying appropriate actions or forces, then it is referred to as a control system. There are numerous control systems around us, that either exist naturally or are engineered by humans. An example of a biological control system is the regulation of glucose in the bloodstream through the production of insulin by the pancreas [AM09]. An engineering example is an air flight control system which assists pilot in flying an aircraft. In manufacturing industry, the desired physical or chemical composition of a product is often achieved by employing certain feedback control mechanism. Even our economies are based on complex interactions between individuals and corporations.

Some control systems have spatially distributed components such that the operation of these distributed components is coordinated by exchange of information to achieve an overall objective. These are commonly known as networked control systems. More precisely, a networked control system has four main components: sensors, actuators, controllers, and the process (or system) to be controlled. These components communicate to perform coordinated actions. For example, our nervous system coordinates voluntary and involuntary actions by transmitting signals between different parts of our body. In the engineering world, networked control systems have existed for several decades in power plants, manufacturing industry, and transportation systems where information is usually collected from different locations and then communicated to a central station using wired connections. Important decisions/assessments are made at the central station which are then communicated to different locations using wired connections. Due to wired connections, the implementation of these systems have been very limited in the past.

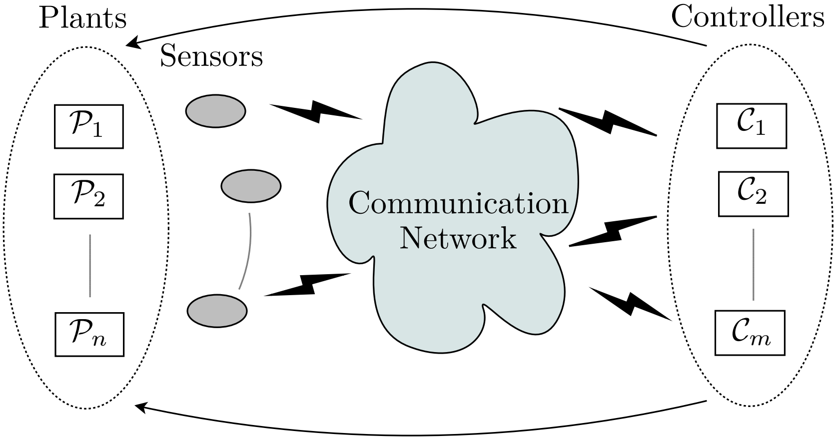


Figure 1.1: Feedback control of plants (dynamical systems) over a wireless network. The sensors monitor the plants and communicate their measurements to remotely situated controllers. The controllers then apply actions to control the plants.

However, the recent advances in wireless access technologies and the rapid development of low cost and energy efficient devices that are capable of sensing, computing, and transmitting, have laid a foundation to build distributed systems that can reliably perform under reasonable cost. The use of wireless technologies in networked control systems may lead to a large scale deployment of distributed components that are mobile, flexible and cost effective. Such system have potential applications in many areas such as machine-to-machine communication for security, surveillance, industrial production, health monitoring, remote surgery, environment management, space missions, and intelligent transport systems.

The replacement of wired connections with wireless connections in networked control systems provide many advantages such as reduced wiring, lower installation and maintenance costs, higher flexibility and adaption capability. However, there are additional challenges in the design and analysis of such systems due to the insertion of less reliable and insecure wireless communication links. In recent years, the researchers in the areas of control theory, communication theory, and computer science have directed substantial attention to address the challenges associated with successful implementation of networked control systems. This is a multi-disciplinary research topic that requires a deep understanding of interaction between information and control sciences. An efficient implementation of networked control systems requires a joint design of control and communication strategies, but unfortunately control engineering and communication engineering have been widely studied separately with very little overlap. The tools, methodologies, and systems developed in these areas are quite mature, but they have very little ground in common. Communication theory primarily focuses on reliable transmission of information and the delay in transmission is of secondary importance. On the other hand, control system designs are robust to uncertainties but very sensitive to delays. It is not simple to merge the systems developed in these areas, therefore, a new and unified approach is needed to understand the fundamental principles and finally build efficient networked control systems. This thesis also makes an effort in this direction by studying some open problems on sensing, estimation, and control over networks.

In this thesis, we consider the scenario depicted in Figure 1.1, where there are a number of plants (systems to be controlled) that are monitored by a group of sensor nodes. These sensor nodes communicate their measurements to remotely situated control units over wireless links. Based on the signals received from the sensors, the remote controllers take appropriate actions to control the plants remotely. Some basic questions we try to answer in this thesis are: How much transmission power the sensors should use to stabilize dynamical systems over a communication network using remote controllers? How do the channel impairments such as noise and interference affect stability of a remote control system? What are the suitable transmission schemes in different network architectures when the objective is to estimate or control state of a dynamical system? Although large networked control systems are the main motivation of this thesis, we focus on small systems with fewer components to understand the fundamental principles. An accurate analysis of small systems can give us right intuitions to understand and build networked control systems on a larger scale.

In the rest of this chapter, we first provide a brief introduction to stochastic control, stochastic networked control, team decision problems, and the problem of stabilization over communication channels. We then give an outline of the main contributions of this thesis. The purpose is to first make the reader familiar with some technical terminologies and basic concepts which are helpful in understanding the key contributions of this thesis as well as the technical discussions in the following chapters.

1.1 Stochastic Control

Consider a discrete-time system with the following state and measurement equations:

$$\begin{aligned} X_{t+1} &= f_t\left(X_t, U_t, W_t\right), \quad t \in \mathbb{N} \\ Y_t &= h_t(X_t, V_t), \end{aligned} \tag{1.1}$$

where $\mathbb{N} = \{0, 1, 2, \ldots\}$, $X_t \in \mathbb{X}$ is the state variable, $U_t \in \mathbb{U}$ is the action variable, $W_t \in \mathbb{W}$ is the system noise (process noise), and $V_t \in \mathbb{V}$ is the measurement noise. The spaces $\mathbb{X}$ and $\mathbb{U}$ are termed as state and action spaces, and the spaces $\mathbb{W}, \mathbb{V}$ are called noise spaces. We assume that $\{\mathbb{X}, \mathbb{U}, \mathbb{W}, \mathbb{V}\}$ are non-empty Borel spaces. The system function f_t and the measurement function h_t are Borel measurable functions. Moreover, the initial state X_0 and the disturbances $\{W_t, V_t\}$ are random objects with distributions μ_{x_0}, μ_{w_t}, and μ_{v_t} respectively. At any time t, the controller has access to the information variable $I_t = \{Y_0, Y_1, \ldots, Y_t, U_0, U_1, \ldots, U_{t-1}\}$ and it employs a policy or strategy γ_t such that

$$\begin{aligned} \gamma_t &: I_t \mapsto U_t \\ \gamma_t &: \mathbb{Y}^{t+1} \times \mathbb{U}^t \to \mathcal{P}(\mathbb{U}), \end{aligned} \tag{1.2}$$

where $\mathcal{P}(\mathbb{U})$ is the space of probability measures defined on $\mathbb{U}$. A control policy is deterministic if $\gamma_t : \mathbb{Y}^{t+1} \times \mathbb{U}^t \to \mathbb{U}$. A control policy is known as *Markov* or *memoryless* if the controller uses only Y_t to generate U_t. A control policy is called *stationary* if $\gamma_t = \gamma$ for all t.

In a stochastic control problem the controller can have different objectives, therefore, the control policy needs to be designed accordingly. For instance, one objective of the controller might be to stabilize the system in some appropriate sense, for example in the mean-square sense: $\sup_t \mathbb{E}\left[X_t^2\right] < \infty$, or in ergodic sense, i.e., $\lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} g(X_t)$ exists, where g is a real-valued Borel measurable function. Another common objective is to design a control policy $\bar{\gamma} := \{\gamma_0, \gamma_1, \ldots \gamma_T\}$ such that the following additive cost function is minimized:

$$J(\bar{\gamma}) = \mathbb{E}\left[\sum_{t=0}^{T} c\left(X_t, U_t\right)\right], \tag{1.3}$$

where $T \in \mathbb{N}$ and $c : \mathbb{X} \times \mathbb{U} \to \mathbb{R}$ is a cost function. In this case, an optimal policy $\bar{\gamma}^\star$ would satisfy $J(\bar{\gamma}^\star) \leq J(\bar{\gamma})$ for all $\bar{\gamma} \in \Gamma$, where Γ is the set of admissible control policies. A control policy is admissible if it is measurable with respect to sigma-algebra generated by the information variable I_t.

In stochastic control problems, the control action U_t can have two effects or roles, with contradictory objectives [Söd02]:

1. Control action u_t should be chosen to make the present cost and the future costs $\mathbb{E}\left[c\left(X_{t+k}, U_{t+k}\right)\right]$ for $k \geq 0$, small. This often implies that u_t has to be small. We refer to this as *action effect.*

2. Control action u_t should be chosen to obtain more precise information about future states. This often implies that u_t has to be large to dominate over the disturbance in the system. This effect is known as *probing.*

Since the *action role* and the *probing role* of control have usually conflicting objectives, an optimal controller is the one that achieves an optimal trade-off. Due to the above two effects, we say that the control has a *dual effect* or a *dual role.* The dual effect of control does not occur when the controller has perfect information of the state, i.e., $Y_t = X_t$. Although in most stochastic control problems with incomplete state information dual effect of control is present, there may exist rare instances where dual effect is not present. Such problems are known as *neutral problems* and an optimal policy for such problems can be obtained in two steps: i) determine the conditional distribution of the state (information state) and express the cost in terms of information state, ii) determine optimal control laws as functions of information state. This property is known as a separation principle or separation of estimation and control [Wit71], i.e., the controller first computes an optimal estimate of the state and then applies action using the state estimate. One popular stochastic control problem where the separation of estimation and control holds is given below.

LQG Control Problem. Consider the following linear system model:

$$\begin{aligned} X_{t+1} &= AX_t + BU_t + W_t \\ Y_t &= CX_t + V_t, \end{aligned} \tag{1.4}$$

where V_t and W_t are jointly Gaussian white noise sequences and A, B, C are matrices of appropriate dimensions. The initial state X_0 is assumed to be Gaussian distributed and independent of the noise variables V_t and W_t. Suppose the objective is to minimize the following quadratic cost function:

$$J(\bar{\gamma}) = \mathbb{E}_{\bar{\gamma}} \left[X_N^T Q_0 X_N + \sum_{t=0}^{N} \left(X_t^T Q_1 X_t + U_t^T Q_2 U_t \right) \right], \tag{1.5}$$

over all admissible control policies, $\bar{\gamma} \in \Gamma$. This is known as the Linear Quadratic Gaussian (LQG) control problem, since the system model is linear, the cost is quadratic and disturbances are Gaussian. The optimal control policy for this problem is given by $U_t = -L_t \hat{X}_t$, where L_t is a time varying matrix and $\hat{X}_t = \mathbb{E}[X_t | I_t]$ is the optimal state estimate that can be generated using Kalman filter. Hence, it is allowed to separate estimation and control in the LQG control optimal. Even if we assume non-Gaussian disturbances and restrict controller to be linear, then optimal linear controller also exhibits the separation structure. The LQG control problem in fact exhibits a stronger property than the separation principle known as certainty equivalence, which says that one can replace all random variables in the system model by their mean values and then determine an optimal controller for this deterministic system [Söd02].

1.2 Stochastic Networked Control

A networked control system is comprised of multiple agents which are often performing two major tasks: i) Generate suitable signals to communicate information, ii) Apply control actions to minimize errors and reduce costs. In the stochastic control framework introduced in the previous section, there is a central agent (controller) with perfect memory (recall). But networked control systems are decentralized in nature. A simple example of a networked control system would be to assume that the measurement function h_t in the state equation (1.1) is not fixed and it has to be designed along with the control function γ_t to minimize some cost function. That is, in this simple stochastic networked control example, the objective is to jointly design measurement and control strategies. Obviously stochastic networked control problems are much harder to solve than stochastic control problems due to their decentralized nature. As a concrete example, we look at the following problem on joint design of measurement and control strategies for a scalar linear systems, that was studied by Bansal and Başar in 1989 [BB89].

Consider the following scalar system:

$$\begin{aligned} X_{t+1} &= \lambda_t X_t + U_t + W_t, \quad t \in \mathbb{N} \\ Y_t &= S_t + V_t, \end{aligned} \tag{1.6}$$

where $S_t = h_t(X_t, Y_0, Y_1, \ldots, Y_{t-1})$, $U_t = \gamma_t(Y_0, Y_1, \ldots, Y_t)$, and $\{h_t, f_t\}$ are Borel measurable functions. The initial state X_0 and the noise variables $\{W_t, V_t\}$ are all independent and zero mean Gaussian distributed. The objective is to jointly design measurement and control policies $\{h_t, \gamma_t\}$ such that the following cost is minimized:

$$J(\bar{h}, \bar{\gamma}) = \mathbb{E}_{\bar{h},\bar{\gamma}} \left[\sum_{t=0}^{N} a_{t+1} X_{t+1}^2 + b_t U_t^2 + q_t S_t^2 \right], \tag{1.7}$$

where $a_t, b_t, q_t > 0$ and $\bar{h} := \{h_0, h_1, \ldots, h_N\}$ and $\bar{\gamma} := \{\gamma_0, \gamma_1, \ldots, \gamma_N\}$. Note that this problem becomes equivalent to the LQG control problem if h_t is fixed to be linear in the state, for which separation structure is optimal. However, for non-linear measurement functions, the optimality of separation structure does not apply in general and this problem is very hard to solve. Interestingly, Bansal and Başar showed that the optimal measurement and control strategies for this problem are actually linear with the following structure:

$$S_t = m_t \left(X_t - \mathbb{E}\left[X_t | Y_0, Y_1, \ldots, Y_{t-1} \right] \right), \quad U_t = -l_t \mathbb{E}\left[X_t | Y_0, Y_1, \ldots, Y_{t-1} \right], \tag{1.8}$$

where m_t and l_t are some scalars (complete description is given in [BB89]). It is noteworthy that the optimal measurement policy is to linearly transmit innovation at each time step. Similar results carry over to infinite horizon problem with discounted cost as shown in [BB89]. To this end, we want to highlight that optimal measurement and control policies for vector valued systems in quadratic Gaussian settings are still unknown.

1.3 Networked Control System as Stochastic Team

A networked control system can be viewed as a stochastic dynamic team where multiple decision makers (agents) strive for a common goal [Rad62, YB13]. These decision makers have the same probabilistic description of the system but usually have different observations or measurements. For example, in the simple networked control example discussed in the previous section, there are two decision makers–a sensor and a controller, and their common objective is to minimize a quadratic cost function. A comprehensive discussion on team decision problems in various settings can be found in Chapter II and Chapter III of [YB13]. In the following, we provide a brief introduction to mathematical framework for team decision problems and some fundamentals which will be useful in understanding the contributions of this thesis.

A stochastic team decision problem has the following five ingredients:

1. Number of decision makers and action sets available at each decision maker.
2. A probabilistic description of the uncertainties in the system.
3. The measurements/observations at the decision makers.
4. A cost function.
5. A solution concept based on which we can compare different policies.

We illustrate this with the help of an example: Consider a sequential[1] dynamic[2] team with N decision makers $\{DM^1, DM^2, \ldots, DM^N\}$, where DM^i has an information space $\mathbb{Y}^i$ and an action space $\mathbb{U}^i$. Let $(\Omega, \mathcal{F}, P)$ be the probability space associated with uncertainties in the system. Consider a decision making sequence where DM^1 acts first, then DM^2 acts, and so on. Let $h_i : \Omega \times \mathbb{U}^1 \times \mathbb{U}^2 \times \cdots \times \mathbb{U}^{i-1} \to \mathbb{Y}^i$ be the measurement function of DM^i and $\gamma_i : \mathbb{Y}^i \to \mathbb{U}^i$ be the action policy of DM^i, such that $Y^i = h_i(\omega, U^1, U^2, \ldots, U^{i-1})$ and $U^i = \gamma_i(Y^i) = \gamma_i(h_i(\omega))$, where $w \in \Omega$, $U^i \in \mathbb{U}^i$ and $\{\mathbb{Y}^i, \mathbb{U}^i\}$ are the information and the action spaces of DM^i. The measurement functions $\{h_1, h_2, \ldots, h_N\}$ for all decision makers are fixed. The action policy of the team $\bar{\gamma} := \{\gamma_1, \gamma_2, \ldots, \gamma_N\}$ has to be designed to minimize an average cost:

$$J(\bar{\gamma}) = \mathbb{E}\left[c(\omega, U^1, U^2, \ldots, U^N)\right], \tag{1.9}$$

where $c : \Omega \times \mathbb{U}^1 \times \mathbb{U}^2 \times \cdots \times \mathbb{U}^N \to \mathbb{R}$ denotes a cost function. A team policy $\bar{\gamma}^\star$ is *optimal* if

$$J(\bar{\gamma}^\star) = \inf_{\bar{\gamma} \in \Gamma} J(\bar{\gamma}), \tag{1.10}$$

where Γ is the set of admissible team policies. Note that an optimal policy may not exist in a stochastic team decision problem. Another weaker notion of optimality is *person-by-person optimality* which is defined as follows. A team strategy $\bar{\gamma}^\star := \{\gamma_1^\star, \gamma_2^\star, \ldots, \gamma_N^\star\}$ is person-by-person optimal if

$$J(\gamma_1^\star, \ldots, \gamma_{i-1}^\star, \gamma_i^\star, \gamma_{i+1}^\star, \ldots, \gamma_N^\star) \le J(\gamma_1^\star, \ldots, \gamma_{i-1}^\star, \gamma_i, \gamma_{i+1}^\star, \ldots, \gamma_N^\star), \tag{1.11}$$

[1]In a sequential team, there exists a predefined order in which the decision makers take actions, whereas in a non-sequential team this order can be uncertain and may depend on other parameters.

[2]In a dynamic team, the information available to a decision maker can be affected by actions of other decision makers, whereas in a static team the actions of one decision maker do not influence information available to the other decision makers.

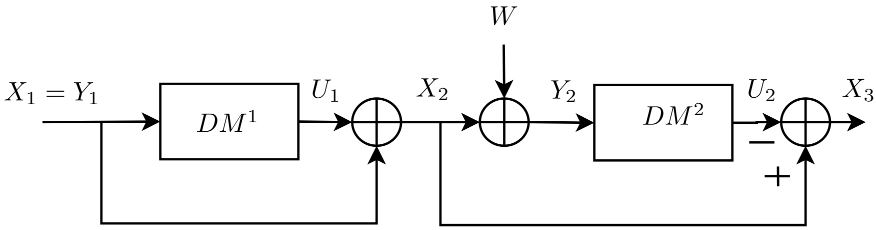


Figure 1.2: Witsenhausen Counterexample.

for all $\gamma_i \in \Gamma_i$ and for all $i \in \{1, 2, \ldots, N\}$. Person-by-person optimality is a necessary but not sufficient condition for global or team optimality.

Since $U^i = \gamma_i(h_i(\omega, U^1, U^2, \ldots, U^{i-1}))$, the cost in (1.9) also depends on the measurement functions $\bar{h} := \{h_1, h_2, \ldots, h_N\}$. We refer to $\bar{h}$ as the information structure of the team problem. The importance of information structure is central in stochastic team decision problems [Wit68, SA74, Ho80, MMRY12, YB13]. In the following we briefly introduce information structures and discuss their effect on the solution of team decision problems.

Broadly speaking, there are three main types of information structures: classical, quasi-classical, and non-classical. A team is said to have a classical information structure if DM^i has access to the information available at DM^k for all $k < i$. In the case of the classical information structure, the information is expanding , i.e., $\mathbb{Y}^1 \subseteq \mathbb{Y}^2 \subseteq \mathbb{Y}^3 \cdots \subseteq \mathbb{Y}^N$. A quasi-classical (or partially nested) information structure is the one in which if DM^k affects the information of DM^j, then DM^j must know the information available at DM^k i.e., $\mathbb{Y}^k \subseteq \mathbb{Y}^j$. An information structure which is neither classical nor quasi-classical is known as a non-classical information structure. In the classical information structure the information is expanding and one can use dynamic programming to find optimal action policies. However, for quasi-classical and non-classical information structures, there does not exist a systematic procedure to find optimal control laws. Only for LQG systems with quasi-classical information structures, optimal control policies are known to be affine [Ho80]. Many team problems are non-classical in nature and unfortunately a systematic theory to address these problems is lacking [Bas08]. In the following we a give an example of a well-known team decision problem with non-classical information pattern.

Witsenhausen's Counterexample. Consider a team problem with two decision makers depicted in Figure 1.2. The decision maker DM^1 observes $Y_1 = X_1$ where $X_1 \sim \mathcal{N}(0, \sigma^2)$ and then applies an action $U_1 = \gamma_1(Y_1)$. The decision maker DM^2 observes $Y_2 = X_1 + U_1 + W$ where $W \sim \mathcal{N}(0, \sigma_w^2)$, and then applies an action $U_2 = \gamma_2(Y_2)$. The objective is to find an optimal team decision policy $\bar{\gamma}^\star = \{\gamma_1^\star, \gamma_2^\star\}$ such that the following cost is minimized:

$$J(\bar{\gamma}) = \mathbb{E}\left[k^2 U_1^2 + X_3^2\right], \tag{1.12}$$

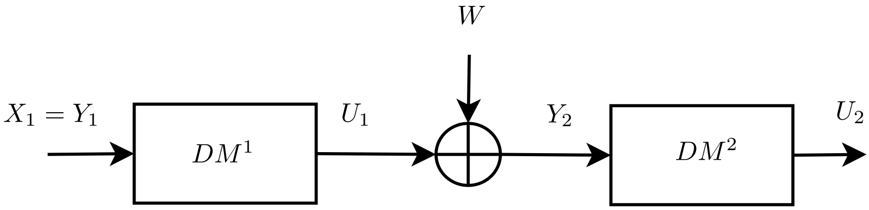


Figure 1.3: Transmission of a Gaussian variable over a Gaussian channel.

where $X_3 = X_2 - U_2$ and $k > 0$. Note that the problem has non-classical information structure because DM^2 does not have access to Y_1. This problem was introduced by Witsenhausen in 1968 [Wit68], even though the problem looks simple but optimal decision policies are unknown till today. Witsenhausen proved that optimal policies for this problem exist and further showed by giving a counter example that the optimal policies are not linear [Wit68].

In problems with a non-classical information pattern, a decision maker can encode some information in his/her action about those variables which are not known to other decision makers, which is commonly known as a signaling incentive. Thus the control has a triple role in team decision problems with non-classical information structures: i) probing role, ii) action role, iii) signaling role. An optimal decision policy is the one that achieves an optimal tradeoff between these three roles of control with conflicting objectives. It has been pointed out that in fact it is not only the non-classical nature but also the structure of cost function that contributes to the difficulty of finding optimal solutions [Bas08]. There exist rare instances where linear policies are actually optimal in non-classical setting. One such example is given below.

Gaussian Test Channel. Consider a team problem with two decision makers depicted in Figure 1.3. The decision maker DM^1 observes $Y_1 = X$ where $X \sim \mathcal{N}(0, \sigma^2)$ and then applies an action $U_1 = \gamma_1(Y_1)$. The decision maker DM^2 observes $Y_2 = U_1 + W$, where $W \sim \mathcal{N}(0, \sigma_w^2)$, and then applies an action $U_2 = \gamma_2(Y_2)$. The objective is to find an optimal team decision policy $\bar{\gamma}^\star = \{\gamma_1^\star, \gamma_2^\star\}$ such that the following cost is minimized:

$$J(\bar{\gamma}) = \mathbb{E}\left[(X - U_2)^2\right], \tag{1.13}$$

subject to the constraint: $\mathbb{E}[U_1^2] \leq P$. Note that the information structure of this problem is non-classical as well, however, an optimal decision policy for this problem is known and it is linear [Ber71, Bas08, YB13].

1.4 Control Over Communication Channels

Consider the scenario depicted in Fig. 1.4, where a plant has to be controlled or stabilized (in some sense) over a communication channel by the joint actions of

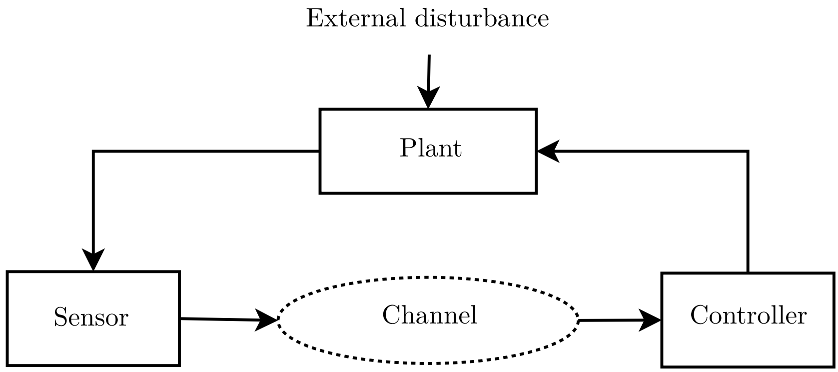


Figure 1.4: Control over a communication channel.

a sensor and a controller. We raise two important questions: i) Given an unstable plant and a communication channel, can we find measurement (sensing) and control policies such that the plant is stable? ii) Given a plant (source) and a communication channel, what are the sensing and control policies that minimize an expected cost function of plant's state process? Obviously to answer such question, one would require a description of the plant and communication channel, and also the notion of stability and cost function that one is interested in. Throughout this thesis, we consider linear time invariant plants, Gaussian communication channels, quadratic cost function, and mean-square stability which requires second moment of plant's state process to be bounded.

As an example, let us consider the problem of mean-square stabilization of a scalar linear time invariant (LTI) plant over an additive white Gaussian noise channel.

Stabilization Over a Gaussian Channel. Consider a scalar discrete time LTI system with the following state equation:

$$X_{t+1} = \lambda X_t + U_t + W_t, \quad t \in \mathbb{N}, \tag{1.14}$$

where $X_t \in \mathbb{R}, U_t \in \mathbb{R}$, and $W_t \in \mathbb{R}$ are state, control, and noise variables. We assume that the open-loop system is unstable ($|\lambda| > 1$) and the initial states X_0 is a random variable with Gaussian distribution. Assume that at any time t the sensor has access to the information set $I_t^S := \{X_0, X_1, \ldots, X_t, U_0, U_1, \ldots, U_{t-1}\}$ and it produces a signal $S_t = f_t(I_t^S)$, where f_t denotes the sensing/encoding policy such that the following average transmit power constraint is ensured: $\mathbb{E}[S_t^2] \leq P$. The signal transmitted by the sensor is received at the controller as $R_t = S_t + Z_t$, where $Z_t \sim \mathcal{N}(0, N)$ denotes Gaussian distributed channel noise variable. Thus, at any time t the controller has access to the information set $I_t^C := \{R_0, R_1, \ldots, R_t\}$ and it applies an action $U_t = \pi_t(I_t^C)$, where π_t denotes the control policy. The common objective of the sensor and the controller is to keep the second moment of the state process bounded, i.e., to ensure $\mathbb{E}\left[X_t^2\right] < \infty$ for all t. This notion of stability is commonly known as mean-square stability or quadratic stability. In [Sah01], it was shown that the scalar plant can be mean-square stabilized over the given scalar Gaussian channel if and only if

$$\log(|\lambda|) < \frac{1}{2}\log\left(1 + \frac{P}{N}\right). \tag{1.15}$$

Furthermore, linear and memoryless (Markov) sensing and control policies are optimal for mean-square stabilization. The fact that linear policies are optimal for quadratic stabilization is not very surprising because we have already discussed an example in Sec.1.2 where linear measurement and control policies minimize a quadratic cost function of state and control variables. More interesting observation at this point is the relationship given in (1.15) between plant parameter λ and the channel parameter $\frac{P}{N}$, which is actually signal-to-noise ratio of the channel. This relationship characterizes signal-to-noise requirement for mean-sqaure stabilization of a scalar LTI plant over a scalar AWGN channel. A communication theorist will also identify that the term on the right hand side (RHS) of the inequality in (1.15) is the Shannon capacity of the AWGN channel.

There exists a diverse literature on the problem of control over communication channels, focusing on different plant models, channel models, performance measures, and design constraints. For instance the plant and the channel models can be either discrete-time or continuous with different network topologies and different assumption on channel impairments. There can be different design constraints such as transmission delay-constraints, transmit power constraints, bandwidth constraint etc. In the following, we provide a brief overview of the literature on control over communication channels.

Bibliographic Notes. Important contributions to control over communication channels include [BB89, Bai99, WB99, NE00, EM01, NE04, Eli04, MS04, TSM04, TM04, MDE06, SM06, MS07, BMF07, MD08, MRFB09, YT09, YB11, FA11, LDA11, Yük12, MCF13]. In the early works [Str65, Won68] it is shown that for linear systems subject to Gaussian noise with linear sensing policies having perfect memory (recall), the optimal control policies are linear and there exists a separation property between estimation and control. However, Witsenhausen showed in [Wit68] via a simple counter example that linear policies may not be optimal when there are more than two or more decision makers (sensors/controllers) without perfect memory (recall). The *Witsenhuasen problem* is still unsolved. Its difficulty lies in its non-classical information structure [Rad62, Wit71, Ho80]. The problem of remotely controlling dynamical systems over communication channels is often studied with methods from stochastic control theory and information theory. The seminal paper by Bansal and Başar [BB89] used fundamental information theoretic arguments to obtain optimal policies for LQG control of a first order plant over a point to point Gaussian channel. Minimum rate requirements for stabilizability of a noiseless scalar plant were first established in [Bai99, WB99] followed by [NE00]. Further rate theorems for stabilization of linear plants over some discrete and continuous alphabet channels can be found in [TM04, BMF07, MD08, CFD08, MRFB09, YTCW09, YB11, FMS10, SM11, Yük12, SGQ10, VSC12]. If a communication channel is used as part of the feedback loop to stabilize a plant, then classical Shannon capacity may not be a relevant metric. In [SM06], Sahai and Mitter introduced the notion of “anytime capacity" based on an operational definition to characterize moment stabilizability of dynamical systems over communication channels. However, the characterization of the anytime capacity for general and multi-terminal networks is a difficult task. The papers [BB89, TSM04, TM04, BMF07, MRFB09, YTCW09, FMS10, YT09, SM11, YB11, SGQ10, VSC12] addressing control over Gaussian channels are more relevant to this thesis. In [BB89] linear sensing and control policies are shown to be optimal for the LQG control of a first order linear plant over a point-to-point Gaussian channel. A necessary condition for stabilization relating eigenvalues of the plant to the capacity of the Gaussian channel first appeared in [TSM04, TM04]. Some important contributions on stabilization over Gaussian channels with average transmit power constraints have been made in [BMF07, MRFB09, FMS10, YTCW09, SM11, KGL11, VSC12]. In [BMF07] sufficient conditions for stabilization of both continuous time and discrete time multi-dimensional plants over a scalar white Gaussian channel were obtained using linear time invariant (LTI) sensing and control schemes. It was shown in [BMF07, FMS10] that under some assumptions there is no loss in using LTI schemes for stabilization. That is the use of non-linear time varying schemes does not allow stabilization over channels with lower signal-to-noise ratio. The stability results were extended to a colored Gaussian channel in [MRFB09]. In [YB11] the authors considered noisy communication links between both sensor–controller and controller–actuator, and presented necessary and sufficient conditions for mean square stability. Stabilization of noiseless LTI plants over parallel white Gaussian channels subject to transmit power constraint has been studied in [YTCW09, SM11, KGL11, VSC12]. The paper [YTCW09] considers output feedback stabilization and [SM11] considers state feedback stabilization, and they both derive necessary and sufficient conditions for stability under a total transmit power constraint.

Although several problems related to control of a linear plant over point-to-point Gaussian channels have been extensively studied, the literature on control of multiple plants over Gaussian networks is scarce. Many important issues such as resource allocation, interference management, cooperation between terminals, distributed designs etc. have been widely studied for communication in classical multi-terminal networks in information theory, however, it is not straightforward to apply those results to the problems where communication channels are used as part of a feedback loop to control dynamical systems. Such problems are hard because a network can have an arbitrary topology and every node within the network can have memory and can employ any transmission strategy. The papers [Tat03] and [GDH+09] have derived conditions for stabilization over networks with digital noiseless channels and analog erasure channels respectively, however, those results do not apply to noisy networks. In [SM06, Yük12] moment stability conditions in terms of error exponents have been established. However, even a single letter expression for channel capacity of some basic Gaussian channels such as relay channel, broadcast channel, and interference channel [CT06] are not known in general.

1.5 Thesis Outline and Contributions

This thesis studies several open problems related to sensing, estimation and control of linear dynamical systems over communication networks. In particular we consider transmission over point-to-point channels, relay channels, multiple-access channels, broadcast channels, and interference channels. We assume that all the nodes in the network have limited transmit power and the communication links between all nodes are modeled as additive white Gaussian channels. The thesis is divided into following three parts, where each part is comprised of two chapters. Part I focuses on control of a linear system (plant) over networks, where different agents in the network can cooperate to communicate state of the system to a remote controller. The common objective of all agents is to stabilize the plant in mean-square sense. Part II considers control of a linear system over basic cascade and parallel Gaussian networks when the objective is to minimize a quadratic cost function of the system's state variables. Part III is a continuation of Part I; it considers stabilization of multiple linear dynamical systems over fundamental multi-user Gaussian channels. The thesis ends with a summary of the key findings and some interesting directions for future research.

Some of the results presented in the thesis have already been published in journals and conferences, and some are under review. Parts of the thesis are adopted from the corresponding research papers nearly verbatim. In the following we give a brief introduction of each chapter along with the reference to the associated papers, so that the reader knows in advance what to expect in each chapter.

Part I: Chapter 2

This chapter studies the problem of mean-square stabilizing a multi-dimensional linear time invariant plant over a multi-dimensional point-to-point Gaussian channel and tries to address the following questions: i) Is there any loss in restricting sensing and control policies to be linear? ii) Should the policies be time-invariant or time-variant? iii) What is an optimal linear scheme? In particular, we show that linear time varying policies are optimal for stabilization in a large class of settings. This chapter is based on

[ZOYS13a] A. A. Zaidi, T. J. Oechtering, S. Yüksel, and M. Skoglund,"Stabilization and Control Over Gaussian Networks," in: B. Bernhardsson, G. Como, A. Ranzter (Eds.), "Information and Control in Networks," *Springer*. (In press)

[ZYOS] A. A. Zaidi, S. Yüksel, T. J. Oechtering and M. Skoglund, "On the tightness of linear policies for stabilization of linear systems over Gaussian networks," submitted to *Systems & Control Letters*, June 2013.

Part I: Chapter 3

This chapter addresses the problem of remotely stabilizing a noisy linear time invariant plant over a Gaussian relay network. The network is comprised of a sensor node, a group of relay nodes that can cooperate to communicate the observations of the plant's state to the remote controller. Necessary as well as sufficient conditions for mean-square stabilization over various network topologies are derived. The sufficient conditions are in most cases obtained using delay-free linear policies and the necessary conditions are obtained using information theoretic tools. Different settings where linear policies are optimal, asymptotically optimal (in certain parameters of the system), and suboptimal have been identified. This chapter is based on

[ZOYS13b] A. A. Zaidi, T. J. Oechtering, S. Yüksel and M. Skoglund, "Stabilization of linear systems over Gaussian networks," *IEEE Transactions Automatic Control*, Accepted July 2013.

[ZOS10a] A. A. Zaidi, T. J. Oechtering and M. Skoglund, "Rate sufficient conditions for closed–loop control over AWGN relay channels," in *Proc. IEEE International Conference on Control and Automation (ICCA)*, June 2010.

[ZOYS10] A. A. Zaidi, T. J. Oechtering, S. Yüksel and M. Skoglund, "Closed-loop control over half-duplex AWGN relay channels," in *Proc. Reglermöte*, June 2010.

[ZOYS11] A. A. Zaidi, T. J. Oechtering, S. Yüksel and M. Skoglund, "Sufficient conditions for closed-loop control over a Gaussian relay channel," in *Proc. American Control Conference (ACC)*, June 2011.

Part II: Chapter 4

This chapter considers the problem of causal transmission of a memoryless Gaussian source over a two-hop memoryless Gaussian relay channel. The source and the relay encoders have average transmit power constraints, and the performance criterion is mean-squared distortion. The main contribution in this chapter is to show that unlike in the case of a point-to-point scalar Gaussian channel, linear encoding schemes are not optimal over a two-hop relay channel in general, extending the sub-optimality results which are known for more than three hops. In some cases, simple three-level quantization policies employed at the source and at the relay can outperform the best linear policies. Further a lower bound on the distortion is derived and it is shown that the distortion bounds derived using cut-set arguments are not tight in general for sensor networks. This chapter is based on

[ZYOS13] A. A. Zaidi, S. Yüksel, T. J. Oechtering and M. Skoglund, "On optimal policies for control and estimation over a Gaussian relay channel," *Automatica*, vol. 49, no. 9, Sep. 2013.

[ZYOS11] A. A. Zaidi, S. Yüksel, T. J. Oechtering and M. Skoglund, "On optimal policies for control and estimation over a Gaussian relay channel," in *Proc. IEEE Conference on Decision and Control (CDC)*, Dec. 2011.

Part II: Chapter 5

This chapter considers a scenario of distributed sensing for networked control systems. A non-linear scheme for distributed sensing and transmission is proposed. The proposed non-linear delay-free sensing and transmission strategy is compared with the well-known amplify-and-forward strategy, using the LQG control cost as a figure of merit. It is demonstrated that the proposed non-linear scheme outperforms the best linear scheme even when there are only two sensors in the network. The proposed sensing and transmission scheme can be implemented with a reasonable complexity and it is shown to be robust to the uncertainties in the knowledge of the sensors about the statistics of the measurement noise and the channel noise. This chapter is based on

[AZWS11] M. Andersson, A. A. Zaidi, N. Wernersson, M. Skoglund, "Nonlinear distributed sensing for closed-loop control over Gaussian channels," in *Proc. IEEE Swedish Communication Technologies Workshop (Swe-CTW)*, Oct. 2011.

Part III: Chapter 6

This chapter studies quadratic stabilization of two linear time invariant plants over noisy multiple-access and broadcast communication channels with arbitrarily distributed initial states. We propose to use communication and control schemes based on the coding schemes introduced by Ozarow et al. [Oza84, OLYC84] for the multiple-access and the broadcast channels with noiseless feedback which are extensions of the Schalkwijk-Kailath coding scheme [SK66]. By employing the proposed communication and control schemes over the multiple-access and the broadcast channels, we derive stability regions those are sufficient for mean square stability of the two linearly controlled LTI systems. This chapter is based on

[ZOS10b] A. A. Zaidi, T. J. Oechtering and M. Skoglund, "Sufficient conditions for closed-loop control over multiple-access and broadcast Channels," in *Proc. IEEE Conference on Decision and Control (CDC)*, Dec. 2010.

Part III: Chapter 7

The problem of feedback stabilization of LTI plants over a Gaussian interference channel is considered. Two plants with arbitrary distributed initial states are monitored by two separate sensors which communicate their measurements to two separate controllers over a Gaussian interference channel under average transmit power constraints. Linear delay-free sensing and control schemes are proposed and sufficient conditions for mean-square stability are derived under these schemes. Necessary conditions for mean square-stabilization are also derived, that are shown to be tight for some system parameters. Further it is shown that linear memoryless sensing and control schemes are optimal for stabilization in some special cases. It is shown that the conventional estimation based control strategies can be quite inefficient in multi-terminal settings, unlike point-to-point channels where they are optimal. This chapter is based on

[ZOYS13c] A. A. Zaidi, T. J. Oechtering, S. Yüksel and M. Skoglund, "Multiplant stabilization over Gaussian interference channel," *IEEE Transactions Automatic Control*, To be submitted.

[ZOS11] A. A. Zaidi, T. J. Oechtering and M. Skoglund, "Closed-loop stabilization over Gaussian interference channel," in *Proc. IFAC World Congress*, Aug. 2011.

[ZOS13] A. A. Zaidi, T. J. Oechtering and M. Skoglund, "On stabilization over a Gaussian interference channel," in *Proc. European Control Conference (ECC)*, July. 2013.

1.6 Contributions Outside the Thesis

The author of the thesis also contributed to the published papers [KZS10, ZKYS09c, ZKYS09b, ZKYS09a, FZFS13, BZOS12, BZOS13a, BZOS13b], which are not included in the thesis. In [KZS10, ZKYS09c, ZKYS09b, ZKYS09a] we propose design of delay-free relaying schemes with the objective of maximizing information transmission rate for some Gaussian relay channel setups, and show that non-linear schemes significantly outperform linear schemes. The papers [KZS10, BZOS12] propose the use of relay for managing interference in delay sensitive applications such as large sensor networks and networked control systems. The paper [FZFS13] proposes a distributed algorithm for power control and interference alignment over MIMO interference networks. In [BZOS13a, BZOS13b] we study stabilization of linear systems over a Gaussian interference relay channel and show that considerable improvement in stability can be achieved by deploying relays in interference networks.

[KZS10] M. N. Khormuji, A. A. Zaidi and M. Skoglund, "Interference management using nonlinear relaying," *IEEE Transactions on Communications*, vol. 58, no. 7, pp. 1294–1930, July 2013.

[ZKYS09c] A. A. Zaidi, M. N. Khormuji, S. Yao and M. Skoglund, "Rate-maximizing mappings for memoryless relaying," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, July 2009.

[ZKYS09b] A. A. Zaidi, M. N. Khormuji, S. Yao and M. Skoglund, "Optimized mappings for dimension-expansion relaying," in *Proc. IEEE Signal Processing Advances in Wireless Communications (SPAWC)*, June 2009.

[ZKYS09a] A. A. Zaidi, M. N. Khormuji, S. Yao and M. Skoglund, "Optimized analog network coding strategies for the white Gaussian multiple-access relay channel," in *Proc. IEEE Information Theory Workshop (ITW)*, Oct. 2009.

[FZFS13] H. Farhadi, A. A. Zaidi, C. Fischione, and M. Skoglund, "Distributed interference alignment and power control for wireless MIMO interference networks with noisy channel state information," in *Proc. IEEE Black Sea Conference on Communications and Networking (BlackSeaCom)*, July 2013.

[BZOS12] I. Bilal, A. A. Zaidi, T. J. Oechtering and M. Skoglund, "Managing interference for stabilization over wireless channels," in *Proc. IEEE International Symposium on Intelligent Control (ISIC)*, Oct. 2012.

[BZOS13a] I. Bilal, A. A. Zaidi, T. J. Oechtering and M. Skoglund, "Feedback stabilization using relay in interference channels," in *Proc. IEEE Signal Processing Advances in Wireless Communications (SPAWC)*, June 2013.

[BZOS13b] I. Bilal, A. A. Zaidi, T. J. Oechtering and M. Skoglund, "An optimized linear scheme for stabilization over multi-user Gaussian networks," in *Proc. Information Theory and Applications Workshop (ITA)*, Feb. 2013.

1.7 Copyright Notice

As detailed in Section 1.5, parts of the material presented in this thesis have been already published by IEEE, Elsevier, and Springer and some parts are submitted to IEEE and Elsevier. IEEE, Elsevier, and Springer hold the copyright of the corresponding published papers and will hold the copyright of the corresponding submitted papers if they are accepted. Material is reused in this thesis with permission.

Part I

Single system: stability

Chapter 2

Point-to-point Channels

This chapter considers a setup where a linear time invariant system (plant) with a random initial state and driven by Gaussian noise has to be remotely stabilized. A sensor node monitors the plant and communicates its observations (measurements) to a remotely situated control unit over a communication link, that is modeled as an additive white Gaussian channel. The common goal of the sensors and the controller is to stabilize the plant in closed-loop. Usually in remote control applications, sensing and transmission under strict delay and power constraints are required. Therefore, we focus on delay-free and energy efficient sensing and transmit schemes throughout the chapter. We study mean-square stabilization of a linear dynamical system over vector Gaussian channels. Some real-time sensing and control schemes are proposed and stabilizability of the plant under those schemes is studied. We try to address the following interesting questions in this chapter: i) Is there any loss in restricting sensing and control policies to be linear? ii) Should the policies be time-invariant or time-variant? iii) What is an optimal linear scheme?

If the goal is stabilization in the sense of ergodicity [Yük12] or similar notions such as stability in probability [SM06,MS07], Shannon capacity is the right measure to tell what is possible or not. But if the goal is stabilization in the sense of second moment, then knowing Shannon capacity is not sufficient [SM06]. However, the Gaussian setting is special due to the possible mean-square optimality of linear policies as a consequence of what is known as source-channel matching [GRV03, BB89, TSM04]. In the control literature, these problems have also been considered where the sufficiency of Shannon capacity being greater than a lower bound has been observed in a class of settings [FMS10,SM11,KGL11], which, however, consider noise-free plants. It is observed in [SM11] that LTI schemes are not optimal for stabilization over parallel channels. For optimal encoding in the unmatched case, linear encoding is also not optimal in general [Pil69, LP76]. One of the main results presented in this chapter states that linear schemes are actually optimal for a wide class of source-channel pairs when the objective is quadratic stability.

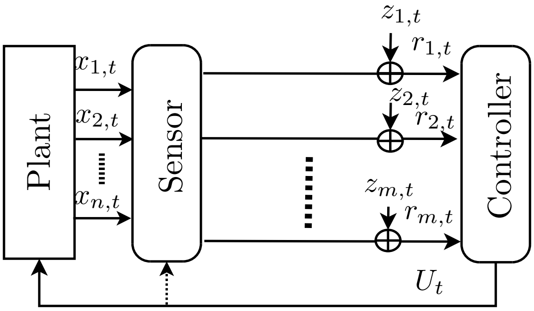


Figure 2.1: Control over parallel Gaussian channels.

2.1 Problem Formulation

Consider the following linear time invariant system:

$$X_{t+1} = AX_t + BU_t + W_t, \qquad t \in \mathbb{N}, \tag{2.1}$$

where $X_t \in \mathbb{R}^n$ is a state process, $U_t \in \mathbb{R}^n$ is a control process, $W_t \in \mathbb{R}^n$ is an independent and identically distributed sequence of Gaussian random variables with zero mean and covariance K_W. The matrices A and B are of appropriate dimensions and the pair (A, B) is controllable. Let $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ denote the eigenvalues of the system matrix A. Without loss of generality we assume that all the eigenvalues of the system matrix are outside the unit disc ($1 \leq |\lambda_i| < \infty$ for all i), i.e., all modes are unstable. The initial state X_0 is a random variable with bounded differential entropy $|h(X_0)| < \infty$ and a given covariance matrix Λ_0 with $\mathrm{Tr}\,[\Lambda_0] < \infty$. Consider the scenario depicted in Fig. 2.1, where a sensor observes an n-dimensional state process and transmits it to a remote controller over m parallel Gaussian channels. At any time instant t, $S_t := [s_{1,t}, s_{2,t}, \ldots, s_{m,t}]$ and $R_t := [r_{1,t}, r_{2,t}, \ldots, r_{m,t}]$ are the input and output of the channel, where $r_{i,t} = s_{i,t} + z_{i,t}$ and $z_{i,t} \sim \mathcal{N}(0, N_i)$ are zero mean white Gaussian noise components with $N_1 \leq N_2 \leq \cdots \leq N_m$. We assume that there is a noiseless causal feedback link from the controller to the sensor and the plant. Let $f_t : \mathbb{R}^{(n+m)t+n} \to \mathbb{R}^m$ denote the sensing policy such that $S_t = f_t(X_{[0,t]}, R_{[0,t-1]})$, where $X_{[0,t]} := \{X_0, X_1, \ldots, X_t\}$ and the sensor is assumed to have an average transmit power constraint $\mathbb{E}[\|S_t\|^2] \leq P_S$. Further, let $\pi_t : \mathbb{R}^{m(t+1)} \to \mathbb{R}^n$ be the controller policy, then we have $U_t = \pi_t\left(R_{[0,t]}\right)$. The common goal of the sensor and the controller is to stabilize the system (2.1) in the mean square sense. The notion of mean-square stability is introduced in Chapter 1, however, in the following we define it for the sake of completeness.

Definition 2.1.1. *A system is said to be mean square stable if there exists $M < \infty$ such that* $\sup_t \mathbb{E}[\|X_t\|^2] < M$.

2.2 Necessary Condition

We first present a necessary condition for *mean-square* stabilization over the point-to-point Gaussian channel depicted in Fig. 2.1.

Theorem 2.2.1. *The linear system in (2.1) can be mean square stabilized over the given parallel Gaussian channel only if*

$$\log\left(|\det(A)|\right) < \frac{1}{2}\sum_{i=1}^{m}\log\left(1+\frac{P_i}{N_i}\right), \tag{2.2}$$

with $P_i = \max\{\gamma - N_i, 0\}$ *and* γ *is chosen such that* $\sum_{i=1}^{m} P_i = P_S$.

Proof. In order to prove Theorem 2.2.1, we make use of the following lemma.

Lemma 2.2.1. *The linear system in* (2.1) *can be mean square stabilized over a channel only if*

$$\log\left(|A|\right) \leq \liminf_{T\to\infty}\frac{1}{T}I\left(S_{[0,T-1]}\to R_{[0,T-1]}\right), \tag{2.3}$$

where $I\left(S_{[0,T-1]}\to R_{[0,T-1]}\right) = \sum_{t=0}^{T-1} I\left(S_{[0,t]}; R_t|R_{[0,t\text{-}1]}\right)$ *denotes the directed information between the sequence of channel input variables* $S_{[0,T-1]}$ *and the sequence of channel output variables* $S_{[0,T-1]}$ *received at the remote controller.*

Proof. The proof is given in Appendix 2.A. □

We can bound the directed information $I\left(S_{[0,T-1]}\to R_{[0,T-1]}\right)$ as

$$\begin{aligned}
&I\left(S_{[0,T-1]}\to R_{[0,T-1]}\right)\\
&\overset{(a)}{\leq} I\left(S_{[0,T-1]}; R_{[0,T-1]}\right)\\
&\overset{(b)}{\leq} \sum_{t=0}^{T-1} I\left(s_{1,t}, s_{2,t},\ldots,s_{m,t}; r_{1,t}, r_{2,t},\ldots,r_{m,t}\right)\\
&= \sum_{t=0}^{T-1}\left[h\left(r_{1,t}, r_{2,t},\ldots,r_{m,t}\right) - h\left(r_{1,t}, r_{2,t},\ldots,r_{m,t}|s_{1,t}, s_{2,t},\ldots,s_{m,t}\right)\right]\\
&= \sum_{t=0}^{T-1}\left[h\left(r_{1,t}, r_{2,t},\ldots,r_{m,t}\right) - h\left(z_{1,t}, z_{2,t},\ldots,z_{m,t}|s_{1,t}, s_{2,t},\ldots,s_{m,t}\right)\right]\\
&= \sum_{t=0}^{T-1}\left[h\left(r_{1,t}, r_{2,t},\ldots,r_{m,t}\right) - h\left(z_{1,t}, z_{2,t},\ldots,z_{m,t}\right)\right]\\
&\overset{(c)}{\leq} \sum_{t=0}^{T-1}\left[\sum_{i=1}^{m} h\left(r_{i,t}\right) - \sum_{i=1}^{m} h\left(z_{i,t}\right)\right]
\end{aligned}$$

$$\stackrel{(d)}{\leq} \sum_{t=0}^{T-1} \left[\sum_{i=1}^{m} \log \left(1 + \frac{P_i}{N_i} \right) \right]$$
$$= \frac{T}{2} \sum_{i=1}^{m} \log \left(1 + \frac{P_i}{N_i} \right), \tag{2.4}$$

where (a) follows from [Mas90, Theorem 1]; (b) follows from the fact that the channels are memoryless and conditioning reduces entropy; (c) follows from conditioning reduces entropy and mutual independence of the noise sequence $\{z_{1,t}, z_{2,t}, \cdots, z_{m,t}\}$; and (d) follows from the fact that the Gaussian distribution maximizes differential entropy for a fixed variance. Now using (2.4) in Lemma 2.2.1 we get the necessary condition given in (2.2). The function $\sum_{i=1}^{L} \log \left(1 + \frac{P_i}{N_i}\right)$ is jointly concave in $\{P_i\}_{i=1}^{m}$, therefore we can solve this optimization problem by the Lagrangian method. The optimal power allocation using the Lagrangian method is given by $P_i = \max\{\gamma - N_i, 0\}$, where γ is chosen such that $\sum_{i=1}^{m} P_i = P_S$, which is the well-known water-filling solution [TV05]. □

We now discuss some sensing and control schemes for stabilization over the given point-to-point Gaussian channel. By employing these schemes, we obtain sufficient conditions for stabilization, which are also presented in the following sections. The corresponding methods for scalar and vector channels are discussed in Sec. 2.3 and Sec. 2.4 respectively.

2.3 Scheme for Scalar Channels

In this section we consider mean square stability of the system in (2.1) over a scalar Gaussian channel, i.e., we assume that $m = 1$ in the system model shown in Fig. 2.1. The state encoder $\mathcal{E}$ observes the n-dimensional state process and transmits it over a one-dimensional Gaussian channel. We restrict our study to the class of encoders that are linear in the observed state with an average transmit power constraint P_S. Therefore at any time t, the signal transmitted by the state encoder is given by $S_t = E_t X_t$, where E_t is an $1 \times n$ row vector. The power constraint at the encoder is given by

$$\mathbb{E}[S_t^2] = E_t \Lambda_t E_t^T \leq P, \tag{2.5}$$

where $\Lambda_t := \mathbb{E}[X_t X_t^T]$. The remotely located controller receives the following signal,

$$R_t = S_t + Z_t, \tag{2.6}$$

where Z_t is an i.i.d. Gaussian variable with zero mean and variance N. The information set available to the controller is $I_t^{\mathcal{C}} = \{R_{[0,t]}, U_{[0,t-1]}\}$. The controller applies an action which is linear in the information set, that is $U_t = m_t I_t^{\mathcal{C}}$. In the following we study mean square stability under the above linear sensing and control scheme.

We have restricted ourselves to linear schemes because they are easy to design and implement. At this point, we highlight some interesting questions that may arise in the reader's mind: i) Is there any loss in restricting sensing and control policies to be linear? ii) Should the policies be time-invariant or time-variant? iii) What is an optimal linear scheme? We try to address these questions in the Sections 2.3.1-2.3.2.

