Chapter 1
Stabilization and Control over Gaussian Networks

Ali A. Zaidi, Tobias J. Oechtering, Serdar Yüksel and Mikael Skoglund

Abstract We provide an overview and some recent results on real-time communication and control over Gaussian channels. In particular, the problem of remote stabilization of linear systems driven by Gaussian noise over Gaussian relay channels is considered. Necessary and sufficient conditions for mean-square stabilization are presented, which reveal signal-to-noise ratio requirements for stabilization which are tight in certain class of settings. Optimal linear policies are constructed, global optimality and sub-optimality of such policies are investigated in a variety of settings. We also consider the design of low-delay sensing and transmit schemes for real-time communication.

1.1 Introduction

In this chapter we consider 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 group of sensor nodes monitor the plant and communicate their observations (measurements) to a remotely situated control unit over wireless links, that are modeled as additive white Gaussian channels. 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 is required.

Ali A. Zaidi
KTH Royal Institute of Technology, Stockholm, Sweden, e-mail: zaidi@kth.se

Tobias J. Oechtering
KTH Royal Institute of Technology, Stockholm, Sweden, e-mail: oech@kth.se

Serdar Yüksel
Queen's University, Kingston, Canada, e-mail: yuksel@mast.queensu.ca

Mikael Skoglund
KTH Royal Institute of Technology, Stockholm, Sweden, e-mail: skoglund@kth.se

Therefore we focus on delay-free and power efficient sensing and transmit schemes throughout the chapter. Our objective is to provide an overview and some recent results on real-time communication and stabilization over Gaussian channels. In order to grasp the fundamental principles, we consider setups with one or two sensor nodes under some basic topologies, however, useful references to more general setups are provided throughout the chapter.

The main focus of this chapter is on mean-square stabilization of a linear dynamical system over some basic Gaussian network settings. Some real-time sensing and transmission schemes are proposed and stabilizability of the plant under those schemes is studied. The chapter is organized as follows. In Section 1.2 the problem of remote stabilization of a discrete-time LTI plant over Gaussian sensor network is formulated. In Section 1.3 a single sensor setup is considered, i.e., stabilization over a point-to-point Gaussian channel. Section 1.4 and Section 1.6 consider two sensor setups, where one sensor node merely acts as a relay for communicating state information to the remote controller. Section 1.4 focuses on mean-square stabilization of an LTI plant in various relaying topologies. Section 1.5 and 1.4 addresses the problem of real-time transmission of a Gaussian source over a Gaussian relay channel for delay-sensitive and energy limited applications such as closed-control over wireless sensor networks. Finally, in Section 1.7 we discuss distributed sensing schemes for control over Gaussian channels. The chapter ends with an overview of the existing literature on the problem of control over Gaussian channels, highlighting the important relevant contributions.

1.2 Remote Stabilization of a Linear System

Consider the following linear time invariant system:

$$X_{t+1} = AX_t + U_t + W_t, \qquad t \in \mathbb{N}, \tag{1.1}$$

where $X_t := [x_{1,t}, x_{2,t}, \cdots, x_{n,t}]^T$ is an $\mathbb{R}^n$-valued state process with an initial Gaussian distribution, $U_t := [u_{1,t}, u_{2,t}, \cdots, u_{n,t}]^T$ is an $\mathbb{R}^n$-valued control process, $W_t := [w_{1,t}, w_{2,t}, \cdots, w_{n,t}]^T$ is an $\mathbb{R}^n$-valued independent and identically distributed sequence of Gaussian random variables with zero mean and covariance K_W, and A is the system matrix of appropriate dimensions. Let $\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ 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 ($|\lambda_i| \geq 1$ for all i), i.e., all modes are unstable. Otherwise unstable modes can be decoupled from the stable modes by a similarity transformation. If the system in (1.1) is one dimensional then A is scalar and we use the notation $A = \lambda$, where $|\lambda| > 1$. The initial state of the system X_0 is assumed to be a random variable with zero mean and covariance Λ_0.

We consider a remote control setup as shown in Fig. 1.1, where a sensor or a group of sensor nodes observe the state process and communicate their observations directly to a remotely situated controller over a wireless channel. In a sensor

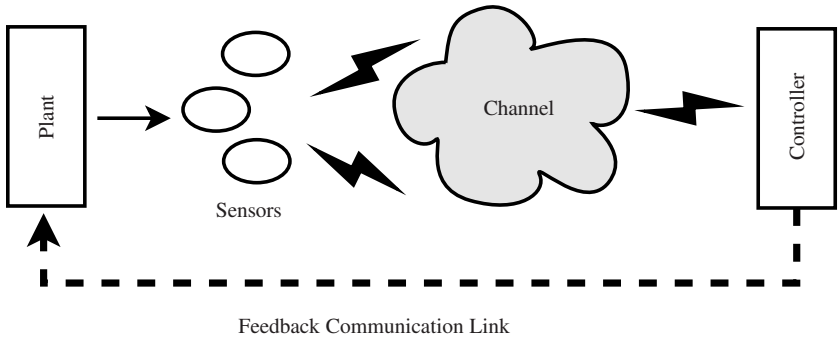


Fig. 1.1 Feedback control over sensor network

network, some nodes can act as relays to support the communication with the remote controller. Upon receiving the signals from the sensors, the remotely located controller aims at stabilizing the system in mean square sense, which is defined as follows.

Definition 1. A system is said to be *mean square stable* if there exists a constant $M < \infty$ such that $\mathbb{E}[\|X_t\|^2] < M$ for all t.

In practice, sensor nodes have limited power to spend. Therefore we assume an average transmit power constraint at each sensor node. The communication links between all agents (sensors and controller) are modeled as independent Gaussian channels. Since control applications are usually quite sensitive to delays, the sensing and transmission schemes are restricted to be delay-free. In order to make the implementation simple we assume that the controller has a separation structure based on the minimum mean-square estimator state estimator. This separation structure is not optimal in general but it makes the design and implementation much simpler, by employing Kalman filter as state estimator.

1.3 Stabilization Over a Point-Point Channel

Consider the scenario shown in Fig. 1.2, where a sensor node $\mathscr{E}$ observes an n-dimensional state process and transmits it to a remote controller $\mathscr{C}$ over an m-dimensional parallel Gaussian channel. We assume that the initial state is zero mean Gaussian distributed. 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 $z_{i,t} \sim \mathscr{N}(0, N_i)$ are zero mean white Gaussian noise components. Let $f_t : \mathbb{R}^{n(t+1)} \to \mathbb{R}^m$ denote the sensing policy such that $S_t = f_t(X_{[0,t]})$, where $X_{[0,t]} := \{X_0, X_1, \ldots, X_t\}$. The sensor is assumed to have an average transmit power constraint $\mathbb{E}[\|S_t\|^2] = \sum_{i=1}^m P_i \le P_S$, where $P_i := \mathbb{E}[(s_{i,t})^2]$. Further, let $\pi_t : \mathbb{R}^{m(t+1)} \to \mathbb{R}^n$ be the

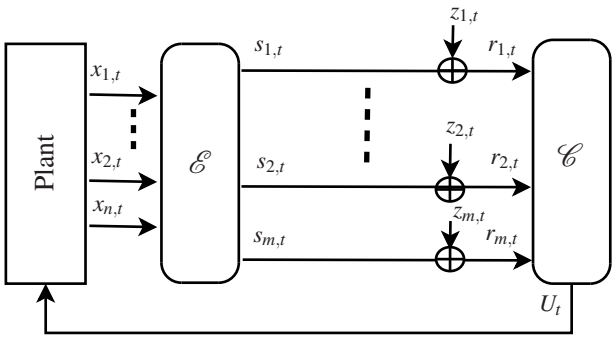


Fig. 1.2 Control over a point-point channel

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 LTI system (1.1) in the mean square sense.

We first present a necessary condition for *mean-square* stabilization over the point-point Gaussian channel depicted in Fig. 1.2.

Theorem 1. *The linear system in (1.1) can be mean square stabilized over the given parallel Gaussian channel only if*

$$\log\left(|A|\right) < \frac{1}{2}\sum_{i=1}^{m}\log\left(1+\frac{P_i}{N_i}\right), \tag{1.2}$$

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

Proof. In order to prove Theorem 1, we make use of the following lemma.

Lemma 1. *([1, Theorem 2.1]) The linear system in* (1.1) *can be mean square stabilized over a channel only if*

$$\log\left(|A|\right) \leq \liminf_{T\to\infty}\frac{1}{T}I\left(\bar{X}_{[0,T-1]} \to R_{[0,T-1]}\right), \tag{1.3}$$

where $\{\bar{X}_t\}$ *is the control free state process given by substituting* $U_t = 0$ *in* (1.1)*,* $R_{[0,T-1]}$ *is the sequence of variables received by the controller over the given channel and* $I\left(\bar{X}_{[0,T-1]} \to R_{[0,T-1]}\right) = \sum_{t=0}^{T-1} I\left(\bar{X}_{[0,t]}; R_t | R_{[0,t\text{-}1]}\right)$ *denotes the directed information.*

Proof. The proof can be found in [1]. This proof essentially follows from the same steps as in Theorem 4.1 of [2], however, with some differences due to the network structure. Similar considerations have appeared in different contexts in [3, 4].

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

$$
\begin{aligned}
I\left(\bar{X}_{[0,T-1]} \to R_{[0,T-1]}\right) &\overset{(a)}{\leq} I\left(\bar{X}_{[0,T-1]}; R_{[0,T-1]}\right) \overset{(b)}{\leq} I\left(S_{[0,T-1]}; R_{[0,T-1]}\right) \\
&\overset{(c)}{\leq} \sum_{t=0}^{T-1} I\left(s_{1,t}, s_{2,t}, \cdots, s_{m,t}; r_{1,t}, r_{2,t}, \cdots, r_{m,t}\right) \\
&= \sum_{t=0}^{T-1} \left[h\left(r_{1,t}, r_{2,t}, \cdots, r_{m,t}\right) - h\left(r_{1,t}, r_{2,t}, \cdots, r_{m,t} | s_{1,t}, s_{2,t}, \cdots, s_{m,t}\right)\right] \\
&= \sum_{t=0}^{T-1} \left[h\left(r_{1,t}, r_{2,t}, \cdots, r_{m,t}\right) - h\left(z_{1,t}, z_{2,t}, \cdots, z_{m,t} | s_{1,t}, s_{2,t}, \cdots, s_{m,t}\right)\right] \\
&= \sum_{t=0}^{T-1} \left[h\left(r_{1,t}, r_{2,t}, \cdots, r_{m,t}\right) - h\left(z_{1,t}, z_{2,t}, \cdots, z_{m,t}\right)\right] \\
&\overset{(d)}{\leq} \sum_{t=0}^{T-1} \left[\sum_{i=1}^{m} h\left(r_{i,t}\right) - \sum_{i=1}^{m} h\left(z_{i,t}\right)\right] \overset{(e)}{\leq} \sum_{t=0}^{T-1} \left[\sum_{i=1}^{m} \log\left(1 + \frac{P_i}{N_i}\right)\right] \\
&= \frac{T}{2} \sum_{i=1}^{m} \log\left(1 + \frac{P_i}{N_i}\right), \qquad (1.4)
\end{aligned}
$$

where (a) follows from [5, Theorem 1]; (b) following data processing inequality with Markov chain $\bar{X}_{[0,T-1]} - S_{[0,T-1]} - R_{[0,T-1]}$; (c) follows from the fact that the channels are memoryless and conditioning reduces entropy; (d) follows from conditioning reduces entropy and mutual independence of the noise sequence $\{z_{1,t}, z_{2,t}, \cdots, z_{m,t}\}$; and (e) follows from the fact that the Gaussian distribution maximizes differential entropy for a fixed variance. Now using (1.4) in Lemma 1 we get the necessary condition given in (1.2). 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.

We now discuss some sensing and control schemes for stabilization over the given point-point Gaussian channel. By employing these schemes, we obtain sufficient conditions for stabilization, which are also presented in the following sections. The schemes for scalar and vector channels are discussed in Sec. 1.3.1 and Sec. 1.3.2 respectively.

1.3.1 Schemes for Scalar Channels

In this section we consider mean square stability of the system in (1.1) over a scalar Gaussian channel, i.e., we assume that $m = 1$ in the system model shown in Fig. 1.2. The state encoder $\mathscr{E}$ 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

$$\mathbb{E}[S_t^2] = E_t \Lambda_t E_t^T \leq P, \tag{1.5}$$

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

$$R_t = S_t + Z_t, \tag{1.6}$$

where Z_t is an i.i.d. Gaussian variable with zero mean and variance N. The information set available to the controller is $I_t^{\mathscr{C}} = \{R_{[0,t]}, U_{[0,t-1]}\}$. The controller applies an action which is linear in the information set, that is $U_t = m_t I_t^{\mathscr{C}}$. 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 1.3.1.1-1.3.1.3.

1.3.1.1 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 $U_t = -A\mathbb{E}[X_t | I_t^{\mathscr{C}}]$, which is optimal for stabilization under the given sensing scheme. Thus the closed loop system is given by

$$\begin{aligned} X_{t+1} &= A\left(X_t - \mathbb{E}[X_t | I_t^{\mathscr{C}}]\right) + W_t \\ &\overset{(a)}{=} A\left(X_t - \Lambda_t E^T \left[E\Lambda_t E^T + \sigma_z^2\right]^{-1} R_t\right) + W_t \\ &\overset{(b)}{=} \left(A_t - \Lambda_t E^T \left[E\Lambda_t E^T + \sigma_z^2\right]^{-1} E\right) X_t + \tilde{Z}_t, \end{aligned} \tag{1.7}$$

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 $\mathbb{E}[X_t | I_t^{\mathscr{C}}] = \mathbb{E}[X_t | R_t] = \mathbb{E}[X_t R_t^T]\mathbb{E}[R_t R_t^T]^{-1} R_t$; 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

$$\Lambda_{t+1} = A\Lambda_t A^T - A\Lambda_t E^T \left[E\Lambda_t E^T + \sigma_z^2\right]^{-1} E\Lambda_t A^T + K_W, \tag{1.8}$$

which is the well-known Riccati equation. In [6], the authors studied such a scheme. According to [6], a noiseless plant can be mean square stabilized by any time-invariant encoding matrix E over a Gaussian channel capacity C as long as the following two conditions are fulfilled: i) $\log\{|A|\} < C$, 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 $A = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)$ with two equal eigenvalues and let $E = \begin{pmatrix} e_1 & e_2 & e_3 \end{pmatrix}$. The observability matrix $\mathscr{O}$ is then given by

$$\mathscr{O} \triangleq \begin{pmatrix} E \\ EA \\ EA^2 \end{pmatrix} = \begin{pmatrix} e_1 & e_2 & e_3 \\ e_1\lambda_1 & e_2\lambda_2 & e_3\lambda_3 \\ e_1\lambda_1^2 & e_3\lambda_2^2 & e_3\lambda_3^2 \end{pmatrix}. \tag{1.9}$$

For the pair (A,E) to observable, the observability matrix $\mathscr{O}$ 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 $\mathscr{O}$ and thus rank of $\mathscr{O}$ can never be made full 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 large power the encoder is allowed to spend. In the following section, we present a linear time varying scheme and show that this scheme can always stabilize the system.

1.3.1.2 Linear Time Variant Scheme

Consider that the linear system (1.1) has to be stabilized over a Gaussian channel having information capacity $C := \frac{1}{2}\log(1+\frac{P}{N})$, which means that the sensor transmits with an average power P and the channel is disturbed by a zero mean Gaussian noise with variance N. In the following we state a sufficient condition for mean-square stability under a linear time varying scheme.

Theorem 2. *The linear system* (1.1) *can be mean-squared stabilized by a linear time-variant scheme over a scalar Gaussian channel of capacity C, if* $\log(|A|) < C$.

Proof. Without loss of generality we assume that the system matrix is in real Jordan. Depending on the nature of eigenvalues, a real Jordan matrix J has the following structure:

$$J = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_p \end{pmatrix}, \quad \text{where} \tag{1.10}$$

$$J_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix} \text{ for } \lambda_i \in \mathbb{R}, \quad J_i = \begin{pmatrix} C_i & I & & \\ & C_i & \ddots & \\ & & \ddots & I \\ & & & C_I \end{pmatrix} \text{ for } \lambda_i = \sigma_i \pm j\omega_i \in \mathbb{C}, \tag{1.11}$$

with $C_i = \begin{pmatrix} \sigma_i & \omega_i \\ -\omega_i & \sigma_i \end{pmatrix}$.

Consider a scheme in which the sensor transmits only one component of the state vector at each time step. Since the system matrix is in Jordan form, it can transmit the state components corresponding to more unstable modes more often. In the following we show that with such a time varying mode-by-mode transmission scheme, the plant can be stabilized if $\log(|A|) < C$. We justify this with the help of two simple examples. In the first example we consider system matrix with repeated real and complex eigenvalues having equal magnitude, where as in the second example we consider eigenvalues with unequal magnitude. These two examples capture the general principle. In the end we outline a general transmit scheme.

