TECHNISCHE UNIVERSITÄT MÜNCHEN
Fachgebiet Methoden der Signalverarbeitung

Optimal Control Over Power Constrained Communication Channels

Peter Breun

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr.-Ing. (Univ. Tokio) Martin Buss
Prüfer der Dissertation:
1. Univ.-Prof. Dr.-Ing. Wolfgang Utschick
2. Univ.-Prof. Dr. rer. nat. Michael Ulbrich

Die Dissertation wurde am 18.04.2012 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 20.02.2013 angenommen.

Acknowledgements

This thesis is the result of a process which lasted in my case almost exactly six years. During this time it has been influenced in many ways by the people I met at the Associate Institute of Signal Processing at TU München and I'd like to use the opportunity to thank them at this point. Each of them made an important contribution to the fact that I really enjoyed these years.

I want to express my gratitude to my supervisor Professor Wolfgang Utschick who laid the foundation for my scientific work more than 10 years ago when I first visited one of his lectures on signal processing. From this time on he set an example as a scientist, teacher and friend. I'm happy to have learned so much from him during those years and to have found someone who put trust in me, supported me in every situation and who always had an open ear.

Guido Dietl and Frank Dietrich were great teachers for me when they supervised the work on my Bachelor's and Diploma thesis, respectively, and became colleagues and friends during the time we worked together at the Associate Institute. The same applies to Michael Joham who never was a supervisor for me but always an inexhaustible source for fruitful discussions and support. Johannes Brehmer and Pedro Tejera join this group of people who set the benchmark for my academic work and who motivated me to follow their example. Thanks to all of you.

I cannot imagine how the years in the research group would have looked like without Christian Guthy, Raphael Hunger and David Schmidt who started their work at the Associate Institute almost at the same time as I did. Thank you for giving me the opportunity to play in a team with such a high scientific level, for the coffee brakes, the jokes and all the other things which made working with you a great time. I'm really happy to have made some good friends in you. I also want to include Michael Botsch at this point. Although he was not an official member of the research group, it felt this way and I can only repeat what I have written so far.

The list of persons which made the Associate Institute of Signal Processing a great place to do research is completed by Andreas Dotzler, Lennart Gerdes (thanks for all the support with stochastic signals), Andreas Gründinger (after Frank and Johannes my third office partner), Christoph Hellings, Alexander Krebs and Maximilian Riemensberger. It was a pleasure to work with such an amount of extraordinary people.

I'd like to thank my parents Helga and Walter for their unconditional support which made it possible for me to go to TU München and for showing me that engineering and mathematics is not just something for the others. My sister Sylvia was lead climbing with these topics, demonstrating what is achievable and giving me the incentive to follow her.

Finally, my special thanks go to Birgit who was always at my side during the last years, for listening patiently to all my complaints when something did not work out, to my Hoorays when it did and for reminding me that there is more to life than science.

Munich, September 2013

Contents

1. Introduction

1.1 Communication in Control Systems

For the transmission of information from a source where it is generated to a destination where it is needed, it is a well known fact that the quality of the *channel* which connects source and destination is the limiting factor for the amount of information that can be transmitted in a certain time interval [1]. According to [1], "[t]he channel is merely the medium used to transmit the signal from transmitter to receiver. It may be a pair of wires, a coaxial cable, a band of radio frequencies, a beam of light, etc." From this point of view, it is clear that the presence of a channel makes it necessary to process signals at the input and the output of this channel. First of all, the information must be represented in a way that is suitable for the channel. For example, when transmitting speech over a wireless channel, the acoustic waves have to be transformed into electromagnetic waves. The received signal has to be converted back to an acoustic signal at the destination. The concrete conversion mechanisms are determined by the parameters of the wireless channels, e. g., its bandwidth or frequency response. Additionally, signals can be corrupted during the transmission because "[w]hatever the physical medium used for transmission of the information, the essential feature is that the transmitted signal is corrupted in a random manner by a variety of possible mechanisms, such as additive *thermal noise* generated by electronic devices; man-made noise [...]; and atmospheric noise [...]" (cf. [2, p. 3]). Thus, it may be necessary to recover the transmitted information from the distorted received signal, often it is even mandatory because the signal of interest is completely masked by noise and distortions. These facts are captured in the generic block diagram shown in Figure 1.1 which taken from [1].

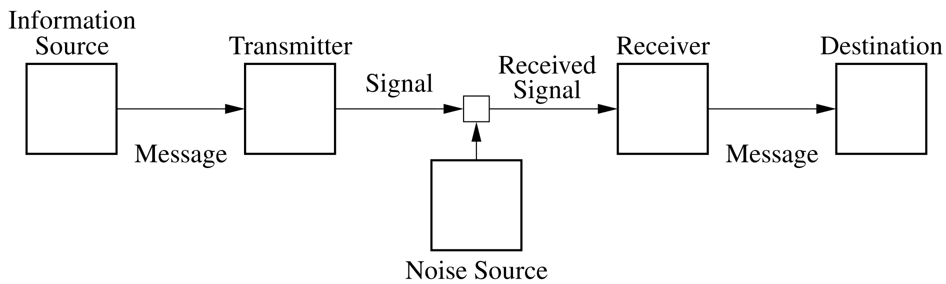


Figure 1.1: Schematic diagram of a general communication system [1].

The problem of information transmission is to design a transmitter-receiver pair such that the message which is generated by the information source can be reconstructed at the destination either perfectly (without an error) or with a distortion which is as small as possible. A major challenge in the design of transmitters and receivers is that, in general, the resources which are available for the transmission of information are limited. It has already been mentioned that the physical medium may offer only a finite bandwidth which restricts the choice of possible signals to band-limited ones. Sometimes the channel can only carry a finite number of signals for the transmission of an uncountable number of messages which makes it necessary to decide which signal represents the message best. Additionally, the power of the signals that amplifiers and electronic devices can handle and generate is finite. Thus, in general, a transmitter has to take a power constraint into account. An even harder restriction might be that the signals are only allowed to have a finite (peak) amplitude. The main point is that due to these limitations, the channel distortions can not be neglected. As an example, assume that a real number x is transmitted over a channel that simply adds an unknown constant n such that the receiver observes the value $x + n$. If it is possible to transmit a real number with arbitrarily large magnitude, a transmitter can amplify x to $x' = \alpha x$ using an arbitrarily large real scalar α. A receiver simply performs the inverse operation to get $x'' = \alpha^{-1} x'$. It is easy to verify that $\lim_{\alpha \to \infty} x'' = x$, i. e., the number x can be reconstructed perfectly. With an amplitude constraint on x', this is not possible and it remains an error of $\alpha^{-1} n$ at the receiver. One can think of more sophisticated transmit and receive schemes, but in general a limitation of communication resources results in imperfect transmission of information.

The term *information* is essential for a communication system, which is obvious as the first block in Figure 1.1 is the information source which generates the messages to be transmitted. For control systems, information also plays a central role. This can be confirmed with almost every textbook concerning (stochastic) control problems (e. g., [3–6]) where one can find discussions about the amount and quality of available information and how it is used for controller design. In order to explain the importance, let us first describe what we understand by a control system (see, e. g., [4, p. 119]). The starting point is a *dynamical system* or *plant* which represents a given physical or technical process. Sensors at the output of the system provide observations about its state[1]. Typically, the dynamical system is desired to show a certain behavior that is different from its intrinsic one. The system itself, i. e., its parameters, structure etc., can not be changed, but it has an input which can be used to manipulate the system in order to achieve the desired dynamical behavior. The output of the system is used to compare the actual with the desired behavior and an appropriate input signal has to be generated if a deviation is observed. The device which maps the observations of the system output to input signals is the *controller* and has to be designed accordingly. This setup is called a *closed loop control*[2] system since a feedback loop between output and input of the plant is created which can easily be identified in Figure 1.2.

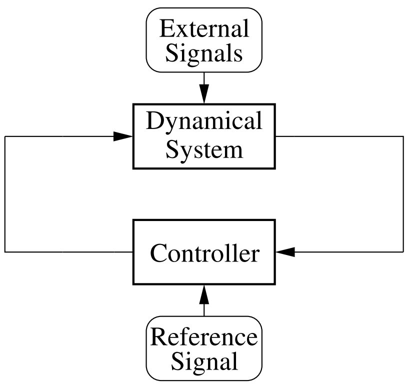


Figure 1.2: Feedback control loop.

[1]At this point, we do not go into detail what the state of a dynamical system is. For the moment, we use it as an abstract term: the state represents the condition of a dynamical system and summarizes all past information that is relevant for the future evolution of the system (see, e. g., [6, p. 2] and [7, pp. 62-63]).

[2]This thesis considers the closed loop case only since open loop control systems can not take into account disturbances or changes of the plant which are not known to the controller (cf. [4, p. 123]). In situations where communication channels are present, such disturbances are inevitable in general.

The feedback loop which is shown in Figure 1.2 also includes *external signals* at the dynamical system and a *reference signal* at the controller. External signals are, e. g., disturbances or input signals generated by the environment of the system, which are not known to and can not be changed by the controller. This is a reason for the introduction of feedback since the controller can compare the desired with the actual behavior of the plant which differs due to the unknown external signals. The reference signal at the controller side is the representation of what has been called "desired behavior" so far, e. g., a given trajectory that should be tracked by the system output signal or the stabilization of an unstable dynamical system.

Using Figure 1.2, it can be seen how and which information flows in a control loop. The controller obtains information about the dynamical system by observing the system output. This is relevant information due to the presence of external signals at the plant which are not known to the controller, i. e., the controller can not determine the system output with the knowledge of the (self generated) system input and the parameters of the dynamical system alone. The information flows from the plant to the controller and is used to determine an appropriate input signal for the plant. In the other direction, i. e., from the controller to the plant, information flows about the desired behavior of the dynamical system, represented by the reference signal and encoded in the control signal which is sent to the system input. In the classical analysis of control systems, the channels which connect the dynamical system and the controller are often assumed to be ideal, i. e., the output of the plant is identical to the input at the controller and the output of the controller is identical to the input of the system, and that all information is available instantaneously without any delay. There are scenarios where this assumption is valid, e. g., when short and shielded wires which offer a high physical bandwidth with low disturbances are used for the connection. But even models used in classical control theory take into account disturbances like those described at the beginning of the section. One example is the measurement or observation noise (cf., e. g., [4, p. 121]) which is present due to the amplifiers that are necessary for the sensing of output signals. In that case, the output of the plant is not identical to the input of the controller but differs by the additive observation noise. Especially when the dynamical system and the controller are spatially separated, their physical connections or channels, as mentioned earlier, introduce additional disturbances which are determined by the concrete model of the channels that are used and the associated limited communication resources.

It can be seen that information is important for a control system, on the one hand information about the actual system behavior (provided by the system output) and on the other hand about the desired behavior (given by a reference signal and the control input). This information has to be exchanged between the dynamical system and the controller. Thus, we identify two sources of information and two destinations which are connected in a feedback loop. Having noticed that the dynamical system and the controller exchange information over communication channels that are, in general, not ideal due to limited communication resources, it is natural to apply the model of the general communication system shown in Figure 1.1 to the two communication links of the feedback control loop depicted in Figure 1.2. The resulting closed loop system is shown in Figure 1.3. In order to distinguish the directions of information flow, the channel which connects the output of the dynamical system with the input of the controller is called the *observation channel*, whereas the channel between the controller and the system input is called the *control channel*. In contrast to Figure 1.1, the channels are depicted as abstract entities because a variety of different channel models can be found in the literature for the present scenario. This point is discussed in more detail in Chapter 2. As a final remark, note that the shaded area in Figure 1.3 indicates that the three included blocks, i. e., receiver, controller and transmitter, can be interpreted as a joint controller-transceiver[3] since all three components can be chosen by the system designer and their connections are assumed to be ideal. Thus, the separation in three distinct blocks is not necessary.

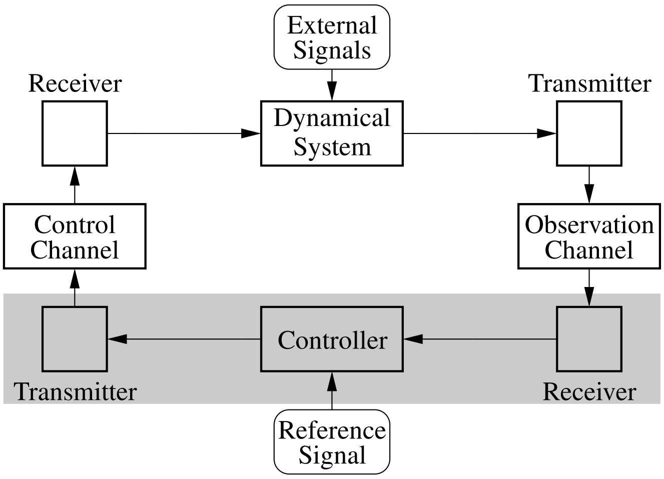


Figure 1.3: Feedback control loop closed over two communication channels.

Since communication channels introduce disturbances, we expect that the performance of a control loop which is closed using such channels is generally worse than that of a loop with ideal connections (cf. Figure 1.2). Nevertheless, the application of transmitters and receivers for the exchange of information offers additional degrees of freedom to reduce the negative effects of the communication channels within the limits that are given by the available communication resources. The task of the system designer is to optimize the transmitters and receivers w.r.t. the performance of the control system. Note that this goal can be different compared to objectives encountered in pure communication systems as shown in Figure 1.1 because there the focus is put on the information itself and not on the purpose it has been transmitted for.

1.2 Networked Control Systems

The control system shown in Figure 1.3 is one of the simplest[4] instances of a so-called Networked Control System (NCS). Following the definition in the guest editorial of [8], a key feature of such systems is that the application of sensors, controllers and actuators is coordinated by "some form of communication network". Typically, these elements are spatially distributed and communicate using wired or wireless links. From this point of view, the term NCS can be applied to a wide range of systems, including sensor networks [9, 10], control over Internet-like data networks [11, 12] or the Internet [13], control of individual systems using a shared communication medium [14–16] or coordination of a group of individual control systems, e. g., a rendezvous in space [17] or formation control of robots [18]. The examples and references given here are far from being exhaustive, but the large interest in and variety of topics in NCSs can be verified with the special issues [8, 19, 20] and the references therein. Note that the expression *network* is used ambiguously. On the one hand, it refers to a wide area network (like the Internet) with advanced communication protocols where

[3]The expression *transceiver* refers to a pair of transmitter and receiver.

[4]Of course, the system model can be further simplified by assuming one of the two channels to be ideal. This would correspond to the case where the controller is either directly located at the input or the output of the dynamical system. Thus, such cases are included in the presented model of the control system.

the control loop uses the network as an application without direct access to the physical parameters of the communication channels that are used. On the other hand, it also refers to a single channel with a direct interaction of the control system and the physical channels. Not surprisingly, the challenges and solution approaches are quite different depending on the network model under consideration. In this thesis we adopt the second point of view in order to explore the interaction of control and communication and the degrees of freedom that are offered by a joint design of controllers, transmitters and receivers.

It is worth noting that some of the effects encountered in the analysis of NCSs are discussed in the literature on control system design for a long time already. One example are so called packet drops (or packet loss, data loss etc.) which describe the effect that in networks information may be lost due to network congestion [21] or because the receiver is not able to reconstruct the transmitted information due to disturbances of the communication channel. The fact that information about the system output or the control input may be lost can be modeled by switching input and output parameters of the dynamical system from a regular mode of operation to a failure where the input or the output is disconnected. The effect of such jump parameters on a control system is studied for more than four decades (see, e. g., [22,23] and references therein). A second example is uncertainty of system parameters which plays an important role when considering physical channel models. The phenomenon of model or parameter uncertainty appears especially in wireless communication systems. Due to the mobility of transmitters, receivers and the environment, the parameters of the channel can change very fast (cf. [2, Chapter 14]) and often it is not possible to predict or estimate these changes accurately. The resulting uncertainty can be interpreted as uncertainty about the parameters of the dynamical system to be controlled. Starting in the 1970s, methods have been developed to describe the effect of model uncertainties on control systems and to design controllers that take into account these uncertainties (e. g., [24, 25], [26] and references therein). As a last example, amplitude constraints on the control signals have been considered since the late 1950s (cf. [27] and references therein) because the effect of saturating actuators can never be avoided in physical control systems. Thus, it is necessary to take into account this kind of non-linearity in the control design procedure in order to guarantee a desired system behavior in the presence of such input constraints.

The fact that uncertainties and constraints that are encountered in NCSs are considered for the design of control systems already for decades does not imply that research on NCSs is not necessary and already covered by the existing results. The feature of NCSs is that the channel and network models provided by information and communication theory define very specific constraints, e. g., on bandwidth, rate, power etc., which have to be taken into account for the optimization of a control system. Thus, the interaction between these constraints and the dynamical behavior of the control system can be explored. Taking this point of view, some fundamental limitations, e. g., on stabilizability of linear systems, have been derived during the last years. The authors of [28] determined the lowest rate[5] that a communication channel which supports only a finite number of input and output signals must provide in order to stabilize an unstable linear dynamical system using feedback control. For the model of additive noise channels, the authors of [29] showed that the stabilization of an unstable linear dynamical system is only possible if the Signal to Noise Ratio (SNR) of the channel is large enough. For the packet drop model and an estimation prob-

[5]The term *rate* is often used in a misleading way in the literature on NCSs. In [28], it has not the general, information theoretic meaning but refers to the number of input (and output) values of a discrete channel that are transmitted instantaneously without an error. This corresponds to the number of quantization levels that are used for the representation of continuous control signals.

lem, the authors of [30] derived upper and lower bounds for the critical probability of a packet loss which leads to an unbounded expected estimation error variance. In all cases, the determined bounds depend solely on the unstable eigenvalues of the system matrix. The main difference of the results is the channel model that is used. Note that for the determination of the mentioned bounds, the authors always considered unstable dynamical systems. The reason is that a stable system can always be left in open loop, i. e., with no control input at all, without the danger of a catastrophic system behavior. In such a case, the minimal requirement for any communication resource is always zero, but the system does not show the desired behavior, i. e., the performance of the control system is poor. For unstable systems, a non-zero control input is mandatory and thus at least some communication resources have to be provided.

1.3 Notation

Throughout the thesis, the following notation is used:

- The set of real numbers is denoted by $\mathbb{R}$, the set of non-negative real numbers by $\mathbb{R}_{+,0}$, and the set of strictly positive real numbers by $\mathbb{R}_{+}$. The set of integer numbers is denoted by $\mathbb{Z}$, of non-negative integers by $\mathbb{N}_0$, and of strictly positive integers by $\mathbb{N}$. Finally, $\mathbb{C}$ is the set of complex numbers.
- Scalars are denoted by upper or lower case letters, vectors by lower case bold letters and matrices by upper case bold letters, e. g., a or A, $\boldsymbol{a}$, and $\boldsymbol{A}$, respectively.
- The constant matrix of dimension $M \times N$ which contains only zeros is denoted by $\mathbf{0}_{M\times N}$ and the corresponding all-zeros vector of dimension M by $\mathbf{0}_M$.
- The identity matrix of dimension $M \times M$ is $\mathbf{I}_M$. The i-th column of this identity matrix is denoted by $\boldsymbol{e}_i^{(M)}$.
- The imaginary unit is denoted by j and Euler's constant by e.
- The operators tr, ${}^{\mathrm{T}}$ and ${}^{\mathrm{H}}$ are the trace, transpose and Hermitian transpose of a matrix (or vector for ${}^{\mathrm{T}}$ and ${}^{\mathrm{H}}$).
- $\mathrm{diag}\,[a_i]_{i=1}^N$ is the diagonal matrix $\sum_{i=1}^N a_i \boldsymbol{e}_i^{(N)} \boldsymbol{e}_i^{(N),\mathrm{T}}$.
- The operator vec stacks the columns of a matrix, i. e., for $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_N] \in \mathbb{R}^{M\times N}$ it holds $\mathrm{vec}\,[\boldsymbol{A}] = \left[\boldsymbol{a}_1^{\mathrm{T}}, \boldsymbol{a}_2^{\mathrm{T}}, \ldots, \boldsymbol{a}_N^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{R}^{MN}$.
- The expectation operator w.r.t. to the distribution of a random variable x is denoted by E_x and the conditional expected value w.r.t. to the conditional distribution of x given the event $\{y = \eta\}$ by $\mathrm{E}_{x|y}\,[\bullet|\,\eta]$, where x and y are jointly distributed random variables (see also Appendix A5.2).
- The expected value of a random vector $\boldsymbol{x}$ is $\boldsymbol{\mu}_{\boldsymbol{x}} = \mathrm{E}_{\boldsymbol{x}}\,[\boldsymbol{x}]$ and its covariance matrix $\boldsymbol{C}_{\boldsymbol{x}} = \mathrm{E}_{\boldsymbol{x}}\left[(\boldsymbol{x}-\boldsymbol{\mu}_{\boldsymbol{x}})(\boldsymbol{x}-\boldsymbol{\mu}_{\boldsymbol{x}})^{\mathrm{T}}\right]$. The correlation matrix is denoted by $\boldsymbol{R}_{\boldsymbol{x}} = \mathrm{E}_{\boldsymbol{x}}\left[\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}\right]$.
- $\mathcal{N}\left(\boldsymbol{\mu}_{\boldsymbol{x}}, \boldsymbol{C}_{\boldsymbol{x}}\right)$ denotes the Gaussian (also called normal) distribution of the real random vector $\boldsymbol{x}$ with expected value $\boldsymbol{\mu}_{\boldsymbol{x}}$ and covariance matrix $\boldsymbol{C}_{\boldsymbol{x}}$.

1.4 Definitions

In this section, we provide standard definitions which are used throughout the thesis.

Definition 1.4.1: Positive (semi)definite matrix
A square matrix $\boldsymbol{A} \in \mathbb{R}^{N \times N}$ is called *positive definite* if it is symmetric, i. e., $\boldsymbol{A} = \boldsymbol{A}^{\mathrm{T}}$, and it holds that

$$\boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x} > 0, \qquad \forall \boldsymbol{x} \in \mathbb{R}^{N}.$$

It is *positive semidefinite* if it is symmetric and it holds that

$$\boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x} \geq 0, \qquad \forall \boldsymbol{x} \in \mathbb{R}^{N}.$$

A necessary and sufficient condition for positive definiteness is that all eigenvalues of $\boldsymbol{A}$ are larger than zero, while for semidefiniteness, all eigenvalues are non-negative. For a positive definite matrix, we use the notation

$$\boldsymbol{A} > \mathbf{0}_{N \times N},$$

whereas a positive semidefinite matrix is denoted by

$$\boldsymbol{A} \geq \mathbf{0}_{N \times N}.$$

Remark: The definition above is restricted to symmetric matrices despite the fact that this is not necessary in general. Since an arbitrary square matrix $\boldsymbol{A} \in \mathbb{R}^{N \times N}$ can be decomposed in its symmetric and anti-symmetric part, i. e.,

$$\boldsymbol{A} = \frac{1}{2}(\boldsymbol{A} + \boldsymbol{A}^{\mathrm{T}}) + \frac{1}{2}(\boldsymbol{A} - \boldsymbol{A}^{\mathrm{T}}),$$

and it holds that $\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{A} - \boldsymbol{A}^{\mathrm{T}})\boldsymbol{x} = 0$, $\forall \boldsymbol{x} \in \mathbb{R}^{N}$, a real matrix is positive (semi)definite if and only if its symmetric part is positive (semi)definite. In the following, we are only interested in the symmetric case, which is the reason for the restriction.

Definition 1.4.2: Matrix inequality
Let $\boldsymbol{A} \in \mathbb{R}^{N \times N}$ and $\boldsymbol{B} \in \mathbb{R}^{N \times N}$ be two symmetric matrices of the same dimension. The matrix inequality

$$\boldsymbol{A} \geq \boldsymbol{B}$$

is equivalent to the condition

$$\boldsymbol{A} - \boldsymbol{B} \geq \mathbf{0}_{N \times N},$$

i. e., that the difference between $\boldsymbol{A}$ and $\boldsymbol{B}$ is positive semidefinite. The definite case is defined analogously.

Definition 1.4.3: Controllability [6, p. 152], [31, Chapter C.3]
Let $\boldsymbol{A} \in \mathbb{R}^{N \times N}$ and $\boldsymbol{B} \in \mathbb{R}^{N \times M}$. The pair $(\boldsymbol{A}, \boldsymbol{B})$ is called controllable if and only if[6] the matrix

$$[\boldsymbol{B}, \boldsymbol{A}\boldsymbol{B}, \boldsymbol{A}^2\boldsymbol{B}, \dots, \boldsymbol{A}^{N-1}\boldsymbol{B}]$$

has full rank N.

Remark: The name *controllability* stems from the interpretation of $\boldsymbol{A}$ and $\boldsymbol{B}$ as the parameters of a controlled dynamical system which is described by the difference equation

$$\boldsymbol{x}_{k+1} = \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{B}\boldsymbol{u}_k, \qquad k \in \mathbb{N}_0.$$

If $(\boldsymbol{A}, \boldsymbol{B})$ is controllable, there exists a sequence of N control inputs $\boldsymbol{u}_0, \boldsymbol{u}_1, \dots, \boldsymbol{u}_{N-1}$ which drives any initial state $\boldsymbol{x}_0$ to an arbitrary state $\boldsymbol{x}_N$ at time index N.

Definition 1.4.4: Observability [6, p. 152], [31, Chapter C.4]
Let $\boldsymbol{A} \in \mathbb{R}^{N \times N}$ and $\boldsymbol{C} \in \mathbb{R}^{M \times N}$. The pair $(\boldsymbol{A}, \boldsymbol{C})$ is called observable if and only if the pair $(\boldsymbol{A}^{\mathrm{T}}, \boldsymbol{C}^{\mathrm{T}})$ is controllable.

Remark: The concept of observability stems from an interpretation of an autonomous dynamical system given by the equations

$$\begin{aligned} \boldsymbol{x}_{k+1} &= \boldsymbol{A}\boldsymbol{x}_k, \\ \boldsymbol{y}_k &= \boldsymbol{C}\boldsymbol{x}_k, \qquad k \in \mathbb{N}_0. \end{aligned}$$

If $(\boldsymbol{A}, \boldsymbol{C})$ is observable, it is possible to determine the initial state $\boldsymbol{x}_0$ exactly from the sequence of state observations $\boldsymbol{y}_0, \boldsymbol{y}_1, \dots, \boldsymbol{y}_{N-1}$.

Definition 1.4.5: Stabilizability [31, Chapter C.3], [6, p. 159]
Let $\boldsymbol{A} \in \mathbb{R}^{N \times N}$ and $\boldsymbol{B} \in \mathbb{R}^{N \times M}$. The pair $(\boldsymbol{A}, \boldsymbol{B})$ is called stabilizable if there exists a matrix $\boldsymbol{L} \in \mathbb{R}^{M \times N}$ such that the eigenvalues of

$$\boldsymbol{A} - \boldsymbol{B}\boldsymbol{L}$$

have magnitude less than 1.

[6] There exist equivalent definitions of controllability, see, e. g., [31, Chapter C.3]

Remark: Stabilizability is a weaker condition than controllability as it only requires the unstable subspace of a dynamical system to be contained in the controllable subspace (cf. [4, p. 461] and [31, p. 763]). This can be illustrated by a system which is stabilizable but not controllable and thus can be represented as (cf. [32, p. 73] and [4, pp. 461-462])

$$\boldsymbol{x}_{k+1} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ & \boldsymbol{A}_{22} \end{bmatrix} \boldsymbol{x}_k + \begin{bmatrix} \boldsymbol{B}_1 \\ \boldsymbol{0}_{U \times M} \end{bmatrix} \boldsymbol{u}_k, \qquad k \in \mathbb{N}_0,$$

with $0 < U < N$ and where the pair $(\boldsymbol{A}_{11}, \boldsymbol{B}_1)$ is controllable. For the system to be stable, the eigenvalues of $\boldsymbol{A}_{22}$ (which represents the uncontrollable subspace) must be less than 1 in magnitude since the corresponding subspace of the state space can not be reached by the control input. The unstable subspace of the system (represented by $\boldsymbol{A}_{11}$ which has eigenvalues with magnitude larger than 1) must be controllable or at least contained in the controllable subspace.

Definition 1.4.6: Detectability [31, Chapter C.4], [6, p. 159]
Let $\boldsymbol{A} \in \mathbb{R}^{N \times N}$ and $\boldsymbol{C} \in \mathbb{R}^{M \times N}$. The pair $(\boldsymbol{A}, \boldsymbol{C})$ is called detectable if the pair $(\boldsymbol{A}^{\mathrm{T}}, \boldsymbol{C}^{\mathrm{T}})$ is stabilizable.

Remark: Analogous to the case of stabilizability, detectability requires the subspace of the system which is not observable to be stable (cf. [4, p. 465]), i. e., the corresponding part of the state vector eventually goes to zero.

Definition 1.4.7: Wide Sense Stationary (WSS) random sequence [33, p. 361]
Let $(\boldsymbol{x}_k : k \in \mathbb{N}_0)$ be a sequence of random vectors. The sequence is called WSS if

$$\mathrm{E}_{\boldsymbol{x}_k}\left[\boldsymbol{x}_k\right] = \boldsymbol{\mu}_{\boldsymbol{x}}, \qquad k \in \mathbb{N}_0$$

and

$$\mathrm{E}_{\boldsymbol{x}_m, \boldsymbol{x}_n}\left[\boldsymbol{x}_m \boldsymbol{x}_n^{\mathrm{T}}\right] = \mathrm{E}_{\boldsymbol{x}_{m+k}, \boldsymbol{x}_{n+k}}\left[\boldsymbol{x}_{m+k} \boldsymbol{x}_{n+k}^{\mathrm{T}}\right], \qquad k \in \mathbb{Z} \text{ and } m, n, m+k, n+k \in \mathbb{N}_0,$$

i. e., if the first and second order moments of the random sequence are shift invariant.

Remark: A WSS random sequence is also called weakly or second-order stationary [33, p. 361].

Definition 1.4.8: Asymptotically WSS random sequence, (cf. [34, p. 392])
Let $(\boldsymbol{x}_k : k \in \mathbb{N}_0)$ be a sequence of random vectors. The sequence is called *asymptotically* WSS if $\mathrm{E}_{\boldsymbol{x}_k}\left[\boldsymbol{x}_k\right]$ and $\mathrm{E}_{\boldsymbol{x}_m, \boldsymbol{x}_n}\left[\boldsymbol{x}_m \boldsymbol{x}_n^{\mathrm{T}}\right]$ exist and are finite for all $m, n, k \in \mathbb{N}_0$ and additionally

$$\lim_{k \to \infty} \mathrm{E}_{\boldsymbol{x}_k}\left[\boldsymbol{x}_k\right] = \boldsymbol{\mu}_{\boldsymbol{x}}$$

and

$$\lim_{k\to\infty} \mathrm{E}_{\boldsymbol{x}_{m+k},\boldsymbol{x}_{n+k}}\left[\boldsymbol{x}_{m+k}\boldsymbol{x}_{n+k}^{\mathrm{T}}\right] = \boldsymbol{R}_{\boldsymbol{x}}(m-n),$$

i. e., if the expected value converges to a constant value and the cross-correlation matrix to a function $\boldsymbol{R}_{\boldsymbol{x}}$ which only depends on the distance $(m-n)$ of the sequence elements $\boldsymbol{x}_{m+k}$ and $\boldsymbol{x}_{n+k}$ for $k \to \infty$.

Remark: Speaking in terms of dynamical systems, an asymptotically WSS random sequence may possess a non-WSS transient phase but eventually reaches a WSS steady state.

Definition 1.4.9: Mean square stability [35]
Consider a dynamical system given by the difference equation

$$\boldsymbol{x}_{k+1} = \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{w}_k, \qquad k \in \mathbb{N}_0,$$

where $\boldsymbol{A}$ is a square matrix, $\boldsymbol{x}_0$ is a random vector with mean $\boldsymbol{\mu}_{\boldsymbol{x}_0}$ and covariance matrix $\boldsymbol{C}_{\boldsymbol{x}_0}$ and $(\boldsymbol{w}_k : k \in \mathbb{N}_0)$ is an (asymptotically) WSS sequence of random vectors. The dynamical system is said to be mean square stable if

$$\lim_{k\to\infty} \mathrm{E}_{\boldsymbol{x}_k}\left[\boldsymbol{x}_k\right] = \boldsymbol{\mu}_{\boldsymbol{x}}$$

and

$$\lim_{k\to\infty} \mathrm{E}_{\boldsymbol{x}_k}\left[\boldsymbol{x}_k\boldsymbol{x}_k^{\mathrm{T}}\right] = \boldsymbol{R}_{\boldsymbol{x}},$$

irrespective of the distribution of the initial state $\boldsymbol{x}_0$, i. e., the first and second order moments of the state vector $\boldsymbol{x}_k$ converge to finite values.

Remark: Note that Definition 1.4.9 is only meaningful for time-invariant systems which are driven by (asymptotically) WSS random sequences. The reason is that even if first and second order moments remain bounded, the limits do not exist for systems with time varying parameters and excitations. In this case, the definition from [36] that a system is mean square stable if

$$\sup_{k\in\mathbb{N}_0} \mathrm{E}_{\boldsymbol{x}_k}\left[\|\boldsymbol{x}_k\|_2^2\right] < \infty$$

is appropriate. It is easy to verify that a system which is mean square stable in the sense of Definition 1.4.9 also fulfills the above condition.

1.5 System Model

Throughout the thesis, we consider discrete-time[7] linear dynamical systems which are described by the state space representation

$$\begin{aligned} \boldsymbol{x}_{k+1} &= \boldsymbol{A}_k\boldsymbol{x}_k + \boldsymbol{B}_k\boldsymbol{u}_k + \boldsymbol{w}_k, \\ \boldsymbol{y}_k &= \boldsymbol{C}_k\boldsymbol{x}_k + \boldsymbol{v}_k, \qquad\qquad k \in \mathbb{N}_0, \end{aligned} \tag{1.1}$$

[7] For the conversion of a continuous-time linear model to discrete-time, see Appendix A1.

where $\boldsymbol{x}_k \in \mathbb{R}^{N_x}$ is the system state, $\boldsymbol{u}_k \in \mathbb{R}^{N_u}$ is the system input and $\boldsymbol{y}_k \in \mathbb{R}^{N_y}$ the observation of the system state, i. e., the output of the dynamical system, at time index k. The parameters of the system, i. e., the system matrix $\boldsymbol{A}_k \in \mathbb{R}^{N_x \times N_x}$, input matrix $\boldsymbol{B}_k \in \mathbb{R}^{N_x \times N_u}$ and output matrix $\boldsymbol{C}_k \in \mathbb{R}^{N_y \times N_x}$, are in general not constant but depend on the index $k \in \mathbb{N}_0$. The time-invariant case when $\boldsymbol{A}_k = \boldsymbol{A}$, $\boldsymbol{B}_k = \boldsymbol{B}$ and $\boldsymbol{C}_k = \boldsymbol{C}$ for all $k \in \mathbb{N}_0$ will be of special importance in the following and we will always assume that the pair $(\boldsymbol{A}, \boldsymbol{B})$ is stabilizable and the pair $(\boldsymbol{A}, \boldsymbol{C})$ is detectable, respectively (see Definitions 1.4.5 and 1.4.6).

The initial state $\boldsymbol{x}_0 \in \mathbb{R}^{N_x}$ is modeled as a Gaussian random vector with $\boldsymbol{x}_0 \sim \mathcal{N}\left(\mathbf{0}_{N_x}, \boldsymbol{C}_{\boldsymbol{x}_0}\right)$. For the sake on simplicity, the process noise $(\boldsymbol{w}_k : k \in \mathbb{N}_0)$ and the observation noise $(\boldsymbol{v}_k : k \in \mathbb{N}_0)$ are assumed to be sequences of independent random vectors which are mutually independent as well as independent of the initial state $\boldsymbol{x}_0$, and their distributions are $\boldsymbol{w}_k \sim \mathcal{N}\left(\mathbf{0}_{N_x}, \boldsymbol{C}_{\boldsymbol{w}_k}\right)$ and $\boldsymbol{v}_k \sim \mathcal{N}\left(\mathbf{0}_{N_y}, \boldsymbol{C}_{\boldsymbol{v}_k}\right)$, respectively, with $k \in \mathbb{N}_0$. When considering time-invariant systems, it will be additionally assumed that the noise sequences are stationary. Together with the independence assumption, this implies that $(\boldsymbol{w}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{v}_k : k \in \mathbb{N}_0)$ with $\boldsymbol{w}_k \sim \mathcal{N}\left(\mathbf{0}_{N_x}, \boldsymbol{C}_{\boldsymbol{w}}\right)$ and $\boldsymbol{v}_k \sim \mathcal{N}\left(\mathbf{0}_{N_y}, \boldsymbol{C}_{\boldsymbol{v}}\right)$ for all $k \in \mathbb{N}_0$ are identically and independently distributed (i.i.d.) sequences of random vectors in the time-invariant case.

Note that the independence assumption of the process and observation noise can be relaxed. If the noise sequences are given by Gauss-Markov sequences, they can be described by a state space model which is driven by a sequence of independent random vectors (cf. Theorem A7.2). Thus, the overall system can be described using a model according to Equation (1.1) by augmenting the model of the original dynamical system with the state space description of the noise sequences (cf., e. g., [37, Chapter 11]). Mutual dependencies of $(\boldsymbol{w}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{v}_k : k \in \mathbb{N}_0)$ can also be included and slightly affect the results presented in Appendix A7. A detailed description of the necessary modifications can be found in, e. g., [38, pp. 69-71] and [31, Chapter 9].

1.6 Channel Model

The channels which are used to exchange information between the dynamical system described above and a controller in a closed loop system are assumed to be linear, memoryless and to introduce additive noise which corrupts the transmitted information. Thus, in the general case, the received vector $\boldsymbol{r}_k \in \mathbb{R}^M$ at time index $k \in \mathbb{N}_0$ at the channel output reads as

$$\boldsymbol{r}_k = \boldsymbol{H}_k \boldsymbol{t}_k + \boldsymbol{n}_k, \tag{1.2}$$

where $\boldsymbol{t}_k \in \mathbb{R}^N$ is the transmitted vector, $\boldsymbol{H}_k \in \mathbb{R}^{M \times N}$ is the matrix which describes the linear channel and $\boldsymbol{n}_k$ is the additive noise described by a sequence of random vectors with properties defined in Section 1.6.2. The sequence of transmit vectors $(\boldsymbol{t}_k : k \in \mathbb{N}_0)$ is a sequence of random vectors where we do not make specific assumptions about its properties at this point despite the existence of finite first and second order moments for all $k \in \mathbb{N}_0$. Figure 1.4 illustrates the channel model.

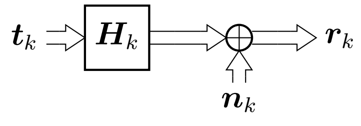


Figure 1.4: General model of a linear memoryless additive noise channel.

Note that the dimensions N and M are not necessarily the same. However, if the channel matrix is known, it can be included in the model of the dynamical system presented in Section 1.5. Assume for example that the observations $\boldsymbol{y}_k$ (cf. Equation 1.1) are transmitted using the channel model of Equation (1.2). An equivalent channel without an explicit channel matrix can be constructed if the system output matrix $\boldsymbol{C}_k$ is replaced by $\boldsymbol{H}_k\boldsymbol{C}_k$ and the observation noise $\boldsymbol{v}_k$ by $\boldsymbol{H}_k\boldsymbol{v}_k$. In this case, the remaining channel is simply represented by the additive noise $\boldsymbol{n}_k$ and the dimensions of the input and the output of this channel are the same.[8]

1.6.1 The Limited Communication Resource

The communication resource which is limited by the specific choice of the communication channel is the power of the transmitted signal.[9] It is given by its variance and limited by the available power $P_{\text{Tx}} \geq 0$ at the transmitter. Thus, in general, the constraint

$$\mathrm{E}_{\boldsymbol{t}_k}\left[\|\boldsymbol{t}_k - \boldsymbol{\mu}_{\boldsymbol{t}_k}\|_2^2\right] = \mathrm{tr}\left[\boldsymbol{C}_{\boldsymbol{t}_k}\right] \leq P_{\text{Tx}}, \qquad k \in \mathbb{N}_0, \tag{1.3}$$

has to be considered. Using this formulation of a communication constraint, it is implicitly assumed that the variance of the additive noise is given and not a design parameter. In the literature on NCSs which considers the presented channel model one can often find an equivalent formulation of the constraint presented in Equation (1.3) which explicitly includes the variance of the channel noise. In this case, the SNR, i. e., the ratio of the power of the transmitted signal and the power of the channel noise, is the limiting factor and assumed to be bounded by some constant $\varphi \geq 0$. Thus, Equation (1.3) becomes

$$\frac{\mathrm{tr}\left[\boldsymbol{C}_{\boldsymbol{t}_k}\right]}{\mathrm{tr}\left[\boldsymbol{C}_{\boldsymbol{n}_k}\right]} \leq \varphi, \qquad k \in \mathbb{N}_0. \tag{1.4}$$

Equation (1.4) is obtained by dividing Equation (1.3) by $\mathrm{tr}\left[\boldsymbol{C}_{\boldsymbol{n}_k}\right]$, i. e., the variance of the channel noise. This shows the equivalence of both representations of the limited communication resource by noting that $P_{\text{Tx}} = \varphi\,\mathrm{tr}\left[\boldsymbol{C}_{\boldsymbol{n}_k}\right]$.