2.3.1 A Linear Time Invariant Scheme

Consider the linear scheme presented above to be time invariant, i.e., at any time t, the encoder output is given by, $S_t = EX_t$. The controller receives $R_t = EX_t + Z_t$, then it runs a Kalman filter to estimate the state and applies the following action $U_t = -A\mathbb{E}[X_t|I_t^{\mathcal{C}}]$, which is optimal for stabilization under the given sensing scheme. Thus the closed-loop system is given by

$$\begin{aligned} X_{t+1} &= A\left(X_t - \mathbb{E}[X_t|I_t^{\mathcal{C}}]\right) + W_t \\ &\overset{(a)}{=} A\left(X_t - \Lambda_t E^T\left[E\Lambda_t E^T + \sigma_z^2\right]^{-1} R_t\right) + W_t \\ &\overset{(b)}{=} A\left(I - \Lambda_t E^T\left[E\Lambda_t E^T + \sigma_z^2\right]^{-1} E\right) X_t + \tilde{Z}_t, \end{aligned} \tag{2.7}$$

where (a) follows from the fact that the control actions whiten the state process and the Gaussian distribution of state process is preserved via linear actions of the encoder and the controller, which results in $\mathbb{E}[X_t|I_t^{\mathcal{C}}] = \mathbb{E}[X_t|R_t] = \mathbb{E}[X_t R_t^T]\mathbb{E}[R_t R_t^T]^{-1} R_t$; and (b) follows by substituting $R_t = EX_t + Z_t$ and summing up all the white Gaussian noise terms and denoting the sum by $\tilde{Z}_t$. The state covariance matrix Λ_t satisfies the following recursion

$$\Lambda_{t+1} = A\Lambda_t A^T - A\Lambda_t E^T\left[E\Lambda_t E^T + \sigma_z^2\right]^{-1} E\Lambda_t A^T + K_W, \tag{2.8}$$

which is the well-known Riccati equation. In [BMF07], the authors studied such a scheme. According to [BMF07], a noiseless plant can be mean square stabilized by any time-invariant encoding matrix E over a Gaussian channel of information capacity[1] C as long as the following two conditions are fulfilled: i) $\log(|\det(A)|) < C$, ii) the pair (A, E) is observable.

We now give a simple example where the LTI scheme fails to stabilize the system. Consider a diagonal system matrix $A = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$ with two equal eigenvalues and let $E = \begin{pmatrix} e_1 & e_2 & e_3 \end{pmatrix}$. The observability matrix $\mathcal{O}$ is then given by

$$\mathcal{O} \triangleq \begin{pmatrix} E \\ EA \\ EA^2 \end{pmatrix} = \begin{pmatrix} e_1 & e_2 & e_3 \\ e_1\lambda_1 & e_2\lambda_2 & e_3\lambda_3 \\ e_1\lambda_1^2 & e_3\lambda_2^2 & e_3\lambda_3^2 \end{pmatrix}. \tag{2.9}$$

[1] The definition of information capacity for Gaussian channels can be found on page 263 in [CT06].

For the pair (A, E) to observable, the observability matrix $\mathcal{O}$ is required to have full rank. In the above example if any two eigenvalues of A are equal, then there can be at most two linearly independent columns in $\mathcal{O}$ and thus $\mathcal{O}$ can never be made of full rank by any choice of E. (One can also use the Hautus–Rosenbrock test for observability.) Therefore, an LTI scheme can never stabilize if two or more eigenvalues of a diagonal system matrix are equal, no matter how much power the encoder is allowed to spend. In Section 2.4 we introduce a linear time varying scheme and show that this scheme can always stabilize the system.

2.3.2 An Optimal Linear Scheme for Stabilization

Consider a linear system with diagonalizable system matrix A and diagonal K_W. For this linear system, we derive an optimal linear time varying sensing policy $E_t^\star$ which minimizes the following cost: $\sum_{i=1}^{t_f} \mathbb{E}\left[\|X_t\|^2\right]$. We have restricted the matrices A and K_W to be diagonal for the ease of analysis. The optimal time varying sensing scheme is presented in the following theorem.

Theorem 2.3.1. *Let* $\tilde{G} := \left[\sqrt{P}, 0, 0, \ldots, 0\right] \in \mathbb{R}^n$, $K_t = A^T\left(I+K_{t+1}\right)A\left(I - \tilde{G}^T\tilde{G}\left(\frac{1}{N+P}\right)\right)$ *with* $K_{t_f} = 0$, *and* π_t *be a unitary matrix such that* $\pi_t^T\left(\Lambda_t^{\frac{1}{2}}A^T\left(I+K_{t+1}\right)A\Lambda_t^{\frac{1}{2}}\right)\pi_t = diag\left(\upsilon_{1,t}, \ldots, \upsilon_{n,t}\right)$ *with* $\upsilon_{1,t} \geq \upsilon_{2,t} \geq \cdots > 0$. *The optimal linear time varying sensing policy is given by:* $E_t^\star = \tilde{G}\pi_t\Lambda_t^{-\frac{1}{2}}$.

Proof. We rewrite the Riccati equation (2.8) as

$$\begin{aligned}
\Lambda_{t+1} &= A\Lambda_t^{\frac{1}{2}}\left(I - \frac{\Lambda_t^{\frac{1}{2}}E_t^T}{\sqrt{N}}\left[\frac{E_t\Lambda_tE_t^T}{\sqrt{N}} + 1\right]^{-1}\frac{E_t\Lambda_t^{\frac{1}{2}}}{\sqrt{N}}\right)\Lambda_t^{\frac{1}{2}}A^T + K_W \\
&\overset{(a)}{=} A\Lambda_t^{\frac{1}{2}}\left(I - C_t^T\left[C_tC_t^T + 1\right]^{-1}C_t\right)\Lambda_t^{\frac{1}{2}}A^T + K_W \\
&\overset{(b)}{=} A\Lambda_t^{\frac{1}{2}}\left[I + C_t^TC_t\right]^{-1}\Lambda_t^{\frac{1}{2}}A^T + K_W, \qquad (2.10)
\end{aligned}$$

where (a) follows from $C_t := \frac{E_t\Lambda_t^{\frac{1}{2}}}{\sqrt{N}}$; and (b) follows from the matrix inversion lemma $[I + UWV]^{-1} = I - U[W^{-1} + VU]^{-1}V$, by choosing $U = C_t^T, W = 1, V = C_t$.

The finite horizon optimal stabilization problem can be stated as:

$$\{C_i^\star\}_{i=0}^{t_f-1} = \arg\min_{\{C_i\}_{i=0}^{t_f-1}: C_tC_t^T \leq \frac{P}{N}} \sum_{t=0}^{t_f-1} \mathrm{Tr}[\Lambda_{t+1}], \qquad (2.11)$$

subject to,

$$\Lambda_{t+1} = A\Lambda_t^{\frac{1}{2}}\left[I + C_t^TC_t\right]^{-1}\Lambda_t^{\frac{1}{2}}A^T + K_W. \qquad (2.12)$$

This is a non-linear dynamic optimization problem. In order to solve this problem we follow a dynamic programming approach. Such an approach has also been considered for continuous time systems in [BB94]. At any time t let the value function be $V_t(\Lambda_t) = \mathrm{Tr}[K_t\Lambda_t + L_t]$. We have to find C_t such that

$$
\begin{aligned}
&\mathrm{Tr}[K_t\Lambda_t + L_t] = \min_{C_t : C_t C_t^T \le \frac{P}{N}} \left\{\mathrm{Tr}\left[\Lambda_{t+1}\right] + \mathrm{Tr}\left[K_{t+1}\Lambda_{t+1} + L_{t+1}\right]\right\}\\
&= \min_{C_t : C_t C_t^T \le \frac{P}{N}} \mathrm{Tr}\left[\left(I + K_{t+1}\right)\Lambda_{t+1} + L_{t+1}\right]\\
&\overset{(a)}{=} \min_{C_t : C_t C_t^T \le \frac{P}{N}} \mathrm{Tr}\left[\left(I+K_{t+1}\right) \times \left(A\Lambda_t^{\frac{1}{2}}\left[I+C_t^T C_t\right]^{-1}\Lambda_t^{\frac{1}{2}}A^T + K_W\right) + L_{t+1}\right]\\
&\overset{(b)}{=} \mathrm{Tr}\left[\left(I+K_{t+1}\right)K_W + L_{t+1}\right] + \min_{C_t : C_t C_t^T \le \frac{P}{N}} \mathrm{Tr}\left[\left(I+K_{t+1}\right)A\Lambda_t^{\frac{1}{2}}\left[I+C_t^T C_t\right]^{-1}\Lambda_t^{\frac{1}{2}}A^T\right]\\
&= \mathrm{Tr}\left[\left(I+K_{t+1}\right)K_W + L_{t+1}\right] + \min_{C_t : C_t C_t^T \le \frac{P}{N}} \mathrm{Tr}\left[\Lambda_t^{\frac{1}{2}}A^T\left(I+K_{t+1}\right)A\Lambda_t^{\frac{1}{2}}\left[I+C_t^T C_t\right]^{-1}\right]\\
&\overset{(c)}{=} \mathrm{Tr}\left[\left(I+K_{t+1}\right)K_W + L_{t+1}\right] + \mathrm{Tr}\left[\Lambda_t^{\frac{1}{2}}A^T\left(I+K_{t+1}\right)A\Lambda_t^{\frac{1}{2}}\left[I+\pi_t G^T G\pi_t^T\right]^{-1}\right]\\
&\overset{(d)}{=} \mathrm{Tr}\left[\left(I+K_{t+1}\right)K_W + L_{t+1}\right] +\\
&\qquad \mathrm{Tr}\left[\Lambda_t^{\frac{1}{2}}A^T\left(I+K_{t+1}\right)A\Lambda_t^{\frac{1}{2}}\left(I-\pi_t G^T\left(1+G\pi_t^T\pi_t G^T\right)^{-1}G\pi_t^T\right)\right]\\
&\overset{(e)}{=} \mathrm{Tr}\left[\left(I+K_{t+1}\right)K_W + L_{t+1}\right] +\\
&\qquad \mathrm{Tr}\left[A^T\left(I+K_{t+1}\right)A\left(\Lambda_t - \Lambda_t^{\frac{1}{2}}\pi_t G^T G\pi_t^T\Lambda_t^{\frac{1}{2}}\left(1+\frac{P}{N}\right)^{-1}\right)\right]\\
&\overset{(f)}{=} \mathrm{Tr}\left[\left(I+K_{t+1}\right)K_W + L_{t+1}\right] + \mathrm{Tr}\left[A^T\left(I+K_{t+1}\right)A\left(I - G^T G\left(\frac{N}{N+P}\right)\right)\Lambda_t\right],
\end{aligned}
\tag{2.13}
$$

where (a) follows by substituting Λ_{t+1} using equation (2.12); (b) follows from the fact that K_{t+1} and L_{t+1} do not depend on C_t; (c) follows from the fact that according to [Bas80] the unique solution to the trace minimization problem given by $C_t^\star = G\pi_t^T$, where $G := \left[\sqrt{\frac{P}{N}}, 0, 0, \ldots, 0\right]$, and π_t is a unitary matrix which diagonalizes $\left(\Lambda_t^{\frac{1}{2}}A^T\left(I+K_{t+1}\right)A\Lambda_t^{\frac{1}{2}}\right)$ such that $\pi_t^T\left(\Lambda_t^{\frac{1}{2}}A^T\left(I+K_{t+1}\right)A\Lambda_t^{\frac{1}{2}}\right)\pi_t = \mathrm{diag}\left(\upsilon_{1,t}, \ldots, \upsilon_{n,t}\right)$ with $\upsilon_{1,t} \ge \upsilon_{2,t} \ge \cdots > 0$; (d) follows from the matrix inversion lemma, $[I + UWV]^{-1} = I - U[W^{-1} + VU]^{-1}V$ by choosing $V = G\pi_t^T, W = 1, U = \pi_t G^T$; (e) follows from $\pi_t\pi_t^T = I$ and $GG^T = \frac{P}{N}$; and (f) follows from the assumption that A and Λ_t are diagonal, which implies that K_{t+1} and π_t are also diagonal. (Diagonality of K_{t+1} will become clear shortly.) Therefore, we have $\Lambda_t^{\frac{1}{2}}\pi_t G^T G\pi_t^T\Lambda_t^{\frac{1}{2}} = \Lambda_t^{\frac{1}{2}}G^T G\Lambda_t^{\frac{1}{2}} = G^T G\Lambda_t$, since $G^T G$ is diagonal. In order to satisfy the above equality (2.13), we choose $K_{t_f} = L_{t_f} = 0$, and $\{K_{t+1}, L_{t+1}\}$ according to

$$K_t = A^T\left(I+K_{t+1}\right)A\left(I - G^T G\left(\frac{N}{N+P}\right)\right),$$
$$L_t = \left(I+K_{t+1}\right)K_W + L_{t+1}. \tag{2.14}$$

We can observe that K_t is also diagonal if A and K_W are diagonal, since $G^T G$ is diagonal. We have found the optimal $C_t^\star = G\pi_t$ and we know that $C_t = \frac{E_t\Lambda_t^{\frac{1}{2}}}{\sqrt{N}}$, therefore, $E_t^\star = \tilde{G}\pi_t\Lambda_t^{-\frac{1}{2}}$ where $\tilde{G} := \sqrt{N}G = \left[\sqrt{P}, 0, 0, \cdots, 0\right]$. □

2.4 A Linear Time Variant Scheme For Vector Channels

In this section we propose a linear time varying scheme for stabilization over vector Gaussian channels and derive sufficient conditions for mean-sqaure stability under the proposed scheme. We also state conditions that guarantee optimality of our proposed scheme.

We first give the scheme for a system with invertible input matrix B, assuming that $B = I$ in (2.1): Consider that the control actions in (2.1) are taken periodically after every K time steps, i.e., at $t = lK - 1$ for $l \in \mathbb{N}$. Under this control strategy, the state equation at $t = lK$ is given by

$$X_{t+K} = A^K X_t + U_{t+K-1} + \sum_{i=0}^{K-1} A^{K-i-1} W_{t+i}. \tag{2.15}$$

For $A^K \in \mathbb{R}^{n\times n}$ there exists a real non-singular matrix T and a real matrix $\tilde{A}$ such that $T^{-1}A^K T = \operatorname{diag}[J_1, \ldots, J_p]$, where J_p is a Jordan block of dimension (algebraic multiplicity) n_p [HJ90]. A Jordan block $J_p \in \mathbb{R}^{n_p\times n_p}$ associated with a real eigenvalue λ of multiplicity n_p has the following form:

$$J_p = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}, \tag{2.16}$$

and a Jordan block $J_p \in \mathbb{R}^{n_p\times n_p}$ associated with a complex conjugate pair of eigenvalues $\lambda = \sigma \pm j\omega$ is given by

$$J_p = \begin{pmatrix} D & I & & \\ & D & \ddots & \\ & & \ddots & I \\ & & & D \end{pmatrix} \tag{2.17}$$

where $D = \begin{pmatrix} \sigma & \omega \\ -\omega & \sigma \end{pmatrix}$. Now apply the linear transformation $\tilde{X}_t = T^{-1}X_t$ such that $T^{-1}A^K T$ is in a real Jordan normal form. Under this transformation, (2.15) is written as

$$\tilde{X}_{t+K} = \tilde{A}\tilde{X}_t + \tilde{U}_{t+K-1} + V_t, \quad \text{for} \quad t = lK, l \in \mathbb{N}, \tag{2.18}$$

where $\tilde{A} := T^{-1}A^K T$, $\tilde{U}_t := T^{-1}U_t$, and $V_t := T^{-1}\sum_{i=0}^{K-1} A^{K-i-1}\bar{W}_{t+i}$. The matrix $\tilde{A}$ is in real Jordan form with eigenvalues $\tilde{\lambda}_i = \lambda_i^K$, where λ_i are the eigenvalues of A.

Now consider that the sensor observes the state vector $\tilde{X}_t := [x_{1,t}, x_{2,t}, \ldots, x_{n,t}]^T$ periodically at $t = lK$. The sensor has access to m parallel Gaussian channels over which it wishes to communicate the state vector X_t to the remote decoder/controller. We propose the following periodic linear transmit strategy. Consider that within each period of K time steps, the encoder linearly transmits different components $x_{i,t}$ of the state vector X_t on different channels such that the following two conditions are satisfied in every time step t:

1. Each channel is used for the transmission of at most one state component, i.e., two components of the state vector are not transmitted over one channel simultaneously.

2. None of the state components is transmitted over more than one channel simultaneously.

Let $k_{ij} \in \mathbb{N}$ be the number of times the j-th channel having information information capacity is used to transmit the state $x_{i,t}$. Under the proposed scheme, we have

$$k_{ij} \geq 0, \quad \sum_{j=1}^{m} k_{ij} \leq K, \quad \sum_{i=1}^{n} k_{ij} \leq K. \tag{2.19}$$

We assume that $x_{i,t}$ is Gaussian distributed due to the following argument: If the initial state $x_{i,0}$ is not Gaussian distributed, then one can perform an initialization step following the idea in [SK66] to make it Gaussian. This initialization step is explained in Appendix 3.B. After this initialization step, the state is always Gaussian distributed since the policies are linear and the noise variables are Gaussian. Therefore, without loss of generality we can assume that $x_{i,t}$ is Gaussian distributed. Let $\hat{x}_{i,t}$ denote the decoder's MMSE estimate of $x_{i,t}$ at the end of each transmission period of K time steps. It is shown in Appendix 2.B that under the proposed linear scheme, the minimum mean-squared error of each state component is given by

$$\mathbb{E}\left[(x_{i,t} - \hat{x}_{i,t})^2\right] = 2^{-2\sum_{j=1}^{m} k_{ij}C_j}\,\mathbb{E}\left[x_{i,t}^2\right]. \tag{2.20}$$

Let us define $\tilde{X}_{est,t} := [\hat{x}_{1,t}, \hat{x}_{2,t}, \ldots, \hat{x}_{n,t}]^T$. The controller then takes the following actions $\tilde{U}_{t+K-1} = -\tilde{A}\tilde{X}_{est,t}$ at $t = lK - 1$. Let $f_{ij} := \frac{k_{ij}}{K}$. It is shown in Sec. 2.4.2

that if we can choose $\{K, k_{ij}\}$ such that $\{f_{ij}\}$ satisfy the conditions given in the following theorem, then the plant is mean-square stable.

Theorem 2.4.1. *The system* (2.1) *can be mean square stabilized over m parallel Gaussian channels using a linear time varying scheme if there exist $f_{ij} \in \mathbb{Q}$ such that $f_{ij} \geq 0$, $\sum_{j=1}^{m} f_{ij} \leq 1$, $\sum_{i=1}^{n} f_{ij} \leq 1$ and*

$$\log(|\lambda_i|) < \sum_{j=1}^{m} f_{ij} C_j, \quad \forall i \in \{1, 2, \ldots, n\}, \tag{2.21}$$

where λ_i are eigenvalues of the system matrix A in (2.1) *and $C_j := \frac{1}{2}\log(1 + \frac{P_j}{N_j})$.*

Proof. The proof is given in Sec. 2.4.2. □

Theorem 2.4.1 holds for a system with a controllable (A, B) pair by the following argument: For a system with controllable (A, B), any input can be realized by n consecutive actions of the controller. Since the state encoder has access to the channel outputs, the encoder-decoder pair can always keep refining estimates of the state components during the K time steps according to the scheme given in the appendix. At the end of K time steps, an estimate $\tilde{X}_{est}$ is available at the controller side. Now the controller wishes to apply an input $-A\tilde{X}_{est}$, which can be realized in following n time steps due to the assumption that the pair (A, B) is controllable. During these n time steps, when the controller is applying actions to realize the input $-A\tilde{X}_{est}$, the encoder-decoder pair will keep on refining the state estimate.

2.4.1 Tightness of the Linear Scheme

Papers [SM11, KGL11] derive conditions for mean-square stabilization of noiseless linear plants over parallel Gaussian channels. The necessary condition in [SM11, Theorem 6] is not tight in general and its achievability is not guaranteed by LTI schemes. The paper [KGL11] proposes a non-linear scheme for a noise-free scalar plant and derives a sufficient condition [KGL11, Theorem 6], which coincides with the necessary condition (2.2). Thus, the non-linear scheme in [KGL11] is optimal for stabilization of noiseless scalar plant. In the following, we present the conditions under which the proposed linear scheme is optimal for mean-square stabilization of noisy multi-dimensional plants over vector Gaussian channels.

Theorem 2.4.2. *The linear scheme is optimal for mean-square stabilizing an n-dimensional plant over m parallel Gaussian channels if there exist $f_{ij} \in \mathbb{Q}$ such that $f_{ij} \geq 0$, $\sum_{j=1}^{m^\star} f_{ij} \leq 1$, $\sum_{i=1}^{n} f_{ij} = 1$ and*

$$\log(|\lambda_i|) < \sum_{j=1}^{m^\star} \frac{f_{ij}}{2} \log\left(1 + \frac{P_j^\star}{N_j}\right), \tag{2.22}$$

for all $i \in \{1, 2, \ldots, n\}$ and $j \in \{1, 2, \ldots, m^\star\}$, where $P_j^\star$ is the optimal power allocation given by the water-filling solution and $m^\star \leq m$ is the number of active channels for which optimal transmit power is non-zero.

Proof. According to the water-filling solution [TV05, pp. 204-205], the optimal power allocation over sub-channels for maximizing information capacity is given by $P_j^\star = \max\{\gamma - N_j, 0\}$, where γ is chosen such that $\sum_{i=j}^{m} P_j^\star = P_S$. Since we have assumed $N_1 \leq N_2 \leq \cdots \leq N_m$, there exists $0 \leq m^\star \leq m$ such that $P_j^\star > 0$ for $j \leq m^\star$ and $P_j^\star = 0$ for $j > m^\star$. Suppose we allocate the powers to the sub-channels according to the water filling solution, i.e., we now have $m^\star$ active parallel channels with capacities $C_j = \frac{1}{2}\log\left(1 + \frac{P_j^\star}{N_j}\right)$ for $1 \leq j \leq m^\star$. According to Theorem 2.4.1, the system in (2.1) is mean-square stable under the linear time varying scheme if there exist $f_{ij} \in \mathbb{Q}$ such that $f_{ij} \geq 0$, $\sum_{j=1}^{m} f_{ij} \leq 1$, $\sum_{i=1}^{n} f_{ij} \leq 1$ and

$$\log\left(|\lambda_i|\right) < \sum_{j=1}^{m^\star} \frac{f_{ij}}{2} \log\left(1 + \frac{P_j^\star}{N_j}\right), \tag{2.23}$$

for all $i \in \{1, 2, \ldots, n\}$ and $j \in \{1, 2, \ldots, m^\star\}$. Summing over $1 \leq i \leq n$, we get

$$\begin{aligned} \sum_{i=1}^{n} \log(|\lambda_i|) &< \sum_{i=1}^{n}\sum_{j=1}^{m} f_{ij} C_{ij} \\ &= \sum_{j=1}^{m} \left(\sum_{i=1}^{n} f_{ij}\right) C_j. \end{aligned} \tag{2.24}$$

Observe that if $\sum_{i=1}^{n} f_{ij} = 1$, then (2.24) can be written as

$$\log\left(|\det\left(A\right)|\right) = \sum_{i=1}^{n} \log(|\lambda_i|) < \frac{1}{2}\sum_{j=1}^{m} \log\left(1 + \frac{P_j^\star}{N_j}\right), \tag{2.25}$$

which is the same condition as (2.2), i.e., the necessary and the sufficient conditions coincide. Thus, the proposed scheme is optimal in the sense that there does not exist any other scheme that can stabilize the plant with a lower transmission power, when there exist coefficients f_{ij} that satisfy the conditions given in Theorem 2.4.2. □

Remark 2.4.1. *A general vector Gaussian channel can be decomposed into an equivalent parallel channel by employing linear pre-processing at the encoder and linear post-processing at the decoder [TV05, pp. 292]. If the equivalent parallel channel satisfies the conditions in Theorem 2.4.2, then the proposed linear scheme is also optimal over the given vector Gaussian channel.*

Remark 2.4.2. *The verification of the existence of coefficients f_{ij} satisfying the conditions given in Theorem 2.4.2 is a linear program and is computationally feasible. In the following we provide some particular instances where linear scheme is* *optimal. In all the following examples, we assume that* $\sum_{i=1}^{n}\log(|\lambda_i|) < \sum_{j=1}^{m} C_j$ *because this is a necessary condition for stabilization.*

1. *If* $n = m$, $\log(\lambda_i) < C_i$, *then we can choose* $f_{ii} = 1$ *and* $f_{ij} = 0$ *for* $j \neq i$.

2. *If* $m = 2$, $n = 3$, *with* $\lambda_1 = \lambda_2 = \lambda_3^l$, $C_2 = lC_1$ *for any* $l \in \mathbb{N}$, *then we can choose,* $f_{11} = f_{21} = 1, f_{31} = 0, f_{12} = f_{22} = 0, f_{32} = 1$.

3. *If* $m = 1$ *(scalar channel), then we can choose* $f_{i1} = \frac{\log(|\lambda_i|)}{\sum_{i=1}^{n}\log(|\lambda_i|)}$ *for all* i.

Remark 2.4.3. *The sufficient condition [KGL11, Theorem 6] for stabilization of a scalar noiseless plant over parallel Gaussian channels with the necessary condition* (2.2)*. Thus the non-linear scheme in [KGL11] is optimal for stabilization of noiseless scalar plant. The non-linear scheme [KGL11] can be combined with the time varying scheme proposed in Sec. 2.4 to get an optimal scheme for stabilization of noiseless multi-dimensional plants. Thus the necessary condition* (2.2) *is also sufficient for stabilization of noiseless multi-dimensional plants.*

2.4.2 Proof of Theorem 2.4.1

Under the proposed linear scheme, we can write (2.18) as

$$\tilde{X}_{t+K} = \tilde{A}\left(\tilde{X}_t - \tilde{X}_{est,t}\right) + V_t, \quad t = lK, l \in \mathbb{N}, \tag{2.26}$$

where $\tilde{A}$ is in real Jordan form with eigenvalues $\tilde{\lambda}_i$. Some of these eigenvalues can be either real or complex and distinct or have algebraic multiplicity. According to (2.26), the state component corresponding to a real distinct eigenvalue $\tilde{\lambda}_i$ is given by

$$x_{i,t+K} = \tilde{\lambda}_i(x_{i,t} - \hat{x}_{i,t}) + v_{i,t}. \tag{2.27}$$

For a complex eigenvalue pair $\tilde{\lambda}_i = \tilde{\lambda}^*_{i+1}$, the state components are given by

$$\begin{aligned} x_{i,t+K} &= \tilde{\sigma}(x_{i,t} - \hat{x}_{i,t}) + \tilde{\omega}(x_{i+1,t} - \hat{x}_{i+1,t}) + v_{i,t}, \\ x_{i+1,t+K} &= -\tilde{\omega}(x_{i,t} - \hat{x}_{i,t}) + \tilde{\sigma}(x_{i+1,t} - \hat{x}_{i+1,t}) + v_{i+1,t}. \end{aligned} \tag{2.28}$$

If $\tilde{\lambda}_r = \tilde{\lambda}_{r+1} = \cdots = \tilde{\lambda}_s = \tilde{\lambda}$ for some $r \leq s$, then according to (2.16) and (2.26) the corresponding states are given by

$$\begin{aligned} x_{i,t+K} &= \tilde{\lambda}(x_{i,t} - \hat{x}_{i,t}) + (x_{i+1,t} - \hat{x}_{i+1,t}) + v_{i,t}, \\ x_{s,t+K} &= \tilde{\lambda}(x_{s,t} - \hat{x}_{s,t}) + v_{s,t}, \end{aligned} \tag{2.29}$$

for $i = r, \ldots, s-1$. Finally, if $\tilde{\lambda}_r = \tilde{\lambda}^*_{r+1} = \tilde{\lambda}_{r+2} = \tilde{\lambda}^*_{r+3} \cdots = \tilde{\lambda}^\star_s = \tilde{\sigma} + j\tilde{\omega}$ for some $r \leq s$, then according to (2.17) and (2.26) the corresponding states are given by

$$\begin{aligned} x_{i,t+K} =& \tilde{\sigma}(x_{i,t} - \hat{x}_{i,t}) + \tilde{\omega}(x_{i+1,t} - \hat{x}_{i+1,t}) + (x_{i+2,t} - \hat{x}_{i+2,t}) \\ &+ v_{i,t}, \quad \text{for} \quad i = r, r+2, \ldots, s-2, \\ x_{i+1,t+K} =& -\tilde{\omega}(x_{i,t} - \hat{x}_{i,t}) + \tilde{\sigma}(x_{i+1,t} - \hat{x}_{i+1,t}) + (x_{i+3,t} - \hat{x}_{i+3,t}) \\ &+ v_{i+1,t}, \quad \text{for} \quad i = r+1, r+3 \ldots, s-1, \\ x_{s-1,t+K} =& \tilde{\sigma}(x_{s\text{-}1,t} - \hat{x}_{s\text{-}1,t}) + \tilde{\omega}(x_{s,t} - \hat{x}_{s,t}) + v_{s\text{-}1,t}, \\ x_{s,t+K} =& -\tilde{\omega}(x_{s\text{-}1,t} - \hat{x}_{s\text{-}1,t}) + \tilde{\sigma}(x_{s,t} - \hat{x}_{s,t}) + v_{s,t}. \end{aligned} \tag{2.30}$$

In the following we find conditions which are sufficient for mean-square stabilization of the state components given by (2.27), (2.28), (2.29), and (2.30), that covers all possible Jordan blocks for the system matrix A. We first consider the state equation (2.27) corresponding to a unique real eigenvalue $\tilde{\lambda}_i$. Let k_{ij} be the number of times the j-th channel is used for the transmission of state $x_{i,t}$ associated with the unique real eigenvalue $\tilde{\lambda}_i$. The second moment of $x_{i,t+K}$ at $t = lK$ is,

$$\begin{aligned} \mathbb{E}\left[x^2_{i,t+K}\right] &= \tilde{\lambda}^2_i \mathbb{E}\left[(x_{i,t} - \hat{x}_{i,t})\right] + \mathbb{E}\left[v^2_{i,t}\right] \\ &\overset{(a)}{=} \tilde{\lambda}^2_i 2^{-2\sum_{j=1}^m k_{ij} C_j} \mathbb{E}\left[x^2_{i,t}\right] + \nu_i, \end{aligned} \tag{2.31}$$

where (a) follows from the (2.20) and $\nu_i := \mathbb{E}\left[v^2_{i,t}\right]$ for $i = 1, 2, \ldots, n$. Since ν_i is bounded, we observe that $\mathbb{E}\left[x^2_{i,t}\right]$ is bounded if $\tilde{\lambda}^2_i 2^{-2\sum_{j=1}^m k_{ij} C_j} < 1$. Thus, the state $x_{i,t}$ is stable if

$$\begin{aligned} &\tilde{\lambda}^2_i 2^{-2\sum_{j=1}^m k_{ij} C_j} < 1 \\ &\Rightarrow \log\left(|\tilde{\lambda}_i|\right) < \sum_{j=1}^m k_{ij} C_j. \end{aligned} \tag{2.32}$$

Next consider the states in (2.28), which are associated with a unique complex eigenvalue pair $\tilde{\lambda}_i, \tilde{\lambda}_{i+1}$. Since $|\tilde{\lambda}_i| = |\tilde{\lambda}_{i+1}|$, we assume that $k_{ij} = k_{i+1j}$, i.e., all the channels are equally used for the transmission of $x_{i,t}$ and $x_{i+1,t}$. The second moments of $x_{i,t+K}$ and $x_{i+1,t+K}$ at $t = lK$ are given by

$$\begin{aligned} \mathbb{E}\left[x^2_{i,t+K}\right] =& 2^{-2\sum_{j=1}^m k_{ij} C_j} \left(\tilde{\sigma}^2 \mathbb{E}\left[x^2_{i,t}\right] + \tilde{\omega}^2 \mathbb{E}\left[x^2_{i+1,t}\right]\right) \\ &+ 2\tilde{\sigma}\tilde{\omega}\mathbb{E}\left[(x_{i,t} - \hat{x}_{i,t})(x_{i+1,t} - \hat{x}_{i+1,t})\right] + \nu_i, \end{aligned} \tag{2.33}$$

$$\begin{aligned} \mathbb{E}\left[x^2_{i+1,t+K}\right] =& 2^{-2\sum_{j=1}^m k_{ij} C_j} \left(\tilde{\omega}^2 \mathbb{E}\left[x^2_{i,t}\right] + \tilde{\sigma}^2 \mathbb{E}\left[x^2_{i+1,t}\right]\right) \\ &- 2\tilde{\sigma}\tilde{\omega}\mathbb{E}\left[(x_{i,t} - \hat{x}_{i,t})(x_{i+1,t} - \hat{x}_{i+1,t})\right] + \nu_{i+1}. \end{aligned} \tag{2.34}$$

By summing (2.33) and (2.34) we get,

$$\begin{aligned} \mathbb{E}\left[x^2_{i,t+K}\right] + \mathbb{E}\left[x^2_{i+1,t+K}\right] =& \left(\tilde{\sigma}^2 + \tilde{\omega}^2\right) 2^{-2\sum_{j=1}^m k_{ij} C_j} \\ &\times \left(\mathbb{E}\left[x^2_{i,t}\right] + \mathbb{E}\left[x^2_{i+1,t}\right]\right) + \nu_i + \nu_{i+1}. \end{aligned} \tag{2.35}$$

The sum $\mathbb{E}\left[x_{i,t+K}^2\right]+\mathbb{E}\left[x_{i+1,t+K}^2\right]$ is bounded if $\left(\tilde{\sigma}^2+\tilde{\omega}^2\right)2^{-2\sum_{j=1}^m k_{ij}C_j}<1$. Since $|\tilde{\lambda}_i|^2=|\tilde{\lambda}_{i+1}|^2=\left(\tilde{\sigma}^2+\tilde{\omega}^2\right)^2$ and $k_{ij}=k_{i+1j}$, we have the following conditions for stabilization:

$$\begin{aligned}\log\left(|\tilde{\lambda}_i|\right)&<\sum_{j=1}^m k_{ij}C_j,\\ \log\left(|\tilde{\lambda}_{i+1}|\right)&<\sum_{j=1}^m k_{i+1j}C_j.\end{aligned}\tag{2.36}$$

Now consider the state components $\{x_{i,t}\}_{i=r}^s$, corresponding to the Jordan block associated with real eigenvalue given in (2.29). Since all the states are equally unstable, we let $k_{rj}=k_{r+1j}=\cdots=k_{sj}=:k_{ij}$. Following the same steps as in (2.31), we can show that $x_{s,t}$ is stable if $\log\left(|\tilde{\lambda}|\right)<\sum_{j=1}^m k_{ij}C_j$. For $i=r,\ldots,s-1$, if we assume that $x_{i+1,t}$ is stable, the second moment of $x_{i,t+K}$ at $t=lK$ can be bounded as

$$\begin{aligned}\mathbb{E}\left[x_{i,t+K}^2\right]&\overset{(a)}{=}\tilde{\lambda}^2\mathbb{E}\left[\left(x_{i,t}-\hat{x}_{i,t}\right)^2\right]+\mathbb{E}\left[\left(x_{i+1,t}-\hat{x}_{i+1,t}\right)^2\right]\\&\quad+2\tilde{\lambda}\mathbb{E}\left[\left(x_{i,t}-\hat{x}_{i,t}\right)\left(x_{i+1,t}-\hat{x}_{i+1,t}\right)\right]+\nu_i\\&\overset{(b)}{=}\tilde{\lambda}^2 2^{-2\sum_{j=1}^m k_{ij}C_j}\mathbb{E}\left[x_{i,t}^2\right]+2^{-2\sum_{j=1}^m k_{ij}C_j}\mathbb{E}\left[x_{i+1,t}^2\right]\\&\quad+2\tilde{\lambda}\mathbb{E}\left[\left(x_{i,t}-\hat{x}_{i,t}\right)\left(x_{i+1,t}-\hat{x}_{i+1,t}\right)\right]+\nu_i\\&\overset{(c)}{\leq}\tilde{\lambda}^2 2^{-2\sum_{j=1}^m k_{ij}C_j}\mathbb{E}\left[x_{i,t}^2\right]+2^{-2\sum_{j=1}^m k_{ij}C_j}\mathbb{E}\left[x_{i+1,t}^2\right]\\&\quad+2\tilde{\lambda}\sqrt{\mathbb{E}\left[\left(x_{i,t}-\hat{x}_{i,t}\right)^2\right]\mathbb{E}\left[\left(x_{i+1,t}-\hat{x}_{i+1,t}\right)^2\right]}+\nu_i\\&=\tilde{\lambda}^2 2^{-2\sum_{j=1}^m k_{ij}C_j}\mathbb{E}\left[x_{i,t}^2\right]+2^{-2\sum_{j=1}^m k_{ij}C_j}\mathbb{E}\left[x_{i+1,t}^2\right]\\&\quad+2\tilde{\lambda}2^{-2\sum_{j=1}^m k_{ij}C_j}\sqrt{\mathbb{E}\left[x_{i,t}^2\right]\mathbb{E}\left[x_{i+1,t}^2\right]}+\nu_i\\&\overset{(d)}{\leq}b_1\mathbb{E}\left[x_{i,t}^2\right]+b_2\sqrt{\mathbb{E}\left[x_{i,t}^2\right]}+b_3,\end{aligned}\tag{2.37}$$

where (a) follows from (2.29); (b) follows from (2.20); (c) follows by Cauchy-Schwarz inequality; (d) follows from the assumption that $x_{i+1,t}$ is stable, i.e., $\mathbb{E}\left[x_{i+1,t}^2\right]<M$ (we have already shown that $x_{s,t}$ is stable if $\log\left(|\tilde{\lambda}|\right)<\sum_{j=1}^m k_{ij}C_j$) and by defining $b_1:=\tilde{\lambda}^2 2^{-2\sum_{j=1}^m k_{ij}C_j}$, $b_2:=2\tilde{\lambda}2^{-2\sum_{j=1}^m k_{ij}C_j}\sqrt{M}$, and $b_3:=2^{-2\sum_{j=1}^m k_{ij}C_j}M+\nu_i$. We now find a condition to ensure convergence of the sequence,

$$\alpha_{t+1}=b_1\alpha_t+b_2\sqrt{\alpha_t}+b_3,\tag{2.38}$$

by making use of the following lemma.

Lemma 2.4.1. *Let $T : \mathbb{R} \mapsto \mathbb{R}$ be a non-decreasing continuous mapping with a unique fixed point $x^\star \in \mathbb{R}$. If there exists $u \leq x^\star \leq v$ such that $T(u) \geq u$ and $T(v) \leq v$, then the sequence generated by $x_{t+1} = T(x_t)$, $t \in \mathbb{N}$ converges starting from any initial value $x_0 \in \mathbb{R}$.*

Proof. The proof is given in Appendix 2.C. □

We observe that the mapping $T(\alpha) = b_1\alpha + b_2\sqrt{\alpha} + b_3$ with $\alpha \geq 0$ is monotonically increasing since $b_1, b_2 > 0$. It will have a unique fixed point $\alpha^\star$ if and only if $b_1 < 1$, since $b_2, b_3 > 0$. Assuming that $b_1 < 1$, there exists $u < \alpha^\star < v$ such that $T(u) \geq u$ and $T(v) \leq v$. Therefore, by Lemma 2.4.1 the sequence $\{\alpha_t\}$ is convergent if $b_1 = \tilde{\lambda}^2 2^{-2\sum_{j=1}^m k_{ij}C_j} < 1 \Rightarrow \log(\tilde{\lambda}) < \sum_{j=1}^m k_{ij}C_j$. Since $|\tilde{\lambda}_i| = \tilde{\lambda}$, $x_{i,t}$ is stable if

$$\log\left(|\tilde{\lambda}_i|\right) < \sum_{j=1}^m k_{ij}C_j. \tag{2.39}$$

Finally, let us consider the states given in (2.29) corresponding to the Jordan block associated with the complex eigenvalues. Since all the state components $\{x_{i,t}\}_{i=r}^{s-1}$ are equally unstable, we fix $k_{rj} = k_{r+1j} = \cdots = k_{sj} = k_{ij}$. Following the same steps as in (2.33), (2.34), and (2.35), we can show that the sum $\mathbb{E}\left[x_{i,t+K}^2\right] + \mathbb{E}\left[x_{i+1,t+K}^2\right]$ is bounded if

$$\begin{aligned} \log\left(|\tilde{\lambda}_{s-1}|\right) &< \sum_{j=1}^m k_{ij}C_j, \\ \log\left(|\tilde{\lambda}_s|\right) &< \sum_{j=1}^m k_{ij}C_j. \end{aligned} \tag{2.40}$$

For $i = r, r+2, r+4, \ldots, s-3$, we have

$$\begin{aligned} \mathbb{E}\left[x_{i,t+K}^2\right] =& 2^{-2\sum_{j=1}^m k_{ij}C_j}\left(\tilde{\sigma}^2\mathbb{E}\left[x_{i,t}^2\right] + \tilde{\omega}^2\mathbb{E}\left[x_{i+1,t}^2\right] + \mathbb{E}\left[x_{i+2,t}^2\right]\right) \\ &+ 2\tilde{\sigma}\tilde{\omega}\mathbb{E}\left[\left(x_{i,t} - \hat{x}_{i,t}\right)\left(x_{i+1,t} - \hat{x}_{i+1,t}\right)\right] \\ &+ 2\tilde{\omega}\mathbb{E}\left[\left(x_{i+1,t} - \hat{x}_{i+1,t}\right)\left(x_{i+2,t} - \hat{x}_{i+2,t}\right)\right] \\ &+ 2\tilde{\sigma}\mathbb{E}\left[\left(x_{i,t} - \hat{x}_{i,t}\right)\left(x_{i+2,t} - \hat{x}_{i+2,t}\right)\right] + \nu_i, \end{aligned} \tag{2.41}$$

$$\begin{aligned} \mathbb{E}\left[x_{i+1,t+K}^2\right] =& 2^{-2\sum_{j=1}^m k_{ij}C_j}\left(\tilde{\omega}^2\mathbb{E}\left[x_{i,t}^2\right] + \tilde{\sigma}^2\mathbb{E}\left[x_{i+1,t}^2\right] + \mathbb{E}\left[x_{i+3,t}^2\right]\right) \\ &- 2\tilde{\sigma}\tilde{\omega}\mathbb{E}\left[\left(x_{i,t} - \hat{x}_{i,t}\right)\left(x_{i+1,t} - \hat{x}_{i+1,t}\right)\right] \\ &+ 2\tilde{\sigma}\mathbb{E}\left[\left(x_{i+1,t} - \hat{x}_{i+1,t}\right)\left(x_{i+3,t} - \hat{x}_{i+3,t}\right)\right] \\ &- 2\tilde{\omega}\mathbb{E}\left[\left(x_{i,t} - \hat{x}_{i,t}\right)\left(x_{i+3,t} - \hat{x}_{i+3,t}\right)\right] + \nu_{i+1}. \end{aligned} \tag{2.42}$$

Assuming that $x_{i+2,t}$ and $x_{i+3,t}$ are stable, the sum $\mathbb{E}\left[x_{i,t+K}^2\right]+\mathbb{E}\left[x_{i+1,t+K}^2\right]$ is bounded by

$$\begin{aligned}
\mathbb{E}\left[x_{i,t+K}^2\right]+\mathbb{E}\left[x_{i+1,t+K}^2\right] &\overset{(a)}{=} \left(\tilde{\sigma}^2+\tilde{\omega}^2\right)2^{-2\sum_{j=1}^{m}k_{ij}C_j}\left(\mathbb{E}\left[x_{i,t}^2\right]+\mathbb{E}\left[x_{i+1,t}^2\right]\right)\\
&\quad+2^{-2\sum_{j=1}^{m}k_{ij}C_j}\left(\mathbb{E}\left[x_{i+2,t}^2\right]+\mathbb{E}\left[x_{i+3,t}^2\right]\right)\\
&\quad+2\tilde{\sigma}\mathbb{E}\left[\left(x_{i,t}-\hat{x}_{i,t}\right)\left(x_{i+2,t}-\hat{x}_{i+2,t}\right)\right]\\
&\quad+2\tilde{\sigma}\mathbb{E}\left[\left(x_{i+1,t}-\hat{x}_{i+1,t}\right)\left(x_{i+3,t}-\hat{x}_{i+3,t}\right)\right]\\
&\quad+2\tilde{\omega}\mathbb{E}\left[\left(x_{i+1,t}-\hat{x}_{i+1,t}\right)\left(x_{i+2,t}-\hat{x}_{i+2,t}\right)\right]\\
&\quad-2\tilde{\omega}\mathbb{E}\left[\left(x_{i,t}-\hat{x}_{i,t}\right)\left(x_{i+3,t}-\hat{x}_{i+3,t}\right)\right]+\nu_i+\nu_{i+1}\\
&\overset{(b)}{\leq} \left(\tilde{\sigma}^2+\tilde{\omega}^2\right)2^{-2\sum_{j=1}^{m}k_{ij}C_j}\left(\mathbb{E}\left[x_{i,t}^2\right]+\mathbb{E}\left[x_{i+1,t}^2\right]\right)\\
&\quad+2^{-2\sum_{j=1}^{m}k_{ij}C_j}\left(\mathbb{E}\left[x_{i+2,t}^2\right]+\mathbb{E}\left[x_{i+3,t}^2\right]\right)\\
&\quad+2\tilde{\sigma}2^{-2\sum_{j=1}^{m}k_{ij}C_j}\sqrt{\mathbb{E}\left[x_{i,t}^2\right]\mathbb{E}\left[x_{i+2,t}^2\right]}\\
&\quad+2\tilde{\sigma}2^{-2\sum_{j=1}^{m}k_{ij}C_j}\sqrt{\mathbb{E}\left[x_{i+1,t}^2\right]\mathbb{E}\left[x_{i+3,t}^2\right]}\\
&\quad+2\tilde{\omega}2^{-2\sum_{j=1}^{m}k_{ij}C_j}\sqrt{\mathbb{E}\left[x_{i+1,t}^2\right]\mathbb{E}\left[x_{i+2,t}^2\right]}\\
&\quad+2\tilde{\omega}2^{-2\sum_{j=1}^{m}k_{ij}C_j}\sqrt{\mathbb{E}\left[x_{i,t}^2\right]\mathbb{E}\left[x_{i+3,t}^2\right]}+\nu_i+\nu_{i+1}\\
&\overset{(c)}{\leq} \left(\tilde{\sigma}^2+\tilde{\omega}^2\right)2^{-2\sum_{j=1}^{m}k_{ij}C_j}\left(\mathbb{E}\left[x_{i,t}^2\right]+\mathbb{E}\left[x_{i+1,t}^2\right]\right)\\
&\quad+2^{-2\sum_{j=1}^{m}k_{ij}C_j}M+2\tilde{\sigma}2^{-2\sum_{j=1}^{m}k_{ij}C_j}M\sqrt{\mathbb{E}\left[x_{i,t}^2\right]}\\
&\quad+2\tilde{\sigma}2^{-2\sum_{j=1}^{m}k_{ij}C_j}M\sqrt{\mathbb{E}\left[x_{i+1,t}^2\right]}\\
&\quad+2\tilde{\omega}2^{-2\sum_{j=1}^{m}k_{ij}C_j}M\sqrt{\mathbb{E}\left[x_{i,t}^2\right]}\\
&\quad+2\tilde{\omega}2^{-2\sum_{j=1}^{m}k_{ij}C_j}M\sqrt{\mathbb{E}\left[x_{i+1,t}^2\right]}+\nu_i+\nu_{i+1}\\
&\overset{(d)}{\leq} b_1\left(\mathbb{E}\left[x_{i,t}^2\right]+\mathbb{E}\left[x_{i+1,t}^2\right]\right)+b_2\sqrt{\mathbb{E}\left[x_{i,t}^2\right]+\mathbb{E}\left[x_{i+1,t}^2\right]}+b_3,
\end{aligned} \tag{2.43}$$

where (a) follows from (2.41) and (2.42); (b) follows from Cauchy-Schwarz inequality; (c) follows from the assumption that $\left(\mathbb{E}\left[x_{i+2,t}^2\right]<+\mathbb{E}\left[x_{i+3,t}^2\right]\right)<M$, and (d) follows by using the following inequality $\sqrt{\mathbb{E}\left[x_{i,t}^2\right]}+\sqrt{\mathbb{E}\left[x_{i+1,t}^2\right]}<2\sqrt{\mathbb{E}\left[x_{i,t}^2\right]+\mathbb{E}\left[x_{i+1,t}^2\right]}$ and defining $b_1:=\left(\tilde{\sigma}^2+\tilde{\omega}^2\right)2^{-2\sum_{j=1}^{m}k_{ij}C_j}$, $b_2:=8\left(\tilde{\sigma}+\tilde{\omega}\right)2^{-2\sum_{j=1}^{m}k_{ij}C_j}M$, and $b_3:=2^{-2\sum_{j=1}^{m}k_{ij}C_j}M+\nu_i+\nu_{i+1}$. If we define $\alpha_t:=\mathbb{E}\left[x_{i,t}^2\right]+\mathbb{E}\left[x_{i+1,t}^2\right]$, then according to (2.43) we get a majorizing sequence that has the same form as (2.38) with the values of b_i given above. Using Lemma

2.4.1 we can show that α_t is convergent if $b_1 = \left(\tilde{\sigma}^2 + \tilde{\omega}^2\right) 2^{-2\sum_{j=1}^{m} k_{ij} C_j} < 1$. Since $\left(\tilde{\sigma}^2 + \tilde{\omega}^2\right) = |\tilde{\lambda}|^2 = |\tilde{\lambda}_i|^2$, we get

$$\log\left(|\tilde{\lambda}_i|\right) < \sum_{j=1}^{m} k_{ij} C_j. \tag{2.44}$$

According to (2.32), (2.36), (2.39), and (2.44), all modes are stable if

$$\log\left(|\tilde{\lambda}_i|\right) < \sum_{j=1}^{m} k_{ij} C_j. \tag{2.45}$$

Since $|\tilde{\lambda}_i| = |\lambda_i|^K$, we can re-write (2.45) as,

$$\begin{aligned} \log\left(|\lambda_i|^K\right) &< \sum_{j=1}^{m} k_{ij} C_j \\ \Rightarrow \log\left(|\lambda_i|\right) &< \sum_{j=1}^{m} f_{ij} C_j, \end{aligned} \tag{2.46}$$

where $f_{ij} := \frac{k_{ij}}{K}$. According to (2.19) $f_{ij} \geq 0$, $\sum_{j=1}^{m} f_{ij} \leq 1$, and $\sum_{i=1}^{n} f_{ij} \leq 1$. This completes the proof. □

2.5 Conclusions

We studied the problem of mean-square stabilization of a *noisy multi-dimensional* linear systems over *vector* Gaussian channels subject to an average transmit power constraint. We observed that linear time invariant schemes are not sufficient in general for stabilization of a multi-dimensional plant even over a scalar point-to-point Gaussian channel. A linear time varying sensing and control scheme is proposed and the conditions which guarantee optimality of the linear scheme are derived. For a given system and a given channel, the optimality of the linear policies can be verified by solving a linear program. We observe that the linear scheme is optimal for a wide class of linear systems and Gaussian channels in the sense that there does not exist any other scheme that can mean-square stabilize the system using a lower transmission power. Interestingly, linear policies are optimal for quadratic stabilization even if the source-channel matching principle does not hold.

Appendix

2.A Proof of Lemma 2.2.1

This proof essentially follows from the same steps as in Theorem 4.1 of [Yük12], however, with some differences due to the network structure. Similar considerations have appeared in different contexts in [MD08,SDO11]. Consider the following series of inequalities:

$$
\begin{aligned}
& I\left(S_{[0,T\text{-}1]} \to R_{[0,T\text{-}1]}\right) \\
& \stackrel{(a)}{=} \sum_{t=0}^{T-1} I\left(S_{[0,t]}; R_t | R_{[0,t\text{-}1]}\right) \\
& \stackrel{(b)}{=} \sum_{t=0}^{T-1} \left(h\left(R_t | R_{[0,t\text{-}1]}\right) - h\left(R_t | S_{[0,t]}, R_{[0,t\text{-}1]}\right)\right) \\
& \stackrel{(c)}{=} \sum_{t=0}^{T-1} \left(h\left(R_t | R_{[0,t\text{-}1]}\right) - h\left(R_t | S_{[0,t]}, R_{[0,t\text{-}1]}, X_t\right)\right) \\
& \stackrel{(d)}{\geq} \sum_{t=0}^{T-1} \left(h\left(R_t | R_{[0,t\text{-}1]}\right) - h\left(R_t | R_{[0,t\text{-}1]}, X_t\right)\right) \\
& \stackrel{(e)}{=} \sum_{t=0}^{T-1} I\left(X_t; R_t | R_{[0,t\text{-}1]}\right) \\
& = I\left(X_0; R_0\right) + \sum_{t=1}^{T-1} I\left(X_t; R_t | R_{[0,t\text{-}1]}\right) \\
& = \sum_{t=1}^{T-1} \left(h\left(X_t | R_{[0,t\text{-}1]}\right) - h\left(X_t | R_{[0,t]}\right)\right) + I\left(X_0; R_0\right) \\
& \stackrel{(f)}{=} \sum_{t=1}^{T-1} \left(h\left(AX_{t-1} + BU_{t-1} + W_{t-1} | R_{[0,t\text{-}1]}\right) - h\left(X_t | R_{[0,t]}\right)\right) + I\left(X_0; R_0\right) \\
& \stackrel{(g)}{=} \sum_{t=1}^{T-1} \left(h\left(AX_{t-1} + W_{t-1} | R_{[0,t\text{-}1]}\right) - h\left(X_t | R_{[0,t]}\right)\right) + I\left(X_0; R_0\right) \\
& \stackrel{(h)}{\geq} \sum_{t=1}^{T-1} \left(h\left(AX_{t-1} + W_{t-1} | R_{[0,t\text{-}1]}, W_{t-1}\right) - h\left(X_t | R_{[0,t]}\right)\right) + I\left(X_0; R_0\right) \\
& \stackrel{(i)}{=} \sum_{t=1}^{T-1} \left(h\left(AX_{t-1} | R_{[0,t\text{-}1]}\right) - h\left(X_t | R_{[0,t]}\right)\right) + I\left(X_0; R_0\right) \\
& \stackrel{(j)}{=} \sum_{t=1}^{T-1} \left(\log\left(\left|\det\left(A\right)\right|\right) + h\left(X_{t-1} | R_{[0,t\text{-}1]}\right) - h\left(X_t | R_{[0,t]}\right)\right) + I\left(X_0; R_0\right) \\
& = (T-1)\log\left(\left|\det\left(A\right)\right|\right) + h\left(X_0 | R_0\right) - h\left(X_{T\text{-}1} | R_{[0,T\text{-}1]}\right) + I\left(X_0; R_0\right) \\
& = h\left(X_0\right) + (T-1)\log\left(\left|\det\left(A\right)\right|\right) - h\left(X_{T\text{-}1} | R_{[0,T\text{-}1]}\right) \\
& \stackrel{(k)}{\geq} h\left(X_0\right) + (T-1)\log\left(\left|\det\left(A\right)\right|\right) - h\left(X_{T\text{-}1}\right)
\end{aligned}
$$

$$\overset{(l)}{\geq} h\left(X_0\right) + (T-1)\log\left(\left|\det\left(A\right)\right|\right) - \log\left(\left(2\pi e\right)^n \left|\det\left(K\right)\right|\right), \tag{2.47}$$

where (a) follows from the definition of directed information; (b) follows from the definition of mutual information; (c) follows from the Markov chain $X_{[0,t]} - \{S_{[0,t]}, R_{[0,t\text{-}1]}\} - R_t$; (d) follows from the fact that conditioning reduces entropy; (e) follows from the property of mutual information; (f) follows from (2.1); (g) follows from $U_{t-1} = \pi_{t-1}(R_{[0,t\text{-}1]})$; (h) follows from the fact that conditioning reduces entropy; (i) follows from $h\left(AX_{t\text{-}1} + W_{t\text{-}1} | R_{[0,t\text{-}1]}, W_{t-1}\right) = h(AX_{t-1}|R_{[0,t\text{-}1]}) + h(W_{t-1}|W_{t-1})$ due to mutual independence of X_t and W_t; (j) follows from $h\left(AX\right) = \log\left(\left|\det\left(A\right)\right|\right) + h(X)$ [CT06, Theorem 8.6.4]; (k) follows from the fact that conditioning reduces entropy; and (m) follows the fact that for a mean square stable system there exists a matrix $K \succ 0$ with $\det\left(\mathbb{E}\left[X_t^T X_t\right]\right) < \det\left(K\right)$ for all t and further for a given covariance matrix the differential entropy is maximized by the Gaussian distribution. Using (2.47) we lower bound the directed information rate as $\liminf_{T\to\infty} \frac{1}{T} I\left(S_{[0,T\text{-}1]} \to R_{[0,T\text{-}1]}\right) \geq \log\left(\left|\det\left(A\right)\right|\right)$.