Example 1 (Repeated eigenvalues with equal magnitude): Consider an LTI system with the following state equation:

$$\bar{X}_{t+1} = A\bar{X}_t + \bar{U}_t + \bar{W}_t, \tag{1.12}$$

where A is an $n \times n$ matrix with eigenvalues λ_i such that $|\lambda_i| = |\lambda_j|$ for all $1 \le i, j \le n$. Assume that the control actions U_t are taken periodically after every n time steps, i.e., at $t = l(n-1)$ for $l = 1, 2, 3, \ldots$. Under this control scheme, the state at times steps $t = ln$ is given by

$$\bar{X}_{t+n} = A^n\bar{X}_t + \bar{U}_{t+n-1} + \sum_{i=0}^{n-1} A^{n-i-1}\bar{W}_{t+i}, \quad t = ln, l \in \mathbb{N}. \tag{1.13}$$

Let T be a linear transformation such that $T^{-1}A^nT$ is in real Jordan form. It is known that such a transformation always exists [7]. Now apply the linear transformation $X_t = T^{-1}\bar{X}_t$, which gives

$$\begin{aligned} X_{t+n} &= T^{-1}A^nTX_t + T^{-1}\bar{U}_{t+n-1} + T^{-1}\sum_{i=0}^{n-1} A^{n-i-1}\bar{W}_{t+i}, \\ &= \tilde{A}X_t + U_{t+n-1} + V_t, \qquad \text{for} \quad t = ln, l \in \mathbb{N}, \end{aligned} \tag{1.14}$$

where we have defined $\tilde{A} := T^{-1}A^nT$, $U_t := T^{-1}\bar{U}_t$, and $V_t := T^{-1}\sum_{i=0}^{n-1} A^{n-i-1}\bar{W}_{t+i}$. The matrix $\tilde{A}$ is in real Jordan form with eigenvalues $\tilde{\lambda}_i = \lambda_i^n$, where λ_i are the eigenvalues of A for $i = 1, 2, \cdots, n$. Now consider the following sensing scheme for stabilization. The sensor periodically observes the state vector $X_t \in \mathbb{R}^n$ after every $n-$ time steps, i.e., at $t, t+n, t+2n, \ldots$. The sensor linearly amplifies each component of the state vector under an average transmit power constraint P and sequentially

transmits n state components over the Gaussian channel. The state vector is thus transmitted to the controller by using the Gaussian channel n times. The controller computes the MMSE estimate of the state vector $\hat{X}_t$ based on the received signals and periodically takes actions after every n time steps, i.e., $U_{t+n-1} = -\tilde{A}\hat{X}_t$. Under the above scheme, we can write (1.14) as

$$X_{t+n} = \tilde{A}\left(X_t - \hat{X}_t\right) + V_t, \quad t = ln, l \in \mathbb{N}, \tag{1.15}$$

In the following we consider an example with $n = 6$ and show that above scheme is sufficient for stabilization.

Consider a six-dimensional plant ($n = 6$) with state vector $X_t \in \mathbb{R}^6$, that is a plant with six poles (eigenvalues). Assume that the matrix $\tilde{A}$ has a real eigenvalue and a complex conjugate pair, each with algebraic multiplicity two. That is we have $\tilde{\lambda}_1 = \tilde{\lambda}_3 = \tilde{\sigma} + j\tilde{\omega}$, $\tilde{\lambda}_2 = \tilde{\lambda}_4 = \tilde{\sigma} - j\tilde{\omega}$, and $\tilde{\lambda}_5 = \tilde{\lambda}_6 = \tilde{\lambda}$. Since $\tilde{A}$ is in real Jordan form, we have

$$\tilde{A} = \begin{pmatrix} \tilde{\sigma} & \tilde{\omega} & 1 & 0 & 0 & 0 \\ -\tilde{\omega} & \tilde{\sigma} & 0 & 1 & 0 & 0 \\ 0 & 0 & \tilde{\sigma} & \tilde{\omega} & 0 & 0 \\ 0 & 0 & -\tilde{\omega} & \tilde{\sigma} & 0 & 0 \\ 0 & 0 & 0 & 0 & \tilde{\lambda} & 1 \\ 0 & 0 & 0 & 0 & 0 & \tilde{\lambda} \end{pmatrix}. \tag{1.16}$$

By substituting $\tilde{A}$ from (1.16) in (1.15), each component of the state vector is given by

$$\begin{aligned}
x_{1,t+n} &= \tilde{\sigma}(x_{1,t} - \hat{x}_{1,t}) + \tilde{\omega}(x_{2,t} - \hat{x}_{2,t}) + (x_{3,t} - \hat{x}_{3,t}) + v_{1,t}, \\
x_{2,t+n} &= -\tilde{\omega}(x_{1,t} - \hat{x}_{1,t}) + \tilde{\sigma}(x_{2,t} - \hat{x}_{2,t}) + (x_{4,t} - \hat{x}_{4,t}) + v_{2,t}, \\
x_{3,t+n} &= \tilde{\sigma}(x_{3,t} - \hat{x}_{3,t}) + \tilde{\omega}(x_{4,t} - \hat{x}_{4,t}) + v_{3,t}, \\
x_{4,t+n} &= -\tilde{\omega}(x_{3,t} - \hat{x}_{3,t}) + \tilde{\sigma}(x_{4,t} - \hat{x}_{4,t}) + v_{4,t}, \\
x_{5,t+n} &= \tilde{\lambda}(x_{5,t} - \hat{x}_{5,t}) + (x_{6,t} - \hat{x}_{6,t}) + v_{5,t}, \\
x_{6,t+n} &= \tilde{\lambda}(x_{6,t} - \hat{x}_{6,t}) + v_{6,t},
\end{aligned} \tag{1.17}$$

We now find conditions for all modes to be stable. We start with the lowest mode. The second moment of $x_{6,t}$ is given by

$$\begin{aligned}
\mathbb{E}\left[x_{6,t+n}^2\right] &= \tilde{\lambda}^2 \mathbb{E}[(x_{6,t} - \hat{x}_{6,t})] + \mathbb{E}\left[v_{6,t}^2\right] \\
&\overset{(a)}{=} \tilde{\lambda}^2 2^{-2C} \mathbb{E}\left[x_{6,t}^2\right] + n_{v,6},
\end{aligned} \tag{1.18}$$

where (a) follows from the linear mean-square estimation of a Gaussian variable over a scalar Gaussian channel of capacity C and $n_{v,k} := \mathbb{E}\left[v_{k,t}^2\right]$ for $k = 1, 2, \cdots, 6$. We observe that $\mathbb{E}\left[x_{6,t}^2\right]$ is bounded if $\tilde{\lambda}^2 2^{-2C} < 1$. Since $|\tilde{\lambda}_6| = \tilde{\lambda}$, the state component $x_{6,t}$ is stable if

$$|\tilde{\lambda}_6|^2 2^{-2C} < 1 \Rightarrow \log\left(|\tilde{\lambda}_6|\right) < C. \tag{1.19}$$

Now consider $x_{5,t}$, whose second moment can be bounded as

$$\begin{aligned}
\mathbb{E}\left[x_{5,t+n}^2\right] &\overset{(a)}{=} \tilde{\lambda}^2 \mathbb{E}\left[\left(x_{5,t}-\hat{x}_{5,t}\right)^2\right] + 2\tilde{\lambda}\mathbb{E}\left[\left(x_{5,t}-\hat{x}_{5,t}\right)\left(x_{6,t}-\hat{x}_{6,t}\right)\right] \\
&\quad + \mathbb{E}\left[\left(x_{6,t}-\hat{x}_{6,t}\right)^2\right] + n_{v,5} \\
&\overset{(b)}{=} \tilde{\lambda}^2 2^{-2C}\mathbb{E}\left[x_{5,t}^2\right] + 2\tilde{\lambda}\mathbb{E}\left[\left(x_{5,t}-\hat{x}_{5,t}\right)\left(x_{6,t}-\hat{x}_{6,t}\right)\right] + 2^{-2C}\mathbb{E}\left[x_{6,t}^2\right] + n_{v,5} \\
&\overset{(c)}{\leq} \tilde{\lambda}^2 2^{-2C}\mathbb{E}\left[x_{6,t}^2\right] + 2\tilde{\lambda}\sqrt{\mathbb{E}\left[\left(x_{5,t}-\hat{x}_{5,t}\right)^2\right]\mathbb{E}\left[\left(x_{6,t}-\hat{x}_{6,t}\right)^2\right]} \\
&\quad + 2^{-2C}\mathbb{E}\left[x_{6,t}^2\right] + n_{v,5} \\
&= \tilde{\lambda}^2 2^{-2C}\mathbb{E}\left[x_{5,t}^2\right] + 2\tilde{\lambda}\sqrt{2^{-2C}\mathbb{E}\left[x_{5,t}^2\right]}\sqrt{2^{-2C}\mathbb{E}\left[x_{6,t}^2\right]} \\
&\quad + 2^{-2C}\mathbb{E}\left[x_{6,t}^2\right] + n_{v,5} \\
&\overset{(d)}{\leq} k_1\mathbb{E}\left[x_{5,t}^2\right] + k_2\sqrt{\mathbb{E}\left[x_{5,t}^2\right]} + k_3,
\end{aligned} \tag{1.20}$$

where (a) follows from (1.17); (b) follows from the linear mean-square estimation of a Gaussian variable over a scalar Gaussian channel of capacity C; (c) follows Cauchy-Schwarz inequality; (d) follows from the fact $\mathbb{E}\left[x_{6,t}^2\right] < M$ (assuming that (1.19) is satisfied) and by defining $k_1 := \tilde{\lambda}^2 2^{-2C}$, $k_2 := 2\tilde{\lambda}2^{-2C}\sqrt{M}$, and $k_3 := 2^{-2C}M + n_{v,5}$. We now want to a find condition which ensures convergence of the following sequence:

$$\alpha_{t+1} = k_1\alpha_t + k_2\sqrt{\alpha_t} + k_3. \tag{1.21}$$

In order to show convergence, we make use of the following lemma.

Lemma 2. *([1, Lemma 6.1]) 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 can be found in [1].

We observe that the mapping $T(\alpha) = k_1\alpha + k_2\sqrt{\alpha} + k_3$ with $\alpha \geq 0$ is monotonically increasing since $k_1, k_2 > 0$. It will have a unique fixed point $\alpha^\star$ if and only if $k_1 < 1$, since $k_2, k_3 > 0$. Assuming that $k_1 < 1$, there exists $u < \alpha^\star < v$ such that $T(u) \geq u$ and $T(v) \leq v$. Therefore by Lemma 2 the sequence $\{\alpha_t\}$ is convergent if $k_1 = \tilde{\lambda}^2 2^{-2C} < 1 \Rightarrow \log(\tilde{\lambda}) < C$. Since $|\tilde{\lambda}_5| = \tilde{\lambda}$, the state $x_{5,t}$ is stable if

$$\log\left(|\tilde{\lambda}_5|\right) < C. \tag{1.22}$$

The second moments of $x_{3,t}$ and $x_{4,t}$ are given by

$$\begin{aligned}\mathbb{E}\left[x_{3,t+n}^2\right] =& \tilde{\sigma}^2 2^{-2C}\mathbb{E}\left[x_{3,t}^2\right] + \tilde{\omega}^2 2^{-2C}\mathbb{E}\left[x_{4,t}^2\right] \\ &+ 2\tilde{\sigma}\tilde{\omega}\mathbb{E}\left[\left(x_{3,t}-\hat{x}_{3,t}\right)\left(x_{4,t}-\hat{x}_{4,t}\right)\right] + n_{v,3} \\ \mathbb{E}\left[x_{4,t+n}^2\right] =& \tilde{\omega}^2 2^{-2C}\mathbb{E}\left[x_{3,t}^2\right] + \tilde{\sigma}^2 2^{-2C}\mathbb{E}\left[x_{4,t}^2\right] \\ &- 2\tilde{\sigma}\tilde{\omega}\mathbb{E}\left[\left(x_{3,t}-\hat{x}_{3,t}\right)\left(x_{4,t}-\hat{x}_{4,t}\right)\right] + n_{v,4}. \end{aligned} \tag{1.23}$$

By using the above equations, we can write

$$\mathbb{E}\left[x_{3,t+n}^2\right] + \mathbb{E}\left[x_{4,t+n}^2\right] = \left(\tilde{\sigma}^2+\tilde{\omega}^2\right)2^{-2C}\left(\mathbb{E}\left[x_{3,t}^2\right] + \mathbb{E}\left[x_{4,t}^2\right]\right) + n_{v,3} + n_{v,4}. \tag{1.24}$$

We observe that the sum $\mathbb{E}\left[x_{3,t+n}^2\right] + \mathbb{E}\left[x_{4,t+n}^2\right]$ is bounded if $\left(\tilde{\sigma}^2+\tilde{\omega}^2\right)2^{-2C} < 1$. Since $|\tilde{\lambda}_3|^2 = |\tilde{\lambda}_4|^2 = \left(\tilde{\sigma}^2+\tilde{\omega}^2\right)^2$, the state components $x_{3,t}$ and $x_{4,t}$ are stable if

$$\log\left(|\tilde{\lambda}_3|\right) < C, \quad \log\left(|\tilde{\lambda}_4|\right) < C. \tag{1.25}$$

Finally, the second moments of $x_{1,t}$ and $x_{2,t}$ are given by

$$\begin{aligned}\mathbb{E}\left[x_{1,t+n}^2\right] =& \tilde{\sigma}^2 2^{-2C}\mathbb{E}\left[x_{1,t}^2\right] + \tilde{\omega}^2 2^{-2C}\mathbb{E}\left[x_{2,t}^2\right] + 2^{-2C}\mathbb{E}\left[x_{3,t}^2\right] \\ &+ 2\tilde{\sigma}\tilde{\omega}\mathbb{E}\left[\left(x_{1,t}-\hat{x}_{1,t}\right)\left(x_{2,t}-\hat{x}_{2,t}\right)\right] + 2\tilde{\sigma}\mathbb{E}\left[\left(x_{1,t}-\hat{x}_{1,t}\right)\left(x_{3,t}-\hat{x}_{3,t}\right)\right] \\ &+ 2\tilde{\omega}\mathbb{E}\left[\left(x_{2,t}-\hat{x}_{2,t}\right)\left(x_{3,t}-\hat{x}_{3,t}\right)\right] + n_{v,1} \\ \mathbb{E}\left[x_{1,t+n}^2\right] =& \tilde{\omega}^2 2^{-2C}\mathbb{E}\left[x_{1,t}^2\right] + \tilde{\sigma}^2 2^{-2C}\mathbb{E}\left[x_{2,t}^2\right] + 2^{-2C}\mathbb{E}\left[x_{4,t}^2\right] \\ &- 2\tilde{\sigma}\tilde{\omega}\mathbb{E}\left[\left(x_{1,t}-\hat{x}_{1,t}\right)\left(x_{2,t}-\hat{x}_{2,t}\right)\right] - 2\tilde{\omega}\mathbb{E}\left[\left(x_{1,t}-\hat{x}_{1,t}\right)\left(x_{4,t}-\hat{x}_{4,t}\right)\right] \\ &+ 2\tilde{\sigma}\mathbb{E}\left[\left(x_{2,t}-\hat{x}_{2,t}\right)\left(x_{4,t}-\hat{x}_{4,t}\right)\right] + n_{v,2} \end{aligned} \tag{1.26}$$

By using the above equations, we can write

$$\begin{aligned}\mathbb{E}\left[x_{1,t+n}^2\right] + \mathbb{E}\left[x_{2,t+n}^2\right] =& \left(\tilde{\sigma}^2+\tilde{\omega}^2\right)2^{-2C}\left(\mathbb{E}\left[x_{1,t}^2\right] + \mathbb{E}\left[x_{2,t}^2\right]\right) \\ &+ 2\tilde{\sigma}\left(\mathbb{E}\left[\left(x_{1,t}-\hat{x}_{1,t}\right)\left(x_{3,t}-\hat{x}_{3,t}\right)\right] + \mathbb{E}\left[\left(x_{2,t}-\hat{x}_{2,t}\right)\left(x_{4,t}-\hat{x}_{4,t}\right)\right]\right) \\ &+ 2\tilde{\omega}\left(\mathbb{E}\left[\left(x_{2,t}-\hat{x}_{2,t}\right)\left(x_{3,t}-\hat{x}_{3,t}\right)\right] - \mathbb{E}\left[\left(x_{1,t}-\hat{x}_{1,t}\right)\left(x_{4,t}-\hat{x}_{4,t}\right)\right]\right) \\ &+ n_{v,1} + n_{v,2} \end{aligned} \tag{1.27}$$

We can now bound $\mathbb{E}\left[x_{1,t+n}^2\right] + \mathbb{E}\left[x_{2,t+n}^2\right]$ as

$$\begin{aligned}\mathbb{E}\left[x_{1,t+n}^2\right]+\mathbb{E}\left[x_{2,t+n}^2\right] &\overset{(a)}{\leq} \left(\tilde{\sigma}^2+\tilde{\omega}^2\right)2^{-2C}\left(\mathbb{E}\left[x_{1,t}^2\right]+\mathbb{E}\left[x_{2,t}^2\right]\right)\\ &\quad+4\tilde{\sigma}^2 2^{-2C}\left(\sqrt{\mathbb{E}\left[x_{1,t}^2\right]\mathbb{E}\left[x_{3,t}^2\right]}+\sqrt{\mathbb{E}\left[x_{2,t}^2\right]\mathbb{E}\left[x_{4,t}^2\right]}\right)\\ &\quad+4\tilde{\omega}^2 2^{-2C}\left(\sqrt{\mathbb{E}\left[x_{2,t}^2\right]\mathbb{E}\left[x_{3,t}^2\right]}+\sqrt{\mathbb{E}\left[x_{1,t}^2\right]\mathbb{E}\left[x_{4,t}^2\right]}\right)\\ &\quad+n_{v,1}+n_{v,2}\\ &\overset{(b)}{\leq} k_1\left(\mathbb{E}\left[x_{1,t}^2\right]+\mathbb{E}\left[x_{2,t}^2\right]\right)+k_2\sqrt{\mathbb{E}\left[x_{1,t}^2\right]+\mathbb{E}\left[x_{2,t}^2\right]}+k_3,\end{aligned} \tag{1.28}$$

where (a) follows from Cauchy-Schwarz inequality; and (b) follows from $\mathbb{E}\left[x_{3,t}^2\right] < M$ and $\mathbb{E}\left[x_{4,t}^2\right] < M$ (assuming that the condition in (1.25) is satisfied) and by defining $k_1 := \left(\tilde{\sigma}^2+\tilde{\omega}^2\right)2^{-2C}$, $k_2 := 16\left(\tilde{\sigma}^2+\tilde{\omega}^2\right)^2 2^{-2C}M$, and $k_3 := n_{v,1}+n_{v,2}$. By using Lemma 2 we can show that $x_{1,t}$ and $x_{2,t}$ are stable if $k_1 = \left(\tilde{\sigma}^2+\tilde{\omega}^2\right)^2 2^{-2C} < 1$. Since $|\tilde{\lambda}_1|^2 = |\tilde{\lambda}_2|^2 = \left(\tilde{\sigma}^2+\tilde{\omega}^2\right)^2$, we get

$$\log\left(|\tilde{\lambda}_1|\right) < C, \quad \log\left(|\tilde{\lambda}_2|\right) < C, \tag{1.29}$$

It follows from (1.29), (1.29), and (1.29) that the system is stable if

$$\sum_{i=1}^{n}\log(|\tilde{\lambda}_i|) = \log\left(|\tilde{A}|\right) < nC. \tag{1.30}$$

Since $|\tilde{A}| = |T^1 A^n T| = |A|^n$, we have

$$\log\left(|A|\right) < C. \tag{1.31}$$

Having shown sufficiency of the linear time variant scheme for the system matrix having equal magnitude eigenvalues with algebraic multiplicity, we next consider an example of a system matrix having eigenvalues with unequal magnitude.