1.6.2 Additive (White) Gaussian Noise

For this channel model only the additive noise sequence $(\boldsymbol{n}_k : k \in \mathbb{N}_0)$ is considered. Thus, Equation (1.2) reduces to

$$\boldsymbol{r}_k = \boldsymbol{t}_k + \boldsymbol{n}_k, \qquad k \in \mathbb{N}_0, \tag{1.5}$$

with transmitted vector $\boldsymbol{t}_k \in \mathbb{R}^M$ and received vector $\boldsymbol{r}_k \in \mathbb{R}^M$. The noise sequence $(\boldsymbol{n}_k : k \in \mathbb{N}_0)$ is assumed to be uncorrelated (white) and Gaussian with $\boldsymbol{n}_k \sim \mathcal{N}\left(\boldsymbol{0}_M, \boldsymbol{C}_{\boldsymbol{n}_k}\right)$ for $k \in \mathbb{N}_0$. Additionally, it is independent of all other random variables which contribute to the sequence of transmit vectors $(\boldsymbol{t}_k : k \in \mathbb{N}_0)$. Due to its properties, the sequence $(\boldsymbol{n}_k : k \in \mathbb{N}_0)$ is referred to as Additive White Gaussian Noise (AWGN).

[8] Similar considerations lead to the treatment of channels with memory and temporarily correlated noise sequences by including their descriptions in the system model given by Equation (1.1).

[9] If the power is not limited, an infinitely large amplification of the transmit signal together with the reverse operation at the receiver can be used to eliminate the additive channel noise (see Section 1.1). Thus, the consideration of a non-zero noise always implies the limitation of the transmit power.

Remark: The consideration of the AWGN case in the context of information exchange of dynamical systems is not as restrictive as it may seem and can be extended to the case of Gauss-Markov noise sequences $(\boldsymbol{n}_k : k \in \mathbb{N}_0)$ which allow for a state space representation (cf. Theorem A7.2)[10]

$$\begin{aligned} \boldsymbol{x}_{k+1}^{(n)} &= \boldsymbol{A}_k^{(n)} \boldsymbol{x}_k^{(n)} + \boldsymbol{B}_k^{(n)} \boldsymbol{\eta}_k, \\ \boldsymbol{n}_k &= \boldsymbol{C}_k^{(n)} \boldsymbol{x}_k^{(n)} + \boldsymbol{D}_k^{(n)} \boldsymbol{\eta}_k, \qquad k \in \mathbb{N}_0, \end{aligned} \tag{1.6}$$

with system state $\boldsymbol{x}_k^{(n)} \in \mathbb{R}^K$, an uncorrelated (white) sequence $(\boldsymbol{\eta}_k : k \in \mathbb{N}_0)$ of zero-mean Gaussian random vectors, i. e., $\boldsymbol{\eta}_k \sim \mathcal{N}\left(\mathbf{0}_L, \boldsymbol{C}_{\boldsymbol{\eta}_k}\right)$, and system parameters $\boldsymbol{A}_k^{(n)}$, $\boldsymbol{B}_k^{(n)}$, $\boldsymbol{C}_k^{(n)}$ and $\boldsymbol{D}_k^{(n)}$ of appropriate dimensions. Additionally, the random vectors $\boldsymbol{\eta}_k$, $k \in \mathbb{N}_0$, are independent of all other random variables under consideration. As an example, assume that the output $\boldsymbol{y}_k$ of the dynamical system introduced in Equation (1.1) is transmitted over the additive noise channel given by Equation (1.5), i. e., $\boldsymbol{t}_k = \boldsymbol{y}_k$, with correlated noise $\boldsymbol{n}_k$ according to Equation (1.6). The output $\boldsymbol{r}_k$ of the channel reads as

$$\begin{aligned} \boldsymbol{r}_k = \boldsymbol{y}_k + \boldsymbol{n}_k &= \boldsymbol{C}_k \boldsymbol{x}_k + \boldsymbol{v}_k + \boldsymbol{C}_k^{(n)} \boldsymbol{x}_k^{(n)} + \boldsymbol{D}_k^{(n)} \boldsymbol{\eta}_k \\ &= \begin{bmatrix} \boldsymbol{C}_k, & \boldsymbol{C}_k^{(n)}, & \boldsymbol{D}_k^{(n)} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_k \\ \boldsymbol{x}_k^{(n)} \\ \boldsymbol{\eta}_k \end{bmatrix} + \boldsymbol{v}_k. \end{aligned} \tag{1.7}$$

This corresponds to the noisy observation of the state $\begin{bmatrix} \boldsymbol{x}_k^{\mathrm{T}}, & \boldsymbol{x}_k^{(n),\mathrm{T}}, & \boldsymbol{\eta}_k^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$ of a system given by the difference equation

$$\begin{bmatrix} \boldsymbol{x}_{k+1} \\ \boldsymbol{x}_{k+1}^{(n)} \\ \boldsymbol{\eta}_{k+1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_k & & \\ & \boldsymbol{A}_k^{(n)} & \boldsymbol{B}_k^{(n)} \\ & & \mathbf{0}_{L \times L} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_k \\ \boldsymbol{x}_k^{(n)} \\ \boldsymbol{\eta}_k \end{bmatrix} + \begin{bmatrix} \boldsymbol{B}_k \\ \mathbf{0}_{K \times N_u} \\ \mathbf{0}_{L \times N_u} \end{bmatrix} \boldsymbol{u}_k + \begin{bmatrix} \boldsymbol{w}_k \\ \mathbf{0}_K \\ \boldsymbol{\eta}_{k+1} \end{bmatrix}. \tag{1.8}$$

Recalling the assumption that the sequences $(\boldsymbol{w}_k : k \in \mathbb{N}_0)$, $(\boldsymbol{\eta}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{v}_k : k \in \mathbb{N}_0)$ are white and mutually independent, it can be seen that the correlated noise case can be included in the AWGN channel model. Similar steps are possible if the correlated noise is added to the system input $\boldsymbol{u}_k$. Thus, only the AWGN case will be considered in the following.

Finally, Figure 1.5 illustrates the application of the AWGN channel model to a control system. It shows a control loop which is closed using two additive noise channels which introduce the disturbances $(\boldsymbol{q}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{n}_k : k \in \mathbb{N}_0)$ and separate the controller from the dynamical system to be controlled (cf. Section 1.5). Comparing with Figure 1.3, we identify the observation and control channel and recognize that the transmitters and receivers for the respective channels are not present in Figure 1.5. Their design subject to the limitation of communication resources (cf. Section 1.6.1) will be discussed in the following chapters.

[10] In Theorem A7.2 the feed-through parameter $\boldsymbol{D}_k$ has not been introduced but can be included in the presented stochastic model.

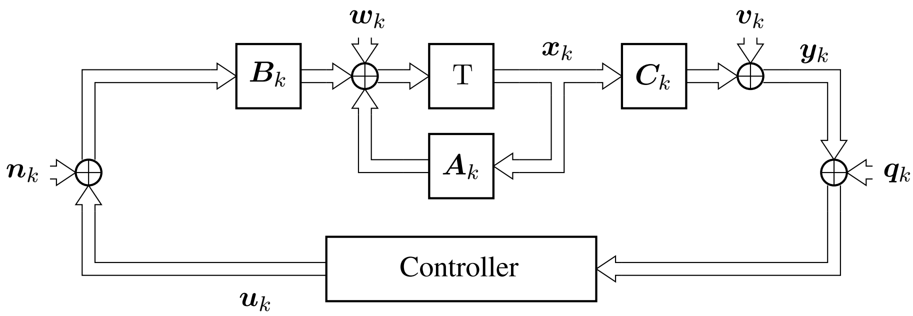


Figure 1.5: Model of the control loop which is closed over two additive noise channels.

2. Challenges in Control Problems Due to Communication Constraints

In the preceding chapter, it has been explained what we understand by a control system and how communication systems and the associated constraints extend the problem of control system design. At this point, we give some concrete examples of limitations and disturbances due to the presence of communication channels in a control loop. The examples cover the most important channel models that are used in the literature on NCSs. Note that for the following examples, we will make use of specific system and channel models that have not been introduced so far, provide solutions without a detailed derivation and sometimes anticipate results from subsequent sections. Nevertheless, the reader who is familiar with stochastic systems and basics in communication theory should be able to follow the explanations and may obtain some intuition about the problems that arise when considering limited communication resources in a control loop.

2.1 Models of the Communication System

2.1.1 Distributed Control Systems

What has not been mentioned so far are the implications of the distributed nature of NCSs (cf. Section 1.2). Having a look at Figure 1.3 and using a general approach, the joint design of transmitters, receivers and the controller can be viewed as a controller design with a structural constraint. In this case, the controller has a distributed configuration where each element has access to different information. One element (in communication terms: the transmitter) has direct access to the system output, one (the receiver) has access to the system input and a third element (the traditional controller) connects the other two using the observation and the control channel. Thus, the different parts of the controller generate the information which is the input to the subsequent parts.

The configuration that defines what information is available at which component of a distributed controller is called *information pattern* [39]. The *classical* information pattern assumes that all components of the controller have access to the same information at the same time. This is of course true if a centralized controller which collects all available information in the feedback loop is used. It is easy to verify that this is not the case for the distributed control system shown in Figure 1.3, i. e., transmitter, receiver and controller have access to different information. The author of [39] gave a very simple example of a control problem with a linear dynamical system and quadratic cost function, but with distributed structure, i. e., with a *non-classical* information pattern. He showed that the optimal control problem is not convex in the optimization variables and that the optimal control scheme must be non-linear.[1] However, with a classical information pattern, the problem is convex and allows for a linear solution in the framework of Linear Quadratic Gaus-

[1] It may be interesting for the reader with a background on communication systems that in [39] it has been pointed out that "[w]hen communications problems are considered as control problems (which they are), the information pattern is never classical since at least two stations, not having access to the same data, are always involved." Thus, the well known fact is stated that optimal communication strategies are in general non-linear and the associated optimization problems are non-convex.

sian (LQG) control. In [40], this example is revisited and presented in a more information theoretic context which provides some additional intuition about the results. As a final remark, the authors of [41] analyzed the property of (non-)convexity in decentralized control and provided the largest known class of convex problems for this case.

The results presented here demonstrate that we can not expect in general to find straightforward solutions to the problem of joint optimization of transmitters, receivers and controllers for a NCS, e. g., by applying convex optimization techniques, and suboptimal approaches may be necessary.

2.1.2 Packet Drops

The packet drop channel model is quite popular in the literature on NCSs for several reasons. First of all, it can be motivated information theoretically [42] as well as from a practical point of view since a lot of general purpose communication systems and data networks are packet switched , i. e., the network can either deliver a packet with all its information or lose the contained information if a packet can not delivered or correctly decoded.[2] A second reason is that such a channel abstracts the physical properties of the communication channel and the question how to allocate the limited communication resources in an optimal way is not part of the control system design. With the packet drop model, the obvious goal of the communication system is to provide a probability of packet loss which is as small as possible. Thus, the design of the communication and the control system is decoupled. Note that despite the fact that the packet drop model typically is the result of a digital, packet switched communication system, the effect of quantization of the transmitted data is neglected in most cases. The standard argument is that a packet can carry such an amount of information that the effect of quantization is extremely small or can be modeled as additive white noise (cf. [44–46]). For a discussion of quantization in NCSs, see Section 2.1.3.

In order the get an intuition why the presence of packet drops has to be taken into account for the design of a control system, consider the following example from [47] with a dynamical system where the input signal can be lost due to packet drops, i. e., without a packet drop the system operates as a closed loop while the loss of the input signal results in an open loop system. We will use the function $\delta(k)$ to describe if the channel loses a packet. It has the value 0 when a packet drop occurs and 1 otherwise. Assume that the dynamical system is linear and described by the difference equation

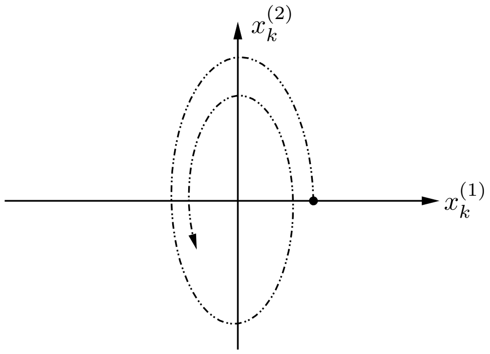
$$\boldsymbol{x}_{k+1} = \boldsymbol{A}_{\delta(k)}\boldsymbol{x}_k, \quad k \in \mathbb{N}_0 \text{ and } \delta(k) \in \{0,1\}, \tag{2.1}$$

where $\boldsymbol{x}_k \in \mathbb{R}^{N_x}$ is the state of the dynamical system at time index k with the initial state $\boldsymbol{x}_0$. The matrix $\boldsymbol{A}_i \in \mathbb{R}^{N_x \times N_x}$, $i \in \{0,1\}$, determines the dynamics of the system where $\boldsymbol{A}_0$ refers to the open loop dynamics when the control input is lost and $\boldsymbol{A}_1$ to the closed loop dynamics when the packet which contains the control signal has been delivered. We can see that due to the packet drops the dynamical system becomes time variant and switches between two possible operation modes. Assume that in both cases, i. e., if the system is operating in open or closed loop mode exclusively, the system is stable in the sense that

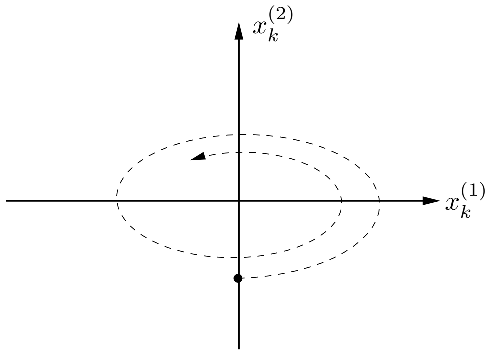
$$\lim_{k\to\infty} \boldsymbol{x}_k = \lim_{k\to\infty} \boldsymbol{A}_i^k \boldsymbol{x}_0 = \boldsymbol{0}_{N_x}, \quad i \in \{0,1\}, \tag{2.2}$$

for all $\boldsymbol{x}_0 \in \mathbb{R}^{N_x}$. An example for such a case is depicted in Figure 2.1 for $N_x = 2$, where $x_k^{(1)}$ and $x_k^{(2)}$ denote the first and the second component of the state vector $\boldsymbol{x}_k$, respectively. It is relatively

[2] For an information theoretic discussion of the infinite alphabet erasure channel see [43, pp. 318-320].



(a) Open loop dynamics with $\boldsymbol{A}_0$.



(b) Closed loop dynamics with $\boldsymbol{A}_1$.

Figure 2.1: Trajectories of a system which is open loop (left) and closed loop (right) stable.

easy to determine a sequence of packet drops, i. e., a switching between the open and the closed loop dynamics, which produces an unstable behavior of the control system. The result is shown in Figure 2.2.

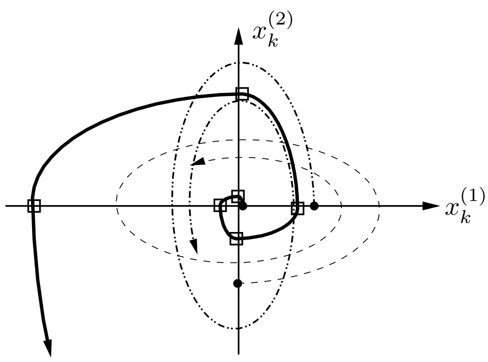


Figure 2.2: Trajectory of a switched system which is open loop (dashed line) and closed loop (dash-dotted line) stable. The system can become unstable by switching (switching points are denoted by □).

The presence of packet drops as a result of limited communication resources of course has negative effects on the performance since the system is operating in open loop for a fraction of the time and thus its behavior differs from the desired one which is realized by closed loop control. But the example shows that the performance can become arbitrarily bad because of the loss of stability. This effect must be taken into account for the design of a controller. Dynamical systems with abruptly changing parameters have been studied in, e. g., in [22, 23, 47, 48], under different aspects like stability and controllability, where the change of the system parameters can be deterministic or driven by a stochastic process. Especially the latter point of view is an interesting starting point for the investigation of NCSs using the packet drop channel model. In this case, the loss of information in a control loop can be modeled as a stochastic change of system parameters between two possible modes, depending on the event of a packet drop. The problem of state estimation with probabilistic loss of measurements of the system output has been investigated in, e. g., [30, 49–51]. The closed loop control problem has been addressed in [11, 45, 52–58], which demonstrates the popularity of the packet drop channel model for the investigation of NCSs.

Nevertheless, in this thesis we do not focus on this channel model because it does not allow for a deeper analysis of the parameters of the physical communication channels and their impact on the control system.

As a final remark, the model of a channel which carries a real number without introducing disturbances but with a certain probability of data loss is a good example why concepts from information theory which solely concentrate on the problem of reliable transmission of information may be inadequate for control system design. It has been pointed out in [42] and is proved[3] in [43, pp. 318-320] that the packet drop channel with infinite input alphabet has infinite Shannon capacity, i. e., it can be used to reliably transmit any amount of information. Nevertheless, this property is not sufficient for the feedback stabilization over packet drop channels (see, e. g., [56]). For a detailed discussion of the properties of communication channels which are compatible with the objective of closed loop control, see [42].

2.1.3 Quantization

A digital communication system offers a finite word length for the data that is transmitted. Typically, this is expressed by the number of bits that can be sent and received reliably per time unit. Thus, it is necessary to evaluate the effect of quantization of the transmitted information in the control system which is generally represented by real numbers. It has already been mentioned in Section 2.1.2 that a possible way of doing this is to simply ignore the effects of quantization or to use an additive noise model like in, e. g., [44, 46, 59]. A different line of research focuses explicitly (and often exclusively) on the effect of quantization as the main representation of limited communication resources. For example, the authors of [36, 60–65] consider the case were a communication channel supports a finite (sometimes relaxed to countable) set of messages which can be transmitted without an error.[4] This introduces the necessity of quantizers which determine a discrete representation of the generally continuous information like sensor measurements and control signals with a finite number of bits, which is the limited resource in this scenario.

One of the major challenges in the analysis and design of NCSs with error-free but bit rate limited channels is the non-linearity of the quantizers which change the closed loop dynamics of the control system significantly and can even lead to a chaotic behavior (cf. [60] and references therein). This complicates the analysis of quantized control systems and the design of optimal quantizers. A second problem occurs due to the boundedness of signals which have to be represented by a finite number of bits. In [66] a linear dynamical system which is driven by Gaussian noise is considered. The control loop is closed over discrete channels with finite input and output alphabets and it is assumed that the control loop which consists of dynamical system, channel encoders and decoders, channels and controller is an irreducible Markov chain. In [66, Theorem 4.2] it is shown that for this scenario, a bounded control input leads to a transient behavior of the Markov chain, which is in conflict with notions of stochastic stability (cf. [67, Section 1.3.1]), e. g., stability in the mean square sense. A possibility to overcome this problem is a zooming type of quantizer which dynamically adapts the range of values that can be quantized [61, 65], i. e., the same number of quantization points is distributed over a large range when large values have to be quantized and over a small range for small values. This means that large values, e. g., of the state or the control input, tolerate a larger quantization error than small ones. This effect has also been observed in [62] where a quantizer is optimized such that a linear dynamical system with scalar

[3] The author of [43] considers an erasure channel with a countably infinite alphabet and it is shown that its capacity is infinitely large. Since this alphabet is a subset of the real numbers, the capacity of an erasure channel with real input is also infinitely large.

[4] This means that the message at the input of the channel is identical to the channel output. A positive effect of this model is that the transmitter has exact knowledge about the information at the receiver which often simplifies the analysis of such systems.

control input is quadratically[5] stabilized. The authors of [62] assume a countable instead of finite number of values for the representation of the control signal and show that the coarsest quantizer for the control signal is logarithmic, i. e., the control input at time index $k \in \mathbb{N}_0$ can represented as

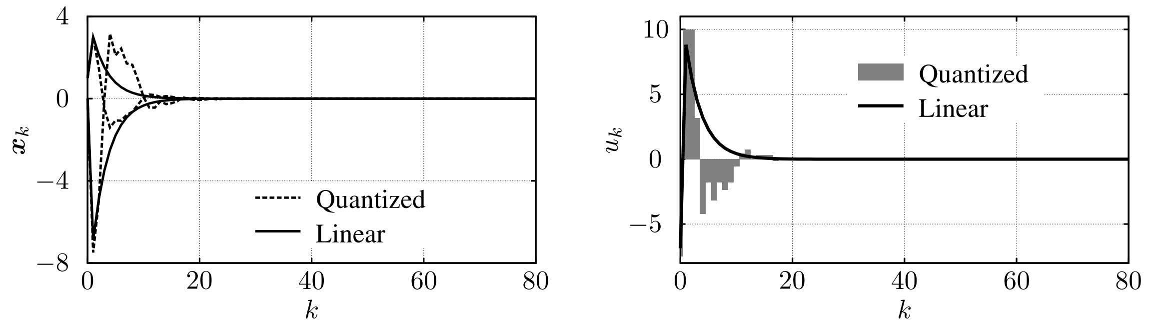
$$u_k = \pm\alpha\rho^i, \quad i \in \mathbb{Z}, \tag{2.3}$$

with the constants $\alpha > 0$ and $0 < \rho < 1$. An important result of [62] is that ρ can not be chosen arbitrarily close to zero but has to be larger than a value that depends on the unstable eigenvalues of the system matrix. The larger the magnitude of these eigenvalues, i. e., for a higher degree of instability, the larger ρ must be which results in a smaller distance between the quantization points. Note that irrespective of the size of ρ, a countably infinite set of values is always available for the quantization. Obviously, this fact alone is not sufficient for a logarithmically quantized controller to stabilize an unstable dynamical system.

The following example illustrates the result of [62]. The model of the dynamical system to be controlled is given by the linear difference equation

$$\boldsymbol{x}_{k+1} = \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{b}u_k, \quad k \in \mathbb{N}_0, \tag{2.4}$$

with the system state $\boldsymbol{x}_k \in \mathbb{R}^{N_x}$, system matrix $\boldsymbol{A} \in \mathbb{R}^{N_x \times N_x}$, system input vector $\boldsymbol{b} \in \mathbb{R}^{N_x}$ and initial state $\boldsymbol{x}_0 \in \mathbb{R}^{N_x}$. For the determination of the control input u_k according to Equation (2.3), the complete knowledge of the state vector $\boldsymbol{x}_k$ is available. For the example, $N_x = 2$ has been chosen and for the applied system matrix[6] we obtain the smallest value of ρ that ensures quadratic stabilizability to be $\rho_{\text{min}} \approx 0.714$. In Figure 2.3, the basis for the logarithmic quantizer is $\rho = 0.75$ which implies quadratic stability. Consequently, it can be observed that the closed loop system with quantized control reaches the zero state in approximately the same number of steps as the one with unquantized control.



(a) System states. (b) Control input.

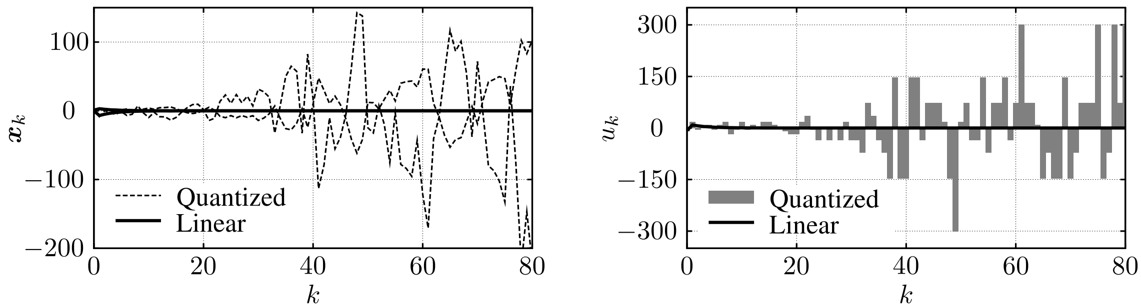
Figure 2.3: System with logarithmically quantized control input. Applied quantizer basis is $\rho = 0.75$, coarsest basis is $\rho_{\text{min}} \approx 0.714$.

The situation changes if the basis is chosen too small. The result for $\rho = 0.49$ is shown in Figure 2.4. Despite the fact that a countably infinite set of quantized values with unrestricted magnitude is available, the closed loop system is not stabilized by the quantized controller.

The purpose of this section is to demonstrate the difficulties that arise in the design of control systems with quantized information. Note a standard assumption for the investigation of control

[5]This stability criterion is also referred to as Lyapunov stability, see, e. g., [7, pp. 177- 180].

[6]The exact system parameters are $\boldsymbol{A} = \begin{bmatrix} 3 & 1 \\ & 2 \end{bmatrix}$ and $\boldsymbol{b}^{\mathrm{T}} = [0, 1]^{\mathrm{T}}$, respectively. The initial state is $\boldsymbol{x}_0^{\mathrm{T}} = [1, 0]^{\mathrm{T}}$.



(a) System states. (b) Control input.

Figure 2.4: System with logarithmically quantized control input. Applied quantizer basis is $\rho = 0.49$, coarsest basis is $\rho_{\text{min}} \approx 0.714$.

systems that include quantization is the absence of any other disturbance due to limited communication resources. One possible implication of this assumption is the necessity of a very reliable communication channel which might be realized by the application of a powerful channel code. The situation becomes much more involved when the possibility of errors is considered, i. e., that a transmitted value is mapped to a wrong value at the receiver. One approach for this scenario can be found in [66].

2.1.4 Model Uncertainty

It has already been mentioned that limited communication resources lead to an imperfect exchange of information, e. g., due to noise, data loss or quantization. In this case, the imperfections of the communication channel directly affect the transmitted signals which carry the information. But this is not the only consequence of imperfect communication. A second problem concerns the knowledge of the parameters of the channels. In order to design a controller with satisfactory performance, it is essential to have a precise model of the dynamical system to be controlled as well as of the channels which connect the system and the controller. It is, for example, important that the control signal at the output of the control channel (i. e., the input of the dynamical system, cf. Figure 1.3) has the value which leads to the desired system behavior. Thus, the signal at the channel input, which is subject to distortions and disturbances, must be chosen accordingly. Without accurate knowledge of the parameters of the channel model, this is not possible. Imperfect channel knowledge is equivalent to uncertainties about the system dynamics, but this uncertainty enters the control system at very specific points and can be modeled according to the channel model under consideration.

The assumption that the model of the dynamical system and its parameters are perfectly known for the controller design can be justified if it is a technical system which is completely determined[7] by a system designer or if there are enough resources like time, bandwidth, power etc. to apply identification techniques (cf., e. g., [68]) for the offline determination of the parameters of a physical system. For the latter case, we have to assume that the system parameters remain constant for a sufficiently long time interval or that it is always possible to repeat the system identification when the parameters change. But these two points are critical for the identification (or estimation) of communication channels. Especially in wireless scenarios, parameters like the channel gain change continuously over time and have to be identified repeatedly using channel estimation

[7]An example is the Internet where TCP is used to control the congestion of the network, see, e. g., [21].

techniques (see, e. g., [69]). The amount of communication resources which is available for the estimation procedure is a limiting factor for the accuracy of the estimate. Assume a model of a communication channel with additive noise and an unknown channel gain such that a transmitted real number x results in a channel output $y = ax + n$, where a is the gain factor and n the noise. The value of a can be estimated by transmitting a so-called training symbol, i. e., a value of x that is known to the transmitter and the receiver, and using the received value for the determination of the unknown constant, e. g., by computing $a_{\text{estimate}} = x^{-1}y = a + x^{-1}n$. Note that with finite communication resources it is in general not possible to obtain a perfect estimate of a. For example, with a constraint on the magnitude of x, there is a residual error in the estimate due to the additive noise. This problem can be circumvented if other resources are not limited. With a reasonable (stochastic) model of the noise, the error can be made arbitrarily small by transmitting an infinitely long sequence of known symbols over the channel and applying estimation methods like least squares, maximum likelihood etc. (cf. [31, Chapters 2 and 3] and [70, Chapters 6-9]). The drawback of this approach is that it trades the limited magnitude of transmit signals with an unlimited amount of time or bandwidth, respectively, in order to transmit an infinitely long sequence of training symbols. But even if we could determine the value of a exactly, this information would only be useful when a remained constant after the training period. Especially in wireless scenarios, the parameters of a communication channel can change fast and it is very likely that parameters that have been determined with training are not the same when transmitting relevant information.

In order to demonstrate why the accurate determination of system parameters is essential for the design of control systems, we show the result of [25] which is called the *uncertainty threshold principle* and considers a dynamical system with imperfect knowledge of its model parameters. The time-variant linear system is described by the scalar difference equation

$$x_{k+1} = a_k x_k + b_k u_k, \qquad k \in \mathbb{N}_0, \tag{2.5}$$

where x_k is the system state at time index k with initial state x_0 and u_k is the control input. In order to model the effect of uncertainty, the parameters a_k and b_k, $k \in \mathbb{N}_0$, are assumed to be given by i.i.d. random sequences which are mutually independent[8]. In more detail, the model for the uncertainty is $a_k \sim \mathcal{N}(\mu_a, c_a)$ and $b_k \sim \mathcal{N}(\mu_b, c_b)$ for all k, where the mean is interpreted as the actual knowledge (or estimate) of the respective parameter while the variance corresponds to the degree of uncertainty, i. e., a parameter is perfectly known if its corresponding variance is zero. If the controller has access to the state x_k at time index k, the optimal control input for the system which minimizes a quadratic cost function reads as (see [25] and [38, pp. 42-53 and Section III.3])

$$u_k = -\left(g_{k+1}\left(c_b + \mu_b^2\right) + r\right)^{-1} g_{k+1}\mu_a\mu_b x_k, \tag{2.6}$$

where g_k, $k \in \{0, 1, \ldots, N\}$, is given by the backward recursion

$$g_k = g_{k+1}\left(c_a + \mu_a^2\right) - g_{k+1}^2\mu_a^2\mu_b^2\left(g_{k+1}\left(c_b + \mu_b^2\right) + r\right)^{-1} + q, \tag{2.7}$$

with initial condition $g_N = 0$. The numbers $N \in \mathbb{N}$, $q > 0$ and $r > 0$ are parameters of the respective optimization problem. We refer to Appendix A6 for a detailed description of the cost function and the meaning of its parameters. The optimal value of the cost function that is achieved by applying the result of Equation (2.6) is given by

$$J^* = g_0 x_0^2. \tag{2.8}$$

[8]This assumption has not been made in [25] where correlations between a_k and b_k are considered. We omit these correlations here for the sake of simplicity.

The value of N determines the number of steps or the *horizon* of the control problem. If the control system operates over a very long horizon, i. e., in the limit for $N \to \infty$, it is of special interest if the recursion in Equation (2.7) converges and does not grow without a limit which would result in an unbounded optimal value of the cost function (cf. Equation 2.8). In [25] it has been shown[9] that the cost function is only bounded if

$$m = c_a + \mu_a^2 - \mu_a^2 \mu_b^2 \left(c_b + \mu_b^2\right)^{-1} < 1. \tag{2.9}$$

Thus, even if the dynamical system is, loosely speaking, stable on average, i. e., if $|\mu_a| < 1$, a large uncertainty of the system parameters, represented by large values of c_a and c_b, respectively, results in unbounded costs for $N \to \infty$ because the sequence g_k does not converge.

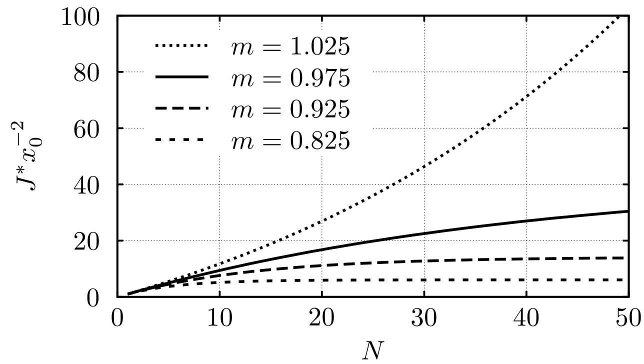


Figure 2.5: Normalized cost (cf. Equation 2.8) of a control problem with uncertain parameters vs. horizon ($q = r = 1$).

Figure 2.5 shows the value of the optimal cost function given by Equation (2.8), normalized w.r.t. x_0^2. The uncertainty of the system parameters is given by $\mu_a = 0.5$, $\mu_b = 1$ and $c_b = 1$. The variance of a_k is chosen to be $c_a \in \{0.7, 0.8, 0.85, 0.9\}$ which results in the threshold parameter $m \in \{0.825, 0.925, 0.975, 1.025\}$ (cf. Equation 2.9). It can be seen that the cost function is bounded as long as $m < 1$, but it grows for increasing horizon without a bound if m is larger than one.

The result presented in this section is only one example of the consideration of dynamical systems which are not perfectly known. The problem of control of and state estimation for systems with uncertain parameters has a long history and can be traced back through more than five decades [25, 71–77]. The references named here model the uncertainty as random variables or processes, respectively, which is consequent from the point of view of stochastic control theory. This is of course not the only possibility to consider imperfect knowledge of system parameters, sometimes there may be even not enough knowledge to provide a suitable stochastic model. A different approach is to assume that the actual dynamical system is an element of a certain set. The uncertainty is modeled by the fact that the system is not known but the set it is taken from. Several approaches to deal with this model description can be found, e. g., in [32, 78, 79].

As a final remark, note that the problem of model uncertainty is not restricted to physical systems which can not be identified with infinite precision. For example, the fact that a coefficient

[9] The result can be verified by assuming that g_k is large and thus neglecting the effect of r in Equation (2.7). The resulting linear time-invariant difference equation converges under the condition of Equation (2.9). Another possibility is to assume that a fixed point of Equation (2.7) exists. The optimal cost is only well defined if the resulting quadratic equation has a positive solution. This also leads to the condition of Equation (2.9).

of a communication channel is not perfectly known can not only be observed in wireless communication scenarios. This uncertainty is also present on the abstract transportation layer in data networks, despite the fact that its parameters are completely determined by the designer and thus can be assumed to be perfectly known. Nevertheless, the data packets of different users that share the network can lead to congestion and packet loss. From the point of view of one user, this can be modeled as a channel coefficient which is either one (packet is delivered) or zero (packet is lost) with a certain probability. In general, this coefficient can not be determined by a transmitter and thus represents a model uncertainty.

2.1.5 Finite Signal-to-Noise Ratio

A fundamental resource for the transmission of information over a communication channel is the power of the signals that carry the information [1]. If a channel is considered which introduces additive noise with constant power, an equivalent representation of the transmit power for such a channel is the so-called SNR [1,2,70], a common term for the ratio of the power of the transmitted signal and the noise power. It is intuitively clear that the more power is available for the transmission of information (or the higher the SNR), the less destructive is the effect of the noise. On the other hand, the amount of available power is typically limited, e. g., due to regulatory reasons or the non-linearity of power amplifiers, which gives rise to the question how to design transmitters and receivers that use the limited communication resource such that the effect of the channel disturbances is minimized.

For the isolated problem of information exchange between one transmitter and one receiver (see Figure 1.1), an increase of the SNR typically[10] results in a better performance of the communication system, but a decrease usually leads to a graceful performance degradation. One example is the Shannon capacity of a scalar channel with real input and AWGN (cf. [1]) which is given by the well known [1, 80] expression $C = \frac{1}{2}\log_2\left(1 + \frac{P}{N}\right)$ bits per transmission, where $P \geq 0$ is the power (or variance) of the channel input signal and $N \geq 0$ the power of the noise. Thus, as long as the signal to noise ratio is larger than zero, the transmission of information is possible. This situation can change when the AWGN channel is used to transmit observations of the state of a dynamical system to a controller in a closed control loop. The critical point is the open loop (in)stability of the dynamical system to be controlled. In the following, we give a simple example of the effect of transmit power limitations in a control system.

Let the system to be controlled be scalar, linear and time-invariant and given by the stochastic difference equation

$$x_{k+1} = ax_k + bu_k + w_k, \quad k \in \mathbb{N}_0, \tag{2.10}$$

with the system state x_k at time index k, initial state $x_0 \sim \mathcal{N}(0, c_{x_0})$, control input u_k and the i.i.d. driving process noise sequence $w_k \sim \mathcal{N}(0, c_w)$. The system parameters a and b are assumed to be known. At the output of the system, it is assumed that the state x_k can be observed without any error at time index k, but has to be transmitted over an AWGN channel to the controller subject to a maximal transmit power of P. In order to control the power of the transmit signal, x_k is multiplied by the factor c_k before transmitting it over the channel. Thus, the controller receives the observation

$$y_k = c_k x_k + n_k, \quad k \in \mathbb{N}_0, \tag{2.11}$$

[10] A properly designed communication system must not react on an increase of the available transmit power with a decrease of performance.

where $n_k \sim \mathcal{N}(0, c_n)$ is the i.i.d. Gaussian channel noise. The controller is chosen such that the infinite horizon LQG average cost function with $q = 1$ and $r = 0$ is minimized. We do not discuss the approach in detail at this point and refer the reader to Appendix A6. Loosely speaking, this controller minimizes the variance of the system state, i. e., the amount of information to be transmitted, and thus helps the transmitter to exploit the available power efficiently. Finally, the scalar c_k is chosen to use the full amount of transmit power, which can be shown to be optimal in the scenario under consideration, i. e.,

$$c_k = \sqrt{\frac{P}{c_{x_k}}}. \tag{2.12}$$

Note that with the system and channel model from Sections 1.5 and 1.6 as well as the choice of the controller, it is possible to determine the sequence of variances c_{x_k}, $k \in \mathbb{N}_0$, which is given by the iteration[11]

$$c_{x_{k+1}} = a^2 \left(c_{x_k} - c_{x_k}^2 \left(c_{x_k} + \frac{c_n}{c_k^2} \right)^{-1} \right) + c_w \tag{2.13}$$

with initial condition c_{x_0} for $k = 0$. Finally, by inserting the factor c_k from Equation (2.12) and some algebraic manipulations, Equation (2.13) can be simplified to

$$\begin{aligned} c_{x_{k+1}} &= a^2 \frac{c_n}{P + c_n} c_{x_k} + c_w \\ &= \left(a^2 \frac{c_n}{P + c_n} \right)^{k+1} c_{x_0} + \sum_{i=0}^{k} \left(a^2 \frac{c_n}{P + c_n} \right)^{i} c_w. \end{aligned} \tag{2.14}$$

Obviously, this sequence of covariances only converges for $a^2 \frac{c_n}{P+c_n} < 1$, otherwise it grows without a bound. This inequality can be rewritten to provide an explicit condition for the transmit power that has to be provided such that the variance of the system state stays bounded:

$$P > (a^2 - 1)c_n. \tag{2.15}$$

The property of a bounded state covariance matrix corresponds to the stability of the system in the mean square sense, see Definition 1.4.9. Thus, in order to obtain a stable control loop, the available transmit power P has to be chosen depending on the system parameter a and the variance of the channel noise.

Equation (2.15) shows the fundamental difference between the control of stable and unstable dynamical systems with finite transmit power. If the system is stable even without control, i. e., if $|a| < 1$, the available transmit power can be reduced down to zero, which is a quite obvious result: a stable system does not need control to be stabilized. On the other hand, dynamical systems which are open loop unstable always need some control input to be stabilized which results in the requirement of a strictly positive transmit power. The amount of the power depends on the degree of instability, i. e., the value of the squared system parameter a.

It can be seen by Equation (2.13) that as long as k is finite, the variance of the system state is also finite and consequently the power constraint can be met with a non-zero transmit scaling. In order to show the effect of the available transmit power on the stability of the control loop, Figure 2.6 shows the asymptotic value of the state variance for $k \to \infty$. As long as the transmit

[11] Equation (2.13) stems from the Riccati iteration for the error covariance matrix of the Kalman filter, see Appendix A7, since the variance of the system state is in this case identical to the variance of the state prediction error.

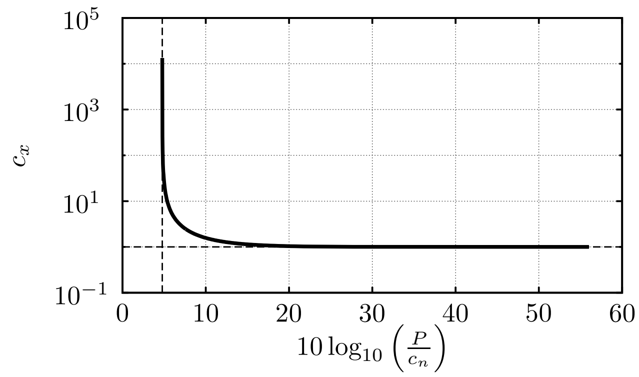


Figure 2.6: Asymptotic variance $c_x = \lim_{k\to\infty} c_{x_k}$ of the system state over the SNR ($a = 2$, $c_w = 1$).

power (or SNR, respectively) is large enough, the closed loop system is mean square stable and thus has a finite variance of the system state. For the chosen parameters $a = 2$, the bound for the SNR, given by $\frac{P}{c_n}$, is 3 (cf. Equation 2.15). Consequently, we observe an unbounded increase of the asymptotic variance of the system state when the SNR approaches this value. On the other hand, the asymptotic value of the state variance for $\frac{P}{c_n} \to \infty$ is (cf. Equation 2.14) $c_w = 1$.

Finally, consider the case when the available transmit power does not suffice to stabilize the scalar system in the mean square sense. For the parameters $a = 2$ and $c_n = 1$, the power must be larger than 3. If we chose instead only 95% of this value, i. e., $P = 2.85$, it can not be expected to obtain a stable closed loop system. Figure 2.7 shows the effect on the variance of the system state. In order to fulfill the power constraint, the transmitter has to chose a scaling factor c_k which leads

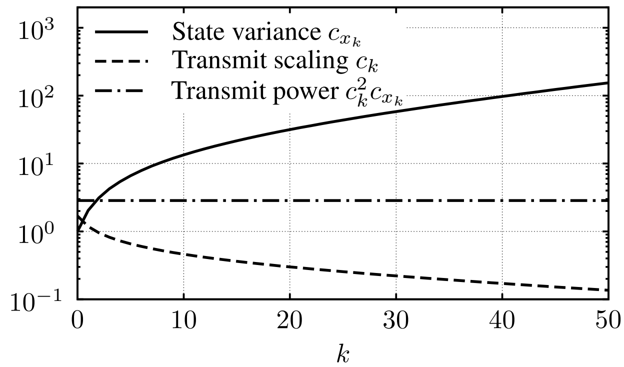


Figure 2.7: Variance of the system state, scalar at the transmitter and variance of the transmit signal for a transmit power that is not sufficient for stabilizability in the mean square sense ($a = 2, c_n = 1, P = 2.85$).

to an increase of the variance of the system state. Consequently, the transmitter further decreases its scaling, leading to a further increase of the state variance and so on. The fact that the transmit power is too low results in an unbounded increase.