2.B Estimation With Noiseless Feedback

For simplification we drop the subscripts of $x_{i,t}$ and consider that a variable x has to be transmitted/estimated over m parallel Gaussian channels, where j-th sub-channel has information capacity $C_j = \frac{1}{2}\log\left(1 + P_j/N_j\right)$ for $j = 1, 2, \ldots, m$. Assume that the encoder uses the first channel k_1 times, the second channel k_2 times, and so on. Consider the following scheme based on the Schalkwijk Kailath (SK) coding scheme [SK66]: At $t = 1$, the encoder transmits $s_{1,1} = \sqrt{\frac{P_1}{\mathbb{E}[x^2]}}x$ over the first channel, the decoder receives $r_{1,1} = s_{1,1} + z_{1,1}$, and computes the MMSE estimate $\hat{x}_1 = \frac{\mathbb{E}[xr_{1,1}]}{\mathbb{E}[r_{1,1}^2]} r_{1,1}$, where $\hat{x}_t$ denotes the estimate of x at time t. Further the estimation error at any time t is denoted as $\epsilon_t := x - \hat{x}_t$ and the encoder can compute the error due to noiseless feedback link. By computation, the variance of ϵ_1 is

$$\begin{aligned}\mathbb{E}[\epsilon_1^2] &= \frac{N_1}{P_1 + N_1}\mathbb{E}[x^2] \\ &= 2^{-C_1}\mathbb{E}[x^2],\end{aligned} \tag{2.48}$$

where $C_1 = \frac{1}{2}\log\left(1 + P_1/N_1\right)$. For $2 \leq t \leq k_1$, the encoder transmits $s_{1,t} = \sqrt{\frac{P_1}{\mathbb{E}[\epsilon_{t\text{-}1}^2]}}\epsilon_{t\text{-}1}$, the decoder estimates $\hat{\epsilon}_{t\text{-}1} = \frac{\mathbb{E}[\epsilon_{t\text{-}1} r_{1,t}]}{\mathbb{E}[r_{1,t}^2]} r_{1,t}$ and updates its estimate of x as $\hat{x}_t = \hat{x}_{t\text{-}1} - \hat{\epsilon}_{t\text{-}1}$. The associated estimation error is, $\epsilon_t = x - \hat{x}_t = \hat{\epsilon}_{t\text{-}1} - \epsilon_{t\text{-}1}$. The variance of the estimation error is computed as $\mathbb{E}[\epsilon_t^2] = 2^{-C_1}\mathbb{E}[\epsilon_{t-1}^2]$, which together with (2.48) implies

$$\mathbb{E}[\epsilon_{k_1}^2] = 2^{-k_1 C_1}\mathbb{E}[x^2]. \tag{2.49}$$

Similarly, for the next k_2 time steps the encoder transmits over the second sub-channel having information capacity C_2, and then over the third channel and so on. At the end of transmission over the j-th sub-channel the variance of estimation error is given by

$$\mathbb{E}[\epsilon^2_{k_1+k_2+\ldots+k_j}] = 2^{-k_j C_j}\mathbb{E}[\epsilon^2_{k_{j\text{-}1}}]. \tag{2.50}$$

Accordingly the estimation error at the end of whole transmission period, i.e., at $t = \sum_{j=1}^{m} k_j$, is given by

$$\begin{aligned}\mathbb{E}[(x-\hat{x}_t)^2] &= 2^{-k_m C_m}\mathbb{E}[\epsilon^2_{k_{m\text{-}1}}] \\ &= 2^{-\sum_{j=1}^{m} k_j C_j}\mathbb{E}[x^2].\end{aligned} \tag{2.51}$$

2.C Proof of Lemma 2.4.1

Assume that $T(x)$ is a non-decreasing mapping with a unique fixed point $x^\star$. Further assume that there exist $u \le x^\star \le v$ such that $T(u) \ge u$ and $T(v) \le v$. Consider a sequence generated by the following iterations: $x_{t+1} = T(x_t)$ with $t \in \mathbb{N}$, $x_0 \in \mathbb{R}$. We want to show that starting from any $x_0 \in \mathbb{R}$, the sequence $\{x_t\}$ converges. There are three possibilities: i) $x_0 = x^\star$, ii) $x_0 > x^\star$, and iii) $x_0 < x^\star$. For $x_0 \in [x^\star, \infty)$ we have $T(x) \le x$, therefore $x_1 = T(x_0) \le x_0$. Since $T(x)$ is non-decreasing, $x_2 = T(x_1) \le T(x_0) = x_1$. Thus, for any $t \in \mathbb{N}$ we have $x_{t+1} = T(x_t) \le T(x_{t-1}) = x_t$. Further, this sequence is lower bounded by $x^\star$ because for any $x_t \in [x^\star, \infty)$, $x^\star = T(x^\star) \le T(x_t) = x_{t+1}$ due to non-decreasing $T(\cdot)$. Thus, the sequence $\{x_t\}$ converges since it is monotonically decreasing and lower bounded by $x^\star$ [Rud76, Theorem 3.14]. For $x \in (-\infty, x^\star]$ we have $T(x) \ge x$, therefore, $x_1 = T(x_0) \le x_0$. Since $T(x)$ is non-decreasing, we have $x_2 = T(x_1) \ge T(x_0) = x_1$. Thus, for any $t \in \mathbb{N}$ we have $x_{t+1} = T(x_t) \ge T(x_{t-1}) = x_t$. Further, this sequence is upper bounded by $x^\star$ because for any $x_t \in (-\infty, x^\star]$, we have $x_{t+1} = T(x_t) \le T(x^\star) = x^\star$ due to non-decreasing $T(\cdot)$. Since $\{x_t\}$ is strictly increasing and upper bounded by $x^\star$ for $x_0 \in [x^\star, \infty)$, it converges [Rud76, Theorem 3.14].

Chapter 3

Relay Networks

This chapter considers a setup in which a sensor node communicates the observations of a linear dynamical system (plant) over a network of wireless nodes to a remote controller in order to stabilize the system in closed-loop. The wireless nodes have transmit and receive capability and we call them *relays*, as they relay the plant's state information to the remote controller. We assume a transmit power constraint on the sensor and relays, and the wireless links between all agents (sensor, relays, and controller) are modeled as Gaussian channels. The objective is to study stabilizability of the plant over Gaussian relay networks.

The existing literature on control over Gaussian channels has mostly focused on situations where there is no intermediate node between the sensor and the remote controller. Problems related to control over Gaussian networks with relay nodes have so far been open. Such problems are hard because a relay network can have an arbitrary topology and every node within the network can have memory and can employ any transmit strategy. The papers [Tat03] and [GDH$^+$09] have derived conditions for stabilization over networks with digital noiseless channels and analog erasure channels respectively, however those results do not apply to noisy networks. In [SM06, Yük12] moment stability conditions in terms of error exponents have been established. However, even a single letter expression for channel capacity of the basic three-node Gaussian relay channel [CT06] is not known in general. In [GV05] Gastpar and Vetterli determined capacity of a large Gaussian relay network in the limit as the number of relays tends to infinity. The problem of control over Gaussian relay channels was introduced in [ZOS10a, ZOYS10] and further studied in [ZOYS11, KGL10]. The papers [ZOS10a, ZOYS10, ZOYS11, KGL10] derived sufficient conditions for mean square stability of a scalar plant by employing linear schemes over Gaussian channels with single relay nodes. In this chapter we consider more general setups with multiple relays and multi-dimensional plants. We also derive necessary conditions along with sufficient conditions and further discuss how good linear policies are for various network topologies.

3.1 Problem Formulation

Consider the same discrete LTI plant model as in Chapter 2, given by the following state equation:

$$X_{t+1} = AX_t + BU_t + W_t, \quad t \in \mathbb{N}, \tag{3.1}$$

where $X_t \in \mathbb{R}^n, U_t \in \mathbb{R}^m$, and $W_t \in \mathbb{R}^n$ are state, control, and plant noise The initial state X_0 is a random variable with bounded differential entropy $|h(X_0)| < \infty$ and a given covariance matrix Λ_0. The plant noise $\{W_t\}$ is a zero mean white Gaussian noise sequence with variance K_W and it is assumed to be independent of the initial state X_0. The matrices A and B are of appropriate dimensions and the pair (A, B) is controllable. Let $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ denote the eigenvalues of the system matrix A. Without loss of generality we assume that all the eigenvalues of the system matrix are outside the unit disc, i.e., $|\lambda_i| \geq 1$. The unstable modes can be decoupled from the stable modes by a similarity transformation. If the system in (3.1) is one-dimensional then A is scalar and we use the notation $A = \lambda$. We consider a remote control setup, where a sensor node observes the state process and transmits it to a remotely situated controller over a network of relay[1] nodes as shown in Fig. 4.1. The communication links between nodes are modeled as white Gaussian channels, which is why we refer to it as a Gaussian network. In order to communicate the observed state value X_t, an encoder $\mathcal{E}$ is lumped with the observer $\mathcal{O}$ and a decoder $\mathcal{D}$ is lumped with the controller $\mathcal{C}$. In addition there are L relay nodes $\{\mathcal{R}_i\}_{i=1}^{L}$ within the channel to support communication from $\mathcal{E}$ to $\mathcal{D}$. At any time instant t, $S_{e,t}$ and R_t are the input and the output of the network and U_t is the control action. Let f_t denote the observer/encoder policy such that $S_{e,t} = f_t(X_{[0,t]}, U_{[0,t-1]})$, where $X_{[0,t]} := \{X_0, X_1, \ldots, X_t\}$ and we have the following average transmit power constraint $\mathbb{E}[S_{e,t}^2] \leq P_S$. Further let π_t denote the decoder/controller policy, then $U_t = \pi_t\left(R_{[0,t]}\right)$. The objective in this paper is to find conditions on the system matrix A so that the plant in (3.1) can be mean square stabilized (cf. Definition 2.1.1) over a given Gaussian network.

3.2 Necessary Condition for Stabilization

In the literature [Eli04, MD08, SDO11, Yük12], there exists a variety of information rate inequalities characterizing fundamental limits on the performance of linear systems controlled over communication channels. We have presented one such result in Lemma 2.2.1 for point-to-point channels, which is also valid for the general network depicted in Fig. 3.1. In the following we state another relationship which also gives a necessary condition for mean square stabilization over the general network depicted in Fig. 3.1.

[1] A relay is a communication device whose sole purpose is to support communication from the information source to the destination. In our setup the relay nodes cooperate to communicate the state process from sensor to the remote controller. If the system design objective is to replace wired connections, then relaying is a vital approach to communicate over longer distances.

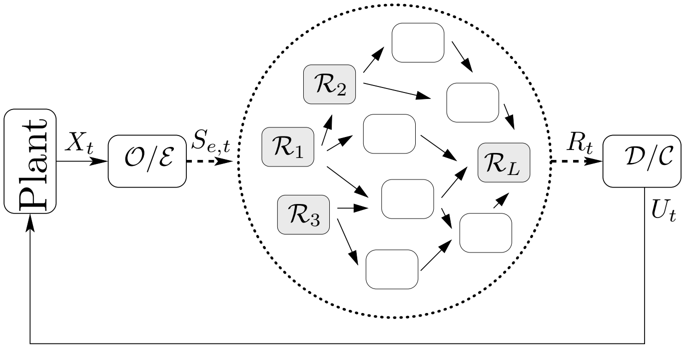


Figure 3.1: The unstable plant has to be controlled over a Gaussian relay network.

Theorem 3.2.1. *If the linear system in* (3.1) *is mean square stable over the Gaussian relay network, then*

$$\log\left(\left|\det\left(A\right)\right|\right) \leq \liminf_{T\to\infty} \frac{1}{T} I(\bar{X}_{[0,T-1]} \to R_{[0,T-1]}), \tag{3.2}$$

where $\{\bar{X}_t\}$ *denotes the uncontrolled state process obtained by substituting* $U_t = 0$ *in* (3.1)*, i.e.,* $\bar{X}_{t+1} = A\bar{X}_t + W_t$*, the notation* $|\det(A)|$ *represents the absolute value of determinant of matrix* A *and*

$$I(\bar{X}_{[0,T-1]} \to R_{[0,T-1]}) = \sum_{t=0}^{T-1} I\left(\bar{X}_{[0,t]}; R_t | R_{[0,t\text{-}1]}\right)$$

is the directed information from the uncontrolled state process $\{\bar{X}_{[0,T-1]}\}$ *to the sequence of variables* $\{R_{[0,T-1]}\}$ *received by the controller over the network of relay nodes.*

Proof. The proof is given in Appendix 3.A, which essentially follows from the same steps as in the proof of Theorem 4.1 in [Yük12], however, with some differences due to the network structure. Similar constructions can also be found in [MD08, SDO11]. □

A general communication network consists of an arbitrary number of nodes with arbitrary communication links. In order to understand the problem of stabilization over a general relay network, we study some basic topologies such as non-orthogonal network, orthogonal network, cascade network, and parallel network (which will be defined later). These topologies serve as the basic building blocks of a large network. In the rest of this chapter, we derive necessary as well as sufficient conditions for mean square stabilization over these fundamental network topologies. Necessary conditions are obtained using Theorem 3.2.1. Sufficient conditions are obtained using delay-free linear sensing and control schemes. We study these fundamental topologies individually so that the proof techniques and the intuitions gained from our study are rich enough to address general Gaussian networks.

3.3 Cascade (Serial) Network

Consider a cascade network comprised of $L-1$ half-duplex relay nodes depicted in Fig. 3.2. The state encoder $\mathcal{E}$ observes the state of the system and transmits its signal to the relay node $\mathcal{R}_1$. The relay node $\mathcal{R}_1$ transmits a signal to the relay node $\mathcal{R}_2$ and so on. Finally the state information is received at the remote decoder/controller $\mathcal{D}$ from $\mathcal{R}_{L-1}$. At any time step t, $S_{e,t}$ is the signal transmitted from $\mathcal{E}$ and $S^i_{r,t}$ is the signal transmitted from $\mathcal{R}_i$, which are given by

$$\begin{aligned} S_{e,t} &= f_t\left(X_{[0,t]}, U_{[0,t-1]}\right) \quad \forall t: t = 1+nL, n \in \mathbb{N}, \qquad S_{e,t} = 0 \quad \text{otherwise}, \\ S^i_{r,t} &= g^i_t\left(Y^i_{[0,t]}\right) \quad \forall t: t = 1+i+nL, n \in \mathbb{N}, \qquad S^i_{r,t} = 0 \quad \text{otherwise}, \end{aligned} \tag{3.3}$$

where $f_t : \mathbb{R}^{2t\text{-}1} \to \mathbb{R}$, $g^i_t : \mathbb{R}^t \to \mathbb{R}$ such that $\mathbb{E}\left[f^2_t\left(X_{[0,t]}, U_{[0,t-1]}\right)\right] = LP_S$, $\mathbb{E}\left[\left(g^i_t\left(Y_{[0,t]}\right)\right)^2\right] = LP^i_r$, $\sum_{i=1}^{L\text{-}1} P^i_r \leq P_R$. The signal received by $\mathcal{R}_i$ is

$$Y^1_t = S_{e,t} + Z^1_t, \quad Y^i_t = S^{i\text{-}1}_{r,t} + Z^i_t \quad \forall t: t = nL+i, n \in \mathbb{N}, \qquad Y^i_t = 0 \quad \text{otherwise}. \tag{3.4}$$

Here $Z^i_t \sim \mathcal{N}(0, N_i)$ denotes mutually independent white Gaussian noise components. Accordingly $\mathcal{D}$ receives $R_t = S^{L\text{-}1}_{r,t} + Z^L_t$ at $t = nL$ and zero otherwise.

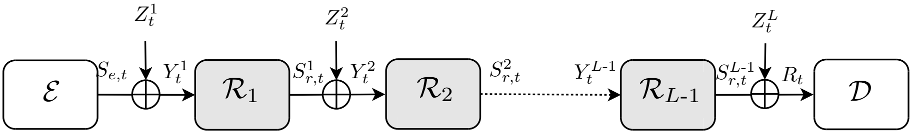


Figure 3.2: A cascade Gaussian network model.

We now present a necessary condition for mean square stability over the given channel.

Theorem 3.3.1. *If the system (3.1) is mean square stable over the* cascade network *then*

$$\log\left(|\det(A)|\right) < \frac{1}{2L}\log\left(1 + L\min\left\{\frac{P_S}{N_1}, \frac{P_R}{\sum_{i=2}^{L} N_i}\right\}\right). \tag{3.5}$$

Proof. We first derive an outer bound on the directed information $I(\bar{X}_{[1,LT]} \to R_{[1,LT]})$ over the given channel and then use Theorem 3.2.1 to find the necessary condition (3.5).

$$
\begin{aligned}
I(\bar{X}_{[1,LT]} \to R_{[1,LT]}) &\overset{(a)}{=} I(\bar{X}_{[1,LT]}; R_{[1,LT]}) \overset{(b)}{\leq} I(\bar{X}_{[1,LT]}; Y^i_{[1,LT]}, R_{[1,LT]}) \\
&= \sum_{t=1}^{LT} I(\bar{X}_{[1,LT]}; R_t, Y^i_t | R_{[1,t\text{-}1]}, Y^i_{[1,t\text{-}1]}) \\
&\overset{(c)}{=} \sum_{t=1}^{LT} \Big(h(R_t, Y^i_t | R_{[1,t-1]}, Y^i_{[1,t\text{-}1]}) \\
&\quad - h(R_t, Y^i_t | R_{[1,t-1]}, Y^i_{[1,t\text{-}1]}, \bar{X}_{[1,LT]}) \Big) \\
&\overset{(d)}{=} \sum_{t=1}^{LT} \Big(h(Y^i_t | R_{[1,t-1]}, Y^i_{[1,t-1]}) + h(R_t | R_{[1,t\text{-}1]}, Y^i_{[1,t]}) \\
&\quad - h(Y^i_t | R_{[1,t\text{-}1]}, Y^i_{[1,t\text{-}1]}, \bar{X}_{[1,LT]}) - h(R_t | R_{[1,t\text{-}1]}, Y^i_{[1,t]}, \bar{X}_{[1,LT]}) \Big) \\
&\overset{(e)}{=} \sum_{t=1}^{LT} \Big(h(Y^i_t | R_{[1,t-1]}, Y^i_{[1,t-1]}) - h(Y^i_t | R_{[1,t\text{-}1]}, Y^i_{[1,t\text{-}1]}, \bar{X}_{[1,LT]}) \\
&\quad + \underbrace{I(R_t; \bar{X}_{[1,LT]} | R_{[1,t\text{-}1]}, Y^i_{[1,t]})}_{=0} \Big) \\
&\overset{(f)}{\leq} \sum_{t=1}^{LT} \Big(h(Y^i_t) - h(Y^i_t | R_{[1,t\text{-}1]}, Y^i_{[1,t\text{-}1]}, \bar{X}_{[1,LT]}) \Big) \\
&\overset{(g)}{\leq} \sum_{t=1}^{LT} \Big(h(Y^i_t) - h(Y^i_t | S^{i\text{-}1}_{r,t}, R_{[1,t\text{-}1]}, Y^i_{[1,t\text{-}1]}, \bar{X}_{[1,LT]}) \Big) \\
&\overset{(h)}{=} \sum_{t=1}^{LT} I(S^{i\text{-}1}_{r,t}; Y^i_t) \\
&\overset{(i)}{=} \sum_{t=0}^{T-1} I(S^{i\text{-}1}_{r,tL+i}; Y^i_{tL+i}) \\
&\overset{(j)}{\leq} \frac{1}{2} \sum_{t=0}^{T-1} \log \left(1 + \frac{LP^{i\text{-}1}_r}{N_i} \right) \\
&= \frac{T}{2} \log \left(1 + \frac{LP^{i\text{-}1}_r}{N_i} \right)
\end{aligned} \tag{3.6}
$$

where (a) follows from [Mas90, Theorem 1]; (b) follows from the fact that adding side information cannot decrease mutual information; (c), (d) and (e) follow from properties of mutual information and differential entropy; (f) follows from conditioning reduces entropy and the following Markov chain $\bar{X}_{[1,LT]} - (Y^i_{[1,t]}, R_{[1,t\text{-}1]}) - R_t$; (g) follow from conditioning reduces entropy; (h) follows from the Markov chain

$Y_t^i - S_{r,t}^{i\text{-}1} - (R_{[1,t\text{-}1]}, Y_{[1,t\text{-}1]}^i, \bar{X}_{[1,LT]})$ due to memoryless channel from $S_{r,t}^{i\text{-}1}$ to Y_t^i; (i) follows from (3.3) and (3.4); and (j) follows from the fact that mutual information of a Gaussian channel is maximized by the Gaussian input distribution [CT06, Theorem 8.6.5]. By following the same steps as (3.6), we can also obtain

$$I(\bar{X}_{[1,LT]} \to R_{[1,LT]}) \leq \frac{T}{2} \log\left(1 + \frac{LP_S}{N_1}\right). \tag{3.7}$$

The directed information $I(\bar{X}_{[1,LT]} \to R_{[1,LT]})$ can also be bounded as

$$\begin{aligned} I\left(\bar{X}_{[1,LT]} \to R_{[1,LT]}\right) &= \sum_{t=1}^{LT} I\left(\bar{X}_{[1,t]}; R_t | R_{[1,t\text{-}1]}\right) \\ &\overset{(a)}{\leq} \sum_{t=1}^{LT} I\left(S_{r,[1,t]}^{L\text{-}1}; R_t | R_{[1,t\text{-}1]}\right) \\ &= I\left(S_{r,[1,LT]}^{L\text{-}1} \to R_{[1,LT]}\right) \\ &\overset{(b)}{\leq} \sum_{t=1}^{LT} I(S_{r,t}^{L\text{-}1}; R_t) \\ &\overset{(c)}{=} \sum_{t=0}^{T-1} I(S_{r,tL+L}^{L\text{-}1}; R_{tL+L}) \\ &\overset{(d)}{\leq} \frac{T}{2} \log\left(1 + \frac{LP_r^{L\text{-}1}}{N_L}\right), \end{aligned} \tag{3.8}$$

where (a) follows from the Markov chain $\bar{X}_{[1,LT]} - (S_{r,[1,t]}^{L\text{-}1}, R_{[1,t\text{-}1]}) - R_{[1,t]}$, (b) follows from [Mas90, Theorem 1]; (c) follows from (3.3) and (3.4); and (d) follows from the fact that mutual information of a Gaussian channel is maximized by the Gaussian input distribution [CT06, Theorem 8.6.5]. Finally using (3.6), (3.7), and (3.8), we have the following bound:

$$\begin{aligned} &I(\bar{X}_{[1,LT]} \to R_{[1,LT]}) \\ &\leq \frac{T}{2} \min\left\{ \log\left(1 + \frac{LP_S}{N_1}\right), \log\left(1 + \frac{LP_r^1}{N_2}\right), \ldots, \log\left(1 + \frac{LP_r^{L\text{-}1}}{N_L}\right) \right\} \\ &\overset{(a)}{=} \frac{T}{2} \log\left(1 + L \min\left\{\frac{P_S}{N_1}, \frac{P_r^1}{N_2}, \ldots, \frac{P_r^{L\text{-}1}}{N_L}\right\}\right) \\ &\leq \frac{T}{2} \log\left(1 + L \min\left\{\frac{P_S}{N_1}, \max_{P_r^i : \sum P_r^i \leq P_R} \min\left\{\frac{P_r^1}{N_2}, \ldots, \frac{P_r^{L\text{-}1}}{N_L}\right\}\right\}\right) \\ &\overset{(b)}{=} \frac{T}{2} \log\left(1 + L \min\left\{\frac{P_S}{N_1}, \frac{P_R}{\sum_{i=2}^{L} N_i}\right\}\right), \end{aligned} \tag{3.9}$$

(a) follows from the fact that $\log(1+x)$ is a monotonically increasing function of x; and (b) follows from the optimal power allocation choice $P_r^i = \frac{N_{i+1}P_R}{\sum_{i=2}^{L} N_i}$. Finally dividing (3.9) by LT and let $T \to \infty$ according to Theorem 3.2.1, we get the necessary condition (3.5). □

We now present a sufficient condition for mean-square stability over the given network.

Theorem 3.3.2. *The scalar linear time invariant system in (3.1) with $A = \lambda$ can be mean square stabilized using a linear scheme over a* cascade network *of L relay nodes if*

$$\log(|\lambda|) < \max_{P_r^i : \sum_{i=1}^{L} P_r^i \le P_R} \frac{1}{2L} \log\left(1 + \frac{LP_S}{LP_S + N_1} \prod_{i=1}^{L-1} \left(\frac{LP_r^i}{LP_r^i + N_{i+1}}\right)\right), \quad (3.10)$$

where the optimal power allocation is given by $P_r^i = \frac{-N_{i+1} + \sqrt{N_{i+1}^2 - \frac{4N_{i+1}}{\gamma}}}{2}$ and $\gamma < 0$ is chosen such that $\sum_{i=1}^{L} P_r^i \le P_R$. When all N_i are equal, the optimal choice is $P_r^i = \frac{P_R}{L-1}$.

Outline of proof: The result can be derived by using a memoryless linear transmission and control scheme. Under linear policies, the overall mapping from the encoder to the controller becomes a scalar Gaussian channel, which has been well studied in the literature (see for example [BB89]). We refer the reader to the proof of Theorem 3.5.2, which contains a detailed derivation for the non-orthogonal network and the proof for this setting is similar. The optimal power allocation follows from the concavity of $\prod_{i=1}^{L-1}\left(\frac{LP_r^i}{LP_r^i + N_{i+1}}\right)$ in $\{P_r^i\}_{i=1}^{L-1}$ and by using the Lagrange multiplier method.

Remark 3.3.1. *For fixed power allocations, as the number of relays L approaches infinity in* (3.5)*, the RHS converges to zero and stabilization becomes impossible. We also note that the ratio between the sufficiency and necessity bounds (ratio between RHS of* (3.3.2) *and RHS of* (3.5)*) converges to zero as the number of relays goes to infinity.*

In the related problem of transmitting a Gaussian source with minimum mean-square distortion, it is shown in Chapter 4 that linear sensing policies are not globally optimal in general when there is one or more relay nodes in cascade. However linear policies are shown to be person-by-person optimal in a single relay setup. According to Chapter 4, simple quantizer based policies can lead to a lower mean-square distortion than the best linear policy.

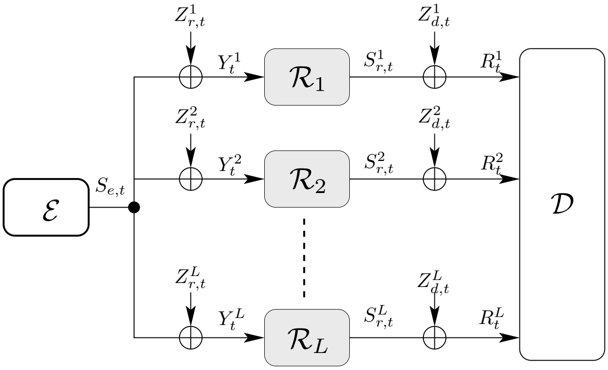


Figure 3.3: Parallel relay network.

3.4 Parallel Network

Consider the network shown in Fig. 3.3, where the signal transmitted by a node does not interfere with the signals transmitted by other nodes, i.e., there are L parallel channels from $\{\mathcal{R}_i\}_{i=1}^L$ to $\mathcal{D}$. We call this setup a *parallel network*, which models a scenario where the signal spaces of the relay nodes are mutually orthogonal. For example the signals may be transmitted in either disjoint frequency bands or in disjoint time slots. In the first transmission phase, the sensor transmits $S_{e,t}$ with an average power $\mathbb{E}\left[S_{e,t}^2\right] = 2P_S$ to the relays and in the second phase all relays simultaneously transmit to the remote controller with average powers $2P_r^i$ such that $\sum_{i=1}^L P_r^i \leq P_R$. Accordingly, the received signals are given by

$$\begin{aligned} Y_t^i &= S_{e,t} + Z_{r,t}^i, \quad R_t^i = S_{r,t}^i = 0, \quad t = 1,3,5,\dots \\ R_t^i &= S_{r,t}^i + Z_{d,t}^i, \quad Y_t^i = S_{e,t} = 0, \quad t = 2,4,6,\dots \end{aligned} \tag{3.11}$$

where $Z_{r,t}^i \sim \mathcal{N}(0, N_r^i)$, $Z_{d,t}^i \sim \mathcal{N}(0, N_d^i)$ denote mutually independent white Gaussian noise variables. In the following we present conditions for mean square stability of the system in (3.1) over the given parallel network.

Theorem 3.4.1. *If the system (3.1) is mean square stable over the* parallel network *then*

$$\log\left(|\det(A)|\right) \leq \frac{1}{4}\min\left\{\log\left(1 + 2\sum_{i=1}^L \frac{P_S}{N_r^i}\right), \sum_{i=1}^L \log\left(1 + \frac{2P_r^i}{N_d^i}\right)\right\}, \tag{3.12}$$

where $P_r^i = \max\{\gamma - N_d^i, 0\}$ *and* γ *is chosen such that* $\sum_{i=1}^L P_r^i = P_R$.

Proof. Following the same steps as in proof of Theorem 3.3.1, we can bound directed information $I(\bar{X}_{[1,2T]} \to R_{[1,2T]})$ over *parallel* relay network as,

$$\begin{aligned}
&I(\bar{X}_{[1,2T]} \to \{R^i_{[1,2T]}\}_{i=1}^L) \\
&\overset{(a)}{\leq} \min\left\{\sum_{t=1}^{2T} I\left(S_{e,t}; \{Y_t^i\}_{i=1}^L\right), \sum_{t=1}^{2T} I\left(\{S^i_{r,t}\}_{i=1}^L; \{R_t^i\}_{i=1}^L\right)\right\} \\
&\overset{(b)}{=} \min\left\{\sum_{t=1}^{T} I\left(S_{e,2t\text{-}1}; \{Y^i_{2t\text{-}1}\}_{i=1}^L\right), \sum_{t=1}^{T} I\left(\{S^i_{r,2t}\}_{i=1}^L; \{R^i_{2t}\}_{i=1}^L\right)\right\} \\
&\overset{(c)}{\leq} \frac{T}{2}\min\left\{\log\left(1+2\sum_{i=1}^L \frac{P_S}{N_r^i}\right), \max_{P_r^i: P_r \geq 0, \sum_i P_r^i \leq P_R} \sum_{i=1}^L \log\left(1+\frac{2P_r^i}{N_d^i}\right)\right\}, \qquad (3.13)
\end{aligned}$$

where (a) follows from the same steps as in (3.6) and (3.8); (b) follows from (3.11); and (c) follows from the fact that Gaussian input distribution maximizes mutual information for a Gaussian channel. The function $\sum_{i=1}^L \log\left(1+\frac{2P_r^i}{N_d^i}\right)$ is jointly concave in $\{P_r^i\}_{i=1}^L$. The optimal power allocation is given by $P_r^i = \max\{\gamma - N_d^i/2, 0\}$, where γ is chosen such that $\sum_{i=1}^L P_r^i = P_R$, which is the well-known water-filling solution [TV05, pp. 204-205]. We obtain (3.12) by using (3.13) in Theorem 3.2.1. ☐

We can obtain a sufficient condition for mean square stability over the *parallel network* using linear policies like previously discussed scenarios, which is stated in the following theorem.

Theorem 3.4.2. *The scalar linear time invariant system in (3.1) with $A = \lambda$ can be mean square stabilized using a linear scheme over the Gaussian* parallel network *if*

$$\log(|\lambda|) < \frac{1}{4}\log\left(1+\sum_{i=1}^L \frac{4P_S P_r^i}{2P_S N_d + 2P_r^i N_r^i + N_d N_r^i}\right). \qquad (3.14)$$

Proof. The above result can be obtained by using a memoryless linear sensing and control scheme and as discussed in the proof of Theorem 3.3.2. ☐

Proposition 3.4.1. *The gap between the necessary and sufficient conditions for a symmetric parallel network with $P_r^i = P_r, N_r^i = N_r$ is a non-decreasing function of the number of relays L and approaches $\frac{1}{4}\log\left(1+\frac{N_d(2P_S+N_r)}{2P_r N_r}\right)$ as L goes to infinity.*

Proof. For $P_r^i = P_r, N_r^i = N_r$, the R.H.S. of (3.14) is evaluated as $\Gamma_{\text{suf}} := \frac{1}{4}\log\left(1+\frac{4LP_SP_r}{2P_SN_d+2P_rN_r+N_dN_r}\right)$ and the R.H.S of (3.12) can be bounded as $\Gamma_{\text{nec}} := \frac{1}{4}\log\left(1+\frac{2LP_S}{N_r}\right)$. The gap is given by

$$\Gamma_{\text{nec}} - \Gamma_{\text{suf}} = \frac{1}{4}\log\left(1+\frac{2P_SN_d(2P_S+N_r)}{4P_SP_rN_r + \frac{N_r(2P_SN_d+2P_rN_r+N_dN_r)}{L}}\right), \qquad (3.15)$$

which is a non-decreasing function of L, approaching $\frac{1}{4}\log\left(1+\frac{N_d(2P_S+N_r)}{2P_rN_r}\right)$ as $L\to\infty$. □

Remark 3.4.1. *If $N_d^i=0$, then $\Gamma_{nec}-\Gamma_{suf}=0$ and the linear scheme is exactly optimal. For $N_r^i=0$, $\Gamma_{suf}:=\frac{1}{4}\log\left(1+\frac{2LP_r}{N_d}\right)$ and $\Gamma_{nec}:=\frac{L}{4}\log\left(1+\frac{2P_r}{N_d}\right)$ according to* (3.12). *Clearly $\lim_{L\to\infty}(\Gamma_{nec}-\Gamma_{suf})=\infty$, showing the inefficiency of the LTI scheme for parallel channels.*

It is known that linear schemes can be sub-optimal for transmission over parallel channels [Vai89, WS09b]. A distributed joint source–channel code is optimal in minimizing mean-square distortion if the following two conditions hold [SVZ98]: i) All channels from the source to the destination send independent information; ii) All channels utilize the capacity, i.e., the source and channel need to be matched. If we use linear policies at the relay nodes then the first condition is not fulfilled because all nodes would be transmitting correlated information. In [YT09] the authors proposed a non-linear scheme for a parallel network of two sensors without relays, in which one sensor transmits only the magnitude of the observed state and the other sensor transmits only the phase of the observed state. The magnitude and phase of the state are shown to be independent and thus the scheme fulfills the first condition of optimality. This nonlinear sensing scheme is shown to outperform the best linear scheme for the LQG control problem in the absence of measurement noise, although the second condition of source-channel matching is not fulfilled. We can use this non-linear scheme together with the initialization step of the Schalkwijk Kailath (SK) type scheme described in Appendix 3.C for the non-orthogonal network, which will ensure source-channel matching by making the outputs of the two sensors Gaussian distributed after the initial transmissions. In Chapter 5 it is shown that linear sensing policies may not be even person-by-person optimal for LQG control over parallel network without relays.

For the special case of *parallel network* with noiseless $\mathcal{E}-\mathcal{R}_i$ links, we have the following necessary and sufficient condition for mean-square stability.

Theorem 3.4.3. *The system (3.1) in absence of process noise ($W_t=0$) can be mean square stabilized over the Gaussian* parallel network *with $Z_{r,t}^i=0$ for all i, only if*

$$\log(|\det(A)|)\le\frac{1}{4}\max_{P_r^i:P_r\ge 0,\sum_i P_r^i\le P_R}\sum_{i=1}^{L}\log\left(1+\frac{2P_r^i}{N_d^i}\right). \tag{3.16}$$

If the inequality is strict, then there exists a non-linear policy leading to mean-square stability.

Proof. The necessity follows from Theorem 3.4.1. The sufficiency part for scalar systems follows from [KGL11, Theorem 6], which is derived using a non-linear scheme. This scheme can be extended to vector systems using a time sharing scheme presented in Sec. 2.4. □

Remark 3.4.2. *The term on RHS of* (3.4.3) *is equal to information capacity of the parallel Gaussian channel. It was shown by Shu and Middleton in [SM11] that for some first order noiseless plants, linear time invariant encoders/decoders cannot achieve information capacity of parallel Gaussian channels. However information capacity required for stabilization can always be achieved by a non-linear time varying scheme as discussed in Remark 2.4.3.*

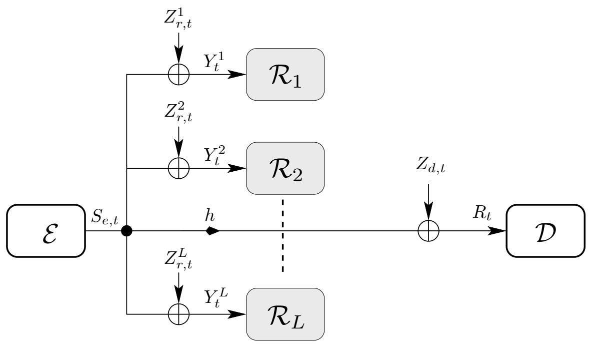
3.5 Non-orthogonal Network

A communication network is said to be *non-orthogonal* if all the communicating nodes transmit signals in overlapping time slots using the same frequency bands. The networks considered in the previous sections are examples of *orthogonal* networks. In this section, we will focus on non-orthogonal networks. A node which is capable of transmitting and receiving signals simultaneously using the same frequency band is known as *full-duplex* while a *half-duplex* node cannot simultaneously receive and transmit signals. In practice it is expensive and hard to a build a communication device which can transmit and receive signals at the same time using the same frequency, due to the self-interference created by the transmitted signal to the received signal. Therefore half-duplex systems are mostly used in practice. In this section we study both half-duplex and full-duplex configurations.

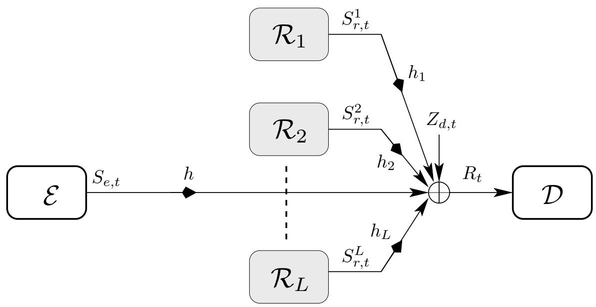
3.5.1 Non-orthogonal Half-duplex Network

A non-orthogonal half-duplex Gaussian network with L relay nodes $\{\mathcal{R}_i\}_{i=1}^L$ is illustrated in Fig. 3.4. The variables $S_{e,t}$ and $S_{r,t}^i$ denote the transmitted signals from the state encoder $\mathcal{E}$ and relay $\mathcal{R}_i$ at any discrete time step t. The variables $Z_{r,t}^i$ and $Z_{d,t}$ denote the mutually independent white Gaussian noise components at the relay node i and $\mathcal{D}$ of the remote control unit, with $Z_{r,t}^i \sim \mathcal{N}(0, N_r^i)$ and $Z_{d,t} \sim \mathcal{N}(0, N_d)$. The noise components $\{Z_{r,t}^i\}_{i=1}^L$ are independent across the relays, i.e., $\mathbb{E}[Z_{r,t}^k Z_{r,t}^i] = 0$ for all $i \neq k$. The information transmission from the state encoder consists of two phases as shown in Fig. 3.4. In the first phase the encoder $\mathcal{E}$ transmits a signal with an average power $2\beta P_S$, where $\beta \in (0, 1]$ is a parameter that adjusts power between the two transmission phases. In this transmission phase all the relay nodes listen but remain silent. In the second transmission phase, the encoder $\mathcal{E}$ and relay nodes $\{\mathcal{R}_i\}_{i=1}^L$ transmit simultaneously. In this second transmission phase, the encoder transmits with an average power $2(1-\beta)P_S$ and the i-th relay node transmits with an average power $2P_r^i$ such that $\sum_{i=1}^L P_r^i \leq P_R$. The input and output of the i-th relay are given by,

$$\begin{aligned} &Y_t^i = S_{e,t} + Z_{r,t}^i, \quad S_{r,t}^i = 0, && t = 1, 3, 5, \dots \\ &Y_t^i = 0, \quad S_{r,t}^i = g_t^i\left(Y_{[0,t-1]}^i\right), && t = 2, 4, 6, \dots \end{aligned} \tag{3.17}$$



(a) First transmission phase.



(b) Second transmission phase.

Figure 3.4: A non-orthogonal half-duplex Gaussian network model.

where $g^i_t : \mathbb{R}^{t+1} \to \mathbb{R}$ is the i-th relay encoding policy such that $\mathbb{E}\left[\left(g^i_t\left(Y^i_{[0,t-1]}\right)\right)^2\right] = 2P^i_r$ and $\sum_{i=1}^{L} P^i_r \leq P_R$. The signal received at the decoder/controller is given by

$$R_t = hS_{e,t} + \sum_{i=1}^{L} h_i S^i_{r,t} + Z_{d,t},$$

where $h, h_i \in \mathbb{R}$ denote the channel gains of $\mathcal{E} - \mathcal{D}$ and $\mathcal{R}_i - \mathcal{D}$ links respectively.

We now present a necessary condition mean-square stabilization over *non-orthogonal half-duplex* relay network.

Theorem 3.5.1. *If the linear system in* (3.1) *is mean-square stable over the* non-orthogonal half-duplex *relay network, then*

$$\log\left(|\det\left(A\right)|\right) \leq \frac{1}{4}\min\left\{\max_{0<\beta\leq 1}\left(\log\left(1+\frac{2h^2(1-\beta)P_S}{N_d}\right)\right.\right.$$

$$+\log\left(1+2\beta P_S\left(\sum_{i=1}^{L}\frac{1}{N_r^i}+\frac{h^2}{N_d}\right)\right)\Bigg), \max_{\substack{0<\beta\leq 1\\ P_r^i:\sum_i P_r^i\leq P_R}}\Bigg(\log\left(1+\frac{2h^2\beta P_S}{N_d}\right)$$

$$+\log\left(1+\frac{1}{N_d}\left(\sum_{i=1}^{L+1}\delta_i^2 P_i+2\sum_{i=1}^{L+1}\sum_{k=i+1}^{L+1}\rho^\star_{i,k}\delta_i\delta_k\sqrt{P_iP_k}\right)\right)\Bigg)\Bigg\}, \quad (3.18)$$

where $\rho^\star_{i,k} := \frac{2(1-\beta)P_S}{\sqrt{(2(1-\beta)P_S+N_i)(2(1-\beta)P_S+N_k)}}$, $P_{L+1} := 2(1-\beta)P_S$, $N_{L+1} := 0$, $\delta_{L+1} := h$, $P_i := 2P_r^i$, $\delta_i := h_i$, $N_i := N_r^i$ *for all* $i = \{1, 2, \ldots, L\}$.

Proof. We first derive an outer bound on the directed information $I(\bar{X}_{[1,LT]} \to R_{[1,LT]})$ over the given channel and then use Theorem 3.2.1 to find the necessary condition (3.18).

$$\begin{aligned}
&I(\bar{X}_{[1,2T]} \to R_{[1,2T]})\\
&\overset{(a)}{=} I\left(\bar{X}_{[1,2T]}; R_{[1,2T]}\right)\\
&\overset{(b)}{\leq} I(\bar{X}_{[1,2T]}; \{Y^i_{[1,2T]}\}_{i=1}^L, R_{[1,2T]})\\
&\overset{(c)}{=} I\left(\bar{X}_{[1,2T]}; \tilde{R}_{[1,2T]}, \{Y^i_{[1,2T]}\}_{i=1}^L\right)\\
&= \sum_{t=1}^{2T} I(\bar{X}_{[1,2T]}; \tilde{R}_t, \{Y^i_t\}_{i=1}^L | \tilde{R}_{[1,t\text{-}1]}, \{Y^i_{[1,t\text{-}1]}\}_{i=1}^L)\\
&\overset{(d)}{\leq} \sum_{t=1}^{2T} I(S_{e,t}; \tilde{R}_t, \{Y^i_t\}_{i=1}^L | \tilde{R}_{[1,t\text{-}1]}, \{Y^i_{[1,t\text{-}1]}\}_{i=1}^L)\\
&\overset{(e)}{\leq} \sum_{t=1}^{2T} I\left(S_{e,t}; \tilde{R}_t, \{Y^i_t\}_{i=1}^L\right)\\
&\overset{(f)}{=} \sum_{t=1}^{T} I\left(S_{e,2t}; \tilde{R}_{2t}\right) + \sum_{t=1}^{T} I\left(S_{e,2t\text{-}1}; \tilde{R}_{2t\text{-}1}, \{Y^i_{2t\text{-}1}\}_{i=1}^L\right)\\
&\overset{(g)}{\leq} \frac{T}{2}\log\left(1+\frac{2h^2(1-\beta)P_S}{N_d}\right) + \frac{T}{2}\log\left(1+2\beta P_S\left(\sum_{i=1}^{L}\frac{1}{N_r^i}+\frac{h^2}{N_d}\right)\right)\\
&\leq \frac{T}{2}\max_{0<\beta\leq 1}\left\{\log\left(1+\frac{2h^2(1-\beta)P_S}{N_d}\right)+\log\left(1+2\beta P_S\left(\sum_{i=1}^{L}\frac{1}{N_r^i}+\frac{h^2}{N_d}\right)\right)\right\},
\end{aligned}\quad (3.19)$$

where (a) follows from [Mas90, Theorem 1]; (b) follows from the fact that adding side information cannot decrease mutual information; (c) follows by defining $\tilde{R}_t := R_t - \sum_{i=1}^{L} h_i S^i_{r,t}$ and from the fact that $S^i_{r,t}$ is a function of $Y^i_{[1,t\text{-}1]}$; (d) follows from the Markov chain $\bar{X}_{[1,2T]} - S_{e,t} - (\tilde{R}_t, \{Y^i_t\}_{i=1}^L)$, since $\bar{X}_{[0,T]}$ is the uncontrolled state process and the fact that the channel between $S_{e,[1,2T]}$ and $(\tilde{R}_{[1,2T]}, \{Y^i_{[1,2T]}\}_{i=1}^L)$ is memoryless due to $\tilde{R}_t = R_t - \sum_{i=1}^L h_i S^i_{r,t}$; (e) follows from the Markov chain $(\tilde{R}_{[1,t\text{-}1]}, \{Y^i_{[1,t\text{-}1]}\}_{i=1}^L) - S_{e,t} - (\tilde{R}_t, \{Y^i_t\}_{i=1}^L)$ and conditioning reduces entropy; (f) follows by separating odd and even indexed terms and $Y^i_{2t} = 0$ according to (3.17); (g) follows from $Y^i_{2t\text{-}1} = S_{e,2t\text{-}1} + Z^i_{r,2t\text{-}1}$, $\tilde{R}_t = S_{e,t} + Z_t$, $\mathbb{E}\left[S^2_{e,2t}\right] = 2(1-\beta)P_S$, $\mathbb{E}\left[S^2_{e,2t\text{-}1}\right] = 2\beta P_S$, and the fact that mutual information of a Gaussian channel is maximized by centered Gaussian input distribution [TV05]. The directed information rate $I(\bar{X}_{[1,2T]} \to R_{[1,2T]})$ can also be bounded as,

$$
\begin{aligned}
& I(\bar{X}_{[1,2T]} \to R_{[1,2T]}) \\
& = \sum_{t=1}^{2T} I(\bar{X}_{[1,t]}; R_t | R_{[1,t\text{-}1]}) \\
& \overset{(a)}{\leq} \sum_{t=1}^{2T} I(S_{e,t}, \{S^i_{r,t}\}_{i=1}^L; R_t | R_{[1,t\text{-}1]}) \\
& \overset{(b)}{\leq} \sum_{t=1}^{2T} I\left(S_{e,t}, \{S^i_{r,t}\}_{i=1}^L; R_t\right) \\
& \overset{(c)}{=} \sum_{t=1}^{T} \left(I\left(S_{e,2t\text{-}1}; R_{2t\text{-}1}\right) + I\left(S_{e,2t}, \{S^i_{r,2t}\}_{i=1}^L; R_{2t}\right)\right) \\
& \overset{(d)}{\leq} \frac{T}{2}\log\left(1 + \frac{2h^2\beta P_S}{N_d}\right) + \frac{T}{2}\log\left(1 + \frac{1}{N_d}\left(\sum_{i=1}^{L+1}\delta_i^2 P_i + 2\sum_{i=1}^{L+1}\sum_{k=i+1}^{L+1}\rho^\star_{i,k}\delta_i\delta_k\sqrt{P_i P_k}\right)\right) \\
& \leq \frac{T}{2} \max_{\substack{0<\beta\leq 1 \\ P^i_r : \sum_i P^i_r \leq P_R}} \left\{ \log\left(1 + \frac{2h^2\beta P_S}{N_d}\right)\right. \\
& \qquad \left. + \log\left(1 + \frac{1}{N_d}\left(\sum_{i=1}^{L+1}\delta_i^2 P_i + 2\sum_{i=1}^{L+1}\sum_{k=i+1}^{L+1}\rho^\star_{i,k}\delta_i\delta_k\sqrt{P_i P_k}\right)\right)\right\}, \qquad (3.20)
\end{aligned}
$$

where $\rho^\star_{i,k} = \frac{2(1-\beta)P_S}{\sqrt{(2(1-\beta)P_S + N_i)(2(1-\beta)P_S + N_k)}}$, $P_{L+1} = 2(1-\beta)P_S$, $N_{L+1} = 0$, $\delta_{L+1} = h$, $P_i = 2P^i_r$, $\delta_i = h_i$, $N_i = N^i_r$ for all $i = \{1, 2, \ldots, L\}$. The inequality (a) follows from the Markov chain $\bar{X}_{[0,t]} - \left(S_{e,t}, \{S^i_{r,t}\}_{i=1}^L\right) - R_t$ due to the memoryless channel between $S_{e,[1,2T]}, \{S^i_{r,[1,2T]}\}_{i=1}^L$ and $R_{[1,2T]}$; (b) follows from the Markov chain $R_{[1,t\text{-}1]} - \left(S_{e,t}, \{S^i_{r,t}\}_{i=1}^L\right) - R_t$ and conditioning reduces entropy; (c) follows by separating the odd and even indexed terms and $S^i_{r,2t\text{-}1} = 0$ according to (3.17); (d) follows from the fact that the first addend on the R.H.S. of (c) is maximized by a centered Gaussian distributed $S_{e,t}$ and the second addend is bounded using a bound presented in [Gas07], where the author studied the problem of transmitting a Gaussian source over a simple sensor network. In order to apply the upper bound given in (48) of [Gas07] to our setup, we consider state encoder $\mathcal{E}$ to be a sensor node with zero observation noise and make the following change of system variables so that our system model becomes equivalent to the one discussed in [Gas07]: $\sigma_S^2 := \alpha_t$, $\delta_i := h_i$, $M := L+1$, $P_i := 2P_r^i$, $\sigma_Z^2 := N_d$, $\sigma_{W,i}^2 := N_r^i$, $\alpha_i = \sqrt{\frac{2(1-\beta)P_S}{\alpha_t}}$ for all i. We finally obtain (3.18) by dividing (3.19) and (3.20) by $2T$ and let $T \to \infty$ according to Theorem 3.2.1. □

We now present a sufficient condition for mean square stability of a scalar plant over the given network, which can be extended to a multi-dimensional plant using the linear time variant scheme given in Sec. 2.4.