Example 2 (Eigenvalues with unequal magnitude): Consider a system matrix A with three eigenvalues, $\lambda_1 \in \mathbb{R}$ and $\lambda_2, \lambda_3 \in \mathbb{C}$ with $|\lambda_1| = |\lambda_2|^2 = |\lambda_3|^2$. For this system, consider the following scheme. The transmission from the sensor to the controller happens periodically, where each transmission period consists for four time slots. In the first two time slots, the state corresponding to λ_1 is transmitted and in the last two slots the states corresponding to λ_2 and λ_3 are transmitted. Note that the sensor is serving more unstable modes more often. In the following we show again that under this transmit scheme, the system is mean square stable if $\log(|A|) < C$.

Let us assume that the transmission period starts at t and ends at $t+4$. At time t, the sensor transmits $x_{1,t}$ and the controller takes action $U_t = [-\lambda_1 \hat{x}_{1,t}, 0, 0]$. At time $t+1$, the sensor transmits $x_{1,t+1}$ and the controller takes action $U_t = [-\lambda_1 \hat{x}_{1,t+1}, 0, 0]$. At time $t+2$, the sensor transmits $x_{2,t}$ and the does not take any action. Like the previous example, we are using such scheme to make the analysis simpler, although it is better to transmit the most recent state and apply control action as early as possible. At time $t+3$, the sensor transmits $x_{3,t}$ and the controller takes the following action: $U_t = [0, -\lambda_1 \hat{x}_{2,t}, -\lambda_1 \hat{x}_{2,t}]$. Under this transmit and control scheme, the second moments of $x_{1,t}$ are given by,

$$\begin{aligned}
\mathbb{E}\left[x_{1,t+1}^2\right] &= \lambda_1^2 2^{-2C} \mathbb{E}\left[x_{1,t}^2\right] + n_1, \quad \mathbb{E}\left[x_{1,t+2}^2\right] = \lambda_1^2 2^{-2C} \mathbb{E}\left[x_{1,t+1}^2\right] + n_1, \\
\mathbb{E}\left[x_{1,t+3}^2\right] &= \lambda_1^2 \mathbb{E}\left[x_{1,t+2}^2\right] + n_1, \quad \mathbb{E}\left[x_{1,t+4}^2\right] = \lambda_1^2 \mathbb{E}\left[x_{1,t+4}^2\right] + n_1,
\end{aligned} \tag{1.32}$$

where n_1 is variance of process noise. Using the above equations, the second moment of the state $x_{1,t}$ at the start of each transmission period is given by,

$$\mathbb{E}\left[x_{1,t+4}^2\right] = \lambda_1^8 2^{-4C} \mathbb{E}\left[x_{1,t}^2\right] + \tilde{n}_1, \quad t = 4l, l \in N, \tag{1.33}$$

where $\tilde{n}_1 = n_1(1 + \lambda_1^2 2^{-2C} + \lambda_1^2 2^{-6C} + \lambda_1^2 2^{-8C})$. Similarly using the approach that was used in the previous example, we can show that

$$\mathbb{E}\left[x_{2,t+4}^2\right] + \mathbb{E}\left[x_{2,t+4}^2\right] = |\lambda_2|^8 2^{-2C} \left(\mathbb{E}\left[x_{2,t+4}^2\right] + \mathbb{E}\left[x_{2,t+4}^2\right]\right) + \tilde{n}_2, \quad t = 4l, l \in N, \tag{1.34}$$

where $\tilde{n}_2$ is the term due to the process noise. From (1.33) and (1.34) we observe that all modes will be stable if

$$\begin{aligned}
&\lambda_1^8 2^{-4C} < 1, \quad \lambda_2^8 2^{-2C} < 1, \quad \lambda_3^8 2^{-2C} < 1 \\
&\Rightarrow \log(|\lambda_1|) < \frac{1}{2}C, \quad \log(|\lambda_2|) = \log(|\lambda_3|) < \frac{1}{4}C \\
&\Rightarrow \sum_{i=1}^{n} \log(|\lambda_i|) = \log(|A|) < C.
\end{aligned} \tag{1.35}$$

For a general n-dimensional system the transmit scheme can be generalized as follows: Choose k, k_m, such that $\frac{k_m}{k} = \frac{\log(|\lambda_m|)}{\sum_{i=1}^{n} \log(|\lambda_i|)}$ for $m = 1, 2, \cdots, n$. The sensor transmits periodically with a period equal to k time slots, in which the state x_m corresponding to λ_m is transmitted k_m times. Note that $\sum_{m=1}^{n} k_m = k$. The system will be stable if $\log(|\lambda_m|) < \frac{\log(|\lambda_m|)}{\sum_{i=1}^{n} \log(|\lambda_i|)} C$ for all $m \in \{1, 2, \cdots, n\}$, which is equivalent to $\sum_{i=1}^{n} \log(|\lambda_i|) = \log(|A|) < C$.

Remark 1. Although we have proved Theorem 2 for Gaussian distributed initial states, it is also valid for other distributions with finite variance. For any non-Gaussian distributed initial state with finite variance, we can use the approach in [8, Section IV] to make the state process Gaussian distributed and then the schemes discussed earlier in this section can be applied.

According to Theorem 2, there is no loss in mean-square stabilizability by restricting the sensing and control scheme to be linear. This makes linear policies a good choice for stabilization over scalar Gaussian channels. Notice, while deriving an achievable stability region, the objective was to keep the second moment of the state process bounded and we did not aim at minimizing the second moment. One might be interested in minimizing the second moment of the state process over a finite or an infinite time horizon. In the following we consider a finite horizon stabilization problem and present an optimal linear scheme.

1.3.1.3 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 3. *Let* $\tilde{G}:=\left[\sqrt{P},0,0,\ldots,0\right]$, $K_t=A^T\left(I+K_{t+1}\right)A\left(I-\tilde{G}^T\tilde{G}\left(\frac{1}{N+P}\right)\right)$ *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 is given by:* $E_t^\star=\tilde{G}\pi_t\Lambda_t^{-\frac{1}{2}}$.

Proof. We rewrite the Riccati equation (1.8) as

$$\begin{aligned}\Lambda_{t+1}=&A\Lambda_t^{\frac{1}{2}}\left(I-\frac{\Lambda_t^{\frac{1}{2}}E_t^T}{\sqrt{N}}\left[\frac{E_t\Lambda_tE_t^T}{\sqrt{N}}+1\right]^{-1}\frac{E_t\Lambda_t^{\frac{1}{2}}}{\sqrt{N}}\right)\Lambda_t^{\frac{1}{2}}A^T+K_W\\ \overset{(a)}{=}&A\Lambda_t^{\frac{1}{2}}\left(I-C_t^T\left[C_tC_t^T+1\right]^{-1}C_t\right)\Lambda_t^{\frac{1}{2}}A^T+K_W\\ \overset{(b)}{=}&A\Lambda_t^{\frac{1}{2}}\left[I+C_t^TC_t\right]^{-1}\Lambda_t^{\frac{1}{2}}A^T+K_W,\end{aligned}\tag{1.36}$$

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:

$$\{C_i^\star\}_{i=0}^{t_f-1}=\arg\min_{\{C_i\}_{i=0}^{t_f-1}:C_tC_t^T\leq\frac{P}{\sigma^2}}\sum_{t=0}^{t_f-1}\mathrm{Tr}[\Lambda_{t+1}],\tag{1.37}$$

subject to,

$$\Lambda_{t+1} = A\Lambda_t^{\frac{1}{2}} \left[I + C_t^T C_t\right]^{-1} \Lambda_t^{\frac{1}{2}} A^T + K_W. \tag{1.38}$$

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 [9]. 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

$$\begin{aligned}
\mathrm{Tr}[K_t\Lambda_t + L_t] &= \min_{C_t : C_t C_t^T \le \frac{P}{N}} \{\mathrm{Tr}[\Lambda_{t+1}] + \mathrm{Tr}[K_{t+1}\Lambda_{t+1} + L_{t+1}]\} \\
&= \min_{C_t : C_t C_t^T \le \frac{P}{N}} \mathrm{Tr}[(I + K_{t+1})\Lambda_{t+1} + L_{t+1}] \\
&\overset{(a)}{=} \min_{C_t : C_t C_t^T \le \frac{P}{N}} \mathrm{Tr}\left[(I+K_{t+1}) \times \left(A\Lambda_t^{\frac{1}{2}} \left[I + C_t^T C_t\right]^{-1} \Lambda_t^{\frac{1}{2}} A^T + K_W\right) + L_{t+1}\right] \\
&\overset{(b)}{=} \mathrm{Tr}[(I+K_{t+1})K_W + L_{t+1}] + \min_{C_t : C_t C_t^T \le \frac{P}{N}} \mathrm{Tr}\left[(I+K_{t+1})A\Lambda_t^{\frac{1}{2}} \left[I + C_t^T C_t\right]^{-1} \Lambda_t^{\frac{1}{2}} A^T\right] \\
&= \mathrm{Tr}[(I+K_{t+1})K_W + L_{t+1}] + \min_{C_t : C_t C_t^T \le \frac{P}{N}} \mathrm{Tr}\left[\Lambda_t^{\frac{1}{2}} A^T (I+K_{t+1}) A\Lambda_t^{\frac{1}{2}} \left[I + C_t^T C_t\right]^{-1}\right] \\
&\overset{(c)}{=} \mathrm{Tr}[(I+K_{t+1})K_W + L_{t+1}] + \mathrm{Tr}\left[\Lambda_t^{\frac{1}{2}} A^T (I+K_{t+1}) A\Lambda_t^{\frac{1}{2}} \left[I + \pi_t G^T G \pi_t^T\right]^{-1}\right] \\
&\overset{(d)}{=} \mathrm{Tr}[(I+K_{t+1})K_W + L_{t+1}] + \\
&\qquad \mathrm{Tr}\left[\Lambda_t^{\frac{1}{2}} A^T (I+K_{t+1}) A\Lambda_t^{\frac{1}{2}} \left(I - \pi_t G^T \left(1 + G\pi_t^T \pi_t G^T\right)^{-1} G\pi_t^T\right)\right] \\
&\overset{(e)}{=} \mathrm{Tr}[(I+K_{t+1})K_W + L_{t+1}] + \\
&\qquad \mathrm{Tr}\left[A^T (I+K_{t+1}) A\left(\Lambda_t - \Lambda_t^{\frac{1}{2}} \pi_t G^T G \pi_t^T \Lambda_t^{\frac{1}{2}} \left(1 + \frac{P}{N}\right)^{-1}\right)\right] \\
&\overset{(f)}{=} \mathrm{Tr}[(I+K_{t+1})K_W + L_{t+1}] + \mathrm{Tr}\left[A^T (I+K_{t+1}) A\left(I - G^T G \left(\frac{N}{N+P}\right)\right)\Lambda_t\right],
\end{aligned} \tag{1.39}$$

where (a) follows by substituting Λ_{t+1} using equation (1.38); (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 [10] the unique solution to the trace minimization problem unique minimizer is 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 $\left(\Lambda_t^{\frac{1}{2}} A^T (I+K_{t+1}) A\Lambda_t^{\frac{1}{2}}\right)$ such that $\pi_t^T \left(\Lambda_t^{\frac{1}{2}} A^T (I+K_{t+1}) A\Lambda_t^{\frac{1}{2}}\right) \pi_t = \mathrm{diag}(\upsilon_{1,t}, \cdots, \upsilon_{N,t})$ 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 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

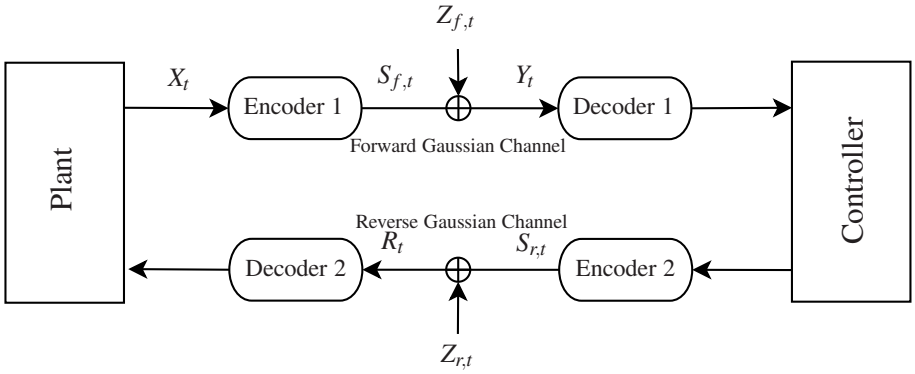


Fig. 1.3 Stabilization over noisy forward and reverse Gaussian channels.

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 (1.39), we choose $K_{t_f} = L_{t_f} = 0$, and $\{K_{t+1}, L_{t+1}\}$ according to

$$K_t = A^T (I + K_{t+1}) A \left(I - G^T G \left(\frac{N}{N+P} \right) \right),$$
$$L_t = (I + K_{t+1}) K_W + L_{t+1}. \tag{1.40}$$

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 $\tilde{G} := \sqrt{N} G = \left[\sqrt{P}, 0, 0, \cdots, 0\right]$.

1.3.1.4 Noisy Feedback Link

So far we have considered the communication link from the controller to the plant to be noiseless. Noiseless communication link from the controller to the plant can be a good assumption for certain scenarios in which the controller is either connected to the plant via a cable or the controller has very large power to spend for transmission of signals over the air. However in some practical situations, it may not be reasonable to model the communication link from the controller to the plant as noiseless. In these situations, the remote controller can be equipped with an encoder to encode the control actions before transmitting them over a noisy channel and the remotely located actuator can be equipped with a decoder, to decode the control actions using the signal received over the noisy channel. In Fig. 1.3 such a setup is shown, where there are two encoders to encode the plant's state and the control actions with average transmit powers $\mathbb{E}\left[\left(S_{f,t}\right)^2\right] = P_f$ and $\mathbb{E}\left[\left(S_{r,t}\right)^2\right] = P_r$ respectively. The forward and the reverse channels are disturbed by white Gaussian noises $Z_{f,t} \sim \mathcal{N}(0, N_f)$

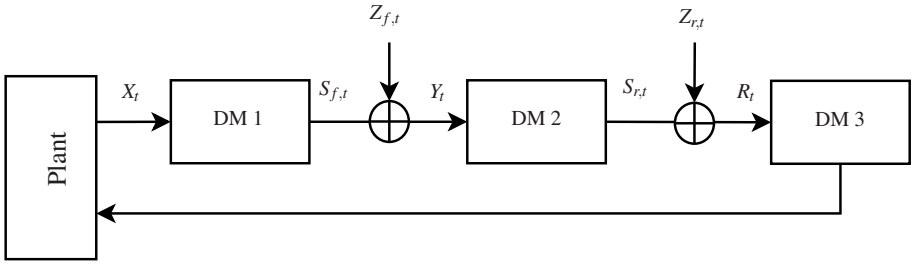


Fig. 1.4 Another representation of the system model in Fig. 1.3.

and $Z_{r,t} \sim \mathcal{N}(0,N_r)$ respectively. Thus the capacities of the forward channel and the reverse channel are given by $C_f := \frac{1}{2}\log\left(1+\frac{P_S}{N_f}\right)$ and $C_r := \frac{1}{2}\log\left(1+\frac{P_r}{N_r}\right)$. In the following present necessary and sufficient conditions for mean square stability over the given channel.