The consideration of power constraints in control problems has a long history. Even the unconstrained optimization of a weighted quadratic cost function in LQG control problems (see, e. g., [4] and Appendix A6) includes the desire to keep the variance of the system states, outputs or control inputs small. An explicit consideration of variance constraints using Lagrange multipliers can be found in [81] where the controller is the only instance which can be used to fulfill the constraints and no additional transmitter like the scaling factor in the example above is present. The authors of [82] take into account power constraints together with the optimization of the measurement (or transmission) strategy. The variance of signals at the system output is the subject of investigation in [83, 84], where [84] formulates the constraints not in terms of power but SNR. A consideration of limited transmit power as an extension to classical LQG control can be found in [85]. A breakthrough for the investigation of linear systems with power or SNR constraints was [29] where a condition similar to Equation (2.15) has been determined for general Single-Input Single-Output (SISO) systems with linear controllers. The work has been extended in [86] with a scaling at the transmitter and in [87] where it has been shown that the limits of stabilizability do not change when arbitrary time varying or non-linear control is used. Contributions which emphasize the role of (convex) optimization techniques for the controller, transmitter and receiver design are [35] and [88], were the authors of [89] and [35] provide a connection between additive noise, data rate limited channels and information theoretic aspects of control with communication constraints which lead to implementable NCSs.

2.2 Scope of the Thesis

The concept of NCSs covers a wide range of models for the dynamical systems and the communication infrastructure that is used for the information exchange of such systems. A common feature which separates NCSs from the classical control system design is to consider explicitly the influence of the communication infrastructure on the dynamical behavior of the control system. Depending on the dynamical system and communication channel model under consideration, different effects can be observed and a suitable analysis and design methodology has to be found.

The intention of this thesis is to investigate challenges and benefits of a *joint design* of control systems, transmitters and receivers on a low level, i. e., with system and channel models which are close to their physical description. From a practical point of view, other models like erasure channels (packet drops, cf. Section 2.1.2) or discrete channels (quantization, cf. Section 2.1.3) can be of larger interest, e. g., if control systems are designed on top of a existing general purpose communication system like a Local Area Network (LAN) or the Internet. With the model of communication imperfections on the higher level of abstraction which are provided by those networks, data loss and quantization are relevant problems to deal with. But despite the practical importance of the investigation of scenarios using these channel models and the theoretical bounds derived from it, the degrees of freedom for a joint design of control and communication are somehow restricted. If data loss is the only imperfection, the underlying communication system must be designed to provide the lowest possible loss probability. If quantization is the only source of distortion, one has to provide the largest possible resolution of the quantizer. Of course there may be a coupling between the resolution of the quantizer which determines the data rate and the probability of data loss in a communication network and it is not trivial to find the optimal trade-off. Nevertheless, the design of the physical communication system is separated from the controller design and shifted to higher levels of abstraction or layers, e. g., the design of control oriented communication and routing protocols for data loss networks [90, 91].

In the following, for the goal of a joint design of control and communication, simple but not trivial system and channel models have been chosen which can be treated in a common framework. In more detail, the models are linear dynamical systems and additive noise channels. Additionally, only linear transmitters and receivers are considered and optimized, although it is known that even in such a simple scenario non-linear communication and control techniques are optimal (cf. [39, 40]). The linear models are compatible in the sense that they can be treated on the same level and in the same optimization framework. Additionally, the consideration of variance-based constraints is quite natural in this framework. Thus, the limited communication resource will be the power of transmit signals or the SNR of the communication channels, respectively, where the former is more common in the communication community and the latter in the control community.

At the time of writing, the author of this thesis was member of a research group with a focus on signal processing for communications. Consequently, the following results are presented with the intention to introduce the special features and problems of a joint investigation of control and communication systems to the signal processing and communication community. This is the reason why some results which might be considered standard in the context of control system design, e. g., LQG optimal control, dynamic programming, optimal estimation using the Kalman filter, Lyapunov and Riccati equations, are introduced in the Appendix. Since they represent the basic tools for the analysis and design of control systems under power or SNR constraints, respectively, we hope that the way of presenting these tools is useful for the reader who is not familiar with them. The reader with a background in the corresponding topics may start directly with the discussion of power (or SNR) constrained LQG control in the following chapter.

3. Optimal Control With Power Constraints

3.1 The Unconstrained Optimization Problem

Consider the problem of the optimal design of a controller which is separated from the dynamical system to be controlled by two communication channels, one which is used for the transmission of the system output to the controller and one for the transmission of the control signal back to the system input. According to the system and channel model which have been presented in Sections 1.5 and 1.6, the dynamical system is assumed to be Linear Time-Invariant (LTI) and the channels are assumed to introduce stationary additive noise into the closed control loop. Figure 3.1 illustrates the scenario.

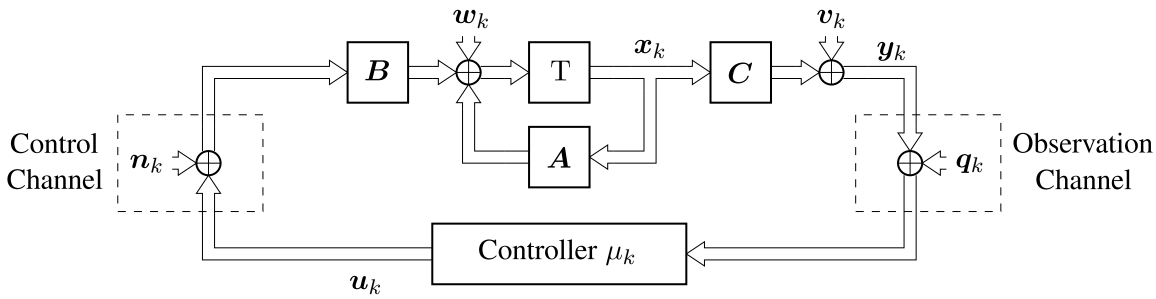


Figure 3.1: Model of the control loop which is closed over two channels with additive noise $\boldsymbol{q}_k$ and $\boldsymbol{n}_k$.

The optimal controller is determined by the solution of the Linear Quadratic Gaussian (LQG) optimization problem (cf. Chapter A6 and Section A6.3)

$$\underset{\mu_0,\mu_1,\mu_2\dots}{\text{minimize}} \lim_{N\to\infty} \frac{1}{N} \operatorname{E}_{\substack{\boldsymbol{x}_0,\boldsymbol{w}_0,\dots,\boldsymbol{w}_{N-1},\boldsymbol{v}_0,\dots,\boldsymbol{v}_{N-1},\\ \boldsymbol{q}_0,\dots,\boldsymbol{q}_{N-1},\boldsymbol{n}_0\dots,\boldsymbol{n}_{N-1}}} \left[\boldsymbol{x}_N^{\mathrm{T}} \boldsymbol{Q}_N \boldsymbol{x}_N + \sum_{n=0}^{N-1} \begin{bmatrix} \boldsymbol{x}_n \\ \boldsymbol{u}_n \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{S} \\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{R} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_n \\ \boldsymbol{u}_n \end{bmatrix} \right] \tag{3.1}$$

$$\begin{aligned}
\text{subject to} \quad \boldsymbol{x}_{k+1} &= \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{B}(\boldsymbol{u}_k + \boldsymbol{n}_k) + \boldsymbol{w}_k, && k \in \mathbb{N}_0, \\
\boldsymbol{y}_k &= \boldsymbol{C}\boldsymbol{x}_k + \boldsymbol{v}_k, && k \in \mathbb{N}_0, \\
\boldsymbol{u}_k &= \mu_k(\mathcal{I}_k), && k \in \mathbb{N}_0, \\
\mathcal{I}_k &= \begin{cases} \{(\boldsymbol{y}_0 + \boldsymbol{q}_0)\}, & \\ \{(\boldsymbol{y}_0 + \boldsymbol{q}_0), (\boldsymbol{y}_1 + \boldsymbol{q}_1), \dots, (\boldsymbol{y}_k + \boldsymbol{q}_k), \boldsymbol{u}_0, \boldsymbol{u}_1, \dots, \boldsymbol{u}_{k-1}\}, & \end{cases} && \begin{matrix} k = 0, \\ k \in \mathbb{N}, \end{matrix}
\end{aligned}$$

where the controller μ_k, $k \in \mathbb{N}_0$, is the function which maps the available information at time index k to the control signal $\boldsymbol{u}_k$. Note that the system input is the sum $\boldsymbol{u}_k + \boldsymbol{n}_k$, $k \in \mathbb{N}_0$, of the control signal and the control channel noise, and the observations for the determination of the control signal are given by the sum $\boldsymbol{y}_k + \boldsymbol{q}_k$, $k \in \mathbb{N}_0$, of the system output and the observation channel noise. Since the additive noises of the communication channels are modeled as random sequences, the expected value in Equation (3.1) is additionally taken w.r.t. the corresponding random vectors.

Recall the assumption that $(\boldsymbol{w}_k : k \in \mathbb{N}_0)$, $(\boldsymbol{v}_k : k \in \mathbb{N}_0)$, $(\boldsymbol{q}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{n}_k : k \in \mathbb{N}_0)$ are mutually independent sequences of independent Gaussian random vectors and are additionally independent of the initial state $\boldsymbol{x}_0$ which is described by a Gaussian random vector. Since the time-invariant case is considered, the noise sequences are also assumed to be stationary. Using these assumptions, it can be seen that the above optimization is a standard LQG problem with process noise $\boldsymbol{B}\boldsymbol{n}_k + \boldsymbol{w}_k$, $k \in \mathbb{N}_0$, and observation noise $\boldsymbol{v}_k + \boldsymbol{q}_k$, $k \in \mathbb{N}_0$, where the corresponding noise sequences also fulfill the above independence assumptions. Thus, the channel noise sequences just change the parameters of the process and observation noise. Note the coupling of the noise and system parameters by the covariance matrix $\boldsymbol{C}_{\boldsymbol{w}} + \boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}$ of the effective driving process noise.

In Section A6.3 it is shown that the solution to the above optimization problem, more precisely the control sequence $(\boldsymbol{u}_k : k \in \mathbb{N}_0)$, is given by

$$\boldsymbol{u}_k = \boldsymbol{L}\hat{\boldsymbol{x}}_k, \qquad k \in \mathbb{N}_0, \tag{3.2}$$

where

$$\boldsymbol{L} = -\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B} + \boldsymbol{R}\right)^{-1}\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} + \boldsymbol{S}^{\mathrm{T}}\right), \tag{3.3}$$

with the stabilizing solution $\boldsymbol{K}$ of the Discrete Algebraic Riccati Equation (DARE) (see Appendix A3)

$$\boldsymbol{K} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B} + \boldsymbol{S}\right)\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B} + \boldsymbol{R}\right)^{-1}\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} + \boldsymbol{S}^{\mathrm{T}}\right) + \boldsymbol{Q}, \tag{3.4}$$

and the state estimate

$$\hat{\boldsymbol{x}}_k = \mathrm{E}_{\boldsymbol{x}_k|\mathcal{I}_k}\left[\boldsymbol{x}_k|\,\mathcal{I}_k\right], \qquad k \in \mathbb{N}_0. \tag{3.5}$$

The estimate $\hat{\boldsymbol{x}}_k$ is computed using the Kalman filter (cf. Appendix A7 and Section A7.2) which also provides the covariance matrix of the asymptotic estimation error

$$\begin{aligned}\boldsymbol{C}_{\tilde{\boldsymbol{x}}} &= \lim_{k\to\infty} \mathrm{E}_{\boldsymbol{x}_k,\mathcal{I}_k}\left[\left(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k\right)\left(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k\right)^{\mathrm{T}}\right] \\ &= \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{C}_{\boldsymbol{q}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}},\end{aligned} \tag{3.6}$$

which is determined by the stabilizing solution of the DARE

$$\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} = \boldsymbol{A}\left(\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{C}_{\boldsymbol{q}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\right)\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{w}} + \boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}. \tag{3.7}$$

Since the observation and control channel represent a communication system, it is of importance for the design of such a system which amount of transmit power has to be provided for the transmission of the system output and the control signals. However, for the determination of the optimal control sequence $(\boldsymbol{u}_k : k \in \mathbb{N}_0)$ (cf. Equation 3.2), no constraints for the variances of $\boldsymbol{y}_k$ and $\boldsymbol{u}_k$, $k \in \mathbb{N}_0$, i. e., the powers of the signals at the input of the observation and the control channel, respectively, have been considered. This means that we have to live with the result of the optimization and provide the corresponding power for the transmission of these signals. In the following, the necessary transmit power for the observation and control channel will be determined using the results of the optimal controller given in Equations (3.2) to (3.7). To this end, we will consider asymptotic variances which are associated to the property of mean square stability (cf. Definition 1.4.9). Note that under the assumption that the DAREs given by Equations (3.4) and (3.7) have stabilizing solutions (see Appendix A3), the closed loop dynamics of the interconnection of the original system and the optimal controller which is based on the Kalman filter are described by a stable dynamical LTI system, see, e. g., [3, pp. 275-276], [4, pp. 542-543] and [5, pp. 300-301]. Thus, if this system is driven by Wide Sense Stationary (WSS) noise sequences, the resulting state sequences $(\boldsymbol{x}_k : k \in \mathbb{N}_0)$ and $(\hat{\boldsymbol{x}}_k : k \in \mathbb{N}_0)$ as well as the associated output sequences $(\boldsymbol{y}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{u}_k : k \in \mathbb{N}_0)$ are asymptotically WSS (see Definition 1.4.8).

Definition 3.1.1: Power of transmit signals
The transmit power of the observation channel is determined by the variance of the observations $\boldsymbol{y}_k$, $k \in \mathbb{N}_0$, at the system output. With the LTI system, the time-invariant controller and the stationary noise sequences, this power is (asymptotically) given by

$$P_1 = \lim_{k\to\infty} \mathrm{E}_{\boldsymbol{y}_k}\left[\|\boldsymbol{y}_k\|_2^2\right] = \mathrm{tr}\left[\boldsymbol{C}\boldsymbol{C}_{\boldsymbol{x}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}}\right]. \tag{3.8}$$

Analogously, the power of the control channel is given by the variance of the control signal $\boldsymbol{u}_k$, $k \in \mathbb{N}_0$, and reads as

$$P_2 = \lim_{k\to\infty} \mathrm{E}_{\boldsymbol{u}_k}\left[\|\boldsymbol{u}_k\|_2^2\right] = \mathrm{tr}\left[\boldsymbol{C}_{\boldsymbol{u}}\right] = \mathrm{tr}\left[\boldsymbol{L}\boldsymbol{C}_{\hat{\boldsymbol{x}}}\boldsymbol{L}^{\mathrm{T}}\right]. \tag{3.9}$$

For both transmit powers, the assumption has been used that all random vectors have zero mean.

The first step for the computation of the covariance matrices $\boldsymbol{C}_{\boldsymbol{x}}$ and $\boldsymbol{C}_{\hat{\boldsymbol{x}}}$ which determine P_1 and P_2 is to insert the optimal control input given by Equation (3.2) in the state equation given by the constrains of the optimization problem (3.1) which results in

$$\begin{aligned} \boldsymbol{x}_{k+1} &= \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{B}\boldsymbol{L}\hat{\boldsymbol{x}} + \boldsymbol{w}_k + \boldsymbol{B}\boldsymbol{n}_k \\ &= (\boldsymbol{A} + \boldsymbol{B}\boldsymbol{L})\hat{\boldsymbol{x}}_k + \boldsymbol{A}\tilde{\boldsymbol{x}}_k + \boldsymbol{w}_k + \boldsymbol{B}\boldsymbol{n}_k, \qquad k \in \mathbb{N}_0, \end{aligned} \tag{3.10}$$

with the state estimation error (cf. Equation 3.6)

$$\tilde{\boldsymbol{x}}_k = \boldsymbol{x}_k - \hat{\boldsymbol{x}}_k, \qquad k \in \mathbb{N}_0. \tag{3.11}$$

A standard result which is also used for the derivation of the Kalman filter (see Section A7.2) is that the optimal state estimate $\hat{\boldsymbol{x}}_k$ and the resulting estimation error $\tilde{\boldsymbol{x}}_k$ are uncorrelated for all $k \in \mathbb{N}_0$, i. e., for the presented scenario it holds that

$$\mathrm{E}\left[\hat{\boldsymbol{x}}_k\tilde{\boldsymbol{x}}_k^{\mathrm{T}}\right] = \boldsymbol{0}_{N_x\times N_x}, \qquad k \in \mathbb{N}_0. \tag{3.12}$$

Using this fact, the asymptotic covariance matrix of the system state is given by,

$$\boldsymbol{C}_{\boldsymbol{x}} = (\boldsymbol{A} + \boldsymbol{B}\boldsymbol{L})\boldsymbol{C}_{\hat{\boldsymbol{x}}}(\boldsymbol{A} + \boldsymbol{B}\boldsymbol{L})^{\mathrm{T}} + \boldsymbol{A}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{w}} + \boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}. \tag{3.13}$$

Due to the property of uncorrelatedness, the covariance matrix above can also be expressed as

$$\boldsymbol{C}_{\boldsymbol{x}} = \boldsymbol{C}_{\hat{\boldsymbol{x}}} + \boldsymbol{C}_{\tilde{\boldsymbol{x}}}, \tag{3.14}$$

which together with Equation (3.13) provides the following Lyapunov equation for the determination of $\boldsymbol{C}_{\hat{\boldsymbol{x}}}$:

$$\boldsymbol{C}_{\hat{\boldsymbol{x}}} = (\boldsymbol{A} + \boldsymbol{B}\boldsymbol{L})\boldsymbol{C}_{\hat{\boldsymbol{x}}}(\boldsymbol{A} + \boldsymbol{B}\boldsymbol{L})^{\mathrm{T}} + \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}, \tag{3.15}$$

where $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} = \boldsymbol{A}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{w}} + \boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}$ (cf. Equations 3.7 and 3.6). Note that the difference $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}$ is positive semidefinite, which can be verified with Equation (3.6), and that the magnitude of all eigenvalues of $\boldsymbol{A} + \boldsymbol{B}\boldsymbol{L}$ is less than one because $\boldsymbol{K}$ is the stabilizing solution of the DARE given by Equation (3.4). Thus, $\boldsymbol{C}_{\hat{\boldsymbol{x}}}$ is the unique positive semidefinite solution of Equation (3.15). Finally, the covariance matrix $\boldsymbol{C}_{\boldsymbol{x}}$ is given by Equation (3.14).

The following example will be used to illustrate the above result. It does not represent a specific physical or technical system and has been chosen randomly in order to visualize and hopefully clarify the presented concepts and the impact of different approaches to the joint treatment of control and communication. Consequently, we will come back to this example repeatedly in the following sections and chapters.

Example 3.1.1 Let the LTI system to be controlled be given by the matrices

$$\boldsymbol{A} = \begin{bmatrix} 0.5 & 0.9 & -0.4 \\ 1.5 & 0.3 & 0.3 \\ -1.5 & -1.3 & 1.2 \end{bmatrix}, \qquad \boldsymbol{B} = \begin{bmatrix} 2.7 & 0.7 \\ -1.3 & 0 \\ 3 & 0.7 \end{bmatrix}, \qquad \boldsymbol{C} = \begin{bmatrix} -0.2 & 1.5 & 1.4 \\ -0.1 & 1.4 & 0.7 \end{bmatrix},$$

i. e., $N_x = 3$ and $N_y = N_u = 2$. The pair $(\boldsymbol{A}, \boldsymbol{B})$ is controllable and $(\boldsymbol{A}, \boldsymbol{C})$ is observable. The matrix $\boldsymbol{A}$ has only real eigenvalues and is unstable, i. e., the magnitude of one eigenvalue is larger than one (approximately 1.96).

The process and observation noise sequences are Gaussian, identically and independently distributed (i.i.d.) with zero mean and covariance matrices

$$\boldsymbol{C}_{\boldsymbol{w}} = \begin{bmatrix} 0.2 & 0 & 0.1 \\ 0 & 0.6 & -0.2 \\ 0.1 & -0.2 & 0.2 \end{bmatrix}, \qquad \boldsymbol{C}_{\boldsymbol{v}} = \boldsymbol{0}_{N_y \times N_y},$$

respectively. Note that we assume the absence of observation noise at this point without loss of generality because it can be combined with the observation channel noise sequence $(\boldsymbol{q}_k : k \in \mathbb{N}_0)$. This sequence describes the effect of communication imperfections in one channel. The other channel is characterized by the random sequence $(\boldsymbol{n}_k : k \in \mathbb{N}_0)$. Both sequences are Gaussian i.i.d. with zero mean and covariance matrices

$$\boldsymbol{C}_{\boldsymbol{q}} = \begin{bmatrix} 1 & -0.1 \\ -0.1 & 0.5 \end{bmatrix}, \qquad \boldsymbol{C}_{\boldsymbol{n}} = \begin{bmatrix} 0.7 & -0.4 \\ -0.4 & 0.3 \end{bmatrix},$$

respectively. Finally, the parameters of the LQG cost function are given by

$$\boldsymbol{Q} = \begin{bmatrix} 7.2 & -0.7 & -1.3 \\ -0.7 & 4.2 & -0.4 \\ -1.3 & -0.4 & 0.35 \end{bmatrix}, \qquad \boldsymbol{R} = \begin{bmatrix} 2.7 & 0.6 \\ 0.6 & 0.3 \end{bmatrix}, \qquad \boldsymbol{S} = \boldsymbol{0}_{3\times 2}.$$

Using all these parameters of the dynamical system, the communication channels and the cost function, the solution of the optimization problem in Equation (3.1) leads to the transmit powers

$$P_1 = \operatorname{tr}\left[\boldsymbol{C}\boldsymbol{C}_{\boldsymbol{x}}\boldsymbol{C}^{\mathrm{T}}\right] \approx 4189 \quad \text{and} \quad P_2 = \operatorname{tr}\left[\boldsymbol{L}\boldsymbol{C}_{\hat{\boldsymbol{x}}}\boldsymbol{L}^{\mathrm{T}}\right] \approx 1661. \tag{3.16}$$

The respective Signal to Noise Ratios (SNRs) are

$$\varphi_1 = \frac{P_1}{\operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{q}}\right]} \approx 2793 \quad \text{and} \quad \varphi_2 = \frac{P_2}{\operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{n}}\right]} \approx 1661. \tag{3.17}$$

In order to compare these values with later results, Figure 3.2 shows the SNRs with a logarithmic scale, i. e.,

$$10\log_{10}(\varphi_1) \approx 34.46 \quad \text{and} \quad 10\log_{10}(\varphi_2) \approx 32.20.$$

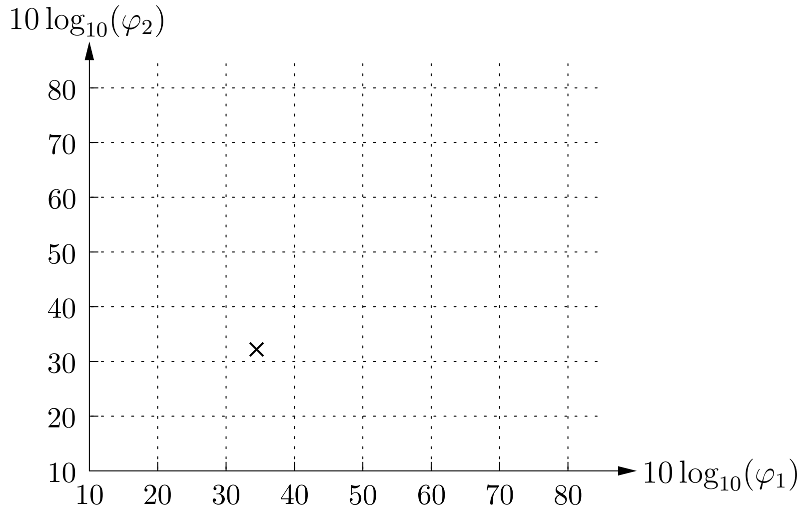


Figure 3.2: SNRs of the observation and control channel for the unconstrained LQG problem.

3.2 The Constrained Optimization Problem

The transmit powers or SNRs, respectively, which are the result of the optimization problem (3.1) without the consideration of any power constraints may be unsatisfactory for many reasons. Assume for example an asymmetric communication scenario where the transmitter at the system output has a much smaller maximum transmit power available than the transmitter at the controller. This is often the case if the system to be controlled is a mobile unit with simple communication equipment and the controller is located at a base station without this restriction. As a second example, consider the case where both transmitters have more transmit power available than the amount which must be provided for the optimal LQG controller. In this case, it is desirable to use the additional communication resources to combat the negative effect of the channel noises.

The obvious way to incorporate the limited communication resources of the transmitters in a control loop which is closed over communication channels is the addition of constraints to the problem (3.1). Let the available transmit powers for the observation and control channel be $P_{\mathrm{Tx},1} \geq 0$ and $P_{\mathrm{Tx},2} \geq 0$, respectively. The corresponding constraints for the actual transmit powers thus are (cf. Equations 3.8 and 3.9)

$$\begin{aligned} \operatorname{tr}\left[\boldsymbol{C}\boldsymbol{C}_{\boldsymbol{x}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}}\right] &\leq P_{\mathrm{Tx},1} \qquad \text{and} \\ \operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{u}}\right] &\leq P_{\mathrm{Tx},2}. \end{aligned} \tag{3.18}$$

In this section we adopt a method which has been proposed, e. g., in [81] and recently in [88], to include the constraints given by Equation (3.18) in the LQG control problem. The basic idea is to reformulate the optimization problem in terms of (asymptotic) covariance matrices. This formulation is equivalent to problem (3.1) if the optimization is carried out within the class of controllers which provide asymptotically WSS state and control sequences $(\boldsymbol{x}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{u}_k : k \in \mathbb{N}_0)$, respectively. Since the optimal controller shown in Equations (3.2) to (3.7) belongs to this class, we expect no loss of optimality due to this restriction.[1] The reason for the reformulation of the

[1]The constrained optimization problem can be solved using the formulation of Equation (3.1) (cf. Chapter 4) which provides the same result as the one shown in the following. Thus, the optimality of the solution is not just a conjecture.

LQG control problem in terms of covariance matrices is that it allows for a direct application of general purpose software for semidefinite optimization for the numerical investigations. We do not claim that this is the most efficient way of solving the problem, but it is not necessary to develop a special algorithm for the numerical evaluation of the presented results and we can focus on the structure of the problem in the following discussions.

We start with the reformulation of the basic LQG problem (3.1) in terms of covariance matrices and include the power constraint in a second step. Since it is assumed that the controller generates jointly asymptotically WSS state and control sequences, the respective covariance matrices converge to their stationary values and the average cost infinite horizon problem becomes

$$\begin{aligned} &\underset{\boldsymbol{C}_{\boldsymbol{x}},\boldsymbol{C}_{\boldsymbol{x},\boldsymbol{u}},\boldsymbol{C}_{\boldsymbol{u}}}{\text{minimize}} \ \operatorname{tr}\left[\begin{bmatrix}\boldsymbol{Q} & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{R}\end{bmatrix}\begin{bmatrix}\boldsymbol{C}_{\boldsymbol{x}} & \boldsymbol{C}_{\boldsymbol{x},\boldsymbol{u}}\\ \boldsymbol{C}_{\boldsymbol{x},\boldsymbol{u}}^{\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{u}}\end{bmatrix}\right] \\ &\text{subject to} \quad \boldsymbol{C}_{\boldsymbol{x}} = \begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\end{bmatrix}\begin{bmatrix}\boldsymbol{C}_{\boldsymbol{x}} & \boldsymbol{C}_{\boldsymbol{x},\boldsymbol{u}}\\ \boldsymbol{C}_{\boldsymbol{x},\boldsymbol{u}}^{\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{u}}\end{bmatrix}\begin{bmatrix}\boldsymbol{A}^{\mathrm{T}}\\ \boldsymbol{B}^{\mathrm{T}}\end{bmatrix} + \boldsymbol{C}_{\boldsymbol{w}} + \boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}, \\ &\qquad\qquad \begin{bmatrix}\boldsymbol{C}_{\boldsymbol{x}} & \boldsymbol{C}_{\boldsymbol{x},\boldsymbol{u}}\\ \boldsymbol{C}_{\boldsymbol{x},\boldsymbol{u}}^{\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{u}}\end{bmatrix} \succeq \mathbf{0}_{(N_x+N_u)\times(N_x+N_u)}, \end{aligned} \tag{3.19}$$

where we have to keep in mind that the control input $\boldsymbol{u}_k$ at time index $k \in \mathbb{N}_0$ is restricted to be a function μ_k of the information $\mathcal{I}_k$. This restriction is not captured by the problem formulation above and it can not be expected that its solution has the required property. In order to include this constraint in the optimization problem (3.19), the fact can be used that the state estimate $\hat{\boldsymbol{x}}_k = \mathrm{E}_{\boldsymbol{x}_k|\mathcal{I}_k}[\boldsymbol{x}_k|\,\mathcal{I}_k]$ and the associated estimation error $\tilde{\boldsymbol{x}}_k = \boldsymbol{x}_k - \hat{\boldsymbol{x}}_k$ are uncorrelated, which results in the property $\boldsymbol{C}_{\boldsymbol{x}} = \boldsymbol{C}_{\hat{\boldsymbol{x}}} + \boldsymbol{C}_{\tilde{\boldsymbol{x}}}$. Additionally, the estimation error and any function of the information for the determination of the estimate are uncorrelated (cf. [33, p. 346]), i. e.,

$$\begin{aligned} \boldsymbol{C}_{\tilde{\boldsymbol{x}},\boldsymbol{u}} &= \lim_{k\to\infty} \mathrm{E}_{\boldsymbol{x}_k,\boldsymbol{u}_k,\mathcal{I}_k}\left[\left(\boldsymbol{x}_k - \mathrm{E}_{\boldsymbol{x}_k|\mathcal{I}_k}[\boldsymbol{x}_k|\,\mathcal{I}_k]\right)\boldsymbol{u}_k^{\mathrm{T}}\right] \\ &= \lim_{k\to\infty} \mathrm{E}_{\boldsymbol{x}_k,\mathcal{I}_k}\left[\left(\boldsymbol{x}_k - \mathrm{E}_{\boldsymbol{x}_k|\mathcal{I}_k}[\boldsymbol{x}_k|\,\mathcal{I}_k]\right)\mu_k(\mathcal{I}_k)^{\mathrm{T}}\right] \\ &= \lim_{k\to\infty} \left(\mathrm{E}_{\boldsymbol{x}_k,\mathcal{I}_k}\left[\boldsymbol{x}_k\mu_k(\mathcal{I}_k)^{\mathrm{T}}\right] - \mathrm{E}_{\mathcal{I}_k}\left[\mathrm{E}_{\boldsymbol{x}_k|\mathcal{I}_k}\left[\boldsymbol{x}_k\mu_k(\mathcal{I}_k)^{\mathrm{T}}\middle|\,\mathcal{I}_k\right]\right]\right) \\ &= \lim_{k\to\infty} \left(\mathrm{E}_{\boldsymbol{x}_k,\mathcal{I}_k}\left[\boldsymbol{x}_k\mu_k(\mathcal{I}_k)^{\mathrm{T}}\right] - \mathrm{E}_{\boldsymbol{x}_k,\mathcal{I}_k}\left[\boldsymbol{x}_k\mu_k(\mathcal{I}_k)^{\mathrm{T}}\right]\right) \\ &= \mathbf{0}_{N_x\times N_u}, \end{aligned} \tag{3.20}$$

where the results of Sections A5.1 and A5.2 have been used. Consequently, we have the identity

$$\boldsymbol{C}_{\boldsymbol{x},\boldsymbol{u}} = \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}. \tag{3.21}$$

Note that the covariance matrix of the system state reads as $\boldsymbol{C}_{\boldsymbol{x}} = \boldsymbol{C}_{\hat{\boldsymbol{x}}} + \boldsymbol{C}_{\tilde{\boldsymbol{x}}}$, where the controller has no impact on the covariance matrix $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}$ of the estimation error (cf. Section A6.2) for the given scenario. Thus, together with Equation (3.21), it can be seen that the LQG problem in Equation (3.19) can be reformulated using the joint covariance matrix of the state *estimate* and the control signal. Since the estimate at time index $k \in \mathbb{N}_0$ is a function of the information $\mathcal{I}_k$, the optimization over the joint covariance matrix ensures that no additional information is required by the optimal controller. Using this formulation of the optimization problem (3.19) and considering the constraints given by Equation (3.18) finally leads to the power constrained LQG control problem

$$\begin{aligned}
\underset{\boldsymbol{C}_{\hat{\boldsymbol{x}}},\boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}},\boldsymbol{C}_{\boldsymbol{u}}}{\text{minimize}} \quad & \operatorname{tr}\left[\begin{bmatrix}\boldsymbol{Q} & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{R}\end{bmatrix}\begin{bmatrix}\boldsymbol{C}_{\hat{\boldsymbol{x}}} & \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}\\ \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{u}}\end{bmatrix}\right] + \operatorname{tr}\left[\boldsymbol{Q}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\right] \qquad (3.22)\\
\text{subject to} \quad & \operatorname{tr}\left[\boldsymbol{C}\boldsymbol{C}_{\hat{\boldsymbol{x}}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}}\right] + \operatorname{tr}\left[\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\boldsymbol{C}^{\mathrm{T}}\right] \le P_{\mathrm{Tx},1},\\
& \operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{u}}\right] \le P_{\mathrm{Tx},2},\\
& \boldsymbol{C}_{\hat{\boldsymbol{x}}} = \begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\end{bmatrix}\begin{bmatrix}\boldsymbol{C}_{\hat{\boldsymbol{x}}} & \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}\\ \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{u}}\end{bmatrix}\begin{bmatrix}\boldsymbol{A}^{\mathrm{T}}\\ \boldsymbol{B}^{\mathrm{T}}\end{bmatrix} + \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}},\\
& \begin{bmatrix}\boldsymbol{C}_{\hat{\boldsymbol{x}}} & \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}\\ \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{u}}\end{bmatrix} \ge \mathbf{0}_{(N_x+N_u)\times(N_x+N_u)},
\end{aligned}$$

where the estimation error covariance matrices $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}$ and $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}$ do not depend on the optimization variables and are obtained from Equations (3.6) and (3.7). For the equality constraint, the fact has been used that $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} = \boldsymbol{A}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{w}} + \boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}$. The problem above consists of a cost function and equality as well as inequality constraints which are linear in the optimization variables and an additional semidefiniteness constraint for the joint covariance matrix of the state estimate and the control input. Thus, it can be solved using standard software for semidefinite optimization. In the following, the toolbox SeDuMi [92] will be used together with the interface YALMIP [93] which allows for a direct implementation of the problem formulation above.

The remaining step is to recover the functions μ_k of the information $\mathcal{I}_k$, $k \in \mathbb{N}_0$, for the computation of the control sequence $(\boldsymbol{u}_k : k \in \mathbb{N}_0)$ from the covariance matrices determined by the optimization problem (3.22). Recall that in the unconstrained case, these functions are given by a constant linear function of the optimal estimate of the system state $\boldsymbol{x}_k$ given $\mathcal{I}_k$, $k \in \mathbb{N}_0$, (cf. Equation 3.2). Assume that in the constrained case, the optimizing covariance matrices are given by $\boldsymbol{C}_{\hat{\boldsymbol{x}}}^*$, $\boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^*$ and $\boldsymbol{C}_{\boldsymbol{u}}^*$. Thus, we only know the joint covariance matrix of the state estimate $\hat{\boldsymbol{x}}_k$ and the control $\boldsymbol{u}_k$ but not the mapping from $\mathcal{I}_k$ to $\boldsymbol{u}_k$. On the other hand, since the state estimate $\hat{\boldsymbol{x}}_k$ is a function of $\mathcal{I}_k$, a mapping from $\hat{\boldsymbol{x}}_k$ to the control $\boldsymbol{u}_k$ which provides the optimal joint covariance matrix solves the problem.

It is not guaranteed that the optimal control input is a linear function of the optimal state estimate alone since, in general, their joint covariance matrix may have full rank. In the following, it will be shown that this is not the case for the optimizing joint covariance matrix given by $\boldsymbol{C}_{\hat{\boldsymbol{x}}}^*$, $\boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^*$ and $\boldsymbol{C}_{\boldsymbol{u}}^*$. To this end, the linear estimate $\hat{\boldsymbol{u}}_k$, $k \in \mathbb{N}_0$, of the control signal given the state estimate is computed, i. e.,

$$\hat{\boldsymbol{u}}_k = \boldsymbol{L}\hat{\boldsymbol{x}}_k, \qquad k \in \mathbb{N}_0, \tag{3.23}$$

where

$$\boldsymbol{L} = \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{*,\mathrm{T}}\boldsymbol{C}_{\hat{\boldsymbol{x}}}^{*,-1}, \tag{3.24}$$

and it is assumed that the inverse exists.[2] Note that $\hat{\boldsymbol{u}}_k$ is the optimal linear estimate of $\boldsymbol{u}_k$ given $\hat{\boldsymbol{x}}_k$ in the mean square sense for the asymptotically stationary case (cf., e. g., [6, pp. 485-486] or [31, Section 3.2]). The estimation error is

$$\tilde{\boldsymbol{u}}_k = \boldsymbol{u}_k - \hat{\boldsymbol{u}}_k, \qquad k \in \mathbb{N}_0, \tag{3.25}$$

[2]Noting that $\hat{\boldsymbol{u}}_k$ is the optimal estimate of $\boldsymbol{u}_k$ given $\hat{\boldsymbol{x}}_k$, it is not necessary to assume the existence of the inverse. Nevertheless, this facilitates the following derivations.

which is (asymptotically) uncorrelated with the estimate $\hat{\boldsymbol{u}}_k$, i. e.,

$$\boldsymbol{C}_{\boldsymbol{u}}^{*} = \boldsymbol{C}_{\hat{\boldsymbol{u}}}^{*} + \boldsymbol{C}_{\tilde{\boldsymbol{u}}}^{*}, \tag{3.26}$$

as well as with the state estimate $\hat{\boldsymbol{x}}_k$, i. e.,

$$\boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{*} = \boldsymbol{C}_{\hat{\boldsymbol{x}},\hat{\boldsymbol{u}}}^{*}. \tag{3.27}$$

Using these results, the optimal joint asymptotic covariance matrix of $\hat{\boldsymbol{x}}_k$ and $\boldsymbol{u}_k$ reads as

$$\begin{bmatrix} \boldsymbol{C}_{\hat{\boldsymbol{x}}}^{*} & \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{*} \\ \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{*,\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{u}}^{*} \end{bmatrix} = \begin{bmatrix} \boldsymbol{C}_{\hat{\boldsymbol{x}}}^{*} & \boldsymbol{C}_{\hat{\boldsymbol{x}}}^{*}\boldsymbol{L}^{\mathrm{T}} \\ \boldsymbol{L}\boldsymbol{C}_{\hat{\boldsymbol{x}}}^{*} & \boldsymbol{L}\boldsymbol{C}_{\hat{\boldsymbol{x}}}^{*}\boldsymbol{L}^{\mathrm{T}} + \boldsymbol{C}_{\tilde{\boldsymbol{u}}}^{*} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{N_x} \\ \boldsymbol{L} \end{bmatrix} \boldsymbol{C}_{\hat{\boldsymbol{x}}}^{*} \begin{bmatrix} \mathbf{I}_{N_x} \\ \boldsymbol{L} \end{bmatrix}^{\mathrm{T}} + \begin{bmatrix} \mathbf{0}_{N_x \times N_x} & \\ & \boldsymbol{C}_{\tilde{\boldsymbol{u}}}^{*} \end{bmatrix}, \tag{3.28}$$

and the optimal value of the cost function (cf. optimization problem 3.22) is given by

$$J^{*} = \mathrm{tr}\left[\begin{bmatrix} \mathbf{I}_{N_x} \\ \boldsymbol{L} \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{S} \\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{R} \end{bmatrix} \begin{bmatrix} \mathbf{I}_{N_x} \\ \boldsymbol{L} \end{bmatrix} \boldsymbol{C}_{\hat{\boldsymbol{x}}}^{*} \right] + \mathrm{tr}\left[\boldsymbol{Q}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{*} \right] + \mathrm{tr}\left[\boldsymbol{R}\boldsymbol{C}_{\tilde{\boldsymbol{u}}}^{*} \right]. \tag{3.29}$$

Additionally, the equality constraint of problem (3.22) can be rewritten as

$$\boldsymbol{C}_{\hat{\boldsymbol{x}}}^{*} = (\boldsymbol{A} + \boldsymbol{B}\boldsymbol{L})\,\boldsymbol{C}_{\hat{\boldsymbol{x}}}^{*}\,(\boldsymbol{A} + \boldsymbol{B}\boldsymbol{L})^{\mathrm{T}} + \boldsymbol{B}\boldsymbol{C}_{\tilde{\boldsymbol{u}}}^{*}\boldsymbol{B}^{\mathrm{T}} + \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}. \tag{3.30}$$

Since Equation (3.30) is a Lyapunov equation, we conclude that all eigenvalues of $\boldsymbol{A} + \boldsymbol{B}\boldsymbol{L}$ have magnitude less than one in order to obtain a positive semidefinite solution $\boldsymbol{C}_{\hat{\boldsymbol{x}}}^{*}$.[3]

It is now easy to verify that the solution of optimization problem (3.22) has the property

$$\boldsymbol{C}_{\tilde{\boldsymbol{u}}}^{*} = \mathbf{0}_{N_u \times N_u}. \tag{3.31}$$

As a first step, note that the solution $\boldsymbol{C}_{\hat{\boldsymbol{x}}}$ of Equation (3.15) is smaller (in the positive semidefinite sense) than $\boldsymbol{C}_{\hat{\boldsymbol{x}}}^{*}$ if $\hat{\boldsymbol{u}}_k$ from Equation (3.23) instead of $\boldsymbol{u}_k$ is applied to the dynamical system since $\boldsymbol{C}_{\tilde{\boldsymbol{u}}}^{*}$ is not present. It is obvious from Equation (3.26) that this holds analogously for the covariance matrix $\boldsymbol{C}_{\hat{\boldsymbol{u}}}^{*}$ compared to $\boldsymbol{C}_{\boldsymbol{u}}^{*}$. Thus, if the transmit power constraints are fulfilled by the optimizing covariance matrices $\boldsymbol{C}_{\hat{\boldsymbol{x}}}^{*}$ and $\boldsymbol{C}_{\boldsymbol{u}}^{*}$, they also hold for the matrices which are obtained by the application of the control signal $\hat{\boldsymbol{u}}_k$ shown in Equation (3.23). Additionally, the value J^{*} of the cost function shown in Equation (3.29) decreases in this case due to the decrease (in the positive semidefinite sense) of $\boldsymbol{C}_{\hat{\boldsymbol{x}}}$ compared to $\boldsymbol{C}_{\hat{\boldsymbol{x}}}^{*}$ and since the positive summand $\mathrm{tr}\left[\boldsymbol{R}\boldsymbol{C}_{\tilde{\boldsymbol{u}}}^{*}\right]$ can be dropped. Consequently, the assumption of $\boldsymbol{C}_{\tilde{\boldsymbol{u}}}^{*} \neq \mathbf{0}_{N_u \times N_u}$ leads to a contradiction since $\boldsymbol{C}_{\hat{\boldsymbol{x}}}^{*}$, $\boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{*}$ and $\boldsymbol{C}_{\boldsymbol{u}}^{*}$ can not be optimal if there exist a controller, given by Equation (3.23), which fulfills the power constraints and at the same time results in a reduction of the cost function. It follows that the optimizing matrices lead to Equation (3.31) and thus a control signal which is a linear function of the optimal state estimate.