Theorem 3.5.2. *The scalar linear time invariant system in (3.1) with $A = \lambda$ can be mean square stabilized using a linear scheme over the* non-orthogonal half-duplex *network if*

$$\log(|\lambda|) < \frac{1}{4} \max_{\substack{0<\beta\leq 1 \\ P_r^i : \sum_i P_r^i \leq P_R}} \left\{ \log\left(1 + \frac{2h^2\beta P_S}{N_d}\right) + \log\left(1 + \frac{\tilde{M}\left(\beta, \{P_r^i\}_{i=1}^L\right)}{\tilde{N}\left(\beta, \{P_r^i\}_{i=1}^L\right)}\right)\right\}, \tag{3.21}$$

where $\tilde{M}\left(\beta, \{P_r^i\}_{i=1}^L\right) = \left(\sqrt{2h^2(1-\beta)P_S} + \sqrt{\frac{2\beta P_S N_d}{(2h^2\beta P_S + N_d)}}\left(\sum_{i=1}^L \sqrt{\frac{2h_i^2 P_r^i}{2\beta P_S + N_r^i}}\right)\right)^2$ *and* $\tilde{N}(\beta, \{P_r^i\}_{i=1}^L) = \sum_{i=1}^L \frac{2h_i^2 P_r^i N_r^i}{2\beta P_S + N_r^i} + N_d$ *are real-valued functions.*

Proof. The proof is given in Appendix 3.B. □

Remark 3.5.1. *An optimal choice of the power allocation parameter β at the state encoder and an optimal power allocation at the relay nodes $\{P_r^i\}_{i=1}^L$ which maximize the term on the right hand side of (3.21) depend on the quality of the $\mathcal{E}-\mathcal{D}$, $\mathcal{E}-\mathcal{R}_i$, and $\mathcal{R}_i - \mathcal{D}$ links. This is a non-convex optimization problem, however it can be transformed into an equivalent convex problem by using the approach in [XCLG08, Appendix A]. This equivalent convex problem can be efficiently solved for optimal $\{P_r^i\}_{i=1}^L$ using the interior point method. For $\beta = 1$, we can analytically obtain the following optimal power allocation using the Lagrangian method:*

$$P_r^i = P_R\left(\frac{h_i^2\left(2P_S + N_r^i\right)}{\left(2P_S N_d + N_r^i N_d + P_R h_i^2 N_r^i\right)^2}\right)\left[\sum_{l=1}^L \frac{h_l^2\left(2P_S + N_r^l\right)}{\left(2P_S N_d + N_r^l N_d + P_R h_l^2 N_r^l\right)^2}\right]^{-1}. \tag{3.22}$$

Remark 3.5.2. *For channels with feedback, directed information is a useful quantity [Mas90, TSM09]. It is shown in Appendix 3.D that the term on the RHS of (3.21) is the information rate over the half-duplex network with noiseless feedback, obtained when running the described closed-loop protocol. Further we show that the* *directed information rate is also equal to the term on the RHS of (3.21). By following the same steps as in Appendix 3.D, one can also obtain relationships between sufficient conditions and information rates under linear policies for the cascade and the parallel relay networks considered in Sec. 3.3 and Sec. 3.4 respectively.*

Remark 3.5.3. *It is interesting to see that the conditions in* (3.18) *and* (3.21) *do not depend on the process noise.*

3.5.2 Noiseless Plant

In the absence of the process noise in (3.1), the state variance of the noiseless system under our proposed scheme can be obtained by substituting $n_w = 0$ in (3.39) and (3.38). That is the state variance of the noiseless plant is given by

$$\alpha_t = \left(\frac{\lambda^2 N}{2h^2 \beta P_S + N} \right) \alpha_{t-1}, \qquad t = 2, 4, 6, ...$$
$$\alpha_t = \left(\frac{\lambda^4 k \tilde{N}(\beta)}{(k_2 + \sqrt{k_1 k})^2 + \tilde{N}(\beta)} \right) \alpha_{t-2}, \qquad t = 3, 5, 7, ... \tag{3.23}$$

Since $\alpha_1 = \frac{\lambda^2 N}{h^2 P_S} \alpha_0 + n_w$, the state variance $\alpha_t \to 0$ as $t \to \infty$ if $\left(\frac{\lambda^4 k \tilde{N}(\beta)}{(k_2 + \sqrt{k_1 k})^2 + \tilde{N}(\beta)} \right) \leq 1$. This is the same condition as in (3.44). Thus by using the proposed linear scheme, we obtain identical sufficient conditions for mean square stability of noisy and noiseless first order LTI systems over the *non-orthogonal half-duplex* sensor network. Although the sufficient conditions are identical, the state variance of the noisy plant does not converge to zero unlike the noiseless plant.

Remark 3.5.4. *In the presence of process noise the inequality* (3.21) *is strict, however it follows from the analysis done above that the equality can be achieved in the absence of process noise.*

A comparison of second moments of the plant state process at three different power levels of process noise is illustrated in Fig. 3.5. We fix network and plant parameters, and plot the second moment of the state process $\mathbb{E}[X_t^2]$ moment as a function of time t for three power levels of process noise, i.e., $n_w = 0$, 0.1, and 0.25. For the given set of channel parameters, *mean square stability* of the system requires $|\lambda| < 1.975$ according to Theorem 3.5.2. We have fixed $\lambda = 1.5$, therefore the second moment stays bounded for all levels of process noise. For $n_w = 0$ the second moment converges to zero, starting from an arbitrary value equal to 0.25. For non-zero values of noise process, the second moment keeps alternating between two different values. This happens due to the first and the second transmission phases in the *half-duplex* network.

As shown in Fig. 3.5, for $n_w = 0.1$ and $n_w = 0.25$ the second moment converges to a unique non-zero value for each transmission phase and thus it keeps alternating between these two unique limit points. We can also observe that the rate of

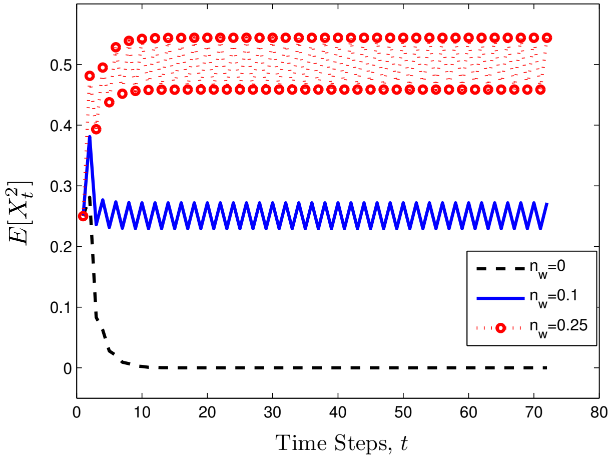


Figure 3.5: Comparison of second moments of the plant state process at three different levels of process noise for fixed network parameters $\{P_S = 1, L = 1, P_r = 1, h = h_1 = 1, \beta = 0.5, N_d = 0.5, N_r^1 = 0.1\}$ and plant parameters $\{\alpha_0 = 0.25, \lambda = 1.5\}$.

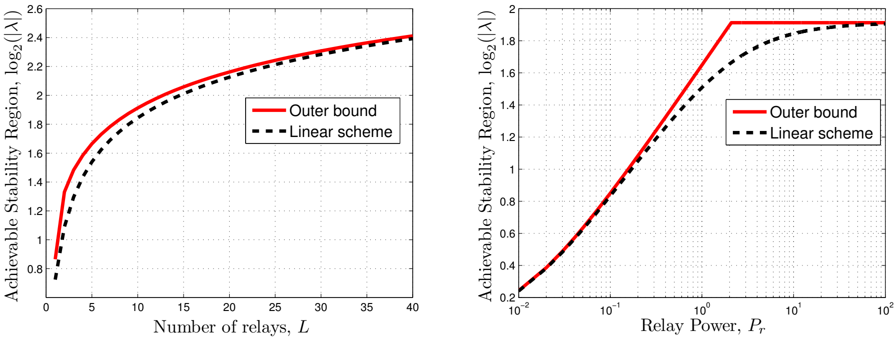
convergence is similar for the three examples, and seems to be unaffected by the power level of process noise.

In the following we study a special case of the general non-orthogonal half-duplex network, where there is no direct communication link between the state encoder and the remote controller.

3.5.3 Two-Hop Network

Consider the half-duplex relay network illustrated in Fig. 3.4 with $h = 0$. The state information is communicated to the remote controller only via the relay nodes. We call this setup a *two-hop* relay network, where the communication from the state encoder to the controller takes place in two hops. In the first hop the relay nodes receive the state information from the state encoder, which then communicate the state information to the controller in the second hop. The controller takes action in alternate time steps upon receiving the state information. We can obtain a sufficient condition for stability over this network by substituting $h = 0, \beta = 1$ in Theorem 3.5.2. Similarly a necessary condition can be obtained from (3.18), where $\beta = 1$ is the maximizer of the first term and $\beta = 0$ is the maximizer of the second term. In the following we evaluate the gap between the sufficient and necessary conditions for a symmetric two hop network.

Proposition 3.5.1. *For a symmetric two-hop network with* $P_r^i = P_r, N_r^i =$



(a) $P_S = 2P_r^i = 10, N_r^i = N_d = 1, h_i = 1$ (b) $L = 10, P_S = 10, N_r^i = N_d = 1, h_i = 1$

Figure 3.6: Comparison of necessary and sufficient conditions for a symmetric two-hop relay network.

$N_r, h_i = c, h = 0, \beta = 1$, *the gap between necessary and sufficient conditions approaches zero as the number of relays L goes to infinity. The gap also monotonically approaches zero as P_r goes to infinity.*

Proof. For $P_r^i = P_r, N_r^i = N_r, h_i = c, h = 0, \beta = 1$ for all i, the R.H.S. of (3.21) is evaluated as $\Gamma_{\text{suf}} := \frac{1}{4}\log\left(1 + \frac{4L^2c^2P_SP_r}{2Lc^2P_rN_r+N_d(2P_S+N_r)}\right)$ and the R.H.S of (3.18) can be bounded as $\Gamma_{\text{nec}} := \frac{1}{4}\log\left(1 + \frac{2LP_S}{N_r}\right)$. The gap between Γ_{suf} and Γ_{nec} is given by

$$\Gamma_{\text{nec}} - \Gamma_{\text{suf}} = \frac{1}{4}\log\left(1 + \frac{\frac{4P_S^2N_d+2P_SN_rN_d}{L}}{4c^2P_SP_rN_r + \frac{2c^2P_rN_r^2}{L} + \frac{N_dN_r(2P_S+N_r)}{L^2}}\right), \tag{3.24}$$

which approaches zero as L goes to infinity. The gap also monotonically approaches zero as P_r tends to infinity. □

In Fig. 3.6 we have plotted Γ_{nec} and Γ_{suf} as functions of L and P_r. These figures show that linear schemes are quite efficient in some regimes.

Remark 3.5.5. *Linear policies can be even exactly optimal in the following special cases: i) If we fix all relaying policies to be linear, then the channel becomes equivalent to a point-point scalar Gaussian channel, for which linear sensing is known to be optimal for LQG control [BB89]. ii) If we fix the state encoder to be linear and assume noiseless causal feedback links from the controller to the relay nodes, then linear policies are optimal for mean-square stabilization over a symmetric* two-hop *relay network, by the following arguments. Since the control actions are available at the relay nodes via noiseless feedback links, there is no dual effect of control,*

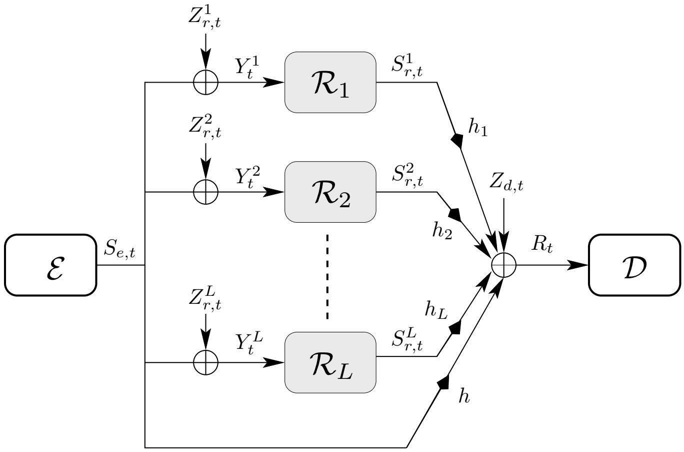


Figure 3.7: A Non-orthogonal Full-duplex Gaussian Network.

i.e., the separation of estimation and control holds. Further by restricting the state encoder to be linear, the relay network becomes equivalent to the Gaussian network studied in [Gas08, Gas07], where it is shown that linear policies are optimal if the network is symmetric.

3.5.4 Non-orthogonal Full-duplex Network

We now consider a non-orthogonal network of L *full-duplex* relay nodes, where all the nodes receive and transmit their signals in every time step, i.e., at any time instant $t \in \mathbb{N}$,

$$S_{e,t} = f_t\left(X_{[0,t]}, U_{[0,t-1]}\right), \qquad S^i_{r,t} = g^i_t\left(Y^i_{[0,t-1]}\right), \quad \forall t \in \mathbb{N},$$

$$Y^i_t = S_{e,t} + Z^i_{r,t}, \qquad R_t = hS_{e,t} + \sum_{i=1}^{L} S^i_{r,t} + Z_{d,t}, \quad \forall t \in \mathbb{N}, \tag{3.25}$$

where $\mathbb{E}\left[(S_{e,t})^2\right] = P_S$, $\mathbb{E}\left[(S^i_{r,t})^2\right] = P^i_r$, and $\sum_{i=1}^{L} P^i_r \leq P_R$.

Theorem 3.5.3. *If the linear system in* (3.1) *is mean-square stable over the* non-orthogonal full-duplex *relay network, then*

$$\log\left(|\det(A)|\right) \leq \frac{1}{2}\min\Bigg\{\log\left(1 + P_S\left(\sum_{i=1}^{L}\frac{1}{N^i_r} + \frac{h^2}{N_d}\right)\right),$$
$$\max_{P^i_r:\sum_i P^i_r \leq P_R}\left(\log\left(1 + \frac{1}{N_d}\left(\sum_{i=1}^{L+1}\delta_i^2 P_i + 2\sum_{i=1}^{L+1}\sum_{k=i+1}^{L+1}\rho^{\star}_{i,k}\delta_i\delta_k\sqrt{P_iP_k}\right)\right)\right)\Bigg\}, \tag{3.26}$$

where $\rho^\star_{i,k} = \frac{P_S}{\sqrt{(P_S+N_i)(P_S+N_k)}}$, $P_{L+1} := P_S$, $N_{L+1} := 0$, $\delta_{L+1} := h$, $P_i := P_r^i$, $\delta_i := h_i$, $N_i := N_r^i$ *for all* $i = \{1, 2, \ldots, L\}$.

Proof. The proof follows exactly in the steps of the proof of Theorem 3.5.1, with an exception that odd and even indexed terms are not treated separately because $\mathbb{E}\left[S_{e,t}^2\right] = P_S$ and $\mathbb{E}\left[(S_{r,t}^i)^2\right] = P_r^i$ for all t. □

Theorem 3.5.4. *The scalar linear time invariant system in (3.1) with* $A = \lambda$ *and* $W_t = 0$ *can be mean square stabilized using a linear scheme over the* non-orthogonal full-duplex *Gaussian network if*

$$\log(|\lambda|) < \frac{1}{2} \max_{P_r^i : \sum_{i=1}^L P_r^i \le P_R} \left\{ \log\left(1 + \left(\sqrt{h^2 P_S} + \eta^\star \sum_{i=1}^L \sqrt{\frac{h_i^2 P_S P_r^i}{P_S + N_r^i}}\right)^2 \left(N_d + \sum_{i=1}^L \frac{h_i^2 P_r^i N_r^i}{P_S + N_r^i}\right)^{-1}\right)\right\}, \tag{3.27}$$

where $\eta^\star$ *is the unique root in the interval* $[0, 1]$ *of the following fourth order polynomial*

$$\left(\sum_{i=1}^L \sqrt{\frac{h_i^2 P_S P_r^i}{(P_S + N_r^i)}}\right)\eta^4 + \left(2hP_S \sum_{i=1}^L \sqrt{\frac{h_i^2 P_r^i}{(P_S + N_r^i)}}\right)\eta^3 + \left(h^2 P_S + N_d + \sum_{i=1}^L \frac{h_i^2 P_r^i N_r^i}{P_S + N_r^i}\right)\eta^2 = \left(N_d + \sum_{i=1}^L \frac{h_i^2 P_r^i N_r^i}{P_S + N_r^i}\right). \tag{3.28}$$

Proof. The proof is given in Appendix 3.C. □

Remark 3.5.6. *The term on the right hand side of the inequality in (3.27) is an achievable rate with which information can be transmitted reliably over the* non-orthogonal full-duplex *relay network. This result is derived for a network with single relay node in [BW09, Theorem 5], however it can be easily extended to problems with multiple relays.*

3.6 Conclusions

We have studied the problem of mean-square stabilization of LTI plants over Gaussian relay networks. Some necessary and sufficient conditions for stabilization are presented which reveal relationships between stabilizability and communication parameters. These results can serve as a useful guideline for a system designer. Necessary conditions have been derived using information theoretic tools, and are not tight in general. Sufficient conditions for stabilization of scalar plants are obtained by employing linear communication and control schemes. We have discussed optimality of linear policies over the given network topologies. In some special cases, linear schemes are shown to be either asymptotically optimal or exactly optimal. It is observed in some network settings that sufficient conditions do not depend on the plant noise and they can be characterized by the directed information rate from the sequence of channel inputs to the sequence of channel outputs.

Appendix

3.A Proof of Necessary Condition

Consider the following series of inequalities:

$$
\begin{aligned}
& I\left(X_{[0,T\text{-}1]} \to R_{[0,T\text{-}1]}\right) \\
& \overset{(a)}{=} \sum_{t=0}^{T-1} I\left(X_{[0,t]}; R_t | R_{[0,t\text{-}1]}\right) \\
& \overset{(b)}{\geq} \sum_{t=0}^{T-1} I\left(X_t; R_t | R_{[0,t\text{-}1]}\right) \\
& = I\left(X_0; R_0\right) + \sum_{t=1}^{T-1} I\left(X_t; R_t | R_{[0,t\text{-}1]}\right) \\
& = I\left(X_0; R_0\right) + \sum_{t=1}^{T-1} \left(h\left(X_t | R_{[0,t\text{-}1]}\right) - h\left(X_t | R_{[0,t]}\right)\right) \\
& \overset{(c)}{=} I\left(X_0; R_0\right) + \sum_{t=1}^{T-1} \left(h\left(AX_{t-1} + BU_{t-1} + W_{t-1} | R_{[0,t\text{-}1]}\right) - h\left(X_t | R_{[0,t]}\right)\right) \\
& \overset{(d)}{=} I\left(X_0; R_0\right) + \sum_{t=1}^{T-1} \left(h\left(AX_{t-1} + W_{t-1} | R_{[0,t\text{-}1]}\right) - h\left(X_t | R_{[0,t]}\right)\right) \\
& \overset{(e)}{\geq} I\left(X_0; R_0\right) + \sum_{t=1}^{T-1} \left(h\left(AX_{t-1} + W_{t-1} | R_{[0,t\text{-}1]}, W_{t-1}\right) - h\left(X_t | R_{[0,t]}\right)\right) \\
& \overset{(f)}{=} I\left(X_0; R_0\right) + \sum_{t=1}^{T-1} \left(h\left(AX_{t-1} | R_{[0,t\text{-}1]}\right) - h\left(X_t | R_{[0,t]}\right)\right) \\
& \overset{(g)}{=} I\left(X_0; R_0\right) + \sum_{t=1}^{T-1} \left(\log\left(|\det\left(A\right)|\right) + h\left(X_{t-1} | R_{[0,t\text{-}1]}\right) - h\left(X_t | R_{[0,t]}\right)\right) \\
& = I\left(X_0; R_0\right) + (T-1)\log\left(|\det\left(A\right)|\right) + h\left(X_0 | R_0\right) - h\left(X_{T\text{-}1} | R_{[0,T\text{-}1]}\right) \\
& = h\left(X_0\right) + (T-1)\log\left(|\det\left(A\right)|\right) - h\left(X_{T\text{-}1} | R_{[0,T\text{-}1]}\right)
\end{aligned}
$$

$$\overset{(h)}{\geq} h\left(X_0\right) + (T-1)\log\left(\left|\det\left(A\right)\right|\right) - h\left(X_{T\text{-}1}\right)$$
$$\overset{(i)}{\geq} h\left(X_0\right) + (T-1)\log\left(\left|\det\left(A\right)\right|\right) - \log\left((2\pi e)^n\left|\det\left(K\right)\right|\right), \tag{3.29}$$

where (a) follows from the definition of directed information; (b) follows from the fact that discarding variables cannot increase mutual information; (c) follows from (3.1); (d) follows from $U_{t-1} = \pi_{t-1}(R_{[0,t\text{-}1]})$; (e) follows from the fact that conditioning reduces entropy; (f) follows from $h\left(AX_{t\text{-}1} + W_{t\text{-}1}|R_{[0,t\text{-}1]}, W_{t-1}\right) = h(AX_{t-1}|R_{[0,t\text{-}1]}, W_{t-1}) = h(AX_{t-1}|R_{[0,t\text{-}1]})$ due to mutual independence of X_t and W_t; (g) follows from $h\left(AX\right) = \log\left(\left|\det\left(A\right)\right|\right) + h(X)$ [CT06, Theorem 8.6.4]; (h) follows from conditioning reduces entropy; and (i) follows the fact that for a mean square stable system there exists a matrix $K \succ 0$ with $\det\left(\mathbb{E}\left[X_t^T X_t\right]\right) < \det\left(K\right)$ for all t and further for a given covariance matrix the differential entropy is maximized by the Gaussian distribution. We can also write

$$\begin{aligned}
I\left(X_{[0,T\text{-}1]} \to R_{[0,t\text{-}1]}\right) &= \sum_{t=0}^{T-1} I\left(X_{[0,t]}; R_t|R_{[0,t\text{-}1]}\right) \\
&\overset{(a)}{=} \sum_{t=0}^{T-1} I\left(\bar{X}_{[0,t]} + \bar{f}\left(U_{[0,t\text{-}1]}\right); R_t|R_{[0,t\text{-}1]}\right) \\
&\overset{(b)}{=} \sum_{t=0}^{T-1} I\left(\bar{X}_{[0,t]}; R_t|R_{[0,t\text{-}1]}\right) \\
&= I\left(\bar{X}_{[0,T\text{-}1]} \to R_{[0,T\text{-}1]}\right),
\end{aligned} \tag{3.30}$$

where (a) follows by defining uncontrolled state process $\bar{X}_{t+1} = A\bar{X}_t + W_t$ and writing the controlled state process $X_{[0,t]}$ as a sum of uncontrolled process $\bar{X}_{[0,t]}$ and a linear function of control actions $U_{[0,t\text{-}1]}$, since the system is linear and control actions are additive; and (b) follows from $U_t = \pi_t(R_{[0,t]})$. From (3.29) and (3.30) we have $\liminf_{T\to\infty}\frac{1}{T}I\left(\bar{X}_{[0,T\text{-}1]} \to R_{[0,T\text{-}1]}\right) \geq \log\left(\left|\det\left(A\right)\right|\right)$, since we have assumed $h(X_0) < \infty$.

3.B Proof of Theorem 3.5.2

In order to prove Theorem 3.5.2 we propose a linear communication and control scheme. This scheme is based on the coding scheme given in [BW09] which is an adaptation of the well-known Schalkwijk–Kailath scheme [SK66]. By employing the proposed linear scheme, we find a condition on the system parameters λ which is sufficient to mean square stabilize the system (3.1). The control and communication scheme for the *half-duplex* network works as follows: If the initial state X_0 is not Gaussian distributed, then we first make the state process Gaussian distributed by performing the following initialization step which was introduced in [ZOYS10].

Initial time step, $t = 0$

At time step $t = 0$, the state encoder $\mathcal{E}$ observes X_0 and it transmits $S_{e,0} = \sqrt{\frac{P_S}{\alpha_0}} X_0$. The decoder $\mathcal{D}$ receives $R_0 = hS_{e,0} + Z_{d,0}$. It estimates X_0 as $\hat{X}_0 = \frac{1}{h}\sqrt{\frac{\alpha_0}{P_S}} R_0 = X_0 + \frac{1}{h}\sqrt{\frac{\alpha_0}{P_S}} Z_{d,0}$. The controller $\mathcal{C}$ then takes an action $U_0 = -\lambda \hat{X}_0$ which results in

$$\begin{aligned} X_1 &= \lambda X_0 + U_0 + W_0 \\ &= \lambda \left(X_0 - \hat{X}_0 \right) + W_0 \\ &= -\frac{\lambda}{h}\sqrt{\frac{\alpha_0}{P_S}} Z_{d,0} + W_0. \end{aligned} \tag{3.31}$$

The new plant state $X_1 \sim \mathcal{N}(0, \alpha_1)$ where $\alpha_1 = \frac{\lambda^2 N_d}{h^2 P_S}\alpha_0 + n_w$.

First transmission phase, $t = 1, 3, 5, ...$

The state encoder $\mathcal{E}$ observes X_t and transmits $S_{e,t} = \sqrt{\frac{2\beta P_S}{\alpha_t}} X_t$. The relay nodes $\{\mathcal{R}_i\}_{i=1}^{L}$ receive this signal over the Gaussian links and do not transmit any signal in this transmission phase due to half-duplex restriction. The decoder $\mathcal{D}$ observes $R_t = hS_{e,t} + Z_{d,t}$ and computes the MMSE estimate of X_t, which is given by

$$\begin{aligned} \hat{X}_t &= \mathbb{E}[X_t | R_{[1,t]}] \\ &\overset{(a)}{=} \mathbb{E}[X_t | R_t] \\ &\overset{(b)}{=} \frac{\mathbb{E}[X_t R_t]}{\mathbb{E}[R_t^2]} R_t \\ &\overset{(c)}{=} \left(\frac{h\sqrt{2\beta P_S \alpha_t}}{2h^2 \beta P_S + N_d} \right) R_t, \end{aligned}$$

where (a) follows from the *orthogonality principle* of MMSE estimation (that is $\mathbb{E}[X_t R_{t-j}] = 0$ for $j \geq 1$) [Hay96]; (b) follows from the fact that the optimum MMSE estimator for a Gaussian variable is linear [Hay96]; and (c) follows from $\mathbb{E}[X_t R_t] = \sqrt{2h^2 \beta P_S \alpha_t}$ and $\mathbb{E}[R_t^2] = 2h^2 \beta P_S + N_d$.
The controller $\mathcal{C}$ takes an action $U_t = -\lambda \hat{X}_t$ which results in $X_{t+1} = \lambda(X_t - \hat{X}_t) + W_t$. The new plant state X_{t+1} is a linear combination of zero mean Gaussian variables $\{X_t, \hat{X}_t, W_t\}$, therefore it is also zero mean Gaussian with the following variance

$$\begin{aligned} \alpha_{t+1} := \mathbb{E}[X_{t+1}^2] &= \lambda^2 \mathbb{E}[(X_t - \hat{X}_t)^2] + \mathbb{E}[W_t^2] \\ &= \lambda^2 \left(\frac{N_d}{2h^2 \beta P_S + N_d} \right) \alpha_t + n_w, \end{aligned} \tag{3.32}$$

where the last equality follows from $\mathbb{E}[X_t\hat{X}_t] = \mathbb{E}[\hat{X}_t^2] = \frac{2h^2\beta P_S\alpha_t}{2h^2\beta P_S + N_d}$ (by computation).

Second transmission phase, $t = 2, 4, 6, ...$

The encoder $\mathcal{E}$ observes X_t and transmits $S_{e,t} = \sqrt{\frac{2(1-\beta)P_S}{\alpha_t}}X_t$. In this phase the relay nodes choose to transmit their own signal to the decoder $\mathcal{D}$ and thus they cannot listen to the signal transmitted from the state encoder due to the half-duplex assumption. Each relay node amplifies the signal that it had received in the previous time step (first transmission phase) under an average transmit power constraint and transmits it to the decoder $\mathcal{D}$. The signal transmitted from the i-th relay node is thus given by $S_{r,t}^i = \sqrt{\frac{2P_r^i}{(2\beta P_S + N_r^i)}}\left(S_{e,t-1} + Z_{r,t-1}^i\right)$. The decoder $\mathcal{D}$ accordingly receives

$$\begin{aligned} R_t &= hS_{e,t} + \sum_{i=1}^{L} h_i S_{r,t}^i + Z_{d,t} \\ &= L_1 X_t + L_2 X_{t-1} + \tilde{Z}_t, \end{aligned} \tag{3.33}$$

where $L_1 = \sqrt{\frac{2(1-\beta)h^2P_S}{\alpha_t}}$, $L_2 = \sum_{i=1}^{L}\sqrt{\frac{4\beta h_i^2 P_S P_r^i}{(2\beta P_S + N_r^i)\alpha_{t-1}}}$, and $\tilde{Z}_t = Z_{d,t} + \sum_{i=1}^{L}\sqrt{\frac{2h_i^2P_r^i}{2\beta P_S + N_r^i}}Z_{r,t-1}^i$ is a white Gaussian noise sequence with zero mean and variance $\tilde{N}(\beta, \{P_r^i\}_{i=1}^L) = N_d + \sum_{i=1}^{L}\frac{2h_i^2P_r^iN_r^i}{2\beta P_S + N_r^i}$. The decoder then computes the MMSE estimate of X_t given all previous channel outputs $\{R_1, R_2, ..., R_t\}$ in the following three steps:

1. Compute the MMSE prediction of R_t from $\{R_1, R_2, ..., R_{t-1}\}$, which is given by $\hat{R}_t = L_2\hat{X}_{t-1}$, where $\hat{X}_{t-1}$ is the MMSE estimate of X_{t-1}.

2. Compute the innovation

$$\begin{aligned} I_t &= R_t - \hat{R}_t \\ &= L_1X_t + L_2(X_{t-1} - \hat{X}_{t-1}) + \tilde{Z}_t \\ &\overset{(a)}{=} \left(\frac{\lambda L_1 + L_2}{\lambda}\right)X_t - \frac{L_2}{\lambda}W_{t-1} + \tilde{Z}_t, \end{aligned} \tag{3.34}$$

   where (a) follows from $X_t = \lambda\left(X_{t-1} - \hat{X}_{t-1}\right) + W_{t-1}$.

3. Compute the MMSE estimate of X_t given $\{R_1, R_2, ..., R_{t-1}, I_t\}$. The state X_t is independent of $\{R_1, R_2, ..., R_{t-1}\}$ given I_t, therefore we can compute the

estimate $\hat{X}_t$ based only on I_t without any loss of optimality, that is,

$$\begin{aligned}\hat{X}_t &= \mathbb{E}[X_t|I_t]\\ &\overset{(a)}{=} \frac{\mathbb{E}[X_t I_t]}{\mathbb{E}[I_t^2]} I_t\\ &\overset{(b)}{=} \frac{\lambda\left(\lambda L_1 + L_2\right)\alpha_t}{\left(\lambda L_1 + L_2\right)^2 \alpha_t + L_2^2 n_w + \lambda^2 \tilde{N}(\beta, P_r)} I_t, \end{aligned} \tag{3.35}$$

where (a) follows from an MMSE estimation of a Gaussian variable; and (b) follows from $\mathbb{E}[X_t I_t] = \left(\frac{\lambda L_1 + L_2}{\lambda}\right)\alpha_t$ and $\mathbb{E}[I_t^2] = \left(\frac{\lambda L_1 + L_2}{\lambda}\right)^2 \alpha_t + \frac{L_2^2 n_w}{\lambda^2} + \tilde{N}(\beta, P_r)$.

The controller $\mathcal{C}$ takes action $U_t = -\lambda \hat{X}_t$ which results in $X_{t+1} = \lambda(X_t - \hat{X}_t) + W_t$. The new plant state X_{t+1} is a linear combination of zero mean Gaussian random variables $\{X_t, \hat{X}_t, W_t\}$, therefore it is also zero mean Gaussian distributed with the following variance,

$$\begin{aligned}\alpha_{t+1} &= \lambda^2 \mathbb{E}[(X_t - \hat{X}_t)^2] + \mathbb{E}[W_t^2]\\ &\overset{(a)}{=} \lambda^2 \alpha_t \left(\frac{L_2^2 n_w + \lambda^2 \tilde{N}(\beta, P_r)}{\left(\lambda L_1 + L_2\right)^2 \alpha_t + L_2^2 n_w + \lambda^2 \tilde{N}(\beta, P_r)}\right) + n_w \qquad (3.36)\\ &\overset{(b)}{=} \lambda^2 \alpha_t \left(\frac{\left(\sum_{i=1}^{L} \sqrt{\frac{4h_i^2 \beta P_S P_r^i}{2\beta P_S + N_r^i}}\right)^2 \frac{n_w}{\alpha_{t-1}} + \lambda^2 \tilde{N}(\beta, P_r)}{\left(\lambda\sqrt{2h^2(1-\beta P_S)} + \sum_{i=1}^{L} \sqrt{\frac{4h_i^2 \beta P_S P_r^i}{2\beta P_S + N_r^i} \frac{\alpha_t}{\alpha_{t-1}}}\right)^2 + \left(\sum_{i=1}^{L} \sqrt{\frac{4h_i^2 \beta P_S P_r^i}{2\beta P_S + N_r^i}}\right)^2 \frac{n_w}{\alpha_{t-1}} + \lambda^2 \tilde{N}(\beta, P_r)}\right) + n_w\\ &\overset{(c)}{=} \lambda^2 \left(\lambda^2 k \alpha_{t-1} + n_w\right) \left(\frac{(n_w k_1) \frac{1}{\alpha_{t-1}} + \lambda^2 \tilde{N}(\beta, P_r)}{\left(\lambda k_2 + \sqrt{\frac{k_1}{\lambda^2}(\lambda^2 k + n_w \frac{1}{\alpha_{t-1}})}\right)^2 + (n_w k_1)\frac{1}{\alpha_{t-1}} + \lambda^2 \tilde{N}(\beta, P_r)}\right) + n_w\\ &= \lambda^2 \left(\lambda^2 k \alpha_{t-1} + n_w\right) \left(\frac{\left(\frac{n_w k_1}{\lambda^2}\right) \frac{1}{\alpha_{t-1}} + \tilde{N}(\beta, P_r)}{\left(k_2 + \sqrt{k_1 k + \frac{n_w k_1}{\lambda^2} \frac{1}{\alpha_{t-1}}}\right)^2 + \left(\frac{n_w k_1}{\lambda^2}\right)\frac{1}{\alpha_{t-1}} + \tilde{N}(\beta, P_r)}\right) + n_w, \qquad (3.37)\end{aligned}$$

where (a) follows from $\mathbb{E}[X_t \hat{X}_t] = \mathbb{E}[\hat{X}_t^2] = \frac{(\lambda L_1 + L_2)^2 \alpha_t}{(\lambda L_1 + L_2)^2 \alpha_t + L_2^2 n_w + \lambda^2 \tilde{N}(\beta, P_r)}$; (b) follows by substituting the values of L_1 and L_2; and (c) by substituting $\frac{\alpha_t}{\alpha_{t-1}}$ using (3.32) and by defining $k := \frac{N}{2h^2\beta P_S+N}$, $k_1 := \left(\sum_{i=1}^{L}\sqrt{\frac{4h_i^2\beta P_S P_r^i}{2\beta P_S+N_r^i}}\right)^2$, $k_2 := \sqrt{2h^2(1-\beta P_S)}$.

We want to find the values of the parameter λ for which the second moment of the state remains bounded. Rewriting (3.32) and (3.37), the variance of the state at any time t is given by

$$\alpha_t = \lambda^2\left(\lambda^2 k\alpha_{t-2}+n_w\right)\underbrace{\left(\frac{\left(\frac{n_w k_1}{\lambda^2}\right)\frac{1}{\alpha_{t-2}}+\tilde{N}(\beta,P_r)}{\left(k_2+\sqrt{k_1k+\frac{n_wk_1}{\lambda^2}\frac{1}{\alpha_{t-2}}}\right)^2+\left(\frac{n_wk_1}{\lambda^2}\right)\frac{1}{\alpha_{t-2}}+\tilde{N}(\beta,P_r)}\right)}_{\triangleq f(\alpha_{t-2})}+n_w$$

$$= \lambda^2\left(\lambda^2 k\alpha_{t-2}+n_w\right)f(\alpha_{t-2})+n_w, \qquad t=3,5,7,... \tag{3.38}$$

$$\alpha_t = \lambda^2\left(\frac{N}{2h^2\beta P_S+N}\right)\alpha_{t-1}+n_w, \qquad t=2,4,6,... \tag{3.39}$$

where $\alpha_1 = \frac{\lambda^2 N}{h^2 P_S}\alpha_0+n_w$. If the odd indexed sub-sequence $\{\alpha_{2t+1}\}$ in (3.38) is bounded, then the even indexed sub-sequence $\{\alpha_{2t}\}$ in (3.39) is also bounded. Thus it is sufficient to consider the odd indexed sub-sequence $\{\alpha_{2t+1}\}$. We will now construct a sequence $\{\alpha'_t\}$ which upper bounds the sub-sequence $\{\alpha_{2t+1}\}$. Then we will derive conditions on the system parameter λ for which the sequence $\{\alpha'_t\}$ stays bounded and consequently the boundedness of $\{\alpha_{2t+1}\}$ will be guaranteed. In order to construct the upper sequence $\{\alpha'_t\}$, we work on the term $f(\alpha_{t-2})$ in (3.38) and make use of the following lemma.

Lemma 3.B.1. *Consider a function* $f(x) = \frac{a+\frac{b}{x}}{(c+\sqrt{d+\frac{b}{x}})^2+a+\frac{b}{x}}$ *defined over the interval* $[0,\infty)$, *where* $0 \le a,b,c,d < \infty$. *The function* $f(x)$ *can be upper bounded as* $f(x) \le f_\infty + \frac{m}{x}$ *for some* $0 < m < \infty$, *where* $f_\infty := \lim_{x\to\infty} f(x) = \frac{a}{(c+\sqrt{d})^2+a}$.

Proof. We want to show that for some $m > 0$, $f(x) \le f_\infty + \frac{m}{x}$ for all $x \in [0,\infty)$, where $f_\infty := \lim_{x\to\infty} f(x) = \frac{a}{(c+\sqrt{d})^2+a}$. To this end, we show that $-f(x)+f_\infty+\frac{m}{x}$ is non-negative for some $m \ge 0$.

$$-f(x)+f_\infty+\frac{m}{x} = -\frac{ax+b}{(c\sqrt{x}+\sqrt{dx+b})^2+ax+b}+\frac{a}{(c+\sqrt{d})^2+a}+\frac{m}{x}$$

$$= \frac{1}{\left(\left(c+\sqrt{d}\right)^2+a\right)\left(\left(c\sqrt{x}+\sqrt{dx+b}\right)^2+ax+b\right)x}$$

$$\times\Bigg[-(ax^2+bx)\left(\left(c+\sqrt{d}\right)^2+a\right)$$

$$+\,ax\left(\left(c\sqrt{x}+\sqrt{dx+b}\right)^2+ax+b\right)$$
$$\left.+m\left(\left(c+\sqrt{d}\right)^2+a\right)\left(\left(c\sqrt{x}+\sqrt{dx+b}\right)^2+ax+b\right)\right] \quad (3.40)$$

The denominator term in (3.40) is always positive for $x \in [0, \infty)$, therefore we focus on the numerator term. The numerator term after simplification is equal to

$$m(c^4 + c^2d + 2c^3d + c^2d + d^2 + 2cd^{\frac{3}{2}} + 2ad + ac^2 + 2ac\sqrt{d} + a^2)x$$
$$- (bc^2 + bd + 2bc\sqrt{d})x + \underbrace{2ac(x\sqrt{dx^2 + bx} - \sqrt{d}x^2)}_{\geq 0} + \underbrace{\vartheta}_{\geq 0}, \quad (3.41)$$

where ϑ is the summation of the remaining terms independent of x, which are all non-negative and therefore their sum is also non-negative. If we choose $m \geq \frac{(bc^2+bd+2bc\sqrt{d})}{(c^4+c^2d+2c^3d+c^2d+d^2+2cd^{\frac{3}{2}}+2ad+ac^2+2ac\sqrt{d}+a^2)}$ in (3.41), the non-negativity of (3.40) in the interval $[0, \infty)$ will be guaranteed. □

Starting from (3.38) and by using the above lemma, we write the following series of inequalities

$$\begin{aligned}
\alpha_t &= \lambda^2\left(\lambda^2 k\alpha_{t-2} + n_w\right) f(\alpha_{t-2}) + n_w \\
&\overset{(a)}{\leq} \lambda^2\left(\lambda^2 k\alpha_{t-2} + n_w\right)\left(f_\infty + \frac{m}{\alpha_{t-2}}\right) + n_w \\
&= \lambda^4 k f_\infty \alpha_{t-2} + \frac{\lambda^2 n_w m}{\alpha_{t-2}} + \lambda^2 n_w f_\infty + \lambda^4 mk + n_w \\
&\overset{(b)}{\leq} \lambda^4 k f_\infty \alpha_{t-2} + \lambda^2 m + \lambda^2 n_w f_\infty + \lambda^4 mk + n_w =: g(\alpha_{t-2}),
\end{aligned} \quad (3.42)$$

where (a) follows from Lemma 3.B.1 and $f_\infty = \lim_{\alpha\to\infty} f(\alpha) = \left(\frac{\tilde{N}(\beta,P_r)}{(k_2+\sqrt{k_1 k})^2+\tilde{N}(\beta,P_r)}\right)$; and (b) follows from the fact that $\alpha_t \geq n_w$ for all t according to (3.39) and (3.38). Since $g(\alpha)$ in (3.42) is a linearly increasing function, it can be used to construct the sequence $\{\alpha'_t\}$, which upper bounds the odd indexed sub-sequence $\{\alpha_{2t+1}\}$ given in (3.38). We construct the sequence $\{\alpha'_t\}$ for all $t \geq 1$ as

$$\begin{aligned}
\alpha_{2t+1} \leq \alpha'_{t+1} := g(\alpha'_t) &\overset{(a)}{=} \lambda^4 k f_\infty \alpha'_t + \lambda^2 m + \lambda^2 n_w f_\infty + \lambda^4 mk + n_w \\
&\overset{(b)}{=} \left(\lambda^4 k f_\infty\right)^t \alpha'_0 + (\lambda^2 m + \lambda^2 n_w f_\infty + \lambda^4 mk + n_w)\sum_{i=0}^{t-1}\left(\lambda^4 k f_\infty\right)^i,
\end{aligned} \quad (3.43)$$

where (a) follows from (3.42) and (b) follows by recursively apply (a).

We observe from (3.43) that if $\left(\lambda^4 k f_\infty\right) = \left(\frac{\lambda^4 k\tilde{N}(\beta,P_r)}{(k_2+\sqrt{k_1 k})^2+\tilde{N}(\beta,P_r)}\right) < 1$, then the sequence $\{\alpha'_t\}$ converges as $t \to \infty$ and consequently the original sequence $\{\alpha_t\}$ is guaranteed to stay bounded. Thus the system in (3.1) can be mean square stabilized if

$$\lambda^4 < \left(\frac{(k_2+\sqrt{k_1 k})^2 + \tilde{N}(\beta,\{P_r^i\}_{i=1}^L)}{k\tilde{N}(\beta,\{P_r^i\}_{i=1}^L)}\right) \tag{3.44}$$

$$\Rightarrow \log(\lambda) < \frac{1}{4}\left(\log\left(1+\frac{2h^2\beta P_S}{N_d}\right) + \log\left(1+\frac{\tilde{M}(\beta,\{P_r^i\}_{i=1}^L)}{\tilde{N}(\beta,\{P_r^i\}_{i=1}^L)}\right)\right), \tag{3.45}$$

where the last equality follows from $k = \frac{N}{2h^2\beta P_S + N}$ and $M(\beta,\{P_r^i\}_{i=1}^L) := (k_2 + \sqrt{k_1 k})^2$. Since the relay nodes amplify the desired signal as well as the noise, which is then superimposed at the decoder to the signal coming directly from the state encoder, the optimal choice of the transmit powers $\{P_r^i\}_{i=1}^L : \sum_{i=1}^L P_r^i \le P_R$ depends on the parameters $\{P_S, \{N_r^i\}_{i=1}^L, N_d, h, h_i, \beta\}$. Moreover, the optimal choice of the power allocation factor β at the state encoder also depends on these parameters. Therefore, we rewrite (3.45) as

$$\log(\lambda) < \frac{1}{4} \max_{\substack{0<\beta\le 1 \\ P_r^i : \sum_i P_r^i \le P_R}} \left\{\log\left(1+\frac{2h^2\beta P_S}{N_d}\right) + \log\left(1+\frac{\tilde{M}\left(\beta,\{P_r^i\}_{i=1}^L\right)}{\tilde{N}\left(\beta,\{P_r^i\}_{i=1}^L\right)}\right)\right\}, \tag{3.46}$$

which completes the proof. □

3.C Proof of Theorem 3.5.4

The scheme for the *full-duplex* network is in principle similar to the scheme for the *half-duplex* network with some modifications to adapt to the full-duplex nature of the relay nodes. A full-duplex node can simultaneously transmit and receive signals, therefore in this scheme the relay nodes transmit in every time slot in contrast to the half-duplex network scenario where the relay nodes transmit in alternate time slots. The initial transmission and control at $t = 0$ in the *full-duplex* scenario is identical to that of the scheme proposed for the *half-duplex* scenario in Appendix 3.B. Therefore according to Appendix 3.B, the plant state $X_1 \sim \mathcal{N}(0, \alpha_1)$, where $\alpha_1 = \frac{\lambda^2 N_d}{h^2 P_S}\alpha_0$. Further transmissions and control actions are as follows.

Time step $t = 1$

The encoder $\mathcal{E}$ observes X_1 and it then transmits $S_{e,1} = \sqrt{P_S/\alpha_1}X_1$ to the decoder $\mathcal{D}$ at the remote control unit. The relay nodes $\{\mathcal{R}_i\}_{i=1}^L$ overhear the signal transmitted by the controller but remain silent. The decoder $\mathcal{D}$ observes $R_1 = hS_{e,1} + Z_1$

and computes the MMSE estimate of X_1, which is given by

$$\begin{aligned}\hat{X}_1 &= \mathbb{E}[X_1|R_1] \\ &\overset{(a)}{=} \frac{\mathbb{E}[X_1 R_1]}{\mathbb{E}[R_1^2]} R_1 \\ &\overset{(b)}{=} \left(\frac{h\sqrt{P_S \alpha_1}}{h^2 P_S + N_d}\right) R_1,\end{aligned}$$

where (a) follows from the fact that the optimum MMSE estimator for a Gaussian random variable is linear [Hay96]; and (b) follows from $\mathbb{E}[X_1 R_1] = \sqrt{h^2 P_S \alpha_1}$ and $\mathbb{E}[R_1^2] = h^2 P_S + N$.
The controller $\mathcal{C}$ takes action $U_1 = -\lambda \hat{X}_1$ which results in $X_2 = \lambda(X_1 - \hat{X}_1) + W_1$. The new plant state X_2 is a linear combination of zero mean Gaussian random variables $\{X_1, \hat{X}_1, W_1\}$, therefore it is also zero mean Gaussian distributed with the following variance

$$\begin{aligned}\alpha_2 := \mathbb{E}[X_2^2] &= \lambda^2 \mathbb{E}[(X_1 - \hat{X}_1)^2] \\ &= \lambda^2 \left(\frac{N_d}{h^2 P_S + N_d}\right) \alpha_1, \qquad (3.47)\end{aligned}$$

where the last equality follows from $\mathbb{E}[X_1 \hat{X}_1] = \mathbb{E}[\hat{X}_1^2] = \frac{h^2 P_S \alpha_1}{h^2 P_S + N_d}$ (by computation).

Time steps $t \geq 2$

The encoder $\mathcal{E}$ observes X_t and it then transmits $S_{e,t} = \sqrt{\frac{P_S}{\alpha_t}} X_t$. The relay nodes simultaneously receive this signal and transmit an amplified version of the signal which they had received in the previous time step. The signal transmitted by the i-th relay node is given by

$$S_{r,t}^i = \sqrt{\frac{P_r^i}{P_S + N_r^i}} \left(S_{e,t-1} + Z_{r,t-1}^i\right), \qquad \text{for all } i \in \{1, 2, \ldots, L\}, \qquad (3.48)$$

where $\mathbb{E}[(S_{r,t}^i)^2] = P_r^i$ are chosen to ensure $\sum_{i=1}^L P_r^i \leq P_R$. Accordingly $\mathcal{D}$ receives

$$\begin{aligned}R_t &= h S_{e,t} + \sum_{i=1}^{L} h_i S_{r,t} + Z_t \\ &= L_1 X_t + L_2 X_{t-1} + \tilde{Z}_t, \qquad (3.49)\end{aligned}$$

where $L_1 = \sqrt{\frac{h^2 P_S}{\alpha_t}}$, $L_2 = \sum_{i=1}^L \sqrt{\frac{h_i^2 P_S P_r^i}{(P_S + N_r^i)\alpha_{t-1}}}$, and $\tilde{Z}_t = Z_t + \sum_{i=1}^L \sqrt{\frac{h_i^2 P_r^i}{P_S + N_r^i}} Z_{r,t-1}^i$ with $\tilde{Z}_t \sim \mathcal{N}(0, \tilde{N})$. The computation of the state MMSE estimate $\hat{X}_t$ and the action taken by controller $U_t = -\lambda \hat{X}_t$ are identical to that of the *half-duplex* network scheme as proposed in Appendix 3.B. Therefore according to (3.36) the variance α_{t+1} of the new plant state X_{t+1} is given by

$$\begin{aligned}
\alpha_{t+1} &= \lambda^2 \alpha_t \left(\frac{\lambda^2 \tilde{N}(P_r)}{\left(\lambda L_1 + L_2\right)^2 \alpha_t + \lambda^2 \tilde{N}(P_r)} \right) \\
&\overset{(a)}{=} \lambda^2 \alpha_t \left(\frac{\lambda^2 \tilde{N}(P_r)}{\left(\lambda \sqrt{h^2 P_S} + \sum_{i=1}^{L} \sqrt{\frac{h_i^2 P_S P_r^i}{(P_S + N_r^i)} \frac{\alpha_t}{\alpha_{t-1}}}\right)^2 + \lambda^2 \tilde{N}(P_r)} \right) \\
&\overset{(b)}{=} \lambda^2 \alpha_t \left(\frac{\tilde{N}(P_r)}{\left(k_2 + \sqrt{\frac{k_1}{\lambda^2} \frac{\alpha_t}{\alpha_{t-1}}}\right)^2 + \tilde{N}(P_r)} \right), \qquad \forall t \geq 2,
\end{aligned} \tag{3.50}$$

where (a) follows by substituting the values of L_1 and L_2; and (b) follows by defining $k_1 = \left(\sum_{i=1}^{L} \sqrt{\frac{h_i^2 P_S P_r^i}{(P_S + N_r^i)}} \right)^2$ and $k_2 = \sqrt{h^2 P_S}$. Our aim is to find a condition on the system parameter λ which is sufficient to ensure that the state variance in (3.50) stays bounded. By defining $\eta_t := \sqrt{\frac{1}{\lambda^2} \frac{\alpha_t}{\alpha_{t-1}}}$, we can rewrite (3.50) as

$$\eta_{t+1} = \sqrt{\frac{\tilde{N}(P_r)}{\left(k_2 + \eta_t \sqrt{k_1}\right)^2 + \tilde{N}(P_r)}}, \qquad \forall t \geq 2. \tag{3.51}$$

In the following we show that the sequence $\{\eta_t\}$ converges to a unique fixed point. The convergence follows from the following lemma.

Lemma 3.C.1. *([BW09, Lemma 2]) Consider the function* $f : x \mapsto \sqrt{\frac{a}{a + p(1+bx)^2}}$ *defined on the closed interval [0, 1] when* $0 \leq a, b, p < \infty$. *The function* $f(\cdot)$ *has exactly one fixed point* $x^* \in [0, 1]$ *and the infinite sequence* $x_0, x_1 = f(x_0), x_2 = f(x_1), \ldots$ *converges to this fixed point for any starting point* $x_0 \in [0, 1]$.

According to Lemma 3.C.1, starting with $\eta_2 = \frac{\alpha_2}{\alpha_1} = \sqrt{\frac{N}{h^2 P_S + N}} \in [0, 1]$, the sequence $\{\eta_t\}$ in (3.51) converges to a fixed point $\eta^\star$. This fixed point is given by the unique solution in the interval [0,1] of the following fourth order polynomial in (3.28), which has been obtained by simplifying (3.51) and substituting the values of $k_1, k_2, \tilde{N}(P_r)$ [BW09]. Having shown that the sequence $\{\eta_t\}$ converges to the unique fixed point, we now find the values of the system parameter λ for which the state variance $\{\alpha_t\}$ converges to a limit point and consequently stays bounded.