Theorem 4. *The linear system* (1.1) *can be mean square stabilized over a noisy forward channel with capacity C_f and a noisy reverse channel with capacity C_r only if*

$$\log(|A|) \leq \min\left\{C_f, C_r\right\}. \tag{1.41}$$

Proof. The system in Fig. 1.3 can be viewed as the system depicted in Fig. 1.4, where the encoders, decoders, and controllers are viewed as decision makers. The encoder in the forward link (Encoder 1) is the first decision maker DM1. The decoder in the forward link (Decoder 1), the controller, and the encoder in the reverse link (Encoder 2) are altogether can be viewed as second decision maker DM2. Finally the DM3 represents decoder of the reverse link (Decoder 2). For the system shown in Fig. 1.4 we know from Lemma 1 that the following condition is necessary for stabilization:

$$\log\left(|A|\right) \leq \liminf_{T\to\infty} \frac{1}{T} I\left(\bar{X}_{[0,T-1]} \to R_{[0,T-1]}\right). \tag{1.42}$$

According to the proof of Theorem 3.1 in [1], the directed information can be bounded as

$$\begin{aligned} I\left(\bar{X}_{[0,T-1]} \to R_{[0,T-1]}\right) &\leq \min\left\{\sum_{t=0}^{T-1} I(S_{f,t};Y_t), \sum_{t=0}^{T-1} I(S_{r,t};R_t)\right\} \\ &\leq \frac{1}{2}\min\left\{\log\left(1+\frac{P_f}{N_f}\right), \log\left(1+\frac{P_r}{N_r}\right)\right\} \\ &= \min\left\{C_f, C_r\right\}, \end{aligned} \tag{1.43}$$

where the last equality follows from the definition of channel capacities. By using (1.43) in (1.42), we get (1.41).

Theorem 4 shows that the reliability of the reverse channel is as important as the forward channel. The necessary condition (1.41) was first obtained for memoryless sensors and controller in [11, Theorem 8.1]. In [11] the authors have also obtained a sufficient condition for stabilization by restricting the encoders and the decoders to be linear and memoryless, which is stated in the following theorem.

Theorem 5. *([11, Theorem 8.1]) The linear system* (1.1) *can be mean square stabilized over a Gaussian forward channel with capacity C_f and a Gaussian reverse channel with capacity C_r using a linear memoryless sensing and control scheme if*

$$\log(|A|) \leq \frac{1}{2}\log\left(\frac{1}{2^{-2C_f}+2^{-2C_r}-2^{-2(C_f+C_r)}}\right). \tag{1.44}$$

In Fig. 1.5 we have fixed forward channel capacity $C_f = 5$ bits/channel use, and have plotted stability region achievable with linear memoryless scheme as a function of reverse channel capacity C_r using (1.44). The figure shows that an LTI plant with system matrix A is stabilizable with linear scheme if $\log(|A|)$ is below the stability curve drawn in the figure. For the sake of comparison, the outer bound on the achievable stability region is also plotted according to (1.41). One can observe that the linear memoryless scheme is quite efficient because its performance gets close to optimal as one of the two links becomes relatively more reliable. In Sec. 1.5 we discuss sub-optimality of linear policies for estimation over multi-hop relay networks. It is shown that a simple three-level quantizer policies can outperform the best linear policy even over a two-hop network. The setup of noisy forward and noisy reverse channel can be viewed as a two-hop network, as illustrated in Fig. 1.4. Use of such non-linear schemes in forward and reverse channels can potentially improve stability of the closed-loop system, however this is yet to explored. Control over noisy forward and reverse channels have been considered also for more general channels in [11].

In the remainder of the chapter we keep the assumption of noiseless link from the controller to the plant, in order to simplify the analysis and to avoid tedious computations. As long as the encoders and decoders are linear and the channels are modeled as Gaussian, the nature of the problem does not change and one can obtain stability results for noisy reverse channels with some straightforward analysis.

1.3.2 Schemes for Vector Channel

In the previous section we showed that linear schemes can achieve the minimum signal-to-noise required for stabilization of multi-dimensional LTI system over a scalar Gaussian channel. That is there is no rate loss in restricting the scheme to be linear. In fact, for the stabilization of scalar plant over a scalar Gaussian channel, linear policies are optimal. However for transmission over vector channels, linear schemes may not be good enough. It is known from the information theory literature [12] that a distributed joint source-channel code is optimal in the MMSE

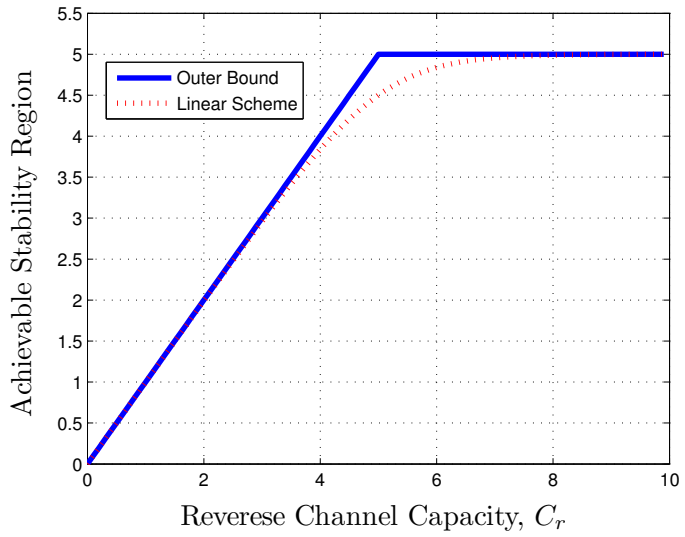


Fig. 1.5 Stability Region

sense, if the following conditions hold: i) The information transmitted on all available channels is independent, ii) Capacity is utilized by all channels (source-channel needs to be matched). By using any linear scheme, it is not possible to make the transmitted signals on parallel channels independent. However independent signals can be transmitted on parallel channels by employing non-linear schemes. Some non-linear sensing schemes for stabilization and control of a scalar system over parallel Gaussian channels are given in [13, 14]. In [13], the authors considered two parallel channels and proposed to send the magnitude of the observed state process on one channel and the phase value (plus or minus) on the second channel. The phase and magnitude of a signal are shown to be independent, thus satisfying the first condition of optimality. Although the second condition of optimality is not met, the proposed non-linear schemes outperforms the best linear scheme. In [14], the authors proposed to use a hybrid digital-analog scheme, in which the state process is quantized and the quantized signal is transmitted on one channel and quantization error is transmitted on the other channel. This scheme can be extended to arbitrary number of parallel channels and it achieves the minimum signal-to-noise ratio requirement for mean-square stabilization of a scalar noiseless plant.

In order to demonstrate inefficiency of linear schemes over vector channels, let us consider the following example.

Example: Consider a scalar LTI system that has to be stabilized over M parallel white Gaussian channels. Assume that the sensor has an average transmit power constraint P_S and all channels are disturbed by noises of equal power, i.e., $z_{i,t} \sim \mathcal{N}(0,N)$ for all $i \in \{1,2,\cdots,M\}$. It can be easily shown that the system can be stabilized over the given channel by a linear sensing and control scheme if $\log(\lambda) <$

$\frac{1}{2}(1+\frac{P}{N})$. However with the non-linear scheme proposed in [14], the achievable stability region is given by $\log(\lambda) < \frac{M}{2}(1+\frac{P}{MN})$. The stability regions achieved by linear and non-linear schemes have been plotted in Fig.1.6 for $P = 1$ and $M = 2, 10$, and 100. Note that in the given example, the achievable stability region of linear scheme is independent of number of available parallel channels M. But with the non-linear scheme, the stability region significantly enlarges as the number of parallel channels increases. This example shows that linear schemes can be very inefficient in some parallel channel settings.

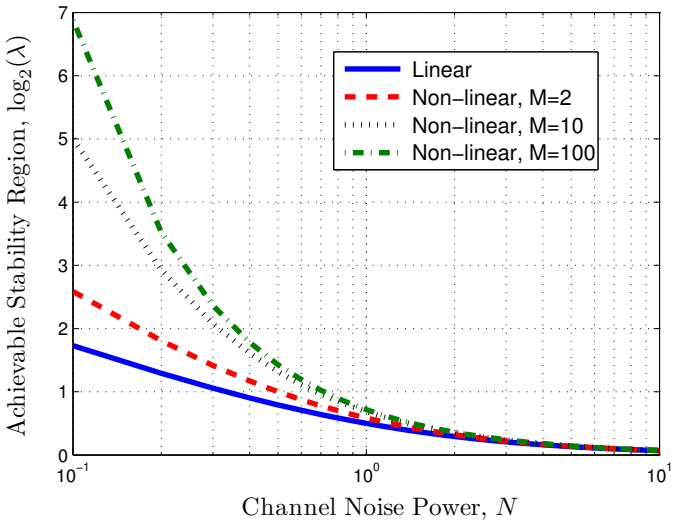


Fig. 1.6 Example: Comparison of linear and non-linear schemes.

The non-linear scheme proposed in [14] works for scalar plants. For stabilization of a multi-dimensional plant, the non-linear scheme of [13, 14] can be used together with the time varying (mode-by-mode transmission) scheme discussed in Sec. 1.3.1.2. Such a non-linear time varying scheme can achieve the minimum rate required for stabilization. An interesting open problem is to determine tight conditions for optimality of linear scheme for stabilization of multi-dimensional systems over vector Gaussian channel. In the following we state a sufficient condition for optimality of a linear time varying scheme proposed in [15].

Theorem 6. *([15, Theorem 3.2]) A linear time varying 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*

$$\log\left(|\lambda_i|\right) < \sum_{j=1}^{m^\star} \frac{f_{ij}}{2} \log\left(1+\frac{P_j^\star}{N_j}\right), \tag{1.45}$$

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 [16, pp. 204-205] and $m^\star \leq m$ is the number of active channels for which optimal transmit power is non-zero.

Some relevant works on the source-channel matching and optimality of linear estimation can be found in [17–23].

1.4 Stabilization Over Relay Channels

In this section we study stabilization of linear systems over Gaussian relay channels. The basic relay channel consists of one sender (source), one receiver (destination), and an intermediate node (relay) whose sole purpose is to help the communication between the source and the destination [24]. The basic three node relay channel is a basic block of a large sensor network where a group of sensor nodes cooperate to communicate information from a source to a destination. In order to understand the problem of stabilization over a general relay network, we study some basic relay network topologies such as non-orthogonal relay channel and orthogonal relay channel. By the orthogonality of the relay channel, we mean that the signal spaces of the encoder and the relay are orthogonal. For example, if the source node and the relay node transmit in disjoint frequency bands or non-overlapping time slots, then the relay channel is considered to be orthogonal. These topologies serve as the basic building blocks of a large network. In practice, the relay node can be either half-duplex or full-duplex. 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. 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.

The problem of control over a basic three node Gaussian relay channel was first introduced in [25, 26], some sufficient conditions for mean square stability were derived. Further related work on control over noisy relay channels can be found in [8, 27].

We know from [28] that the concept of Shannon capacity is not sufficient to characterize moment stability. Moreover even for the general three node Gaussian relay channel, a single-letter expression for Shannon capacity is still not known. In [29] Gastpar and Vetterli determined capacity of a particular large Gaussian relay network in the limit as the number of relays tends to infinity. The achievable information rate over the relay channel depends on the processing strategy of the relay. The most well known relaying strategies are amplify-and-forward (AF), compress-and-forward, and decode-and-forward [30]. The AF strategy is well suited for delay sensitive control applications and is therefore addressed here.

In this section we discuss mean-square stabilization of the system in (1.1) over some fundamental relay channels such as non-orthogonal half-duplex relay chan-

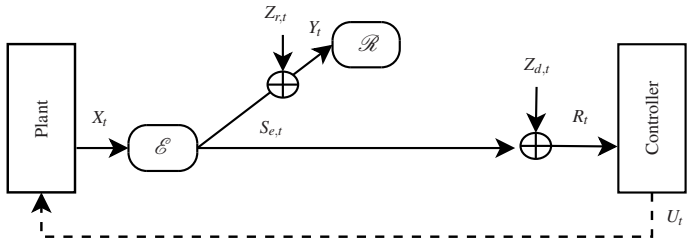
nels, non-orthogonal full-duplex relay channel, and orthogonal relay channel in Sec. 1.4.1, Sec. 1.4.2, and Sec. 1.4.3 respectively. For each relay channel, we present necessary conditions and sufficient conditions for stabilization. In Sec. 1.4.4 we briefly compare achievable stability regions using linear schemes over these basic relay channels.

1.4.1 Non-orthogonal Half-duplex relay

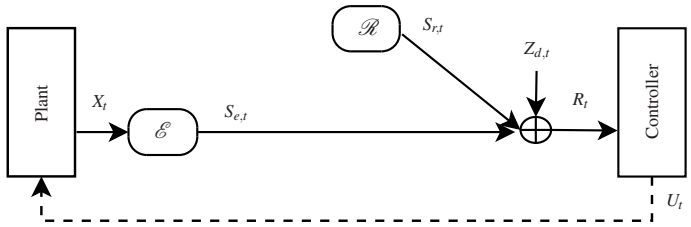
Consider a non-orthogonal half-duplex Gaussian relay channel shown in Fig. 1.7. A sensor node (state encoder) $\mathscr{E}$ senses the state of the plant and transmits it to the remote controller, while another sensor node $\mathscr{R}$ acts as a relay to support communication from $\mathscr{E}$ to the controller. The state encoder and the relay transmit in the same frequency band and the relay node is assumed to be half-duplex, i.e., it cannot transmit and receive signals simultaneously. In [8] the authors proposed a transmission protocol having two transmission phases, as shown in Fig. 1.7. The signals transmitted by $\mathscr{E}$ and $\mathscr{R}$ are denoted as $S_{e,t}$ and $S_{r,t}$ respectively. The variables $Z_{r,t}$ and Z_t denote two mutually independent white Gaussian noise components with zero mean and variances N_r and N respectively. In the first transmission phase (odd time steps), $\mathscr{E}$ transmits signal with an average power $2\beta P_S$, where $0 < \beta \leq 1$ is a parameter that distributes power between the two transmission phases. In this transmission phase the relay $\mathscr{R}$ receives a noisy signal Y_t from the encoder but it does not transmit any signal. In the second transmission phase (even time steps) both the encoder $\mathscr{E}$ and the relay $\mathscr{R}$ transmit with average powers $2(1-\beta)P_S$ and P_r respectively. The relay node employs amplify-and-forward (linear) strategy, where amplification at the relay is done under an average power constraint P_R. The multiplicative gain of the $\mathscr{E}-\mathscr{D}$ link is assumed to be a constant $h \in \mathbb{R}$ and the gain of the $\mathscr{R}-\mathscr{D}$ link is assumed to be one, without loss of generality. The presence of relay node can be more useful in scenarios where the direct link is weaker i.e., $|h|$ is small. The controller in the first transmission phase receives $R_t = hS_{e,t} + Z_t$ and in the second phase receives $R_t = hS_{e,t} + S_{r,t} + Z_t$. At any time, the controller estimates the present state of the plant using all the signals it has received so far, and then takes an action to stabilize the plant using its state estimate. In the following we discuss a linear control and communication scheme based on the above transmission protocol and give necessary and sufficient conditions for stabilization. The communication and control scheme is presented in Sec. 1.4.1.1 and the mean-square stability of the system under the given scheme is analyzed in Sec. 1.4.1.2.

1.4.1.1 Sensing and Control Scheme

The control and communication scheme has an initialization step, which is done to make the state distribution Gaussian. This initialization step works as follows. The encoder $\mathscr{E}$ observes X_0 and transmits $S_{e,0} = \sqrt{\frac{P_S}{\alpha_0}} X_0$. The controller $\mathscr{D}$ receives



(a) First transmission phase.



(b) Second transmission phase.

Fig. 1.7 Half-duplex AWGN relay channel.