Remark: The case of $\boldsymbol{C}_{\tilde{\boldsymbol{u}}}^{*} \neq \mathbf{0}_{N_u \times N_u}$ can be interpreted as a controller which artificially adds independent noise to the control input given by Equation (3.23). It is intuitively clear that this leads to an increase of the transmit power of the control signal and an increase of the variance of the system state which again leads to a larger transmit power for the observation channel and to a larger cost.

[3] The solution $\boldsymbol{C}_{\hat{\boldsymbol{x}}}^{*}$ is positive semidefinite due to the constraint of the optimization problem. Additionally, the matrix $\boldsymbol{\Phi} = \boldsymbol{B}\boldsymbol{C}_{\tilde{\boldsymbol{u}}}^{*}\boldsymbol{B}^{\mathrm{T}} + \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}$ is positive semidefinite. Assume now that $(\boldsymbol{A} + \boldsymbol{B}\boldsymbol{L})^{\mathrm{T}}$ has a real eigenvalue $|\lambda| > 1$ with corresponding eigenvector $\boldsymbol{z}$. Multiplying Equation (3.30) with $\boldsymbol{z}^{\mathrm{T}}$ from the left and $\boldsymbol{z}$ from the right results in $\boldsymbol{z}^{\mathrm{T}}\boldsymbol{C}_{\hat{\boldsymbol{x}}}^{*}\boldsymbol{z} = \lambda^{2}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{C}_{\hat{\boldsymbol{x}}}^{*}\boldsymbol{z} + \boldsymbol{z}^{\mathrm{T}}\boldsymbol{\Phi}\boldsymbol{z}$. This equality can not be fulfilled for $|\lambda| > 1$ which leads to a contradiction. The case of complex eigenvalues can be treated similarly.

The presented approach is essentially the one shown in [88], where additionally the finite horizon case is considered. The author of [88] formulates the dual of the constrained optimization problem in order to obtain the solution. This is not done here in order to keep things simple and since the focus of this section is not to develop algorithms for the determination of the optimal controller. Instead, a compact formulation for the problem has been chosen which allows for a direct application of general purpose software for semidefinite optimization for the numerical evaluation of power constrained LQG problems. Finally, note that the derivation of the optimal control sequence $(\boldsymbol{u}_k : k \in \mathbb{N}_0)$ from the joint covariance matrix of the state estimate and the control signal is different to the approach shown in [88] and relies on estimation theoretic arguments.

3.3 Feasibility

In the preceding section, a convex optimization problem for the determination of the LQG controller with power constraints has been presented in terms of the joint covariance matrix of the system state or its estimate, respectively, and the control input (see problem 3.22). Additionally, given the optimal joint covariance matrix, it has been shown how to determine the actual control sequence $(\boldsymbol{u}_k : k \in \mathbb{N}_0)$ (see Equations 3.23 and 3.24). Nevertheless, until now no statement has been made about the existence of such a solution. The critical point which may prevent the feasibility of optimization problem (3.22) are the power constraints, given by the constants $P_{\mathrm{Tx},1}$ and $P_{\mathrm{Tx},2}$. Without these constraints, the problem is always feasible if the system to be controlled is stabilizable and detectable. In this case, the solution is the standard LQG controller.

It is possible to construct a very simple example where the constrained LQG problem is not feasible. Assume that the system to be controlled is unstable, i. e., the matrix $\boldsymbol{A}$ has eigenvalues with magnitude larger than one, and that $P_{\mathrm{Tx},2} = 0$. In this case, the control sequence $(\boldsymbol{u}_k : k \in \mathbb{N}_0)$ must also be zero because its power, which is non-negative, is restricted to be zero.[4] With no control input, the system remains unstable which results in an unbounded covariance matrix of the system state.[5] Note that this example can still be feasible for stable systems since they have a bounded asymptotic covariance matrix of the system state even without a control input. This covariance matrix determines the power of the signal at the system output if no control is applied. If this power is smaller than $P_{\mathrm{Tx},1}$, the problem is feasible, otherwise not.

The discussion above shows that there is a clear distinction between stable and unstable systems for the determination of the feasibility of the constrained LQG control problem. Thus, these two cases are discussed separately in the following. Nevertheless, in both cases a closed form expression for the set of feasible values of $(P_{\mathrm{Tx},1}, P_{\mathrm{Tx},2})$ is not available in general which makes it necessary to determine this set numerically. The corresponding problem to be solved is the determination of the smallest values of transmit powers which allow for a solution of optimization problem (3.22). The simple reason is that if the problem is feasible for given values of $P_{\mathrm{Tx},1}$ and $P_{\mathrm{Tx},2}$, it is obviously also feasible if one or both powers are increased. Thus, transmit powers have to be determined which can not be further decreased.

In order to formalize the problem, we notice that the value of the cost function of problem (3.22) is not important for the question about minimal transmit powers as long as it is finite. Thus,

[4] We do not consider the case that $\boldsymbol{u}_k$, $k \in \mathbb{N}_0$, with variance zero can take values from a set with probability zero.

[5] More technically, the equality constraint which determines the asymptotic covariance matrix of the state estimate can not be fulfilled with a positive semidefinite matrix.

only the equality constraint which describes the system dynamics, i. e.,

$$\boldsymbol{C}_{\hat{\boldsymbol{x}}} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \end{bmatrix} \begin{bmatrix} \boldsymbol{C}_{\hat{\boldsymbol{x}}} & \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}} \\ \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{u}} \end{bmatrix} \begin{bmatrix} \boldsymbol{A}^{\mathrm{T}} \\ \boldsymbol{B}^{\mathrm{T}} \end{bmatrix} + \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}, \tag{3.32}$$

and the semidefiniteness constraint of the original LQG control problem have to be fulfilled for the determination of the minimal values of the transmit powers P_1 and P_2, respectively, which can be achieved by an LQG controller without the consideration of an additional performance criterion. These minimal values determine the lower bounds for $P_{\mathrm{Tx},1}$ and $P_{\mathrm{Tx},2}$ such that the general power constrained LQG problem (see optimization problem 3.22) is feasible. Unfortunately, the transmit powers of the observation and the control channel are coupled by the joint distribution of the system state and the control signal which can be seen by Equation (3.32). Thus, it can not be expected in general that there exists a minimum[6] pair (P_1, P_2) of achievable transmit powers in the sense that all other achievable pairs of available transmit powers are larger than (P_1, P_2) in both components. For example, it is intuitive that a smaller transmit power for the control channel will lead to an increase of the variance of the system state due to the reduced control effort and thus to a larger power of the observations at the system output. But despite the fact that it is in general not possible to determine a pair of achievable transmit powers (P_1, P_2) such that all other achievable pairs are *larger in both* components, it is possible to determine pairs (P_1, P_2) of achievable transmit powers such that there exist no pairs which are *smaller in both* components. Such values are called *Pareto optimal* [94, Section 4.7.3]. They have the property that, loosely speaking, it is not possible to decrease the value of one component of (P_1, P_2) without the necessity to increase the other one. The set $\mathcal{P}$ of Pareto optimal values of (P_1, P_2) determines the boundary of the set of feasible transmit powers $(P_{\mathrm{Tx},1}, P_{\mathrm{Tx},2})$ for the constrained LQG control problem (3.22). Thus, the optimization problem is feasible if the set $\mathcal{P}$ contains at least one pair of achievable transmit powers which is smaller in both components than (or equal to) the pair $(P_{\mathrm{Tx},1}, P_{\mathrm{Tx},2})$ of transmit powers which are available for the actual optimization problem.

One possible approach to the determination of Pareto optimal values of transmit powers is the so-called *scalarization* (see [94, Section 4.7.4]) of the problem of jointly minimizing P_1 and P_2, i. e., the transmit power of the observation channel and the control channel, subject to the system dynamics. Instead of the joint minimization, a weighted sum of the transmit powers is considered by the optimization problem

$$\begin{aligned} \underset{\boldsymbol{C}_{\boldsymbol{x}},\boldsymbol{C}_{\hat{\boldsymbol{x}}},\boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}},\boldsymbol{C}_{\boldsymbol{u}}}{\text{minimize}} \quad & \rho\, \mathrm{tr}\left[\boldsymbol{C}\boldsymbol{C}_{\boldsymbol{x}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}}\right] + (1-\rho)\,\mathrm{tr}\left[\boldsymbol{C}_{\boldsymbol{u}}\right] \\ \text{subject to} \quad & \boldsymbol{C}_{\hat{\boldsymbol{x}}} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \end{bmatrix} \begin{bmatrix} \boldsymbol{C}_{\hat{\boldsymbol{x}}} & \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}} \\ \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{u}} \end{bmatrix} \begin{bmatrix} \boldsymbol{A}^{\mathrm{T}} \\ \boldsymbol{B}^{\mathrm{T}} \end{bmatrix} + \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}, \\ & \boldsymbol{C}_{\boldsymbol{x}} = \boldsymbol{C}_{\hat{\boldsymbol{x}}} + \boldsymbol{C}_{\tilde{\boldsymbol{x}}}, \\ & \begin{bmatrix} \boldsymbol{C}_{\hat{\boldsymbol{x}}} & \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}} \\ \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{u}} \end{bmatrix} \geq \mathbf{0}_{(N_x+N_u)\times(N_x+N_u)}, \end{aligned} \tag{3.33}$$

with the weighting factor $\rho \in [0, 1]$ and where Definition 3.1.1 has been used for the transmit powers P_1 and P_2. Note that the minimization above fits into the framework of unconstrained LQG optimization problems (see Section 3.1 and Equation 3.19) by choosing $\boldsymbol{Q} = \rho\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}$, $\boldsymbol{R} = (1-\rho)\mathbf{I}_{N_u}$ and $\boldsymbol{S} = \mathbf{0}_{N_x \times N_u}$. A problem formulation using both $\boldsymbol{C}_{\boldsymbol{x}}$ and $\boldsymbol{C}_{\hat{\boldsymbol{x}}}$, although redundant, has been

[6]For a formal definition of minimum and minimal points see, e. g., [94, Section 2.4.2].

chosen in order to make the minimization of $P_1 = \mathrm{tr}\left[\boldsymbol{C}\boldsymbol{C}_{\boldsymbol{x}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}}\right]$ explicit. The reason for the introduction of $\boldsymbol{C}_{\hat{\boldsymbol{x}}}$ is explained in Section 3.2.

In the following, the minimizing covariance matrices for a specific value of ρ in the optimization problem (3.33) are denoted by $\boldsymbol{C}_{\boldsymbol{x}}^{(\rho)}$, $\boldsymbol{C}_{\hat{\boldsymbol{x}}}^{(\rho)}$, $\boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{(\rho)}$ and $\boldsymbol{C}_{\boldsymbol{u}}^{(\rho)}$, respectively. Choosing $\rho = 1$ in problem (3.33) results in the smallest value of P_1 without taking into account the value of P_2 and thus in the individual lower bound for $P_{\mathrm{Tx},1}$. For $\rho = 0$, we get the analogous result for P_2 and $P_{\mathrm{Tx},2}$. The trade-off between the decrease of one transmit power and the corresponding increase of the other is performed by varying the value of ρ between zero and one. Putting a larger weight on the power of the observation channel will typically result in a lower value of this power at the expense of a larger power of the control signal, and vice versa. Thus, different elements of the set $\mathcal{P}$ of Pareto optimal values are determined by the optimization problem (3.33) using different values for ρ. Note that in our case, all Pareto optimal values can be determined by varying $\rho \in [0, 1]$ since both transmit powers P_1 and P_2 as well as the constraints of problem (3.33) are convex functions of the optimization variables (see, e. g., [94, pp. 179-180]). This also leads to the fact that the trade-off curve of Pareto optimal values of the transmit powers is convex (cf. Appendix A8).

3.3.1 Stable Systems

It has already been mentioned that a fundamental difference between stable and unstable dynamical systems is that the latter ones require a strictly positive value of $P_{\mathrm{Tx},2}$ because they always need some stabilizing control input. Since the former ones are stable even in the case of no control, we expect that the lowest possible value of $P_{\mathrm{Tx},2}$ which is feasible is zero. In fact this is true which can be verified by considering optimization problem (3.33) with $\rho = 0$. The minimal value of the cost function is zero by choosing $\boldsymbol{C}_{\boldsymbol{u}}^{(0)} = \boldsymbol{0}_{N_u \times N_u}$. The constraints are fulfilled with $\boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{(0)} = \boldsymbol{0}_{N_x \times N_u}$ and the covariance matrix of the state estimate which is determined by the Lyapunov equation

$$\boldsymbol{C}_{\hat{\boldsymbol{x}}}^{(0)} = \boldsymbol{A}\boldsymbol{C}_{\hat{\boldsymbol{x}}}^{(0)}\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{C}_{\hat{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}. \tag{3.34}$$

This matrix is positive semidefinite since all eigenvalues of $\boldsymbol{A}$ have a magnitude less than one. The corresponding covariance matrix of the system state is determined by

$$\boldsymbol{C}_{\boldsymbol{x}}^{(0)} = \boldsymbol{A}\boldsymbol{C}_{\boldsymbol{x}}^{(0)}\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{w}}, \tag{3.35}$$

which can be verified using the system equation $\boldsymbol{x}_{k+1} = \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{B}\boldsymbol{n}_k + \boldsymbol{w}_k$ when the control input is zero. Note that the corresponding transmit power P_1 of the observation channel is larger than zero, i. e., even if the controller is effectively switched of by setting $P_{\mathrm{Tx},2}$ to zero, it is still necessary to provide a positive amount of power for the transmission of the observations at the system output to the controller. This behavior results from the system model which offers no possibility to stop the transmission of observations in this case.

Example 3.3.1 In order to illustrate the region of feasible transmit powers, the parameters given in Example 3.1.1 are used, but the system matrix $\boldsymbol{A}$ is replaced by $\boldsymbol{A}_{\text{stable}} = \frac{1}{3}\boldsymbol{A}$, which has only eigenvalues with magnitude less than one. The optimization problem (3.33) is solved for 200 values of ρ which are equidistantly sampled from the interval $[0, 1]$. The resulting transmit powers $P_1 = \mathrm{tr}\left[\boldsymbol{C}\boldsymbol{C}_{\boldsymbol{x}}^{(\rho)}\boldsymbol{C}^{\mathrm{T}}\right] + \mathrm{tr}\left[\boldsymbol{C}_{\boldsymbol{v}}\right]$ and $P_2 = \mathrm{tr}\left[\boldsymbol{C}_{\boldsymbol{u}}^{(\rho)}\right]$ are shown in Figure 3.3.

It can be observed that the trade-off curve of Pareto optimal values of the two transmit powers is convex and that the minimal value of P_2 is zero which corresponds to the case where the controller

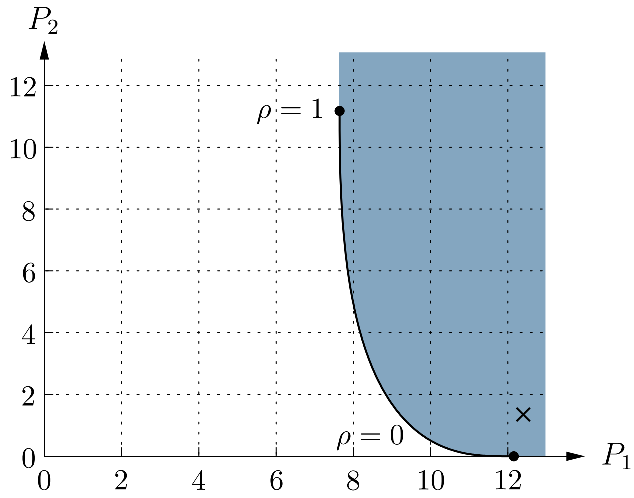


Figure 3.3: The set of Pareto optimal values (solid line) of transmit powers for the observation and control channel and feasible values (shaded area) for the constrained LQG control problem. The values of the transmit powers for $\rho = 0$ and $\rho = 1$ are denoted by •, the transmit powers for the unconstrained LQG problem by $\times$.

is switched off. The shaded area shows all values of $P_{\mathrm{Tx},1}$ and $P_{\mathrm{Tx},2}$, respectively, which result in a feasible power constrained LQG control problem (cf. optimization problem 3.22).

Besides the illustration of the results of this section, Example 3.3.1 shows one fundamental shortcoming of the presented approach to the consideration of power constraints in LQG control. Even if the controller is switched off, i. e., when $P_{\mathrm{Tx},2} = 0$ and consequently $P_2 = 0$, the observations at the system output are still transmitted to the controller which results in a strictly positive transmit power for the observation channel. This power is wasted since the controller does not use the transmitted information. An obvious step would be to stop the transmission of observations in this case which corresponds to a transmit power of zero. Unfortunately, the system model does not allow for such a step because there is no degree of freedom at the system output to control the transmission. This degree of freedom, the transmitter, will be considered in Chapter 4.

3.3.2 Unstable Systems

The discussion about the feasibility of the power constrained LQG control problem (3.22) for unstable dynamical systems is in many ways analogous to the one concerning stable systems. Due to the coupling of the transmit powers of the observation and the control channel (cf. Equation 3.32), it is again necessary to resort to the optimization problem (3.33) in order to determine Pareto optimal values of transmit powers. Those values represent the boundary of the set of feasible values of $(P_{\mathrm{Tx},1}, P_{\mathrm{Tx},2})$ for the power constraints.

The fundamental difference between the consideration of stable and unstable dynamical systems is that for the latter ones, the values of $P_{\mathrm{Tx},1}$ and $P_{\mathrm{Tx},2}$ which are the available transmit powers of the observation and the control channel must *both* be strictly positive. The intuitive reason is the fact that unstable systems always need a stabilizing control input in order to have a bounded covariance matrix of the system state. Therefore, the power of the control channel must be larger than zero. Additionally, the power of the observation channel must be larger than zero in order to provide the controller with information about the system state which is needed to stabilize the system. It can be seen that the information exchange between the dynamical system and the controller in both directions is essential for the task of stabilization, which leads to the fact that all transmit powers in the closed loop system have to be strictly positive.

Example 3.3.2 Example 3.3.1 is revisited at this point with the original system matrix from Example 3.1.1, i. e., with an unstable dynamical system. For the determination of Pareto optimal values of the transmit powers for the observation and the control channel, 1000 values of the weighting factor ρ of optimization problem (3.33) are sampled equidistantly from the interval $\rho \in [0, 1]$. Note that for the given parameters of the dynamical system and the noise sequences, the resulting transmit powers vary over more than two orders of magnitude. Thus, a logarithmic scale has been chosen in Figure 3.4 for the presentation of the results. Finally, since it is a common quantity in communication systems, not the transmit power but the SNR is shown in this figure, i. e.,

$$\varphi_1 = \operatorname{tr}\left[\boldsymbol{C_q}\right]^{-1} P_1 \qquad \text{and} \qquad \varphi_2 = \operatorname{tr}\left[\boldsymbol{C_n}\right]^{-1} P_2,$$

where $P_1 = \operatorname{tr}\left[\boldsymbol{C}\boldsymbol{C}_{\boldsymbol{x}}^{(\rho)}\boldsymbol{C}^{\mathrm{T}}\right] + \operatorname{tr}\left[\boldsymbol{C_v}\right]$ and $P_2 = \operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{u}}^{(\rho)}\right]$ are given by the solution of the optimization problem (3.33) for a given weighting factor ρ. The SNRs φ_1 and φ_2 which correspond to Pareto optimal values of the transmit powers P_1 and P_2 are also Pareto optimal. This can be seen by noting that the SNRs and transmit powers are proportional, which means that the minimization of the weighted sum of transmit powers with weighting factor ρ is equivalent to the minimization of the weighted sum of SNRs with weighting factor $\rho^{(\mathrm{SNR})}$ which is a function of ρ, $\operatorname{tr}\left[\boldsymbol{C_q}\right]$ and $\operatorname{tr}\left[\boldsymbol{C_n}\right]$.[7] As in Example 3.3.1, the set of Pareto optimal values of transmit powers or SNRs, respectively, represents the boundary of feasible values for the constrained LQG control problem (3.22).

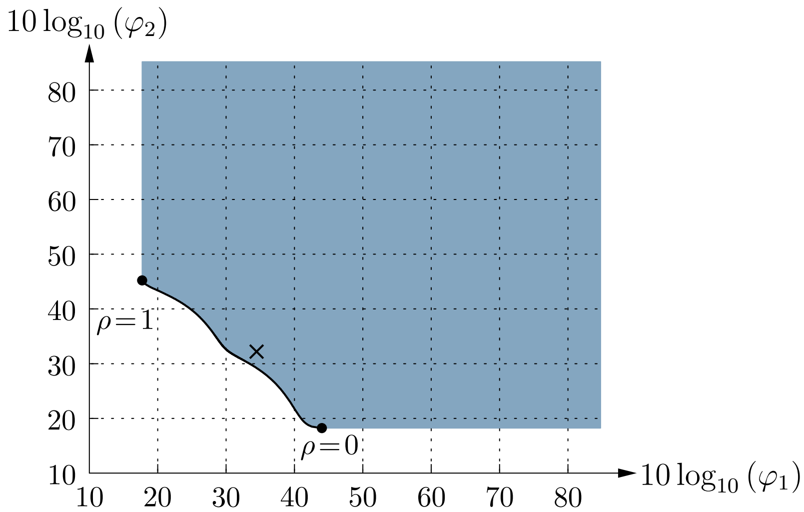


Figure 3.4: The set of Pareto optimal values (solid line) of SNRs for the observation and control channel and feasible values (shaded area) for the constrained LQG control problem. The values of the SNRs for $\rho = 0$ and $\rho = 1$ are denoted by $\bullet$, the point of SNRs for the unconstrained LQG solution is denoted by $\times$.

[7] In fact, $\rho^{(\mathrm{SNR})} = \left(1 - \rho + \frac{\operatorname{tr}[\boldsymbol{C_q}]}{\operatorname{tr}[\boldsymbol{C_n}]}\rho\right)^{-1} \frac{\operatorname{tr}[\boldsymbol{C_q}]}{\operatorname{tr}[\boldsymbol{C_n}]}\rho$.

Note that the trade-off curve of Pareto optimal values in Figure 3.4 does not seem to be convex due to the logarithmic scales. In fact, it is convex, which is ensured by the result shown in Appendix A8 and confirmed by Figure 3.5 which shows a detail of Figure 3.4 with linear scale.

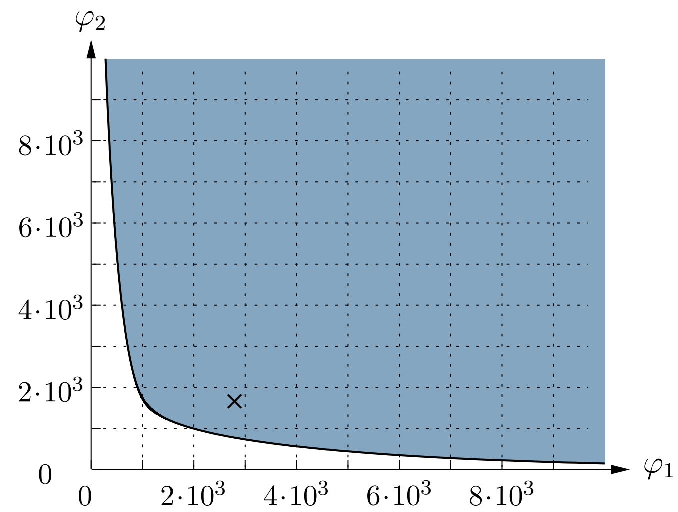


Figure 3.5: Detail of Figure 3.4 with linear scale where $10\log_{10}(\varphi_1)$ and $10\log_{10}(\varphi_2)$ are in the interval $[10, 40]$.

3.4 Solution of the Optimization Problem

The solution of the power constrained LQG control problem (3.22) is, at least from a theoretical point of view, not a problem since it can be determined numerically with standard software packages due to the convexity of the optimization problem. Thus, in the preceding section, the question about the feasibility of the problem (3.22), i. e., about the transmit powers $P_{\mathrm{Tx},1}$ and $P_{\mathrm{Tx},2}$ that have to be provided for the transmission of observations and control signals, has been treated. In order to answer this question, Pareto optimal values of transmit powers have been determined since no pairs (P_1, P_2) can be realized which are smaller than those values in both components. Larger values of $P_{\mathrm{Tx},1}$ and $P_{\mathrm{Tx},2}$ always lead to a feasible problem due to the existence of a controller, determined by the optimization problem (3.33), which realizes Pareto optimal values of transmit powers that obviously fulfill those power constraints.

Since the solution of the power constrained LQG problem is in principle obtained relatively easy, it is of interest to have a closer look on some of its properties. Aspects which have not been discussed so far are the influence of the transmit power constraints on the value of the LQG cost function in problem (3.22) and the set of transmit powers that can actually be achieved by the optimization of the controller for the considered scenario (see Figure 3.1). One important observation has already been made in Example 3.3.1 where we saw that even for the case of a stable dynamical system where the controller is switched of by setting $P_{\mathrm{Tx},2} = 0$, the corresponding transmit power P_1 of the observation channel which is limited by $P_{\mathrm{Tx},1}$ is larger than zero. Thus, even not used for any control action, observations of the system state are transmitted from the system output to the controller and the corresponding transmit power is wasted. The reason for this behavior is that the controller is the only entity which has an influence on the system performance, represented by the LQG cost function, and on the transmit powers for the control channel and the observation channel. The former power is directly determined by the choice of the control signal while the latter one is the result of the response of the dynamical system to the control action. Thus, it is the task of the controller alone to optimize the system performance while simultaneously taking into account the communication restrictions, i. e., the transmit power constraints. We will see later that this does not only lead to the fact that transmit power may be wasted, but also that an increase of transmit power does not always lead to an increase in system performance.

3.4.1 Inequality Constraints and Achievable Transmit Powers

Recall the power constrained LQG problem which has already been presented in Section 3.2:

$$
\begin{aligned}
\underset{\boldsymbol{C}_{\hat{\boldsymbol{x}}},\boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}},\boldsymbol{C}_{\boldsymbol{u}}}{\text{minimize}}\ & \operatorname{tr}\left[\begin{bmatrix}\boldsymbol{Q} & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{R}\end{bmatrix}\begin{bmatrix}\boldsymbol{C}_{\hat{\boldsymbol{x}}} & \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}\\ \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{u}}\end{bmatrix}\right] + \operatorname{tr}\left[\boldsymbol{Q}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\right] \qquad (3.36)\\
\text{subject to}\ & \operatorname{tr}\left[\boldsymbol{C}\boldsymbol{C}_{\hat{\boldsymbol{x}}}\boldsymbol{C}^{\mathrm{T}}+\boldsymbol{C}_{\boldsymbol{v}}\right] + \operatorname{tr}\left[\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\boldsymbol{C}^{\mathrm{T}}\right] \le P_{\mathrm{Tx},1},\\
& \operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{u}}\right] \le P_{\mathrm{Tx},2},\\
& \boldsymbol{C}_{\hat{\boldsymbol{x}}} = \begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\end{bmatrix}\begin{bmatrix}\boldsymbol{C}_{\hat{\boldsymbol{x}}} & \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}\\ \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{u}}\end{bmatrix}\begin{bmatrix}\boldsymbol{A}^{\mathrm{T}}\\ \boldsymbol{B}^{\mathrm{T}}\end{bmatrix} + \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}},\\
& \begin{bmatrix}\boldsymbol{C}_{\hat{\boldsymbol{x}}} & \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}\\ \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{u}}\end{bmatrix} \ge \boldsymbol{0}_{(N_x+N_u)\times(N_x+N_u)}.
\end{aligned}
$$

In Section 3.3 it has been shown that the transmit powers P_1 and P_2 of the observation and control channel can not be made arbitrarily small and thus the optimization problem above is only feasible if $P_{\mathrm{Tx},1}$ and $P_{\mathrm{Tx},2}$ are large enough. But the question which transmit powers can be achieved by a controller that is determined by optimization problem (3.36) is not only interesting for small values of $P_{\mathrm{Tx},1}$ and $P_{\mathrm{Tx},2}$ but also for large ones. In order to understand this, consider the optimization of a controller using the LQG approach without any power constraints. The solution leads to specific values of P_1 and P_2, respectively, which have been determined, e. g., in Example 3.1.1 (cf. Equation 3.16) and are depicted in Figures 3.2, 3.4 and 3.5. But if $P_{\mathrm{Tx},1}$ and $P_{\mathrm{Tx},2}$ are larger than the corresponding values of P_1 and P_2 of the unconstrained optimization, the transmit power constraints are not active. In other words, if more transmit power is provided than needed by the unconstrained LQG controller, it will not be used. From the perspective of a communication system designer this is an unsatisfactory result because additional communication resources can be used to improve the quality of the transmitted information. In our case, the additional transmit power could be used to reduce the negative effect of the channel noise. Nevertheless, this is obviously not possible for the present scenario.

It has already been mentioned earlier that the transmit powers of the observation and control channel are coupled by the state equation shown in Equation (3.32). Thus, it seems reasonable that the set of unachievable transmit powers is not only determined by power constraints which are too restrictive, but also by constraints which require the transmit power of one communication channel to be small while the other channel is provided with a large amount of transmit power. The following example illustrates this intuition.

Example 3.4.1 The parameters of the dynamical system to be controlled, the noise sequences and the LQG cost function are again given by Example 3.1.1. Without any power constraints (or equivalently, for $P_{\text{Tx},1} \to \infty$ and $P_{\text{Tx},2} \to \infty$), the optimization problem (3.36) results in the standard LQG controller with the associated transmit powers (cf. Figure 3.2). Let these to powers be $P_{\text{LQG},1}$ for the observation channel and $P_{\text{LQG},2}$ for the control channel. Since the associated solution results in the minimal value of the LQG cost function, any power constraints where both values of $P_{\text{Tx},1}$ and $P_{\text{Tx},2}$ are larger than $P_{\text{LQG},1}$ and $P_{\text{LQG},2}$, respectively, will not be active. The additional transmit power can not be used to decrease the value of the cost function.

In order to evaluate the largest values of the transmit powers which are the result of a constrained optimization of the LQG cost function, the following variants of the original problem (3.36) are considered:

$$\begin{aligned} \underset{\boldsymbol{C}_{\hat{\boldsymbol{x}}},\boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}},\boldsymbol{C}_{\boldsymbol{u}}}{\text{minimize}} \quad & \operatorname{tr}\left[\begin{bmatrix}\boldsymbol{Q} & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{R}\end{bmatrix}\begin{bmatrix}\boldsymbol{C}_{\hat{\boldsymbol{x}}} & \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}\\ \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{u}}\end{bmatrix}\right] + \operatorname{tr}\left[\boldsymbol{Q}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\right] \\ \text{subject to} \quad & \operatorname{tr}\left[\boldsymbol{C}\boldsymbol{C}_{\hat{\boldsymbol{x}}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}}\right] + \operatorname{tr}\left[\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\boldsymbol{C}^{\mathrm{T}}\right] \leq P_{\text{Tx},1}, \\ & \boldsymbol{C}_{\hat{\boldsymbol{x}}} = \begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\end{bmatrix}\begin{bmatrix}\boldsymbol{C}_{\hat{\boldsymbol{x}}} & \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}\\ \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{u}}\end{bmatrix}\begin{bmatrix}\boldsymbol{A}^{\mathrm{T}}\\ \boldsymbol{B}^{\mathrm{T}}\end{bmatrix} + \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}, \\ & \begin{bmatrix}\boldsymbol{C}_{\hat{\boldsymbol{x}}} & \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}\\ \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{u}}\end{bmatrix} \geq \mathbf{0}_{(N_x+N_u)\times(N_x+N_u)}, \end{aligned} \tag{3.37}$$

i. e., no constraint is imposed on the power of the control channel. Thus, the controller can use as much power as desired in order to minimize the value of the cost function while satisfying the power constraint of the observation channel. Analogously, dropping the constraint on the power of the observation channel results in the optimization problem

$$\begin{aligned} \underset{\boldsymbol{C}_{\hat{\boldsymbol{x}}},\boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}},\boldsymbol{C}_{\boldsymbol{u}}}{\text{minimize}} \quad & \operatorname{tr}\left[\begin{bmatrix}\boldsymbol{Q} & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{R}\end{bmatrix}\begin{bmatrix}\boldsymbol{C}_{\hat{\boldsymbol{x}}} & \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}\\ \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{u}}\end{bmatrix}\right] + \operatorname{tr}\left[\boldsymbol{Q}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\right] \\ \text{subject to} \quad & \operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{u}}\right] \leq P_{\text{Tx},2}, \\ & \boldsymbol{C}_{\hat{\boldsymbol{x}}} = \begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\end{bmatrix}\begin{bmatrix}\boldsymbol{C}_{\hat{\boldsymbol{x}}} & \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}\\ \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{u}}\end{bmatrix}\begin{bmatrix}\boldsymbol{A}^{\mathrm{T}}\\ \boldsymbol{B}^{\mathrm{T}}\end{bmatrix} + \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}, \\ & \begin{bmatrix}\boldsymbol{C}_{\hat{\boldsymbol{x}}} & \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}\\ \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{u}}\end{bmatrix} \geq \mathbf{0}_{(N_x+N_u)\times(N_x+N_u)}, \end{aligned} \tag{3.38}$$

which allows for an arbitrary amount of transmit power for the observation channel to minimize the cost function while satisfying the power constraint of the control channel.

For the optimization problem (3.37), the relevant interval $\mathcal{P}_1$ of values for $P_{\text{Tx},1}$ is between $P_{\text{LQG},1}$ and the lowest Pareto optimal value of P_1 which is determined by the optimization problem (3.33) for $\rho = 1$. Lower values of $P_{\text{Tx},1}$ obviously render the problem (3.37) infeasible, while values larger than $P_{\text{LQG},1}$ lead to an inactive power constraint. Analogously, the relevant interval $\mathcal{P}_2$ of values for $P_{\text{Tx},2}$ in the optimization problem (3.38) is between $P_{\text{LQG},2}$ and the Pareto optimal value of P_2 determined by the optimization (3.33) for $\rho = 0$. For the results which are shown in Figure 3.6, 200 logarithmically spaced values have been sampled from the intervals $\mathcal{P}_1$ and $\mathcal{P}_2$,

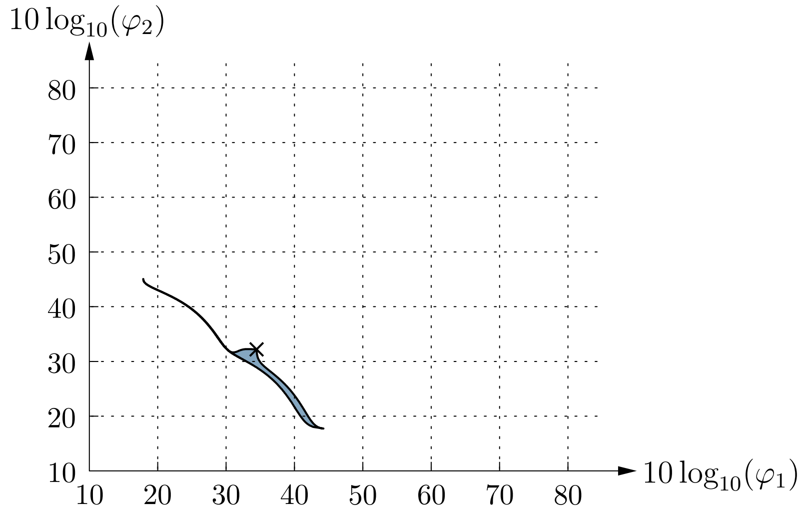


Figure 3.6: The SNR region for the inequality constrained LQG problem (3.22).

respectively. Additionally, the Pareto optimal values of the SNRs (see Figure 3.4) and the SNR point of the unconstrained LQG controller, denoted by $\times$, are shown.

It can be seen that the region of SNRs $\varphi_1 = \operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{q}}\right]^{-1} P_1$ and $\varphi_2 = \operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{n}}\right]^{-1} P_2$ which is the result of the optimization problem (3.22) between the lowest possible, Pareto optimal values and the maximal values when no or one power constraint is considered for the controller optimization is relatively small. Thus, the potential to improve the system performance by increasing the amount of available communication resources of one or both channels in the control loop is rather limited.

Example 3.4.1 demonstrates the inherent contradiction between the goal of minimizing the LQG cost function and of using the communication resource transmit power for the improvement of the value of the cost when the controller is the only entity which has an impact on all these quantities. Both the cost function and the transmit powers depend on the joint covariance matrix of the system state and the control signal (see problem 3.36). This means that the controller has only limited possibilities to reduce the value of the cost function while at the same time increasing the value of one or both transmit powers.

3.4.2 Equality Constraints and Shortcomings of the Solution

In the last section, we saw that the available transmit powers for the observation and the control channel, given by $P_{\mathrm{Tx},1}$ and $P_{\mathrm{Tx},2}$, may not be used for the optimization of the system performance and the corresponding inequality constraints of the optimization problem (3.36) are not active. An ad hoc approach to change this behavior is to replace the inequality constraints by equalities. This ensures that the available transmit power is used. Unfortunately, this is not a good idea in general since the modification can only lead to an increase of the LQG cost function, i. e., a decrease of system performance. The simple reason is that an inactive inequality constraint means that the optimal value of the cost function is achieved without using the full transmit power. If we force the constraint to hold with equality, the resulting solution can not exhibit a decrease of the optimal value of the cost function.