Rewriting (3.50) as

$$\begin{aligned}\alpha_{t+1} &= \left(\frac{\lambda^2\tilde{N}(P_r)}{\left(k_2+\eta_t\sqrt{k_1}\right)^2+\tilde{N}(P_r)}\right)\alpha_t \\ &= \left(\prod_{i=2}^{t}\left(\frac{\tilde{N}(P_r)}{\left(k_2+\eta_i\sqrt{k_1}\right)^2+\tilde{N}(P_r)}\right)\right)\lambda^{2t-2}\alpha_2 \\ &= \left(\mu(t)\lambda^{2t}\right)\nu \\ &= \left(\lambda^{t\left[\frac{1}{t}\log_\lambda(\mu(t))+2\right]}\right)\nu, \end{aligned} \tag{3.52}$$

where $\mu(t) := \prod_{i=2}^{t}\left(\frac{\tilde{N}(P_r)}{\left(k_2+\eta_i\sqrt{k_1}\right)^2+\tilde{N}(P_r)}\right)$ and $\nu := \lambda^{-2}\alpha_2$. We observe that $\alpha_t \to 0$ as $t \to \infty$ if

$$\lim_{t\to\infty}\left[\frac{1}{t}\log_\lambda\left(\mu\left(t\right)\right)+2\right]<0, \tag{3.53}$$

where the existence of the limit follows from convergence of the sequence $\{\eta_i\}$. Since $\log_\lambda\left(\mu\left(t\right)\right) = \frac{\log(\mu(t))}{\log(\lambda)}$, we can rewrite (3.53) as

$$\begin{aligned}\log\left(\lambda\right) &< \lim_{t\to\infty}\frac{1}{2t}\log\left(\frac{1}{\mu\left(t\right)}\right) \\ &\overset{(a)}{=} \lim_{t\to\infty}\frac{1}{2t}\log\left(\prod_{i=2}^{t}\left(1+\frac{\left(k_2+\eta_i\sqrt{k_1}\right)^2}{\tilde{N}(P_r)}\right)\right) \\ &= \lim_{t\to\infty}\frac{1}{2t}\sum_{i=2}^{t}\log\left(1+\frac{\left(k_2+\eta_i\sqrt{k_1}\right)^2}{\tilde{N}(P_r)}\right) \\ &\overset{(b)}{=} \frac{1}{2}\log\left(1+\frac{\left(k_2+\eta^\star\sqrt{k_1}\right)^2}{\tilde{N}(P_r)}\right) \\ &\overset{(c)}{=} \frac{1}{2}\log\left(1+\left(\sqrt{h^2P_S}+\eta^\star\sum_{i=1}^{L}\sqrt{\frac{h_i^2P_SP_r^i}{P_S+N_r^i}}\right)^2\left(N_d+\sum_{i=1}^{L}\frac{h_i^2P_r^iN_r^i}{P_S+N_r^i}\right)^{-1}\right), \end{aligned} \tag{3.54}$$

where (a) follows by substituting the value of $\mu(t)$ from (3.52); (b) follows by convergence of the sequence $\{\eta_i\}$ to $\eta^\star$ and using Cesaro mean theorem [Har49]; and (c) follows by substituting the values of k_1, k_2, and $\tilde{N}(P_r)$. This completes the proof of Theorem 3.5.4. □

3.D Proof of Remark 3.5.2 on Information Rate

The given scheme can be seen as a point-point communication channel, where $R_{2t\text{-}1}$ is the channel output corresponding to the input $S_{e,2t\text{-}1}$ and I_{2t} is the channel output corresponding to the input $S_{e,2t}$ for $t = 1, 2, 3, \ldots$. The information rate is given by

$$\begin{aligned}
&\lim_{T\to\infty} \frac{1}{2T} I\left(\{S_{e,2t\text{-}1}\}_{t=1}^{T}, \{S_{e,2t}\}_{t=1}^{T}; \{R_{2t\text{-}1}\}_{t=1}^{T}, \{I_{2t}\}_{t=1}^{T}\right) \\
&\quad = \lim_{T\to\infty} \frac{1}{2T}\Bigg[h\left(\{R_{2t\text{-}1}\}_{t=1}^{T}, \{I_{2t}\}_{t=1}^{T}\right) \\
&\qquad - h\left(\{R_{2t\text{-}1}\}_{t=1}^{T}, \{I_{2t}\}_{t=1}^{T} \,|\, \{S_{e,2t\text{-}1}\}_{t=1}^{T}, \{S_{e,2t}\}_{t=1}^{T}\right)\Bigg] \\
&\quad \overset{(a)}{=} \lim_{T\to\infty} \frac{1}{2T}\left[\sum_{t=1}^{T}\Big(h\left(R_{2t\text{-}1}\right) + h\left(I_{2t}\right) - h\left(R_{2t\text{-}1}|S_{e,2t\text{-}1}\right) - h\left(I_{2t}|S_{e,2t}\right)\Big)\right] \\
&\quad \overset{(b)}{=} \lim_{T\to\infty} \frac{1}{2T}\Big[T\big(h\left(R_{2t\text{-}1}\right) + h\left(I_{2t}\right) - h\left(R_{2t\text{-}1}|S_{e,2t\text{-}1}\right) - h\left(I_{2t}|S_{e,2t}\right)\big)\Big] \\
&\quad = \frac{1}{2}\left(I\left(S_{e,2t\text{-}1}; R_{2t\text{-}1}\right) + I\left(S_{e,2t}; I_{2t}\right)\right), \qquad (3.55)
\end{aligned}$$

where (a) follows from the fact that $P(I_{2t}, R_{2t\text{-}1}|S_{e,2t}, S_{e,2t\text{-}1}) = P(I_{2t}|S_{e,2t})P(R_{2t\text{-}1}|S_{e,2t\text{-}1})$, the channel is memoryless, the random variables are Gaussian and $\mathbb{E}[R_{2l\text{-}1}R_{2k\text{-}1}] = \mathbb{E}[I_{2l}I_{2k}] = 0$ for $k \neq l$, and $\mathbb{E}[R_{2l\text{-}1}I_{2k}] = 0$ for all $l, k = 1, 2, 3, \ldots$; and (b) follows from the fact that $R_{2t\text{-}1}$ and I_{2t} are both sequences of i.i.d. variables. For the first transmission phase the mutual information between the transmitted variable and the received variable is given by

$$\begin{aligned}
I\left(S_{e,2t\text{-}1}; R_{2t\text{-}1}\right) &= h(R_{2t\text{-}1}) - h(R_{2t\text{-}1}|S_{e,2t\text{-}1}) \\
&= h(R_{2t\text{-}1}) - h(Z_{2t\text{-}1}) \\
&\overset{(a)}{=} \frac{1}{2}\log\left(1 + \frac{2h^2\beta P_S}{N}\right), \qquad (3.56)
\end{aligned}$$

where (a) follows from $R_{2t\text{-}1} \sim \mathcal{N}(0, 2h^2\beta P_S + N)$ and $Z_{2t\text{-}1} \sim \mathcal{N}(0, N)$. In the second phase the decoder computes the innovation I_t according to (3.34). The mutual information between the transmitted variable and the innovation variable is then given by

$$\begin{aligned}
I\left(S_{e,2t}; I_{2t}\right) &= h(I_{2t}) - h(I_{2t}|S_{e,2t}) \\
&= h(I_{2t}) - h(\tilde{Z}_{2t}) \\
&\overset{(a)}{=} \frac{1}{2}\log\left(1 + \frac{\tilde{M}(\beta, P_r)}{\tilde{N}(\beta, P_r)}\right), \qquad (3.57)
\end{aligned}$$

where (a) follows from $I_{2t} \sim \mathcal{N}(0, \tilde{M}(\beta, P_r) + \tilde{N}(\beta, P_r))$ and $\tilde{Z}_{2t} \sim \mathcal{N}(0, \tilde{N}(\beta, P_r))$. From (3.56), (3.57), and (3.55) the corresponding information rate is equal to

$$\frac{1}{4}\left(\log\left(1 + \frac{2h^2\beta P_S}{N}\right) + \log\left(1 + \frac{\tilde{M}(\beta, P_r)}{\tilde{N}(\beta, P_r)}\right)\right). \qquad (3.58)$$

Likewise, the directed information rate for the relay channel under discussion is given by

$$\begin{aligned}
&\lim_{T\to\infty} \frac{1}{2T} I\left(\{S_{e,2t\text{-}1}\}_{t=1}^{T}, \{S_{e,2t}\}_{t=1}^{T} \to \{R_{2t\text{-}1}\}_{t=1}^{T}, \{I_{2t}\}_{t=1}^{T}\right) \\
&\overset{(a)}{=} \lim_{T\to\infty} \frac{1}{2T} \left[\sum_{t=1}^{T} \left(I\left(S_{e,2t\text{-}1}; R_{2t\text{-}1}\right) + I\left(S_{e,2t}; I_{2t}\right)\right)\right] \\
&\overset{(b)}{=} \lim_{T\to\infty} \frac{1}{2T} \left[T\left(I\left(S_{e,2t\text{-}1}; R_{2t\text{-}1}\right) + I\left(S_{e,2t}; I_{2t}\right)\right)\right] \\
&= \frac{1}{2}\left(I\left(S_{e,2t\text{-}1}; R_{2t\text{-}1}\right) + I\left(S_{e,2t}; I_{2t}\right)\right),
\end{aligned} \tag{3.59}$$

where (a) follows from the [Mas90, Theorem 2] and independence of the channel output sequence $\{R_{2t\text{-}1}, I_{2t}\}_{t=1}^{T}$; and (b) follows from the fact that $R_{2t\text{-}1}$ and I_{2t} are both sequences of i.i.d. variables. Comparing (3.59) and (3.55), the directed information rate is equal to the information rate which is due to independence of the channel output sequence.

Part II

Single system: cost-minimization

Chapter 4

Cascade Network

It is well-known that linear encoding is optimal for transmission of a Gaussian source over a point-to-point scalar Gaussian channel when the distortion measure is the mean squared error and when the source and channel bandwidths match [Jr.65, Ber71]. From [AG87, DT62, Fin65, BB87] we know that linear policies are also optimal if the encoder observes a noisy version of a Gaussian source. Moreover in [Gas08] Gastpar has shown that a linear (uncoded) scheme is even optimal in a simple Gaussian sensor network setting where each sensor node observes a noisy version of a Gaussian source and all the sensor nodes simultaneously transmit over a multiple-access Gaussian channel. For the problem of mean-square stabilization of a linear system over Gaussian networks, we concluded in Chapter 2 that linear policies can even be optimal when the source-channel matching principle does not hold.

In this chapter we show that linear policies are not optimal in general for the transmission of a Gaussian source over a Gaussian channel comprised of one or more relay nodes connected in cascade, when the distortion measure is the mean squared error. A special case of this problem was studied by Lipsa and Martins in [LM11] where they provided counterexamples based on binary quantizers to show that linear policies are not optimal when the number of number of relays are greater than or equal to three. Further we discuss that linear encoding policies are person-by-person optimal[1] for a two-hop relay channel. However they do not guarantee global optimality because the given team problem is concave in the encoding policies, as observed in [YL11]. We also derive a lower bound on distortion which is not tight in general.

This is a team decision problem under non-classical information structure [Ho80]. Some such problems are very difficult depending on the cost function and the information structure. The problem under study in this chapter is a variation of the Witsenhausen's problem [Wit68] with the addition of a relay encoder, which

[1] In a team decision problem, policies of decision makers are person-by-person optimal if there is no incentive for any decision maker to uni-laterally deviate from its policy given others' policies [Ho80].

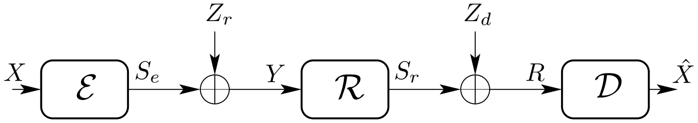


Figure 4.1: System model.

is unsolved till today.

The problem of causal transmission over a two-hop relay channel is motivated by control applications, where the sensor measurements of a dynamical system are transmitted via a relay node to a remote controller. Control of linear systems over various types of relay channels is studied in Chapter 3, where we used linear schemes to derive conditions on mean-square stability.

4.1 System Model

Consider a physical phenomenon characterized by a sequence of independent and identically distributed real valued Gaussian random variables $\{X_n\}_{n\in\mathbb{Z}_+}$ having zero mean and variance σ_x^2, where n denotes a discrete time index. We wish to instantly communicate this physical phenomenon to a remote destination over a two-hop relay channel with as high fidelity as possible. The system model is illustrated in Fig. 4.1. According to the figure, at a discrete time $n \in \mathbb{Z}_+$ the source encoder $\mathcal{E}$ observes X_n and produces $S_{e,n} = f_{1,n}(\{X_i\}_{i=1}^n)$ suitable for transmission, where $f_{1,n} : \mathbb{R}^n \mapsto \mathbb{R}$ is a causal measurable mapping. The mapping $f_{1,n}$ has to satisfy the following average power constraint,

$$\mathbb{E}[S_{e,n}^2] \leq P_S. \tag{4.1}$$

The signal $S_{e,n}$ is then observed in noise by the relay node $\mathcal{R}$ as $Y_n = S_{e,n} + Z_{r,n}$, where $\{Z_{r,n}\}_{n\in\mathbb{Z}_+}$ is a zero mean white Gaussian noise sequence of variance N_r. Since there is no direct link from the source encoder to the destination, we neglect transmission and processing delays at the relay, i.e., the relay node applies a causal mapping on the received signal $f_{2,n} : \mathbb{R}^n \mapsto \mathbb{R}$ to produce $S_{r,n} = f_{2,n}(\{Y_i\}_{i=1}^n)$ under the power constraint,

$$\mathbb{E}[S_{r,n}^2] \leq P_r. \tag{4.2}$$

The signal $S_{r,n}$ is then transmitted over a Gaussian channel. Accordingly the destination node $\mathcal{D}$ receives $R_n = S_{r,n} + Z_{d,n}$, where $\{Z_{d,n}\}_{n\in\mathbb{Z}_+}$ is a zero mean white Gaussian noise sequence of variance N_d. Upon receiving R_n the decoder wishes to reconstruct the transmitted variable X_n by applying a mapping $g_n : \mathbb{R}^n \mapsto \mathbb{R}$ to produce $\hat{X}_n = g_n(\{R_i\}_{i=1}^n)$. We define the signal-to-noise ratios of $\mathcal{E}$–$\mathcal{R}$ and $\mathcal{R}$–$\mathcal{D}$ links as $\gamma_r := P_S/N_r$ and $\gamma_d := P_r/N_d$ respectively. The encoder, the relay, and the decoder are all causal and delay-free (zero delay). The objective is to choose the encoder, relay, and decoder mappings such that following distortion

$$D = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}[(X_n - \hat{X}_n)^2] \tag{4.3}$$

is minimized subject to the constraints in (4.1) and (4.2).

4.2 Distortion Lower Bound

In this section we derive a lower bound on the distortion. Consider the following series of (in)equalities:

$$\begin{aligned}
\frac{1}{N}\sum_{n=1}^{N} I(X_n;\hat{X}_n) &\overset{(a)}{=} \frac{1}{N}\left(\sum_{n=1}^{N} h(X_n) - \sum_{n=1}^{N} h(X_n|\hat{X}_n)\right)\\
&\overset{(b)}{=} \frac{1}{N}\left(h(X_{[1,n]}) - \sum_{n=1}^{N} h(X_n|\hat{X}_n)\right)\\
&\overset{(c)}{\leq} \frac{1}{N}\left(h(X_{[1,n]}) - \sum_{n=1}^{N} h(X_n|\hat{X}_{[1,n]}, X_{[1,n\text{-}1]})\right)\\
&= \frac{1}{N}\left(h(X_{[1,n]}) - h(X_{[1,n]}|\hat{X}_{[1,n]})\right)\\
&= \frac{1}{N} I(X_{[1,n]};\hat{X}_{[1,n]})\\
&\overset{(d)}{\leq} \frac{1}{N} I(S_{e,[1,n]};R_{[1,n]})\\
&\overset{(e)}{\leq} \frac{1}{N} \min\{I(S_{e,[1,n]};Y_{[1,n]}), I(S_{r,[1,n]};R_{[1,n]})\}\\
&\overset{(f)}{\leq} \frac{1}{N} \min\left\{\sum_{n=1}^{N} I(S_{e,n};Y_n), \sum_{n=1}^{N} I(S_{r,n};R_n)\right\}\\
&\overset{(g)}{\leq} \frac{1}{2} \min\left\{\log\left(1+\frac{P_S}{N_r}\right), \log\left(1+\frac{P_r}{N_d}\right)\right\},
\end{aligned} \tag{4.4}$$

where (a) follows from the definition of mutual information; (b) follows from independence of the sequence X_n^N; (c) follows from conditioning reduces entropy; (d) and (e) follow from the data processing inequality [CT06, Theorem 2.8.1] with Markov chain $X^N - S_e^N - R^N - \hat{X}^N$; (f) follows from the fact that the channels are memoryless and conditioning reduces entropy; and (g) follows from the fact that mutual information is maximized by Gaussian distribution [CT06, Theorem 8.6.5].

Further consider the following inequalities:

$$\begin{aligned}\frac{1}{N}\sum_{n=1}^{N} I(X_n;\hat{X}_n) &\overset{(a)}{\geq} \frac{1}{2N}\sum_{n=1}^{N}\log\left(\frac{\sigma_x^2}{\mathbb{E}[(X_n-\hat{X}_n)^2]}\right)\\ &\overset{(b)}{\geq} \frac{1}{2}\log\left(\sigma_x^2\right)-\frac{1}{2}\log\left(\frac{1}{N}\sum_{n=1}^{N}\mathbb{E}\left[(X_n-\hat{X}_n)^2\right]\right), \end{aligned} \tag{4.5}$$

where (a) follows from the rate distortion theorem for an i.i.d. Gaussian source [CT06]; and (b) follows from the concavity of the logarithm function and Jensen's inequality. From (4.4) and (4.5), we obtain the following lower bound:

$$\begin{aligned} D &= \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\mathbb{E}\left[(X_n-\hat{X}_n)^2\right]\\ &\geq \sigma_x^2\max\left\{\frac{N_r}{P_S+N_r},\frac{N_d}{P_r+N_d}\right\}\\ &= \frac{\sigma_x^2}{1+\min\{\gamma_r,\gamma_d\}}. \end{aligned} \tag{4.6}$$

Remark 4.2.1. *The lower bound has been obtained without using causality constraints. Due to the presence of the two channel noise components (Z_r and Z_d), we have*

$$\max_{P_{S_e}} I(S_e;R) < \min\{\max_{P_{S_e}} I(S_e;Y), \max_{P_{S_r}} I(S_r;R)\}.$$

As a result of this strict inequality, the bound (4.6) *is not tight if we restrict the encoding policies to be memoryless*[2]*. This bound will be tight for memoryless policies when the variance of any of the two channel noise components is zero. This observation can also be made from [BB89]. In [YT07, Theorem 3.5] the authors discussed that $I(S_e;R)$ is strictly lower than the capacity of a two-hop relay channel which follows from block coding arguments and cut-set bound. This tells us that the distortion bounds obtained using cut-set arguments are not tight in general for relay networks with memoryless policies. Typically optimal causal mappings are memoryless for the transmission of an i.i.d. source [WV83], however it needs to be shown for the given two-hop relay channel.*

4.3 Linear Policies

In this section we find the optimal linear policies and show that linear policies are person-by-person optimal. Moreover, it is shown that person-by-person optimality of linear policies does not guarantee global optimality.

[2]A policy that only uses the current input at any time n.

4.3.1 Optimal Linear Encoding

Typically when a source is memoryless and the encoders are causal, the optimal encoders are memoryless [WV83]. This can be easily verified by showing that if we transmit a linear combination of the current and the previous source observations, then the previous observations will only contribute to noise as the source is memoryless. We therefore restrict our study to memoryless linear policies, in the sense that the encoders merely transmit a scaled version of the received signal. That is, the source and the relay encoders transmit the following signals:

$$S_{e,n} = \sqrt{\frac{a_n}{\sigma_x^2}} X_n, \quad S_{r,n} = \sqrt{\frac{b_n}{a_n + N_r}} Y_n,$$

where $a_n, b_n \in \mathbb{R}_+$ are time varying gain coefficients which are chosen such that the transmit power constraints in (4.1) and (4.2) are satisfied, i.e., $a_n \leq P_S$ and $b_n \leq P_r$. The decoder accordingly receives

$$R_n = \sqrt{\frac{a_n b_n}{\sigma_x^2(a_n + N_r)}} X_n + \sqrt{\frac{b_n}{a_n + N_r}} Z_{r,n} + Z_{d,n},$$

and computes the MMSE estimate according to $\hat{X}_n = \mathbb{E}[X_n|R^n] = \mathbb{E}[X_n|R_n]$, where we have used the fact that the $\{R_n, R^{n-k}\}$ are mutually independent for all $k \neq n$. Since X_n is Gaussian, the distortion is given by

$$\mathbb{E}[(X_n - \hat{X}_n)^2] = \sigma_x^2 \left(1 - \frac{a_n b_n}{(a_n + N_r)(b_n + N_d)}\right),$$

$$\Rightarrow D_L = \lim_{N\to\infty} \frac{\sigma_x^2}{N} \sum_{n=1}^{N} \left(1 - \frac{a_n b_n}{(a_n + N_r)(b_n + N_d)}\right). \tag{4.7}$$

The optimal choice of the gain coefficients $0 < a_n \leq P_S$, $0 < b_n \leq P_r$, which minimizes (4.7) is $\{a_n^\star = P_S, b_n^\star = P_r\}$. This choice of the gain coefficients leads to

$$D_L^\star = \sigma_x^2 \left(1 - \frac{1}{(1+\gamma_r)(1+\gamma_d)}\right). \tag{4.8}$$

We have so far found a strict lower bound on the distortion in (4.6) and an upper bound in (4.8) using the best linear scheme. However we still do not know how good linear policies are? In the following we show that the linear policies are person-by-personal optimal, however they do not guarantee global optimality.

4.3.2 Person-by-person Optimality of Linear Policies

Let us fix the source encoder to be linear. Given a linear and memoryless policy at the source encoder, we now find an optimal relaying policy which minimizes the distortion $\mathbb{E}[(X_n - \mathbb{E}[X_n|R_{[1,n]}])^2]$ at time n. We can rewrite the distortion at time n as

$$\begin{aligned}
&\mathbb{E}\left[\left(X_n - \mathbb{E}\left[X_n|R_{[1,n]}\right]\right)^2\right] \\
&\quad \overset{(a)}{=} \mathbb{E}\left[\left(X_n - \mathbb{E}\left[X_n|Y_{[1,n]}\right]\right)^2\right] + \mathbb{E}\left[\left(\mathbb{E}\left[X_n|Y_{[1,n]}\right] - \mathbb{E}\left[X_n|R_{[1,n]}\right]\right)^2\right] \\
&\quad \overset{(b)}{=} \mathbb{E}\left[(X_n - c_nY_n)^2\right] + \mathbb{E}\left[\left(c_nY_n - \mathbb{E}\left[X_n|R_{[1,n]}\right]\right)^2\right],
\end{aligned} \tag{4.9}$$

where (a) follows from $X_n - Y_{[1,n]} - R_{[1,n]}$ and $\mathbb{E}[(X_n - \mathbb{E}[X_n|Y_{[1,n]}])(\mathbb{E}[X_n|Y_{[1,n]}] - \mathbb{E}[X_n|R_{[1,n]}])] = 0$ (by the orthogonality principle); and (b) follows from the fact that the source encoder is linear and memoryless and the MMSE estimation of a Gaussian variable is linear, i.e., $\mathbb{E}[X_n|Y_{[1,n]}] = c_nY_n$ where c_n is a scalar. According to (4.9), the optimal relaying policy is the one which minimizes $E[(c_nY_n - E[X_n|R_{[1,n]}])^2]$ since the first term in the distortion function is independent of the relaying policy. This problem was studied in [BB87], from which it follows that an optimal relay encoding policy is linear and memoryless if we fix the source encoder to be linear memoryless. This observation can also be made from [GJ07, Yük10, Gas08, AG87, DT62, Fin65]. Now if we fix the relay encoder policy to be linear and memoryless, one can observe that the problem becomes equivalent to the transmission of a Gaussian source over a point to point Gaussian channel subject to an average power constraint, for which it is well-known that linear (memoryless) encoding is optimal in the sense of minimizing mean-squared distortion [Jr.65, Ber71]. Hence, linear encoding policies are person-by-person optimal.

4.3.3 Concavity of the Team Problem

In a decentralized team optimization problem person-by-person optimal solutions are globally optimal if the cost function is convex in the policies of the decision makers and the cost function satisfies differentiability conditions in the policies [Rad62]. Let P be an observation channel from the input variable X at the source encoder to the channel output variable R such that $P(\cdot|x)$ is a probability measure on the Borel σ-algebra $\mathcal{B}(\mathbb{R})$ on $\mathbb{R}$ for every $x \in \mathbb{R}$, and $P(A|\cdot) : \mathbb{R} \mapsto [0 : 1]$ is a Borel measurable function for every $A \in \mathcal{B}(\mathbb{R})$. That is, $P(\cdot|x)$ is a stochastic kernel. Similarly we define P_1 as an observation channel from the variable X to the variable Y, and P_2 as an observation channel from the variable Y to the variable R. From [YL11] it follows that the distortion in (4.3) is concave in the joint observation channel $P(A|x) = \int_{\mathbb{R}} P_2(A|y)P_1(dy|x)$ for every $A \in \mathcal{B}(\mathbb{R})$. If the encoding policies are viewed as stochastic kernels, then the individual observation channels P_1 and P_2 are given by convolutions of Gaussian distributions with the encoding policies, i.e., $P_1 = P_{S_e|X} * \mathcal{N}(0, N_r)$ and $P_2 = P_{S_r|Y} * \mathcal{N}(0, N_d)$. Since the distortion is concave in the joint channel and the individual channels are affine in the source and the relay encoding policies, the distortion is concave in the encoding policies $(P_{S_e|X}, P_{S_r|Y})$

and the optimal policies have to be deterministic if they exist. This implies that the person-by-person optimal encoding policies do not guarantee global optimality when policies are viewed as stochastic kernels [YL11].

4.4 Counterexample: Non-linear Policies

In this section we provide a simple counter example to show that linear policies are not optimal for causal transmission of a Gaussian source over the given two-hop relay channel. Consider the following time invariant policies at the source encoder and the relay encoder respectively:

$$f_1(x) = \left\{ \begin{array}{ccc} a, & \text{for} & x > m_1 \\ 0, & \text{for} & |x| \leq m_1 \\ -a, & \text{for} & x < -m_1 \end{array} \right\}, \tag{4.10}$$

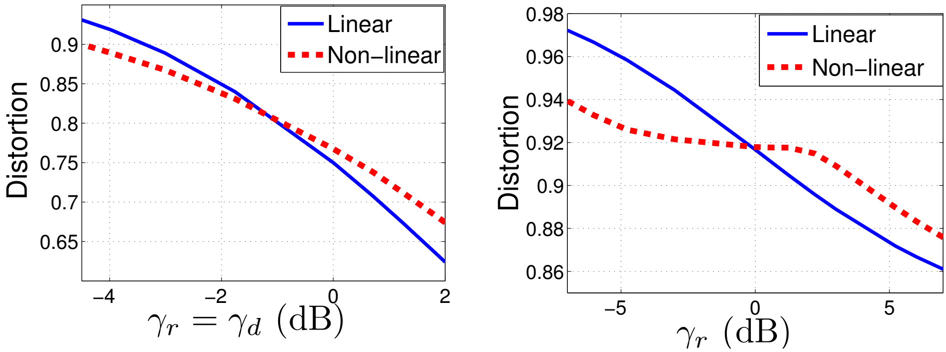
$$f_2(y) = \left\{ \begin{array}{ccc} b, & \text{for} & y > m_2 \\ 0, & \text{for} & |y| \leq m_2 \\ -b, & \text{for} & y < -m_2 \end{array} \right\}, \tag{4.11}$$

with $a, b \in \mathbb{R}_+$. In (4.10) and (4.11) we have dropped the index n for the sake of simplicity without any loss as we are considering time invariant policies. According to these policies, the signals observed at the relay and the destination are respectively given by

$$Y = \left\{ \begin{array}{ccc} a + Z_r, & \text{for} & X > m_1 \\ Z_r, & \text{for} & |X| \leq m_1 \\ -a + Z_r, & \text{for} & X < -m_1 \end{array} \right\}, \tag{4.12}$$

$$R = \left\{ \begin{array}{ccc} b + Z_d, & \text{for} & Y > m_2 \\ Z_d, & \text{for} & |Y| \leq m_2 \\ -b + Z_d, & \text{for} & Y < -m_2 \end{array} \right\}. \tag{4.13}$$

The policies in (4.10) and (4.11) have to satisfy the transmit power constraints in (4.1) and (4.2). In Appendix 4.A we have obtained conditions on $a, b \in \mathbb{R}_+$ to ensure these power constraints, which are given by (4.17) and (4.20) respectively. For these non-linear policies, the MMSE decoder $g(R)$ and the corresponding distortion D_{NL} are given in (4.27) and (4.28). We can numerically compute D_{NL} using (4.28), (4.17), and (4.20) for any fixed values of the parameters $\{\sigma_x^2, P_S, P_r, N_d, N_r, m_1, m_2\}$. We now give two examples to demonstrate that the proposed simple non-linear scheme can outperform the best linear scheme. In the following examples we fix the values of the system parameters and then numerically compute the distortion for non-linear and linear policies according to (4.28) and (4.8) respectively. We also evaluate the lower bound in (4.6), however the reader should keep in mind that the bound is not tight in general.



(a) $\sigma_x^2 = P_S = P_r = 1$ (b) $\sigma_x^2 = P_S = P_r = 1, N_d = 5$

Figure 4.2: Comparison of linear and non-linear schemes.

Example 1: Fixing $\sigma_x^2 = P_S = P_r = 1$, $N_r = N_d = 4$, $m_1 = 2.45$, $m_2 = 6.84$, $a = 8.36$, and $b = 9.25$ we get: $D_{NL} = 0.926$, $D_L^\star = 0.96$, and $D = 0.8$.
Example 2: Fixing $\sigma_x^2 = P_S = P_r = 1$, $N_r = N_d = 10$, $m_1 = 2.85$, $m_2 = 12.05$, $a = 15.12$, and $b = 16.25$ we get: $D_{NL} = 0.964$, $D_L^\star = 0.992$, and $D = 0.909$.
These examples validate that linear policies are not optimal in general for the given two-hop relay channel when the source and the relay node have individual power constraints. Let us now consider a total power, $\mathbb{E}[S_{e,n}^2] + \mathbb{E}[S_{r,n}^2] = P$. It can be easily shown that the distortion is minimized for the linear policies by an equal power allocation $\mathbb{E}[S_{e,n}^2] = \mathbb{E}[S_{r,n}^2] = \frac{P}{2}$ if $N_r = N_d$. In the above two counterexamples equal power constraints and noise variances are imposed on the source and relay nodes. Therefore, linear policies are still suboptimal even if a total transmit power constraint is imposed on the source and the relay, instead of separate power constraints. Note that for more than one relay nodes linked in cascade (multi-hop channel), the end-to-end distortion can be written as sum of distortions suffered at each hop since the source input is memoryless and by applying the orthogonality property, as we did for the two-hop case in (4.9). Therefore linear policies are sub-optimal for multi-hop relaying in general.

The proposed non-linear scheme is not always better than the optimal linear scheme as demonstrated in Fig. 4.2, where we have plotted distortion achieved with the non-linear and the optimal linear schemes as functions of signal-to-noise ratios for some fixed parameters. The non-linear scheme outperforms the linear scheme in low SNR regions, however there might exist better non-linear strategies which may outperform the linear strategy also in high SNR regions. When the channels are very noisy, one intuition on why the proposed non-linear strategy is superior may be that it does not amplify the large values of channel noise at its input unlike the linear (amplify-and-forward) strategy. We note that, in [LM11] a two-level quantizer was used when the number of relays were greater than two. In our setting the result also holds for a single relay, which generalizes and implies the results of [LM11]. The reason for selecting symmetric quantizers is due to the fact that symmetry in distribution is preserved when symmetric functions are applied to sources with symmetric distributions. Moreover with centering the quantizer at zero, the encoders can utilize the available transmit power in an efficient way by transmitting signals with power equal to zero more often.

4.5 Conclusion

We studied the problem of mean-square estimation of a Gaussian source over a two-hop Gaussian relay channel with average transmit power constraints. A lower bound on the distortion was derived. We observed that the distortion bounds obtained using cut-set arguments are not tight in general for sensor networks if we restrict policies to be memoryless. Further it was shown that linear policies are person-by-person optimal over the given two-hop relay channel. However person-by-person optimality of the linear policies does not guarantee global optimality due to concavity property of the distortion function in the observation channel. A simple three-level quantizer function was shown to outperform the best linear scheme in some cases, thus validating that linear policies are not optimal in general. This observation is in accordance with the previously known results for non-classical information structures [Ho80].

Appendix

4.A Transmit Power Constraints

The parameter $a \in \mathbb{R}_+$ in (4.10) is chosen such that

$$\begin{aligned} P_S &\geq \mathbb{E}\left[f_1^2(X)\right] \\ &\overset{(a)}{=} a^2\left(\int_{-\infty}^{-m_1} p(x)dx + \int_{m_1}^{\infty} p(x)dx\right) \\ &\overset{(b)}{=} 2a^2 Q\left(\frac{m_1}{\sigma_x^2}\right) \\ &\Rightarrow a \leq \sqrt{\frac{P_S}{2Q\left(\frac{m_1}{\sigma_x^2}\right)}}, \end{aligned} \tag{4.17}$$

where (a) follows from (4.10); (b) follows from $p(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{x^2}{2\sigma_x^2}}$, $Q(x) := \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{\tau^2}{2}} d\tau = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-x} e^{-\frac{\tau^2}{2}} d\tau$. We now compute $p(y)$ to find the condition on $b \in \mathbb{R}_+$ in (4.11) which ensures $\mathbb{E}\left[S_r^2\right] \leq P_r$. From (4.12) we have

$$p(y|x) = \left\{ \begin{array}{c} \frac{1}{\sqrt{2\pi N_r}} e^{-\frac{(y-a)^2}{2N_r}}, \text{if} \quad x > m_1 \\ \frac{1}{\sqrt{2\pi N_r}} e^{-\frac{y^2}{2N_r}}, \text{ if } \quad |x| \leq m_1 \\ \frac{1}{\sqrt{2\pi N_r}} e^{-\frac{(y+a)^2}{2N_r}}, \text{ if } \quad x < -m_1 \end{array} \right\}. \tag{4.18}$$

The marginal pdf $p(y)$ can now be computed as

$$\begin{aligned} p(y) &= \int_{\mathbb{R}} p(y|x)p(x)dx \\ &= \frac{1}{\sqrt{2\pi N_r}} \bigg(e^{\frac{-(y+a)^2}{2N_r}} \int_{-\infty}^{-m_1} p(x)dx + e^{\frac{-y^2}{2N_r}} \int_{-m_1}^{m_1} p(x)dx \\ &+ e^{\frac{-(y-a)^2}{2N_r}} \int_{m_1}^{\infty} p(x)dx \bigg) \\ &= \frac{1}{\sqrt{2\pi N_r}} \left[\left(e^{\frac{-(y+a)^2}{2N_r}} + e^{\frac{-(y-a)^2}{2N_r}} \right) Q\left(\frac{m_1}{\sigma_x^2}\right) + e^{\frac{-y^2}{2N_r}} \left(1 - 2Q\left(\frac{m_1}{\sigma_x^2}\right) \right) \right]. \end{aligned} \tag{4.19}$$

The condition on b is obtained as follows:

$$\begin{aligned} P_r &\geq \mathbb{E}\left[f_2^2(Y)\right] \\ &\overset{(a)}{=} b^2 \left(\int_{-\infty}^{-m_2} p(y)dy + \int_{m_2}^{\infty} p(y)dy \right) \\ &\overset{(b)}{=} 2b^2 \int_{-\infty}^{-m_2} p(y)dy \\ &\overset{(c)}{=} 2b^2 \Bigg[Q\left(\frac{m_1}{\sigma_x^2}\right) \left\{ \frac{1}{\sqrt{2\pi N_r}} \int_{-\infty}^{-m_2} e^{\frac{-(y+a)^2}{2N_r}} dy + \frac{1}{\sqrt{2\pi N_r}} \int_{-\infty}^{-m_2} e^{\frac{-(y-a)^2}{2N_r}} dy \right\} \\ &+ \left(1 - 2Q\left(\frac{m_1}{\sigma_x^2}\right) \right) \left\{ \frac{1}{\sqrt{2\pi N_r}} \int_{-\infty}^{-m_2} e^{\frac{-y^2}{2N_r}} dy \right\} \Bigg] \\ &\overset{(d)}{=} 2b^2 \left[Q\left(\frac{m_1}{\sigma_x^2}\right) \left(Q\left(\frac{m_2-a}{\sqrt{N_r}}\right) + Q\left(\frac{m_2+a}{\sqrt{N_r}}\right) \right) + Q\left(\frac{m_2}{\sqrt{N_r}}\right) \left(1 - 2Q\left(\frac{m_1}{\sigma_x^2}\right) \right) \right] \\ &\overset{(e)}{=}: 2b^2 \kappa(m_1, m_2, a, \sigma_x^2, N_r) \\ &\Rightarrow b \leq \sqrt{\frac{P_r}{2\kappa(m_1, m_2, a, \sigma_x^2, N_r)}}, \end{aligned} \tag{4.20}$$

where (a) follows from (4.11); (b) follows from symmetry of $p(y)$ around origin; (c) follows by substituting $p(y)$ from (4.19); (d) follows by the definition of $Q(\cdot)$ function; and (e) follows by defining $\kappa(m_1, m_2, a, \sigma_x^2, N_r)$.

4.B Distortion Calculation

Since we have the following Markov chain $X - Y - R$, $p(x,r)$ is given by

$$p(x,r) = \int_{\mathbb{R}} p(r|y)p(y|x)p(x)dy, \tag{4.21}$$

where $p(r|y)$ follows from (4.13), that is

$$p(r|y) = \left\{ \begin{array}{ll} \frac{1}{\sqrt{2\pi N_d}} e^{-\frac{(r-b)^2}{2N_d}}, & \text{if} \quad y > m_2 \\ \frac{1}{\sqrt{2\pi N_d}} e^{-\frac{r^2}{2N_d}}, & \text{if} \quad |y| \leq m_2 \\ \frac{1}{\sqrt{2\pi N_d}} e^{-\frac{(r+b)^2}{2N_d}}, & \text{if} \quad y < -m_2 \end{array} \right\}. \tag{4.22}$$

From (4.18) we see that $p(y|x)$ is defined over the three disjoint intervals of x (i.e., $x < m_1, |x| \leq m_1, x > m_1$). Due to this, the joint pdf $p(x,r)$ is also defined over these three intervals. For the interval $x < -m_1$,

$$\begin{aligned} p(x,r) &\overset{(a)}{=} \frac{p(x)}{\sqrt{2\pi N_d}} \Bigg[\frac{e^{\frac{-(r+b)^2}{2N_d}}}{\sqrt{2\pi N_r}} \int_{-\infty}^{-m_2} e^{\frac{-(y+a)^2}{2N_r}} dy + \frac{e^{\frac{-r^2}{2N_d}}}{\sqrt{2\pi N_r}} \int_{-m_2}^{m_2} e^{\frac{-(y+a)^2}{2N_r}} dy \\ &\quad + \frac{e^{\frac{-(r-b)^2}{2N_d}}}{\sqrt{2\pi N_r}} \int_{m_2}^{\infty} e^{\frac{-(y+a)^2}{2N_r}} dy \Bigg] \\ &= \frac{p(x)}{\sqrt{2\pi N_d}} \Bigg[e^{\frac{-(r+b)^2}{2N_d}} Q\left(\frac{m_2 - a}{\sqrt{N_r}}\right) + e^{\frac{-(r-b)^2}{2N_d}} Q\left(\frac{m_2 + a}{\sqrt{N_r}}\right) \\ &\quad + e^{\frac{-r^2}{2N_d}} \left\{ 1 - Q\left(\frac{m_2 - a}{\sqrt{N_r}}\right) - Q\left(\frac{m_2 + a}{\sqrt{N_r}}\right) \right\} \Bigg] \\ &\overset{(b)}{=:} p(x)l_1(r), \end{aligned} \tag{4.23}$$

where (a) follows from (4.21) and (4.18) and (4.22); and (b) follows by defining $l_1(r)$. Similarly for $|x| \leq m_1$,

$$\begin{aligned} p(x,r) &= \frac{p(x)}{\sqrt{2\pi N_d}} \left[e^{\frac{-r^2}{2N_d}} + \left\{ e^{\frac{-(r+b)^2}{2N_d}} + e^{\frac{-(r-b)^2}{2N_d}} - 2e^{\frac{-r^2}{2N_d}} \right\} Q\left(\frac{m_2}{\sqrt{N_r}}\right) \right] \\ &=: p(x)l_2(r), \end{aligned} \tag{4.24}$$

and for $x > m_1$,

$$p(x,r) = \frac{p(x)}{\sqrt{2\pi N_d}} \Bigg[e^{\frac{-(r+b)^2}{2N_d}} Q\left(\frac{m_2 + a}{\sqrt{N_r}}\right) + e^{\frac{-r^2}{2N_d}} \left\{ 1 - Q\left(\frac{m_2 - a}{\sqrt{N_r}}\right) - Q\left(\frac{m_2 + a}{\sqrt{N_r}}\right) \right\}$$

$$+ e^{\frac{-(r-b)^2}{2N_d}} Q\left(\frac{m_2 - a}{\sqrt{N_r}}\right)\Bigg]$$
$$=: p(x)l_3(r). \tag{4.25}$$

From (4.23), (4.24), and (4.25), we compute

$$\begin{aligned} p(r) &= \int_{\mathbb{R}} p(x,r)dx \\ &= l_1(r)\int_{-\infty}^{-m_1} p(x)dx + l_2(r)\int_{-m_1}^{m_1} p(x)dx + l_3(r)\int_{m_1}^{\infty} p(x)dx \\ &= (l_1(r) + l_3(r))\, Q\left(\frac{m_1}{\sigma_x^2}\right) + l_2(r)\left(1 - 2Q\left(\frac{m_1}{\sigma_x^2}\right)\right). \end{aligned} \tag{4.26}$$

The MMSE estimator can now be computed using (4.23), (4.24), (4.25), and (4.26) as follows.

$$\begin{aligned} \mathbb{E}[X|R=r] &= \frac{1}{p(r)}\int xp(x,r)dx \\ &= \frac{1}{p(r)}\left(l_1(r)\int_{-\infty}^{-m_1} xp(x)dx + l_2(r)\int_{-m_1}^{m_1} xp(x)dx + l_3(r)\int_{m_1}^{\infty} xp(x)dx\right) \\ &\overset{(a)}{=} \frac{1}{p(r)}\,(l_3(r) - l_1(r))\int_{m_1}^{\infty} xp(x)dx \\ &= \frac{1}{p(r)}\,(l_3(r) - l_1(r))\sqrt{\frac{\sigma_x^2}{2\pi}}\exp\left(-\frac{{m_1}^2}{2\sigma_x^2}\right) \\ &=: g(r), \end{aligned} \tag{4.27}$$

where (a) follows from $\int_{-m_1}^{m_1} xp(x)dx = 0$. The associated mean-squared error is

$$\begin{aligned} D_{NL} &:= \int_{\mathbb{R}^2} (x - g(r))^2\, p(x,r)d(x,r) \\ &= \int_{-\infty}^{\infty}\Bigg(l_1(r)\int_{-\infty}^{-m_1}(x - g(r))^2\, p(x)dx + l_2(r)\int_{-m_1}^{m_1}(x - g(r))^2\, p(x)dx \\ &\quad + l_3(r)\int_{m_1}^{\infty}(x - g(r))^2\, p(x)dx\Bigg)dr. \end{aligned} \tag{4.28}$$

Chapter 5

Parallel Network

This chapter studies a scenario where the state of a linear dynamical system is observed in noise by several sensors that convey their measurements over orthogonal wireless channels to a remote controller. We model communication links as white Gaussian channels and restrict our study to scalar LTI systems driven by Gaussian noise. The objective is to design sensing and control schemes that minimize a quadratic function of the state variables.

For a single sensor setup, a linear sensing and control strategy has been shown to be optimal for control of a scalar-valued system (plant) over a scalar Gaussian channel by Bansal and Başar in [BB89]. This single sensor setup has been briefly discussed in Sec. 1.2. Furthermore in [YT09], Yüksel and Tatikonda showed via a counter-example that linear schemes are not globally optimal when there are multiple sensors transmitting their signals to a remote controller over parallel Gaussian channels. In this chapter we show that linear schemes are not even person-by-person optimal for a parallel network of two sensors, unlike the two-hop cascade network where linear schemes were shown to be person-by-person optimal in Chapter 4. We propose a non-linear sensing and control scheme for closed-loop control over parallel Gaussian channels using multiple sensors. The proposed scheme is instantaneous (i.e., delay-free) and it can be implemented with reasonable complexity. Furthermore, we show that the suggested approach is robust to the uncertainty in the knowledge of the powers of the measurement and the channel noises, which is generally a crucial aspect in designing practical systems.

5.1 Problem Formulation

In the earlier chapters (Chapter 2 and Chapter 3) we considered LTI systems in absence of measurement noise. In this chapter, we consider a more genral setup where sensors observe noisy measurements of a system's state. Consider the following scalar discrete time LTI system (plant):

$$X_{t+1} = \lambda X_t + U_t + W_t, \quad t \in \mathbb{N} \tag{5.1}$$

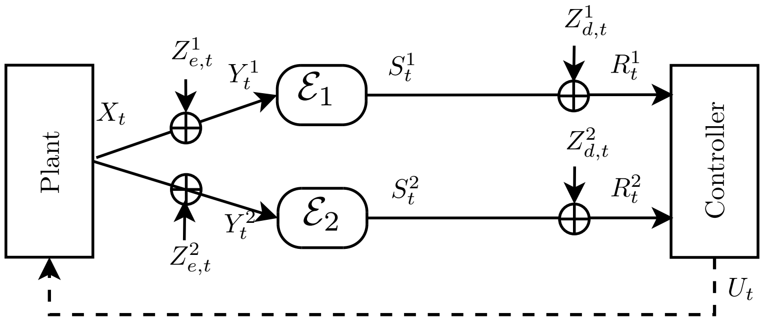


Figure 5.1: Control over a parallel sensor network.

where X_t is the state process of the system, U_t is the control process, and W_t is the process noise. The initial state X_0 is an unknown random variable drawn according to a zero-mean Gaussian distribution with variance σ_x^2. The process noise W_t is assumed to be i.i.d. zero-mean Gaussian distributed with variance n_w. For the sake of simplicity, we consider a two sensor setup shown in Fig. 5.1. The state X_t is observed in noise by the sensors $\mathcal{E}_1$ and $\mathcal{E}_2$ as

$$Y_t^i = X_t + Z_{e,t}^i, \quad i = 1, 2, \tag{5.2}$$

where $Z_{e,t}^1$ and $Z_{e,t}^2$ are two i.i.d. measurement noise components, which are Gaussian distributed with zero means and variances N_e^1 and N_e^2 respectively. Based on their noisy observations, the two sensors transmit the following signals,

$$S_t^i = f_{i,t}(Y_t^i), \quad i = 1, 2, \tag{5.3}$$

subject to the following power constraints:

$$E[(S_t^i)^2] \leq P_i, \qquad i = 1, 2. \tag{5.4}$$

Accordingly the remote controller receives

$$R_t^i = S_t^i + Z_{d,t}^i, \quad i = 1, 2, \tag{5.5}$$

where $Z_{d,t}^i$, $i = 1, 2$, are mutually independent zero-mean Gaussian noise variables with powers N_d^1 and N_d^2 respectively. We have assumed orthogonal channels from the sensors to the controller, therefore there is no interference between the two received signals (i.e., we have two parallel Gaussian channels from the sensors to the controller). Based on the received signals, the controller takes an action $U_t = \pi_t\left(R_{[0,t]}^1, R_{[0,t]}^2\right)$. The objective is to minimize the following finite horizon quadratic cost function:

$$J_T = \mathbb{E}\left[\sum_{t=1}^{T} X_t^2\right], \tag{5.6}$$

where the expectation is taken over the initial state X_0, the process noise W_t, the measurement noise $Z_{e,t}^i$, and the channel noise $Z_{d,t}^i$.

5.2 Sensing and Control Schemes

In this section we propose a non-linear distributed sensing and control scheme. For the sake of reference, we first give a linear sensing scheme.

5.2.1 Linear Sensing Scheme

The reference scheme is the well-known amplify-and-forward strategy, in which the sensor nodes amplify the received signals subject to average power constraints and then transmit them to the remote controller. The transmitted signals from the two sensors are given by

$$S_t^1 = \eta_{1,t} Y_t^1, \tag{5.7}$$

$$S_t^2 = \eta_{2,t} Y_t^2, \tag{5.8}$$

where the parameters $\eta_{1,t}, \eta_{2,t}$ are chosen such that the power constraints (5.4) are fulfilled. The optimal control scheme is given by $U_t = -a\hat{X}_t$, where the state estimate $\hat{X}_t$ is generated using a Kalman filter [Ber05].

5.2.2 Non-linear Sensing Scheme

The non-linear distributed sensing scheme works as follows. At any time t, the signals transmitted by the two sensors are given by,

$$S_t^1 = \eta_{1,t} Y_t^1, \tag{5.9}$$

$$S_t^2 = \eta_{2,t} \left(Y_t^2 - \Delta_t \left\lfloor \frac{Y_t^2}{\Delta_t} \right\rceil \right), \tag{5.10}$$

where $\lfloor \cdot \rceil$ denotes rounding to the nearest integer. A pictorial illustration of this non-linear scheme is given in Fig. 5.2. The parameter Δ_t controls the length of each period in the periodic sawtooth function shown in Fig. 5.2. The values $\Delta_{[0,t]}$ are chosen such that the cost function J_T in (5.6) is minimized. The procedure of choosing $\Delta_{[0,t]}$ is given in Sec. 5.2.4. The parameters $\{\eta_{1,t}, \eta_{2,t}, \Delta_{[0,t]}\}$ are chosen such that the average transmit power constraints (5.4) are met.

This non-linear scheme was first introduced in [WS09b] for the transmission of a Gaussian source over parallel Gaussian channels. Here we have extended this scheme for closed-loop control of a dynamical system driven Gaussian noise over parallel Gaussian channels.

5.2.3 Control Scheme for the Non-linear Sensing Scheme

The controller is assumed to have a separation structure where it first computes an estimate of the state $\hat{X}_t$ and then takes action using the state estimate, i.e., $U_t = -\lambda\hat{X}_t$. Since the computation of optimal MMSE state estimate based on all

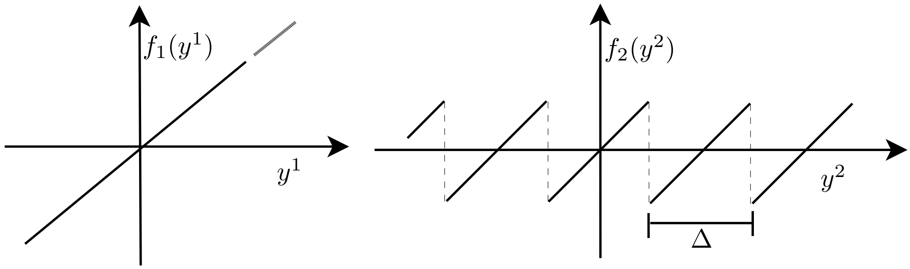


Figure 5.2: Non-linear distributed sensing scheme.

previously received signals $\left\{R^1_{[0,t]}, R^2_{[0,t]}\right\}$ is not practical, we propose the following sub-optimal algorithm.

1. Compute estimates $\tilde{X}_{0|t}, \dots, \tilde{X}_{t|t}$ of $X_0, \dots, X_t$, where $\tilde{X}_{k|t}$ denotes estimate of X_k at time t based on the previous state estimate $\hat{X}_{t-1}$ and R^1_t using a Kalman filter cf. Kalman Filter 1 in the Fig. 5.3.

2. Assume that $|(\tilde{X}_{s|t} - Y^2_s - Z^2_{d,s})/\eta_s| \leq \Delta_s/2\ \forall s$ and compute the Maximum Likelihood estimates $\hat{Y}^2_s$ as (cf. ML decoder in Fig. 5.3):

$$\hat{Y}^2_s = \text{argmin}_{Y_s \in \mathcal{Y}}((S^2(Y_s) - R^2_s)^2), \tag{5.11}$$

   where $\mathcal{Y} = \{Y_s : |\tilde{X}_{s|t} - Y_s| \leq \eta_s \Delta_s/2\}$.

3. Finally assume that the estimates $\hat{Y}^2_s$ had been linearly encoded (multiplied by η^t_0) and find the estimate $\hat{X}_t$ from a Kalman filter using $\{R^1_{[0,t]}, \eta_{[0,t]}, U_{[0,t-1]}\}$ and $U_{[0,t-1]}$ as input cf. Kalman Filter 2 in Fig. 5.3.

5.2.4 Choosing Δ_t.

We propose the following procedure to choose the parameters $\{\Delta_t\}$ in the sawtooth sensor mapping:

- For time step $t = 0$, we choose Δ_0 so as to minimize $E[(\hat{X}_0 - X_0)^2]$. This is done using methods similar to the ones in [WS09a].

- For time step $t = 1$, we fix Δ_0 to the value found in step 1, and simulate the system up to $t = 1$ for different values of Δ_1. We then choose the value of Δ_1 that minimizes $E[(\hat{X}_1 - X_1)^2]$.

- Similarly for any time $t = s$, we fix $\Delta_0, \dots, \Delta_{s-1}$ to the values found in the previous steps and simulate the system up to $t = s$ for different Δ_s. We then choose the Δ_s that minimizes $E[(\hat{X}_s - X_s)^2]$.

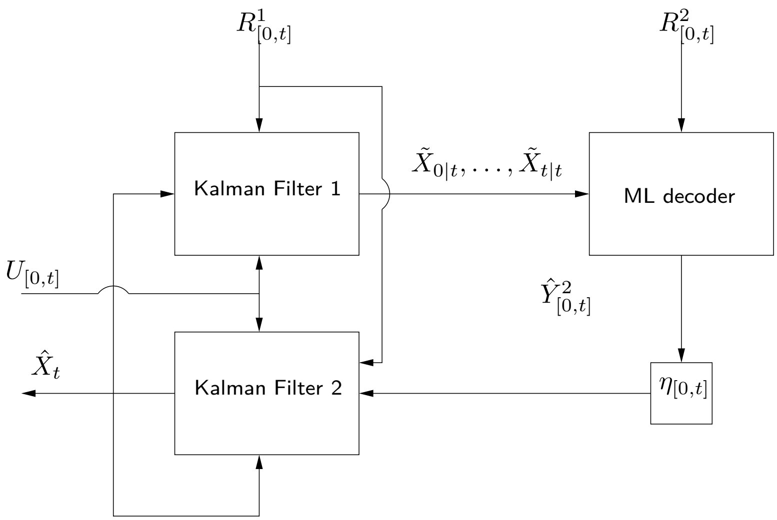


Figure 5.3: State estimator for the non-linear distributed sensing scheme.

5.3 Performance Analysis

Let Z_t denote the estimation error in $\hat{X}_t$, i.e.

$$\hat{X}_t = X_t + Z_t. \tag{5.12}$$

Now we can rewrite the state equation (5.1) as

$$\begin{aligned} X_{t+1} &= \lambda X_t + U_t + W_t \\ &= -\lambda Z_t + W_t, \end{aligned} \tag{5.13}$$

which follows from $U_t = -\lambda \hat{X}_t$. The cost function (5.6) is then given by

$$\begin{aligned} L_T &= E\left[\sum_{t=1}^{T} X_t^2\right] \\ &= E\left[\sum_{t=0}^{T-1} (-\lambda Z_t + W_t)^2\right] \\ &= \sum_{t=0}^{T-1} \left(\lambda^2 \sigma_{Z_t}^2 + n_w\right), \end{aligned} \tag{5.14}$$

which shows that the performance of the system will depend on the following two terms: $\lambda^2\sigma_{Z_t}^2$ and n_w. The second term n_w arises from the process noise which we will not be able to affect. The first term $\lambda^2\sigma_{Z_t}^2$ arises from the estimation error Z_t.

Hence, improving the estimation of X_t, i.e., lowering $\sigma^2_{Z_t}$ will lower the value of the cost function as previously stated. However, one might also suspect that the induced distribution of Z_t affects the cost L_T since it changes the distribution of X_t. This is only implicitly true, the created MSE $E[(X_t - \hat{X}_t)^2]$ of the proposed scheme in (5.9)–(5.10) will, approximately, not depend on the distribution of X_t given a fixed system S^1 and S^2, see [WS09a]. However, the distribution of X_t will affect $P(\mathbf{\Delta}_0^t, t)$ which in turn will affect the scaling parameter of the output and this parameter naturally affects the MSE.