$R_0 = hS_{e,0} + Z_0$ and estimates the initial state as

$$\hat{X}_0 = \frac{1}{h}\sqrt{\frac{\alpha_0}{P_S}}R_0 = X_0 + \frac{1}{h}\sqrt{\frac{\alpha_0}{P_S}}Z_0.$$

The controller $\mathscr{C}$ then takes an action $U_0 = -\lambda\hat{X}_0$ which results in

$$X_1 = \lambda X_0 + U_0 + W_0 = \lambda\left(X_0 - \hat{X}_0\right) + W_0 = -\frac{\lambda}{h}\sqrt{\frac{\alpha_0}{P_S}}Z_0 + W_0. \tag{1.46}$$

The new plant state X_1 is Gaussian distributed with zero mean and variance $\alpha_1 = \frac{\lambda^2 N}{h^2 P_S}\alpha_0 + n_w$. This initialization step is not required if the initial state is already Gaussian distributed. After the initialization, further transmissions are divided into two separate transmission phases as discussed earlier. In the first transmission phase, i.e., for $t = 1,3,5,\ldots$, the encoder $\mathscr{E}$ transmits $S_{e,t} = \sqrt{\frac{2\beta P_S}{\alpha_t}}X_t$ to the controller. The relay $\mathscr{R}$ operates in the receiving mode, i.e., it receives the signal transmitted by $\mathscr{E}$. The controller $\mathscr{C}$ observes $R_t = hS_{e,t} + Z_t$ and computes the MMSE estimate of X_t based on $R_{[0,t]}$. It can be shown that $\mathbb{E}[X_t R_{t-j}] = 0$ for $j \geq 1$, therefore the optimal MMSE estimator uses only the latest received signal R_t to estimate the state. Further, the optimal estimator is linear in the received signal due to the Gaussian distribution. The optimal MMSE state estimate is computed as $\hat{X}_t = \left(\frac{h\sqrt{2\beta P_S \alpha_t}}{2h^2\beta P_S + N}\right)R_t$. Based on the estimate $\hat{X}_t$, the controller $\mathscr{C}$ 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. The variance of X_{t+1} can be computed as,

$$\alpha_{t+1} := \mathbb{E}[X_{t+1}^2] = \lambda^2 \mathbb{E}[(X_t - \hat{X}_t)^2] + \mathbb{E}[W_t^2] = \lambda^2 \left(\frac{N}{2h^2 \beta P_S + N} \right) \alpha_t + n_w. \quad (1.47)$$

In the second transmission phase, i.e., for $t = 2,4,6,\ldots$, the encoder $\mathscr{E}$ transmits $S_{e,t} = \sqrt{\frac{2(1-\beta)P_S}{\alpha_t}} X_t$ to the controller. The relay now operates in the transmitting mode. It amplifies the previously received signal under the average transmit power constraint and transmits the following signal to the controller, $S_{r,t} = \sqrt{\frac{P_r}{(2\beta P_S + N_r)}} (S_{e,t-1} + Z_{r,t-1})$. The controller $\mathscr{C}$ thus receives a linear combination of the signal transmitted from the relay and the state encoder. The signal received at the controller is given by

$$R_t = hS_{e,t} + S_{r,t} + Z_t = L_1 X_t + L_2 X_{t-1} + \tilde{Z}_t, \quad (1.48)$$

where $L_1 = \sqrt{\frac{2(1-\beta)h^2 P_S}{\alpha_t}}$, $L_2 = \sqrt{\frac{2\beta P_S P_r}{(2\beta P_S + N_r)\alpha_{t-1}}}$, and $\tilde{Z}_t = Z_t + \sqrt{\frac{P_r}{2\beta P_S + N_r}} Z_{r,t-1}$ with $\tilde{Z}_t \sim \mathscr{N}(0, \tilde{N}(\beta, P_r))$. Next, the controller computes the MMSE estimate of X_t given all previous channel outputs $R_{[0,t]}$ in the following three steps: i) Compute the MMSE prediction of R_t from $R_{[0,t]}$ as $\hat{R}_t = L_2 \hat{X}_{t-1}$, where $\hat{X}_{t-1}$ is the MMSE estimate of X_{t-1}, ii) Compute the innovation as $I_t = R_t - \hat{R}_t$, and iii) Estimate the state using only the innovation I_t as $\hat{X}_t = \mathbb{E}[X_t | I_t]$. Note that this is the optimum MMSE estimate since X_t is independent of $\{R_1, R_2, ..., R_{t-1}\}$ due orthogonality property of MMSE estimation. The optimal MMSE state estimate is computed as

$$\hat{X}_t = \mathbb{E}[X_t | I_t] = \frac{\lambda (\lambda L_1 + L_2) \alpha_t}{(\lambda L_1 + L_2)^2 \alpha_t + L_2^2 n_w + \lambda^2 \tilde{N}(\beta, P_r)} I_t. \quad (1.49)$$

Based on the state estimate, the controller $\mathscr{C}$ takes an action $U_t = -\lambda \hat{X}_t$ that results in $X_{t+1} = \lambda (X_t - \hat{X}_t) + W_t$. The new plant state X_{t+1} is linear combination of zero mean Gaussian variables $\{X_t, \hat{X}_t, W_t\}$, therefore it is also zero mean Gaussian distributed. The variance of the new plant state X_{t+1} follows from simple computations as,

$$\begin{aligned}
\alpha_{t+1} &= \lambda^2 \mathbb{E}[(X_t - \hat{X}_t)^2] + \mathbb{E}[W_t^2] = \lambda^2 \alpha_t \left(\frac{L_2^2 n_w + \lambda^2 \tilde{N}(\beta, P_r)}{(\lambda L_1 + L_2)^2 \alpha_t + L_2^2 n_w + \lambda^2 \tilde{N}(\beta, P_r)} \right) + n_w \\
&= \lambda^2 \left(\lambda^2 k \alpha_{t-1} + n_w \right) \left(\frac{\left(\frac{n_w k_1}{\lambda^2} \right) \frac{1}{\alpha_{t-1}} + \tilde{N}(\beta, P_r)}{\left(k_2 + \sqrt{k_1 k + \frac{n_w k_1}{\lambda^2} \frac{1}{\alpha_{t-1}}} \right)^2 + \left(\frac{n_w k_1}{\lambda^2} \right) \frac{1}{\alpha_{t-1}} + \tilde{N}(\beta, P_r)} \right) + n_w,
\end{aligned} \quad (1.50)$$

where the last equality follows by substituting α_t from (1.47) and by defining $k := \frac{N}{2h^2\beta P_S+N}$, $k_1 := \frac{2\beta P_S P_r}{2\beta P_S+N_r}$, $k_2 := q\sqrt{2h^2(1-\beta)P_S)}$. Having presented the sensing and control scheme, we now discuss stability of the plant under the given scheme.

1.4.1.2 Stability Analysis

We wish to find the values of the system parameter λ for which the second moment of the state remains bounded, i.e., the sequence $\{\alpha_t\}$ has to be bounded. Rewriting (1.47) and (1.50), the variance of the state at any time t is given by

$$\alpha_t = \lambda^2\left(\frac{N}{2h^2\beta P_S+N}\right)\alpha_{t-1}+n_w, \qquad t=2,4,6,... \tag{1.51}$$

$$\alpha_t = \lambda^2\left(\lambda^2 k\alpha_{t-2}+n_w\right)f(\alpha_{t-2})+n_w, \qquad t=3,5,7,... \tag{1.52}$$

where $\alpha_1 = \frac{\lambda^2 N}{h^2 P_S}\alpha_0+n_w$ and $f(\alpha_{t-2}) \triangleq \left(\frac{(\frac{n_w k_1}{\lambda^2})\frac{1}{\alpha_{t-2}}+\tilde{N}(\beta,P_r)}{(k_2+\sqrt{k_1 k+\frac{n_w k_1}{\lambda^2}\frac{1}{\alpha_{t-2}}})^2+(\frac{n_w k_1}{\lambda^2})\frac{1}{\alpha_{t-2}}+\tilde{N}(\beta,P_r)}\right)$. If the odd indexed sub-sequence $\{\alpha_{2t+1}\}$ in (1.52) is bounded, then the even indexed sub-sequence $\{\alpha_{2t}\}$ in (1.51) is also bounded. Therefore it is sufficient to consider the odd indexed sub-sequence $\{\alpha_{2t+1}\}$. A complicated structure of $f(\alpha_t)$ in (1.52) makes it difficult to find a condition on λ for which this sequence is bounded. Therefore in [8] we use the following approach. We construct a sequence $\{\alpha'_t\}$ which upper bounds the sub-sequence $\{\alpha_{2t+1}\}$ and is easier to analyze. Then we derive a condition on the system parameter λ for which the sequence $\{\alpha'_t\}$ converges to a limit point as $t\to\infty$ and consequently the boundedness of $\{\alpha_{2t+1}\}$ is guaranteed. We will show later that there is no loss in considering the majorizing sequence instead of the original sequence. The detailed analysis is given in [8] and here we merely give the condition under which the system is stable:

$$\lambda^4 < \left(\frac{(k_2+\sqrt{k_1 k})^2+\tilde{N}(\beta,P_r)}{k\tilde{N}(\beta,P_r)}\right) \tag{1.53}$$

$$\Rightarrow \log(\lambda) <= \frac{1}{4}\left(\log\left(1+\frac{2h^2\beta P_S}{N}\right)+\log\left(1+\frac{\tilde{M}(\beta,P_r)}{\tilde{N}(\beta,P_r)}\right)\right), \tag{1.54}$$

where in the last equality we substituted $k=\frac{N}{2h^2\beta P_S+N}$ and $M(\beta,P_r)=(k_2+\sqrt{k_1 k})^2$ in order to show the dependencies on the average relay power P_r and the power allocation parameter β at the encoder. Since the relay node amplifies the desired signal as well as the noise which is then superimposed at the decoder to the signal coming directly from the encoder, an optimal choice of the relay transmit power $0\le P_r\le P_R$ depends on the relay channel parameters $\{P_S,N_r,N,h,\beta\}$. Moreover an optimal choice of the power allocation factor β at the encoder also depends on the relay channel parameters $\{P_S,P_r,N_r,N,h\}$. Therefore we can rewrite (1.54) as,

$$\log(\lambda) < \frac{1}{4} \max_{\substack{0<\beta\leq 1 \\ 0\leq P_r \leq P_R}} \left(\log\left(1+\frac{2h^2\beta P_S}{N}\right) + \log\left(1+\frac{\tilde{M}(\beta,P_r)}{\tilde{N}(\beta,P_r)}\right)\right), \tag{1.55}$$

which is a sufficient condition for mean-square stability of a scalar plant.

It is interesting to see that the sufficient condition for mean square stability does not depend on the process noise. This provides motivation to study stabilizability of the system in (1.1) without process noise, i.e., $W_t = 0$. In the absence of the process noise in (1.1), the state variance of the noiseless system at any time step t is then given by substituting $n_w = 0$ in (1.50), that is

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

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 $\left(\frac{\lambda^4 k\tilde{N}(\beta,P_r)}{(k_2+\sqrt{k_1 k})^2+\tilde{N}(\beta,P_r)}\right) < 1$. This is the same condition as in (1.53). Thus by using the proposed linear coding and control scheme, we obtain identical sufficient conditions for mean square stability of noisy and noiseless first LTI system over *half-duplex* relay channel. Although the sufficient conditions are identical, the state variance in the noisy plant scenario cannot converge to zero unlike the noiseless scenario.

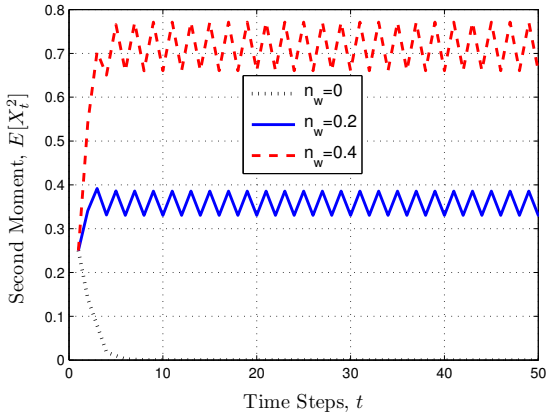


Fig. 1.8 Comparison of second moments of the plant state process at three different levels of process noise.

A comparison of the second moments of the plant's state process at three different power levels of the process noise is illustrated in Fig. 1.8. In this figure, we have fixed the relay channel parameters $\{P_S = 2, P_r = 2, h = 1, \beta = 0.5, N = 0.5, N_r = 0.1\}$, the plant parameters $\{\alpha_0 = 0.25, \lambda = 1.5\}$, and have plotted the second mo-

ment $\mathbb{E}[X_t^2]$ of the state process as a function of time t for three power levels of the process noise, i.e., $n_w = 0$, 0.2, and 0.4. For the given set of channel parameters, *mean square stability* of the system requires $\lambda < 1.975$ according to Theorem 7. In Fig. 1.8 we have fixed $\lambda = 1.5$ (i.e., less than 1.975), therefore starting from an arbitrary initial value the second moment of the state process stays bounded for all levels of the process noise. For $n_w = 0$ the second moment converges to zero, starting from an initial value equal to 0.25 as shown in Fig. 1.8. For non-zero values of process, the second moment keeps alternating between two different values. This happens due to the first and the second transmission phases . As shown in Fig. 1.8, for $n_w = 0.2$ and $n_w = 0.4$ 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. In Fig. 1.8 we can also observe that the rate of convergence is similar in the three examples, and seems to be unaffected by the power level of the process noise.

The sufficient condition for a multi-dimensional plant can be obtained by using the time varying (mode-by-mode transmission) scheme proposed in Sec. 1.3.1.2 together with the linear scheme used for the scalar plant above. With the similar analysis as above, we can prove the following theorem.

Theorem 7. *([8, Theorem 3.1]) The linear time invariant system in (1.1) can be mean square stabilized over the* half-duplex *AWGN relay channel if*

$$\log(|A|) < \frac{1}{4} \max_{\substack{0<\beta\leq 1 \\ 0\leq P_r\leq P_R}} \left(\log\left(1+\frac{2h^2\beta P_S}{N}\right) + \log\left(1+\frac{\tilde{M}(\beta,P_r)}{\tilde{N}(\beta,P_r)}\right)\right), \qquad (1.56)$$

where $\tilde{N}(\beta,P_r) = \frac{P_r N_r}{2\beta P_S + N_r} + N$, $\beta \in [0,1]$, *and* $\tilde{M}(\beta,P_r) = \left(\sqrt{2h^2(1-\beta)P_S} + \sqrt{\frac{2\beta P_S P_r N}{(2\beta P_S+N_r)(2h^2\beta P_S+N)}}\right)^2$.

Remark 2. It has been shown in [8][Appendix I] that the term on the right hand side of (1.56) is the information rate over the half-duplex AWGN relay channel with noiseless feedback.

Optimal choices of the power allocation parameter β at the encoder and the relay transmit power P_r which maximize the term on the right hand side of (1.56) depend on the quality (i.e., SNR) of $\mathscr{E}-\mathscr{D}$, $\mathscr{E}-\mathscr{R}$, and $\mathscr{R}-\mathscr{D}$ links. To illustrate this, we have plotted optimal relay power $P_r^\star$ as a function of the relay noise power N_r for fixed values of $P_R = 2, P_S = 10, N = 1$ in Fig. 1.9, with the help of numerical computations. We observe that for low values of N_r, the relay uses all available power. As the N_r increases, $P_r^\star$ decreases because the relay is using an amplify-and-forward strategy in which noise also gets amplified along with the signal of interest (state information). Eventually $P_r^\star$ goes to zero for very high values of N_r, indicating the fact that the relay is not useful anymore if linear strategy is employed at the relay. However non-linear strategies might be useful for high values of N_r. Some non-linear relaying protocols have been proposed in [31] for the given half-duplex non-orthogonal relay channel, which give significantly higher transmission

rates than the linear relaying. Such nonlinear schemes [31] can potentially enlarge the achievable stability region if one uses them for remote stabilization. However a careful analysis is yet to be carried out to quantify the gains one can obtain in terms of stability with those non-linear relaying strategies.

The optimal choices of power allocation parameter β has not been plotted, however, we have observed via numerical experiments that $\beta = 0.5$ is usually a good choice which corresponds to an equal power allocation to the two transmission phases. For very low values of N_r (i.e., very reliable $\mathscr{E} - \mathscr{R}$ link), an optimal β can be slightly greater than 0.5, which is due to the reason that the communication via the relay can be more helpful.

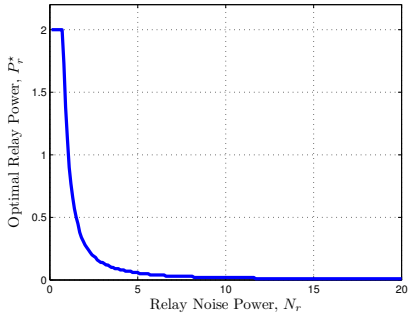


Fig. 1.9 Optimal relay power P_r for $P_R = 2, P_S = 10, N = 1$.

We now consider a special case, where there is no direct communication link from the encoder to the decoder and the information can be communicated only via the relay. We call this setup as a *two-hop* relay channel, since the communication from the sensor to the controller takes places in two hops- the first hop is from the $\mathscr{E}$ to $\mathscr{R}$ and the second hop is from $\mathscr{R}$ to $\mathscr{C}$. The half-duplex relay channel discussed earlier becomes *two-hop* if $h = 0$. Naturally for this case we choose $\beta = 1$ and $P_r = P_R$ and obtain the following sufficient condition for stabilization.

Corollary 1. *([8, Corollary 3.2]) The linear time invariant system in (1.1) can be mean square stabilized over a* two-hop *half-duplex AWGN relay channel if*

$$\log(|A|) < \frac{1}{4}\log\left(1 + \frac{2P_S P_R}{P_R N_r + N(2P_S + N_r)}\right). \tag{1.57}$$

For a setup which is equivalent to the *two-hop* relay channel, we find a necessary condition in [11, Theorem 4.1] which reads as

$$\log(|A|) < \frac{1}{4}\min\left\{\log\left(1 + \frac{2P_S}{N_r}\right), \log\left(1 + \frac{P_R}{N}\right)\right\}.$$

The condition in (1.57) becomes both necessary and sufficient if either the $\mathscr{E}-\mathscr{R}$ link is noiseless ($N_r = 0$) or the $\mathscr{R}-\mathscr{D}$ link is noiseless ($N = 0$).