Despite the fact that it results in a worse system performance, we investigate the case of equality constraints for the transmit powers of the observation and the control channel in order to obtain some insights in the effects of power constraints on the LQG control problem. Until now, we only relied on the fact that the constrained optimization problem (3.36) is convex to make the statement that a solution exists and that the global optimum can be determined relatively easy using standard software packages [92, 93]. Nevertheless, we investigate some steps of the problem solution in more detail at this point. In order to keep things simple, only one power constraint is considered with equality while the other one is dropped. The equality constraint is reformulated as two inequality constraints, i. e., we consider the optimization problem

$$\begin{aligned} \underset{\boldsymbol{C}_{\hat{\boldsymbol{x}}},\boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}},\boldsymbol{C}_{\boldsymbol{u}}}{\text{minimize}} \quad & \operatorname{tr}\left[\begin{bmatrix}\boldsymbol{Q} & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{R}\end{bmatrix}\begin{bmatrix}\boldsymbol{C}_{\hat{\boldsymbol{x}}} & \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}\\ \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{u}}\end{bmatrix}\right] + \operatorname{tr}\left[\boldsymbol{Q}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\right] \\ \text{subject to} \quad & \operatorname{tr}\left[\boldsymbol{C}\boldsymbol{C}_{\hat{\boldsymbol{x}}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}}\right] + \operatorname{tr}\left[\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\boldsymbol{C}^{\mathrm{T}}\right] \leq P_{\mathrm{Tx},1}, \\ & \operatorname{tr}\left[\boldsymbol{C}\boldsymbol{C}_{\hat{\boldsymbol{x}}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}}\right] + \operatorname{tr}\left[\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\boldsymbol{C}^{\mathrm{T}}\right] \geq P_{\mathrm{Tx},1}, \\ & \boldsymbol{C}_{\hat{\boldsymbol{x}}} = \begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\end{bmatrix}\begin{bmatrix}\boldsymbol{C}_{\hat{\boldsymbol{x}}} & \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}\\ \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{u}}\end{bmatrix}\begin{bmatrix}\boldsymbol{A}^{\mathrm{T}}\\ \boldsymbol{B}^{\mathrm{T}}\end{bmatrix} + \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}, \\ & \begin{bmatrix}\boldsymbol{C}_{\hat{\boldsymbol{x}}} & \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}\\ \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{u}}\end{bmatrix} \succeq \boldsymbol{0}_{(N_x+N_u)\times(N_x+N_u)}. \end{aligned} \tag{3.39}$$

This problem is again convex and if it is feasible, the transmit power P_1 of the observation channel has the value $P_{\mathrm{Tx},1}$. To understand the effects of the power constraint on the LQG control problem, we use the method of Lagrangian duality (see, e. g., [94, Chapter 5] or [95, Chapter 6]) to include the power constraints in the cost function and thus obtain the optimization problem for the determination of the Lagrange dual function:

$$\begin{aligned} \underset{\boldsymbol{C}_{\hat{\boldsymbol{x}}},\boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}},\boldsymbol{C}_{\boldsymbol{u}}}{\text{minimize}} \quad & \operatorname{tr}\left[\begin{bmatrix}\left(\boldsymbol{Q} + (\lambda_1 - \lambda_2)\,\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}\right) & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{R}\end{bmatrix}\begin{bmatrix}\boldsymbol{C}_{\hat{\boldsymbol{x}}} & \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}\\ \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{u}}\end{bmatrix}\right] + \operatorname{tr}\left[\boldsymbol{Q}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\right] \\ & \qquad + (\lambda_1 - \lambda_2)\left(\operatorname{tr}\left[\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}}\right] - P_{\mathrm{Tx},1}\right) \\ \text{subject to} \quad & \boldsymbol{C}_{\hat{\boldsymbol{x}}} = \begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\end{bmatrix}\begin{bmatrix}\boldsymbol{C}_{\hat{\boldsymbol{x}}} & \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}\\ \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{u}}\end{bmatrix}\begin{bmatrix}\boldsymbol{A}^{\mathrm{T}}\\ \boldsymbol{B}^{\mathrm{T}}\end{bmatrix} + \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}, \\ & \begin{bmatrix}\boldsymbol{C}_{\hat{\boldsymbol{x}}} & \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}\\ \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{u}}\end{bmatrix} \succeq \boldsymbol{0}_{(N_x+N_u)\times(N_x+N_u)}, \end{aligned} \tag{3.40}$$

where $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ are the Lagrange multipliers associated with the two inequality constraints for P_1.[8] Note that for given values of the multipliers, the minimization above corresponds to an unconstrained LQG optimization (cf. problem 3.19), where the weighting matrix for the system state $\boldsymbol{Q}$ is replaced by the matrix $\boldsymbol{Q} + (\lambda_1 - \lambda_2)\,\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}$. This replacement is quite intuitive when considering the case that $\lambda_1 > 0$ and $\lambda_2 = 0$, i. e., the constraint which forces a larger value of P_1 to meet the more restrictive value $P_{\mathrm{Tx},1}$ is active. The loading of $\boldsymbol{Q}$ with the positive semidefinite matrix $\lambda_1\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}$ shifts the solution of (3.40) towards covariance matrices with a lower value of $P_1 = \operatorname{tr}\left[\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{C}_{\hat{\boldsymbol{x}}}\right] + \operatorname{tr}\left[\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}}\right]$. As long as the problem is feasible, the value of λ_1 is increased until the power constraint is met by the solution of the original problem (3.39).

[8] Obviously, the two non-negative multipliers can be combined to one real-valued one since the two inequality power constraints are effectively one equality constraint.

If the transmit power for the unconstrained LQG controller is smaller than $P_{\mathrm{Tx},1}$, the second inequality constraint of the optimization problem (3.39) is active, i. e., the Lagrange multiplier λ_1 of problem (3.40) is zero while $\lambda_2 > 0$. It can be seen that in order to increase the power of the transmit signal for the observation channel, the weighting matrix $\boldsymbol{Q}$ gets a *negative* loading with $-\lambda_2 \boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}$. Consequently, solutions of (3.39) are preferred which result in larger values of P_1. Note that the negative value of $\lambda_1 - \lambda_2$ can even lead to an indefinite weighting matrix $\boldsymbol{Q} + (\lambda_1 - \lambda_2)\, \boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}$. This is of special interest since for the indefinite case the cost function of the problem (3.40) which is linear in the optimization variables is not bounded from below in general for positive semidefinite covariance matrices of the state estimate and the control signal. The corresponding solution of the minimization is not useful due to the fact that the variance of at least one of the system states is not bounded from above in this case.

In the following, we investigate the case of $\lambda_1 = 0$ and $\lambda_2 > 0$, i. e., it is necessary to increase the transmit power P_1 to meet the value $P_{\mathrm{Tx},1}$ with equality. The goal is the determination of the largest value of λ_2 for which the minimum of the optimization problem (3.40) has a finite value. Additionally, a numerical method for the solution of this problem is presented. This maximal value limits the range of the optimal value of the dual variable λ_2. Similarly to the preceding parts of this chapter, the determination of the maximal value of λ_2 is formulated as a convex optimization problem. In order to keep the notation simple, we define the joint covariance matrix of the state estimate and the control signal as

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{C}_{\hat{\boldsymbol{x}}} & \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}} \\ \boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{u}}^{\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{u}} \end{bmatrix}. \tag{3.41}$$

Additionally, define the matrix $\boldsymbol{E} = [\mathbf{I}_{N_x},\ \mathbf{0}_{N_x \times N_u}] \in \{0,1\}^{N_x \times (N_x + N_u)}$. Consequently, the equality constraint of the optimization problem (3.40) can be rewritten as

$$\boldsymbol{E}\boldsymbol{X}\boldsymbol{E}^{\mathrm{T}} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \end{bmatrix} \boldsymbol{X} \begin{bmatrix} \boldsymbol{A}^{\mathrm{T}} \\ \boldsymbol{B}^{\mathrm{T}} \end{bmatrix} + \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}. \tag{3.42}$$

Note that if $\boldsymbol{X} \geq \mathbf{0}_{(N_x+N_u)\times(N_x+N_u)}$ fulfills Equation (3.42), i. e., $\boldsymbol{X}$ is a feasible matrix for the problem (3.40), the matrix $\boldsymbol{X} + \xi\boldsymbol{Y}$, where $\xi \geq 0$ and $\boldsymbol{Y} \geq \mathbf{0}_{(N_x+N_u)\times(N_x+N_u)}$ fulfills the equation

$$\boldsymbol{E}\boldsymbol{Y}\boldsymbol{E}^{\mathrm{T}} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \end{bmatrix} \boldsymbol{Y} \begin{bmatrix} \boldsymbol{A}^{\mathrm{T}} \\ \boldsymbol{B}^{\mathrm{T}} \end{bmatrix}, \tag{3.43}$$

is also feasible. This can easily be verified by adding Equation (3.42) and Equation (3.43) which is scaled by $\xi \geq 0$. Note that with the stabilizability assumption, besides the trivial all-zeros solution, Equation (3.43) has infinitely many solutions since it corresponds to the state equation of a closed loop system where the state of the system to be controlled is known to the controller.[9]

In order to determine the largest value of λ_2 such that the optimization problem (3.40) has a bounded solution, its cost function which is evaluated at a feasible matrix $\boldsymbol{X} + \xi\boldsymbol{Y}$ is rewritten as

$$\operatorname{tr}\left[\begin{bmatrix} (\boldsymbol{Q} - \lambda_2 \boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}) & \boldsymbol{S} \\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{R} \end{bmatrix} \boldsymbol{X}\right] + \xi \left(\operatorname{tr}\left[\begin{bmatrix} \boldsymbol{Q} & \boldsymbol{S} \\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{R} \end{bmatrix} \boldsymbol{Y}\right] - \lambda_2 \operatorname{tr}\left[\begin{bmatrix} \boldsymbol{C}^{\mathrm{T}}\boldsymbol{C} & \mathbf{0}_{N_x\times N_u} \\ \mathbf{0}_{N_u\times N_x} & \mathbf{0}_{N_u\times N_u} \end{bmatrix} \boldsymbol{Y}\right] \right), \tag{3.44}$$

[9] Let the state equation be $\boldsymbol{x}_{k+1} = \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{B}(\boldsymbol{u}_k + \boldsymbol{n}_k)$, $k \in \mathbb{N}_0$, with the standard assumptions of the LQG scenario and let the controller be $\boldsymbol{u}_k = \boldsymbol{L}\boldsymbol{x}_k$ such that the matrix $\boldsymbol{A} + \boldsymbol{B}\boldsymbol{L}$ has only eigenvalues with magnitudes less than one. If the asymptotic covariance matrix of this system is partitioned according to Equation (3.41), it can be rewritten as Equation (3.43). Due to the freedom of the choice of $\boldsymbol{L}$ and of the stationary covariance matrix of $(\boldsymbol{n}_k : k \in \mathbb{N}_0)$, we get infinitely many non-trivial solutions of Equation (3.43).

where only the terms which depend on the optimization variables have been considered and where $\lambda_1 = 0$ has been used. Obviously, the value of Equation (3.44) is only bounded from below for given values of $\boldsymbol{X}$ and $\boldsymbol{Y}$ if the difference in the parentheses is positive since it can be scaled with an arbitrary large scalar $\xi \geq 0$. Consequently, the largest value of λ_2 is determined by the largest absolute value of the negative trace expression in Equation (3.44) for a fixed value of the first trace expression of the difference. This value of λ_2 is given by the solution of the following optimization problem, where the choice of the positive semidefinite matrix $\boldsymbol{Y}$ is restricted by Equation (3.43):

$$\begin{aligned} \underset{\boldsymbol{Y}}{\text{maximize}} \;\; & \operatorname{tr}\left[\begin{bmatrix} \boldsymbol{C}^{\mathrm{T}}\boldsymbol{C} & \mathbf{0}_{N_x\times N_u} \\ \mathbf{0}_{N_u\times N_x} & \mathbf{0}_{N_u\times N_u} \end{bmatrix} \boldsymbol{Y}\right] \\ \text{subject to} \;\; & \boldsymbol{E}\boldsymbol{Y}\boldsymbol{E}^{\mathrm{T}} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \end{bmatrix} \boldsymbol{Y} \begin{bmatrix} \boldsymbol{A}^{\mathrm{T}} \\ \boldsymbol{B}^{\mathrm{T}} \end{bmatrix}, \\ & \operatorname{tr}\left[\begin{bmatrix} \boldsymbol{Q} & \boldsymbol{S} \\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{R} \end{bmatrix} \boldsymbol{Y}\right] = 1, \\ & \boldsymbol{Y} \succeq \mathbf{0}_{(N_x+N_u)\times(N_x+N_u)}. \end{aligned} \tag{3.45}$$

Note that the constant 1 for the constraint which fixes the value of the first summand inside the parentheses in Equation (3.44) is chosen arbitrarily. Nevertheless, with this choice, the maximal value of λ_2 is given by $(\operatorname{tr}\left[\boldsymbol{C}\boldsymbol{E}\boldsymbol{Y}^{*}\boldsymbol{E}^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}\right])^{-1}$, where $\boldsymbol{Y}^{*}$ is the optimizing matrix of the convex maximization problem (3.45).

The same approach as the one presented above can be used to analyze an equality constraint for the transmit power P_2 of the control channel or equality constraints for both channels. In any case, the requirement to use the full available transmit powers $P_{\mathrm{Tx},1}$ and $P_{\mathrm{Tx},2}$, respectively, can effectively lead to a negative loading of the weighting matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$, respectively, for the unconstrained LQG problem which has to be solved for the determination of the Lagrange dual function (cf. optimization problem 3.40). As a result, the variance of the control signal and of the system state, respectively, increases in this case which leads to a larger LQG cost compared to the scenario where no (or inactive) inequality constraints are considered. This is the shortcoming of controller solutions which are obtained by equality constrained LQG problems. Transmit powers which are larger than those obtained by unconstrained or inequality constrained optimization problems can only be realized by increasing the variance of the system states which is in general in conflict with the primary goal of LQG control. Additionally, the increased SNR (recall that the noise variances of the observation and control channel are assumed to be fixed) does not lead to a smaller error of the state estimate which is an integral part of all controller solutions obtained in this chapter and which also contributes to the LQG cost (cf., e. g., optimization problems 3.36 and 3.39). The simple reason is that all increases of SNRs are due to the choice of the control signal $(\boldsymbol{u}_k : k \in \mathbb{N}_0)$ which has no influence on the estimation error and thus the cost function in the scenario under consideration (see Section A6.2). Consequently, the intuition that larger SNRs for the transmission of signals over a noisy channel lead to smaller estimation errors at the receiver side does not apply here.

Example 3.4.2 To conclude this section, the following example presents several results obtained so far. Again, the parameters of the dynamical system and the noise sequences of Example 3.1.1 are used. In order to determine the maximal region of transmit powers P_1 and P_2 which can be achieved with the equality constrained LQG approach, the optimization problem (3.39) is solved with the specific choice of $\boldsymbol{Q} = \mathbf{0}_{N_x \times N_x}$ and $\boldsymbol{S} = \mathbf{0}_{N_x \times N_u}$. For the weighting matrix of the control signal, we chose $\boldsymbol{R} = \mathbf{I}_{N_u}$ for the determination of the minimal value of P_2 when P_1 is fixed to $P_{\text{Tx},1}$. Analogously, $\boldsymbol{R} = -\mathbf{I}_{N_u}$ is applied for the maximization of P_2 for the given value of P_1. The transmit powers that can be realized for a given constraint on P_1 must obviously lie between these minimal and maximal values. For the constraint for P_1, i. e., $P_{\text{Tx},1}$, 250 logarithmically spaced values between the minimal value of approx 92.1 which leads to a feasible optimization problem (see Section 3.3) and 10^9 have been evaluated. The resulting region of SNRs, which is determined by the minimal and maximal values of P_2, is shown in Figure 3.7. Additionally, the maximal region of transmit powers that can be realized with inequality constraints (cf. Figure 3.6) and the point of transmit powers of the unconstrained LQG controller with the original weighting matrices $\boldsymbol{Q}$, $\boldsymbol{R}$ and $\boldsymbol{S}$ from Example 3.1.1 (cf. Figure 3.2) are shown.

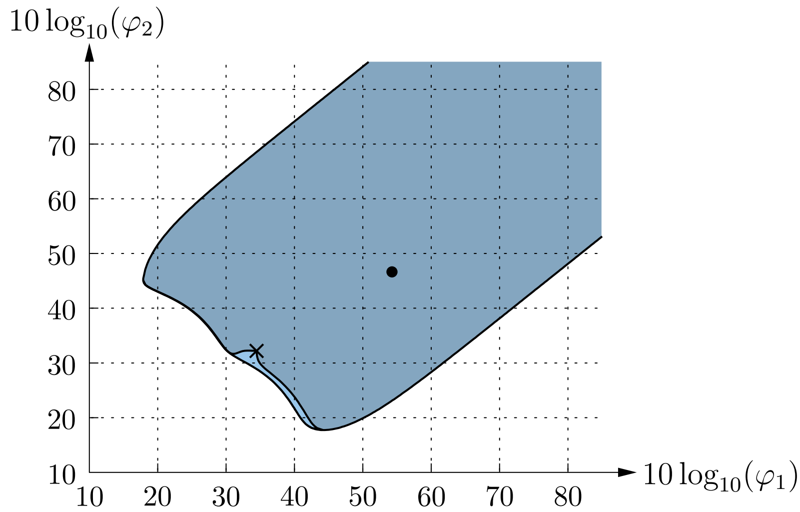


Figure 3.7: The SNR region for the equality constrained LQG problem.

In Figure 3.7, it can be seen that the region of SNRs φ_1 and φ_2 (cf. Equation 3.17) for the inequality constrained LQG control problem lies inside the region for the equality constrained optimization and shares the Pareto optimal set of transmit powers. Nevertheless, the larger set of SNRs is achieved at the expense of a larger value of the cost function, i. e., of worse performance. This is illustrated by the solution of the optimization problem (3.39) where the value of $P_{\text{Tx},1}$ is set to $4 \cdot 10^5$, i. e., $10\log_{10}(\operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{q}}\right]^{-1} P_{\text{Tx},1}) \approx 54.26$, and where the original matrices $\boldsymbol{Q}$, $\boldsymbol{R}$ and $\boldsymbol{S}$ from Example 3.1.1 are used. Since the equality constraint can be fulfilled, we have $P_1 = P_{\text{Tx},1}$. The controller which minimizes the LQG cost subject to this constraint results in a transmit power for the control channel of $P_2 \approx 4.61 \cdot 10^4$ with the respective SNR of $10\log_{10}(\operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{n}}\right]^{-1} P_2) \approx 46.63$. The cost associated with this SNR point, depicted by $\bullet$ in Figure 3.7, is approximately $2.40 \cdot 10^4$. In contrast, the cost of the unconstrained LQG solution which leads to the SNR point depicted by $\times$ is approximately 799. Thus, the cost is increased by a factor of 30 in order to provide the desired values of transmit power. This increase is due to the negative contribution of the Lagrange multipliers $\lambda_1 - \lambda_2 \approx -0.0591$ in the optimization problem (3.40) which leads to the fact that the weighting matrix $\boldsymbol{Q} + (\lambda_1 - \lambda_2)\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}$ is indefinite. Considering the solution of the problem (3.40), the magnitude of the largest eigenvalue of the associated matrix $\boldsymbol{A} + \boldsymbol{B}\boldsymbol{L}$ which describes the closed loop dynamics of the controlled system and where $\boldsymbol{L}$ is given by Equation (3.24) is approximately 0.995. Thus, the closed loop system is stable, but the eigenvalue which is close to one leads to a large variance of the system states.

Finally, the largest value of λ_2 for the case when λ_1 is zero and which is determined by the optimization problem (3.45) is approximately 0.0591. The difference between this value and the one determined above for the specific power constraint is in the order of 10^{-6}.

3.5 Discussion

In this chapter, we presented an approach to LQG controller design with additional transmit power constraints which is based on convex, or precisely, semidefinite optimization. For the reformulation of the classical LQG optimization problem, the fact has been used that for LTI systems with stationary Gaussian disturbances which are stabilized by a stationary LQG controller (and with the respective stabilizability and observability assumptions), the joint distribution of the system state and the control signal converges to an asymptotically stationary distribution. Since it is again Gaussian due to the linearity of the closed loop system, the determination of the optimal controller is equivalent to the determination of the optimal covariance matrix of the stationary distribution. The mean vector has not been considered because it is zero due to the simplifying model assumptions. The reformulation of the LQG problem in terms of covariance matrices has the appealing property that transmit power constraints can be included directly in the optimization of the cost function. Besides the semidefiniteness constraint for the covariance matrices, the cost function and all constraints are linear in the optimization variables and thus a numerical solution can be obtained relatively easy using standard software packages for semidefinite optimization.

There are several examples of constrained LQG controller designs which use a description of the optimization problem in terms of covariance matrices. In [81] (see also [96]), trace constraints for the covariance matrices of system output variables are considered and the Lagrangian associated with the constrained optimization problem is formulated. In order to determine the optimal value of the dual variables, a quasi-Newton method is applied to the associated system of non-linear equations. Nevertheless, the covariance matrix of the system state or its estimate, respectively, is not optimized directly but is the solution of a Lyapunov equation (cf. Equation 3.13 or 3.15, respectively) which depends explicitly on the optimal feedback gain $\boldsymbol{L}$ (cf. Equation 3.3). This gain matrix is determined by the solution of a DARE, see Equations (3.3) and (3.4).

A second example which uses a systematic approach to the constrained LQG control problem in the context of convex optimization is [78, Chapter 12 and Section 14.5]. The authors consider only the continuous-time case and formulate the LQG optimization in the frequency domain, i. e., in terms of transfer functions or power spectral densities, respectively (cf. Section A6.3). Consequently, the variance constraints for output signals are taken into account by the norms of the associated transfer functions. Unfortunately, in order to apply the proposed optimization algorithms, it is necessary to find finite dimensional approximations for the transfer functions under consideration [78, Chapter 15].

The authors of [83] do not restrict themselves to variance or power constraints, i. e., constraints for the (weighted) trace of covariance matrices of output signals or the norm of the associated transfer functions, but consider matrix inequalities for these covariance matrices. Although not mentioned explicitly, a projected gradient algorithm is used to determine the optimal value of the dual variables which are associated with the matrix constraints. Again, the controller itself is determined conventionally, i. e., according to Equations (3.3) and (3.4), where the case of output feedback which makes it necessary to estimate the system state is also taken into account.

The approach introduced in [88] is the basis for the results presented in this chapter. The optimization problem for the determination of the optimal LQG controller is not formulated in terms of a linear function of the system state or its estimate, respectively, but of the joint covariance matrix of those quantities and the control signal. The author of [88] analyzes the structure of the finite horizon LQG problem using the dual of the original optimization problem and establishes relations to the classical solution methods for the optimal LQG controller. Since this has not been the focus of this chapter, we relied on the approach from [88] for the infinite horizon case to highlight the convexity of the constrained optimization problem and to provide a formulation which can be implemented almost directly using software packages for semidefinite optimization. Additionally, we provided a discussion about the feasibility of the problem and some basic properties of the region of transmit powers that can be realized using inequality and equality constraints. The analysis of the latter case led to the consideration of LQG control problems with indefinite weighting matrices and presented a condition which determines if the optimization problem is well-posed, i. e., has a cost function which is bounded from below within the constraint set. Note that we did not intend to provide a detailed analysis of the indefinite LQG control problem and thus refer the interested reader to, e. g., [97, 98].

The investigation of the transmit powers which can be realized using the power constrained LQG approach showed some shortcomings of this way to consider limited communication resources. One example are dynamical systems which are stable even without control. In this case, the minimal power which must be provided for the transmission of the control signal is obviously zero (see Section 3.3.1). Since the controller is switched off, it is not necessary to transmit the observations at the system output to the controller, i. e., also this transmit power could be set to zero. Unfortunately, this is not possible since the controller is the only degree of freedom which has an influence on all transmit powers in the control loop. If it is switched off, the observations at the system output require a certain amount of power which is determined by the open loop behavior of the dynamical system and is larger than zero in general. A second example is the benefit of available transmit power which exceeds the requirements of the optimal, unconstrained LQG controller. This power can only be used if the variance of the system states and their estimates (which determines the power of the control signal) is increased (see Sections 3.4.1 and 3.4.2). Since these quantities are also elements of the LQG cost function, the performance of a system which amplifies its transmit powers beyond the point that is determined by an unconstrained controller will perform worse than without this increase. These two examples show that additional degrees of freedom, i. e., transmitters and receivers for the signals in the control loop, have to be introduced in order to use the available communication resources efficiently. A simple extension of the system model shown in Figure 3.1 is presented in the following chapter.

4. Joint Design of Controller and Scalar Transceiver

It has been shown in the preceding chapter how the optimal Linear Quadratic Gaussian (LQG) controller which takes into account the limited communication resources, i. e., transmit powers, for the exchange of information in a closed loop control system, is obtained by the solution of a convex optimization problem. Nevertheless, the investigation of power constraints for the standard LQG system model (see Figure 3.1) revealed some undesired properties of the solution. For example, if the controller can be switched off without loosing stability (which is the case for open loop stable dynamical systems), it is reasonable to stop the transmission of observations from the system output to the controller. Consequently, it should be possible to reduce the transmit power at the output of the dynamical system to zero. As a second example, assume that the controller is allowed to use more power for the transmission of control signals than the optimal solution of the LQG problem without constraints. Since the input matrix of the dynamical system $\boldsymbol{B}$ is fixed, it is in general not possible to simply amplify the control signal due to the potential loss of stability of the closed loop system. This behavior becomes clear if we consider a stabilizing controller, represented by the matrix $\boldsymbol{L}$ (cf. Equation 3.3), such that the closed loop matrix $\boldsymbol{A} + \boldsymbol{B}\boldsymbol{L}$ has only eigenvalues with magnitude less than one. If the control signal is amplified using a scaling factor $c > 0$ for the controller gain $\boldsymbol{L}$, it is obvious that the resulting closed loop matrix $\boldsymbol{A}+\boldsymbol{B}c\boldsymbol{L}$ will have eigenvalues with magnitude larger than one if c is chosen too large and thus the system becomes unstable.

4.1 Extension of the System Model

The two examples mentioned above show that it is desirable to introduce two degrees of freedom for the closed loop control system shown in Figure 3.1. One should allow for a scaling of the transmit signal of the observation channel such that, in the limit, it can be switched off when it is not needed by the controller. The other one should enable the receiver of the control channel to scale the received signal such that the transmitting controller can adjust its transmit power without the negative effects for the stability of the closed loop system. The most simple instances of these degrees of freedom are scalars at the transmitter of the observation channel and the receiver of the control channel, respectively, which fulfill the requirements mentioned above. The modification of the system model from Chapter 3 using the two scalars $t > 0$ and $g > 0$ is shown in Figure 4.1.

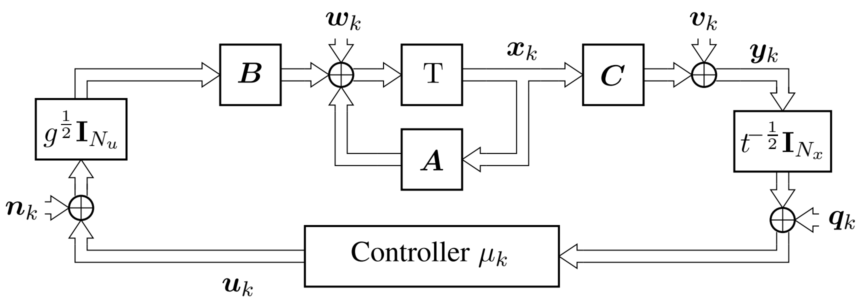


Figure 4.1: Model of the control loop which is closed over two channels with additive noise $\boldsymbol{q}_k$ and $\boldsymbol{n}_k$. At the system input and output, the receive scaling $g^{\frac{1}{2}}$ and the transmit scaling $t^{-\frac{1}{2}}$, respectively, is introduced.

The case of a zero transmit or receive scaling is not considered since it leads to an undetectable or unstabilizable dynamical system, respectively.[1] Note that the transmit and receive signals are scaled by the square root of t^{-1} and g, respectively, The reason for this choice will become clear in the following. Furthermore, the restriction to the positive root represents no loss of generality since the optimal controller, which has to take into account the scaling of the system input and output, has the possibility to change the sign of its input and output signals if the negative root is chosen. This has no effect on the system performance or the transmit powers. In the following, we refer to t as transmitter and to g as receiver. A pair of transmitter and receiver is called transceiver.

In Chapter 3, the state and observation equation shown in Equation (1.1) have been modified for the problem (3.1) to take into account the channel noises, i. e., the communication channels have been incorporated into the system model for the optimization of the LQG controller. The analogous steps are necessary now to include the transmitter t and receiver g. The resulting LQG optimization problem then reads as

$$\underset{\mu_0,\mu_1,\mu_2\dots}{\text{minimize}} \lim_{N\to\infty} \frac{1}{N} \mathrm{E}\left[\boldsymbol{x}_N^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{x}_N+\sum_{n=0}^{N-1}\begin{bmatrix}\boldsymbol{x}_n\\ g^{\frac{1}{2}}\boldsymbol{u}_n\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{Q} & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{R}\end{bmatrix}\begin{bmatrix}\boldsymbol{x}_n\\ g^{\frac{1}{2}}\boldsymbol{u}_n\end{bmatrix}\right] + g\,\mathrm{tr}\left[\boldsymbol{R}\boldsymbol{C}_{\boldsymbol{n}}\right] \tag{4.1}$$

$$\begin{aligned}
\text{subject to}\quad \boldsymbol{x}_{k+1} &= \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{B}g^{\frac{1}{2}}(\boldsymbol{u}_k+\boldsymbol{n}_k)+\boldsymbol{w}_k, && k\in\mathbb{N}_0,\\
\boldsymbol{y}_k &= \boldsymbol{C}\boldsymbol{x}_k+\boldsymbol{v}_k, && k\in\mathbb{N}_0,\\
\boldsymbol{u}_k &= \mu_k(\mathcal{I}_k), && k\in\mathbb{N}_0,\\
\mathcal{I}_k &= \begin{cases}\{(t^{-\frac{1}{2}}\boldsymbol{y}_0+\boldsymbol{q}_0)\}, & k=0,\\ \{(t^{-\frac{1}{2}}\boldsymbol{y}_0+\boldsymbol{q}_0),\dots,(t^{-\frac{1}{2}}\boldsymbol{y}_k+\boldsymbol{q}_k),\boldsymbol{u}_0,\dots,\boldsymbol{u}_{k-1}\}, & k\in\mathbb{N},\end{cases}
\end{aligned}$$

where the standard LQG assumptions are made (cf. Appendix A6) and where the indices of the expected value (cf. Equation 3.1) have been dropped in order to simplify the notation. Together with the channel model assumptions (cf. Section 1.6.2), $(\boldsymbol{w}_k : k\in\mathbb{N}_0)$, $(\boldsymbol{v}_k : k\in\mathbb{N}_0)$, $(\boldsymbol{q}_k : k\in\mathbb{N}_0)$ and $(\boldsymbol{n}_k : k\in\mathbb{N}_0)$ are mutually independent identically and independently distributed (i.i.d.) sequences of Gaussian random vectors which are independent of the initial state $\boldsymbol{x}_0$ which is also a Gaussian random vector. Note that we do not optimize w.r.t. the transmitter t and the receiver g but assume the two scalars to be fixed for the moment. It can be seen that for the optimization problem (4.1), the input of the dynamical system at time index $k\in\mathbb{N}_0$ is now given by $g^{\frac{1}{2}}(\boldsymbol{u}_k+\boldsymbol{n}_k)$. The information $\mathcal{I}_k$ which is available to the controller at time index $k\in\mathbb{N}_0$ contains the noisy observations $t^{-\frac{1}{2}}\boldsymbol{y}_i+\boldsymbol{q}_i$, $i\in\{0,1,\dots,k\}$, which depend on the transmitter $t^{-\frac{1}{2}}$.

The fact that the cost function of the optimization problem (4.1) takes into account the scaled control signal $g^{\frac{1}{2}}\boldsymbol{u}_n$, $n\in\mathbb{N}_0$, and gets the additional term $g\,\mathrm{tr}\left[\boldsymbol{R}\boldsymbol{C}_{\boldsymbol{n}}\right]$ needs some explanations. The LQG cost captures the effort of the control signal which is actually applied to the dynamical system, i. e., $g^{\frac{1}{2}}(\boldsymbol{u}_n+\boldsymbol{n}_n)$, $n\in\mathbb{N}_0$, for the present scenario. Consequently, it takes into account the cost associated with $g^{\frac{1}{2}}\boldsymbol{u}_n$, $n\in\mathbb{N}_0$, rather than $\boldsymbol{u}_n$. Additionally, the noise $g^{\frac{1}{2}}\boldsymbol{n}_n$, $n\in\mathbb{N}_0$, is included in the cost function since it is fed in the dynamical system and thus contributes to the control effort. Due to the model assumptions, the noise sequence $(\boldsymbol{n}_n : n\in\mathbb{N}_0)$ is uncorrelated with all other random variables describing the scenario under consideration. The increase of the cost because of this disturbance is thus given by $g\,\mathrm{tr}\left[\boldsymbol{R}\boldsymbol{C}_{\boldsymbol{n}}\right]$.

[1]This corresponds to a disconnection in the closed control loop which can be considered for open loop stable systems but is of no interest for unstable ones.

4.2 Optimization of Controller and Transceiver With Power Constraints

4.2.1 Optimal Controller With Fixed Transceiver

Before we start with the consideration of power constraints for the optimization problem (4.1) and the determination of the optimal values of the transmitter and receiver, the solution of the unconstrained LQG problem according to Appendix A6 which has also been presented in Equations (3.2) – (3.7) is investigated for the case of given values of t and g. For the adaption of the results, it is useful to incorporate the transmitter and receiver in the system model and in the parameters of the noise sequences and the cost function. Having a closer look at the cost function and the constraints of the optimization problem (4.1), it can be seen that the weighting matrix for the cross products of the system state and the control signal is not $\boldsymbol{S}$ but effectively $g^{\frac{1}{2}}\boldsymbol{S}$ by pulling the scalar $g^{\frac{1}{2}}$ in the joint weighting matrix, while for the control signal alone $\boldsymbol{R}$ has to be replaced by $g\boldsymbol{R}$. The system input matrix becomes $g^{\frac{1}{2}}\boldsymbol{B}$ instead of $\boldsymbol{B}$. Shifting the scalar $t^{-\frac{1}{2}}$ from the information set $\mathcal{I}_k$, $k \in \mathbb{N}_0$, to the observation equation which provides $\boldsymbol{y}_k$, $k \in \mathbb{N}_0$, leads to a replacement of the system output matrix $\boldsymbol{C}$ by $t^{-\frac{1}{2}}\boldsymbol{C}$. Finally, due to the scaling, the covariance matrix of the observation noise is $t^{-1}\boldsymbol{C}_{\boldsymbol{v}}$ instead of $\boldsymbol{C}_{\boldsymbol{v}}$. Using these replacements, it is straightforward to see by comparison with Equations (3.2) – (3.7) that the optimal control sequence is given by

$$\boldsymbol{u}_k = \boldsymbol{L}\hat{\boldsymbol{x}}_k, \qquad k \in \mathbb{N}_0, \tag{4.2}$$

where the controller gain reads as

$$\begin{aligned}\boldsymbol{L} &= -\left(g\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B} + g\boldsymbol{R}\right)^{-1}\left(g^{\frac{1}{2}}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} + g^{\frac{1}{2}}\boldsymbol{S}^{\mathrm{T}}\right)\\ &= -g^{-\frac{1}{2}}\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B} + \boldsymbol{R}\right)^{-1}\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} + \boldsymbol{S}^{\mathrm{T}}\right),\end{aligned} \tag{4.3}$$

and where we used the fact that $g \neq 0$. This expression depends on the stabilizing solution of the Discrete Algebraic Riccati Equation (DARE)

$$\begin{aligned}\boldsymbol{K} &= \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \left(g^{\frac{1}{2}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B} + g^{\frac{1}{2}}\boldsymbol{S}\right)\left(g\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B} + g\boldsymbol{R}\right)^{-1}\left(g^{\frac{1}{2}}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} + g^{\frac{1}{2}}\boldsymbol{S}^{\mathrm{T}}\right) + \boldsymbol{Q}\\ &= \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B} + \boldsymbol{S}\right)\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B} + \boldsymbol{R}\right)^{-1}\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} + \boldsymbol{S}^{\mathrm{T}}\right) + \boldsymbol{Q}.\end{aligned} \tag{4.4}$$

The second lines of Equation (4.3) and (4.4), respectively, are valid since the non-zero scalar g can be pulled out of the inverse. Thus, the DARE does not depend on the receiver scaling. Comparing the controller solution in the equations above with the solution shown in Equations (3.2) – (3.4), it can be seen that we get a controller gain $\boldsymbol{L}$ that is almost identical to the case when no receiver scaling is present. The only difference is that with the receiver $g^{\frac{1}{2}}$, the control signal is simply scaled by the inverse $g^{-\frac{1}{2}}$.

Since the overall system is linear using the scalar transmitter and receiver and all disturbances are Gaussian, the estimate $\hat{\boldsymbol{x}}_k = \mathrm{E}_{\boldsymbol{x}_k|\mathcal{I}_k}\left[\boldsymbol{x}_k|\,\mathcal{I}_k\right]$, $k \in \mathbb{N}_0$, of the system state is computed using the Kalman filter (see Appendix A7), where its estimation error covariance matrix now reads as

$$\begin{aligned}\boldsymbol{C}_{\tilde{\boldsymbol{x}}} &= \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - t^{-\frac{1}{2}}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\left(t^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}} + t^{-1}\boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{C}_{\boldsymbol{q}}\right)^{-1}t^{-\frac{1}{2}}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\\ &= \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}} + t\boldsymbol{C}_{\boldsymbol{q}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\end{aligned} \tag{4.5}$$

and depends on the stabilizing solution of the DARE

$$\begin{aligned}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} &= \boldsymbol{A}\Big(\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - t^{-\frac{1}{2}}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\left(t^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}} + t^{-1}\boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{C}_{\boldsymbol{q}}\right)^{-1}t^{-\frac{1}{2}}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\Big)\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{w}} + g\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\\ &= \boldsymbol{A}\left(\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}} + t\boldsymbol{C}_{\boldsymbol{q}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\right)\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{w}} + g\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}.\end{aligned} \tag{4.6}$$

Remark: The results above show the equivalence of two different points of view in the context of LQG control with power or Signal to Noise Ratio (SNR) constraints. It is common in the literature on Networked Control Systems (NCSs) to consider the SNR as the limited communication resource and not the transmit power for an additive noise channel with fixed noise variance, see, e. g., [29,46, 84]. One reason is that in a control loop, the variances of transmit signals like the system output or the control signal, are directly related to the variance of disturbances which are fed into the system, which makes it natural to investigate the ratio of these variances and not the transmit power alone. Another reason is the additive noise approximation for noise free but quantized communication channels, where the resolution of the quantizer determines the channel SNR [44, 46].

Having a closer look at the DAREs in Equations (4.4) and (4.6), it can be seen that they correspond to the solution of an LQG optimization for the control system shown in Figure 4.2. In this case, no transmit or receive scaling factors are present but the channel noises are scaled with $t^{\frac{1}{2}}$ and $g^{\frac{1}{2}}$, respectively, which leads to the noise covariance matrices $t\boldsymbol{C}_{\boldsymbol{q}}$ and $g\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}$ in Equation (4.6). From this point of view, not the transmit and receive scaling factors are design parameters but the variances of the noise sequences in the communication channels. This is the reason for the specific choice of the transmitter and receiver shown in Figure 4.1: the variances of the observation and control channel noise depend linearly on the scaling factors. Note that also for the scenario shown in Figure 4.2, the covariance matrices $\boldsymbol{C}_{\boldsymbol{q}}$ and $\boldsymbol{C}_{\boldsymbol{n}}$ are assumed to be given and fixed.

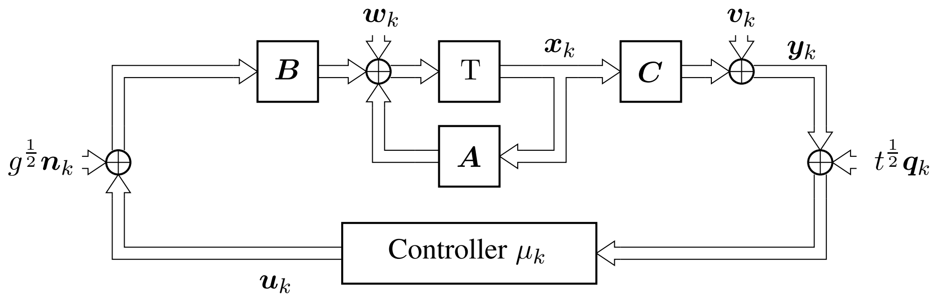


Figure 4.2: Equivalent model of the control loop which is closed over two channels with additive noise $\boldsymbol{q}_k$ and $\boldsymbol{n}_k$ where the transmitter $t^{\frac{1}{2}}$ and receiver $g^{\frac{1}{2}}$ are interpreted as parameters of the channel noise variances.

Considering the limited communication resources in an NCS, a power constraint for the scenario shown in Figure 4.1, where it is assumed that the channel noise variance is fixed, translates into an SNR constraint for the scenario in Figure 4.2. In the former case, let $P_{\mathrm{Tx},1}$ be the available transmit power for the transmission of observations to the controller. The corresponding constraint thus reads as $t^{-1}\,\mathrm{tr}\left[\boldsymbol{C}\boldsymbol{C}_{\boldsymbol{x}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}}\right] \leq P_{\mathrm{Tx},1}$ (cf. Equation 3.18). In the latter case, let φ_1 be the maximal value of the SNR of the observation channel. The constraint now becomes $\frac{\mathrm{tr}\left[\boldsymbol{C}\boldsymbol{C}_{\boldsymbol{x}}\boldsymbol{C}^{\mathrm{T}}+\boldsymbol{C}_{\boldsymbol{v}}\right]}{\mathrm{tr}\left[t\boldsymbol{C}_{\boldsymbol{q}}\right]} \leq \varphi_1$ which is equivalent to $t^{-1}\,\mathrm{tr}\left[\boldsymbol{C}\boldsymbol{C}_{\boldsymbol{x}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}}\right] \leq \varphi_1\,\mathrm{tr}\left[\boldsymbol{C}_{\boldsymbol{q}}\right]$ and thus to the power constraint.

The introduction of an inverse scaling at the input and the output of an additive noise communication channel and the interpretation of an equivalent change of the noise statistics has been discussed, e. g., in [86], [35] and also [44]. Nevertheless, no statement is made that the restriction to an inverse pair of transmitters and receivers represents no loss of generality when the optimal controller is used. Other works like [46] and [99] do not make use of explicit transmitters and receivers but consider the variance of the channel noise as a design parameter, for example if the disturbance models a quantization error which depends on the choice of the actual quantizer.

Since it is an LQG problem, it is straightforward to determine the optimal value J^*_∞ of the optimization problem (4.1). With the modifications due to the channel noises, the transmitter t and the receiver g, this optimal value reads as (cf. Appendix A6.3)

$$J^*_\infty = \operatorname{tr}\left[\boldsymbol{K}\left(\boldsymbol{C}_{\boldsymbol{w}} + g\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\right)\right] + \operatorname{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\right] + g\operatorname{tr}\left[\boldsymbol{R}\boldsymbol{C}_{\boldsymbol{n}}\right], \tag{4.7}$$

where $\boldsymbol{P} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \boldsymbol{K} + \boldsymbol{Q}$. Note that $\boldsymbol{K}$ and $\boldsymbol{P}$ do not depend on t and g, whereas the covariance matrix $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}$ of the state estimation error (see Equation 4.5) is a function of these parameters. Since the transmitter and the receiver have an influence on the *additional* observation noise, given by $t\boldsymbol{C}_{\boldsymbol{q}}$, and process noise, given by $g\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}$, respectively, it is intuitively clear that an increase of g or t results in a larger estimation error and thus in a monotonic increase of J^*_∞. This intuition is confirmed by [100] where it is shown that if $\tilde{t}\boldsymbol{C}_{\boldsymbol{q}} \geq t\boldsymbol{C}_{\boldsymbol{q}}$ and $\tilde{g}\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}} \geq g\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}$, then $\tilde{\boldsymbol{C}}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}} \geq \boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}$, where $\tilde{\boldsymbol{C}}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}$ corresponds to the DARE with parameters $\tilde{t}\boldsymbol{C}_{\boldsymbol{q}}$ and $\tilde{g}\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}$. Additionally, the results in Appendix A4 show that the derivative of $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}$ w.r.t. t and g is positive semidefinite, i. e., $\operatorname{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\right]$ increases monotonically with t and g.