The impact of the process noise W_t on our system is interesting. A high process noise variance n_w will directly make the objective function larger as seen above, but it will also make the correlation between the two sensor measurements larger. If we have no process noise then $X_t = -\lambda Z_{t-1}$ which, given good encoders and decoders, will be small. Thus, the impact of measurement noise will be higher, taking away the correlation between Y_t^1 and Y_t^2 and most of the benefits of non-linear transmission scheme. For low process noise we would thus expect that using non-linear encoders at time $t = 0$ and linear encoders for the other time steps (corresponding to $\Delta_t \to \infty$ for $t > 0$) will be optimal.

The impact of the channel noise variances N_d^i on the performance was studied in [WS09a]. There it was shown that either low or high N_d will lead to a linear system working as well as the nonlinear. In order to verify our claims, we perform numerical simulations for different values of the system parameters.

Numerical Simulations:

In Figures 5.4–5.6 we present results from three simulations. For all simulations, we choose the system parameter $\lambda = 1.2$, the initial state variance $\sigma_x^2 = 5$ and the time-horizon $T = 3$. Further we only consider the cases with equal power constraints at the sensors and equal noise variances, i.e., $N_e^1 = N_e^2 = N_e$, $N_d^1 = N_d^2 = N_d$, and $P_1 = P_2$. All optimized values for $\Delta_{[0,T-1]}$ can be found in Table 5.1.

The results from the first simulation are shown in Figure 5.4. Here we vary the channel noise variance N_d while keeping the other parameters fixed. For the optimal nonlinear and the linear curves we have optimized the encoders and decoder for the actual SNR used in the simulation. For the mismatched nonlinear curve we used the encoder optimized for SNR = 9 dB, but the decoder used the true SNR of the channel. This was done in order to see how robust the encoders are to SNR mismatch. We see that the nonlinear system gives a power gain over the linear system with up to 2 dB, and also that the system is very robust to SNR mismatch.

In the second simulation we use the same parameters as in the first simulation, but here the process noise variance is $n_w = 3$. The mismatched system is again optimized for SNR = 9 dB. For this simulation we also see a 2 dB gain, and again the system is robust to SNR mismatch.

The results from the third simulation are shown in Figure 5.6. Here we keep the channel noise variance N_d fixed, while instead varying the measurement noise variance N_e. As in the first two simulations we show results for both optimized and

Table 5.1: Optimized values of $\Delta_{[0,T-1]}$ for different choices of system and channel parameters.

SNR (dB)	N_e	n_w	Δ_0	Δ_1	Δ_2
6	0.001	1	10.6	6.2	4.2
9	0.001	1	5.6	2.4	2.4
12	0.001	1	4	1.8	1.8
15	0.001	1	3	1.4	1.4
18	0.001	1	2.2	1	1
21	0.001	1	1.6	1	1
6	0.001	3	10.5	7.5	6.75
9	0.001	3	5.5	4.5	4.5
12	0.001	3	4	3	3
15	0.001	3	3	2.5	2.5
18	0.001	3	2.25	1.75	1.75
21	0.001	3	1.75	1.25	1.25
10	0.01	0	5	∞	∞
10	0.06	0	5.6	∞	∞
10	0.11	0	6.2	∞	∞
10	0.16	0	6.8	∞	∞
10	0.21	0	7.2	∞	∞
10	0.26	0	7.8	∞	∞
10	0.31	0	8.4	∞	∞
10	0.36	0	8.8	∞	∞
10	0.41	0	9.2	∞	∞
10	0.46	0	9.8	∞	∞
10	0.51	0	10.2	∞	∞

mismatched $\Delta_{[0,T-1]}$. The mismatched $\Delta_{[0,T-1]}$ were optimized for $N_e = 0.16$.

5.4 Conclusions

We proposed a distributed non-linear sensing scheme for closed-loop control system with two sensors measuring the plant's state. The proposed sensing and transmission scheme is delay-free, robust to the knowledge of noise statistics at the sensors, and can be implemented with reasonable complexity. The non-linear sensing has been shown to significantly outperform the best linear strategy. Conceptually, this

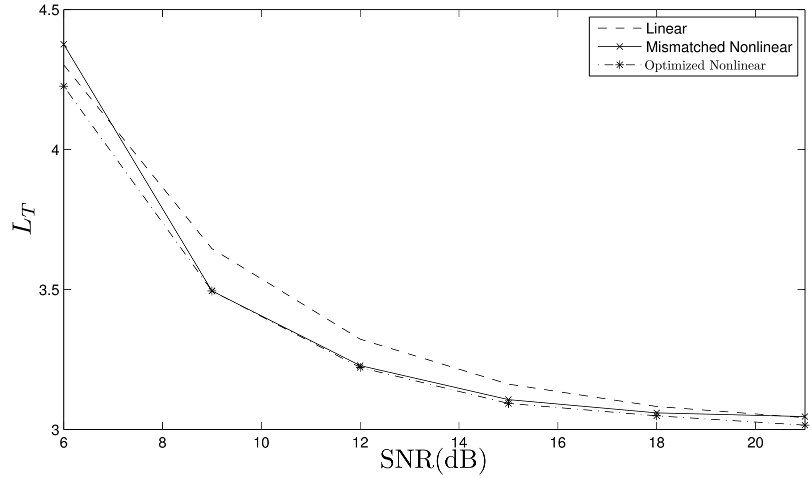


Figure 5.4: Systems with $N_e = 0.001$ and $n_w = 1$

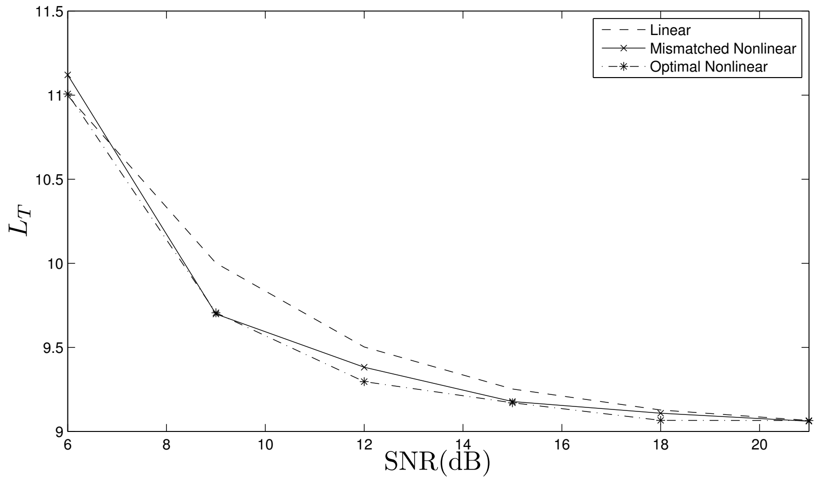


Figure 5.5: Systems with $N_e = 0.001$ and $n_w = 3$

scheme can be easily generalized to an arbitrary number of sensors by employing a linear mapping at the first sensor node and sawtooth mappings at the remaining sensor nodes with successively decreasing time periods Δ_t. How the number of sensor nodes will affect the system performance compared to the best linear scheme is

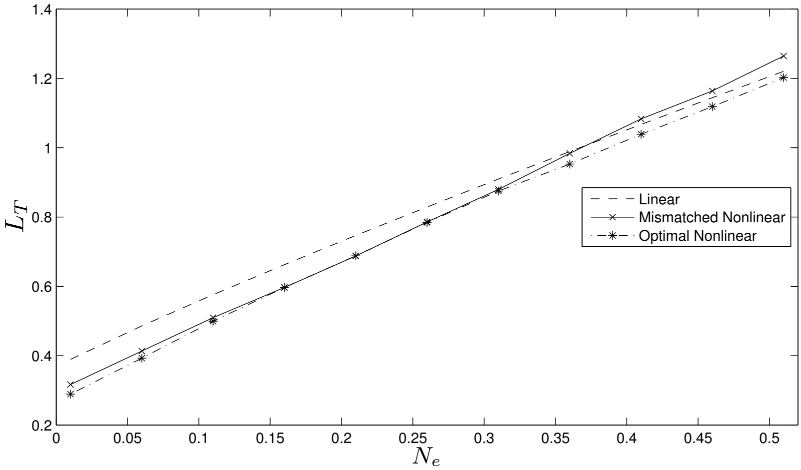


Figure 5.6: System with SNR = 10 dB and $n_w = 0$

yet to be studied.

Part III

Multiple systems: stability

Chapter 6

Multiple-access and Broadcast Channels

6.1 Introduction

In the previous chapters we focused on control of a single LTI system over various communication network settings, with stabilization and cost-minimization as the two main objectives. This chapter, together with Chapter 7, investigates control of multiple plants over basic multi-user Gaussian channels such as multiple-access channel, broadcast channel, and interference channel. In this chapter we focus on the problem of mean-square stabilization of two LTI plants over two-user multiple-access and broadcast channels. The two or more user multiple-access channel is the communication channel where two or more sources transmit their messages to a common destination [CT06]. By the stabilization over the multiple-access channel we mean that there exist two separate sensors to sense the states and a single remote controller to stabilize the two plants i.e., a multi-sensor joint controller setup.The capacity region of the two-user memoryless Gaussian multiple-access channel with noiseless feedback is found in [Oza84], which is relevant to the problem of closed-loop control over the multiple-access channel. The two or more user broadcast channel is the communication channel where one sender transmits messages to two or more destinations [CT06]. Stabilization over broadcast channel refers to a joint sensor multi-controller setup i.e., there exists a common sensor to jointly observe the states of the two plants and there are two separate remote controllers in order to stabilize them. The capacity region of the Gaussian broadcast channel with and without feedback is not known [CT06]. For the problem of closed-loop control, the broadcast channel with feedback is more relevant. In [OLYC84] Ozarow et al. provided an achievable rate region over the two user memoryless Gaussian broadcast channel with noiseless feedback which is highly relevant to the problem considered in this chapter. The coding schemes proposed by Ozarow et al. in [Oza84, OLYC84] for the memoryless Gaussian multiple-access and the broadcast channels with noiseless are extensions of the well-known Schalkwijk-Kailath coding scheme for memoryless Gaussian point-to-point communication channel with noiseless feedback [SK66].

In Chapter 2 and Chapter 3, we used linear schemes schemes for deriving sufficient conditions for closed-loop stabilization of a single LTI system over various Gaussian network topologies. This chapter now considers extends the study to two LTI plants that have to be remotely stabilized over the white Gaussian multiple-access and the white Gaussian broadcast channels. We use linear and memoryless communication and control schemes based on Ozarow's coding schemes [Oza84, OLYC84] to derive regions which are sufficient for mean square stability of the two plants.

6.2 Problem Setup

Consider two scalar noise-free LTI systems whose state equations are given by

$$X_{t+1}^i = \lambda_i X_t^i + U_t^i, \quad i = 1, 2, \tag{6.1}$$

where $X_t^i \in \mathbb{R}$ and $U_t^i \in \mathbb{R}$ are state and control processes of the plant i. We assume that the open-loop systems are unstable ($\lambda_i > 1$) and the initial states X_0^i are random variables with arbitrary probability distributions having second moment $\alpha_{i,0} = \mathbb{E}[(X_0^i)^2]$ and correlation coefficient $\rho_0 = \frac{\mathbb{E}[X_0^1 X_0^2]}{\sqrt{\alpha_{1,0}\alpha_{2,0}}}$. We study the problem of remotely controlling the two unstable systems over white Gaussian broadcast and multiple-access channels.

6.2.1 Control Over Multiple-access Channel

The setup for control over a multiple-access channel is depicted in Fig. 6.1. There are separate observers $\mathcal{O}_1$ and $\mathcal{O}_2$ for the two plants, and there is a common control unit $\mathcal{C}$ situated at remote location. In order to communicate the observed state values to the controller, an encoder $\mathcal{E}_i$ is lumped with $\mathcal{O}_i$ and a decoder $\mathcal{D}$ is lumped with the controller. At any time instant t, the encoders $\mathcal{E}_1$ and $\mathcal{E}_2$ transmit S_t^1 and S_t^2 respectively, and the decoder $\mathcal{D}$ receives $R_t = S_t^1 + S_t^2 + Z_t$, where $Z_t \sim \mathcal{N}(0, N)$ is the white noise component. Let $f_{i,t}$ denote the observer/encoder policy for the plant i, then we have $S_t^i = f_{i,t}(\{X_k^i\}_{k=0}^t)$ which must satisfy an average power constraint $\lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} \mathbb{E}[(S_t^i)^2] \le P_i$. Further let $\gamma_{i,t}$ denote the decoder/controller policy, then $U_t^i = \gamma_{i,t}(\{R_k\}_{k=0}^t)$.

6.2.2 Control Over Broadcast Channel

The setup for control over a broadcast channel is depicted in Fig. 6.2. There is a common observer $\mathcal{O}$ and separate controllers $\mathcal{C}_1$ and $\mathcal{C}_2$ for the two plants. In order to communicate the observed state values to the controllers, an encoder $\mathcal{E}$ is lumped with the observer and the decoders $\mathcal{D}_1$ and $\mathcal{D}_2$ are lumped with the respective controllers. At any time instant t, the encoder transmits S_t and the decoder $\mathcal{D}_i$ receives $R_t^i = S_t + Z_t + Z_t^i$, where $Z_t^i \sim \mathcal{N}(0, N_i)$ and $Z_t \sim \mathcal{N}(0, N)$ are the mutually independent white noise components. The noise component Z_t in the broadcast

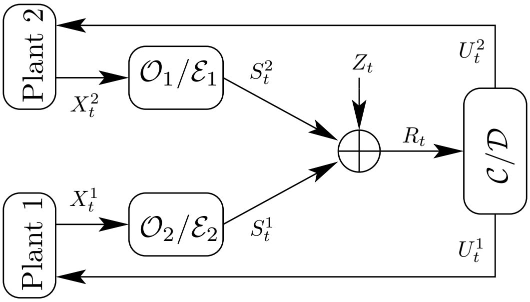


Figure 6.1: The two unstable LTI plants have to be controlled over the white Gaussian multiple access channel. There are two sensors to separately sense the states of the two plants and there is a remote common control unit.

channel can model a common noise or interference in the two signals. Let f_t denote the observer/encoder policy, then we have $S_t = f_t(\{X_{1,k}\}_{k=0}^t, \{X_{2,k}\}_{k=0}^t)$ which must satisfy an average power constraint $\lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} \mathbb{E}[S_t^2] \leq P$. Further let $\gamma_{i,t}$ denote the decoder/controller policy, then $U_t^i = \gamma_{i,t}(\{R_{i,k}\}_{k=0}^t)$.

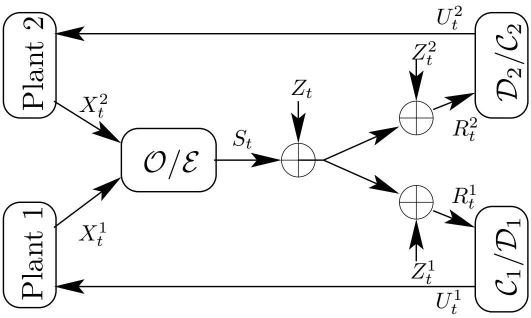


Figure 6.2: The two unstable LTI plants have to be controlled over the white Gaussian broadcast channel. There is a common sensor to jointly sense the states of the two plants and there are remotely located separate control units.

6.2.3 Mean-square Stability

For a noise-free plant, we modify Definition 2.1.1 (definition of mean square stability for a noisy plant) as follows.

Definition 6.2.1. *A noiseless system is said to be* mean square stable *if and only if*

$$\lim_{t\to\infty} \mathbb{E}[X_t^2] = 0,$$

regardless of the initial state X_0.

6.3 Main Results

We will first present our results in a comprehensive fashion and then provide the proofs in the next section.

6.3.1 Stability Results for the Multiple-access Channel

Theorem 6.3.1. *The two scalar LTI systems in (6.1) can be mean square stabilized over the memoryless white Gaussian multiple access channel if the systems' parameters $\{\lambda_1, \lambda_2\}$ satisfy the following inequalities*

$$\begin{aligned} \log(\lambda_1) &< \frac{1}{2}\log\left(1 + \frac{P_1\left(1-\rho^{\star 2}\right)}{N}\right), \\ \log(\lambda_2) &< \frac{1}{2}\log\left(1 + \frac{P_2\left(1-\rho^{\star 2}\right)}{N}\right), \end{aligned} \tag{6.2}$$

where $\rho^\star$ is the root in the open interval $(0,1)$ of the following fourth order polynomial

$$\left(P_1\left(1-\rho^2\right)+N\right)\left(P_2\left(1-\rho^2\right)+N\right) = \left(P_1 + P_2 + 2\rho\sqrt{P_1P_2} + N\right)N. \tag{6.3}$$

Proof. The proof is given in Sec. 6.4.1. □

Remark 6.3.1. *It can be shown that for fully correlated initial states, i.e., $\rho_0 = 1$, the stability conditions are given by*

$$\log(\lambda_i) < \frac{1}{2}\log\left(1 + \frac{P_1 + P_2 + 2\sqrt{P_1P_2}}{N}\right), \quad i = 1, 2.$$

Remark 6.3.2. *The terms on the right hand side in (6.2) correspond to the sum-rate optimal achievable rate pair for the two sources over the white Gaussian multiple-access channel with noiseless feedback [Oza84]. However, the stability region in (6.2) is smaller than the capacity region in [Oza84]. This is because to ensure second moment stability the coding scheme has to have exponential error decay.*

6.3.2 Stability Results for the Broadcast Channel

In the broadcast channel there is a joint encoder with an output power constraint contrary to the multiple-access channel where the two encoders have individual power constraints. Therefore the joint encoder in the broadcast channel has freedom to tradeoff between the powers allocated to the transmission of the observed states of the two plants.

Theorem 6.3.2. *The two scalar LTI systems in (6.1) can be mean square stabilized over the memoryless white Gaussian broadcast channel if the systems' parameters* $\{\lambda_1, \lambda_2\}$ *satisfy the following inequalities*

$$\begin{aligned} \log(\lambda_1) &< \frac{1}{2}\log\left(\frac{D^\star(N+N_1+P)}{D^\star(N+N_1)+g^2P(1-\rho^\star)}\right), \\ \log(\lambda_2) &< \frac{1}{2}\log\left(\frac{D^\star(N+N_2+P)}{D^\star(N+N_2)+P(1-\rho^\star)}\right), \end{aligned} \tag{6.4}$$

where $D^\star = 1+g^2+2g\rho^\star$*,* $g \geq 0$*, and* $\rho^\star$ *is the largest root in the open interval* $(0,1)$ *of the following polynomial*

$$\rho = -\frac{\big(D(N\Sigma+N_1N_2)\rho - gP\Sigma(1-\rho^2)\big)}{\sqrt{(\Pi\,(D(N+N_1)+g^2P(1-\rho^2))\,(D(N+N_2)+g^2P(1-\rho^2)))}} \tag{6.5}$$

where $\Pi = (P+N+N_1)(P+N+N_2)$ *and* $\Sigma = P+N+N_1+N_2$*.*

Proof. The proof is given in Sec. 6.4.2. □

Remark 6.3.3. *The terms on the right hand side in (6.4) is an achievable rate pair for the two decoders over the white Gaussian broadcast channel with noiseless feedback [OLYC84].*

Remark 6.3.4. *If the noise components* Z_t^1 *and* Z_t^2 *are zero in the broadcast channel model, then the two controllers receive the same signal and this setup is equivalent to having a joint controller. Therefore the stability region for the joint-sensor joint-controller case can be obtained by setting* $N_1 = N_2 = 0$ *in (6.4).*

The parameter g in (6.4) can tradeoff between the stabilizability of the two plants and thus we can obtain a stability region for the given channel parameters by increasing g from zero to less than infinity. Fig. 6.3 shows some examples of stability regions for $P = 10$. The solid line shows the boundary of the stability region when $N = 0$ and $N_1 = N_2 = 1$, the dashed line shows the boundary of the stability region when $N = N_1 = N_2 = 0.5$, and the dotted line shows the boundary of the stability region when $N = 1$ and $N_1 = N_2 = 0$. In these examples $N+N_i = 1$, and we can observe that the individual noise components $\{Z_t^1, Z_t^2\}$ are less harmful than the common noise component Z_t due to diversity effect. For comparison we also show the stability region when the encoder separately serves the two plants in

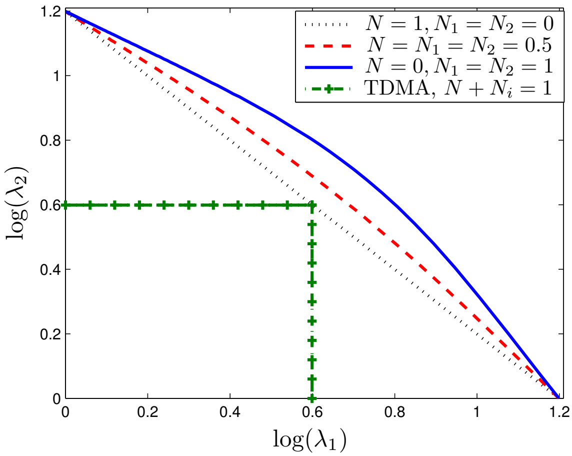


Figure 6.3: Illustration of the stability regions for the broadcast channel.

alternate time steps, i.e., in each time step there is a point-to-point communication link from the encoder to one of the controllers. For this case the necessary and sufficient conditions for mean square stability can be found in [ZOS10a], which are given by

$$\log(\lambda_i) < \frac{1}{4} \log\left(1 + \frac{P}{N + N_i}\right) \quad \text{for all } i \in \{1, 2\}.$$

The boundary of the rectangular stability region defined by the above inequalities is shown in Fig. 6.3 for $P = 10$ and $N + N_i = 1$.

6.4 Proofs

In order to prove Theorems 6.3.1 and 6.3.2, we propose to use coding schemes in [Oza84] and [OLYC84]. These schemes are based on Schalkwijk–Kailath coding scheme [SK66]. By employing the proposed coding schemes over the given broadcast and multiple-access channels, we then find conditions on the system parameters $\{\lambda_1, \lambda_2\}$ which are sufficient to mean square stabilize the systems in (6.1).

6.4.1 Proof of Theorem 6.3.1

The scheme for the white Gaussian multiple-access channel works as follows.

Initial time steps, $t = 0, 1$

Similar to the schemes for the channels (point-to-point and relay channels) studied in the previous chapters, we perform an initialization to make the plants' states Gaussian distributed. Initially, the two encoders transmit the observed state values in alternate time slots to the respective controllers. The first two disjoint transmissions in time make the plants' states Gaussian distributed regardless of the distribution of their initial states, which will be explained shortly. However if the initial states are already Gaussian, then the following disjoint initial transmissions are not needed.

At time step $t = 0$, the encoder $\mathcal{E}_1$ observes X_0^1 and transmits $S_0^1 = \sqrt{\frac{P_1}{\alpha_{1,0}}} X_0^1$. The encoder $\mathcal{E}_2$ stays quiet, i.e., $S_0^2 = 0$. The decoder $\mathcal{D}$ receives $R_0 = S_0^1 + Z_0$. It then estimates X_0^1 as

$$\begin{aligned}\hat{X}_0^1 &= \sqrt{\frac{\alpha_{1,0}}{P_1}} R_0 \\ &= X_0^1 + \sqrt{\frac{\alpha_{1,0}}{P_1}} Z_0.\end{aligned}$$

The controller $\mathcal{C}$ then takes an action $U_0^1 = -\lambda_1 \hat{X}_0^1$ for the plant 1, which results in $X_1^1 = \lambda_1(X_0^1 - \hat{X}_0^1)$. The state $X_1^1 \sim \mathcal{N}(0, \alpha_{1,1})$ with $\alpha_{1,1} = \lambda_1^2 \frac{\alpha_{1,0} N}{P_1}$. The controller does not take any action for the plant 2, therefore $X_1^2 = \lambda_2 X_0^2$ with $\alpha_{2,1} = \lambda_2^2 \alpha_{2,0}$.

At time step $t = 1$, the encoder $\mathcal{E}_1$ stays quiet. The encoder $\mathcal{E}_2$ observes X_1^2 and transmits $S_1^2 = \sqrt{\frac{P_2}{\alpha_{2,1}}} X_1^2$. The decoder $\mathcal{D}$ receives $R_1 = S_1^2 + Z_1$. It then estimates X_1^2 as

$$\begin{aligned}\hat{X}_1^2 &= \sqrt{\frac{\alpha_{2,1}}{P_2}} R_1 \\ &= X_1^2 + \sqrt{\frac{\alpha_{2,1}}{P_2}} Z_1.\end{aligned}$$

The controller $\mathcal{C}$ then takes an action $U_1^2 = -\lambda_2 \hat{X}_1^2$ for the plant 2, which results in $X_2^2 = \lambda_2(X_1^2 - \hat{X}_1^2)$. The state $X_2^2 \sim \mathcal{N}(0, \alpha_{2,2})$. For the plant 1, the controller does not take any action $U_1^1 = 0$, therefore $X_2^1 = \lambda_1 X_1^1$ and $X_2^1 \sim \mathcal{N}(0, \alpha_{1,2})$.

It is noteworthy that due to non-overlapping initial transmissions by the two encoders, the states X_2^1 and X_2^2 are now zero mean Gaussian variables with correlation coefficient $\rho_2 = \frac{\mathbb{E}[X_2^1 X_2^2]}{\sqrt{\alpha_{1,2}\alpha_{2,2}}}$ equal to zero[1]. Henceforth the two encoders will transmit their signals simultaneously.

[1]The states in the second time step become uncorrelated irrespective of the value of the correlation between the initial states. This scheme does not exploit correlation between the initial states and thus the stability region obtained is independent of the correlation of the initial states.

Further time steps $t \geq 2$

The two encoders $\mathcal{E}_1$ and $\mathcal{E}_2$ observe X_t^1 and X_t^2, and they respectively transmit

$$S_t^1 = \sqrt{\frac{P_1}{\alpha_{1,t}}} X_t^1,$$
$$S_t^2 = \sqrt{\frac{P_2}{\alpha_{2,t}}} X_t^2 \text{sgn}(\rho_t),$$

where $\rho_t = \frac{\mathbb{E}\left[(X_t^1 - \mathbb{E}[X_t^1])(X_t^2 - \mathbb{E}[X_t^2])\right]}{\sqrt{\alpha_{1,t}\alpha_{2,t}}}$ and $\text{sgn}(\rho_t) = 1$ if $\rho_t \geq 0$ and $\text{sgn}(\rho_t) = -1$ if $\rho_t < 0$.
The decoder $\mathcal{D}$ receives $R_t = S_t^1 + S_t^2 + Z_t$. It then computes an MMSE estimate of the state of the plant i as

$$\begin{aligned} \hat{X}_t^i &= \mathbb{E}\left[X_t^i | R_t\right] \\ &\overset{(a)}{=} \frac{\mathbb{E}\left[R_t X_t^i\right]}{\mathbb{E}\left[(R_t)^2\right]} R_t, \end{aligned} \tag{6.6}$$

where (a) follows from the fact that the optimum MMSE of the Gaussian variable is linear [Hay96]; and we have

$$\begin{aligned} \mathbb{E}\left[X_t^1 R_t\right] &= \sqrt{\alpha_{1,t}} \left(\sqrt{P_1} + \sqrt{P_2}|\rho_t|\right), \\ \mathbb{E}[X_t^2 R_t] &= \sqrt{\alpha_{2,t}} \left(\sqrt{P_2} + \sqrt{P_1}|\rho_t|\right) \text{sgn}(\rho_t), \\ \mathbb{E}[R_t^2] &= P_1 + P_2 + 2|\rho_t|\sqrt{P_1 P_2} + N. \end{aligned} \tag{6.7}$$

The controller $\mathcal{C}$ takes an action $U_t^i = -\lambda_i \hat{X}_t^i$ for the plant i, which results in $X_{t+1}^i = \lambda_i (X_t^i - \hat{X}_t^i)$. The mean values of the states are

$$\begin{aligned} \mathbb{E}\left[X_t^i\right] &= \mathbb{E}\left[\lambda_i \left(X_t^i - \hat{X}_t^i\right)\right] \\ &\overset{(a)}{=} \lambda_i \mathbb{E}\left[X_t^i - \frac{\mathbb{E}\left[R_t X_t^i\right]}{\mathbb{E}\left[(R_t)^2\right]} R_t\right] \\ &\overset{(b)}{=} 0, \end{aligned} \tag{6.8}$$

where (a) follows from (6.6); and (b) follows from $\mathbb{E}[X_{i,2}] = 0$ and by recursively using (a). The variance of the state X_{t+1}^i is given by

$$\alpha_{i,t+1} := \mathbb{E}\left[(X_{t+1}^i)^2\right] = \lambda_i^2 \mathbb{E}\left[\left(X_t^i - \hat{X}_t^i\right)^2\right]$$

$$= \lambda_i^2 \mathbb{E}\left[\left(X_t^i - \frac{\mathbb{E}\left[R_t X_t^i\right]}{\mathbb{E}\left[(R_t)^2\right]} R_t\right)^2\right]$$

$$= \lambda_i^2 \left(\mathbb{E}\left[\left(X_t^i\right)^2\right] - \frac{\left(\mathbb{E}\left[R_t X_t^i\right]\right)^2}{\mathbb{E}\left[(R_t)^2\right]}\right). \tag{6.9}$$

By using (6.7) in (6.9) we get the following recursive equations

$$\alpha_{1,t+1} = \alpha_{1,t}\lambda_1^2 \left(\frac{N + P_2(1-\rho_t^2)}{P_1 + P_2 + 2|\rho_t|\sqrt{P_1P_2} + N}\right) \tag{6.10}$$

$$\alpha_{2,t+1} = \alpha_{2,t}\lambda_2^2 \left(\frac{N + P_1(1-\rho_t^2)}{P_1 + P_2 + 2|\rho_t|\sqrt{P_1P_2} + N}\right) \tag{6.11}$$

The cross-correlation between the states is given by

$$\mathbb{E}[X_{t+1}^1 X_{t+1}^2] = \mathbb{E}\left[\lambda_1\left(X_t^1 - \hat{X}_t^1\right)\lambda_2\left(X_t^2 - \hat{X}_t^2\right)\right]$$

$$\stackrel{(a)}{=} \lambda_1\lambda_2\left(\mathbb{E}\left[X_t^1X_t^2\right] - \frac{\mathbb{E}\left[X_t^1R_t\right]\mathbb{E}\left[X_t^2R_t\right]}{\mathbb{E}\left[(R_t)^2\right]}\right)$$

$$\stackrel{(b)}{=} \lambda_1\lambda_2\sqrt{\alpha_{1,t}\alpha_{2,t}}\left(\frac{N\rho_t - \operatorname{sgn}(\rho_t)\sqrt{P_1P_2}(1-\rho_t^2)}{P_1 + P_2 + 2|\rho_t|\sqrt{P_1P_2} + N}\right), \tag{6.12}$$

where (a) follows from $\mathbb{E}[\hat{X}_t^1 X_t^2] = \mathbb{E}[\hat{X}_t^2 X_t^1] = \mathbb{E}[\hat{X}_t^1 \hat{X}_t^2] = \frac{\mathbb{E}[X_t^1 R_t]\mathbb{E}[X_t^1 R_t]}{\mathbb{E}[R_t^2]}$; and (b) follows from (6.7). The correlation coefficient is then given by

$$\rho_{t+1} = \frac{\mathbb{E}\left[X_{t+1}^1 X_{t+1}^2\right]}{\sqrt{\alpha_{1,t}\alpha_{2,t}}}$$

$$\stackrel{(a)}{=} \lambda_1\lambda_2\sqrt{\frac{\alpha_{1,t}\alpha_{2,t}}{\alpha_{1,t+1}\alpha_{2,t+1}}}\left(\frac{N\rho_t - \operatorname{sgn}(\rho_t)\sqrt{P_1P_2}(1-\rho_t^2)}{P_1 + P_2 + 2|\rho_t|\sqrt{P_1P_2} + N}\right)$$

$$\stackrel{(b)}{=} \frac{N\rho_t - \operatorname{sgn}(\rho_t)\sqrt{P_1P_2}\left(1-\rho_t^2\right)}{\sqrt{\left(N + P_1(1-\rho_t^2)\right)\left(N + P_2\left(1-\rho_t^2\right)\right)}} \quad \forall t \geq 2, \tag{6.13}$$

where (a) follows from (6.12); and (b) follows from (6.10) and (6.11). It has been shown in [Oza84] that for (6.13) there exists a $\rho^\star$ such that if $\rho_t = \rho^\star$ then $\rho_{t+k} = (-1)^k\rho^\star$ for all $k \geq 0$, where $\rho^\star$ is the root in the open interval $(0,1)$ of the following fourth order polynomial.

$$\left(P_1\left(1-\rho^2\right) + N\right)\left(P_2\left(1-\rho^2\right) + N\right) = \left(P_1 + P_2 + 2\rho\sqrt{P_1P_2} + N\right)N. \tag{6.14}$$

If we modify the control actions such that ρ_2 becomes equal to $\rho^\star$ instead of zero, then ρ_t will be equal to $(-1)^t\rho^\star$ for all $t \geq 2$. Suppose in the time step $t = 1$ the controller takes the actions $U_1^1 = m$ and $U_1^2 = -\lambda_2 \hat{X}_1^2 + m$, where m is a Gaussian variable with zero mean and variance σ_m^2. By varying σ_m^2 the correlation coefficient ρ_2 can be made equal to any value between zero and one. Therefore by choosing σ_m^2 such that $\rho_2 = \rho^\star$, we can rewrite (6.10) and (6.11) as

$$\begin{aligned}\alpha_{i,t} &= \alpha_{i,2}\left(\lambda_i^2 \frac{N + P_i(1-\rho^\star)}{P_1 + P_2 + 2|\rho^\star|\sqrt{P_1P_2} + N}\right)^{t-2} \\ &= \alpha_{i,2}\left(\lambda_i^2 \frac{N}{N + P_i(1-\rho^\star)}\right)^{t-2}, \end{aligned} \tag{6.15}$$

where the last equality follows from (6.14). We observe from (6.15) that $\alpha_{i,t} \to 0$ as $t \to \infty$ if

$$\begin{aligned}&\left(\lambda_i^2 \frac{N}{N + P_i(1-\rho^\star)}\right) < 1 \\ &\Rightarrow \log(\lambda_i) < \frac{1}{2}\log\left(1 + \frac{P_i\left(1-\rho^{\star 2}\right)}{N}\right),\end{aligned}$$

which completes the proof. □

6.4.2 Proof of Theorem 6.3.2

The communication and control scheme for the white Gaussian broadcast channel is in principle similar to that of the multiple-access channel where in the beginning the encoder separately transmit the states of the two plants in order to make them Gaussian. Thereafter the Gaussian distributed states are transmitted jointly. This scheme works as follows.

Initial time steps, $t = 0, 1$

In the first two time steps the encoder transmits state observations of each plant separately. These separate initial transmissions make plant states Gaussian distributed regardless of the distribution of their initial states. However if the initial states are already Gaussian, then the following separate initial transmissions are not needed.

At time step $t = 0$ the encoder ignores X_0^2 and transmits X_0^1 as $S_0 = \sqrt{\frac{P}{\alpha_{1,0}}} X_0^1$. The decoder $\mathcal{D}_1$ receives $R_0^1 = S_0 + Z_0 + Z_0^1$. It then estimates X_0^1 as

$$\begin{aligned}\hat{X}_0^1 &= \sqrt{\frac{\alpha_{1,0}}{P}} R_0^1 \\ &= X_0^1 + \sqrt{\frac{\alpha_{1,0}}{P}}(Z_0 + Z_0^1).\end{aligned}$$

The controller $\mathcal{C}_1$ then takes an action $U_{1,0} = -\lambda_1 \hat{X}_0^1$ for the plant 1, which results in $X_1^1 = \lambda_1(X_0^1 - \hat{X}_0^1)$, where $X_1^1 \sim \mathcal{N}(0, \alpha_{1,1})$ with $\alpha_{1,1} = \lambda_1^2 \frac{\alpha_{1,0}(N+N_1)}{P}$. The controller $\mathcal{C}_2$ does not take any action for the plant 2, therefore $X_1^2 = \lambda_2 X_0^2$ with $\alpha_{2,1} = \lambda_2^2 \alpha_{2,0}$.

At time step $t = 1$ the encoder $\mathcal{E}$ ignores X_1^1 and transmits only X_1^2, i.e., $S_1 = \sqrt{\frac{P}{\alpha_{2,1}}} X_1^2$. The decoder $\mathcal{D}_2$ receives $R_1^2 = S_1 + Z_1 + Z_1^2$. It then estimates X_1^2 as

$$\begin{aligned}\hat{X}_1^2 &= \sqrt{\frac{\alpha_{2,1}}{P}} R_1^2 \\ &= X_1^2 + \sqrt{\frac{\alpha_{2,1}}{P}} (Z_1 + Z_1^2).\end{aligned}$$

The controller $\mathcal{C}_2$ then takes an action $U_{2,1} = -\lambda_2 \hat{X}_1^2$ for the plant 2, which results in $X_2^2 = \lambda_2(X_1^2 - \hat{X}_1^2)$, where $X_2^2 \sim \mathcal{N}(0, \alpha_{2,2})$. The controller $\mathcal{C}_1$ does not take any action for the plant 1, i.e., $U_{1,1} = 0$, therefore $X_2^1 = \lambda_1 X_1^1$ and $X_2^1 \sim \mathcal{N}(0, \alpha_{1,2})$.

Similar to the multiple-access channel, the states X_2^1 and X_2^2 are now zero mean Gaussian variables with correlation coefficient $\rho_2 = \frac{\mathbb{E}[X_2^1 X_2^2]}{\sqrt{\alpha_{1,2}\alpha_{2,2}}}$ equal to zero[2]. Henceforth the encoder will serve both plants simultaneously.

Further time steps, $t \geq 2$

The encoder $\mathcal{E}$ observes X_t^1 and X_t^2, and it transmits

$$S_t = \sqrt{\frac{P}{D_t}} \left(\frac{X_t^1}{\sqrt{\alpha_{1,t}}} + g \frac{X_t^2}{\sqrt{\alpha_{2,t}}} \operatorname{sgn}(\rho_t) \right), \tag{6.16}$$

where $D_t = 1 + g^2 + 2g|\rho_t|$, $g \geq 0$, $\rho_t = \frac{\mathbb{E}\left[(X_t^1 - \mathbb{E}[X_t^1])(X_t^2 - \mathbb{E}[X_t^2])\right]}{\sqrt{\alpha_{1,t}\alpha_{2,t}}}$, and $\operatorname{sgn}(\rho_t) = 1$ if $\rho_t \geq 0$ and $\operatorname{sgn}(\rho_t) = -1$ if $\rho_t < 0$.
The decoder $\mathcal{D}_i$ receives $R_t^i = S_t + Z_t + Z_t^i$. It then computes an MMSE estimate of the state of the plant i as

$$\begin{aligned}\hat{X}_t^i &= \mathbb{E}\left[X_t^i | \{R_{i,k}\}_{k=0}^t\right] \\ &\stackrel{(a)}{=} \mathbb{E}\left[X_t^i | R_t^i\right] \\ &\stackrel{(b)}{=} \frac{\mathbb{E}\left[R_t^i X_t^i\right]}{\mathbb{E}\left[\left(R_t^i\right)^2\right]} R_t^i,\end{aligned} \tag{6.17}$$

where (a) follows from $\mathbb{E}[X_t^i R_{i,k}] = 0$ for all $k < t$ and X_t^i and R_t^i are Gaussian variables; (b) follows from the fact that the optimum MMSE of the Gaussian variable

[2]The states in the second time step become uncorrelated irrespective of the value of the correlation between the initial states. This scheme does not exploit correlation between the initial states and thus the stability region obtained is independent of the correlation of the initial states.

is linear [Hay96]; and we have

$$\mathbb{E}\left[X_t^1 R_t^1\right] = \sqrt{\frac{P\alpha_{1,t}}{D_t}}\left(1 + g|\rho_t|\right),$$
$$\mathbb{E}[X_t^2 R_t^2] = \sqrt{\frac{P\alpha_{2,t}}{D_t}}\left(\rho_t + g\,\mathrm{sgn}\left(\rho_t\right)\right),$$
$$\mathbb{E}[\left(R_t^i\right)^2] = P + N + N_i. \tag{6.18}$$

The controller $\mathcal{C}_i$ takes an action $U_t^i = -\lambda_i \hat{X}_t^i$ for the plant i, which results in $X_{t+1}^i = \lambda_i(X_t^i - \hat{X}_t^i)$. The mean values of the states are

$$\begin{aligned}\mathbb{E}\left[X_{t+1}^i\right] &= \mathbb{E}\left[\lambda_i\left(X_t^i - \hat{X}_t^i\right)\right]\\ &\overset{(a)}{=} \lambda_i \mathbb{E}\left[X_t^i - \frac{\mathbb{E}\left[R_t^i X_t^i\right]}{\mathbb{E}[\left(R_t^i\right)^2]} R_t^i\right]\\ &\overset{(b)}{=} 0,\end{aligned}$$

where (a) follows from (6.17); and (b) follows from $\mathbb{E}[X_{i,2}] = 0$ and by recursively using (a). The variance of the state X_{t+1}^i is given by

$$\begin{aligned}\alpha_{i,t+1} := \mathbb{E}\left[\left(X_{t+1}^i\right)^2\right] &= \lambda_i^2 \mathbb{E}\left[\left(X_t^i - \hat{X}_t^i\right)^2\right]\\ &= \lambda_i^2\left(\mathbb{E}\left[\left(X_t^i\right)^2\right] - \frac{\left(\mathbb{E}\left[R_t^i X_t^i\right]\right)^2}{\mathbb{E}[(R_t^i)^2]}\right).\end{aligned} \tag{6.19}$$

By using (6.18) in (6.19) we get the following recursive equations.

$$\alpha_{1,t+1} = \alpha_{1,t}\lambda_1^2\left(\frac{D_t(N+N_1) + g^2 P(1-\rho_t^2)}{D_t(P+N+N_1)}\right) \tag{6.20}$$
$$\alpha_{2,t+1} = \alpha_{2,t}\lambda_2^2\left(\frac{D_t(N+N_2) + P(1-\rho_t^2)}{D_t(P+N+N_2)}\right). \tag{6.21}$$

The cross-correlation between the states is given by

$$\begin{aligned}\mathbb{E}\left[X_{t+1}^1 X_{t+1}^2\right] &= \mathbb{E}\left[\lambda_1\left(X_t^1 - \hat{X}_t^1\right)\lambda_2\left(X_t^2 - \hat{X}_t^2\right)\right]\\ &= \lambda_1\lambda_2\left(\mathbb{E}\left[X_t^1 X_t^2\right] - 2\mathbb{E}\left[\hat{X}_t^1 X_t^2\right] + \mathbb{E}\left[\hat{X}_t^1 \hat{X}_t^2\right]\right)\\ &\overset{(a)}{=} \lambda_1\lambda_2\left(\frac{\mathbb{E}\left[X_t^1 X_t^2\right]\Pi - \mathbb{E}\left[R_t^1 X_t^1\right]\mathbb{E}\left[R_t^2 X_t^2\right]\Sigma}{\Pi}\right)\\ &\overset{(b)}{=} \lambda_1\lambda_2\sqrt{\alpha_{1,t}\alpha_{2,t}}\left(\rho_t - \frac{P}{D_t\Pi}\left(\rho_t + g\left|\rho_t\right|\rho_t + g\,\mathrm{sgn}(\rho_t) + g\rho_t\right)\Sigma\right),\end{aligned} \tag{6.22}$$

where (a) follows from $\mathbb{E}[\hat{X}_t^1 X_t^2] = \frac{\mathbb{E}[R_t^1 X_t^1]\mathbb{E}[R_t^2 X_t^2]}{P+N+N_1}$, $\mathbb{E}[X_t^1 \hat{X}_t^2] = \frac{\mathbb{E}[R_t^1 X_t^1]\mathbb{E}[R_t^2 X_t^2]}{P+N+N_2}$, $\mathbb{E}[\hat{X}_t^1 \hat{X}_t^2] = \frac{\mathbb{E}[R_t^1 X_t^1]\mathbb{E}[R_t^2 X_t^2](P+N)}{(P+N+N_1)(P+N+N_2)}$, $\Pi \triangleq (P+N+N_1)(P+N+N_2)$, and $\Sigma \triangleq (P+N+N_1+N_2)$; and (b) follows from (6.18). Now we can write a recursive equation for the correlation coefficient ρ_t by using (6.20), (6.21) and (6.22), as

$$\begin{aligned}\rho_{t+1} &= \frac{\mathbb{E}\left[X_{t+1}^1 X_{t+1}^2\right]}{\sqrt{\alpha_{1,t+1}\alpha_{2,t+1}}} \\ &= \frac{\left(D_t\left(N\Sigma + N_1N_2\right)\rho_t - gP\Sigma\left(1-\rho_t^2\right)\operatorname{sgn}(\rho_t)\right)}{\sqrt{\left(\Pi\left(D_t\left(N+N_1\right)+g^2P\left(1-\rho_t^2\right)\right)\left(D_t\left(N+N_2\right)+g^2P\left(1-\rho_t^2\right)\right)\right)}}\end{aligned}$$

It has been shown in [OLYC84] that for the above recursive equation there exists a $\rho^\star$ such that if $\rho_t = \rho^\star$ then $\rho_{t+k} = (-1)^k\rho^\star$ for all $k \geq 0$, where $\rho^\star$ is the largest root in the open interval $(0,1)$ of the polynomial given in (6.5). If we modify our encoding scheme such that ρ_2 becomes equal to $\rho^\star$ instead of zero, then ρ_t will be equal to $(-1)^t\rho^\star$ for all $t \geq 2$. Suppose in the initial transmissions (i.e., $t = 0,1$) the encoder transmits $S_0 = \sqrt{\frac{P}{\alpha_{1,0}}}X_0^1 + m$ and $S_1 = \sqrt{\frac{P}{\alpha_{2,1}}}X_1^2 + m$, where m is a Gaussian variable with zero mean and variance σ_m^2. In this way ρ_2 can take on value between zero and one by varying σ_m^2. Thus by choosing σ_m^2 such that $\rho_2 = \rho^\star$, we can rewrite (6.10) and (6.11) as

$$\alpha_{1,t} = \alpha_{1,2}\left(\lambda_1^2\frac{D^\star(N+N_1)+g^2P(1-\rho^{\star 2})}{D^\star(P+N+N_1)}\right)^{t-2} \tag{6.23}$$

$$\alpha_{2,t} = \alpha_{2,2}\left(\lambda_2^2\frac{D^\star(N+N_2)+P(1-\rho^{\star 2})}{D^\star(P+N+N_2)}\right)^{t-2} \tag{6.24}$$

Although in the modified encoding scheme we have violated the average power constraint for the first two transmissions, its effect can be neglected for infinite time horizon. We observe from (6.23) that $\alpha_{1,t} \to 0$ as $t \to \infty$ if

$$\begin{aligned}&\left(\lambda_1^2\frac{D^\star(N+N_1)+g^2P(1-\rho^{\star 2})}{D^\star(P+N+N_1)}\right) < 1 \\ &\Rightarrow \log(\lambda_1) < \frac{1}{2}\log\left(\frac{D^\star(P+N+N_1)}{D^\star(N+N_1)+g^2P(1-\rho^{\star 2})}\right).\end{aligned}$$

Similarly it follows from (6.24) that $\alpha_{2,t} \to 0$ as $t \to \infty$ if

$$\log(\lambda_2) < \frac{1}{2}\log\left(\frac{D^\star(P+N+N_2)}{D^\star(N+N_2)+P(1-\rho^{\star 2})}\right),$$

which completes the proof. □

Remark 6.4.1. *The sensing and control schemes presented above for stabilization of a scalar plant over multiple-access and broadcast channels can be extended to* *vector-valued plants by using the time sharing approach of mode by mode transmission proposed in Chapter 2. In this time sharing scheme, the sensor transmits only one component of the state vector at each time step such that the state components corresponding to more unstable states are transmitted more often. Thus, channel is time shared between different state components according to their level of instability.*

6.5 Conclusion

We studied the problem of mean square stabilizing two discrete time scalar LTI systems in closed-loop via control over white Gaussian multiple-access and broadcast communication channels. We proposed to use simple linear communication and control schemes and derived sufficient conditions for stability under the proposed schemes. The stability regions obtained are associated with the achievable rate regions for the given channels with noiseless feedback due to the fact Ozarow's scheme provides an exponential error decay. Our results reveal relationships between mean square stability of the two plants and the communication channels' parameters, i.e., average power consumed by the encoder(s) and the average power of the noise components in different links. We showed that the proposed schemes significantly outperform the TDMA schemes. We showed that achievable stability regions are enlarged when the initial states of the two plants are fully correlated. For broadcast channels, individual noise component seems to be less harmful than common noise components which models an external interference disturbing the signals received at both controllers.

Chapter 7

Interference Channels

7.1 Introduction

In a networked control system there are various agents such as sensors, controllers, actuators, and plants. These agents need to communicate to meet certain control objectives, and in many applications this communication should preferably take place over wireless links to reduce cabling cost and to provide flexible and mobile solutions. A major hurdle in implementing wireless networked control systems is the interference which happens due to the cross-talk between various agents while using shared communication resources. There are also external sources of interference such as other radio devices communicating in the neighborhood. In certain situations, interference can be a major factor limiting performance of a networked control system. Therefore, it is essential to study and understand the behavior of networked control systems subject to interference. This chapter makes an effort in this direction by providing necessary conditions and sufficient conditions for mean-square stabilization over a symmetric Gaussian interference channel. The gaps between the necessary conditions and the sufficient conditions are evaluated. It is demonstrated that delay-free linear sensing and control schemes can be close to optimal in some regimes. Moreover, in certain special cases they are shown to be exactly optimal.

7.2 Problem Setup

As in Chapter 6, we consider two noise-free LTI plants whose state equations are given by,

$$X_{t+1}^i = A_i X_t^i + U_t^i, \quad i = 1, 2, \tag{7.1}$$

where $X_t^i \in \mathbb{R}^n$ and $U_t^i \in \mathbb{R}^n$, are the state and the control processes of the i-th plant. We assume that the initial state X_0^i is a random variable with an arbitrary probability distribution and a given covariance matrix Λ_0^i with $\mathrm{Tr}[\Lambda_0^i] < \infty$. Let $\{\lambda_{i,1}, \lambda_{i,2}, \ldots, \lambda_{i,n}\}$ denote the eigenvalues of the system matrix A_i. Without loss

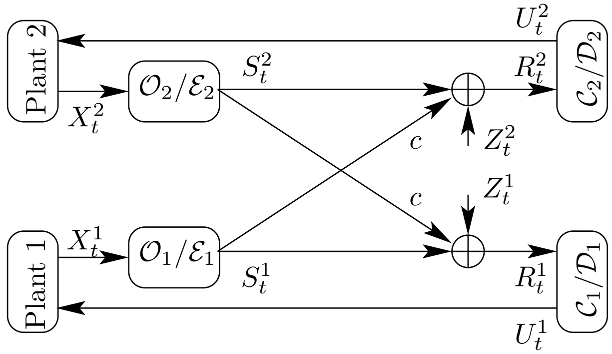


Figure 7.1: System Model.

of generality we assume that all the eigenvalues of the system matrix A_i are outside the unit disc, i.e., $|\lambda_{i,j}| > 1$ for all i, j. The unstable modes can be decoupled from the stable modes by a similarity transformation. If the system in (7.1) is one dimensional, then A_i is scalar and we use the notation $A_i = \lambda_i$, where $|\lambda_i| > 1$.

The setup for control over symmetric Gaussian interference channel is depicted in Fig. 7.1. There are two separate observers $\{\mathcal{O}_1, \mathcal{O}_2\}$ and separate controllers $\{\mathcal{C}_1, \mathcal{C}_2\}$ for the two plants. In order to communicate the observed state values to the controllers, an encoder $\mathcal{E}_i$ is lumped with the observer $\mathcal{O}_i$ and a decoder $\mathcal{D}_i$ is lumped with the controller $\mathcal{C}_i$. At any time instant t, the encoders $\mathcal{E}_1$ and $\mathcal{E}_2$ transmit S_t^1 and S_t^2 respectively. The decoders $\mathcal{D}_1$ and $\mathcal{D}_2$ respectively receive

$$R_t^1 = S_t^1 + cS_t^2 + Z_t^1,$$
$$R_t^2 = S_t^2 + cS_t^1 + Z_t^2,$$

where $c \in \mathbb{R}^+$ is the cross channel gain, and $Z_t^1 \sim \mathcal{N}(0, N)$ and $Z_t^2 \sim \mathcal{N}(0, N)$ are white noise components with a fixed cross-correlation coefficient $\rho_z \triangleq \frac{\mathbb{E}[Z_t^1 Z_t^2]}{N} \in [-1, 1]$. The cross-correlation between the two noise components models a common noise or common interference in the two signals. Let f_t^i denote the ith observer/encoder policy, then we have $S_t^i = f_t^i\left(X_{[0,t]}^i\right)$ where $X_{[0,t]}^i := \{X_1^i, X_2^i, \ldots, X_t^i\}$. The sensors must satisfy an average transmit power constraint $\mathbb{E}\left[(S_t^1)^2\right] = \mathbb{E}\left[(S_t^2)^2\right] = P$. We define the correlation between S_t^1 and S_t^2 as $\tilde{\rho}_t := \frac{\mathbb{E}[S_t^1 S_t^2]}{P} \in [-1, 1]$. Further, let π_t^i denote the i-th decoder/controller policy, then $U_t^i = \pi_t^i\left(R_{[0,t]}^i\right)$. The common objective of the sensors and the controllers is to stabilize their respective plants in the mean-square sense (cf. Definition 6.2.1).

Rest of the chapter is organized as follows: In Sec. 7.3 necessary conditions for mean-square stability are presented. In Sec. 7.4 some linear memoryless sensing and control schemes are introduced and the conditions that guarantee system stability under those schemes are presented. Sec. 7.5 gives some special cases where these schemes are optimal. Further discussions on how interference channel parameters affect stabilizability are given in Sec. 7.6. The proofs of the necessity and sufficiency results are given in Sec. 7.7 and Sec. 7.8 respectively. Finally, Sec. 7.9 concludes the main findings of this chapter and highlights some interesting directions for future research on this topic.