Consider a *two-hop* relay channel with a causal noiseless feedback link from the controller to the relay. For this setup, the condition in (1.57) becomes necessary and sufficient if we restrict the encoder to be linear in the state. This result is an application of a result in [32, 33]. It follows from the following arguments:

For the *two-hop* relaying scenario with a noiseless causal feedback link from the controller to the relay, we have a partially nested type of information pattern. It is known that the separation of estimation and control holds for such an information pattern and there is no dual effect of control [34]. The optimal control strategy using dynamic programming is $U_t = -\lambda \mathbb{E}[X_t|R_0^t]$, where $R_0^t = \{R_i, 0 \le i \le t\}$. By applying the optimal control action, the plant's state at any time t is given by $X_{t+1} = \lambda (X_t - E[X_t|R_0^t]) + W_t$. If we restrict the state encoder policy to be linear, then an innovation (memoryless) encoder is optimal since the control actions whiten the state process. Given a linear and memoryless policy at the encoder, let us now find an optimal relaying policy which minimizes $\mathbb{E}[X_{t+1}^2] = \lambda^2 \mathbb{E}[(X_t - E[X_t|R_0^t])^2] + \mathbb{E}[W_t^2]$. The cost to be minimized is

$$\begin{aligned}\mathbb{E}[(X_t - E[X_t|R_0^t])^2] &\overset{(a)}{=} \mathbb{E}\left[(X_t - E[X_t|Y_0^t])^2\right] + \mathbb{E}\left[(E[X_t|Y_0^t] - E[X_t|R_0^t])^2\right] \\ &\overset{(b)}{=} \mathbb{E}\left[(X_t - E[X_t|Y_0^t])^2\right] + \mathbb{E}\left[(cY_t - E[X_t|R_0^t])^2\right] \qquad (1.58)\end{aligned}$$

where (a) follows from $\mathbb{E}[(X_t - E[X_t|Y_0^t])(E[X_t|Y_0^t] - E[X_t|R_0^t])] = 0$ (by the orthogonality principle of MMSE estimation); and (b) follows from the fact that the encoder transmits only innovation at each time step and the MMSE estimation of a Gaussian variable is linear, i.e. $E[X_t|Y_0^t] = cY_t$, where c is a scalar. An optimal relaying policy is the one which minimizes $E[(cY_t - E[X_t|R_0^t])^2]$, since the remaining term of the cost function in (1.58) is independent of the relaying policy. A similar problem was studied in [35], from which it follows that an optimal relaying policy is linear and memoryless. We have earlier obtained the sufficient condition in (1.57) by using optimal linear (memoryless) communication and control policies, therefore this condition is also necessary provided that the encoder is constrained to be linear in the state. Moreover, if we restrict the relay to be linear, then the two-hop relay channel becomes equivalent to a scalar Gaussian channel. For this channel, it has been shown earlier that a linear scheme is optimal. Therefore if the relay is restricted to be linear in the received signal, then the condition in (1.57) becomes necessary and sufficient.

1.4.2 Non-orthogonal Full-duplex Relay Channel

Although a half-duplex relay node is easier to build compared to a full-duplex node, there is some loss in the performance due to its inability to communicate simultaneously with state encoder and the controller. A full-duplex system can be

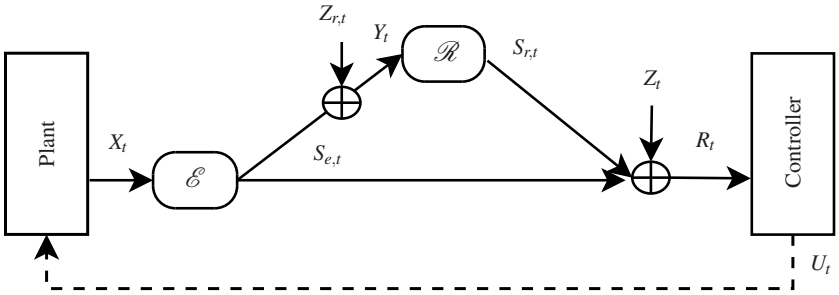


Fig. 1.10 Non-orthogonal full-duplex AWGN relay channel.

realized by placing transmit and receive antennas far enough to ensure sufficient isolation and/or by incorporating some interference cancelation schemes in analog and/or digital domain. In the following we consider remote stabilization of linear plant over non-orthogonal full-duplex Gaussian relay channel depicted in Fig. 1.10. The variables $\{Z_{r,t}, Z_{d,t}\}$ denote mutually independent white noise components with $Z_{r,t} \sim \mathscr{N}(0,N_r)$ and $Z_{d,t} \sim \mathscr{N}(0,N)$. The gain of $\mathscr{R}-\mathscr{C}$ link is denoted by h. At time step t, the encoder $\mathscr{E}$ inputs $S_{e,t}$ to the relay channel with an average power P_S. The relay simultaneously listens to $S_{e,t}$ and transmits $S_{r,t}$ which is the amplified version of the noisy signal received in the time step $t-1$. The amplification at the relay is done under an average power constraint P_r, where $0 \leq P_r \leq P_R$. That is, the relay transmits, $S_{r,t} = \sqrt{\frac{P_r}{P_S+N_R}}(S_{e,t-1}+Z_{r,t-1})$. The controller receives $R_t = S_{e,t} + hS_{r,t} + Z_t$, computes the MMSE estimate of the state as, $\hat{X}_t = \mathbb{E}[X_t|R_{[0,t]}]$, and then applies an action to stabilize the system, as done in the half-duplex case. The stabilization of linear plant over full-duplex Gaussian relay channel under a linear scheme has been in studied in [25]. In the following we present sufficient condition for stabilization under the best linear scheme.

Theorem 8. *([25, Theorem 6]) The linear system in (1.1) can be mean square stabilized over the* non-orthogonal full-duplex *AWGN relay channel if*

$$\log(|A|) < \frac{1}{2}\max_{0\leq P_r\leq P_R}\log\left(1+\frac{\left(\sqrt{P_S(P_S+N_R)}+\eta^\star h\sqrt{P_SP_R}\right)^2}{h^2P_RN_R+N(P_S+N_R)}\right), \tag{1.59}$$

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

$$\left(\frac{h^2P_SP_R}{P_S+N_R}\right)\eta^4+\left(2hP_S\sqrt{\frac{P_R}{P_S+N_R}}\right)\eta^3+\left(P_S+N+\frac{h^2P_RN_R}{P_S+N_R}\right)\eta^2 = \left(N+\frac{h^2P_RN_R}{P_S+N_R}\right).$$

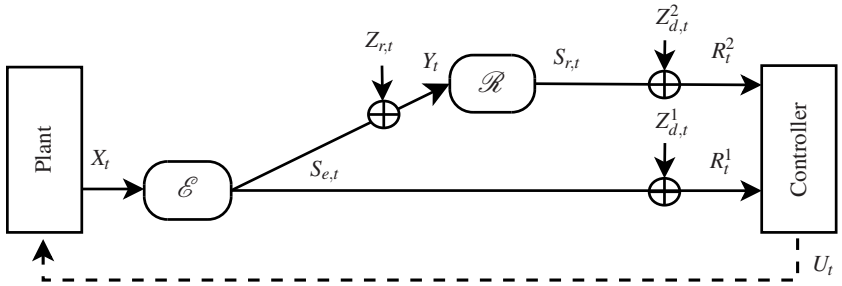


Fig. 1.11 Orthogonal half-duplex AWGN relay channel.

Proof. This theorem can be proved by employing linear sensing and control policies, and by following the same analysis as in the proof of Theorem 7. The detailed proof for a scalar plant can be found in [25]. The result can be extended to multi-dimensional systems using a similar time-sharing (mode-by-mode) transmission scheme as discussed in Sec. 1.3.1.2.

1.4.3 Orthogonal Relay Channel

An orthogonal AWGN relay channel is depicted in Fig. 1.11. The variables $\{Z_{r,t}, Z_{1,t}, Z_{2,t}\}$ denote mutually independent white noise components with $Z_{r,t} \sim \mathscr{N}(0,N_r)$ and $Z^i_{d,t} \sim \mathscr{N}(0,N_i)$ for $i \in \{1,2\}$. At any discrete time step t the encoder $\mathscr{E}$ inputs $S_{e,t}$ to the relay channel with an average power P_S. The relay observes $S_{e,t}$ in noise, amplifies it under an average power constraint P_R and forwards it to the controller. Accordingly, the relay transmits

$$S_{r,t} = \alpha(S_{e,t} + Z_{r,t}) = \sqrt{\frac{P_R}{P_S + N_R}}(S_{e,t} + Z_{r,t}), \tag{1.60}$$

where the amplification factor α is chosen equal to $\sqrt{\frac{P_R}{P_S+N_R}}$ in order to satisfy the average power constraint, i.e., $\mathbb{E}[S^2_{r,t}] \leq P_R$. The output of the relay channel at the decoder $\mathscr{D}$ is $\{R^1_t, R^2_t\}$, which is given by

$$\begin{aligned} R^1_t &= S_{e,t} + Z^1_{d,t}, \\ R^2_t &= S_{r,t} + Z^2_{d,t} = \alpha S_{e,t} + \tilde{Z}_t, \end{aligned} \tag{1.61}$$

where $\tilde{Z}_t \sim \mathscr{N}(0, \alpha^2 N_r + N_2)$.

By using a linear sensing and control scheme over the given channel (as done in the previous sections), we obtain the following sufficient condition for stabilization.

Theorem 9. *([25, Theorem 2]) The linear time invariant system in (1.1) can be mean square stabilized over the* orthogonal half-duplex *AWGN relay channel if*

$$\log(|A|) < \frac{1}{2}\log\left(1+\frac{P_S(P_SN_2+P_RN_R+N_2N_R+P_RN_1)}{N_1(P_RN_R+P_SN_2+N_RN_2)}\right). \tag{1.62}$$

Proof. This theorem can be proved by employing linear sensing and control policies, and by following the same analysis as in the proof of Theorem 7. The detailed proof for a scalar plant can be found in [25].

By using information theoretic arguments, we can obtain the following necessary condition for stabilization .

Theorem 10. *The linear time invariant system in (1.1) can be mean square stabilized over the* orthogonal half-duplex *AWGN relay channel only if*

$$\log(|A|) < \frac{1}{2}\min\left\{\log\left(1+\frac{P_S}{N_1}\right)+\log\left(1+\frac{P_R}{N_2}\right), \log\left(1+\frac{P_S}{N_1}+\frac{P_S}{N_r}\right)\right\}. \tag{1.63}$$

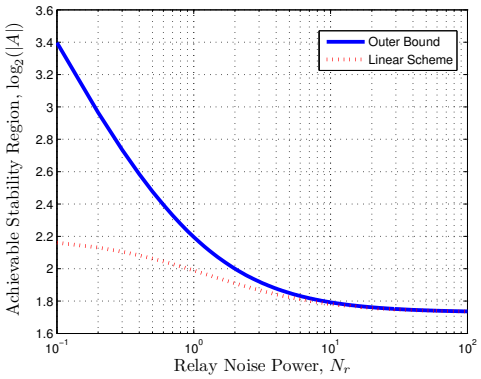


Fig. 1.12 Achievable Stability Region for $P_S = P_R = 10, N_1 = N_2 = 1$.

In order to see the performance of the proposed linear scheme over the orthogonal relay channel, we plot two achievable stability region in Fig. 1.12 and Fig. 1.13 as functions of N_r and N_2 according to (1.62). For comparison we also show the outer bound on stability region using (1.63). We can observe that the linear scheme usually performs good when either N_r is much greater than N_2 or when N_2 is much greater than N_r, i.e., when one of the two links (either $\mathscr{E}-\mathscr{R}$ or $\mathscr{R}-\mathscr{C}$) is much stronger than the other. However in some regimes, there is a large gap between stability region achieved by the linear scheme and the outer bound, indicating that the linear schemes can be highly suboptimal in general for orthogonal Gaussian relay networks. In Sec. 1.6 we will present some non-linear relaying schemes for real-time transmission of a Gaussian source over an orthogonal Gaussian relay channel. Those non-linear schemes significantly outperform linear schemes, which makes them potential candidates to be used in remote control or stabilization over Gaussian relay networks.

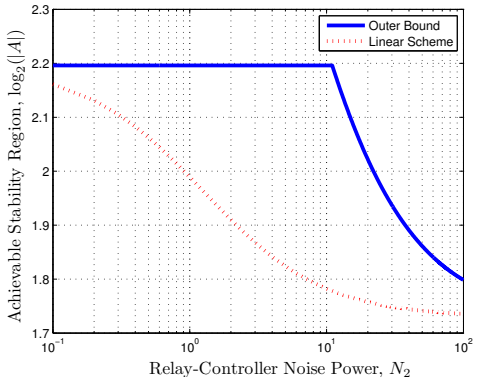


Fig. 1.13 Achievable Stability Region for $P_S = P_R = 10, N_1 = N_r = 1$.

1.4.4 Comparison of Relaying Topologies

We have so far studied following three relaying topologies: i) Non-orthogonal half-duplex relay channel, ii) Non-orthogonal full-duplex relay channel, iii) Orthogonal relay channel. In Fig.1.14 we make a comparison of achievable stability regions over these network topologies. We fix $P_S = 10, P_R = 1, N = N_1 = N_2 = 1$ and show the stability regions that are achieved with linear schemes as functions of increasing relay noise power N_r, according to Theorem 7, Theorem 8, and Theorem 9. Since in the half-duplex setting the relay transmits in alternate time steps, it can transmit with double power compared to the full-duplex case. Therefore for the full-duplex relay and the orthogonal relay channels we have used $P_R = 1$, whereas for the half-duplex relay channel we have used $P_R = 2$ while plotting achievable stability regions in Fig.1.14. The figure shows that the full-duplex relaying is superior to the half-duplex relaying. The reader should keep in mind that the full-duplex sensor nodes are usually more expensive due to the implementation issues discussed earlier. Moreover, we can observe that the orthogonal relaying outperforms non-orthogonal relaying for higher values of N_r, however this gain is obtained at the cost of using more channel resources (for example using extra bandwidth).

1.5 Sub-optimality of Linear Policies for Multi-hop Networks

It is known from [36, 37] that linear schemes are not optimal in general for estimation and control over Gaussian multi-hop relay networks. The paper [36] considers the problem of transmission of an i.i.d. Gaussian source over most basic two-hop relay channel illustrated in Fig. 1.5. The problem formulation is as follows: Consider a sequence of independent and identically distributed real-valued Gaussian random variables $\{X_t\}_{t\in\mathbb{N}}$ having zero mean and variance σ_x^2, where t denotes a

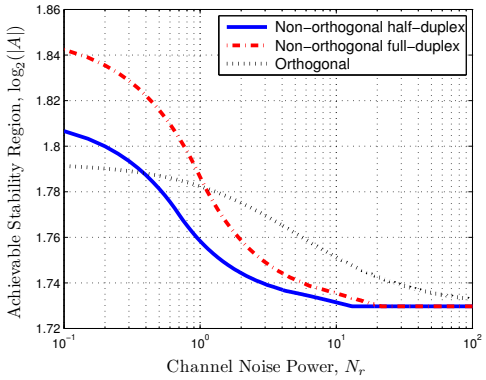


Fig. 1.14 Example: Comparison of linear and non-linear schemes.

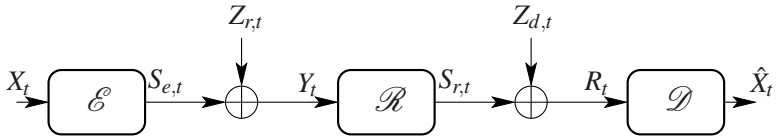


Fig. 1.15 Real-time transmission of a memoryless Gaussian source over a two-hop relay channel.

discrete time index. According to the figure, at a discrete time $t \in \mathbb{N}$ the source encoder $\mathscr{E}$ observes X_t and produces $S_{e,t} = f_{1,t}(X_{[0,t]})$ suitable for transmission, where $f_{1,t} : \mathbb{R}^t \mapsto \mathbb{R}$ is a causal measurable mapping. The mapping $f_{1,t}$ has to satisfy the following average power constraint,

$$\mathbb{E}[S_{e,t}^2] \leq P_S. \tag{1.64}$$

The signal $S_{e,t}$ is then observed in noise by the relay node $\mathscr{R}$ as $Y_t = S_{e,t} + Z_{r,t}$, where $\{Z_{r,t}\}_{t\in\mathbb{N}}$ 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 $f_{2,t} : \mathbb{R}^t \mapsto \mathbb{R}$ to produce $S_{r,t} = f_{2,t}(Y_{[0,t]})$ under the power constraint,

$$\mathbb{E}[S_{r,t}^2] \leq P_r. \tag{1.65}$$

The signal $S_{r,t}$ is then transmitted over a Gaussian channel. Accordingly the destination node $\mathscr{D}$ receives $R_t = S_{r,t} + Z_{d,t}$, where $\{Z_{d,t}\}_{t\in\mathbb{N}}$ is a zero mean white Gaussian noise sequence of variance N. Upon receiving R_t the decoder wishes to reconstruct the transmitted variable X_t by applying a mapping $g_t : \mathbb{R}^t \mapsto \mathbb{R}$ to produce $\hat{X}_t = g_t(R_{[0,t]})$. Let us define the signal-to-noise ratios of the $\mathscr{E}$–$\mathscr{R}$ and $\mathscr{R}$–$\mathscr{D}$ links as $\gamma_r := P_S/N_r$ and $\gamma_d := P_r/N$ 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

$$D = \limsup_{T\to\infty} \frac{1}{T+1} \sum_{t=0}^{T} \mathbb{E}\left[\left(X_t - \hat{X}_t\right)^2\right] \tag{1.66}$$

is minimized subject to the power constraints.

It has been shown in [36] that the following simple time invariant non-linear source and relay policies can beat the best linear scheme in some cases:

$$f_{2,t}(y_t) = \left\{ \begin{array}{ccc} b, & \text{for} & y_t > m_2 \\ 0, & \text{for} & |y_t| \le m_2 \\ -b, & \text{for} & y_t < -m_2 \end{array} \right\},$$

$$f_{1,t}(x_t) = \left\{ \begin{array}{ccc} a, & \text{for} & x_t > m_1 \\ 0, & \text{for} & |x_t| \le m_1 \\ -a, & \text{for} & x_t < -m_1 \end{array} \right\},$$

The functions $f_{1,t}(\cdot)$ and $f_{2,t}(\cdot)$ functions are illustrated in Figure 1.16.