4.2.2 Joint Optimization

Having established the basic results for the LQG control problem with fixed transmitters and receivers, we now turn to the optimization of these parameters when additionally power constraints have to be taken into account. The optimization of the transmitter and receiver obviously only makes sense with these constraints because otherwise t and g can be chosen to be arbitrarily small. This would lead to the elimination of the negative effect of the channel noises on the control performance by increasing all transmit powers towards infinity. In order to be compatible with the infinite horizon average cost function of the optimization problem (4.1), the power constraints are also formulated as averages over the infinite horizon.

Definition 4.2.1: Power of transmit signals
For the system model depicted in Figure 4.1, the transmit power P_1 of the observation channel and P_2 of the control channel is given by

$$P_1 = \lim_{N\to\infty}\frac{1}{N}\operatorname{E}\left[t^{-1}\left(\sum_{n=0}^{N-1}\boldsymbol{x}_n^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{x}_n + \boldsymbol{v}_n^{\mathrm{T}}\boldsymbol{v}_n\right)\right] \quad \text{and} \tag{4.8}$$

$$P_2 = \lim_{N\to\infty}\frac{1}{N}\operatorname{E}\left[\sum_{n=0}^{N-1}\boldsymbol{u}_n^{\mathrm{T}}\boldsymbol{u}_n\right], \tag{4.9}$$

respectively, where the expected values are w.r.t. all random variables of the system and channel models. For P_1, the cross terms of the system state and the observation noise have been omitted since their expected value is zero due to the model assumptions.

Note that if the distributions of the system state and the control signal converge to stationary values (which is the case for the following results), we obtain expressions for the transmit powers analogous to Equations (3.8) and (3.9), which are formulated in terms of the asymptotic covariance matrices of the system state and the control signal. Let the available power for the observation channel be given by $P_{\mathrm{Tx},1} > 0$ and for the control channel by $P_{\mathrm{Tx},2} > 0$. With the constraints that

$P_1 \leq P_{\mathrm{Tx},1}$ and $P_2 \leq P_{\mathrm{Tx},2}$, the optimization problem (4.1) becomes

$$\begin{aligned}
\underset{t,g,\mu_0,\mu_1,\mu_2\ldots}{\text{minimize}} \quad & \lim_{N\to\infty} \frac{1}{N} \mathrm{E}\left[\boldsymbol{x}_N^{\mathrm{T}} \boldsymbol{Q}_N \boldsymbol{x}_N + \sum_{n=0}^{N-1} \begin{bmatrix} \boldsymbol{x}_n \\ g^{\frac{1}{2}} \boldsymbol{u}_n \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{S} \\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{R} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_n \\ g^{\frac{1}{2}} \boldsymbol{u}_n \end{bmatrix}\right] + g \operatorname{tr}\left[\boldsymbol{R}\boldsymbol{C}_{\boldsymbol{n}}\right] \qquad (4.10) \\
\text{subject to} \quad \boldsymbol{x}_{k+1} &= \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{B} g^{\frac{1}{2}} (\boldsymbol{u}_k + \boldsymbol{n}_k) + \boldsymbol{w}_k, \qquad k \in \mathbb{N}_0, \\
\boldsymbol{y}_k &= \boldsymbol{C}\boldsymbol{x}_k + \boldsymbol{v}_k, \qquad k \in \mathbb{N}_0, \\
\boldsymbol{u}_k &= \mu_k(\mathcal{I}_k), \qquad k \in \mathbb{N}_0, \\
\mathcal{I}_k &= \begin{cases} \{(t^{-\frac{1}{2}} \boldsymbol{y}_0 + \boldsymbol{q}_0)\}, & k = 0, \\ \{(t^{-\frac{1}{2}} \boldsymbol{y}_0 + \boldsymbol{q}_0), \ldots, (t^{-\frac{1}{2}} \boldsymbol{y}_k + \boldsymbol{q}_k), \boldsymbol{u}_0, \ldots, \boldsymbol{u}_{k-1}\}, & k \in \mathbb{N}, \end{cases} \\
t &> 0, \; g > 0, \\
\lim_{N\to\infty} \frac{1}{N} & \mathrm{E}\left[t^{-1} \left(\sum_{n=0}^{N-1} \boldsymbol{x}_n^{\mathrm{T}} \boldsymbol{C}^{\mathrm{T}} \boldsymbol{C} \boldsymbol{x}_n + \boldsymbol{v}_n^{\mathrm{T}} \boldsymbol{v}_n\right)\right] \leq P_{\mathrm{Tx},1}, \\
\lim_{N\to\infty} \frac{1}{N} & \mathrm{E}\left[\sum_{n=0}^{N-1} \boldsymbol{u}_n^{\mathrm{T}} \boldsymbol{u}_n\right] \leq P_{\mathrm{Tx},2},
\end{aligned}$$

where the minimization now is additionally w.r.t. t and g. Note that the positivity constraint for the transmitter and the receiver is explicitly shown since otherwise the problem would not be well-posed, e. g., if the weighting matrix $\boldsymbol{R}$ for the control variable is scaled by a negative value.

For fixed values of t and g, the minimization in (4.10) w.r.t. to the controller, i. e., μ_k, $k \in \mathbb{N}_0$, reduces to the problem discussed in Chapter 3. In this case, the controller is the only degree of freedom to minimize the LQG cost subject to the power constraints. With the additional transmit and receive scaling, an interaction between the controller and the transceivers is introduced. For example, the choice of a smaller value of t leads to a smaller error of the state estimate which is computed by the optimal controller (cf. Equations 4.5 and 4.6) and thus has the potential to reduce the variance of the system state and the associated observations. In contrast to this reduction, the observations are multiplied with $t^{-\frac{1}{2}}$ before transmission, which corresponds to an increase of transmit power. It follows that the choice of the optimal value of t, and also of g, such that the LQG cost is minimized and the transmit power constraints are fulfilled is not obvious.

In order to get a better insight in the interaction of the controller and the transceivers, we formulate the Lagrangian which is associated with the problem (4.10) where only the transmit power constraints are dualized, i. e., attached to the cost function using the Lagrange multipliers $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$. To this end, the constraint which is associated with $P_{\mathrm{Tx},1}$ is multiplied by t on both sides of the inequality, whereas the constraint associated with $P_{\mathrm{Tx},2}$ is multiplied by g. This step does not change the validity of the power constraints because t as well as g is positive. The reason for this equivalent formulation of the constraints will become clear in the following. Let J_∞ denote the cost function of the optimization problem (4.10) and P_1 and P_2 the variances of the transmit signals for the observation and the control channel, respectively (cf. Equations 4.8 and 4.9). Then, the Lagrangian associated with the optimization problem (4.10) is given by

$$\begin{aligned}
L &= J_\infty + \lambda_1 t (P_1 - P_{\mathrm{Tx},1}) + \lambda_2 g (P_2 - P_{\mathrm{Tx},2}) \\
&= \lim_{N\to\infty} \frac{1}{N} \mathrm{E}\left[\boldsymbol{x}_N^{\mathrm{T}} \boldsymbol{Q}_N \boldsymbol{x}_N + \sum_{n=0}^{N-1} \begin{bmatrix} \boldsymbol{x}_n \\ \boldsymbol{u}_n \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \boldsymbol{Q} + \lambda_1 \boldsymbol{C}^{\mathrm{T}} \boldsymbol{C} & g^{\frac{1}{2}} \boldsymbol{S} \\ g^{\frac{1}{2}} \boldsymbol{S}^{\mathrm{T}} & g\boldsymbol{R} + \lambda_2 g \mathbf{I}_{N_u} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_n \\ \boldsymbol{u}_n \end{bmatrix}\right] \qquad (4.11) \\
&\qquad + \lambda_1 \operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{v}}\right] + g \operatorname{tr}\left[\boldsymbol{R}\boldsymbol{C}_{\boldsymbol{n}}\right] - \lambda_1 t P_{\mathrm{Tx},1} - \lambda_2 g P_{\mathrm{Tx},2}.
\end{aligned}$$

Note that for fixed positive (non-negative) values of t, g, λ_1 and λ_2, Equation (4.11) is again the cost function of an LQG control problem where the weighting matrix of the system state is given by $\boldsymbol{Q} + \lambda_1 \boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}$ and of the control variable by $g\boldsymbol{R} + \lambda_2 g \mathbf{I}_{N_u}$. Consequently, the minimization of L w.r.t. μ_k, $k \in \mathbb{N}_0$, for fixed values of t, g, λ_1 and λ_2 subject to the state and observation equation and the constraint that the control variable at time index k is a function of $\mathcal{I}_k$ alone, i. e., the problem

$$\underset{\mu_0,\mu_1,\mu_2\cdots}{\text{minimize}}\; L \tag{4.12}$$

$$\begin{aligned}
\text{subject to}\;\; \boldsymbol{x}_{k+1} &= \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{B}g^{\frac{1}{2}}(\boldsymbol{u}_k + \boldsymbol{n}_k) + \boldsymbol{w}_k, && k \in \mathbb{N}_0,\\
\boldsymbol{y}_k &= \boldsymbol{C}\boldsymbol{x}_k + \boldsymbol{v}_k, && k \in \mathbb{N}_0,\\
\boldsymbol{u}_k &= \mu_k(\mathcal{I}_k), && k \in \mathbb{N}_0,\\
\mathcal{I}_k &= \begin{cases} \{(t^{-\frac{1}{2}}\boldsymbol{y}_0 + \boldsymbol{q}_0)\}, & k = 0,\\ \{(t^{-\frac{1}{2}}\boldsymbol{y}_0 + \boldsymbol{q}_0), \ldots, (t^{-\frac{1}{2}}\boldsymbol{y}_k + \boldsymbol{q}_k), \boldsymbol{u}_0, \ldots, \boldsymbol{u}_{k-1}\}, & k \in \mathbb{N},\end{cases}
\end{aligned}$$

is readily solved and provides the optimal value (cf. Equation 4.7)

$$\begin{aligned} L^*(t,g,\lambda_1,\lambda_2) = \operatorname{tr}\left[\boldsymbol{K}\left(\boldsymbol{C}_{\boldsymbol{w}} + g\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\right)\right] + \operatorname{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\right] + \lambda_1 \operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{v}}\right] + g\operatorname{tr}\left[\boldsymbol{R}\boldsymbol{C}_{\boldsymbol{n}}\right] \\ - \lambda_1 t P_{\mathrm{Tx},1} - \lambda_2 g P_{\mathrm{Tx},2}, \end{aligned} \tag{4.13}$$

where it is explicitly shown that L^* is a function of t, g, λ_1 and λ_2. This result can be verified by comparison with the respective expressions of Equation (4.1) and (4.12). Note that L^* is only defined for positive (non-negative) arguments. Using the results from Equations (4.4) – (4.6), L^* depends on the stabilizing solutions of the DAREs

$$\boldsymbol{K} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B} + \boldsymbol{S}\right)\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B} + \boldsymbol{R} + \lambda_2 \mathbf{I}_{N_u}\right)^{-1}\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} + \boldsymbol{S}^{\mathrm{T}}\right) + \boldsymbol{Q} + \lambda_1 \boldsymbol{C}^{\mathrm{T}}\boldsymbol{C} \tag{4.14}$$

and

$$\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} = \boldsymbol{A}\left(\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}} + t\boldsymbol{C}_{\boldsymbol{q}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\right)\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{w}} + g\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}, \tag{4.15}$$

where $\boldsymbol{P} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \boldsymbol{K} + \boldsymbol{Q} + \lambda_1 \boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}$ and $\boldsymbol{C}_{\tilde{\boldsymbol{x}}} = \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}} + t\boldsymbol{C}_{\boldsymbol{q}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}$.

For fixed values of t and g, e. g., $\bar{t}$ and $\bar{g}$, the optimization problem (4.10) is convex (see Chapter 3) and $L^*_{\bar{t},\bar{g}}(\lambda_1,\lambda_2) = L^*(\bar{t},\bar{g},\lambda_1,\lambda_2)$ represents the dual function associated with it.[2] Thus, an approach to obtain a solution for given transmit and receive scaling factors is to maximize the dual function w.r.t. the non-negative dual variables λ_1 and λ_2 (see, e. g., [94, Chapter 5] and [95, Chapter 6]). This requires that strong duality holds, i. e., that the optimal value of the problem (4.10) (evaluated at $\bar{t}$ and $\bar{g}$) and the supremum of $L^*_{\bar{t},\bar{g}}$ w.r.t. $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ are equal. A sufficient condition for this is Slater's constraint qualification (see, e. g., [94, Section 5.2.3] and [95, Section 5.3.1]), i. e., the existence of a controller such that the power constraints hold with strict inequality while the equality constraints are satisfied. This condition is fulfilled if at least one pair (P_1, P_2) of Pareto optimal transmit powers exists such that $P_{\mathrm{Tx},1} > P_1$ and $P_{\mathrm{Tx},2} > P_2$ (see Section 3.3). In the following, we assume that this is the case, i. e., that $P_{\mathrm{Tx},1}$ and $P_{\mathrm{Tx},2}$ are chosen large enough, where the lowest possible values are determined according to Section 3.3. Note that if the available transmit powers $P_{\mathrm{Tx},1}$ or $P_{\mathrm{Tx},2}$ are too small, i. e., the power constraints are not feasible, the dual function $L^*_{\bar{t},\bar{g}}$ is unbounded from above.

[2]The indices $\bar{t}$ and $\bar{g}$ are used to emphasize that transmit and receive scaling are assumed to be fixed at this point.

Equations (4.13), (4.14) and (4.15) show how the controller and transceiver optimization are coupled. The optimal controller gain (cf. Equation 4.3)

$$\boldsymbol{L} = -g^{-\frac{1}{2}} \left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B} + \boldsymbol{R} + \lambda_2 \mathbf{I}_{N_u}\right)^{-1} \left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} + \boldsymbol{S}^{\mathrm{T}}\right) \tag{4.16}$$

depends, up to the inverse of the receiver scaling $g^{\frac{1}{2}}$, only on the Lagrange multipliers λ_1 and λ_2 via the solution of the DARE given in Equation (4.14). Thus, for fixed values of t and g, the controller uses λ_1 and λ_2 to fulfill the power constraints at the expense of an increased LQG cost by shifting the weighting matrix $\boldsymbol{Q}$ towards $\lambda_1 \boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}$ (which represents the transmit power of the observation channel) and $\boldsymbol{R}$ towards $\lambda_2 \mathbf{I}_{N_u}$ (which represents the transmit power of the control channel). The transmit and receive scaling acts differently. Since $\boldsymbol{u}_k = \boldsymbol{L}\hat{\boldsymbol{x}}_k$, $k \in \mathbb{N}_0$, (cf. Equation 4.2 and 4.16), where $\hat{\boldsymbol{x}}_k$ is the optimal estimate of the system state at time index k, a change of t and g has an immediate impact on the transmit powers since the respective transmit signals are scaled accordingly, which can be verified with Equations (4.8) and (4.9). A secondary effect is due to the dependence of the variance of the estimation error on the parameters t and g, see Equations (4.5) and (4.6). Recalling the equality $\boldsymbol{x}_k = \hat{\boldsymbol{x}}_k + \tilde{\boldsymbol{x}}_k$, $k \in \mathbb{N}_0$, where $\hat{\boldsymbol{x}}_k$ and $\tilde{\boldsymbol{x}}_k$ are uncorrelated (see Appendix A7.2), we observe that a change of the variance of the estimation error $\tilde{\boldsymbol{x}}_k$ has also an effect on the variance of $\boldsymbol{x}_k$ and $\hat{\boldsymbol{x}}_k$ and thus on P_1 and P_2.

It can be seen that there are two ways to manage the transmit powers of the communication channels in the LQG control problem. One of changing the controller objective from the original performance criterion towards the transmit powers, which is realized by the adaption of λ_1 and λ_2, and the other one of scaling the controller and system output signals while taking into account the resulting variance of the state estimation error. These approaches are coupled by the function L^* (cf. Equation 4.13) in the subtractive terms $\lambda_1 t P_{\mathrm{Tx},1}$ and $\lambda_2 g P_{\mathrm{Tx},2}$. This specific representation of the function L^* is due to the multiplication of the power constraints of the optimization problem (4.10) by the respective transmit and receive scaling for the construction of the Lagrangian L. The equivalent formulation of the constraints results in the fact that the multipliers λ_1 and λ_2 only have an effect on the controller gain (cf. Equations 4.16 and 4.14) which determines the closed loop dynamics via the matrix $\boldsymbol{A}_{\mathrm{cl}} = \boldsymbol{A} + \boldsymbol{B}g^{\frac{1}{2}}\boldsymbol{L}$ (see Equation 3.10 and A36). With Equation (4.16) it follows that this matrix does not depend on t and g. The transceivers have an immediate effect on the transmit powers due to the scaling of the respective transmitted signals and the resulting variance of the state estimation error (cf. Equation 4.15).

4.2.3 Properties of the Optimization Problem

With the formulation of the problem (4.10), it is immediate to see that the joint optimization of the controller and the scalar transceivers is non-convex because the cost function as well as the state equation (which is an equality constraint) contain the product of the optimization variables $g^{\frac{1}{2}}$ and $\boldsymbol{u}_k$, $k \in \mathbb{N}_0$. Additionally, the constraint concerning the transmit power of the observation channel is quadratic in the system state but multiplied by t^{-1}, which represents also a non-convex function of the optimization variables. Note that this property remains unchanged for a reformulation of the problem without transmitter and receiver where the channel noise variance becomes a design parameter and the power constraints are converted to SNR constraints like in [46, 99] (see Figure 4.2). Even if the optimization problem is formulated in this context, we obtain exactly the same function L^* as in Equation (4.13). Consequently, we can not expect to determine a solution of (4.10) by minimizing $L^*(t, g, \lambda_1, \lambda_2)$ (cf. Equation 4.13) w.r.t. $t > 0$ and $g > 0$ in order to obtain the dual function[3] (see, e. g., [94, Section 5.1.2] and [95, Section 6.1])

$$D(\lambda_1, \lambda_2) = \inf_{t>0, g>0} L^*(t, g, \lambda_1, \lambda_2) \tag{4.17}$$

for the problem (4.10), and then maximizing $D(\lambda_1, \lambda_2)$ w.r.t. $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$. The reason is that the dual function represents a lower bound for the optimal value of the original (primal) optimization problem which is in general not tight for the non-convex case. Thus, the largest lower bound, i. e., the supremum of $D(\lambda_1, \lambda_2)$ w.r.t. $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$, is smaller than the optimal value of the problem (4.10) and we have a so-called *duality gap*. Nevertheless, Equation (4.13) allows for an investigation of the characteristics of the dual function for the specific problem at hand. To this end, we use the properties of the solution of a DARE with parameters which depend on scalar variables (see Appendix A4). Applying these results to the error covariance matrix $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}$ shown in Equation (4.15), it can be seen that its first derivative w.r.t. t and g, respectively, is positive semidefinite, whereas the second derivatives are negative semidefinite. These properties carry over to the error covariance matrix $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}$, which can be verified using Equation (4.5) and applying the steps analogous to Appendix A4 to the right hand side of the expression. Thus, we observe that

$$\frac{\partial^2 L^*}{\partial t^2} = \mathrm{tr}\left[\boldsymbol{P}\frac{\partial^2 \boldsymbol{C}_{\tilde{\boldsymbol{x}}}}{\partial t^2}\right] \leq 0 \qquad \text{and} \qquad \frac{\partial^2 L^*}{\partial g^2} = \mathrm{tr}\left[\boldsymbol{P}\frac{\partial^2 \boldsymbol{C}_{\tilde{\boldsymbol{x}}}}{\partial g^2}\right] \leq 0, \tag{4.18}$$

i. e., $L^*(t, g, \lambda_1, \lambda_2)$ is a concave function of t and g, respectively, because all terms of L^* except for $\mathrm{tr}\,[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}]$ are linear functions of the transmitter and receiver scaling. Since we are looking for the minimal value of a concave function over all non-negative values of t and g, depending on the values of λ_1, λ_2, $P_{\mathrm{Tx},1}$ and $P_{\mathrm{Tx},2}$, the infimum of L^* w.r.t. t and g is either $-\infty$ (with $t \to \infty$ or $g \to \infty$) or the finite positive value which is obtained by setting $t = 0$ and $g = 0$. The latter case which can easily constructed by setting $\lambda_1 = 0$ and $\lambda_2 = 0$ does not provide a valid solution, but the resulting infimum is approached in the limit for $t \to 0$ and $g \to 0$. This corresponds to an unbounded increase of the transmit powers for the observation and control channel (recall that the scaling factor at the system output is $t^{-\frac{1}{2}}$ and that the controller gain is scaled by $g^{-\frac{1}{2}}$, see Equation 4.16) and thus is not a feasible solution. Due to the concavity, the infimum of $L^*(t, g, \lambda_1, \lambda_2)$ w.r.t. t and g is $-\infty$ if $\lim_{t\to\infty} \frac{\partial L^*}{\partial t} < 0$ or $\lim_{g\to\infty} \frac{\partial L^*}{\partial g} < 0$. Since all summands of the concave function L^* except for $\mathrm{tr}\,[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}]$ are linear in t and g and the state estimation error increases monotonically with these two variables, this trace expression can at most behave linearly in the limit for $t \to \infty$ and $g \to \infty$. Thus, if $P_{\mathrm{Tx},1}$ or $P_{\mathrm{Tx},2}$ is large enough, the respective part of L^* with negative slope, i. e., $-\lambda_1 t P_{\mathrm{Tx},1}$ or $-\lambda_2 g P_{\mathrm{Tx},2}$, dominates for $t \to \infty$ or $g \to \infty$, which leads to the infimal value $-\infty$. This case corresponds to scaling factors $t^{-\frac{1}{2}}$ and $g^{-\frac{1}{2}}$ (recall that the optimal controller gain $\boldsymbol{L}$ contains the scaling factor $g^{-\frac{1}{2}}$) at the input of the observation and control channel, respectively, which go to zero. Thus, the variance of the receiver noise is amplified towards infinity together with the value of the LQG cost function, which is obviously not optimal if enough transmit power is available.

The fact that the infimum of the function L^* w.r.t. $t > 0$ and $g > 0$ corresponds either to an infeasible or obviously not optimal solution has two reasons. The main one is the non-convexity of the optimization problem (4.10). Due to this property, the dual function, i. e., the infimum of L^* w.r.t. t and g, provides only a lower bound for the optimal value of the original optimization problem which is not tight. The value $-\infty$ is a trivial lower bound, whereas the case $t \to 0$ and

[3] The dual function is determined using the infimum because the minimum may not exist.

$g \to 0$ provides a lower bound by violating the power constraints. This behavior is due to the second reason, i. e., the equivalent reformulation of the power constraints for the construction of the Lagrangian L (cf. Equation 4.11). For example, the constraints

$$P_1 \leq P_{\mathrm{Tx},1} \quad \text{and} \quad t > 0, \tag{4.19}$$

where P_1 is given by Equation (4.8), are equivalently expressed as

$$tP_1 \leq tP_{\mathrm{Tx},1} \quad \text{and} \quad t > 0. \tag{4.20}$$

The constraints for the power of the control channel are treated analogously. However, for $t \to 0$ or $g \to 0$, the corresponding transmit powers grow towards infinity whereas the product tP_1 or gP_2 converges[4] to a constant value. Thus, the Lagrangian associated with the formulation of the transmit power constraints according to Equation (4.20) does not grow towards infinity for $t \to 0$ or $g \to 0$. On the other hand, if the transmit powers P_1 and P_2 are feasible for $t \to \infty$ and $g \to \infty$, i. e., smaller than or equal to $P_{\mathrm{Tx},1}$ and $P_{\mathrm{Tx},2}$, respectively, the Lagrangian associated with the constraints according to Equation (4.20) may not be bounded from below even though the LQG cost grows towards infinity due to the unbounded noise amplification at the receivers of the observation and the control channel. In this case, the magnitude of the slope of $\lambda_1 t(P_1 - P_{\mathrm{Tx},1})$ is larger than the slope of the LQG cost function in the limit for $t \to \infty$ (and analogously for the power constraint of the control channel with $g \to \infty$).

Note that for both cases, i. e., for $t \to 0$ and $g \to 0$ or $t \to \infty$ and $g \to \infty$, the Lagrangian associated with the original formulation of the transmit power constraints according to Equation (4.19) grows towards infinity, either because of the unbounded transmit power or the unbounded cost function. Thus, the infimum w.r.t. t and g corresponds in general to finite and non-zero values of the transmit and receive scaling factors. Nevertheless, the optimization problem (4.10) remains non-convex and it is even harder to determine the dual function $D(\lambda_1, \lambda_2)$ because the associated function L^* has less exploitable properties, which is demonstrated in the following example.

Example 4.2.1 In order to illustrate the properties of the Lagrangian associated with the optimization problem (4.10) and of the function L^* (see Equation 4.13) as well as the effect of the different formulations of the transmit power constraints on this function, we use the parameters from Example 3.1.1. The available transmit powers $P_{\mathrm{Tx},1}$ and $P_{\mathrm{Tx},2}$ of the observation and the control channel are given by

$$\log_{10}\left(\operatorname{tr}\left[\boldsymbol{C_q}\right]^{-1} P_{\mathrm{Tx},1}\right) = \log_{10}\left(\operatorname{tr}\left[\boldsymbol{C_n}\right]^{-1} P_{\mathrm{Tx},2}\right) = 31.1,$$

where the actual transmit powers P_1 and P_2 are determined according to Equations (4.8) and (4.9), respectively, and the variance of the channel noises is assumed to be fixed. Additionally, the receive scaling factor g of the control channel is chosen to have the fixed value

$$\bar{g} = 2.00565,$$

and the Lagrange multipliers for the function L^* from Equations (4.11) and (4.13) are fixed to

$$\bar{\lambda}_1 = 9.44165 \quad \text{and} \quad \bar{\lambda}_2 = 24.0299.$$

[4]In the limit, the noise of the respective communication channel does not contribute to the state estimation error any more. Thus, the transmit powers of these channels are proportional to t^{-1} or g^{-1}, respectively.

Thus, $L^*_{\bar{g},\bar{\lambda}_1,\bar{\lambda}_2}(t) = L^*(t,\bar{g},\bar{\lambda}_1,\bar{\lambda}_2)$ is only a function of t in order to show the dependence on a single optimization variable. Note that the specific values of g, λ_1 and λ_2 are optimal for the problem (4.10) with the given system parameters, where we do not comment at this point on how they have been obtained. This question is treated in Section 4.3.4. Nevertheless, it can be seen in Figure 4.3 that the function $L^*_{\bar{g},\bar{\lambda}_1,\bar{\lambda}_2}$ is concave in $t > 0$. Its maximum is attained at

$$\bar{t} \approx 2.42593,$$

which is, together with $\bar{g}$, $\bar{\lambda}_1$ and $\bar{\lambda}_2$ and the resulting controller (the controller gain is given by Equation 4.16), the solution of problem (4.10). Again, we refer to Section 4.3.4 for a detailed treatment of the determination of these values.

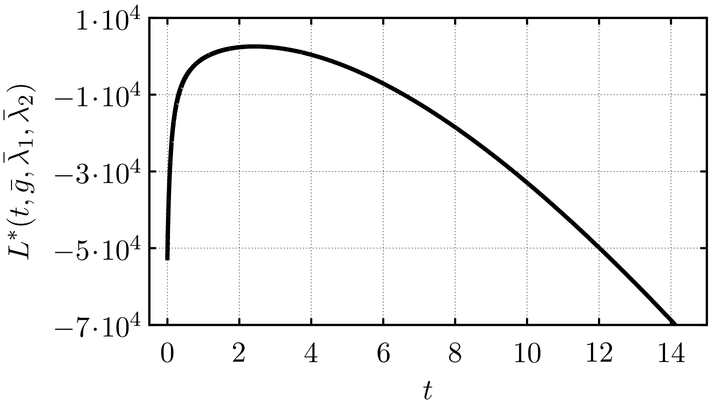


Figure 4.3: The function $L^*_{\bar{g},\bar{\lambda}_1,\bar{\lambda}_2}(t) = L^*(t,\bar{g},\bar{\lambda}_1,\bar{\lambda}_2)$ which has been obtained from the Lagrangian associated with problem (4.10) and the formulation of power constraints according to Equation (4.20). L^* is evaluated at $\bar{g}$, $\bar{\lambda}_1$ and $\bar{\lambda}_2$ and depicted as a function of t.

The picture changes if the Lagrangian associated with the optimization problem (4.10) is not constructed with the alternative formulation of the power constraints (see Equations 4.20 and 4.11) but with the original one, i. e., according to Equation (4.19), which results in the Lagrangian

$$\bar{L} = J_\infty + \nu_1(P_1 - P_{\mathrm{Tx},1}) + \nu_2(P_2 - P_{\mathrm{Tx},2}). \tag{4.21}$$

The Lagrange multipliers are now $\nu_1 \geq 0$ and $\nu_2 \geq 0$ in order to emphasize the difference to the Lagrangian L. The function $\bar{L}^*$ is obtained analogously to L^* by minimizing $\bar{L}$ w.r.t. the controller functions μ_k, $k \in \mathbb{N}_0$, which is again an LQG control problem (cf. optimization problem 4.12). Finally, we define

$$\bar{\nu}_1 = \bar{t}\bar{\lambda}_1 \quad \text{and} \quad \bar{\nu}_2 = \bar{g}\bar{\lambda}_2.$$

Figure 4.4 shows the respective values of $\bar{L}^*_{\bar{g},\bar{\nu}_1,\bar{\nu}_2}(t) = \bar{L}^*(t,\bar{g},\bar{\nu}_1,\bar{\nu}_2)$ for $t > 0$. As expected, the function increases for $t \to 0$, which is due to the increase of transmit power, and for $t \to \infty$, which is due to the increase of the cost function, i. e., J_∞.

It can also be seen in Figure 4.4 that $\bar{L}^*_{\bar{g},\bar{\nu}_1,\bar{\nu}_2}$ is neither convex nor concave and has three local minima. The global minimum is attained at $\bar{t}$ and is identical to the maximum of $L^*_{\bar{g},\bar{\lambda}_1,\bar{\lambda}_2}$ at $\bar{t}$. The fact that the value of both functions at $\bar{t}$ is the same is a result of the specific choice of $\bar{\nu}_1$ and $\bar{\nu}_2$. Since L (cf. Equation 4.11) and $\bar{L}$ (cf. Equation 4.21) only differ by the multiplication of the power constraints with t and g, respectively, both functions are identical if they are evaluated at $\nu_1 = \lambda_1 t$

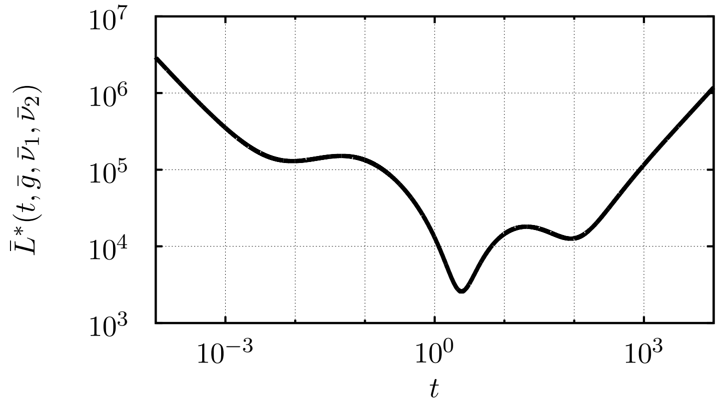


Figure 4.4: The function $\bar{L}^*$ which has been obtained from the Lagrangian associated with problem (4.10) and the formulation of power constraints according to Equation (4.19). $\bar{L}^*$ is evaluated at $\bar{g}$, $\bar{\nu}_1$ and $\bar{\nu}_2$ and depicted as a function of t.

and $\nu_2 = \lambda_2 g$, which is also true for the specific choice of these variables in the this example. The observation that $L^*_{\bar{g},\bar{\lambda}_1,\bar{\lambda}_2}(\bar{t})$ is the maximum of $L^*_{\bar{g},\bar{\lambda}_1,\bar{\lambda}_2}$ whereas $\bar{L}^*_{\bar{g},\bar{\nu}_1,\bar{\nu}_2}(\bar{t})$ with $\bar{\nu}_1 = \bar{t}\bar{\lambda}_1$ and $\bar{\nu}_2 = \bar{g}\bar{\lambda}_2$ is the global minimum of $\bar{L}^*_{\bar{g},\bar{\nu}_1,\bar{\nu}_2}$ is due to the fact that $\bar{t}$, $\bar{g}$, $\bar{\lambda}_1$ and $\bar{\lambda}_2$ actually solve the optimization problem (4.10). Since these values are optimal, the so-called *Karush-Kuhn-Tucker (KKT) conditions* (see, e. g., [95, Section 4.2.13]), which are necessary for a locally optimal feasible point, must hold at $t = \bar{t}$, i. e., the derivatives of $L^*_{\bar{g},\bar{\lambda}_1,\bar{\lambda}_2}$ and $\bar{L}^*_{\bar{g},\bar{\nu}_1,\bar{\nu}_2}$ w.r.t. t must vanish at $\bar{t}$. Although these conditions are in general not sufficient for optimality, the determination of KKT points provides candidates for the solution of the original optimization problem. If such an approach is taken, the Lagrangian L seems to be the better choice due to the concavity, at least individually in t and g, and thus the absence of local minima.

4.3 Solution of the Optimization Problem

If the values of the transmit and receive scaling factors t and g are assumed to be fixed, the resulting optimization of the LQG controller under power constraints is a convex optimization problem and the results of Chapter 3 can be readily applied. Unfortunately, we have seen in the preceding section that the joint optimization of controller, transmitter and receiver is a non-convex problem and it is not obvious how to determine the solution. Even the question if the power constrained optimization problem is feasible, i. e., if a solution exists such that the transmit powers of the observation and the control channel are smaller than or equal to a predetermined value, has to be asked again. If the problem is feasible for a specific choice of t and g, it is not ensured that this is true for arbitrary values, e. g., if t and g are chosen to be so small such that the resulting amplification of the transmit signals can not be compensated by the controller.

4.3.1 Feasibility

As in Chapter 3, it is of interest how the available transmit powers $P_{\text{Tx},1}$ and $P_{\text{Tx},2}$ for the observation and the control channel, respectively, can be chosen such that the power constrained LQG optimization problem (4.10) is feasible. In order to answer this question, we determined Pareto optimal values of the transmit powers in Section 3.3, i. e., pairs of $(P_{\text{Tx},1}, P_{\text{Tx},2})$ such that no controller exist which achieves a transmit power for the observation channel which is smaller than $P_{\mathrm{Tx},1}$ and at the same time a transmit power of the control channel which is smaller than $P_{\mathrm{Tx},2}$. Such points have been computed by minimizing the weighted sum of both transmit powers. For the convex setting in Chapter 3, all Pareto optimal transmit powers can be calculated using this approach (see Section 3.3). However, with the additional degrees of freedom provided by the transmit and receive scaling factors t and g, the problem of finding Pareto optimal transmit powers becomes more involved because, like the original problem (4.10), the joint minimization of the weighted sum of the transmit powers w.r.t. the controller and the transceiver is a non-convex problem. This will be discussed in more detail in Section 4.3.2.

In Chapter, 3 the feasibility question has been treated by the exploration of the set of Pareto optimal transmit powers. However, it is not always practical to determine this set (or approximate it by computing a sufficient number of Pareto optimal transmit powers) in order to decide if a certain pair $(P_{\mathrm{Tx},1}, P_{\mathrm{Tx},2})$ of available transmit powers leads to a feasible optimization problem for a specific choice of t and g. Thus, the approach described in [78, Section 14.5] can be used to decide if a tuple $(t, g, P_{\mathrm{Tx},1}, P_{\mathrm{Tx},2})$ is feasible. The basic idea is to replace the LQG cost function for the power constrained controller optimization by a constant, e. g., the value 0, and to consider the Lagrangian associated with this modified problem. With these changes, Equation (4.11) becomes

$$\begin{aligned} L_{\mathrm{feas}} &= 0 + \lambda_1 t (P_1 - P_{\mathrm{Tx},1}) + \lambda_2 g (P_2 - P_{\mathrm{Tx},2}) \\ &= \lim_{N\to\infty} \frac{1}{N} \mathrm{E}\left[\sum_{n=0}^{N-1} \begin{bmatrix} \boldsymbol{x}_n \\ \boldsymbol{u}_n \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \lambda_1 \boldsymbol{C}^{\mathrm{T}}\boldsymbol{C} & \\ & \lambda_2 g \mathbf{I}_{N_u} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_n \\ \boldsymbol{u}_n \end{bmatrix}\right] + \lambda_1 \operatorname{tr}[\boldsymbol{C}_{\boldsymbol{v}}] - \lambda_1 t P_{\mathrm{Tx},1} - \lambda_2 g P_{\mathrm{Tx},2}, \end{aligned} \tag{4.22}$$

and the minimization of L_{feas} w.r.t. the controller μ_k, $k \in \mathbb{N}_0$, analogous to the optimization problem (4.12), results in the solution

$$L_{\mathrm{feas}}^{*}(\lambda_1, \lambda_2) = \operatorname{tr}\left[\boldsymbol{K}\left(\boldsymbol{C}_{\boldsymbol{w}} + g\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\right)\right] + \operatorname{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\right] + \lambda_1 \operatorname{tr}[\boldsymbol{C}_{\boldsymbol{v}}] - \lambda_1 t P_{\mathrm{Tx},1} - \lambda_2 g P_{\mathrm{Tx},2}, \tag{4.23}$$

which depends on the stabilizing solutions of the DAREs

$$\boldsymbol{K} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B}\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B} + \lambda_2 \mathbf{I}_{N_u}\right)^{-1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} + \lambda_1 \boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}, \tag{4.24}$$

and

$$\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} = \boldsymbol{A}\left(\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}} + t\boldsymbol{C}_{\boldsymbol{q}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\right)\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{w}} + g\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}. \tag{4.25}$$

Additionally, $\boldsymbol{P} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \boldsymbol{K} + \lambda_1 \boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}$ and $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}$ is given by Equation (4.5). Note that compared to Equation (4.13), the arguments t and g have been omitted in Equation (4.23) because they are not considered to be optimization variables at this point and are interpreted as given system parameters. Consequently, L_{feas}^{*} is the dual function associated with the problem of minimizing a constant cost function subject to the constraints of the optimization problem (4.10). Although the cost function does not depend on the optimization variables in this case, the dual function can be used to determine the feasibility, i. e., the existence of a controller such that the constraints of problem (4.10) are satisfied for fixed values of t and g. To this end, the dual function L_{feas}^{*} is maximized w.r.t. λ_1 and λ_2. The usual constraint that $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ is extended by the additional (convex) constraint $\lambda_1 + \lambda_2 = 1$. This normalization is introduced because the multiplication of the Lagrangian L_{feas} by any positive value simply results in the scaling of the respective minimum

L_{feas}^* with the same value. Note that the three constraints for the Lagrange multipliers can be equivalently replaced by $\lambda_1 = \rho$, $\lambda_2 = 1 - \rho$ and $\rho \in [0, 1]$. Using this formulation, the dual of the feasibility problem is given by

$$\underset{\rho}{\text{maximize}}\, L_{\text{feas}}^*(\rho, 1-\rho) \quad \text{subject to} \quad \rho \in [0, 1], \tag{4.26}$$

which is a convex problem because a concave function (the dual function is always concave) is maximized over a convex set. Due to the dependence of the solution $\boldsymbol{K}$ of the DARE in Equation (4.24) on the Lagrange multipliers λ_1 and λ_2 or ρ, respectively, the results of Appendix A4 can be used to determine the first and second derivative of the dual function L_{feas}^* w.r.t. ρ, which allows for a variety of possibilities for the solution of problem (4.26), e. g., a gradient ascent, the Newton algorithm or line search approaches like bisection.

The sign of the optimal value of problem (4.26) tells us if the power constrained LQG optimization is feasible for given values of t, g, $P_{\text{Tx},1}$ and $P_{\text{Tx},2}$. A controller which satisfies the power constraints exists if and only if the maximum determined by (4.26) is smaller than or equal to zero. For a detailed treatment of this statement we refer to [78, Section 14.5]. However, its validity is quite intuitive. Note that the minimization of the Lagrangian L_{feas} with $\lambda_1 = \rho$ and $\lambda_2 = 1 - \rho$, where $\rho \in [0, 1]$, w.r.t. to the controller is essentially identical to optimization problem (3.33) which is used to determine Pareto optimal transmit powers. The only difference is that the system model is extended by t and g and that the cost function differs in the subtractive terms $-\rho t P_{\text{Tx},1}$ and $-(1-\rho) g P_{\text{Tx},2}$ which do not depend on the controller. Consequently, $L_{\text{feas}}^*(\rho, 1-\rho)$ provides the difference between the weighted sum power $\rho t P_1 + (1-\rho) g P_2$, which is minimized by a specific controller, and the respective weighted sum of the available transmit powers, given by $\rho t P_{\text{Tx},1} + (1-\rho) g P_{\text{Tx},2}$. If this difference is positive for some value of ρ, at least one of the power constraints must be violated because the actual transmit power is larger than the available one. On the other hand, if even the maximal difference, given by the maximum of L_{feas}^* w.r.t. $\rho \in [0, 1]$, is negative (or zero), there exists at least one controller which fulfills the power constraints, i. e., the controller that minimizes L_{feas}, and thus provides a Pareto optimal pair of transmit powers.