7.3 Necessary Conditions for Stabilization

In the following theorem we present necessary conditions for mean-square stabilization.

Theorem 7.3.1. *The two LTI systems in (7.1) can be mean square stabilized over the given symmetric Gaussian interference channel only if*

$$\log\left(\left|\det\left(A_1\right)\right|\right) \leq \frac{1}{2}\log\left(1+\frac{P\left(1+c\right)^2}{N}\right), \tag{7.2}$$

$$\log\left(\left|\det\left(A_2\right)\right|\right) \leq \frac{1}{2}\log\left(1+\frac{P\left(1+c\right)^2}{N}\right), \tag{7.3}$$

$$\log\left(\left|\det\left(A_1\right)\right|\right)+\log\left(\left|\det\left(A_2\right)\right|\right) \leq \frac{1}{2}\max_{0\leq\rho\leq1}\left\{\log\left(1+\frac{P\left(1+c^2+2c\rho\right)}{N}\right)\right.$$
$$\left.+\log\left(\frac{N\left(1-\rho_z^2\right)+P\left(1+c^2-2c\rho_z\right)\left(1-\rho^2\right)}{\left(1-\rho_z^2\right)\left(Pc^2\left(1-\rho^2\right)+N\right)}\right)\right\}. \tag{7.4}$$

Proof. The proof is given in Sec. 7.7. □

Remark 7.3.1. *It is interesting to note that the terms on the RHS of* (7.2) *and* (7.3) *are increasing functions of cross-channel gain c, which indicates that interference may improve stabilizability. However, these are only necessary conditions. In the following sections we will also present sufficient conditions and show that achievable stability region can actually enlarge in the presence of strong interference.*

Remark 7.3.2. *In order to see the effect of noise correlation on the necessary conditions, let us consider a special case with $c = 1$. The second addend on the RHS of* (7.4) *is then given by,*

$$\log\left(\frac{N\left(1-\rho_z^2\right)+P\left(1+c^2-2c\rho_z\right)\left(1-\rho^2\right)}{\left(1-\rho_z^2\right)\left(Pc^2\left(1-\rho^2\right)+N\right)}\right)=\log\left(\frac{N+\frac{2P\left(1-\rho^2\right)}{\left(1+\rho_z\right)}}{\left(P\left(1-\rho^2\right)+N\right)}\right), \tag{7.5}$$

which is a decreasing function of ρ_z, indicating that higher noise correlation may be more harmful.

Remark 7.3.3. *For $\rho_z = c = 1$, the interference channel becomes equivalent to a multiple access channel studied in Chapter 6. Therefore, by substituting $\rho_z = c = 1$ in* (7.4) *(we use* (7.5) *while evaluating* (7.4) *to avoid a division by zero), we get the following necessary condition for mean square stabilization over a symmetric Gaussian multiple access channel:*

$$\log\left(\left|\det\left(A_1\right)\right|\right)+\log\left(\left|\det\left(A_2\right)\right|\right) \leq \frac{1}{2}\log\left(1+\frac{P\left(1+c\right)^2}{N}\right), \tag{7.6}$$

which complements the sufficient conditions derived in Chapter 6.

Remark 7.3.4. *The term of on the RHS of* (7.4) *is also an outer bound on the symmetric capacity of a two-user Gaussian interference channel with noiseless feedback [ST11]. This result is presented in [ST11, Theorem 1] for $\rho_z = 0$ and $N = 1$.*

7.4 Sufficient Conditions for Stabilization

In this section, we propose to use delay-free sensing and control schemes for stabilization of scalar LTI plants over the given Gaussian interference channel. By employing the proposed schemes, we obtain sufficient conditions for mean square stability.

7.4.1 Proposed Sensing and Control Schemes

At any time t, the encoders $\{\mathcal{E}_1, \mathcal{E}_2\}$ observe the state of their corresponding plants and respectively transmit,

$$S_t^1 = \sqrt{\frac{P}{\alpha_{1,t}}}X_t^1,$$

$$S_t^2 = \mathrm{sgn}(\rho_t)\sqrt{\frac{P}{\alpha_{2,t}}}X_t^2,$$

where $\alpha_{i,t} := \mathbb{E}\left[(X_t^i)^2\right]$ is the second moment of the i-th state variable, $\rho_t := \frac{\mathbb{E}\left[X_t^1 X_t^2\right]}{\sqrt{\alpha_{1,t}\alpha_{2,t}}}$ is the correlation coefficient of the two state variables and $\mathrm{sgn}(\rho_t) = -1$ when $\rho_t < 0$ and $\mathrm{sgn}(\rho_t) = 1$ when $\rho_t \geq 0$. It is shown in Sec. 7.8.1 that in the steady state either $\mathrm{sgn}(\rho_t) = (-1)^t$ or $\mathrm{sgn}(\rho_t) = 1$ for all t. The sensor $\mathcal{E}_2$ can compute $\mathrm{sgn}(\rho_t)$ off-line depending on the values of channel parameters, therefore the two sensors are not required to cooperate. This transmit scheme is based on the coding scheme introduced by Ozarow for Gaussian multiple-access channel with noiseless feedback [Oza84], which was later used by Kramer for transmission over interference networks [KGG05].

The i-th controller $\mathcal{C}_i$ receives the signal R_t^i and aims at stabilizing the i-th plant in mean square sense. We consider the following two memoryless linear control strategies.

MMSE based Control Scheme

In this scheme, the i-th controller computes a memoryless MMSE estimate of the state and applies an action which is linear in the state estimate, i.e., the control actions are given by

$$U_t^i = -\lambda_i \mathbb{E}\left[X_t^i | R_t^i\right], \qquad \text{for } i = 1, 2.$$

We refer to this scheme as *MMSE based control scheme.*

Optimized Control Scheme

In this scheme, the actions of $\{\mathcal{C}_1, \mathcal{C}_2\}$ at any time t are given by

$$\begin{aligned} U_t^1 &= -\lambda_1 k \sqrt{\alpha_{1,t}} R_t^1, \\ U_t^2 &= -\lambda_2 k \sqrt{\alpha_{2,t}} \operatorname{sgn}(\rho_t) R_t^2, \end{aligned}$$

where $\alpha_{i,t} = \mathbb{E}\left[(X_t^i)^2\right]$, $\rho_t = \frac{\mathbb{E}\left[X_t^1 X_t^2\right]}{\sqrt{\alpha_{1,t}\alpha_{2,t}}}$, and $k \in \mathbb{R}^+$ is a design parameter which can be optimized to maximize achievable stability region. We refer to this scheme as *optimized control scheme.*

For the sake of reference we also consider a TDMA transmit scheme in which the two sensors transmit in disjoint time slots such that there is no cross-talk.

7.4.2 TDMA scheme

Time division multiple-access (TDMA) scheme works as follows: Let $l, l_1, l_2 \in \mathbb{N}$ such that $l = l_1 + l_2$. Now consider a periodic transmission scheme with a time period equal to l time steps. Assume that during each transmission period, sensor $\mathcal{E}_1$ transmits in the first l_1 time steps and then sensor $\mathcal{E}_2$ transmits in the remaining l_2 time steps. The signals transmitted by the two sensors are given by,

$$\begin{aligned} S_t^1 &= \sqrt{\frac{P}{\alpha_{1,t}}} X_t^1, \quad S_t^2 = 0, \quad t \in \{1 + nl, 2 + nl, \ldots, l_1 + nl\}, n \in \mathbb{N}, \\ S_t^2 &= \sqrt{\frac{P}{\alpha_{2,t}}} X_t^2, \quad S_t^1 = 0, \quad t \in \{l_1 + 1 + nl, l_1 + 2 + nl, \ldots, nl\}, n \in \mathbb{N}. \end{aligned} \tag{7.7}$$

Suppose that the controllers also apply actions in disjoint time steps, given as follows:

$$\begin{aligned} U_t^1 &= -\lambda_1 \mathbb{E}\left[X_t^1 | R_t^1\right], \quad U_t^2 = 0, \quad t \in \{1 + nl, 2 + nl, \ldots, l_1 + nl\}, n \in \mathbb{N}, \\ U_t^2 &= -\lambda_2 \mathbb{E}\left[X_t^2 | R_t^2\right], \quad U_t^1 = 0, \quad t \in \{l_1 + 1 + nl, l_1 + 2 + nl, \ldots, nl\}, n \in \mathbb{N}. \end{aligned} \tag{7.8}$$

This is an optimal control strategy given that sensors are transmitting in non-overlapping time slots. It will be shown shortly that in this scheme the parameters

$l, l_1, l_2 \in \mathbb{N}$ can be chosen arbitrarily to achieve different points in the stability region.

Remark 7.4.1. *The sensing and control schemes presented above for scalar systems can be applied to vector systems using the time sharing approach of mode by mode transmission proposed in Chapter 2 for stabilization of multi-dimensional plants over Gaussian channels. In this time sharing scheme, the sensor transmits only one component of the state vector at each time step such that the state components corresponding to more unstable states are transmitted more often. Thus, channel is time shared between different state components according to their level of instability.*

7.4.3 Achievable Stability Regions

We now present the stability regions that can be obtained by employing the proposed sensing and control schemes.

Theorem 7.4.1. *The two scalar LTI systems in (7.1) with $A_i = \lambda_i$ can be mean square stabilized over the given Gaussian interference channel using* MMSE based scheme *if*

$$\log(|\lambda_i|) < \frac{1}{2}\log\left(\frac{P\left(1+c^2+2c\rho^\star\right)+N}{Pc^2\left(1-\rho^{\star 2}\right)+N}\right), \tag{7.9}$$

where $\rho^\star$ is the largest among all the roots in the interval $[0,1]$ of the following two fourth order polynomials

$$\begin{aligned} f_1(\rho) &:= \rho^4 + a_3\rho^3 + a_2\rho^2 + a_1\rho + a_0, \\ f_2(\rho) &:= \rho^4 + b_3\rho^3 + b_2\rho^2 + b_1\rho + b_0, \end{aligned} \tag{7.10}$$

where

$$\begin{aligned} a_3 &= \frac{N}{2cP}, \quad a_2 = -2 - \frac{N(4+c\rho_z)}{2c^2P}, \\ a_1 &= -\frac{N(1+2c^2+2c\rho_z)}{2c^3P} - \frac{N^2}{c^3P^2}, \\ a_0 &= 1 + \frac{N(2c-\rho_z)}{2c^3P}, \quad b_3 = \frac{2c^2P+2P+N}{2cP}, \\ b_2 &= \frac{N\rho_z}{2cP}, \quad b_1 = -\frac{(1+c^2)}{c} - \frac{N(1+2\rho_z-2c^2)}{2c^3P}, \\ b_0 &= -1 - \frac{N(2c-\rho_z)}{2c^3P}. \end{aligned}$$

Proof. The proof is given in Sec. 7.8.1. □

Theorem 7.4.2. *The two LTI systems in* (7.1) *with* $A_i = \lambda_i$ *can be mean square stabilized over the given Gaussian interference channel using* optimized scheme *if*

$$\log(|\lambda_i|) < \max_k -\frac{1}{2}\log\Big(1 + k^2\left(P\left(1+c^2+2c\rho\right)+N\right) - 2k\sqrt{P}\left(1+c\rho\right)\Big) \quad (7.11)$$

where $k \in \mathbb{R}^+$ *and* ρ *is the unique root in the interval* $[0,1]$ *of the following quadratic equation:*

$$\rho^2 + \beta\rho + \gamma = 0, \quad (7.12)$$

where
$\beta = \frac{\left(2\left(k\sqrt{P}-1\right)^2 + k^2\left(2c^2P+N\right)\right)}{2ck\sqrt{P}\left(k\sqrt{P}-1\right)}$, $\gamma = \frac{k^2N\rho_z}{2ck\sqrt{P}\left(k\sqrt{P}-1\right)} + 1$ *if* $kN\rho_z < 2c\sqrt{P}(1-k\sqrt{P})$,
and $\beta = \frac{k^2N}{2ck\sqrt{P}\left(k\sqrt{P}-1\right)}$, $\gamma = -\frac{k^2N\rho_z}{2ck\sqrt{P}\left(k\sqrt{P}-1\right)} - 1$ *if* $kN\rho_z > 2c\sqrt{P}(1-k\sqrt{P})$.

Proof. The proof is given in Sec. 7.8.2. □

Remark 7.4.2. *It is shown in Sec. 7.8 that the MMSE based scheme is a special case of the optimized scheme, therefore, if a plant is stable under the MMSE based control scheme then it is also stable under the optimized control scheme but the reverse is usually not true.*

Theorem 7.4.3. *The two LTI systems in* (7.1) *with* $A_i = \lambda_i$ *can be mean square stabilized over the given Gaussian interference channel using a TDMA scheme if*

$$\log(|\lambda_1|) + \log(|\lambda_2|) < \frac{1}{2}\log\left(1 + \frac{P}{N}\right). \quad (7.13)$$

Proof. Following the same steps as in the proof of Theorem 7.4.1, we can obtain the following sufficient conditions for mean square stability under the TDMA scheme.

$$\log(|\lambda_1|) < \frac{l_1}{2l}\log\left(1 + \frac{P}{N}\right), \quad (7.14)$$

$$\log(|\lambda_2|) < \frac{l_2}{2l}\log\left(1 + \frac{P}{N}\right). \quad (7.15)$$

The conditions (7.14) and (7.14) imply the condition (7.13) because the parameters $l, l_1, l_2 \in \mathbb{N}$ can be chosen arbitrarily. □

7.5 Optimality of Linear Policies

The problem at hand has a non-classical information structure, where there are four decision makers (two sensors and two controllers). It is well-known that for such problems linear policies are not optimal in general. As an example, we discussed the Witsenhausen problem [Wit68] in Section 1.3, in which there are two decision makers and linear policies are shown to be sub-optimal. In the following we give some special cases where linear memoryless policies are actually optimal for mean-square stabilization over the Gaussian interference channel.

Proposition 7.5.1. *The MMSE based linear memoryless scheme is optimal for mean-square stabilization over the given symmetric Gaussian channel if* $\rho_z = 0$*,* $c < 1$*, and* $P = \frac{2N}{(1-c^2)^2}$*.*

Proof. If we substitute $\rho_z = 0$, $P = \frac{2N}{(1-c^2)^2}$, and $\rho_t = c$ in (7.40), then after some algebraic simplifications we have $\rho_{t+1} = -\mathrm{sgn}(\rho_t)c$ which implies $\rho^\star = c$. By evaluating (7.9) for $\rho^\star = c$ and $P = \frac{2N}{(1-c^2)^2}$, we get the following sufficient condition for stability under the MMSE based scheme:

$$\log(|\lambda_i|) < \frac{1}{2}\log\left(\frac{3+c^2}{1-c^2}\right). \tag{7.16}$$

Now consider the necessary condition for stability given in (7.4). It can be shown that for $\rho_z = 0$, $P = \frac{2N}{(1-c^2)^2}$, the term on the RHS in (7.4) is maximized by choosing $\rho = c$ and the necessary condition is then given by

$$\log(|\lambda_1|) + \log(|\lambda_2|) < \log\left(\frac{3+c^2}{1-c^2}\right). \tag{7.17}$$

The necessary condition in (7.17) coincides with the sufficient condition in (7.16) for a symmetric setting with $|\lambda_1| = |\lambda_2|$. □

Proposition 7.5.2. *The MMSE based linear memoryless scheme is optimal for mean square stabilization over the given symmetric Gaussian channel with noise correlation* $\rho_z = 1$ *if the initial states of two plants are fully correlated or anti-correlated, i.e.,* $\rho_0 = \pm 1$*.*

Proof. For this special case with $\rho_0 = \pm 1$, we modify the initial transmission phase discussed in Sec. 7.8.1 where the sensors transmit initial states $\{X_0^1, X_0^2\}$ in alternate time steps. Instead suppose that both sensors transmit their initial states simultaneously as $S_0^1 = \sqrt{\frac{P}{\alpha_{1,0}}}X_0^1$ and $S_0^2 = \sqrt{\frac{P}{\alpha_{2,0}}}X_0^2\mathrm{sgn}(\rho_0)$. It is trivial to show that after this initial transmission, $\{X_1^1, X_1^2\}$ are zero mean Gaussian distributed when $\rho_z = 1$. Furthermore, $\rho_1 = 1$ if $\rho_0 = 1$ and $\rho_1 = -1$ if $\rho_0 = -1$. Following the transmission scheme given in Sec. 7.8.1 for the remaining time steps, we get $|\rho_t| = 1$ for all t. By substituting $\rho^\star = 1$ in (7.9) we get

$$\log(|\lambda_i|) < \frac{1}{2}\log\left(1 + \frac{P(1+c)^2}{N}\right), \quad i = 1, 2,$$

which coincide with the necessary conditions given in (7.2) and (7.3). □

Proposition 7.5.3. *The optimized linear memoryless scheme is asymptotically optimal for mean-square stabilization over the given symmetric Gaussian channel if the noise correlation* $\rho_z = -1$.

Proof. In the optimized scheme if we choose $k = \frac{1}{\sqrt{P(1+c)}}$, then for $\rho_z = -1$ the recursion of the state correlation coefficient in (7.40) simplifies to $\rho_{t+1} = -\mathrm{sgn}(\rho_t)$. This implies $\rho^\star = 1$. Now, for $k = \frac{1}{\sqrt{P(1+c)}}$ and $\rho^\star = 1$, the sufficient condition in (7.9) becomes

$$\log(|\lambda_i|) < \frac{1}{2}\log\left(\frac{P(1+c)^2}{N}\right) \quad i = 1, 2. \tag{7.18}$$

By comparing the above condition with the necessary conditions in (7.2) and (7.3), we observe that the gap between the sufficient and necessary conditions approaches zero as c goes to infinity. □

Proposition 7.5.4. *The MMSE based linear memoryless scheme is optimal for mean-square stabilization over the given symmetric Gaussian channel if the noise correlation* $\rho_z = 1$ *and* $2c(1 + \frac{c^2 P}{N}) < 1$.

Proof. It is shown at the end of Appendix 7.C that $\rho^\star = 1$ when $\rho_z = 1$ and $2c(1 + \frac{c^2 P}{N}) < 1$. By substituting $\rho^\star = 1$ in (7.9) and comparing with (7.2), we observe that the scheme is optimal for this special case. □

7.6 Numerical Results and Discussions

The stabilizability of the plants depends on the interference channel parameters such as average transmit power P, noise power N, noise cross-correlation ρ_z, and cross channel gain c. Therefore, it is interesting to study the effect of these channel parameters on the behavior of the two systems under our proposed schemes and gain further insight on how good these linear memoryless schemes are. In this section, we investigate stabilizability of the two plants under different values of noise cross-correlation and cross channel gain with fixed transmit and noise powers.

In Fig. 7.2–Fig. 7.4 we fix $P = 20$, $N = 1$, $\rho_z \in \{-1, 0, 0.5\}$, and plot achievable stability regions for symmetric plants ($|\lambda_1| = |\lambda_2|$) as functions of interference parameter c according to Theorem 7.4.1, Theorem 7.4.2, and Theorem 7.4.3. For the sake of reference we also plot the outer bound on achievable stability region using Theorem 7.3.1. We observe that the proposed linear memoryless schemes perform quite close to optimal in low to moderate interference regimes. The optimized control scheme always outperforms the other two schemes. For $\rho_z = -1$, the optimized scheme seems to be optimal for all the values of cross-channel gain c according to Fig. 7.2. We have already shown in Proposition 7.5.3 that optimized scheme is asymptotically optimal when $\rho_z = -1$, however, it might be that it is exactly optimal as indicated by Fig. 7.2. It is interesting that under the proposed schemes

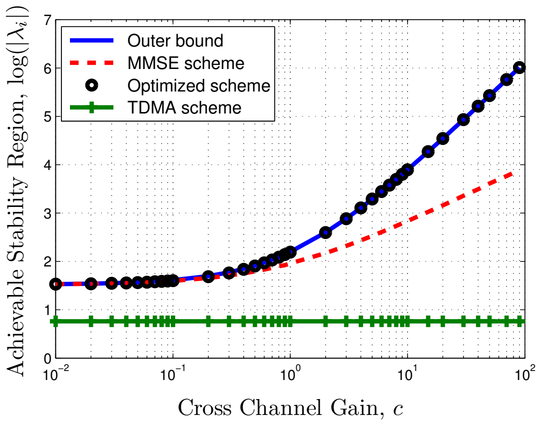


Figure 7.2: $P = 20,\ N = 1,\ \rho_z = -1$.

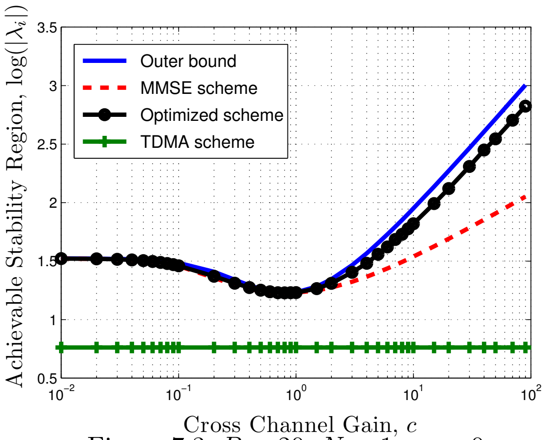


Figure 7.3: $P = 20,\ N = 1,\ \rho_z = 0$.

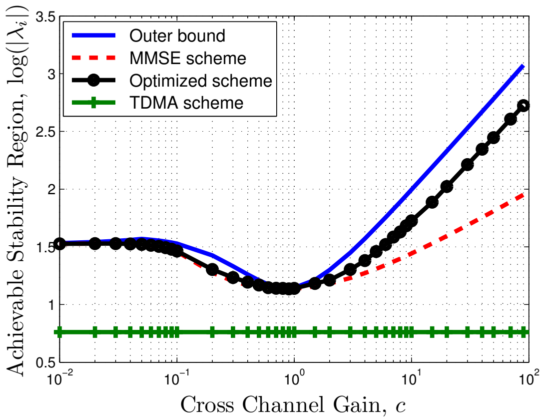


Figure 7.4: $P = 20,\ N = 1,\ \rho_z = 0.5$.

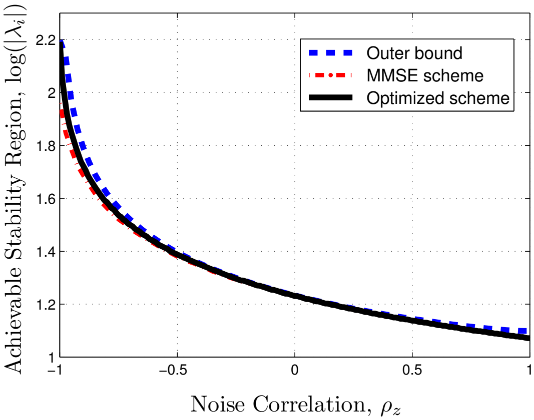


Figure 7.5: $P = 20$, $N = 1$, $c = 1$.

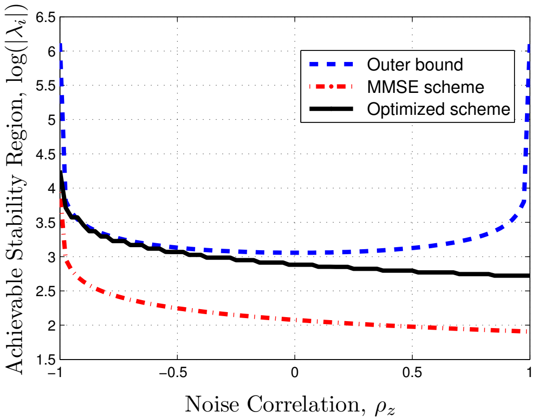


Figure 7.6: $P = 20$, $N = 1$, $c = 100$.

stabilizability significantly improves as the interference gets very strong. This result is in line with already known results in information theory, where it has been shown that the transmission rates over interference channel can be significantly improved in the presence of very strong interference [Sat81]. One reason for this is that the state processes of the two plants become highly correlated in strong interference scenarios due to dominant cross talk. Furthermore, we observe an increasing gap between the optimized scheme and the MMSE based scheme as the interference increases beyond certain threshold. This indicates that the conventional MMSE based control strategy can be quite inefficient in multi-sensor multi-controller settings, even when the sensors are restricted to be linear.

In Fig. 7.5 and Fig. 7.6 we fix $P = 20$, $N = 1$, and plot achievable stability regions as functions of noise correlation ρ_z for a moderate cross-channel gain (interference) $c = 1$ and a high cross-channel gain $c = 100$, respectively. In these examples, stability region reduces by increasing ρ_z from -1 to 1. For anti-correlated noise variables i.e., $\rho_z = -1$, there is a dramatic boost in the stabilizability. A similar behavior was observed in [GLW10], where the authors showed that the sum-rate capacity over symmetric Gaussian interference channel can be doubled with feedback in high SNR when the noise components are anti-correlated.

7.7 Necessity Proofs

In order to prove conditions (7.2) and (7.3), we make use of the following Lemma.

Lemma 7.7.1. *The i-th linear system in* (7.1) *can be mean square stabilized over the Gaussian interference channel only if*

$$\log\left(\left|\det\left(A_i\right)\right|\right) \leq \liminf_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T\text{-}1} I\left(X_t^i; R_t^i | R_{[0,t\text{-}1]}^i\right), \tag{7.19}$$

where $I\left(X_t^i; R_t^i | R_{[0,t\text{-}1]}^i\right)$ denotes mutual information between state variable of the i-th plant state X_t^i and the received variable R_t^i, conditioned on the sequence of received variables $R_{[0,t\text{-}1]}^i$.

Proof. The proof can be found in Appendix 7.A. □

In the following we obtain an upper bound on the term $\sum_{t=0}^{T\text{-}1} I\left(X_t^i; R_t^i | R_{[0,t\text{-}1]}^i\right)$ and then use Lemma 7.7.1 to derive the necessary conditions given in (7.2) and (7.3). The directed information can be bounded as,

$$\begin{aligned}
\sum_{t=0}^{T-1} I\left(X_t^i; R_t^i | R_{[0,t\text{-}1]}^i\right) &\overset{(a)}{\leq} \sum_{t=0}^{T-1} I\left(S_{[0,t]}^i; R_t^i | R_{[0,t\text{-}1]}^i\right) \\
&\overset{(b)}{\leq} I\left(S_{[0,T-1]}^i; R_{[0,T-1]}^i\right) \\
&\overset{(c)}{\leq} I\left(S_{[0,T-1]}^1, S_{[0,T-1]}^2; R_{[0,T-1]}^i\right) \\
&= h\left(R_{[0,T-1]}^i\right) - h\left(R_{[0,T-1]}^i | S_{[0,T-1]}^1, S_{[0,T-1]}^2\right) \\
&\overset{(d)}{=} h\left(R_{[0,T-1]}^i\right) - h\left(Z_{[0,T-1]}^i\right) \\
&\overset{(e)}{=} \sum_{t=0}^{T-1}\left[h\left(R_t^i | R_{[0,t-1]}^i\right) - h\left(Z_t^i\right)\right] \\
&\overset{(f)}{\leq} \sum_{t=0}^{T-1}\left[h\left(R_t^i\right) - h\left(Z_t^i\right)\right]
\end{aligned}$$

$$\overset{(g)}{\leq} \sum_{t=0}^{T-1} \frac{1}{2} \log \left(1 + \frac{P + c^2 P + 2c\tilde{\rho}_t P}{N}\right)$$
$$\overset{(h)}{\leq} \frac{T}{2} \log \left(1 + \frac{P\,(1+c)^2}{N}\right), \tag{7.20}$$

where (a) follows from the Markov chain $X_t^i - \{S_{[0,t]}^i, R_{[0,t\text{-}1]}^i\} - R_t^i$; (b) and (c) follow from the fact that adding more variables cannot decrease mutual information; (d) follows from $R_t^i = S_t^i + cS_t^j + Z_t^i$ and mutual independence of $\{Z_t^1, Z_t^2\}$ and $\{S_t^1, S_t^2\}$; (e) follows from the assumption that $Z_{[0,T\text{-}1]}^i$ is a sequence of independent variables; (f) follows from the fact that conditioning reduces entropy; (g) follows from

$$h\left(Z_t^i\right) = \frac{1}{2}\log\left(2\pi e N\right),$$
$$h\left(R_t^i\right) \leq \frac{1}{2}\log\left(2\pi e\left(P + c^2 P + 2c\tilde{\rho}_t P\right)\right),$$

where equality holds if we assume Gaussian distributed variables since the Gaussian distribution maximizes differential entropy for a given variance [CT06, Theorem 8.6.5]; and (h) follows by the maximization of the RHS of (j) subject to $-1 \leq \tilde{\rho}_t \leq 1$ for all t. The maximum value is attained by choosing $\tilde{\rho}_t = 1$ if $c \geq 0$ and $\tilde{\rho}_t = -1$ if $c \leq 0$. Now by using (7.20) in Lemma 7.7.1, we get the necessary conditions for mean square stabilization given in (7.2) and (7.3).

In the following, we derive the necessary condition (7.4) using a genie-aided bound approach [ST09]. Consider a superior system where the controllers have some side (extra) information. We define $Y_t^i := cS_t^i + Z_t^j$ for $i \neq j$ and assume that at any time t, the i-th controller has access to $Y_{[0,t]}^i$ in addition to $R_{[0,t]}^i$. In the following, we obtain a necessary condition for mean square stabilization of this superior system. Note that a condition which is necessary for the stabilization of this superior system is also necessary for the stabilization of the actual system. Similar approaches have been used in information theory community to derive outer bounds on capacity regions [GC82, ST09]. We will again use Lemma 7.7.1 to derive a necessary condition for the superior system. The following lemma reveals some functional relationship between different variables in the superior system, which will be useful in the derivation of the necessary condition.

Lemma 7.7.2. *For the superior system with side information $Y_t^i := hS_t^i + Z_t^j$ at the i-th controller, we have the following relationships for all $i, j \in \{1, 2\}$ and $i \neq j$:*

$$X_t^i = A_i^t X_0^i + \mu_t^i\left(R_{[0,t\text{-}1]}^i\right), \tag{7.21}$$

$$S_t^i = \nu_t^i\left(X_0^i, Y_{[0,t\text{-}1]}^j\right), \tag{7.22}$$

$$S_t^i = \upsilon_t^i\left(X_0^i, R_{[0,t\text{-}1]}^i\right), \tag{7.23}$$

where $\mu_t^i : \mathbb{R}^{t\text{-}1} \to \mathbb{R}$, $\nu_t^i : \mathbb{R}^t \to \mathbb{R}$, and $\upsilon_t^i : \mathbb{R}^t \to \mathbb{R}$.

Proof. The proof is given in Appendix 7.B. $\square$

The term $\sum_{t=0}^{T-1} I\left(X_t^i; R_t^i | R^i_{[0,t\text{-}1]}\right)$ can also be written as,

$$\begin{aligned}
&\sum_{t=0}^{T-1} I\left(X_t^i; R_t^i | R^i_{[0,t\text{-}1]}\right)\\
&\quad= \sum_{t=0}^{T-1} h\left(X_t^i | R^i_{[0,t\text{-}1]}\right) - h\left(X_t^i | R^i_{[0,t]}\right)\\
&\quad\stackrel{(a)}{=} \sum_{t=0}^{T-1} h\left(A_i^t X_0^i + \mu_t^i\left(R^i_{[0,t\text{-}1]}\right)|R^i_{[0,t\text{-}1]}\right) - h\left(A_i^t X_0^i + \mu_t^i\left(R^i_{[0,t\text{-}1]}\right)|R^i_{[0,t]}\right)\\
&\quad= \sum_{t=0}^{T-1} h\left(A_i^t X_0^i | R^i_{[0,t\text{-}1]}\right) - h\left(A_i^t X_0^i | R^i_{[0,t]}\right)\\
&\quad= I\left(A_i^t X_0^i; R^i_{[0,T\text{-}1]}\right)\\
&\quad\stackrel{(b)}{=} I\left(X_0^i; R^i_{[0,T\text{-}1]}\right),
\end{aligned} \tag{7.24}$$

where (a) follows by substituting X_t^i from (7.21); and (b) follows from the fact that A_i is invertible and the mutual information between two variables is invariant with respect to any reversible transformation of one of the variables [Gal68, Page 31]. Now by using (7.24) in Lemma 7.7.1, the i-th plant can be mean square stabilized only if

$$\log\left(|\det\left(A_i\right)|\right) \le \liminf_{T\to\infty} \frac{1}{T} I\left(X_0^i; R^i_{[0,T-1]}\right). \tag{7.25}$$

Thus, the two plants can be mean square stabilized only if

$$\begin{aligned}
&\log\left(|\det\left(A_1\right)|\right) + \log\left(|\det\left(A_2\right)|\right)\\
&\qquad\stackrel{(a)}{\le} \liminf_{T\to\infty} \frac{1}{T} I\left(X_0^1; R^1_{[0,T\text{-}1]}\right) + \liminf_{T\to\infty} \frac{1}{T} I\left(X_0^2; R^2_{[0,T\text{-}1]}\right)\\
&\qquad\stackrel{(b)}{\le} \liminf_{T\to\infty} \frac{1}{T}\left\{I\left(X_0^1; R^1_{[0,T\text{-}1]}\right) + I\left(X_0^2; R^2_{[0,T\text{-}1]}\right)\right\},
\end{aligned} \tag{7.26}$$

where (a) follows from (7.25) and (b) follows from the fact that limit inferior satisfies superadditivity. We can now bound the sum $I\left(X_0^1; R^1_{[0,T\text{-}1]}\right) + I\left(X_0^2; R^2_{[0,T\text{-}1]}\right)$ as

$$\begin{aligned}
&I\left(X_0^1; R^1_{[0,T-1]}\right) + I\left(X_0^2; R^2_{[0,T-1]}\right)\\
&\stackrel{(a)}{\le} I\left(X_0^1; R^1_{[0,T-1]}, Y^1_{[0,T-1]}, X_0^2\right) + I\left(X_0^2; R^2_{[0,T-1]}\right)
\end{aligned}$$

$$\stackrel{(b)}{=} I\left(X_0^1; R^1_{[0,T-1]}, Y^1_{[0,T-1]} \big| X_0^2\right) + I\left(X_0^2; R^2_{[0,T-1]}\right)$$
$$\stackrel{(c)}{=} h\left(R^1_{[0,T\text{-}1]}, Y^1_{[0,T\text{-}1]} \big| X_0^2\right) - h\left(R^1_{[0,T\text{-}1]}, Y^1_{[0,T\text{-}1]} \big| X_0^1, X_0^2\right) + h\left(R^2_{[0,T\text{-}1]}\right)$$
$$- h\left(R^2_{[0,T\text{-}1]} \big| X_0^2\right), \tag{7.27}$$

where (a) follows from the fact that adding side information cannot decrease mutual information; (b) follows from the assumption that X_0^1 and X_0^2 are mutually independent; and (c) follows by writing mutual information in terms of differential entropies. The differential entropy $h\left(R^1_{[0,T\text{-}1]}, Y^1_{[0,T\text{-}1]}|X_0^1, X_0^2\right)$ in (7.27) can be simplified as

$$\begin{aligned}
h\left(R^1_{[0,T\text{-}1]}, Y^1_{[0,T\text{-}1]}|X_0^1, X_0^2\right) &= \sum_{t=0}^{T\text{-}1} h\left(R_t^1, Y_t^1|X_0^1, X_0^2, R^1_{[0,t\text{-}1]}, Y^1_{[0,t\text{-}1]}\right) \\
&\stackrel{(a)}{=} \sum_{t=0}^{T\text{-}1} h\left(R_t^1, Y_t^1|X_0^1, X_0^2, R^1_{[0,t\text{-}1]}, Y^1_{[0,t\text{-}1]}, S_t^1, S_t^2\right) \\
&\stackrel{(b)}{=} \sum_{t=0}^{T\text{-}1} h\left(Z_t^1, Z_t^2|X_0^1, X_0^2, R^1_{[0,t\text{-}1]}, Y^1_{[0,t\text{-}1]}, S_t^1, S_t^2\right) \\
&\stackrel{(c)}{=} \sum_{t=0}^{T\text{-}1} h\left(Z_t^1, Z_t^2\right),
\end{aligned} \tag{7.28}$$

where (a) follows from (7.22); (b) follows from $R_t^1 = S_t^1 + cS_t^2 + Z_t^1$ and $Y_t^1 = cS_t^1 + Z_t^2$; and (c) follows from the fact that the channels are memoryless. The entropy term $h\left(R^2_{[0,T\text{-}1]}|X_0^2\right)$ in (7.27) can be expressed as

$$\begin{aligned}
h\left(R^2_{[0,T\text{-}1]}|X_0^2\right) &= \sum_{t=0}^{T\text{-}1} h\left(R_t^2|X_0^2, R^2_{[0,t\text{-}1]}\right) \\
&\stackrel{(a)}{=} \sum_{t=0}^{T\text{-}1} h\left(S_t^2 + Y_t^1|X_0^2, R^2_{[0,t\text{-}1]}, S^2_{[0,t]}\right) \\
&= \sum_{t=0}^{T\text{-}1} h\left(Y_t^1|X_0^2, R^2_{[0,t\text{-}1]}, S^2_{[0,t]}\right) \\
&\stackrel{(b)}{=} \sum_{t=0}^{T\text{-}1} h\left(Y_t^1|X_0^2, R^2_{[0,t\text{-}1]}, S^2_{[0,t]}, Y^1_{[0,t\text{-}1]}\right) \\
&\stackrel{(c)}{=} \sum_{t=0}^{T\text{-}1} h\left(Y_t^1|X_0^2, Y^1_{[0,t\text{-}1]}\right) = h\left(Y^1_{[0,T\text{-}1]}|X_0^2\right),
\end{aligned} \tag{7.29}$$

where (a) follows from $R_t^2 = S_t^2 + Y_t^1$ and (7.23); (b) follows from $Y_t^1 = R_t^2 - S_t^2$; and (c) follows from $R_t^2 = S_t^2 + Y_t^1$ and (7.22). We can now re-write (7.27) as,

$$
\begin{aligned}
& I\left(X_0^1; R^1_{[0,T-1]}\right) + I\left(X_0^2; R^2_{[0,T-1]}\right) \\
& \stackrel{(a)}{\leq} h\left(R^1_{[0,T\text{-}1]}, Y^1_{[0,T\text{-}1]} | X_0^2\right) - \sum_{t=0}^{T\text{-}1} h\left(Z_t^1, Z_t^2\right) + h\left(R^2_{[0,T\text{-}1]}\right) - h\left(Y^1_{[0,T\text{-}1]} | X_0^2\right) \\
& = h\left(R^1_{[0,T\text{-}1]} | Y^1_{[0,T\text{-}1]}, X_0^2\right) + h\left(R^2_{[0,T\text{-}1]}\right) - \sum_{t=0}^{T\text{-}1} h\left(Z_t^1, Z_t^2\right) \\
& \stackrel{(b)}{=} h\left(R^1_{[0,T\text{-}1]} | Y^1_{[0,T\text{-}1]}, X_0^2, S^2_{[0,T\text{-}1]}\right) + h\left(R^2_{[0,T\text{-}1]}\right) - \sum_{t=0}^{T\text{-}1} h\left(Z_t^1, Z_t^2\right) \\
& \stackrel{(c)}{\leq} \sum_{t=0}^{T-1} \left[h\left(R_t^1 | Y_t^1, S_t^2\right) + h\left(R_t^2\right) - h\left(Z_t^1, Z_t^2\right)\right] \\
& \stackrel{(d)}{\leq} \sum_{t=0}^{T-1} \left[\frac{1}{2}\log\left(1 + \frac{P\left(1 + c^2 + 2c\rho_t\right)}{N}\right)\right. \\
& \qquad\qquad \left. + \log\left(\frac{N\left(1-\rho_z^2\right) + P\left(1 + c^2 - 2c\rho_z\right)\left(1-\rho_t^2\right)}{\left(1-\rho_z^2\right)\left(Pc^2\left(1-\rho_t^2\right) + N\right)}\right)\right] \\
& \stackrel{(e)}{\leq} \sum_{t=0}^{T-1} \max_{0\leq\rho_t\leq 1} \left[\frac{1}{2}\log\left(1 + \frac{P\left(1 + c^2 + 2c\rho_t\right)}{N}\right)\right. \\
& \qquad\qquad \left. + \log\left(\frac{N\left(1-\rho_z^2\right) + P\left(1 + c^2 - 2c\rho_z\right)\left(1-\rho_t^2\right)}{\left(1-\rho_z^2\right)\left(Pc^2\left(1-\rho_t^2\right) + N\right)}\right)\right] \\
& = \frac{T}{2} \max_{0\leq\rho\leq 1} \left\{\log\left(1 + \frac{P\left(1 + c^2 + 2c\rho\right)}{N}\right)\right. \\
& \qquad\qquad \left. + \log\left(\frac{N\left(1-\rho_z^2\right) + P\left(1 + c^2 - 2c\rho_z\right)\left(1-\rho^2\right)}{\left(1-\rho_z^2\right)\left(Pc^2\left(1-\rho^2\right) + N\right)}\right)\right\},
\end{aligned} \tag{7.30}
$$

where (a) follows from (7.28) and (7.29); (b) follows from (7.22); (c) follows since conditioning from reduces entropy; and (d) from

$$h\left(Z_t^1, Z_t^2\right) = \frac{1}{2}\log\left((2\pi e)^2 N^2 (1-\rho_z^2)\right), \tag{7.31}$$

$$h(R_t^2) \leq \frac{1}{2}\log\left(2\pi e\left(N + P\left(1 + c^2 + 2c\rho_t\right)\right)\right), \tag{7.32}$$

$$h\left(R_t^1 | Y_t^1, S_t^2\right) \leq \frac{1}{2}\log\left(2\pi e \frac{N^2\left(1-\rho_z^2\right) + PN\left(1 + c^2 - 2c\rho_z\right)\left(1-\rho_t^2\right)}{Pc^2\left(1-\rho_t^2\right) + N}\right), \tag{7.33}$$

where the computation of (7.31) follows from the assumption that noise variables are Gaussian and the inequality (7.32) follows from the fact that the Gaussian distribution maximizes differential entropy for a given variance. The inequality (7.33) is obtained as follows. We first bound the conditional variance of R_t^1 given (Y_t^1, S_t^2) as

$$\begin{aligned}
&\mathrm{Var}\left[R_t^1 | Y_t^1, S_t^2\right] \\
&\leq \mathbb{E}\left[(R_t^1)^2\right] - \mathbb{E}\left[R_t^1 (Y_t^1, S_t^2)\right] \left(\mathbb{E}\left[(Y_t^1, S_t^2)^T (Y_t^1, S_t^2)\right]\right)^{-1} \left(\mathbb{E}\left[R_t^1 (Y_t^1, S_t^2)\right]\right)^T \\
&= \frac{N\left(1-\rho_z^2\right) + P\left(1+c^2-2c\rho_z\right)\left(1-\rho_t^2\right)}{\left(1-\rho_z^2\right)\left(Pc^2\left(1-\rho_t^2\right)+N\right)},
\end{aligned} \tag{7.34}$$

where the equality is achieved with multivariate Gaussian distribution. Further we know that the Gaussian distribution maximizes entropy for a given variance, which gives the inequality (7.33). This completes the proof.

7.8 Sufficiency Proofs

In order to prove Theorem 7.4.1 and Theorem 7.4.2 we find conditions on the system parameters $\{\lambda_1, \lambda_2\}$ which are sufficient to mean square stabilize the system in (7.1) under the sensing and control schemes proposed in Sec. 7.4.

7.8.1 Proof of Theorem 7.4.1

The control and communication scheme for the interference channel works as follows.

Initial time steps, $t = 0, 1$

Similar to the schemes for all the channels (point-to-point, relay, broadcast, MAC) studied in the previous chapters, we perform an initialization to make the plants' states Gaussian distributed. In this initialization phase, the two encoders transmit the observed state values in alternate time slots to the respective controllers. The first two disjoint transmissions in time make the plants' states Gaussian distributed regardless of the distribution of their initial states, which will be explained shortly. However, if the initial states are already Gaussian, then the following disjoint initial transmissions are not needed.

At time step $t = 0$, the encoder $\mathcal{E}_1$ observes X_0^1 and transmits $S_0^1 = \sqrt{\frac{P}{\alpha_{1,0}}} X_0^1$. The encoder $\mathcal{E}_2$ does not transmit, i.e., $S_0^2 = 0$. The decoder $\mathcal{D}_1$ receives $R_0^1 = S_0^1 + Z_0^1$. It then estimates X_0^1 as

$$\begin{aligned}
\hat{X}_0^1 &= \sqrt{\frac{\alpha_{1,0}}{P}} R_0^1 \\
&= X_0^1 + \sqrt{\frac{\alpha_{1,0}}{P}} Z_0^1.
\end{aligned}$$

The controller $\mathcal{C}_1$ then takes an action $U_0^1 = -\lambda_1 \hat{X}_0^1$ for the plant 1, which results in $X_1^1 = \lambda_1(X_0^1 - \hat{X}_0^1)$. The state $X_1^1 \sim \mathcal{N}(0, \alpha_{1,1})$ with $\alpha_{1,1} = \lambda_1^2 \frac{\alpha_{1,0} N}{P}$. The controller does not take any action for the plant 2, therefore $X_1^2 = \lambda_2 X_0^2$ with $\alpha_{2,1} = \lambda_2^2 \alpha_{2,0}$.

At time step $t = 1$, the encoder $\mathcal{E}_1$ does not transmit any signal. The encoder $\mathcal{E}_2$ observes X_1^2 and transmits $S_1^2 = \sqrt{\frac{P}{\alpha_{2,1}}} X_1^2$. The decoder $\mathcal{D}_2$ receives $R_1^2 = S_1^2 + Z_1^2$. It then estimates X_1^2 as

$$\begin{aligned}\hat{X}_1^2 &= \sqrt{\frac{\alpha_{2,1}}{P}} R_1^2 \\ &= X_1^2 + \sqrt{\frac{\alpha_{2,1}}{P}} Z_1^2.\end{aligned}$$

The controller $\mathcal{C}_2$ then takes an action $U_1^2 = -\lambda_2 \hat{X}_1^2$ for the plant 2, which results in $X_2^2 = \lambda_2(X_1^2 - \hat{X}_1^2)$, where $X_2^2 \sim \mathcal{N}(0, \alpha_{2,2})$. For the plant 1, the controller does not take any action $U_1^1 = 0$, therefore $X_2^1 = \lambda_1 X_1^1$ and $X_2^1 \sim \mathcal{N}(0, \alpha_{1,2})$.

It is noteworthy that due to non-overlapping initial transmissions by the two encoders, the states X_2^1 and X_2^2 are now zero mean Gaussian variables with correlation coefficient $\rho_2 = \frac{\mathbb{E}[X_2^1 X_2^2]}{\sqrt{\alpha_{1,2}\alpha_{2,2}}}$ equal to zero[1]. Henceforth the two encoders will transmit their signals simultaneously.

Further time steps $t \geq 2$

The two encoders $\mathcal{E}_1$ and $\mathcal{E}_2$ observe X_t^1 and X_t^2, and they respectively transmit

$$\begin{aligned}S_t^1 &= \sqrt{\frac{P}{\alpha_{1,t}}} X_t^1, \\ S_t^2 &= \sqrt{\frac{P}{\alpha_{2,t}}} X_t^2 \mathrm{sgn}(\rho_t),\end{aligned}$$

where $\rho_t = \frac{\mathbb{E}\left[(X_t^1 - \mathbb{E}[X_t^1])(X_t^2 - \mathbb{E}[X_t^2])\right]}{\sqrt{\alpha_{1,t}\alpha_{2,t}}}$ and $\mathrm{sgn}(\rho_t) = 1$ if $\rho_t \geq 0$ and $\mathrm{sgn}(\rho_t) = -1$ if $\rho_t < 0$.

In accordance, the decoder $\mathcal{D}_1$ receives $R_t^1 = S_t^1 + cS_t^2 + Z_t^1$ and the decoder $\mathcal{D}_2$ receives $R_t^2 = S_t^2 + cS_t^1 + Z_t^2$. The decoder $\mathcal{D}_i$ then computes a memoryless [2] MMSE

[1]The states in the second time step become uncorrelated irrespective of the value of the correlation between the initial states. This scheme does not exploit correlation between the initial states and thus the stability region obtained is independent of the correlation of the initial states.

[2]The memoryless estimator is not optimal since the channel outputs are correlated. Therefore we expect that an improvement might be possible if we use full memory in the estimator. However the analysis becomes complicated by considering full LMMSE estimation.

estimate of the state of the plant i as

$$\begin{aligned}\hat{X}_t^i &= \mathbb{E}[X_t^i | R_t^i] \\ &\overset{(a)}{=} \frac{\mathbb{E}[R_t^i X_t^i]}{\mathbb{E}\left[\left(R_t^i\right)^2\right]} R_t^i, \end{aligned} \tag{7.35}$$

where (a) follows from the fact that the optimum MMSE of the Gaussian variable is linear [Hay96]; and we have

$$\begin{aligned}\mathbb{E}\left[X_t^1 R_t^1\right] &= \sqrt{P\alpha_{1,t}}\left(1 + c|\rho_t|\right), \\ \mathbb{E}[X_t^2 R_t^2] &= \sqrt{P\alpha_{2,t}}\left(1 + c|\rho_t|\right)\operatorname{sgn}(\rho_t), \\ \mathbb{E}[(R_t^i)^2] &= P(1 + c^2 + 2c|\rho_t|) + N. \end{aligned} \tag{7.36}$$

The controller $\mathcal{C}_i$ takes an action $U_t^i = -\lambda_i \hat{X}_t^i$ for the plant i, which results in $X_{t+1}^i = \lambda_i(X_t^i - \hat{X}_t^i)$. The mean values of the states are

$$\begin{aligned}\mathbb{E}[X_{t+1}^i] &= \mathbb{E}\left[\lambda_i\left(X_t^i - \hat{X}_t^i\right)\right] \\ &\overset{(a)}{=} \lambda_i \mathbb{E}\left[X_t^i - \frac{\mathbb{E}[R_t^i X_t^i]}{\mathbb{E}[(R_t^i)^2]} R_t^i\right] \\ &\overset{(b)}{=} 0, \end{aligned} \tag{7.37}$$

where (a) follows from (7.35); and (b) follows from $\mathbb{E}[X_2^i] = 0$ and by recursively using (a). The variance of the state X_{t+1}^i is given by

$$\begin{aligned}\alpha_{i,t+1} &:= \mathbb{E}[(X_{t+1}^i)^2] \\ &= \lambda_i^2\left(\mathbb{E}[(X_t^i)^2] - \frac{(\mathbb{E}[R_t^i X_t^i])^2}{\mathbb{E}[(R_t^i)^2]}\right) \end{aligned} \tag{7.38}$$

By using (7.36) in (7.38) we get the following recursive equations

$$\alpha_{i,t+1} = \alpha_{i,t}\lambda_i^2\left(\frac{Pc^2\left(1 - |\rho_t|^2\right) + N}{P(1 + c^2 + 2c|\rho_t|) + N}\right). \tag{7.39}$$

The cross-correlation coefficient ρ_t between the two state processes for all $t \geq 3$ is given by

$$\begin{aligned}\rho_{t+1} = \frac{\mathbb{E}[X_{t+1}^1 X_{t+1}^2]}{\sqrt{\alpha_{1,t+1}\alpha_{2,t+1}}} &= \frac{1}{\sqrt{\alpha_{1,t+1}\alpha_{2,t+1}}}\mathbb{E}\left[\lambda_1\left(X_t^1 - \hat{X}_t^1\right)\lambda_2\left(X_t^2 - \hat{X}_t^2\right)\right] \\ &\overset{(a)}{=} \frac{\lambda_1\lambda_2}{\sqrt{\alpha_{1,t+1}\alpha_{2,t+1}}}\left(\mathbb{E}[X_t^1 X_t^2] - \frac{\mathbb{E}[X_t^1 R_t^2]\mathbb{E}[X_t^2 R_t^2]}{\mathbb{E}[(R_t^2)^2]} - \frac{\mathbb{E}[X_t^2 R_t^1]\mathbb{E}[X_t^1 R_t^1]}{\mathbb{E}[(R_t^1)^2]}\right.\end{aligned}$$

$$+\frac{\mathbb{E}[X_t^1R_t^1]\mathbb{E}[X_t^2R_t^2]\mathbb{E}[R_t^1R_t^2]}{\mathbb{E}[R_t^1]\mathbb{E}[R_t^2]}\Bigg)$$

$$\stackrel{(b)}{=} \lambda_1\lambda_2\sqrt{\frac{\alpha_{1,t}\alpha_{2,t}}{\alpha_{1,t+1}\alpha_{2,t+1}}}\Bigg(\rho_t - 2\frac{P\text{sgn}(\rho_t)\,(c+|\rho_t|)\,(1+c|\rho_t|)}{P(1+c^2+ch|\rho_t|)+N}$$
$$+\frac{P\text{sgn}(\rho_t)\,(1+c|\rho_t|)^2\left(2cP+P|\rho_t|(1+c^2)+N\rho_z\right)}{\left(P(1+c^2+2c|\rho_t|)+N\right)^2}\Bigg)$$

$$\stackrel{(c)}{=} \text{sgn}(\rho_t)\left(\frac{P(1+c^2+2c|\rho_t|)+N}{Pc^2\,(1-|\rho_t|^2)+N}\right)\times\Bigg(|\rho_t| - 2\frac{P\,(c+|\rho_t|)\,(1+c|\rho_t|)}{P(1+c^2+2c|\rho_t|)+N}$$
$$+\frac{P\,(1+c|\rho_t|)^2\left(2cP+P|\rho_t|(1+c^2)+N\rho_z\right)}{\left(P(1+c^2+2c|\rho_t|)+N\right)^2}\Bigg)$$

$$\stackrel{(d)}{=}: g(\rho_t)\text{sgn}(\rho_t), \qquad \forall t\geq 2, \tag{7.40}$$

where (a) follows from

$$\mathbb{E}[X_t^1\hat{X}_t^2] = \frac{\mathbb{E}[X_t^1R_t^2]\mathbb{E}[X_t^2R_t^2]}{\mathbb{E}[(R_t^2)^2]},$$
$$\mathbb{E}[X_t^2\hat{X}_t^1] = \frac{\mathbb{E}[X_t^2R_t^1]\mathbb{E}[X_t^1R_t^1]}{\mathbb{E}[(R_t^1)^2]},$$
$$\mathbb{E}[\hat{X}_t^1\hat{X}_t^2] = \frac{\mathbb{E}[X_t^1R_t^1]\mathbb{E}[X_t^2R_t^2]\mathbb{E}[R_t^1R_t^2]}{\mathbb{E}[(R_t^1)^2]\mathbb{E}[(R_t^2)^2]}, \tag{7.41}$$

(b) follows from

$$\mathbb{E}[X_t^1R_t^2] = \sqrt{P\alpha_{1,t}}(c+|\rho_t|),$$
$$\mathbb{E}[X_t^2R_t^1] = \sqrt{P\alpha_{2,t}}(c+|\rho_t|)\text{sgn}(\rho_t),$$
$$\mathbb{E}[R_t^1R_t^2] = 2cP+P|\rho_t|(1+c^2)+N\rho_z \tag{7.42}$$

(c) follows from (7.39); and (d) follows by defining $g(\rho_t)$.