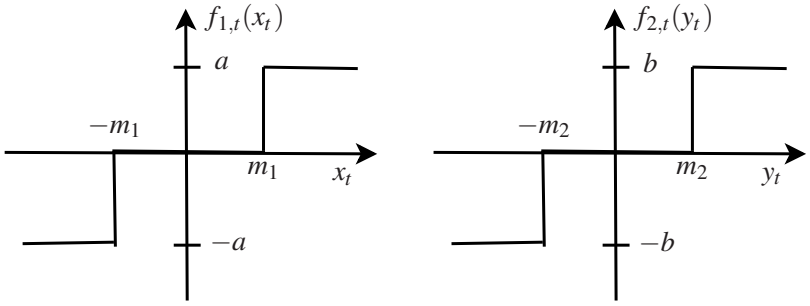
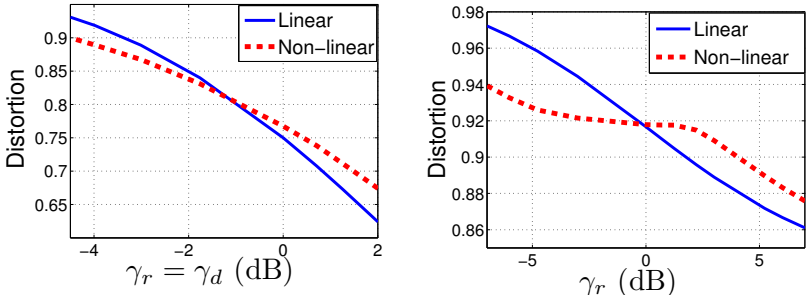


Fig. 1.16 Source and Relay Policies.

In Figure 1.17 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. These figures demonstrate that the simple three-level quantizer policies can outperform the best linear policies. The proposed non-linear scheme is not always better than the optimal linear scheme as demonstrated in Fig. 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, the proposed non-linear strategy is superior because it does not amplify the large values of channel noise at its input unlike the linear (amplify-and-forward) strategy. When linear schemes are employed in multi-hop relay networks, noise is accumulated in every hop, whereas non-linear schemes can suppress noise. In Sec. 1.6 we discuss an algorithm to numerically op-

timize the source and relay mappings for an orthogonal relay channel. One can use a similar approach to numerically optimize the source and relay mappings for the given two-hop relay channel as well.



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

Fig. 1.17 Comparison of the linear and the non-linear schemes.

The motivation for choosing three-level quantizer policies comes from the the multi-stage decision problem studied in [37] where a binary quantizer was shown to beat the best linear policy for five or more stages. The given two-hop relaying setup corresponds to three stages. It is however not known whether binary quantizers are always worse than the best linear scheme for the given three-stage problem. The intuition for choosing symmetric quantizer comes from 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.

1.6 Real-time Transmission Over an Orthogonal Relay Channel

In this section we consider real-time transmission of a memoryless source over an orthogonal half-duplex Gaussian relay channel. The system model is depicted in Fig. 1.18, where a sensor node $\mathscr{E}$ observes a Gaussian variable $X \sim \mathscr{N}(0, \sigma_x^2)$ and transmits it to the destination $\mathscr{D}$ over a Gaussian channel. An intermediate sensor node $\mathscr{R}$ called relay, overhears the signal transmitted from sensor $\mathscr{E}$ and relays its received information to the destination $\mathscr{D}$ over an orthogonal channel. We assume that the relay is half-duplex, i.e., it cannot simultaneously receive and transmit signals. For real-time coding of i.i.d. sources, memoryless coding is optimal [33]. Therefore we consider memoryless encoders. For each source sample X_i, the source encoder $\mathscr{E}$ uses channel K_1 times and the relay encoder $\mathscr{R}$ uses channel K_2, where K_1, K_2 are positive integers. Thus the transmission of source each sample takes $K = K_1 + K_2$

channel uses. For each source sample, $\mathscr{E}$ transmits $\mathbf{S}_e = f_e(X)$, where $f_e : \mathbb{R} \mapsto \mathbb{R}^{K_1}$, subject to an average power constraint $\mathbb{E}[f_e^2(X)] \leq P_S$. The relay and the destination accordingly receive,

$$\begin{aligned} \mathbf{Y} &= \mathbf{S}_e + \mathbf{Z}_r, \\ \mathbf{R}^1 &= \mathbf{S}_e + \mathbf{Z}_d^1, \end{aligned} \tag{1.67}$$

where $\mathbf{Z}_r, \mathbf{Z}_d^1 \in \mathbb{R}^{K_1}$ are mutually independent zero mean white Gaussian noise vectors with $\mathbb{E}\left[\mathbf{Z}_d^1 \left(\mathbf{Z}_d^1\right)^T\right] = N_d^1 I$ and $\mathbb{E}\left[\mathbf{Z}_r \mathbf{Z}_r^T\right] = N_r I$. Upon receiving $\mathbf{Y}$, $\mathscr{R}$ transmits $\mathbf{S}_r = f_r(\mathbf{Y})$, where $f_r : \mathbb{R}^{K_1} \mapsto \mathbb{R}^{K_2}$ is subject to an average power constraint $\mathbb{E}[f_r^2(\mathbf{Y})] \leq P_r$. Accordingly the destination $\mathscr{D}$ receives $\mathbf{R}^2 = \mathbf{S}_r + \mathbf{Z}_d^2$, where $\mathbf{Z}_d^2 \in \mathbb{R}^{K_2}$ is a white Gaussian noise vector with $\mathbb{E}\left[\mathbf{Z}_d^2 \left(\mathbf{Z}_d^2\right)^T\right] = N_d^2 I$. After receiving signals from both $\mathscr{E}$ and $\mathscr{R}$, the decoder $\mathscr{D}$ reconstructs X as $\hat{X} = f_d(\mathbf{R}^1, \mathbf{R}^2)$ where the mapping $f_d : \mathbb{R}^K \mapsto \mathbb{R}$ is chosen such that the mean-squared error $\mathbb{E}\left[\left(X - \hat{X}\right)^2\right]$ is minimized.

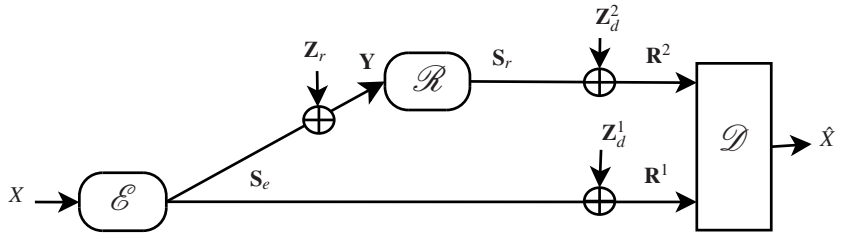
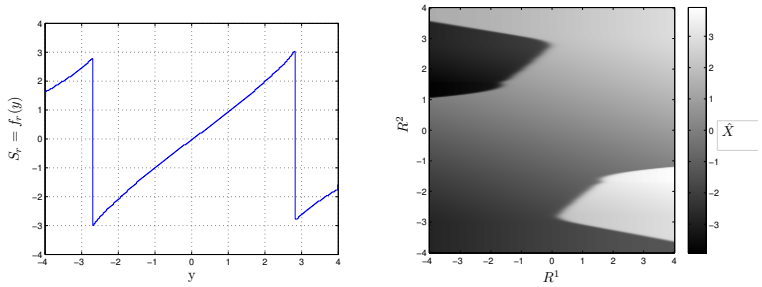


Fig. 1.18 Real-time transmission of a Gaussian source over an orthogonal half-duplex Gaussian relay channel.

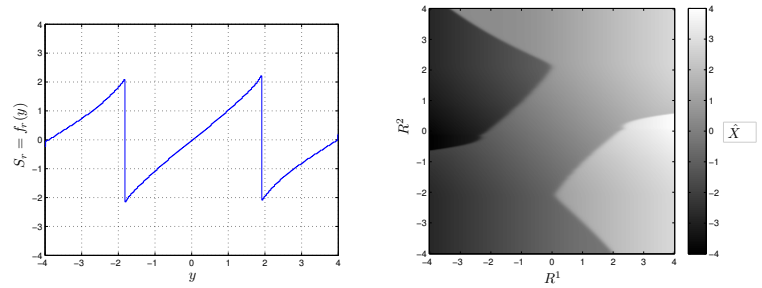
The problem of real-time transmission of a Gaussian source over the given three node orthogonal Gaussian relay channel has been studied in [38], where the authors propose an algorithm to numerically optimize the source, relay, and destination mappings. Since there are three mappings to be optimized for a given set of channel parameters, the design algorithm in [38] uses a common strategy of optimizing one mapping at a time while keeping the other two fixed. Moreover each dimension of the channel space is discretized into equally spaced points in the design algorithm. The optimized mappings obtained in [38] are in general non-linear and are shown to provide significant gains over linear mappings in terms lower achievable distortion. In the following we use the design algorithm of [38] to optimize mappings for some fixed channel parameters. We give examples of optimized mappings for only two cases: i) $K_1 = K_2 = 1$, and ii) $K_1 = 1, K_2 = 2$. Let us define the SNRs of the $\mathscr{E} - \mathscr{D}$, $\mathscr{E} - \mathscr{R}$, and $\mathscr{R} - \mathscr{D}$ links as $\gamma_{ed} := P_S/N_d^1$, $\gamma_{er} := P_S/N_r$, and $\gamma_{rd} := P_r/N_d^2$ respectively. For the case $K_1 = K_2 = 1$, we provide some examples of optimized mappings in Fig. 1.19-1.21. In these examples we have fixed the source mapping to be linear and optimized the other two mappings (relay and decoder mappings) for

different values of signal-to-noise ratios as given in the captions of the respective figures. We observe that these optimized relay mappings are non-invertible and have an almost periodic like behavior. Several input values are mapped to the same output value; this way of reusing output values can be seen as Wyner-Ziv type compression. This reuse of output values (an almost periodic behavior) makes the relay mappings more power efficient. Such non-invertible mappings have become possible due to the availability of the side information via the $\mathscr{E} - \mathscr{D}$ link. In Fig. 1.19-1.21 we have also plotted decoder mappings, which basically estimate the source X using the two received signals R^1 and R^2. The decision regions along with the reconstructions $\hat{X}$ are also shown in the figures. From these examples of optimized mappings, we observe that the number of periods in the relay mappings increase as the reliability of side information increases, thus making the relay more power efficient.



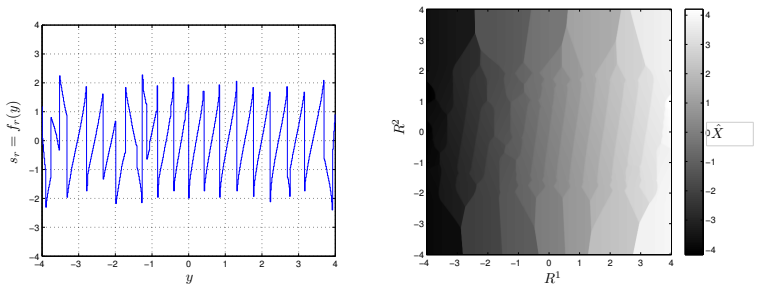
(a) Relay mapping (b) Decoder mapping

Fig. 1.19 Relay and decoder mappings ($K_1 = K_2 = 1$) optimized for $\sigma_x^2 = P_s = P_r = 1$ and $\gamma_{ed} = 5$ dB, $\gamma_{er} = 10$ dB, and $\gamma_{rd} = 25$ dB.



(a) Relay mapping (b) Decoder mapping

Fig. 1.20 Relay and decoder mappings ($K_1 = K_2 = 1$) optimized for $\sigma_x^2 = P_s = P_r = 1$ and $\gamma_{ed} = 10$ dB, $\gamma_{er} = 10$ dB, and $\gamma_{rd} = 25$ dB.



(a) Relay mapping (b) Decoder mapping

Fig. 1.21 Relay and decoder mappings ($K_1 = K_2 = 1$) optimized for $\sigma_x^2 = P_s = P_r = 1$ and $\gamma_{ed} = 20$ dB, $\gamma_{er} = 30$ dB, and $\gamma_{rd} = 25$ dB.

In Fig. 1.22 and Fig. 1.23, we give two examples of optimized relay mappings for $K_1 = 1, K_2 = 2$. That is the case where the relay performs an expansion — from its one-dimensional input to its two-dimensional output. Once again, there is a reuse of the output symbols which is only possible due to the side information from the direct link. As reliability of the side information increases, the reuse of the same output values also increases. The mappings have a spiral like shape. As reliability of direct link increases, the reuse of same output values also increase. Looking at the spiral from above, a similarity to the polynomial based source–channel codes proposed in [39, 40] can be seen.

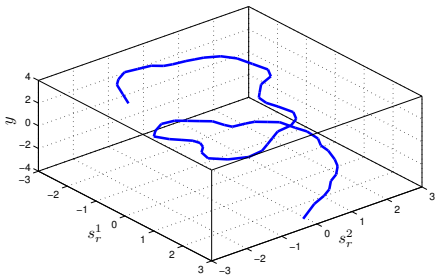


Fig. 1.22 Relay mapping ($K_1 = 1, K_2 = 2$) optimized for $\sigma_x^2 = P_s = P_r = 1$ and $\gamma_{ed} = 5$ dB, $\gamma_{er} = 15$ dB, and $\gamma_{rd} = 10$ dB.

In order to see the gains of non-linear optimized mappings over linear mappings in terms of achievable distortion, we refer the reader to [38]. In [38] the authors have analyzed the performance in detail for various values of channel dimensions (K_1 and K_2) and signal-to-noise ratios. It is observed that with these optimized mappings significantly lower distortion can be achieved, which makes them very useful for

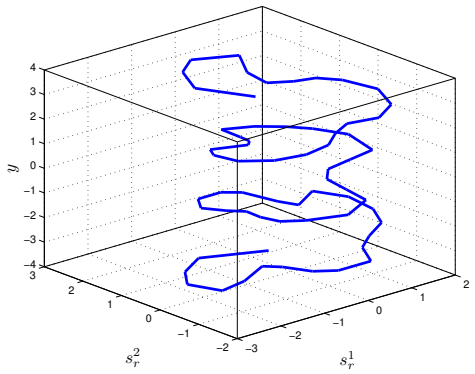


Fig. 1.23 Relay mapping ($K_1 = 1, K_2 = 2$) optimized for $\sigma_x^2 = P_s = P_r = 1$ and $\gamma_{ed} = 15$ dB, $\gamma_{er} = 15$ dB, and $\gamma_{rd} = 10$ dB.

large sensor networks and remote control scenarios where delay is a critical factor and transmit powers are limited.

1.7 Distributed Sensing for Control

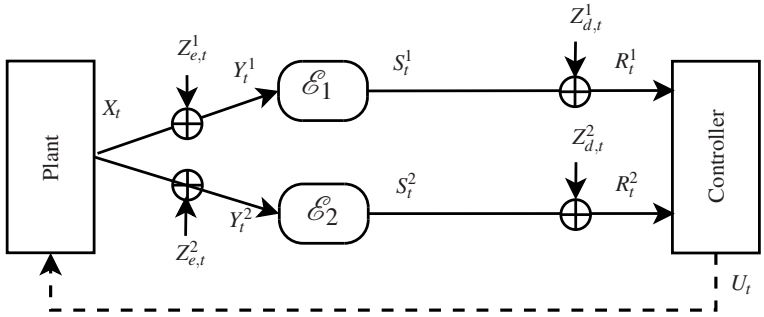


Fig. 1.24 A closed-loop control system with state measurements transmitted over wireless channels.

In this section we consider a multi-sensor setup, where multiple sensors in parallel observe noisy versions of the state process and communicate their observations to a remotely situated controller over orthogonal (parallel) channels. This scenario is different from the one studied earlier since each sensor has access to a different observation due to the addition of the measurement noise. The sensors then transmit their local observations to the controller. The schemes of [13, 14] can also be used

for distributed sensing. However it is not known how useful they are in the presence of measurement noise.

For the sake of simplicity, consider a two sensor setup shown in Fig. 1.24 with a scalar plant whose state equation is given by (1.1) with $A = \lambda$. The state X_t is observed in noise by the sensors $\mathscr{E}_1$ and $\mathscr{E}_2$ as

$$Y_t^i = X_t + Z_{e,t}^i, \quad i = 1,2, \tag{1.68}$$

where $Z_{e,t}^1$ and $Z_{e,t}^1$ are two i.i.d. mutually independent 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,

$$S_t^i = f_{i,t}(Y_t^i), \quad i = 1,2, \tag{1.69}$$

subject to the following power constraints:

$$E[(S_t^i)^2] \leq P_i, \qquad i = 1,2. \tag{1.70}$$

Accordingly the remote controller receives

$$R_t^i = S_t^i + Z_{d,t}^i, \quad i = 1,2, \tag{1.71}$$

where $Z_{d,t}^i$, $i = 1,2$, are independent and i.i.d. zero-mean Gaussian with power N_d^i. 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 sink node). 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:

$$J_T = \mathbb{E}\left[\sum_{t=1}^{T} X_t^2\right], \tag{1.72}$$

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 $Z_{d,t}^i$.