In order to get an intuition about the possible choices of t and g such that the power constrained LQG controller optimization is feasible and to provide some insights for possible approaches to the optimal choice of the transmit and receive scaling, the following example applies the feasibility test described above to a set of sampled values of t and g.

Example 4.3.1 The parameters of the dynamical system as well of the noise sequences which drive the system are again taken from Example 3.1.1. With the fixed channel noise variances, the available transmit powers $P_{\text{Tx},1}$ and $P_{\text{Tx},2}$ are determined by the respective SNRs of the observation and the control channel which are chosen to be

$$10 \log_{10}\left(\operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{q}}\right]^{-1} P_{\text{Tx},1}\right) = 10 \log_{10}\left(\operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{n}}\right]^{-1} P_{\text{Tx},2}\right) = 31.1.$$

Note that the transmit powers $P_{\text{Tx},1}$ and $P_{\text{Tx},2}$ lead to an infeasible optimization problem for $t = 1$ and $g = 1$, where this choice of t and g corresponds to the case considered in problem (3.33), since they lie outside the region shown in Figure 3.4. Nevertheless, the additional degrees of freedom provided by the scaling of transmit and receive signals allow us to determine a solution which satisfies the power constraints. To this end, the intervals $t \in [0, 14]$ and $g \in [0, 7]$ are sampled equidistantly to a set of $5 \cdot 10^5$ different pairs of transmit and receive scaling factors. The feasibility test described above is applied to each pair. Figure 4.5 shows the result of this procedure, where a feasible point is depicted by a dot and only a subset of 2500 of such points which are randomly chosen are shown due to the large amount of data. Approximately 25% of all tested pairs of t and g resulted in a feasible optimization problem.

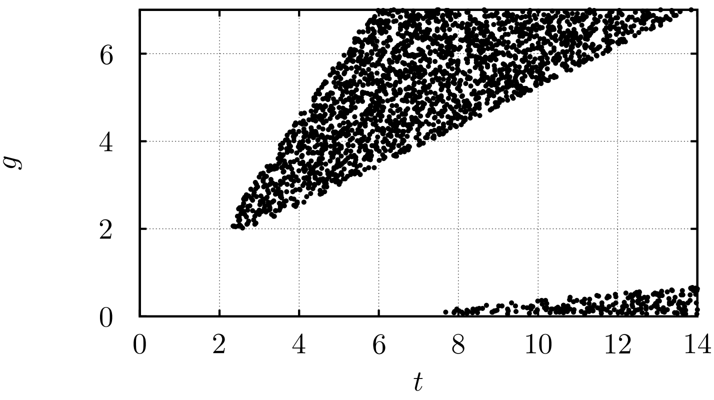


Figure 4.5: Feasibility of the optimization problem (4.10) for $10\log_{10}\left(\operatorname{tr}\left[\boldsymbol{C_q}\right]^{-1} P_{\mathrm{Tx},1}\right) = 31.1$ and $10\log_{10}\left(\operatorname{tr}\left[\boldsymbol{C_n}\right]^{-1} P_{\mathrm{Tx},2}\right) = 31.1$ and given values t and g. A feasible pair (t, g) is denoted by a dot.

It is interesting to see that the set of transmit and receive scaling factors which lead to a feasible power constrained controller optimization seems to consist of two disjoint regions. This is a problem if a local optimization algorithm is applied to the determination of optimal values of t and g due to the dependence of the resulting local optimum on the initialization of such an algorithm. Of course this is not surprising since the problem of joint controller and transceiver optimization is non-convex. A second observation is that the shape of the region seems to imply that if a certain point $(\bar{t}, \bar{g})$ is feasible, then the point $(\beta\bar{t}, \beta\bar{g})$ with $\beta \geq 1$, i. e., a pair with the same ratio of transmit and receive scaling factors which are not smaller that $\bar{t}$ and $\bar{g}$, respectively, also leads to a feasible optimization problem. We will see in the following section that the ratio of t and g is actually of importance for the joint optimization of controller and transceiver with power constraints.

4.3.2 Achievable Transmit Powers

In the preceding section, the feasibility of the power constrained LQG control problem for fixed values of the transmit and receive scaling factors t and g, respectively, and available transmit powers $P_{\mathrm{Tx},1}$ and $P_{\mathrm{Tx},2}$ has been treated. At this point, we come back to the question of how small the values of $P_{\mathrm{Tx},1}$ and $P_{\mathrm{Tx},2}$ can be chosen if the degrees of freedom provided by the transceiver are used and t and g are adapted in order to minimize the LQG cost function while satisfying the transmit power constraints. As a first step to answer the question, we determine Pareto optimal transmit powers analogous to the approach presented in Section 3.3, i. e., the weighted sum of the transmit powers of the observation and the control channel is minimized. The difference is that now this sum is minimized w.r.t. the LQG controller and additionally the transmit and receive scaling factors, which turns the convex optimization problem from Section 3.3 into a non-convex one.

Before we consider the optimization problem for the determination of Pareto optimal values of the transmit powers, the state and observation equation which describe the dynamical system to be controlled and which have been used above, e. g., for the optimization problem (4.10), are equivalently reformulated. This facilitates the line of argumentation in the following. Define the scaled state of the dynamical system $\boldsymbol{z}_k = g^{-\frac{1}{2}}\boldsymbol{x}_k$, $k \in \mathbb{N}_0$, and the scaled system output $\boldsymbol{\eta}_k = g^{-\frac{1}{2}}\boldsymbol{y}_k$, $k \in \mathbb{N}_0$. This step is not problematic since $g > 0$, i. e., $g \neq 0$. Using this notation, the corresponding dynamical system is described by the state and observation equation

$$\begin{aligned} \boldsymbol{z}_{k+1} &= \boldsymbol{A}\boldsymbol{z}_k + \boldsymbol{B}\left(\boldsymbol{u}_k + \boldsymbol{n}_k\right) + g^{-\frac{1}{2}}\boldsymbol{w}_k, \\ \boldsymbol{\eta}_k &= \boldsymbol{C}\boldsymbol{z}_k + g^{-\frac{1}{2}}\boldsymbol{v}_k. \end{aligned} \tag{4.27}$$

Up to the scaling of the system output with $g^{-\frac{1}{2}}$, the system above has the same input-output behavior as the original one with state $\boldsymbol{x}_k$ and the scaling factor $g^{\frac{1}{2}}$ at the system input. The dynamical system described by Equation (4.27) is the result of pulling $g^{\frac{1}{2}}$ from the input to the output of the system. Finally, define the ratio of the transmit and receive scaling as

$$\alpha = \frac{t}{g}. \tag{4.28}$$

Using these definitions, the system model which has been introduced at the beginning of this chapter and depicted in Figure 4.1 is equivalently represented by Figure 4.6. The transmit power

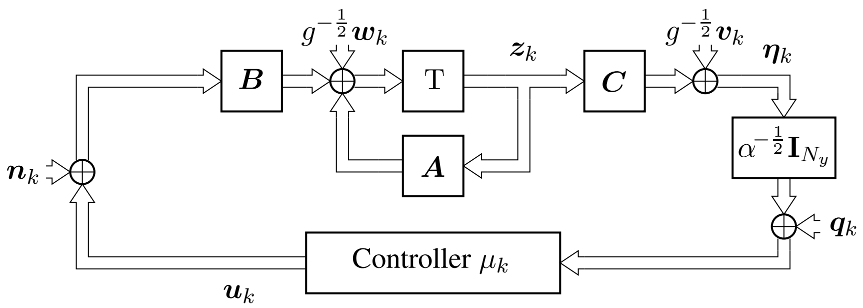


Figure 4.6: Model of the control loop which is closed over two channels with additive noise $\boldsymbol{q}_k$ and $\boldsymbol{n}_k$. The scaling factor $g^{\frac{1}{2}}$ is shifted from the input to the output of the system, resulting in the transmit scaling factor $\alpha^{-\frac{1}{2}} = \left(\frac{t}{g}\right)^{-\frac{1}{2}}$.

of the observation channel which is given by Equation (4.8) now reads as

$$P_1 = \lim_{N\to\infty}\frac{1}{N}\,\mathrm{E}\left[\alpha^{-1}\left(\sum_{n=0}^{N-1}\boldsymbol{z}_n^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{z}_n + g^{-1}\boldsymbol{v}_n^{\mathrm{T}}\boldsymbol{v}_n\right)\right], \tag{4.29}$$

whereas the transmit power P_2 of the control channel is given by Equation (4.9). Analogous to the optimization problem (3.33), we use the cost function $\rho P_1 + (1-\rho)P_2$ with $\rho \in [0,1]$, i. e., the weighted sum of the transmit powers, for the determination of Pareto optimal values of these powers. With the reformulation of the system model from above, this cost function fits in the LQG framework with $\boldsymbol{Q} = \rho\alpha^{-1}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}$, $\boldsymbol{R} = (1-\rho)\mathbf{I}_{N_u}$ and $\boldsymbol{S} = \mathbf{0}_{N_x\times N_u}$. Consequently, the minimization of the weighted sum of the transmit powers reads as

$$\underset{\alpha,g,\mu_0,\mu_1,\mu_2\ldots}{\text{minimize}} \lim_{N\to\infty}\frac{1}{N}\,\mathrm{E}\left[\sum_{n=0}^{N-1}\begin{bmatrix}\boldsymbol{z}_n\\ \boldsymbol{u}_n\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\rho\alpha^{-1}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C} & \\ & (1-\rho)\mathbf{I}_{N_u}\end{bmatrix}\begin{bmatrix}\boldsymbol{z}_n\\ \boldsymbol{u}_n\end{bmatrix}\right]+\rho(\alpha g)^{-1}\operatorname{tr}[\boldsymbol{C}_{\boldsymbol{v}}] \tag{4.30}$$

$$\begin{aligned}
\text{subject to}\quad & \boldsymbol{z}_{k+1}=\boldsymbol{A}\boldsymbol{z}_k+\boldsymbol{B}(\boldsymbol{u}_k+\boldsymbol{n}_k)+g^{-\frac{1}{2}}\boldsymbol{w}_k, && k\in\mathbb{N}_0,\\
& \boldsymbol{\eta}_k=\boldsymbol{C}\boldsymbol{z}_k+g^{-\frac{1}{2}}\boldsymbol{v}_k, && k\in\mathbb{N}_0,\\
& \boldsymbol{u}_k=\mu_k(\mathcal{I}_k), && k\in\mathbb{N}_0,\\
& \mathcal{I}_k=\begin{cases}\{(\alpha^{-\frac{1}{2}}\boldsymbol{\eta}_0+\boldsymbol{q}_0)\}, & k=0,\\ \{(\alpha^{-\frac{1}{2}}\boldsymbol{\eta}_0+\boldsymbol{q}_0),\ldots,(\alpha^{-\frac{1}{2}}\boldsymbol{\eta}_k+\boldsymbol{q}_k),\boldsymbol{u}_0,\ldots,\boldsymbol{u}_{k-1}\}, & k\in\mathbb{N},\end{cases}\\
& \alpha>0,\quad g>0.
\end{aligned}$$

Since this minimization is an LQG control problem without power constraints for fixed values of α and g, the minimum w.r.t. μ_k, $k\in\mathbb{N}_0$, is known and given by (cf. Appendix A6.3)

$$J_\infty^*(\alpha,g)=\operatorname{tr}\left[\boldsymbol{K}\left(g^{-1}\boldsymbol{C}_{\boldsymbol{w}}+\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\right)\right]+\operatorname{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}\right]+\rho(\alpha g)^{-1}\operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{v}}\right], \tag{4.31}$$

which depends on the stabilizing[5] solutions of

$$\boldsymbol{K}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A}-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B}\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B}+(1-\rho)\mathbf{I}_{N_u}\right)^{-1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A}+\rho\alpha^{-1}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C} \tag{4.32}$$

and

$$\begin{aligned}
\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}&=\boldsymbol{A}\left(\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}-\alpha^{-1}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\left(\alpha^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}\boldsymbol{C}+(\alpha g)^{-1}\boldsymbol{C}_{\boldsymbol{v}}+\boldsymbol{C}_{\boldsymbol{q}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}\right)\boldsymbol{A}^{\mathrm{T}}+g^{-1}\boldsymbol{C}_{\boldsymbol{w}}+\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\\
&=\boldsymbol{A}\left(\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}-\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}\boldsymbol{C}+g^{-1}\boldsymbol{C}_{\boldsymbol{v}}+\alpha\boldsymbol{C}_{\boldsymbol{q}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}\right)\boldsymbol{A}^{\mathrm{T}}+g^{-1}\boldsymbol{C}_{\boldsymbol{w}}+\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}},
\end{aligned} \tag{4.33}$$

where $\boldsymbol{P}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A}-\boldsymbol{K}+\rho\alpha^{-1}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}$ and $\boldsymbol{C}_{\tilde{\boldsymbol{z}}}$ is the covariance matrix of the (scaled) state estimation error which is given by $\boldsymbol{C}_{\tilde{\boldsymbol{z}}}=\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}-\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}\boldsymbol{C}+g^{-1}\boldsymbol{C}_{\boldsymbol{v}}+\alpha\boldsymbol{C}_{\boldsymbol{q}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}$. Note that J_∞^* in Equation (4.31) is denoted as a function of α and g in order to emphasize the fact that it still has to be minimized w.r.t. $\alpha>0$ and $g>0$. Having a closer look at the cost function in Equation (4.31) and the expressions of the covariance matrices $\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}$ and $\boldsymbol{C}_{\tilde{\boldsymbol{z}}}$ of the state estimation error (cf. Equation 4.33) it can be seen that the minimum of J_∞^*, or precisely the infimum, w.r.t. g for a given value of $\alpha>0$ is achieved for $g\to\infty$, which corresponds to a dynamical system with no process or observation noise. This statement is easy to verify since the summands $\operatorname{tr}[\boldsymbol{K}g^{-1}\boldsymbol{C}_{\boldsymbol{w}}]$ and $\rho(\alpha g)^{-1}\operatorname{tr}[\boldsymbol{C}_{\boldsymbol{v}}]$ decrease monotonically towards zero in this case. Additionally, the estimation error variance and thus $\operatorname{tr}[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}]$ also decreases monotonically if the noise covariance matrices $\boldsymbol{C}_{\boldsymbol{w}}$ and $\boldsymbol{C}_{\boldsymbol{v}}$ in Equation (4.33) are scaled down by g^{-1}, which can be confirmed using the results in Appendix A4 and [100]. Note that for a fixed value of α, letting $g\to\infty$ implies together with Equation (4.28) that also $t\to\infty$. Keeping in mind that this limit corresponds to the consideration of a system with no process and observation noise, we are effectively left with a scenario where

[5] We have to keep in mind that the closed loop system is required to be stable. For the specific case considered here, the choice of $\rho=1$ can lead to the fact that Equation (4.32) has a non-stabilizing positive semidefinite solution which minimizes the cost function given by Equation (4.31), see, e. g., Appendix A3 and [101].

only the noise sequences $(\boldsymbol{q}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{n}_k : k \in \mathbb{N}_0)$ of the observation and control channel, respectively, drive the closed loop system and thus contribute to the transmit powers of both channels. Such a scenario has been investigated, e. g., in [29]. With the definition of

$$\begin{aligned} I(\alpha) &= \inf_{g>0} J_\infty^*(\alpha, g) \\ &= \operatorname{tr}\left[\boldsymbol{K}\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\right] + \operatorname{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}\right], \end{aligned} \tag{4.34}$$

where the estimation error covariance matrix $\boldsymbol{C}_{\tilde{\boldsymbol{z}}} = \boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}\boldsymbol{C} + \alpha\boldsymbol{C}_{\boldsymbol{q}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}$ is determined using Equation (4.33) with g^{-1} replaced by 0 and where $\boldsymbol{K}$ and $\boldsymbol{P}$ are determined according to Equation (4.32), the remaining part for the solution of optimization problem (4.30) is

$$\underset{\alpha}{\text{minimize}} \;\; I(\alpha) \quad \text{subject to} \quad \alpha > 0. \tag{4.35}$$

Note that despite the fact that the infimum of the weighted sum of transmit powers which is given by I is finite for $\alpha > 0$, the resulting variance of the original system state $\boldsymbol{x}_k$, $k \in \mathbb{N}_0$, and thus of the control signal $\boldsymbol{u}_k$, $k \in \mathbb{N}_0$, is unbounded from above for $g \to \infty$ (together with $t \to \infty$). This is easy to see because the input of the dynamical system is scaled by $g^{\frac{1}{2}}$ which leads to an unbounded amplification of the control channel noise $(\boldsymbol{n}_k : k \in \mathbb{N}_0)$ (see Figures 4.1 and 4.2). Analogously, at least for unstable dynamical systems, the estimation error of the system state grows towards infinity if the observations at the output of the system are scaled by $t^{-\frac{1}{2}}$ which goes to zero. Consequently, the performance of the closed loop control system which is measured by an LQG cost function becomes arbitrarily worse in general if it is desired to achieve Pareto optimal transmit powers.

Although the minimization (4.35) is only w.r.t. to one real variable, its solution is not straightforward, which is expected due to the non-convexity of the original problem (4.30) and illustrated by the following example.

Example 4.3.2 Consider the model parameters provided by Example 3.1.1. The optimization problem (4.30) is solved for the parameter $\rho = \frac{2}{3}$ and a given value of the variable α. Thus, the solution provides the function $I(\alpha)$ shown in Equation (4.34), i. e., the infimum w.r.t. $g > 0$ of

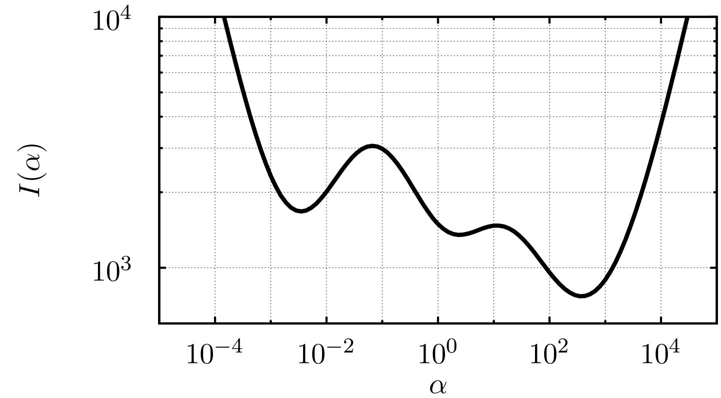


Figure 4.7: Infimum of the weighted sum power $\rho P_1 + (1-\rho)P_2$ with parameter $\rho = \frac{2}{3}$ for $g \to \infty$ (cf. optimization problem 4.30 and Equation 4.34).

the weighted sum of the transmit powers which has been optimized w.r.t. the LQG controller μ_k, $k \in \mathbb{N}_0$. In order to get an impression about the final minimization w.r.t. to $\alpha > 0$, this function is evaluated for a large range of values of α and shown in Figure 4.7.

Note that in the limit for $\alpha \to 0$ and $\alpha \to \infty$, I behaves as expected, i. e., is not bounded from above, because either the transmit power of the observation or the control channel is scaled by a factor which goes to infinity. Between these two limits, the optimized weighted sum I of the transmit powers exhibits one global minimum and additionally two local minima.

The example above shows that it is not straightforward to determine the solution of the minimization problem (4.35). One approach is of course a sampling of the function I for a finite set of values of the optimization variable α. It is easy to implement, but has the general problem of an appropriate selection of the sampling range and the sampling intervals. Additionally, it is not guaranteed that the global optimum of I is actually determined. In order to avoid these problems and to obtain a certificate about the optimality of a solution candidate of problem (4.35), the framework of monotonic optimization is introduced and applied to the problem of the determination of Pareto optimal transmit powers.

4.3.3 Determination of Achievable Powers by Monotonic Optimization

For the determination of the solution of the problems (4.30) and (4.35), respectively, efficient algorithms for convex optimization can not be used due to the lack of convexity. The minimization of the weighted sum of the transmit powers w.r.t. the ratio $\alpha > 0$ of the transmit and receive scaling factors does not even offer a quasi-convex structure but the respective cost function exhibits several local minima (cf. Figure 4.7). Thus, line search algorithms like the golden section method or bisection to determine the root of the derivative (see, e. g., [95, Chapter 8]) are not applicable for the determination of the global minimum. Fortunately, there is usable structure left which allows us to apply the global optimization framework of monotonic optimization.

In [102], the authors propose several algorithms to exploit the monotonicity of the cost function and the constraints of an optimization problem to obtain globally optimal solutions when no other simplifying structure is present. In this thesis, we concentrate on the branch and bound approach which is based on a simple idea. By an exhaustive partitioning, i. e., branching, of the set the optimization variables can be taken from, the optimal value of the problem under consideration is determined. In order to reduce complexity, bounds for each set of the partition are generated which allow for the decision if the optimal value of the problem lies within a certain set or not.

The following example of the approach is used to illustrate the idea and the necessary steps of the branch and bound procedure. Assume that the optimization variable α is element of an interval $A = [a, b]$, $a < b$, which is a non-empty subset of $\mathbb{R}$, and we want to determine the value of α within this interval which minimizes a function I of α. Assume further that this function can be expressed as $I = I_1 - I_2$, where both I_1 and I_2 are monotonically increasing functions of α, i. e., for $\alpha_1 < \alpha_2$ it follows that $I_k(\alpha_1) \leq I_k(\alpha_2)$, $k \in \{1, 2\}$. No additional properties are required. In the branching step, the interval A is partitioned into a number of subsets which cover the original set. For the sake of simplicity, let this partition be $A = A_1 \cup A_2$ with $A_1 = [a, c]$, $A_2 = [c, b]$ and $a < c < b$. Now, the function I is evaluated at one point in each subset in order to obtain an upper bound for the minimal value of the function within this interval. One possibility is to use a boundary point of each interval, e. g., the upper bounds are determined by $\overline{I}_{A_1} = I(a)$ for A_1 and $\overline{I}_{A_2} = I(c)$ for A_2. The next step of the bounding procedure is of great importance and determines a non-trivial lower bound for the function I for each subset. At this point, the monotonicity of I_1 and I_2 is used. Since both functions are increasing, a lower bound for I is obtained by evaluating the positive summand I_1 at the lower boundary point of the considered interval and the negative summand I_2 at the upper boundary point. This minimizes the positive contribution to I and maximizes the summand which is subtracted. The corresponding lower bounds in our example are thus given by $\underline{I}_{A_1} = I_1(a) - I_2(c)$ for A_1 and $\underline{I}_{A_2} = I_1(c) - I_2(b)$ for A_2. The upper and lower bounds for the minimal value of I within each subset of A are now used to determine if the optimal value of this function over the whole interval A can lie in one of its subsets or not. Even if the global minimum is not known, it is not possible that it lies in a subset A_k of A which has a lower bound $\underline{I}_{A_k}$ with a larger value than the smallest upper bound obtained so far, i. e., if $\underline{I}_{A_k} > \min_n \overline{I}_{A_n}$. Such sets can be excluded from any further investigation of the minimization problem. With the remaining subsets of A, the branching and bounding procedure described above is repeated. Due to the increasing refinement of the partitioning of A with each step of the branching, the global minimum of I w.r.t. $\alpha \in A$ is approximated with increasing accuracy, provided that if a subset of A collapses to a single point, the value of the corresponding lower bound converges to the actual value of the function I at this point. This property of the lower bound is called consistency (cf. [102, Section 7.5]).

Note that the approach described above can be interpreted as a sampling of the function I within the interval A. The difference to a conventional sampling approach is that there is no a priori decision about the sampling method, e. g., equidistant, logarithmic or random, or the resolution, i. e., the number of sampling points which are evaluated. Additionally, the branch and bound procedure provides information about subsets of A which do not need to be sampled because they do not contain the minimizer of I. Nevertheless, the most important feature is that the presented approach allows for a statement about the (sub-)optimality of the result of the branch and bound process. Due to the availability of upper bounds, i. e., actually sampled values of I, and lower bounds for the optimal value of I, the gap between the smallest upper and smallest lower bound at a certain point of the branching process tells us how far we are away from the global optimum. Thus, we get a certificate for the optimality of a sampled point which is not available for any standard sampling approach. The branching process can be stopped if a desired accuracy is obtained.

4.3.3.1 Minimization of Weighted Sum of Transmit Powers

After the short description of the branch and bound method which exploits the underlying monotonicity of an optimization problem in order to find a solution, we come back to the computation of Pareto optimal values for the transmit powers in a control loop which has been discussed in Section 4.3.2. Such values can be determined by minimizing the weighted sum of the transmit powers of the observation and the control channel. We have seen that the minimization w.r.t. to the controller μ_k, $k \in \mathbb{N}_0$, is a standard LQG problem and that the infimum of the weighted sum of the powers is achieved in the limit for $g \to \infty$. At this point, we start with the remaining optimization problem (4.35) for the determination of the optimal ratio α of the transmit and receive scaling factors t and g (cf. Equation 4.28). However, it is not possible to determine the lower bound for the branch and bound approach as described above because the function $I(\alpha) = \operatorname{tr}\left[\boldsymbol{K}\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\right] + \operatorname{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}\right]$ (see Equation 4.34) can not be simply rewritten as a difference $I_1 - I_2$ of increasing functions. Thus, we use the fact that the matrix $\boldsymbol{P} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \boldsymbol{K} + \rho\alpha^{-1}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}$ allows us to obtain an alternative expression for I, i. e.,

$$
\begin{aligned}
I(\alpha) &= \operatorname{tr}\left[\boldsymbol{K}\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\right] + \operatorname{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}\right] \\
&= \operatorname{tr}\left[\boldsymbol{K}\left(\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}} + \boldsymbol{A}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}\boldsymbol{A}^{\mathrm{T}} - \boldsymbol{C}_{\tilde{\boldsymbol{z}}}\right)\right] + \rho\alpha^{-1}\operatorname{tr}\left[\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}\boldsymbol{C}^{\mathrm{T}}\right] \\
&= \operatorname{tr}\left[\boldsymbol{K}\left(\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{z}}}\right)\right] + \rho\alpha^{-1}\operatorname{tr}\left[\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}\boldsymbol{C}^{\mathrm{T}}\right],
\end{aligned}
\tag{4.36}
$$

where $\boldsymbol{K}$ is given by Equation (4.32) and $\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{z}}}$ by Equation (4.33) with g^{-1} replaced by 0. The covariance matrix of the scaled state estimation error is $\boldsymbol{C}_{\tilde{\boldsymbol{z}}} = \boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{z}}} - \boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{z}}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{z}}}\boldsymbol{C}+\alpha\boldsymbol{C}_{\boldsymbol{q}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{z}}}$. Using the results of Appendix A4, it can be seen that the matrix $\boldsymbol{K}$ is increasing in α^{-1}, i.e., $\frac{\partial \boldsymbol{K}}{\partial(\alpha^{-1})} \geq \boldsymbol{0}_{N_x\times N_x}$, and thus decreasing in α. On the other hand, $\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{z}}}$ is increasing in α, i.e., $\frac{\partial \boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{z}}}}{\partial\alpha} \geq \boldsymbol{0}_{N_x\times N_x}$, which is also true for $\boldsymbol{C}_{\tilde{\boldsymbol{z}}}$. These monotonicity results are used to obtain lower bounds for the branch and bound algorithm. Let $\alpha \in A = [\underline{\alpha}, \overline{\alpha}]$ with $0 < \underline{\alpha} \leq \overline{\alpha}$. Then, it holds

$$\begin{aligned} I(\alpha) &= \operatorname{tr}\left[\boldsymbol{K}(\alpha)\left(\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{z}}}(\alpha) - \boldsymbol{C}_{\tilde{\boldsymbol{z}}}(\alpha)\right)\right] + \rho\alpha^{-1}\operatorname{tr}\left[\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}(\alpha)\boldsymbol{C}^{\mathrm{T}}\right] \\ &\geq \operatorname{tr}\left[\boldsymbol{K}(\overline{\alpha})\left(\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{z}}}(\alpha) - \boldsymbol{C}_{\tilde{\boldsymbol{z}}}(\alpha)\right)\right] + \rho\overline{\alpha}^{-1}\operatorname{tr}\left[\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}(\alpha)\boldsymbol{C}^{\mathrm{T}}\right] \\ &= \operatorname{tr}\left[\boldsymbol{K}(\overline{\alpha})\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\right] + \operatorname{tr}\left[\boldsymbol{P}(\overline{\alpha})\boldsymbol{C}_{\tilde{\boldsymbol{z}}}(\alpha)\right] \\ &\geq \operatorname{tr}\left[\boldsymbol{K}(\overline{\alpha})\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\right] + \operatorname{tr}\left[\boldsymbol{P}(\overline{\alpha})\boldsymbol{C}_{\tilde{\boldsymbol{z}}}(\underline{\alpha})\right] \\ &= \underline{I}_A. \end{aligned} \tag{4.37}$$

Note that in Equation (4.37) the matrices $\boldsymbol{K}$, $\boldsymbol{P}$, $\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{z}}}$ and $\boldsymbol{C}_{\tilde{\boldsymbol{z}}}$ are explicitly denoted as functions of α in order to illustrate which element of I is lower bounded using the corresponding boundary point of A. It can be seen that $\underline{I}_A$ is a lower bound for $I(\alpha)$ for all $\alpha \in A$ since it only depends on the boundary points of A. Thus, it is also a lower bound for the minimal value of I within this interval. In order to obtain an upper bound for the minimum, I is evaluated at some $\alpha \in A$, e.g., at $\overline{\alpha}$, which results in

$$\overline{I}_A = I(\overline{\alpha}). \tag{4.38}$$

Remark: The specific choice of the lower bound $\underline{I}_A$ given by Equation (4.37) can be interpreted as a relaxation of the original optimization problem (4.30). Note that for $\underline{I}_A$, the matrices $\boldsymbol{K}$ and $\boldsymbol{P}$ are evaluated at $\overline{\alpha}$. Since these matrices depend on the parameters of the LQG cost function (see Appendix A6.3) it can be seen that this contribution to the underestimate of I is due to the replacement of the cost function, i.e., the weighted sum of the transmit powers, by a function which is smaller than or equal to the original cost for each given state sequence $(\boldsymbol{x}_k : k \in \mathbb{N}_0)$ and control sequence $(\boldsymbol{u}_k : k \in \mathbb{N}_0)$. The covariance matrix of the estimation error is evaluated at $\underline{\alpha}$ for the determination of $\underline{I}_A$. With Equation (4.33) it can be seen that this corresponds to a relaxation of problem (4.30) by replacing the scaled covariance matrix $\alpha\boldsymbol{C}_{\boldsymbol{q}}$ of the noise sequence in the observation channel with the smaller one $\underline{\alpha}\boldsymbol{C}_{\boldsymbol{q}}$ which leads to a smaller state estimation error and thus a smaller value of the optimum of (4.30). With this interpretation, the determination of the lower bound $\underline{I}_A$ fits into the framework discussed in [103] where the optimization problem under consideration exhibits a convex structure for some of the optimization variables and a monotonic structure for the other ones. In our case, the problem is a convex LQG controller optimization for fixed values of g and α, and monotonic w.r.t. those scalar variables for a fixed controller.

Note that in order to determine the global minimum of I for $\alpha > 0$, the lower bound given in Equation (4.37) needs to be consistent, which means that if the interval $A = [\underline{\alpha}, \overline{\alpha}]$ collapses to a single point α, i.e., for $\underline{\alpha} \to \alpha$ and $\overline{\alpha} \to \alpha$, it is required that $\underline{I}_A \to I(\alpha)$. This is the case if the solutions $\boldsymbol{K}$ and $\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{z}}}$ of the corresponding DAREs are continuous w.r.t. the optimization variable α. Concerning the continuity of such solutions w.r.t. to the parameters of the dynamical system and the LQG cost function, we refer to [104]. The continuity of the stabilizing solution of a DARE is shown under the assumption of left invertibility of the corresponding dynamical system, which means that its input is uniquely determined by its output (see, e. g., [105]). Keeping in mind the LQG cost function and the optimal estimation of the system state, the parameters of the DAREs for the determination of $\boldsymbol{K}$ and $\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}$ correspond, in the context of [104], to left invertible systems for $\rho \neq 1$ and if $\boldsymbol{C}_{\boldsymbol{q}}$ has full rank. Using the generic formulation of the DARE in Appendix A3, this means that the matrix $\boldsymbol{R}$ has to be invertible.

Before the branch and bound algorithm is applied to the problem (4.35), we have to deal with the fact that the optimization variable α is constrained to be larger than zero, implying that it can become arbitrarily large. The set $A =]0, \infty[$ is not suited for a branching procedure, which is the reason why we map it to the interval $B =]0, 1[$ by

$$\beta = \frac{\alpha}{1+\alpha}, \qquad \alpha > 0, \tag{4.39}$$

or equivalently $\alpha = (1-\beta)^{-1}\beta$ with $\beta \in]0, 1[$. The branching is then performed w.r.t. β. Note that we use the initial set $B = [\underline{\beta}, \overline{\beta}] = [0, 1]$ for the search of the global optimum of I. The inclusion of the boundary points is not a problem for the determination of the lower bound $\underline{I}_A$ (cf. Equation 4.37) since the matrices $\boldsymbol{K}$ and $\boldsymbol{P}$ depend on $\overline{\alpha}^{-1} = \overline{\beta}^{-1}(1-\overline{\beta})$, which is zero at $\overline{\beta} = 1$, whereas the estimation error covariance matrix of the system state depends on $\underline{\alpha} = (1-\underline{\beta})^{-1}\underline{\beta}$, which is also zero for $\underline{\beta} = 0$. Consequently, the lower bound for I can be evaluated for $B = [0, 1]$ and any subset of it. For the upper bound, one can either chose a different value than $\underline{\alpha}$ or $\overline{\alpha}$ if they correspond to $\underline{\beta} = 0$ or $\overline{\beta} = 1$, respectively, or the trivial upper bound $\overline{I}_A = \infty$ is used in such a case. Finally, we are in the position to apply Algorithm 4.1 to the optimization problem (4.35).

Algorithm 4.1 Branch and bound approach for weighted sum power minimization (cf. [102, 103])

Select a desired relative accuracy $\varepsilon > 0$
Use as initial partition $\mathbb{S}_1 = \mathbb{P}_1 = B = [0, 1]$
$k = 1$
$\underline{I}^* = 0$ and $\overline{I}^* = \infty$
while $1 - \frac{\underline{I}^*}{\overline{I}^*} \geq \varepsilon$ **do**
  Compute the lower bound $\underline{I}_A$ for each $A \in \mathbb{P}_k$
  Compute the upper bound $\overline{I}_A$ for each $A \in \mathbb{P}_k$
  Determine the smallest upper bound $\overline{I}^* = \min_{A \in \mathbb{S}_k} \overline{I}_A$
  Remove every $A \in \mathbb{S}_k$ with $\underline{I}_A \geq \overline{I}^*$ from $\mathbb{S}_k$ and let the set of remaining members of $\mathbb{S}_k$ be $\mathbb{R}_k$
  Determine $\underline{I}^* = \min_{A \in \mathbb{R}_k} \underline{I}_A$ as well as $B = \operatorname{argmin}_{A \in \mathbb{R}_k} \underline{I}_A$ (if more than one minimizer is present, choose one randomly)
  Determine the partition $\mathbb{P}_{k+1} = \left\{ \left[\underline{\beta}, \frac{1}{2}\left(\underline{\beta} + \overline{\beta}\right)\right], \left[\frac{1}{2}\left(\underline{\beta} + \overline{\beta}\right), \overline{\beta}, \right] \right\}$ of $B = [\underline{\beta}, \overline{\beta}]$
  $\mathbb{S}_{k+1} = (\mathbb{R}_k \setminus B) \cup \mathbb{P}_{k+1}$
  $k \leftarrow k + 1$
end while

For the above algorithm, the sets $\mathbb{S}_k$, $\mathbb{R}_k$ and $\mathbb{P}_k$, $k \in \mathbb{N}$, are introduced. In the k-th iteration, $\mathbb{R}_k$ is the set of remaining intervals after removing all subsets of $[0, 1]$ with a corresponding lower bound that is larger than the best upper bound obtained so far, meaning that they can not contain the optimizer of problem (4.35). $\mathbb{S}_k$ is determined by replacing one of its elements by a subdivision of the respective interval. Finally, $\mathbb{P}_k$ contains the intervals which are the result of a subdivision of the set B which corresponds to the smallest lower bound obtained so far.

The rule for the subdivision of a given set $B = [\underline{\beta}, \overline{\beta}]$ is a simple bisection (cf. [102]), which can be seen in line 11 of Algorithm 4.1. This subdivision rule has the property that an interval eventually collapses to a single point. For the subdivision in iteration k, the set B with the smallest corresponding lower bound is chosen (cf. line 10). The set $\mathbb{S}_{k+1}$ for the next iteration is then determined by removing B from the set of remaining intervals and replacing it with its subdivision in line 12. In the next iteration, the new lower and upper bounds for the sets which have been created by the subdivision are determined (see lines 6 and 7). Finally, the intervals which can be excluded from further steps are determined using the best upper bound $\overline{I}^*$ which might have changed due to the subdivision of the interval B in the last iteration (see line 8 and 9). The accuracy of the result after each iteration is measured by the gap between the smallest upper bound and the smallest lower bound (which is a lower bound for the global optimum), relative to the best value of I obtained so far, which given by the smallest upper bound $\overline{I}^*$.

Example 4.3.3 Using the parameters given in Example 3.1.1, Algorithm 4.1 (with a relative accuracy of $\varepsilon = 10^{-3}$) is applied to the minimization of the function I (cf. Equation 4.34) which is depicted in Figure 4.7 for $\rho = \frac{2}{3}$. For the determination of Pareto optimal values of the powers P_1 and P_2, 1000 values of $\rho = (1+\theta)^{-1}$ are considered, where θ is sampled logarithmically in the interval $[10^{-10}, 10^{10}]$. In Figure 4.8 the resulting values of the SNRs $\varphi_1 = \operatorname{tr}\left[\boldsymbol{C_q}\right]^{-1} P_1$ and $\varphi_2 = \operatorname{tr}\left[\boldsymbol{C_n}\right]^{-1} P_2$ are shown as a solid line. Since all pairs of SNRs below this line are not feasible, the shaded area shows the region of SNRs which correspond to feasible transmit powers.

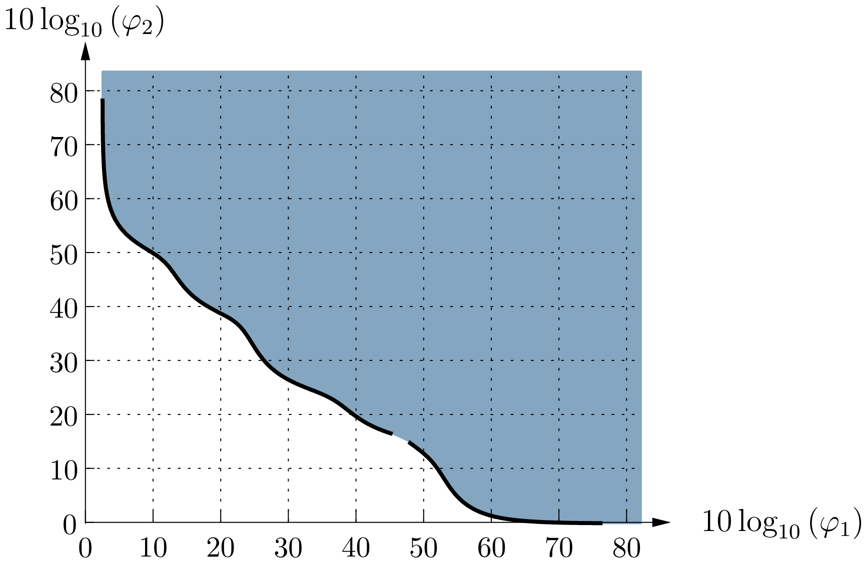


Figure 4.8: Outer bound (solid line) of the region of feasible SNRs (shaded area) for the power constrained LQG problem with optimal selection of the parameter α.

The region of feasible transmit powers or SNRs, respectively, for the power constrained LQG control problem with no transmit and receive scaling which is shown in Figure 3.7 lies inside the shaded area in Figure 4.8. This is not unexpected because the scenario discussed in Chapter 3 is a special case of the scenario considered here where the transmit and receive scaling factors are fixed to $t = g = 1$. If the transceiver is not optimized, the resulting optimal value of the optimization problem (3.33) is larger than the corresponding optimum of problem (4.30) in general.