Now we wish to find conditions on the parameters $\{\lambda_1,\lambda_2\}$ which ensure mean square stability of the two systems in (7.1) over the given white Gaussian interference channel. In order to find the values of the parameters $\{\lambda_1,\lambda_2\}$ for which the variance of the two state processes given by (7.39) can be made equal to zero as time goes to infinity, we make use of the following lemma.

Lemma 7.8.1. *For the recursive equation in* (7.40) *there exists at least one* $\rho^\star \in [0,1]$ *such that if* $|\rho_t| = \rho^\star$ *then* $|\rho_{t+k}| = \rho^\star$ *for all* $k\geq 0$, *where* $\rho^\star$ *is a root of one of the two polynomials* $\{f_1(\rho), f_2(\rho)\}$ *given in* (7.10). *Further, if* $\rho^\star$ *is a root of* $f_1(\rho)$, *then* $\rho_{t+k} = (-1)^k\rho^\star$, *and if* $\rho^\star$ *is a root of* $f_2(\rho)$, *then* $\rho_{t+k} = \rho^\star$ *for all* $k\geq 0$.

Proof. The proof can be found in Appendix 7.C. □

If we modify our encoding scheme such that ρ_2 becomes equal to $\rho^\star$ instead of zero, then $|\rho_t|$ will be equal to $\rho^\star$ for all $t \geq 2$. This modification[3] in the encoding scheme can be done as follows. Suppose in the initial transmissions (i.e., $t = 0, 1$) the two encoders transmit $S_0^1 = \sqrt{\frac{P}{\alpha_{1,0}}} X_0^1 + m$ and $S_1^2 = \sqrt{\frac{P}{\alpha_{2,1}}} X_1^2 + m$, where m is a Gaussian variable with zero mean and variance σ_m^2. In this way ρ_2 can take on any value between zero and one by varying σ_m^2. Thus, by choosing σ_m^2 such that $\rho_2 = \rho^\star$, we can rewrite (7.39) as

$$\begin{aligned}\alpha_{i,t+1} &= \alpha_{i,t}\lambda_i^2 \left(\frac{Pc^2\left(1 - \rho_t^{\star 2}\right) + N}{P(1 + c^2 + 2c\rho_t^\star) + N} \right) \\ &= \alpha_{i,2} \left(\lambda_i^2 \frac{Pc^2\left(1 - \rho_t^{\star 2}\right) + N}{P(1 + c^2 + 2c\rho_t^\star) + N} \right)^{t-2}. \end{aligned} \tag{7.43}$$

Although in the modified encoding scheme we have violated the average power constraint for the first two transmissions, its effect can be neglected for infinite time horizon. We observe from (7.43) that $\alpha_{i,t} \to 0$ as $t \to \infty$ if

$$\begin{aligned}&\left(\lambda_i^2 \frac{Pc^2\left(1 - \rho_t^{\star 2}\right) + N}{P(1 + c^2 + 2c\rho_t^\star) + N} \right) < 1 \\ &\Rightarrow \log(\lambda_i) < \frac{1}{2} \log \left(\frac{P\left(1 + c^2 + 2c\rho^\star\right) + N}{Pc^2\left(1 - \rho^{\star 2}\right) + N} \right), \end{aligned} \tag{7.44}$$

for $i \in \{1, 2\}$. The term on the right hand side in (7.44) is a monotonically increasing function of $\rho^\star$, therefore we choose $\rho^\star$ to be the largest among all roots in $[0, 1]$ of the two polynomials $\{f_1(\rho), f_2(\rho)\}$. This completes the proof of Theorem 7.4.1.

7.8.2 Proof of Theorem 7.4.2

The proof of Theorem 7.4.2 follows the proof of Theorem 7.4.1, since the transmit strategy in the *optimized scheme* is exactly the same as in the MMSE based scheme. The two schemes only differ in the control strategy. Therefore the expectations given in (7.36) and (7.42) are also valid for the *optimized scheme*. The variance of the

[3] At this point we modify the encoding scheme in order to artificially guarantee convergence of ρ_t to a fixed point. Numerical experiments suggest that ρ_t always converges to a fixed point starting from an arbitrary ρ_2, and this fixed point is unique.

state X_{t+1}^i under the optimized scheme is given by

$$\begin{aligned}
\alpha_{i,t+1} &:= \mathbb{E}[(X_{t+1}^i)^2] \\
&= \mathbb{E}\left[\left(\lambda_i X_t^i + U_t^i\right)^2\right] \\
&= \lambda_i^2 \mathbb{E}\left[\left(X_t^i\right)^2\right] + \mathbb{E}\left[\left(U_t^i\right)^2\right] + 2\lambda_i \mathbb{E}\left[\left(X_t^i U_t^i\right)\right] \\
&= \lambda_i^2 \alpha_{i,t}\left(1 + k^2\left(P\left(1 + c^2 + 2c|\rho_t|\right) + N\right) - 2k\sqrt{P}\left(1 + c|\rho_t|\right)\right), \qquad (7.45)
\end{aligned}$$

where the last equality follows from (7.7) and (7.36). The cross-correlation coefficient ρ_t between the two state processes for all $t \geq 3$ is given by

$$\begin{aligned}
\rho_{t+1} &= \frac{\mathbb{E}[X_{t+1}^1 X_{t+1}^2]}{\sqrt{\alpha_{1,t+1}\alpha_{2,t+1}}} \\
&\overset{(a)}{=} \frac{\lambda_1\lambda_2}{\sqrt{\alpha_{1,t+1}\alpha_{2,t+1}}}\mathbb{E}\left[\left(X_t^1 - k\sqrt{\alpha_{1,t}}R_t^1\right)\left(X_t^2 - k\sqrt{\alpha_{1,t}}\mathrm{sgn}(\rho_t)R_t^2\right)\right] \\
&= \frac{\lambda_1\lambda_2}{\sqrt{\alpha_{1,t+1}\alpha_{2,t+1}}}\mathbb{E}\Bigg[X_t^1 X_t^2 - k\sqrt{\alpha_{2,t}}\mathrm{sgn}(\rho_t)X_t^1 R_t^2 - k\sqrt{\alpha_{1,t}}X_t^2 R_t^1 \\
&\qquad + k^2\sqrt{\alpha_{1,t}\alpha_{2,t}}\mathrm{sgn}(\rho_t)R_t^1 R_t^2\Bigg] \\
&\overset{(b)}{=} \mathrm{sgn}(\rho_t)\lambda_1\lambda_2\sqrt{\frac{\alpha_{1,t}\alpha_{2,t}}{\alpha_{1,t+1}\alpha_{2,t+1}}}\Bigg(|\rho_t| + k^2 P\left(|\rho_t| + c^2|\rho_t| + 2c\right) \\
&\qquad - 2k\sqrt{P}\left(c + |\rho_t|\right)\Bigg) \\
&\overset{(c)}{=} \mathrm{sgn}(\rho_t)\frac{\left(|\rho_t| + k^2 P\left(|\rho_t| + c^2|\rho_t| + 2c\right) + k^2 N\rho_z - 2k\sqrt{P}\left(c + |\rho_t|\right)\right)}{\left(1 + k^2\left(P\left(1 + c^2 + 2c|\rho_t|\right) + N\right) - 2k\sqrt{P}\left(1 + c|\rho_t|\right)\right)} \\
&=: \mathrm{sgn}(\rho_t)g(\rho_t) \qquad (7.46)
\end{aligned}$$

where (a) follows from (7.7); (b) follows from (7.42); (c) follows from (7.45); and (d) follows by defining $g(\rho_t)$.

Now we wish to find conditions on the parameters $\{\lambda_1, \lambda_2\}$ which ensure mean square stability of the two systems in (7.1) over the given white Gaussian interference channel. In order to find the values of the parameters $\{\lambda_1, \lambda_2\}$ for which the variance of the two state processes given by (7.45) can be made equal to zero as time goes to infinity, we make use of the following lemma.

Lemma 7.8.2. *For the recursive equation in* (7.46) *there exists at least one* $\rho^\star \in [0, 1]$ *such that if* $|\rho_t| = \rho^\star$ *then* $|\rho_{t+l}| = \rho^\star$ *for all* $l \geq 0$*, where* $\rho^\star$ *is a solution of the quadratic given in* (7.12).

Proof. The proof can be found in Appendix 7.D. □

Like the proof of Theorem 7.4.1 we can modify our encoding scheme such that ρ_2 becomes equal to $\rho^\star$, then $|\rho_t|$ will be equal to $\rho^\star$ for all $t \geq 2$. This modification in the encoding scheme can be done as follows. Suppose in the initial transmissions (i.e., $t = 0, 1$) the two encoders transmit $S_0^1 = \sqrt{\frac{P}{\alpha_{1,0}}} X_0^1 + m$ and $S_1^2 = \sqrt{\frac{P}{\alpha_{2,1}}} X_1^2 + m$, where m is a Gaussian variable with zero mean and variance σ_m^2. In this way ρ_2 can take on any value between zero and one by varying σ_m^2. Thus by choosing σ_m^2 such that $\rho_2 = \rho^\star$, we can rewrite (7.45) as

$$\alpha_{i,t+1} = \alpha_{i,t}\lambda_i^2 \left(1 + k^2 \left(P\left(1 + c^2 + 2c|\rho^\star|\right) + N\right) - 2k\sqrt{P}\left(1 + c|\rho^\star|\right)\right). \quad (7.47)$$

We observe from (7.47) that $\alpha_{i,t} \to 0$ as $t \to \infty$ if

$$\lambda_i^2 \left(1 + k^2 \left(P\left(1 + c^2 + 2c|\rho^\star|\right) + N\right) - 2k\sqrt{P}\left(1 + c|\rho^\star|\right)\right) < 1$$
$$\Rightarrow \log(\lambda_i) < \frac{1}{2}\log\left(\frac{1}{1 + k^2\left(P\left(1 + c^2 + 2c|\rho^\star|\right) + N\right) - 2k\sqrt{P}\left(1 + c|\rho^\star|\right)}\right). \quad (7.48)$$

We have that for any choice of $k \in \mathbb{R}_+$ in the control functions, the two plants are mean square stable if (7.48) is satisfied. In order to maximize achievable stability region, we maximize the RHS of inequality in (7.48) over the design parameter $k \in \mathbb{R}^+$. This completes the proof of Theorem 7.4.2. Note that if we choose $k = \frac{\sqrt{P}(1+c|\rho^\star|)}{P(1+c^2+2c|\rho^\star|)}$, then the optimized scheme becomes equivalent to the MMSE based scheme. Therefore the optimized scheme performs at-least as good as the MMSE based scheme.

7.9 Conclusions

We have derived necessary as well as sufficient conditions for stabilization of LTI plants over a symmetric Gaussian interference channel. Sufficient conditions are obtained by employing linear memoryless sensing and control schemes. We showed that linear memoryless schemes perform quite well over a large range of system parameters. In some special cases they are exactly optimal or nearly optimal. We introduced an optimized linear scheme, in which the controllers can be optimized to maximize achievable stability region. This scheme always outperforms the conventional MMSE based (minimum variance) controller. MMSE based scheme can be quite inefficient in strong interference, as we observed an increasing gap between achievable stability regions of the two schemes as a function of cross-channel gain. This shows that MMSE based control scheme are not suitable for multi-terminal setups, unlike point-point Gaussian channels where it is actually optimal to use MMSE based scheme. Moreover, we also observed that TDMA based transmit scheme is very inefficient for stabilization in multi-plant multi-controller settings.

The necessary and sufficient conditions derived in the chapter reveal relationships between stabilizability and interference channel parameters such as transmit power, channel noise power, interference power, and correlation between the noises at the two controllers. From these relationships, we made some interesting observations: stability region significantly enlarges as the interference gets very strong. At high interference, the state processes of the two plants become highly correlated and thus the two sensors start cooperating by transmitting correlated outputs. Moreover, we have observed that cross-correlation between the noise variables at the two controllers can significantly effect stabilizability. Negative correlation usually helps and positive correlation hurts.

The stability results provided in this chapter can be extended for non-symmetric interference channel using the proposed schemes with further computations. One can also extend results for a setup where the links from the controllers to the plants are also white Gaussian communication channels. For this setup we can have an encoder at each control unit to encode the control action and a decoder at each plant to decode the transmitted value of the control action. As long as the encoders, the decoders, and the controllers are linear, the nature of the problem does not change and the stability results can be obtained cf. [YB11]. For future studies, an interesting step would be to find an optimal linear controller for the given problem. In particular, it would be interesting to find out how much improvement can be obtained by employing memory in the transmission and control schemes. Another interesting direction for future research would be to study LQG control problem in Gaussian interference networks. Clearly, one expects that estimation based controllers would be even worse for LQG control problem. It would be interesting to see the effect of the interference level on the quadratic cost. Can we achieve lower cost in the high interference regime similarly as we observe an improved stability in the high interference regime in this work?

Appendix

7.A Proof of Lemma 7.7.1

The proof of Lemma 7.7.1 follows from the same steps as the proof of Lemma 2.2.1, however, with minor differences due to zero process noise and the presence of two plants. Consider the following series of equalities:

$$\begin{aligned}
&\sum_{t=0}^{T-1} I\left(X_t^i; R_t^i | R_{[0,t\text{-}1]}^i\right) \\
&\quad = I\left(X_0^i; R_0^i\right) + \sum_{t=1}^{T-1} I\left(X_t^i; R_t^i | R_{[0,t\text{-}1]}^i\right)
\end{aligned}$$

$$
\begin{aligned}
&\overset{(b)}{=} I\left(X_0^i; R_0^i\right) + \sum_{t=1}^{T-1}\left(h\left(X_t^i | R_{[0,t\text{-}1]}^i\right) - h\left(X_t^i | R_{[0,t]}^i\right)\right)\\
&\overset{(c)}{=} \sum_{t=1}^{T-1}\left(h\left(A_i X_{t\text{-}1}^i + U_{t\text{-}1}^i | R_{[0,t\text{-}1]}^i\right) - h\left(X_t^i | R_{[0,t]}^i\right)\right) + I\left(X_0^i; R_0^i\right)\\
&\overset{(d)}{=} \sum_{t=1}^{T-1}\left(\log\left(|\det\left(A_i\right)|\right) + h\left(X_{t\text{-}1}^i | R_{[0,t\text{-}1]}^i\right) - h\left(X_t^i | R_{[0,t]}^i\right)\right) + I\left(X_0^i; R_0^i\right)\\
&= T\log\left(|\det\left(A_i\right)|\right) + h\left(X_0^i | R_0^i\right) - h\left(X_{T\text{-}1}^i | R_{[0,T\text{-}1]}^i\right) + I\left(X_0^i; R_0^i\right)\\
&= h\left(X_0^i\right) + (T-1)\log\left(|\det\left(A_i\right)|\right) - h\left(X_{T\text{-}1}^i | R_{[0,T\text{-}1]}^i\right), \qquad (7.49)
\end{aligned}
$$

where (a) follows from the definition of directed information [Mas90]; (b) follows by writing mutual information in terms of differential entropies; (c) follows from (7.1); (d) follows from the fact that for a matrix A and a random variable X, we have $h(AX) = h(X) + \log(|\det(A)|)$ [CT06, Theorem 8.6.4, (8.71)]; Using (7.49) the directed information rate is given by

$$
\begin{aligned}
&\liminf_{T\to\infty}\frac{1}{T} I\left(X_{[0,T\text{-}1]}^i \to R_{[0,T\text{-}1]}^i\right)\\
&\qquad = \liminf_{T\to\infty}\frac{1}{T}\left((T-1)\log\left(|\det\left(A_i\right)|\right) + h\left(X_0^i\right) - h\left(X_{T\text{-}1}^i | R_{[0,T\text{-}1]}^i\right)\right)\\
&\qquad \overset{(a)}{=} \log\left(|\det\left(A_i\right)|\right) - \limsup_{T\to\infty}\frac{1}{T} h\left(X_{T\text{-}1}^i | R_{[0,T\text{-}1]}^i\right)\\
&\qquad \overset{(b)}{\geq} \log\left(|\det\left(A_i\right)|\right) - \limsup_{T\to\infty}\frac{1}{T}\log\left((2\pi e)^n |K|\right)\\
&\qquad = \log\left(|\det\left(A_i\right)|\right),
\end{aligned}
$$

where the inequality (a) follows from $h\left(X_0^i\right) < \infty$; and (b) follows from the fact that for a mean square stable system there exists a matrix $K \succ 0$ with $\Lambda_t \preceq K$ for all t. Further, we know that for a given covariance matrix K the differential entropy is maximized by the Gaussian distribution.

7.B Proof of Lemma 7.7.2

Consider the following series of equalities

$$
\begin{aligned}
X_{t+1}^i &\overset{(a)}{=} A_i X_t^i + U_t^i\\
&\overset{(b)}{=} A_i^{t+1} X_0^i + \sum_{k=0}^{t} A^{t-k} U_k^i\\
&= A_i^{t+1} X_0^i + \sum_{k=0}^{t} A^k \pi_k^i\left(R_{[0,k]}^i\right) \qquad (7.50)
\end{aligned}
$$

$$= A_i^{t+1} X_0^i + \sum_{k=0}^{t} A^k \pi_k^i \left(S_{[0,k]}^i + Y_{[0,k]}^i \right), \tag{7.51}$$

where (a) follows from (7.1); (b) follows by recursively applying (a); (7.50) follows from $U_t^i = \pi_t^i \left(R_{[0,t]}^i \right)$; and (7.51) follows from $R_t^i = S_t^i + hS_t^j + Z_t^i$ and $Y_t^i = hS_t^i + Z_t^j$ for $i \neq j$. From (7.50) we see that $X_t^i = A_i^t X_0^i + \mu_t^i \left(R_{[0,t\text{-}1]}^i \right)$, where $\mu_t^i : \mathbb{R}^{t\text{-}1} \to \mathbb{R}$. Since $S_t^i = f_t^i \left(X_{[0,t]}^i \right)$ and $X_t^i = A_i^t X_0^i + \mu_t^i \left(R_{[0,t\text{-}1]}^i \right)$, we have $S_t^i = \upsilon_t^i \left(X_0^i, R_{[0,t\text{-}1]}^i \right)$, where $\upsilon_t^i : \mathbb{R}^t \to \mathbb{R}$. Moreover, S_t^i can also be written as

$$\begin{aligned} S_t^i &= f_t^i \left(X_{[0,t]}^i \right) \\ &\overset{(a)}{=} g_t^i \left(X_0^i, S_{[0,t-1]}^i, Y_{[0,t-1]}^i \right) \\ &\overset{(b)}{=} \nu_t^i \left(X_0^i, Y_{[0,t-1]}^i \right), \end{aligned} \tag{7.52}$$

where (a) follows from (7.51) and by defining $g_t^i : \mathbb{R}^{2t\text{-}1} \to \mathbb{R}$; and (b) follows from recursively applying (a) and by defining $\nu_t^i : \mathbb{R}^t \to \mathbb{R}$.

7.C Proof of lemma 7.8.1

Consider the recursive equation $\rho_{t+1} = \text{sgn}(\rho_t) g(\rho_t)$ given in (7.40). Based on $g(\rho)$ we define two polynomials $\{f_1(\rho), f_2(\rho)\}$ as,

$$\begin{aligned} f_1(\rho) &:= l_1(\rho + g(\rho)), \\ f_2(\rho) &:= l_2(-\rho + g(\rho)), \end{aligned} \tag{7.53}$$

where l_1 and l_2 are two non-zero scalars chosen such that the leading coefficients of the two polynomials are equal to one, i.e., $f_1(\rho)$ and $f_2(\rho)$ are monic polynomials. With some algebraic manipulations, these polynomials are given in a simplified form in (7.10). Suppose that there exists at least one $\rho^\star \in [0, 1]$ which is a root of one of the two polynomials $\{f_1(\rho), f_2(\rho)\}$. (In other words we are assuming that there exists at least one $\rho^\star$ in the interval $[0, 1]$ which is either a solution to $-\rho = g(\rho)$ or a solution to $\rho = g(\rho)$.) Then if at any time $|\rho_t| = \rho^\star$ and $\rho^\star$ is a root of $f_1(\rho)$ or $f_2(\rho)$, then will we have $|\rho_{t+k}|$ equal to $\rho^\star$ for all $k \geq 0$. If $\rho^\star$ is a root of $f_1(\rho)$ (or a solution to $-\rho = g(\rho)$), then $\rho_{t+k} = (-1)^k \rho^\star$; and if $\rho^\star$ is a root of $f_2(\rho)$ (or a solution to $\rho = g(\rho)$), then $\rho_{t+k} = \rho^\star$ for all $k \geq 0$. Therefore in order to prove Lemma 7.8.1, we need to show existence of a root in the interval [0,1] of at least of one of the two polynomials $\{f_1(\rho), f_2(\rho)\}$.

We first consider the polynomial $f_1(\rho)$. Since $f_1(\rho)$ is a continuous function, it will have at least one root in the interval $[0, 1]$ if it changes sign within this interval. That is at least one root exists in $[0, 1]$ if we have either $\{f_1(0) \geq 0, f_1(1) \leq 0\}$ or

$\{f_1(0) \leq 0, f_1(1) \geq 0\}$. By evaluating $f_1(\rho)$ for $\rho = 0, 1$, we get

$$f_1(0) = 1 + \frac{N(2h - \rho_z)}{2h^3 P},$$
$$f_1(1) = -\frac{PN\left(h^2(\rho_z + 1) + 2h(\rho_z + 1)(\rho_z + 1) + 2N^2\right)}{2h^3 P^2},$$

where $\rho_z \in [-1, 1]$ and $P, N, h > 0$. For all $\rho_z \in [-1, 1]$, we have $f_1(1) < 0$. Further $f_1(0) > 0$ for $-1 \leq \rho_z \leq 2h\left(1 + \frac{h^2 P}{N}\right)$, and $f_1(0) < 0$ for $2h\left(1 + \frac{h^2 P}{N}\right) < \rho_z \leq 1$. Therefore the existence of a root of $f_1(\rho)$ in the interval $[0, 1]$ is guaranteed for all $\rho_z \in [-1, 2h(1 + \frac{h^2 P}{N})]$. Since we could not show the existence of a root in $[0, 1]$ of $f_1(\rho)$ when $\rho_z > 2h(1 + \frac{h^2 P}{N})$, we now investigate the polynomial $f_2(\rho)$ for all $\rho_z > 2h(1 + \frac{h^2 P}{N})$. By evaluating $f_2(\rho)$ for $\rho = 0, 1$, we get

$$f_2(0) = -\left(1 + \frac{N(2h - \rho_z)}{2h^3 P}\right),$$
$$f_2(1) = \frac{N(1 + h)^2 (\rho_z - 1)}{2h^3 P}.$$

We observe that $f_2(1) \leq 0$ for $\rho_z \in [-1, 1]$, and $f_2(0) > 0$ for $\rho_z > 2h(1 + \frac{h^2 P}{N})$. Therefore a root in $[0, 1]$ of $f_2(\rho)$ exists when $\rho_z > 2h(1 + \frac{h^2 P}{N})$. Hence we have shown the existence of a root in the interval [0,1] of at least of one of the two polynomials for all values of $\{P, h, N, \rho_z\}$. Further we observe that $f_2(1) = 0$ for $\rho_z = 1$, from which Theorem 7.5.4 follows.

7.D Proof of lemma 7.8.2

We define the following functions based on the recursive equation in (7.46),

$$J_1(\rho) := \rho + g(\rho),$$
$$J_2(\rho) := -\rho + g(\rho),$$

where $g(\rho)$ is given in (7.46). A value $\rho = \rho^\star$ for which either $J_1(\rho) = 0$ or $J_2(\rho) = 0$ translates to $\rho = |g(\rho)|$. Now if the recursive function $\rho_{t+1} = \text{sgn}(\rho) g(\rho)$ achieves a the point $\rho^\star$ at time t_1 then $|\rho_{t+1}| = \rho^\star$ for all $t \geq t_1$. If $\rho^\star$ is root of $J_1(\rho)$ then $\rho_{t+n} = (-1)^n \rho^\star$, and if $\rho^\star$ is root of $J_2(\rho)$ then $\rho_{t+n} = \rho^\star$.

To find the existence of a root, we evaluate the functions $\{J_1(\rho), J_2(\rho)\}$ for $\rho = 0$ and $\rho = 1$.

$$J_1(1) = 1 + \frac{k^2 N \rho_z + \left(1 - k\sqrt{P}(1 + c)\right)^2}{k^2 N + \left(1 - k\sqrt{P}(1 + c)\right)^2},$$

$$J_2(1) = -\frac{k^2 N (1-\rho_z)}{k^2 N + \left(1 - k\sqrt{P}(1+c)\right)^2},$$

$$J_1(0) = J_2(0) = g(0) = \frac{2ck\sqrt{P}\left(k\sqrt{P}-1\right) + k^2 N \rho_z}{k^2 (c^2 P + N) + \left(k\sqrt{P}-1\right)^2}.$$

It is seen that $J_1(1) > 0$ and $J_2(1) \leq 0$ for all channel parameters $c, P, N \in \mathbb{R}$, $\rho_z \in [-1, 1]$. Since $J_1(0) = J_2(0) = g(0)$, it follows that there is always a root $\rho^\star \in [0, 1]$ in one of the two functions $\{J_1(\rho), J_2(\rho)\}$. If $2ck\sqrt{P}\left(k\sqrt{P}-1\right) + k^2 N \rho_z < 0$, then $J_1(0) < 0$and $\rho^\star \in [0, 1]$ is a root of J_1. If $2ck\sqrt{P}\left(k\sqrt{P}-1\right) + k^2 N \rho_z > 0$, then $J_2(0) < 0$ and $\rho^\star \in [0, 1]$ is a root of J_2. This leads to the conclusion that a fixed point $\rho^\star$ always exists. Furthermore, with some algebraic simplifications it can be shown that $\rho^\star$ is the unique solution of

$$\rho^2 + \beta\rho + \gamma = 0, \tag{7.54}$$

where $\beta = \frac{\left(2\left(k\sqrt{P}-1\right)^2 + k^2\left(2c^2P+N\right)\right)}{2ck\sqrt{P}\left(k\sqrt{P}-1\right)}$, $\gamma = \frac{k^2 N \rho_z}{2ck\sqrt{P}\left(k\sqrt{P}-1\right)} + 1$ if $kN\rho_z < 2c\sqrt{P}(1 - k\sqrt{P})$, and $\beta = \frac{k^2 N}{2ck\sqrt{P}\left(k\sqrt{P}-1\right)}$, $\gamma = -\frac{k^2 N \rho_z}{2ck\sqrt{P}\left(k\sqrt{P}-1\right)} - 1$ if $kN\rho_z > 2c\sqrt{P}(1 - k\sqrt{P})$.

Chapter 8

Conclusions

The thesis considered remote stabilization and control of linear time invariant systems over communication networks. The results were presented in three main parts addressing several open problems in sensing and control of LTI systems over various network topologies with different objectives. Although large networked control systems are the main motivation, we have focused on systems with fewer components to understand the fundamental principles. An in-depth analysis have provided us with intuition to guide the understanding and design of networked control systems. In this chapter we highlight the main findings of the thesis and provide some directions for future research on this subject.

In parts I and III, we studied mean-square stabilization of LTI systems over some fundamental abstractions of communication scenarios such as point-to-point channels, relay channels, broadcast channels, multiple-access channels, and interference channels. We derived necessary conditions as well as sufficient conditions for stabilization over these channels. Necessary conditions are derived using information theoretic arguments. Sufficient conditions are derived using delay-free sensing and control methods that are suitable for delay-sensitive control applications. By comparing these necessary and sufficient conditions, we realize how good delay-free policies are in various network settings. In the following we summarize our key findings. We observed that linear time invariant sensing and control policies are insufficient for stabilization of a plant even over a scalar point-to-point Gaussian channel. However a linear time varying scheme is shown to be optimal for stabilization of LTI plants over a large class of vector Gaussian channels, even when the source channel matching principle does not hold. For a given system and a given channel, the optimality of the linear policies can be verified by solving a linear program. Further, we discussed that a delay-free non-linear time varying scheme policy is always optimal for stabilization of noiseless LTI plants over point-to-point Gaussian channels. Then we extended our study to stabilization of an LTI plant over relay networks, where relays are allowed to employ arbitrary transmit policies under an average power constraint and arbitrary control policies without any constraint. We considered three fundamental relay network topologies (cascade, parallel, non-orthogonal) that are building blocks of a more general network. We observed that delay-free linear sensing and control strategies are asymptotically optimal for a symmetric two-hop non-orthogonal relay network, with respect to the number of relay nodes or the transmit powers at the relays. For a two-hop parallel relay network, we observed a non-decreasing gap between the sufficient conditions derived using linear policies and the necessary conditions, as a function of the number of relay nodes in the network. The gap remains bounded as the number of relays approaches infinity. A very useful observation is that the delay-free linear strategies become increasingly inefficient for stabilization as the number of hops in a relay network increases, motivating the need for non-linear schemes with or without memory. Finally, we studied stabilization of multiple plants over multi-user Gaussian channels without intermediate relay nodes. We observed that delay-free linear time varying sensing and control approaches are still good choice for multi-user channels in most cases, when designed appropriately. In fact in some very special cases, delay-free linear schemes are exactly optimal. An important observation is that TDMA based sensing methods and state estimation based control strategies are not suitable candidates for stabilization of multiple plants over multi-user networks. We introduced an optimized delay-free linear control scheme that can significantly outperform the traditional state estimation (MMSE) based scheme over a wide range of communication channel parameters. The gap between the stability region achieved by the optimized scheme and the stability region achieved by the MMSE based scheme increases as a function of the cross-channel gain (or the interference power).

The necessary and sufficient condition derived in this thesis not only explain how good delay-free sensing and control policies are in various network settings, but they also shed light on the relationship between stabilizability of plant(s) and the communication network parameters. They quantify the effect of channel parameters (transmit power, noise powers, channel gains, interference) on stability. For instance, they tell us how communication resources such as transmit powers should be allocated in the network in order to stabilize the given plant(s). For parallel networks, we observed that an optimal strategy might be to transmit information to the remote controller using only on a subset of the available channels. Similarly in non-orthogonal relay networks, it might be more beneficial to turn off certain relay nodes than employing an AF (linear) strategy if the information received at the relays is very noisy. For interference channels, we have made an interesting observation that interference within the communication network can actually be useful when the objective is to stabilize the plants. In all network settings, we have observed a close connection between information rates and stability that can be achieved over a given channel, motivating the use of information theory in understanding the problems in networked control systems. In particular, we have discussed relationships between unstable modes of the plant and directed information rates between the sequence of channel inputs (or state variables) and the sequence of channel outputs. By upper bounding these information rates, we have derived necessary conditions for stabilization in various network settings. Moreover, we have also shown that the directed information rates under linear policies can characterize sufficient conditions for stabilization.

In Part II of the thesis, we studied how good or bad linear policies are when the objective is to minimize a quadratic cost function of the state process. We considered two basic network settings: i) cascade network and ii) parallel network. For a three node cascade network, we showed that although linear schemes are person-by-person optimal, they are not globally optimal in general. We proposed a non-linear transmission strategy based on a three-level quantization function. The proposed non-linear scheme outperforms the best linear scheme in low SNR regions. When the channels are very noisy, one intuition on why the proposed non-linear strategy is superior may be that it does not amplify the large values of channel noise at its input unlike the linear (amplify-and-forward) strategy. For parallel networks, linear policies are not even person-by-person optimal. One reason for inefficiency of linear sensing schemes is that the signals transmitted on the parallel channels under linear schemes can be highly correlated if the measurement noise powers at the senors are low. We have proposed a non-linear sensing scheme based on a sawtooth function that outperforms the best linear scheme by reducing correlation between the information transmitted on different channels. For the problem of stabilization over parallel networks, we have observed in Part I of the thesis that delay-free linear policies are optimal for a large class of plant and channel pairs. However, this result does not carry over to the case when the objective is optimization (cost-minimization) and not stabilization. Thus, the sensing and control policies which are suitable for stabilization over a network, may not be suitable when we have an optimization (cost minimization) problem. That is, a good choice of sensing and control scheme should be based not only on the information structure of the system but also on the system's objective.

Our quest for understanding open problems related to sensing and control over Gaussian networks has given rise to several interesting questions and at the same time made us aware of some very fundamental problems that are still unsolved. In the following we highlight some important problems that we think would be interesting to explore in the future. This thesis made some progress on the problem of stabilizing a multi-dimensional plant over a multi-dimensional Gaussian channel, however the problem of finding optimal sensing and control polices that minimize a quadratic cost function of the state and control process remains open in both static and dynamic settings. In the dynamic case, even optimal linear policies are not known. Our study on stabilization and control over relay networks was one of the earliest, where we used amplify-and-forward based relaying strategies. We realized that for relay networks, amplify-and-forward relaying strategies become increasingly inefficient as the number of relays connected in cascade increases. One may look for better methods. In fact, in Chapter 4 we have seen examples of simple quantization based policies that worked well for minimizing end-to-end distortion in cascade relay channels. One can use similar policies for stabilization over cascade relay networks. Similarly the non-linear sensing policies introduced in Chapter 5 can be used for stabilization over parallel relay networks. We have realized that it's a challenging task to derive sufficient conditions for stabilization under these non-linear policies. One extension to Chapter 4 would be to find optimal linear policies for transmission of a vector Gaussian source over a vector Gaussian cascade relay channel. Of course it is also interesting to show if optimal policies exist. For multi-user channels, we know very little today. One natural step would be to look for optimal delay-free linear policies, which are not known even in the scalar settings. Extensions to schemes with memory and non-linear methods would be further steps. For interference channels, it would be interesting to formulate an LQG control problem and then see if the presence of strong interference can lead to lower cost, as we have observed in the case of stabilization. We observed that stabilization is not affected by an additive process noise for point-to-point channels. It would be very useful to show if this is also valid for general Gaussian networks. Another potential extension to this thesis would be to consider some sort of cooperation between remote controllers in multi-user settings.

As the author of the thesis, it would be interesting for me if someone pursues the above mentioned research problems or other problems related to this thesis. I would be more than happy to discuss and collaborate in studying problems of mutual interest.

Bibliography

[AG87] E. Ayanoglu and R. M. Gray, "The design of joint source and channel trellis waveform coders," *IEEE Trans. Inf. Theory*, vol. 33, no. 6, pp. 855–865, 1987.

[AM09] K. J. Aström and R. M. Murray, *Feedback Systems: An Introduction for Scientists and Engineers.* Princeton University Press, 2009.

[AZWS11] M. Andersson, A. A. Zaidi, N. Wernersson, and M. Skoglund, "Nonlinear distributed sensing for closed-loop control over Gaussian channels," in *IEEE Swe–CTW*, October 2011, pp. 19–23.

[Bai99] J. Baillieul, "Feedback designs for controlling device arrays with communication channel bandwidth constraints," in *ARO Workshop on Smart Structures*, Aug. 1999.

[Bas80] T. Basar, "A trace minimization problem with applications in joint estimation and control under nonclassical information," *Journal of Optimization Theory and Applications*, vol. 31, no. 3, pp. 343–359, 1980.

[Bas08] T. Basar, "Variations on the theme of the witsenhausen counterexample," in *IEEE CDC*, 2008, pp. 1614–1619.

[BB87] R. Bansal and T. Başar, "Solutions to a class of linear-quadratic-gaussian LQG stochastic team problems with nonclassical information," *Systems Control Letters*, vol. 9, no. 2, pp. 125–130, 1987.

[BB89] ——, "Simultaneous design of measurement and control strategies for stochastic systems with feedback," *Automatica*, vol. 25, no. 5, pp. 679–694, 1989.

[BB94] T. Başar and R. Bansal, "Optimum design of measurement channels and control policies for linear-quadratic stochastic systems," *European Journal of Operations Research*, vol. 73, no. 2, pp. 226–236, March 1994.

[Ber71] T. Berger, *Rate Distortion Theory: A mathematical basis for data compression.* Englewood Cliffs, NJ: Prentice-Hall, 1971.

[Ber05] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, 3rd ed. Athena Scientific, 2005.

[BMF07] J. H. Braslavsky, R. H. Middleton, and J. S. Freudenberg, "Feedback stabilization over signal-to-noise ratio constrained channels," *IEEE Trans. Autom. Control*, vol. 52, no. 8, pp. 1391–1403, Aug. 2007.

[BW09] S. Bross and M. Wigger, "On the relay channel with receiver-transmitter feedback," *IEEE Trans. Inf. Theory*, vol. 55, no. 1, pp. 275–291, 2009.

[BZOS12] I. Bilal, A. A. Zaidi, T. J. Oechtering, and M. Skoglund, "Managing interference for stabilization over wireless channels," in *Intelligent Control (ISIC), 2012 IEEE International Symposium on*, 2012, pp. 933–938.

[BZOS13a] ——, "Feedback stabilization using relay in interference channels," in *IEEE SPAWC 2013*, June 2013.

[BZOS13b] ——, "An optimized linear scheme for stabilization over multi-user gaussian networks," in *ITA*, Feb. 2013.

[CFD08] C. D. Charalambous, A. Farhadi, and S. Z. Denic, "Control of continuous-time linear Gaussian systems over additive Gaussian wireless fading channels: A separation principle," *IEEE Trans. Autom. Control*, vol. 53, no. 4, pp. 1013–1019, May 2008.

[CT06] T. Cover and J. Thomas, *Elements of information theory.* John Wiley Sons, Inc., 2006.

[DT62] R. L. Dobrushin and B. S. Tsybakov, "Information transmission with additional nosie," *IRE Trans. Inf. Theory*, vol. 18, pp. 293–304, 1962.

[Eli04] N. Elia, "When Bode meets Shannon: control–oriented feedback communication schemes," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1477–1488, 2004.

[EM01] N. Elia and S. K. Mitter, "Stabilization of linear systems with limited information," *IEEE Trans. Autom. Control*, vol. 46, no. 9, pp. 1384–1400, Sep. 2001.

[FA11] A. Farhadi and N. U. Ahmed, "Suboptimal decentralized control over noisy communication channels," *Systems Control Letters*, vol. 60, pp. 285–293, April 2011.

[FD] "The free dictionary," http://http://www.thefreedictionary.com/system, accessed: 2013-10-01.

[Fin65] T. Fine, "Optimum mean-square quantization of a noisy input," *IEEE Trans. Inf. Theory*, vol. 11, pp. 293–294, April 1965.

[FMS10] J. S. Freudenberg, R. H. Middleton, and V. Solo, "Stabilization and disturbance attenuation over a Gaussian communication channel," *IEEE Trans. Autom. Control*, vol. 55, no. 3, pp. 795–799, March 2010.

[FZFS13] H. Farhadi, A. A. Zaidi, C. Fischione, and M. Skoglund, "Distributed interference alignment and power control with imperfect channel knowledge," in *IEEE BlackSeaCom*, July 2013.

[Gal68] R. G. Gallager, *Information Theory and Reliable Communication.* John Wiley and Sons, 1968.

[Gas07] M. Gastpar, "Uncoded transmission is exactly optimal for a simple Gaussian sensor network," in *Information Theory and Applications Workshop*, 2007.

[Gas08] ——, "Uncoded transmission is exactly optimal for a simple Gaussian sensor network," *IEEE Trans. Inf. Theory*, vol. 54, pp. 5247–5251, 2008.

[GC82] A. E. Gamal and M. Costa, "The capacity region of a class of deterministic interference channels," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 343–346, March 1982.

[GDH+09] V. Gupta, A. F. Dana, J. P. Hespanha, R. M. Murray, and B. Hassibi, "Data transmission over networks for estimation and control," *IEEE Trans. Autom. Control*, vol. 54, no. 8, pp. 1807–1819, Aug. 2009.

[GJ07] K. S. Gomadam and S. A. Jafar, "Optimal relay functionality for SNR maximization in memoryless relay networks," *IEEE Journal on Selected Areas in Communications 25*, vol. 25, no. 2, pp. 390–401, 2007.

[GLW10] M. Gastpar, A. Lapidoth, and M. Wigger, "When feedback doubles the prelog in AWGN networks," *IEEE Trans. Inform. Theory*, Submitted 2010. [Online]. Available: http://arxiv.org/abs/1003.6082

[GRV03] M. Gastpar, B. Rimoldi, and M. Vetterli, "To code, or not to code: lossy source-channel communication revisited," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1147–1158, May 2003.

[GV05] M. Gastpar and M. Vetterli, "On the capacity of large Gaussian relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 765–779, March 2005.

[Har49] G. H. Hardy, *Divergent Series.* Oxford University Press, 1949.

[Hay96] M. Hayes, *Statistical digital signal processing and modelling.* John Wiley Sons, Inc., 1996.

[HJ90] R. A. Horn and C. R. Johnson, *Matrix Analysis.* Cambridge University Press, 1990.

[Ho80] Y. C. Ho, "Team decision theory and information structures," *Proceedings of the IEEE*, vol. 68, no. 6, pp. 644–654, 1980.

[Jr.65] T. G. Jr., "Theoretical limitations on the transmission of data from analog sources," *IEEE Trans. Inf. Theory*, vol. 11, no. 4, pp. 558–567, Oct. 1965.

[KGG05] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, pp. 3037–3063, 2005.

[KGL10] U. Kumar, V. Gupta, and J. N. Laneman, "Sufficient conditions for stabilizability over Gaussian relay channel and cascade channels," in *IEEE CDC*, December 2010, pp. 4765–4770.

[KGL11] ——, "On stability across a Gaussian product channel," in *IEEE CDC*, December 2011, pp. 3142–3147.

[KZS10] M. N. Khormuji, A. A. Zaidi, and M. Skoglund, "Interference management using nonlinear relaying," *IEEE Trans. Commun.*, vol. 58, no. 7, pp. 1924 –1930, July 2010.

[LDA11] A. S. Leong, S. Dey, and J. Anand, "Optimal LQG control over continuous fading channels," in *18th IFAC World Congress*, August 2011.

[LM11] G. M. Lipsa and N. C. Martins, "Optimal memoryless control in Gaussian noise: A simple counterexample," *Automatica*, vol. 47, pp. 552–558, March 2011.

[LP76] K.-H. Lee and D. P. Petersen, "Optimal linear coding for vector channels," *IEEE Trans. Commun.*, vol. 24, no. 12, pp. 1283–1290, Dec. 1976.

[Mas90] J. L. Massey, "Causality, feedback and directed information," in *IEEE ISITA*, 1990.

[MCF13] P. Minero, L. Coviello, and M. Franceschetti, "Stabilization over markov feedback channels: The general case," *IEEE Trans. Autom. Control*, vol. 58, no. 2, pp. 349–362, Feb. 2013.

[MD08] N. C. Martins and M. A. Dahleh, "Feedback control in the presence of noisy channels: "Bode-like" fundamental limitations of performance," *IEEE Trans. Autom. Control*, vol. 53, no. 7, pp. 1604–1615, Aug. 2008.

[MDE06] N. C. Martins, M. A. Dahleh, and N. Elia, "Feedback stabilization of uncertain systems in the presence of a direct link," *IEEE Trans. Autom. Control*, vol. 51, no. 3, pp. 438–447, March 2006.

[MMRY12] A. Mahajan, N. C. Martins, M. C. Rotkowitz, and S. Yüksel, “Information structures in optimal decentralized control,” in *IEEE CDC*, 2012, pp. 1291–1306.

[MRFB09] R. H. Middleton, A. J. Rojas, J. S. Freudenberg, and J. H. Braslavsky, “Feedback stabilization over a first order moving average Gaussian noise channel,” *IEEE Trans. Autom. Control*, vol. 54, no. 1, pp. 163–167, Jan. 2009.

[MS04] A. S. Matveev and A. V. Savkin, “The problem of LQG optimal control via a limited capacity communication channel,” *Systems & Control Letters*, vol. 53, no. 1, pp. 51–64, Sep. 2004.

[MS07] ——, “An analogue of shannon information theory for detection and stabilization via noisy discrete communication channels,” *SIAM J. Control Optim.*, vol. 46, no. 4, pp. 1323–1367, 2007.

[NE00] G. N. Nair and R. J. Evans, “Stabilization with data-rate-limited feedback: tightest attainable bounds,” *Systems & Control Letters*, vol. 41, no. 1, pp. 49–56, Sep. 2000.

[NE04] ——, “Stabilizability of stochastic linear systems with finite feedback data rates,” *SIAM J. Control Optim.*, vol. 43, no. 2, pp. 413–436, 2004.

[OLYC84] L. Ozarow and S. Leung-Yan-Cheong, “An achievable region and outer bound for the gaussian broadcast channel with feedback,” *IEEE Trans. Inf. Theory*, vol. 30, no. 4, pp. 667—-671, July 1984.

[Oza84] L. H. Ozarow, “The capacity of the white gaussian multiple-access channel with feed-back,” *IEEE Trans. Inf. Theory*, vol. 30, no. 4, pp. 623–629, July 1984.

[Pil69] R. J. Pilc, “The optimum linear modulator for a Gaussian source used with a Gaussian channel,” *The Bell System Tech. Journal*, pp. 3075–3089, Nov. 1969.

[Rad62] R. Radner, “Team decision problems,” *Annals of Mathematical Statistics*, vol. 33, pp. 857–881, 1962.

[Rud76] W. Rudin, *Principles of Mathematical Analysis.* McGraw-Hill Book Company, 1976.

[SA74] N. R. Sandell and M. Athans, “Solution of some nonclassical LQG stochastic decision problems,” *IEEE Trans. on Automatic Control*, vol. 19, no. 2, pp. 108–116, 1974.

[Sah01] A. Sahai, “Anytime information theory,” Ph.D. dissertation, MIT Massachusetts Institute of Technology, 2001.

[Sat81] H. Sato, "The capacity of the Gaussian interference channel under strong interference," *IEEE Trans. Inform. Theory*, vol. 27, no. 6, pp. 786 – 788, nov. 1981.

[SDO11] E. I. Silva, M. S. Derpich, and J. Ostergaard, "A framework for control system design subject to average data-rate constraints," *IEEE Trans. Autom. Control*, vol. 56, no. 8, pp. 1886–1899, Aug. 2011.

[SGQ10] E. I. Silva, G. C. Goodwin, and D. E. Quevedo, "Control system design subject to SNR constraints," *Automatica*, vol. 46, no. 2, pp. 428–436, 2010.

[SK66] Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback–I: No bandwidth constraint," *IEEE Trans. Inf. Theory*, vol. 12, no. 2, pp. 172–182, 1966.

[SM06] A. Sahai and S. Mitter, "The necessity and sufficiency of anytime capacity for stabilization of a linear system over noisy communication links–part I: Scalar systems," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3369–3395, 2006.

[SM11] Z. Shu and R. H. Middleton, "Stabilization over power-constrained parallel Gaussian channels," *IEEE Trans. Autom. Control*, vol. 56, no. 7, pp. 1718–1724, July 2011.

[Söd02] T. Söderström, *Discrete-time Stochastic Systems.* Springer, 2002.

[ST09] C. Suh and D. Tse, "Symmetric feedback capacity of the Gaussian interference channel to within one bit," in *IEEE ISIT*, 2009, pp. 1609–1613.

[ST11] ——, "Feedback capacity of the Gaussian interference channel to within 2 bits," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2667–2685, 2011.

[Str65] C. Striebel, "Sufficient statistics in the optimum control of stochastic systems," *J. Math. Anal. Appl.*, vol. 12, pp. 576–592, 1965.

[SVZ98] S. Shamai, S. Verdu, and R. Zamir, "Systematic lossy source/channel coding," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 564–579, March 1998.

[Tat03] S. Tatikonda, "Some scaling properties of large distributed control systems," in *IEEE CDC*, Dec. 2003, pp. 3142–3147.

[TM04] S. Tatikonda and S. Mitter, "Control over noisy channels," *IEEE Trans. Autom. Control*, vol. 49, no. 7, pp. 1196–1201, 2004.

[TSM04] S. Tatikonda, A. Sahai, and S. Mitter, "Stochastic linear control over a communication channel," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1549–1561, Sept. 2004.

[TSM09] S. Tatikonda, S., and Mitter, "The capacity of channels with feedback," *IEEE Trans. Inf. Theory*, vol. 55, no. 1, pp. 323–349, Jan. 2009.

[TV05] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication.* Cambridge University Press, 2005.

[Vai89] V. A. Vaishampayan, "Combined source-channel coding for bandlimited waveform channels," Ph.D. dissertation, University of Maryland, 1989.

[VSC12] F. J. Vargas, E. I. Silva, and J. Chen, "Stabilization of TITO systems over parallel SNR-constraint AWN channels," in *Proc. 7th IFAC Symposium on Robust Control Design*, 2012.

[WB99] W. S. Wong and R. W. Brockett, "Systems with finite communication bandwidth constraints-II: stabilization with limited information feedback," *IEEE Trans. Autom. Control*, vol. 44, no. 5, pp. 1049–1053, May 1999.

[Wit68] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM J. Control*, vol. 6, pp. 131–147, 1968.

[Wit71] ——, "Separation of estimation and control for discrete time systems," *Proceedings of the IEEE*, vol. 59, no. 11, pp. 1557–1566, Nov. 1971.

[Won68] W. M. Wonham, "On the separation theorem of stochastic control," *SIAM J. Control.*, vol. 6, pp. 312–326, 1968.

[WS09a] N. Wernersson and M. Skoglund, "Nonlinear coding and estimation for correlated data in wireless sensor networks," *Communications, IEEE Transactions on*, vol. 57, no. 10, pp. 2932 –2939, october 2009.

[WS09b] ——, "Nonlinear coding and estimation for correlated data in wireless sensor networks," *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 2932–2939, October 2009.

[WV83] J. Walrand and P. Varaiya, "Optimal causal coding-decoding problems," *IEEE Trans. Inform. Theory*, vol. 29, pp. 814–820, 1983.

[XCLG08] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A. J. Goldsmith, "Linear coherent decentralized estimation," *IEEE Trans. Signal Processing*, vol. 56, no. 2, pp. 757–770, Feb. 2008.

[YB11] S. Yüksel and T. Başar, "Control over noisy forward and reverse channels," *IEEE Trans. Autom. Control*, vol. 56, pp. 1014–1029, May 2011.

[YB13] S. Yüksel and T. Başar, *Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints.* Boston, MA: Birkhäuser, 2013.

[YL11] S. Yüksel and T. Linder, "Optimization and convergence of observation channels in stochastic control," in *ACC*, June 2011, pp. 637–642.

[YT07] S. Yüksel and S. Tatikonda, "Distributed sensing and control over communication channels and a counterexample," in *Allerton Conference on Communication, Control, and Computing*, 2007, pp. 260–267.

[YT09] ——, "A counterexample in distributed optimal sensing and control," *IEEE Trans. Autom. Control*, vol. 54, no. 4, 2009.

[YTCW09] L. Yiqian, E. Tuncel, J. Chen, and S. Weizhou, "Optimal tracking performance of discrete-time systems over an additive white noise channel," in *Proc. IEEE CDC*, 2009.

[Yük10] S. Yüksel, "On optimal causal coding of partially observed markov sources under classical and non-classical information structures," in *IEEE ISIT*, 2010, pp. 81–85.

[Yük12] ——, "Characterization of information channels for asymptotic mean stationarity and stochastic stability of non-stationary/unstable linear systems," *IEEE Trans. Inf. Theory*, vol. 58, no. 10, pp. 6332–6354, Oct. 2012.

[ZKYS09a] A. A. Zaidi, M. N. Khormuji, S. Yao, and M. Skoglund, "Optimized analog network coding strategies for the white Gaussian multiple-access relay channel," in *IEEE ITW*, Oct. 2009, pp. 460–464.

[ZKYS09b] ——, "Optimized mappings for dimension-expansion relaying," in *IEEE SPAWC*, June 2009, pp. 474–478.

[ZKYS09c] ——, "Rate-maximizing mappings for memoryless relaying," in *Proc. IEEE ISIT*, 2009.

[ZOS10a] A. A. Zaidi, T. J. Oechtering, and M. Skoglund, "Rate sufficient conditions for closed–loop control over AWGN relay channels," in *IEEE ICCA*, June 2010, pp. 602–607.

[ZOS10b] ——, "Sufficient conditions for closed-loop control over multiple-access and broadcast channels," in *IEEE CDC*, 2010.

[ZOS11] ——, "Closed-loop stabilization over Gaussian interference channel," in *18th IFAC World Congress*, August 2011.

[ZOS13] ——, "On stabilization over a gaussian interference channel," in *IEEE ECC*, 2013.

[ZOYS10] A. A. Zaidi, T. J. Oechtering, S. Yüksel, and M. Skoglund, "Closed-loop control over half-duplex AWGN relay channels," in *Reglermöte*, June 2010.

[ZOYS11] ——, "Sufficient conditions for closed-loop control over a Gaussian relay channel," in *IEEE ACC*, June 2011, pp. 2240–2245.

[ZOYS13a] ——, "Stabilization and control over Gaussian networks," in *Information and Control in Networks*, B. Bernhardsson, G. Como, and A. Rantzer, Eds. Springer, 2013.

[ZOYS13b] ——, "Stabilization of linear systems over Gaussian networks," *IEEE Trans. Autom. Control*, Accepted, July 2013.

[ZOYS13c] ——, "Multi-plant stabilization of linear systems over Gaussian intereference channel," *IEEE Trans. Autom. Control*, Submitted 2013.

[ZYOS] A. A. Zaidi, S. Yüksel, T. J. Oechtering, and M. Skoglund, "On the tightness of linear policies for stabilization of linear systems over Gaussian networks," *Systems & Control Letters*, submitted Feb. 2013.

[ZYOS11] ——, "On optimal policies for control and estimation over a Gaussian relay channel," in *IEEE CDC*, 2011.

[ZYOS13] ——, "On optimal policies for control and estimation over a Gaussian relay channel," *Automatica*, vol. 49, no. 9, pp. 2892–2897, Sep. 2013.