In the following we present a non-linear distributed sensing scheme which outperforms the best linear scheme. This scheme was first introduced in [21] for the transmission of a Gaussian source over orthogonal Gaussian channels and was later used in control context in [41].

1.7.1 Sensing Scheme

The non-linear distributed sending and control scheme works as follows. The signals transmitted by the two sensors are given by,

$$S_t^1 = \eta_t Y_t^1, \tag{1.73}$$

$$S_t^2 = \eta_t \left(Y_t^2 - \Delta_t \left\lfloor \frac{Y_t^2}{\Delta_t} \right\rceil \right), \tag{1.74}$$

where $\lfloor \cdot \rceil$ denotes rounding to the nearest integer. The parameter Δ_t controls the length of each period in the periodic sawtooth function. The values $\Delta_{[0,t]}$ are chosen such that the cost function J_T in (1.72) is minimized. The procedure of choosing $\Delta_{[0,t]}$ can be found in [41]. The parameters $\{\eta_t, \Delta_{[0,t]}\}$ are chosen such that the average transmit power constraints (1.70) are met.

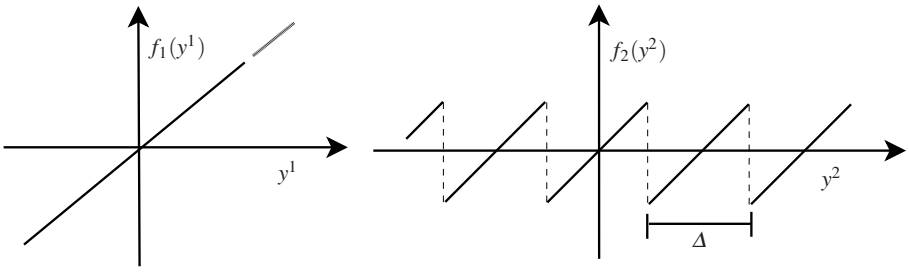


Fig. 1.25 Non-linear distributed sensing.

1.7.2 Control Scheme

The controller is assumed to have a separation structure where it first computes an estimation of the state and then take action using the state estimate. Since the computation of optimal MMSE estimate based on all previously received signals $\left\{R^1_{[0,t]}, R^2_{[0,t]}\right\}$ is not practical, the following sub-optimal algorithm is proposed in [41].

1. Compute estimates $\tilde{X}_{0|t}, \cdots, \tilde{X}_{t|t}$ of $X_0, \cdots, X_t$ based on the previous estimate $\hat{X}_{t-1}$ and R_t^1 using a Kalman filter cf. Kalman Filter 1 in the Fig. 1.26.
2. Assume that $|(\tilde{X}_{s|t} - Y_s^2 - Z_{d,s}^2)/\eta_s| \leq \Delta_s/2 \; \forall s$ and compute the Maximum Likelihood estimates $\hat{Y}_s^2$ as (cf. ML decoder in Fig. 1.26):

$$\hat{Y}_s^2 = \text{argmin}_{Y_s \in \mathscr{Y}}((S^2(Y_s) - R_s^2)^2), \tag{1.75}$$

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

3. Finally assume that the estimates $\hat{Y}_s^2$ had been linearly encoded (multiplied by η_0^t) and find the estimate $\hat{X}_t$ from a Kalman filter using $\{R_{[0,t]}^1, \eta_{[0,t]}, U_{[0,t-1]}\}$ and $U_{[0,t-1]}$ as input cf. Kalman Filter 2 in Fig. 1.26.

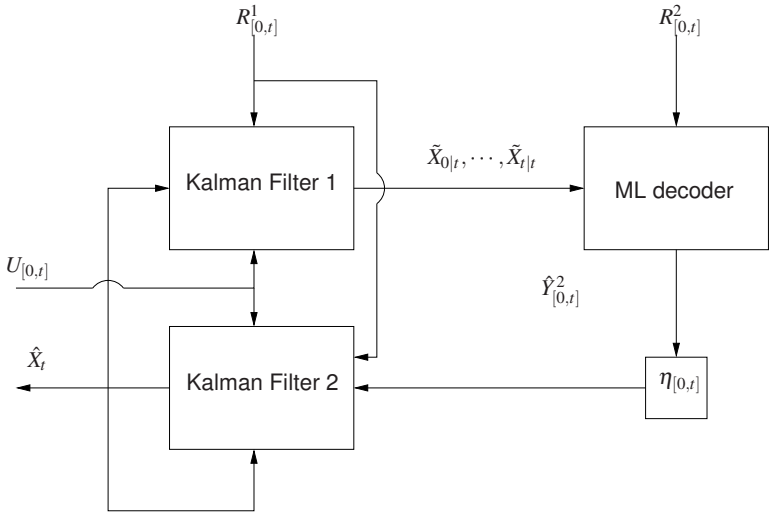
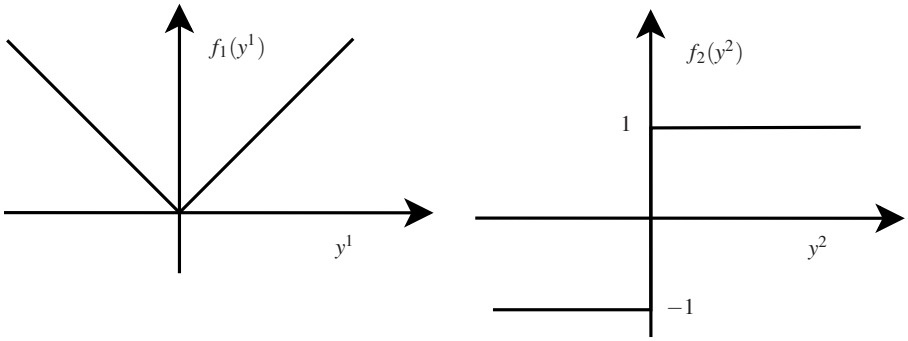


Fig. 1.26 State estimator for the non-linear distributed sensing scheme.

The above non-linear sensing and control scheme is delay-free and can be implemented with reasonable complexity. This non-linear scheme has been shown to outperform the best linear strategy in [41]. Furthermore it is robust to the knowledge of noise statistics at the sensors as demonstrated in [41]. Intuitively, this scheme can be easily extended 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 yet to be studied.

Another non-linear distributed sensing scheme has been proposed in [13] for the two-sensor setup, where one sensor transmits magnitude of the received signal and the other sensor transmits phase value of the received signal. The mappings employed by the two distributed sensors are shown in Fig. 1.27. It has been shown in [13] that the output of these two sensor mappings are mutually independent and thus enable us to send independent information over the two parallel channels. This non-linear scheme has been shown to outperform linear scheme in absence of measurement noise. A careful comparison of these two non-linear sensing schemes discussed in this section is yet to be made. There might be certain regimes where one scheme may perform better than the other.



1st Sensor Mapping. 2nd Sensor Mapping.

Fig. 1.27 Non-linear distributed sensing.

1.8 Bibliographic Notes

There exists a diverse literature on the problem of control and real-time communication over Gaussian channels, focusing on different models, objectives, 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 Gaussian noise. There can be different design constraints such as transmission delay-constraints, sum and individual power constraints, average and peak power constraints, bandwidth constraint etc. And the commonly studied objectives are minimizing a quadratic cost function of the state and the control variables, achieving moment stability or invariant state distribution on the state of the plant. In this chapter, the discussion was mostly limited to the problem of mean-square stabilization of an LTI discrete-time plant and real-time communication over some specific discrete-time white Gaussian channels with average transmit power constraints. In the following we highlight some of the important and related research contributions on the problem of control over Gaussian channels.

Some of the earliest papers addressing the control of linear systems over Gaussian channels include [42, 43]. These papers show 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 in [44], Witsenhausen showed via a simple counter example that linear policies may not be optimal when there are more two or more decision makers (sensors/controllers) without perfect memory (recall). At this point we emphasize the importance of information structures and recommend some fundamental papers on stochastic team decision problems [45–48]. The information structure can be classical, quasiclassical, and non-classical. The problem studied in [42, 43] falls in the class of classical information structure, for which linear control policies were shown to be optimal. In the quasiclassical information structure, decision maker A effects the information of decision maker B, and the decision

maker B knows what is known by decision maker A. For LQG systems having a quasiclassical information structure, linear policies have been shown to be optimal, for example see [45, 47]. The Witsenhausen problem [44] has a non-classical information structure in which decision maker A effects the information of decision maker B, but the decision maker B does not have access to what is known by decision maker A. The Witsenhuasen problem is unsolved till today, which indicates the hardness of such problems. However for some LQG systems with non-classical information structures, linear policies have been shown to be optimal, for example see [34, 35, 49, 50]. These papers have used tools from information theory. The paper [35] studies the problem of a causal memoryless transmission of a noisy Gaussian source over a Gaussian channel and shows that linear coding and decoding policies are optimal. The optimality of linear sensing and control policies for a first order scalar LTI system with an objective of minimizing a quadratic cost function of state and control variables was established in [34], where some concepts from rate distortion theory were used. Another paper [51] used tools from source coding and channel coding to establish necessary condition for stabilization over a large class of communication channel including a memoryless Gaussian channel. The authors of [20] found the conditions under which separation property between estimation and control holds for LQG problems where there is a communication link (for example a memoryless Gaussian channel) between the sensor and the controller. Moreover they introduced a framework of sequential rate distortion theory for designing the encoders and the decoders. In [6, 52–54] the reader can find relevant studies on signal-to-noise ratio requirements for stabilization over some Gaussian channel models. The papers [55, 56] have proposed some techniques for designing linear controllers for Gaussian channels. Some recent results on control over Gaussian fading channels can be found in [57, 58].

The problem of control over communication channels is closely related to the problem of communication over channels with feedback. In [59] a general equivalence was shown between feedback stabilization over an analog communication channel and a communication scheme for channels with noiseless feedback. This communication scheme is a generalization of Schalkwijk-Kailath coding scheme [60] for a single user channel. And for multi-user channels such as broadcast, multiple-access and interference channels, this scheme is a generalization of coding scheme given in [61–63]. Using the communication schemes proposed in [61–63] for multi-user Gaussian channels with noiseless feedback, necessary and sufficient conditions for stabilization of multiple plants over multi-user Gaussian channels are obtained in [64–66]. A decentralized design of linear sensors and controllers over white Gaussian channels with the objective mean square stability is studied in [67]. A comprehensive study of stabilization and optimization of networked control, and information structures is present in [48] and [68].

Acknowledgements

We would like to thank the editors, B. Bernhardsson, G. Como, and A. Rantzer, for giving us an opportunity to write this chapter. We are also very grateful to Johanness Kron (formerly Johanness Karlsson) for performing numerical simulations to generate the figures included in Section 1.6 of this chapter. Some of these results are part of his PhD thesis.

References

1. A. A. Zaidi, T. J. Oechtering, S. Yüksel, and M. Skoglund, "Stabilization of linear systems over Gaussian networks," *IEEE Trans. Autom. Control*, Submitted, June 2012.
2. S. Yüksel, "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.
3. 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.
4. 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.
5. J. L. Massey, "Causality, feedback and directed information," in *IEEE ISITA*, 1990.
6. 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.
7. R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1990.
8. A. A. Zaidi, T. J. Oechtering, S. Yüksel, and M. Skoglund, "Sufficient conditions for closed-loop control over a Gaussian relay channel," in *IEEE ACC*, June 2011, pp. 2240–2245.
9. 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.
10. T. Başar, "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.
11. S. Yüksel and T. Başar, "Control over noisy forward and reverse channels," *IEEE Trans. Autom. Control*, vol. 56, pp. 1014–1029, May 2011.
12. 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.
13. S. Yüksel and S. Tatikonda, "A counterexample in distributed optimal sensing and control," *IEEE Trans. Autom. Control*, vol. 54, no. 4, 2009.
14. U. Kumar, V. Gupta, and J. N. Laneman, "On stability across a Gaussian product channel," in *IEEE CDC*, December 2011, pp. 3142–3147.
15. 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.
16. D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
17. 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.

18. 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.
19. 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.
20. 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.
21. N. Wernersson and M. Skoglund, "Nonlinear coding and estimation for correlated data in wireless sensor networks," *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 2932–2939, October 2009.
22. Y. Wu and S. Verdú, "Functional properties of MMSE," in *IEEE ISIT*, 2010, pp. 1453 – 1457.
23. E. Akyol, K. Viswanatha, and K. Rose, "On conditions for linearity of optimal estimation," *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3497–3508, June 2012.
24. T. Cover and J. Thomas, *Elements of information theory*. John Wiley Sons, Inc., 2006.
25. 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.
26. A. A. Zaidi, T. J. Oechtering, S. Yüksel, and M. Skoglund, "Closed-loop control over half-duplex AWGN relay channels," in *Reglermöte*, June 2010.
27. 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.
28. 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.
29. 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.
30. G. Kramer, I. Maric, and R. Yates, "Cooperative communications," *Foundations and Trends in Networking*, vol. 1, no. 3–4, pp. 271 –425, 2006.
31. A. A. Zaidi, M. N. Khormuji, and M. Skoglund, "Nonlinear transmission strategies for a general half-duplex AWGN relay channel," in *IEEE Swe–CTW*, October 2011, pp. 58–61.
32. 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.
33. ——, "On optimal causal coding of partially observed markov sources in single and multiterminal settings," in *IEEE Trans. Inform. Theory*, vol. 59, Jan. 2013, pp. 424–437.
34. R. Bansal and T. Başar, "Simultaneous design of measurement and control strategies for stochastic systems with feedback," *Automatica*, vol. 25, no. 5, pp. 679–694, 1989.
35. ——, "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.
36. A. A. Zaidi, S. Yüksel, T. J. Oechtering, and M. Skoglund, "On optimal policies for control and estimation over a Gaussian relay channel," in *IEEE CDC*, 2011.
37. G. M. Lipsa and N. C. Martins, "Optimal memoryless control in Gaussian noise: A simple counterexample," *Automatica*, vol. 47, pp. 552–558, March 2011.
38. J. Karlsson and M. Skoglund, "Optimized low-delay source-channel-relay mapping," *IEEE Trans. Commun.*, vol. 58, no. 5, pp. 1397–1404, 2010.
39. N. W. M. Skoglund and T. Ramstad, "Polynomial based analog source-channel codes," *IEEE Trans. Commun.*, vol. 57, no. 9, pp. 2600–2606, Sept. 2009.
40. J. Karlsson and M. Skoglund, "Analog distributed source-channel coding using sinusoids," in *IEEE ISWCS*, Sept. 2009, pp. 279–282.
41. 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.
42. C. Striebel, "Sufficient statistics in the optimum control of stochastic systems," *J. Math. Anal. Appl.*, vol. 12, pp. 576–592, 1962.
43. W. M. Wonham, "On the separation theorem of stochastic control," *SIAM J. Control.*, vol. 6, pp. 312–326, 1968.
44. H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM J. Control*, vol. 6, pp. 131–147, 1968.

45. R. Radner, "Team decision problems," *Annals of Mathematical Statistics*, vol. 33, pp. 857–881, 1962.
46. H. S. Witsenhausen, "Separation of estimation and control for discrete time systems," *Proceedings of the IEEE*, vol. 59, no. 11, pp. 1557–1566, Nov. 1971.
47. Y. C. Ho, "Team decision theory and information structures," *Proceedings of the IEEE*, vol. 68, no. 6, pp. 644–654, 1980.
48. S. Yüksel and T. Başar, *Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints*. Boston, MA: Birkhäuser, 2013.
49. R. Bansal and T. Başar, "Stochastic teams with nonclassical information revisited: When is an affine law optimal?" *IEEE Trans. Automat. Control*, vol. 32, no. 6, pp. 554– 559, June 1987.
50. S. Yüksel, "Stochastic nestedness and the belief sharing information pattern," *IEEE Trans. Automat. Control*, vol. 54, no. 12, pp. 2773–2786, Dec. 2009.
51. S. Tatikonda and S. Mitter, "Control over noisy channels," *IEEE Trans. Autom. Control*, vol. 49, no. 7, pp. 1196–1201, 2004.
52. 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.
53. 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.
54. 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.
55. 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.
56. E. Johannesson, A. Rantzer, and B. Bernhardsson, "Optimal linear control for channels with signal-to-noise ratio constraints," in *ACC*, July 2011, pp. 521–526.
57. 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.
58. A. S. Leong, S. Dey, and J. Anand, "Optimal LQG control over continuous fading channels," in *18th IFAC World Congress*, August 2011.
59. N. Elia, "When Bode meets Shannon: control–oriented feedback communication schemes," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1477–1488, 2004.
60. 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.
61. 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.
62. 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.
63. G. Kramer, "Feedback strategies for white gaussian interference networks," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1423–1438, June 2002.
64. A. A. Zaidi, T. J. Oechtering, and M. Skoglund, "Sufficient conditions for closed-loop control over multiple-access and broadcast channels," in *IEEE CDC*, 2010.
65. ——, "Closed-loop stabilization over Gaussian interference channel," in *18th IFAC World Congress*, August 2011.
66. ——, "On stabilization over a Gaussian interference channel," in *IEEE ECC*, 2013.
67. A. Farhadi and N. U. Ahmed, "Suboptimal decentralized control over noisy communication channels," *Systems & Control Letters*, vol. 60, no. 4, pp. 285–293, April 2011.
68. A. S.Matveev and A. V. Savkin, *Estimation and Control over Communication Networks*. Birkhäuser, Boston, 2008.