Note that the solid line in Figure 4.8 is not connected but shows a gap. This is no inaccuracy but relates to a discontinuity of the value of α which minimizes the cost function I given by Equation (4.34). Let $\alpha^* = (1 - \beta^*)^{-1}\beta^* = \operatorname{argmin}_{\alpha>0} I(\alpha)$ be the optimizing value which has been obtained by the application of Algorithm 4.1. In Figure 4.9, α^* is depicted as a function of $\theta = \rho^{-1}(1 - \rho)$, where θ has been introduced above. It can be seen that for $\theta \approx 3000$, the

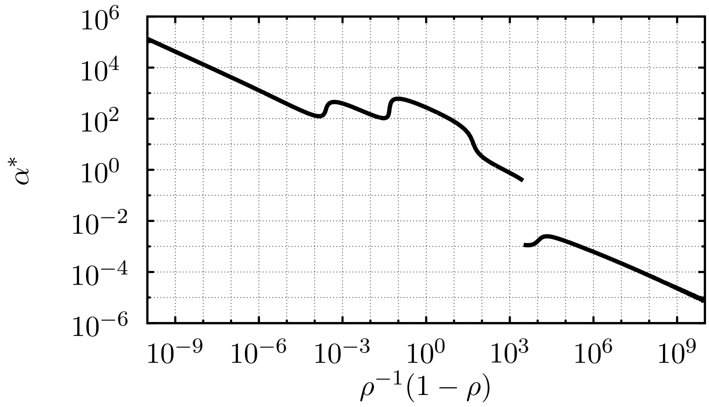


Figure 4.9: Optimal value α^* of the scaling parameter α over $\theta = \rho^{-1}(1 - \rho)$, where ρ is the weighting factor for the minimization of the weighted sum of the transmit powers (cf. optimization problem 4.30).

optimal value of α jumps over more than two orders of magnitude. This jump leads to the gap in the set of Pareto optimal transmit powers which can be observed in Figure 4.8 and is the result of the non-convexity of the optimization problem (4.30). In this case, the approach of minimizing the weighted sum of the transmit powers P_1 and P_2 can still be used to determine Pareto optimal values, but it is generally not possible to determine all of them if the set of feasible transmit powers is not convex (see, e. g., [94, Section 4.7.4]). Besides the non-convexity, a more precise explanation for the discontinuity of the optimal value α^* which is shown in Figure 4.9 is possible. Recall Example 4.3.2 where we have seen that the cost function I exhibits one global minimum and two additional local minima for a specific value of the weighting factor ρ. This configuration, i. e., location and number of local minima, varies for changing values of ρ, which becomes important for $\theta \approx 3000$ or $\rho \approx 3.33 \cdot 10^{-4}$, respectively. Figure 4.10 shows the function I near the global optimum for two different values of ρ which are close to the value mentioned above. Choosing $\rho \approx 3.21 \cdot 10^{-4}$ results in the minimal value which is slightly less than 60, but the nearest local minimum is only a little bit larger. For increasing values of ρ, the values of the global and the local minimum get closer and finally switch their position, i. e., the global optimum becomes a local one and vice versa. The resulting cost function for $\rho \approx 3.66 \cdot 10^{-4}$ is shown by the solid line in Figure 4.10. Thus, the optimizer of I exhibits a jump, which has been observed in Figure 4.9.

Example 4.3.3 demonstrates that the set of feasible transmit powers for the joint optimization of controller and scalar transceiver is not convex, which has been expected due to the properties of the original optimization problem (4.10). Consequently, the approach of the weighted minimization of the transmit powers in the observation and the control channel to determine Pareto optimal values only provides an outer bound for the set of feasible transmit powers. This bound is tight for the solution of problem (4.10) with a specific choice of the weighting factor ρ, i. e., the corresponding transmit powers can be approached arbitrarily close for $t \to \infty$ and $g \to \infty$ with a fixed ratio of both values. For a gap in the set of Pareto optimal transmit powers like in Figure 4.8, the bound is

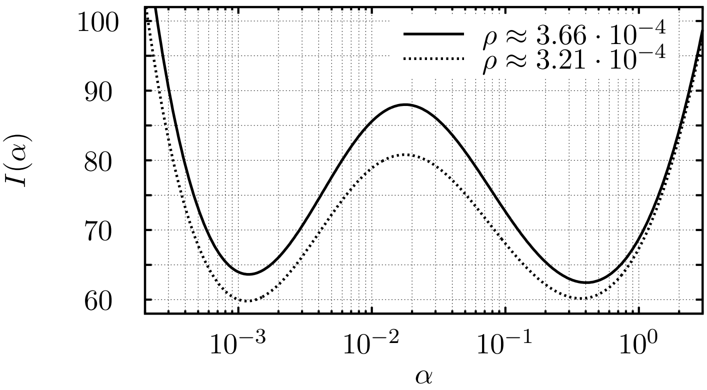


Figure 4.10: Weighted sum power for different values of ρ. For $\rho \approx 3.33 \cdot 10^{-4}$, the optimal value of the scaling parameter α jumps from approximately 0.4 to approximately $1.2 \cdot 10^{-3}$.

generally not tight, i. e., transmit powers which are obtained by a linear interpolation between the boundary points of the gap are not feasible.

4.3.3.2 Constrained Minimization of Transmit Power

Due to the non-convexity which has been observed in the last section, the approach of minimizing the weighted sum of the transmit powers of the observation and the control channel is not suitable for the determination of the whole set of Pareto optimal transmit powers. This set is of special interest because it provides the bound for feasible power constraints for the optimization problem (4.10). In order to determine Pareto optimal values of the transmit powers also for the non-convex part of the set of feasible transmit powers, an approach is proposed which is slightly more involved than the unconstrained minimization of the weighted sum of transmit powers presented in the preceding section. The idea is to minimize only one of the transmit powers, i. e., P_1 for the observation channel or P_2 for the control channel, while taking into account a constraint for the other one. Loosely speaking, we try to make one transmit power as small as possible while keeping the other one below a given value. Using the formulations introduced earlier, the resulting optimization problem reads as

$$\underset{t,g,\mu_0,\mu_1,\mu_2\dots}{\text{minimize}} \lim_{N\to\infty} \frac{1}{N} \operatorname{E}\left[t^{-1}\left(\sum_{n=0}^{N-1} \boldsymbol{x}_n^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{x}_n + \boldsymbol{v}_n^{\mathrm{T}}\boldsymbol{v}_n\right)\right] \tag{4.40}$$

$$\begin{aligned}
\text{subject to}\quad \boldsymbol{x}_{k+1} &= \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{B}g^{\frac{1}{2}}(\boldsymbol{u}_k + \boldsymbol{n}_k) + \boldsymbol{w}_k, && k \in \mathbb{N}_0,\\
\boldsymbol{y}_k &= \boldsymbol{C}\boldsymbol{x}_k + \boldsymbol{v}_k, && k \in \mathbb{N}_0,\\
\boldsymbol{u}_k &= \mu_k(\mathcal{I}_k), && k \in \mathbb{N}_0,\\
\mathcal{I}_k &= \begin{cases} \{(t^{-\frac{1}{2}}\boldsymbol{y}_0 + \boldsymbol{q}_0)\}, & k = 0,\\ \{(t^{-\frac{1}{2}}\boldsymbol{y}_0 + \boldsymbol{q}_0), \dots, (t^{-\frac{1}{2}}\boldsymbol{y}_k + \boldsymbol{q}_k), \boldsymbol{u}_0, \dots, \boldsymbol{u}_{k-1}\}, & k \in \mathbb{N},\end{cases}\\
t &> 0,\ g > 0,\\
\lim_{N\to\infty}\frac{1}{N}\operatorname{E}&\left[\sum_{n=0}^{N-1}\boldsymbol{u}_n^{\mathrm{T}}\boldsymbol{u}_n\right] \le P_{\mathrm{Tx},2},
\end{aligned}$$

where the transmit power P_1 of the observation channel (see Equation 4.8) is minimized subject to the controller, the transmitter t and the receiver g while taking into account that the transmit power P_2 of the control channel (see Equation 4.9) must not exceed the available power $P_{\text{Tx},2}$. For the solution of problem (4.40), we reformulate the system state and observation equation equivalently analogous to Equation (4.27). Note that the reformulation introduced there could also be used at this point, but the expressions derived in the following result in a convenient structure of the problem which has already been observed in Section 4.2.2. Additionally, the problem setting becomes comparable to a very similar approach presented in [46]. This point will be discussed later.

Analogously to Equation (4.27), define the scaled system state $\boldsymbol{z}_k = t^{-\frac{1}{2}}\boldsymbol{x}_k$, $k \in \mathbb{N}_0$, which is not problematic since $t \neq 0$ due to the constraints of problem (4.40). The resulting state and observation equation of the dynamical system to be controlled thus become

$$\begin{aligned} \boldsymbol{z}_{k+1} &= \boldsymbol{A}\boldsymbol{z}_k + \alpha^{-\frac{1}{2}}\boldsymbol{B}\left(\boldsymbol{u}_k + \boldsymbol{n}_k\right) + t^{-\frac{1}{2}}\boldsymbol{w}_k, \\ \boldsymbol{\eta}_k &= \boldsymbol{C}\boldsymbol{z}_k + t^{-\frac{1}{2}}\boldsymbol{v}_k, \end{aligned} \tag{4.41}$$

where $\alpha = \frac{t}{g}$ as in Equation (4.28) and $\boldsymbol{\eta}_k = t^{-\frac{1}{2}}\boldsymbol{y}_k$, $k \in \mathbb{N}_0$. The reformulation introduced above corresponds to a dynamical system which is driven by process noise with covariance matrix $t^{-1}\boldsymbol{C}_{\boldsymbol{w}}$, exhibits observation noise with covariance matrix $t^{-1}\boldsymbol{C}_{\boldsymbol{v}}$ and scales its input signal by $\alpha^{-\frac{1}{2}}$. Figure 4.11 shows the respective closed loop control system.

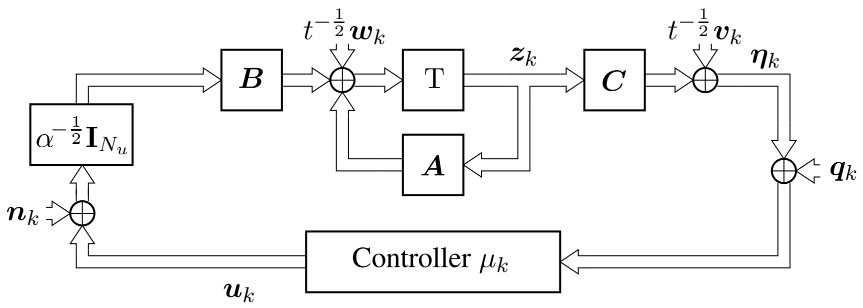


Figure 4.11: Model of the control loop which is closed over two channels with additive noise $\boldsymbol{q}_k$ and $\boldsymbol{n}_k$. The scaling factor $t^{-\frac{1}{2}}$ is shifted from the output to the input of the system, resulting in the receive scaling factor $\alpha^{-\frac{1}{2}} = \left(\frac{t}{g}\right)^{-\frac{1}{2}}$.

With Equation (4.41) the transmit power P_1 of the observation channel reads as

$$P_1 = \lim_{N\to\infty} \frac{1}{N}\,\mathrm{E}\left[\sum_{n=0}^{N-1} \boldsymbol{z}_n^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{z}_n + t^{-1}\boldsymbol{v}_n^{\mathrm{T}}\boldsymbol{v}_n\right]. \tag{4.42}$$

For reasons which will become clear soon, the inequality constraint is multiplied with α^{-1}, which is a valid step due to the constraints that $t > 0$ and $g > 0$, i. e., $\alpha > 0$. Thus, we require that

$$\lim_{N\to\infty} \frac{1}{N}\,\mathrm{E}\left[\sum_{n=0}^{N-1} \alpha^{-1}\boldsymbol{u}_n^{\mathrm{T}}\boldsymbol{u}_n\right] \leq \alpha^{-1}P_{\text{Tx},2}, \tag{4.43}$$

with $\alpha > 0$.

Putting together all previous steps, the optimization problem (4.40) becomes

$$\begin{aligned}
&\underset{\alpha,t,\mu_0,\mu_1,\mu_2\dots}{\text{minimize}} \ \lim_{N\to\infty}\frac{1}{N}\,\mathrm{E}\left[\sum_{n=0}^{N-1}\boldsymbol{z}_n^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{z}_n+t^{-1}\boldsymbol{v}_n^{\mathrm{T}}\boldsymbol{v}_n\right] && (4.44)\\
&\text{subject to} \quad \boldsymbol{z}_{k+1} = \boldsymbol{A}\boldsymbol{z}_k+\alpha^{-\frac{1}{2}}\boldsymbol{B}(\boldsymbol{u}_k+\boldsymbol{n}_k)+t^{-\frac{1}{2}}\boldsymbol{w}_k, && k\in\mathbb{N}_0,\\
&\qquad\qquad \boldsymbol{\eta}_k = \boldsymbol{C}\boldsymbol{z}_k+t^{-\frac{1}{2}}\boldsymbol{v}_k, && k\in\mathbb{N}_0,\\
&\qquad\qquad \boldsymbol{u}_k = \mu_k(\mathcal{I}_k), && k\in\mathbb{N}_0,\\
&\qquad\qquad \mathcal{I}_k = \begin{cases}\{(\boldsymbol{\eta}_0+\boldsymbol{q}_0)\}, & k=0,\\ \{(\boldsymbol{\eta}_0+\boldsymbol{q}_0),\dots,(\boldsymbol{\eta}_k+\boldsymbol{q}_k),\boldsymbol{u}_0,\dots,\boldsymbol{u}_{k-1}\}, & k\in\mathbb{N},\end{cases}\\
&\qquad\qquad \alpha > 0,\ t>0,\\
&\qquad\qquad \lim_{N\to\infty}\frac{1}{N}\,\mathrm{E}\left[\sum_{n=0}^{N-1}\alpha^{-1}\boldsymbol{u}_n^{\mathrm{T}}\boldsymbol{u}_n\right]\le\alpha^{-1}P_{\mathrm{Tx},2}.
\end{aligned}$$

For fixed values of α and t, this optimization is a power constrained LQG control problem which can be solved using the approach from Chapter 3 or the with the help of Lagrangian duality. At this point the latter approach will be used since is provides insights in the joint optimization of the controller and the transceiver as well as the corresponding solution. Taking into account the power constraint by the multiplier $\lambda\ge 0$, the Lagrangian associated with problem (4.44) is given by

$$L=\lim_{N\to\infty}\frac{1}{N}\,\mathrm{E}\left[\sum_{n=0}^{N-1}\boldsymbol{z}_n^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{z}_n+\lambda\alpha^{-1}\boldsymbol{u}_n^{\mathrm{T}}\boldsymbol{u}_n+t^{-1}\boldsymbol{v}_n^{\mathrm{T}}\boldsymbol{v}_n\right]-\lambda\alpha^{-1}P_{\mathrm{Tx},2}, \tag{4.45}$$

and its minimum w.r.t. μ_k, $k\in\mathbb{N}_0$, and the remaining constraints of problem (4.44) reads as

$$L^*(\alpha,t,\lambda)=\mathrm{tr}\left[\boldsymbol{K}\left(t^{-1}\boldsymbol{C}_{\boldsymbol{w}}+\alpha^{-1}\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\right)\right]+\mathrm{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}\right]+\mathrm{tr}\left[t^{-1}\boldsymbol{C}_{\boldsymbol{v}}\right]-\lambda\alpha^{-1}P_{\mathrm{Tx},2}, \tag{4.46}$$

where

$$\begin{aligned}\boldsymbol{K}&=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A}-\alpha^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B}\left(\alpha^{-1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B}+\lambda\alpha^{-1}\mathbf{I}_{N_u}\right)^{-1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A}+\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}\\
&=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A}-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B}\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B}+\lambda\mathbf{I}_{N_u}\right)^{-1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A}+\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}\end{aligned} \tag{4.47}$$

and

$$\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}=\boldsymbol{A}\Big(\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}-\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}+t^{-1}\boldsymbol{C}_{\boldsymbol{v}}+\boldsymbol{C}_{\boldsymbol{q}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}\Big)\boldsymbol{A}^{\mathrm{T}}+t^{-1}\boldsymbol{C}_{\boldsymbol{w}}+\alpha^{-1}\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}. \tag{4.48}$$

Finally, the matrices $\boldsymbol{P}$ and $\boldsymbol{C}_{\tilde{\boldsymbol{z}}}$ in Equation (4.46) are given by $\boldsymbol{P}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A}-\boldsymbol{K}+\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}$ and $\boldsymbol{C}_{\tilde{\boldsymbol{z}}}=\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}-\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}+t^{-1}\boldsymbol{C}_{\boldsymbol{v}}+\boldsymbol{C}_{\boldsymbol{q}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}$, respectively. We make an observation analogous to Section 4.3.2 and the optimization problem (4.30), i. e., the infimum of problem (4.44) is achieved for $t\to\infty$. This is intuitively clear since this limit corresponds to a dynamical system without process and observation noise, i. e., with $\boldsymbol{C}_{\boldsymbol{w}}=\mathbf{0}_{N_x\times N_x}$ and $\boldsymbol{C}_{\boldsymbol{v}}=\mathbf{0}_{N_y\times N_y}$. For a formal argumentation, assume that problem (4.44) is feasible and that strong duality holds. Then, the solution for fixed values of α and t is found by maximizing L^* w.r.t. $\lambda\ge 0$. Now, for any given values of λ and α, $L^*(\alpha,t,\lambda)$ is a monotonically increasing function of t^{-1}. This is obvious for the summands $\mathrm{tr}\left[\boldsymbol{K}t^{-1}\boldsymbol{C}_{\boldsymbol{w}}\right]$ and $\mathrm{tr}\left[t^{-1}\boldsymbol{C}_{\boldsymbol{v}}\right]$ (cf. Equation 4.46) and can be shown for $\mathrm{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}\right]$ using the results from Appendix A4, i. e., the derivative of $\operatorname{tr}[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}]$ w.r.t. t^{-1} is non-negative. Since the monotonicity w.r.t. t^{-1} holds for each λ, also the maximum of $L^*(\alpha, t, \lambda)$ w.r.t. $\lambda \geq 0$ increases with t^{-1}. Consequently, the infimum of problem (4.44) for a fixed value of α is achieved for $t \to \infty$ due to the assumption of strong duality, i. e., the optimal value of (4.44) is found by maximizing the dual function w.r.t. $\lambda \geq 0$. Thus, in the following we consider

$$\begin{aligned} G(\alpha, \lambda) &= \inf_{t>0} L^*(\alpha, t, \lambda) \\ &= \operatorname{tr}\left[\boldsymbol{K}\alpha^{-1}\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\right] + \operatorname{tr}[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}] - \lambda\alpha^{-1}P_{\mathrm{Tx},2}, \end{aligned} \tag{4.49}$$

where $\boldsymbol{K}$ (cf. Equation 4.47) and $\boldsymbol{P}$ are given above and the covariance matrix of the state estimation error reads as $\boldsymbol{C}_{\tilde{\boldsymbol{z}}} = \boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{q}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}$ with

$$\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}} = \boldsymbol{A}\left(\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{q}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}^{\mathrm{P}}\right)\boldsymbol{A}^{\mathrm{T}} + \alpha^{-1}\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}. \tag{4.50}$$

Comparing the function G from Equation (4.49) with L^* shown in Equation (4.13), it can be seen that the results of Section 4.2.3 are also applicable at this point, i. e., G is concave in α^{-1} for a given value of $\lambda \geq 0$. It follows that problem (4.44) can not be solved by minimizing G w.r.t. $\alpha > 0$ and maximizing the resulting dual function w.r.t. $\lambda \geq 0$.[6] We will thus solve the optimization problem (4.44) by keeping α fixed, which reduces the problem to a convex optimization and provides the function

$$I(\alpha) = \sup_{\lambda \geq 0} G(\alpha, \lambda). \tag{4.51}$$

Then, the value of α which minimizes I is determined using the monotonic optimization framework introduced above. The following example shows why this step is necessary.

Example 4.3.4 With the parameters given in Example 3.1.1, the optimization problem (4.44) is solved for fixed values of α, where $\alpha \in [2 \cdot 10^{-4}, \alpha_{\text{max}}]$. The value $\alpha_{\text{max}} \approx 967$ is the largest value such that the minimization of P_1 with the constraint $P_2 \leq P_{\mathrm{Tx},2} \approx 1288$ is feasible, which is equivalent to the SNR constraint $10\log_{10}\left(\operatorname{tr}[\boldsymbol{C}_{\boldsymbol{n}}]^{-1} P_2\right) \leq 31.1$ from Example 4.3.1. For the determination of α_{max}, the feasibility test introduced in Section 4.3.1 can be used to perform a bisection on α.[7]

Figure 4.12 shows the resulting transmit powers P_1, i. e., the cost function of problem (4.44), and P_2, i. e., the power which is constrained by $P_{\mathrm{Tx},2}$. It can be seen that the transmit power constraint is always satisfied but not active for small values of α. The global minimum of $P_1 \approx 559$ is achieved for $\alpha \approx 492$, but there are additionally two local minima for smaller values of α.

Example 4.3.4 demonstrates that due to the lack of convexity or quasi-convexity, it is not straightforward to determine the solution of problem (4.44) when α is also considered as an optimization variable. It is of course possible to sample values of α and to chose the one which leads to the smallest value of P_1, but this implies the difficulties of, e. g., the selection of the sampling range

[6] This means that strong duality holds for a fixed value of α but not if this scalar is also an optimization variable.

[7] A faster approach is a fixed point iteration for the determination of the root of L^*_{feas} (cf. Equation 4.23) w.r.t. α^{-1}, where this function has to be adapted to the problem at hand. The concavity in α^{-1} can be used to show the convergence of the iteration.

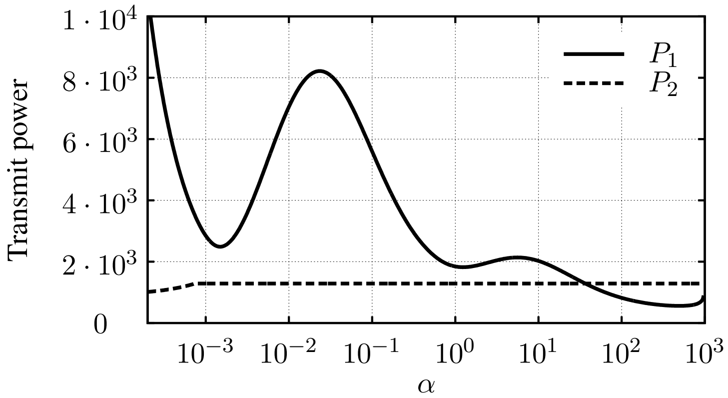


Figure 4.12: Minimal value of $P_1 = I(\alpha)$ for a given value of α with the constraint $P_2 \leq P_{\text{Tx},2}$, where $P_{\text{Tx},2} \approx 1288$. For small values of α the constraint is not active, for large values the problem is not feasible.

and sampling interval, and does not provide information about the optimality of the best sampled solution. Thus, we resort to the monotonic optimization framework introduced in Section 4.3.3.

Due to the assumption that strong duality holds, $I(\alpha)$ defined in Equation (4.51) provides the optimal value of problem (4.44) for a fixed value of α. In order to apply the branch and bound approach, which is the basis of Algorithm 4.1, to the minimization of $I(\alpha)$ w.r.t. $\alpha > 0$, it is necessary to derive a lower bound for the optimal value of (4.44) when α is restricted to lie in an interval $A \subset \mathbb{R}_+$. In contrast to Section 4.3.3.1, we are dealing with a constrained optimization problem and thus use the function G given by Equation (4.49) for the derivation. Let $\alpha \in A = [\underline{\alpha}, \overline{\alpha}]$ with $0 < \underline{\alpha} \leq \overline{\alpha}$. Then, it holds for every $\lambda \geq 0$ that

$$\begin{aligned} G(\alpha,\lambda) &= \operatorname{tr}\left[\boldsymbol{K}\alpha^{-1}\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\right] + \operatorname{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}(\alpha^{-1})\right] - \lambda\alpha^{-1}P_{\text{Tx},2} \\ &\geq \operatorname{tr}\left[\boldsymbol{K}\overline{\alpha}^{-1}\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\right] + \operatorname{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}(\overline{\alpha}^{-1})\right] - \lambda\alpha^{-1}P_{\text{Tx},2} \\ &\geq \operatorname{tr}\left[\boldsymbol{K}\overline{\alpha}^{-1}\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\right] + \operatorname{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{z}}}(\overline{\alpha}^{-1})\right] - \lambda\underline{\alpha}^{-1}P_{\text{Tx},2} = \underline{G}_A(\lambda). \end{aligned} \tag{4.52}$$

Note that the matrices $\boldsymbol{K}$ and $\boldsymbol{P}$ do not depend on α (cf. Equation 4.47) and that the error covariance matrix $\boldsymbol{C}_{\tilde{\boldsymbol{z}}}$ is explicitly denoted as a function of α^{-1} (cf. Equation 4.50). The monotonicity of the first and the last summand of Equation (4.52) which has been used for the lower bound is obvious, whereas the monotonicity of the second summand has been discussed earlier and can be shown using the results of Appendix A4. It can be seen that for each value of the Lagrange multiplier $\lambda \geq 0$, $\underline{G}_A(\lambda)$ is a lower bound for $G(\alpha,\lambda)$ which holds for all $\alpha \in A$. Consequently, $\underline{G}_A$ can now be used to lower bound the optimal value of problem (4.44) by $\underline{I}_A$ when α is taken from the interval A:

$$I(\alpha) = \sup_{\lambda\geq 0} G(\alpha,\lambda) \geq G(\alpha,\underline{\lambda}) \geq \underline{G}_A(\underline{\lambda}) = \sup_{\lambda\geq 0} \underline{G}_A(\lambda) = \underline{I}_A, \qquad \alpha \in A. \tag{4.53}$$

Here, $\underline{\lambda}$ denotes the maximizer of $\underline{G}_A(\lambda)$, where it is assumed that $\underline{\lambda}$ exists and is finite. Note that even the relaxed problem for the determination of the lower bound $\underline{I}_A$ may not be feasible, which means that $\underline{I}_A = I(\alpha) = \infty$. In this case, the intermediate steps in Equation (4.53) have to be ignored. In order to obtain an upper bound $\overline{I}_A$, the optimization problem (4.44) is solved for some value of $\alpha \in A = [\underline{\alpha}, \overline{\alpha}]$, e. g.,

$$\overline{I}_A = I(\overline{\alpha}). \tag{4.54}$$

Again, if the corresponding optimization problem is not feasible, we have $\overline{I}_A = \infty$.

Remark: The lower bound $\underline{G}_A$ (cf. Equation 4.52) which is a function of $\lambda \geq 0$ can be interpreted as the dual function which corresponds to a relaxation of the optimization problem (4.44) when α is constrained to lie in the interval A. This relaxation has two elements. The first one is the scaling of the covariance matrix of the driving noise process from $\alpha^{-1}\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}$ down to $\overline{\alpha}^{-1}\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}$. The second element is the relaxation of the transmit power constraint from $P_2 \leq P_{\mathrm{Tx},2}$ to the less restrictive requirement $P_2 \leq \underline{\alpha}^{-1}\overline{\alpha}P_{\mathrm{Tx},2}$. With the results of Appendix A9, it can be seen that the associated lower bound $\underline{I}_A$ for the optimal value $I(\alpha)$ which is derived from $\underline{G}_A$ (cf. Equation 4.53) is consistent, i. e., for $\underline{\alpha} \to \alpha$ and $\overline{\alpha} \to \alpha$, it holds that $\underline{I}_A \to I(\alpha)$. The required properties for the derivations in Appendix A9 are the monotonicity and continuity of the function $\underline{G}_A$ in Equation (4.52) w.r.t. $\underline{\alpha}$ and $\overline{\alpha}$, respectively. The monotonicity follows from, e. g., [100], whereas the continuity can be shown using the results from [104] under the assumption that $\boldsymbol{C}_{\boldsymbol{q}}$ in Equation (4.50) has full rank.

Basically, we are now in the position to apply the branch and bound Algorithm 4.1 using the mapping of the parameter $\alpha \in A = \mathbb{R}_+$ to $\beta \in B =]0,1[$ with (cf. Equation 4.39)

$$\beta = \frac{\alpha}{1+\alpha}, \qquad \alpha > 0. \tag{4.55}$$

Like in Section 4.3.3.1, we include the boundary points of the interval B, i. e., we will use $B = [\underline{\beta}, \overline{\beta}] = [0,1]$. For the evaluation of $\underline{I}_A$ and $\overline{I}_A$, the larger boundary point $\overline{\beta} = 1$ is not problematic because $\underline{I}_A$ and $\overline{I}_A$ are actually evaluated at $\overline{\alpha}^{-1} = \overline{\beta}^{-1}(1-\overline{\beta})$, which is zero at $\overline{\beta} = 1$. For the smaller boundary point $\underline{\beta} = 0$, we use the convention that $0 \cdot \infty = 0$. In this case, the function $\underline{G}_A$ given by Equation (4.52) has a finite positive value for $\lambda = 0$ and is unbounded from below for $\lambda > 0$ when evaluated at the boundary point $\underline{\beta} = 0$. Thus, we obtain $\underline{I}_A = \underline{G}_A(0)$ if $\underline{\alpha} = 0$.[8] Finally, because we have to deal with a constrained optimization problem, it is necessary to perform a feasibility check of the problems for the determination of $\underline{I}_A$ and $\overline{I}_A$, e. g., using the approach discussed in Section 4.3.1. If the corresponding optimization problems are not feasible, we set $\underline{I}_A = \infty$ or $\overline{I}_A = \infty$, respectively. A second possibility is to determine the value $\alpha_{\max}$ which has been introduced in Example 4.3.4 and to initialize Algorithm 4.1 with the set $B = [0, \beta_{\max}]$, where $\beta_{\max} = (1+\alpha_{\max})^{-1}\alpha_{\max}$.

For the following example, Algorithm 4.1 is applied to the optimization problem (4.44), where the feasibility issue is taken into account by setting the upper and lower bounds to infinity if the corresponding optimization problem is not feasible. Of special interest is the determination of Pareto optimal transmit powers which can not be computed with the approach of Section 4.3.3.1, i. e., the gap in the solid line shown in Figure 4.8.

Example 4.3.5 As in Example 4.3.3, the system and noise parameters of Example 3.1.1 are used in the following. Algorithm 4.1 is applied to solve optimization problem (4.44) with a relative accuracy of $\varepsilon = 10^{-3}$. In order to fill the gap in the set of Pareto optimal transmit powers shown in Figure 4.8, 100 values of $P_{\mathrm{Tx},2}$ which determine the constraint for the transmit power of the control channel are selected between 28.2 and 46.8. Figure 4.13 shows the resulting Pareto optimal transmit powers as a solid line. The boundary points of the gap which has been observed in Figure 4.8 are denoted by dots, and the linear interpolation between these two points is indicated by the

[8]This is confirmed by the interpretation that the lower bound $\underline{I}_A$ corresponds to a relaxed optimization problem. For $\underline{\alpha} \to 0$, the constraint on P_2 is effectively removed which leads to an inactive power constraint with $\lambda = 0$.

dashed line. Note that Figure 4.13 depicts the SNRs $\varphi_1 = \mathrm{tr}\left[\boldsymbol{C}_{\boldsymbol{q}}\right]^{-1} P_1$ and $\varphi_2 = \mathrm{tr}\left[\boldsymbol{C}_{\boldsymbol{n}}\right]^{-1} P_2$ with linear scale. Thus, it can be seen that the set of feasible transmit powers is not convex.

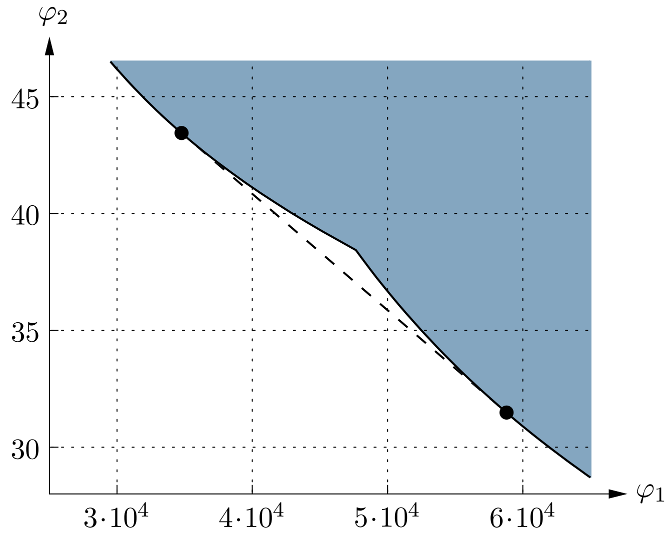


Figure 4.13: Pareto optimal transmit powers with optimal scaling ratio α (solid line). Since the optimization problem is not convex, the region of feasible transmit powers is not a convex set (dashed line).

Remark: The scenario which is depicted in Figure 4.11 and investigated in Example 4.3.4 fits into the framework discussed in [46]. For the situation considered there, the transmit power P_1 which is minimized with problem (4.44) has to be interpreted as a tracking error variance and the power constraint $P_2 \leq P_{\mathrm{Tx},2}$ corresponds to an SNR constraint. Although the limiting case for $t \to \infty$ is discussed, Example 4.3.4 demonstrates that the associated optimization problem may exhibit several local minima. Consequently, the minimization based on a line search which is proposed in [46] does not guarantee to find the global optimum since it requires at least quasi-convexity of the optimization problem.[9] The sampling approach in [58] is in principle capable to determine the optimum, but suffers from problems like the selection of the sampling range and the sampling interval. Additionally, it does not provide information about the optimality of the sampled values, which is available for the monotonic optimization approach.

4.3.4 Solution of the Joint Optimization of Controller and Transceiver

After the discussions in the preceding sections about the properties and the feasibility of the joint optimization of an LQG controller and scalar transceivers, we come back to the original goal, i. e., the solution of the optimization problem (4.10). To summarize the task, it is required to minimize a standard infinite horizon LQG cost function, which is characterized by the weighting matrices $\boldsymbol{Q}$, $\boldsymbol{R}$ and $\boldsymbol{S}$ (see Appendix A6), while satisfying the transmit power constraints of the observation and the control channel, given by the available transmit powers $P_{\mathrm{Tx},1}$ and $P_{\mathrm{Tx},2}$, respectively. The degrees of freedom which are added to the conventional power constrained LQG control problem (see Chapter 3) are the scaling factors $t^{-\frac{1}{2}}$ and $g^{\frac{1}{2}}$ at the output and the input, respectively, of the system to be controlled (cf. Figure 4.1).

[9] The proposed approach which requires quasi-convexity is the golden section method, see, e. g., [95, Section 8.1].

We start with the solution of the joint controller and transceiver optimization by restating the problem to be solved, i. e.,

$$
\begin{aligned}
\underset{t,g,\mu_0,\mu_1,\mu_2\ldots}{\text{minimize}} \quad & \lim_{N\to\infty}\frac{1}{N}\,\mathrm{E}\left[\boldsymbol{x}_N^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{x}_N+\sum_{n=0}^{N-1}\begin{bmatrix}\boldsymbol{x}_n\\ g^{\frac{1}{2}}\boldsymbol{u}_n\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{Q} & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{R}\end{bmatrix}\begin{bmatrix}\boldsymbol{x}_n\\ g^{\frac{1}{2}}\boldsymbol{u}_n\end{bmatrix}\right]+g\,\mathrm{tr}\left[\boldsymbol{R}\boldsymbol{C}_{\boldsymbol{n}}\right] \qquad (4.56)\\
\text{subject to}\quad & \boldsymbol{x}_{k+1} = \boldsymbol{A}\boldsymbol{x}_k+\boldsymbol{B}g^{\frac{1}{2}}(\boldsymbol{u}_k+\boldsymbol{n}_k)+\boldsymbol{w}_k, \qquad k\in\mathbb{N}_0,\\
& \boldsymbol{y}_k = \boldsymbol{C}\boldsymbol{x}_k+\boldsymbol{v}_k, \qquad k\in\mathbb{N}_0,\\
& \boldsymbol{u}_k = \mu_k(\mathcal{I}_k), \qquad k\in\mathbb{N}_0,\\
& \mathcal{I}_k = \begin{cases}\{(t^{-\frac{1}{2}}\boldsymbol{y}_0+\boldsymbol{q}_0)\}, & k=0,\\ \{(t^{-\frac{1}{2}}\boldsymbol{y}_0+\boldsymbol{q}_0),\ldots,(t^{-\frac{1}{2}}\boldsymbol{y}_k+\boldsymbol{q}_k),\boldsymbol{u}_0,\ldots,\boldsymbol{u}_{k-1}\}, & k\in\mathbb{N},\end{cases}\\
& t > 0,\ g>0,\\
& \lim_{N\to\infty}\frac{1}{N}\,\mathrm{E}\left[\left(\sum_{n=0}^{N-1}\boldsymbol{x}_n^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{x}_n+\boldsymbol{v}_n^{\mathrm{T}}\boldsymbol{v}_n\right)\right]\le tP_{\mathrm{Tx},1},\\
& \lim_{N\to\infty}\frac{1}{N}\,\mathrm{E}\left[g\sum_{n=0}^{N-1}\boldsymbol{u}_n^{\mathrm{T}}\boldsymbol{u}_n\right]\le gP_{\mathrm{Tx},2}.
\end{aligned}
$$

Note that the transmit power constraints (see Equations 4.8 and 4.9) are equivalently reformulated by multiplying them by $t>0$ and $g>0$, respectively. For given values of t and g, the power constrained optimization of the controller is a convex problem and we have chosen a solution method based on Lagrangian duality in the preceding sections, but other methods for convex optimization can also be used. Since the minimization of the Lagrangian in Equation (4.11) corresponds to an LQG control problem, the minimum is readily obtained and reads as (cf. problem 4.12 and Equation 4.13)

$$
\begin{aligned}
L^*(t,g,\lambda_1,\lambda_2) = \mathrm{tr}\left[\boldsymbol{K}\left(\boldsymbol{C}_{\boldsymbol{w}}+g\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\right)\right]+\mathrm{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\right]+\lambda_1\,\mathrm{tr}\left[\boldsymbol{C}_{\boldsymbol{v}}\right]+g\,\mathrm{tr}\left[\boldsymbol{R}\boldsymbol{C}_{\boldsymbol{n}}\right]&\\
-\lambda_1 tP_{\mathrm{Tx},1}-\lambda_2 gP_{\mathrm{Tx},2},&
\end{aligned}
\tag{4.57}
$$

where the matrices $\boldsymbol{K}$ and $\boldsymbol{P}$ are determined by the DARE shown in Equation (4.14) and where the covariance matrix of the state estimation error $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}$ depends on the solution $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}$ of the DARE from Equation (4.15). Assuming that strong duality holds, the solution of problem (4.56) for fixed values of t and g is found by maximizing the dual function L^* w.r.t. $\lambda_1\ge 0$ and $\lambda_2\ge 0$. Thus, the optimal value in this case is given by

$$
I(t,g)=\sup_{\substack{\lambda_1\ge 0\\ \lambda_2\ge 0}} L^*(t,g,\lambda_1,\lambda_2). \tag{4.58}
$$

This representation of the optimal value of problem (4.56) with fixed values of t and g is the basis for the remaining optimization w.r.t. these two variables.

Due to the properties of problem (4.56) which have been described in Section 4.2.3, we resort again to the monotonic optimization approach introduced in Section 4.3.3. More precisely, Algorithm 4.1 is applied but has to be slightly modified because the remaining minimization w.r.t. t and g is a two dimensional problem. Thus, the branch and bound procedure is performed on two dimensional sets, but the determination of upper and lower bounds for the optimal value of problem (4.56) when the pair of transmit and receive scaling factors is constrained to a set $A\subset\mathbb{R}_+^2$ is completed analogously to Section 4.3.3.2. Let this set be $A = [\underline{t}, \overline{t}] \times [\underline{g}, \overline{g}]$, i. e., the scaling factor t is constrained to lie in the interval $[\underline{t}, \overline{t}]$ with $0 < \underline{t} \leq \overline{t}$ and at the same time we require that g lies in the interval $[\underline{g}, \overline{g}]$ with $0 < \underline{g} \leq \overline{g}$. For all values of the transmit and receive scaling factors within this set, a lower bound for the dual function L^* is given by

$$\begin{aligned} L^*(t, g, \lambda_1, \lambda_2) &= \operatorname{tr}\left[\boldsymbol{K}\left(\boldsymbol{C_w}+g\boldsymbol{B}\boldsymbol{C_n}\boldsymbol{B}^{\mathrm{T}}\right)\right] + \operatorname{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}(t, g)\right] + \lambda_1 \operatorname{tr}\left[\boldsymbol{C_v}\right] + g \operatorname{tr}\left[\boldsymbol{R}\boldsymbol{C_n}\right] \\ &\qquad - \lambda_1 t P_{\mathrm{Tx},1} - \lambda_2 g P_{\mathrm{Tx},2} \\ &\geq \operatorname{tr}\left[\boldsymbol{K}\left(\boldsymbol{C_w}+\underline{g}\boldsymbol{B}\boldsymbol{C_n}\boldsymbol{B}^{\mathrm{T}}\right)\right] + \operatorname{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}(\underline{t}, \underline{g})\right] + \lambda_1 \operatorname{tr}\left[\boldsymbol{C_v}\right] + \underline{g} \operatorname{tr}\left[\boldsymbol{R}\boldsymbol{C_n}\right] \\ &\qquad - \lambda_1 \overline{t} P_{\mathrm{Tx},1} - \lambda_2 \overline{g} P_{\mathrm{Tx},2} \\ &= \underline{L}_A^*(\lambda_1, \lambda_2), \end{aligned} \tag{4.59}$$

where the covariance matrix $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}(t, g)$ is explicitly denoted as a function of t and g. In order to obtain the inequality above, all summands which depend linearly on the optimization variables are evaluated at the smallest possible value within the set A if they represent a positive contribution to L^* whereas the largest value is inserted for negative contributions. Since the error covariance matrix $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}$ is a monotonically increasing function of t and g, the respective summand $\operatorname{tr}[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}]$ is lower bounded using $\underline{t}$ and $\underline{g}$. At this point we repeat the steps from Equation (4.53) with the lower bound given by Equation (4.59) which holds for every pair $(t, g) \in A = [\underline{t}, \overline{t}] \times [\underline{g}, \overline{g}]$ and for each $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$. Thus, we obtain a lower bound $\underline{I}_A$ for the optimal value of the optimization problem (4.56), which is given by $I(t, g)$ (cf. Equation 4.58), with the restriction that $(t, g) \in A$:

$$I(t, g) = \sup_{\substack{\lambda_1 \geq 0 \\ \lambda_2 \geq 0}} L^*(t, g, \lambda_1, \lambda_2) \geq L^*(t, g, \underline{\lambda}_1, \underline{\lambda}_2) \geq \underline{L}_A^*(\underline{\lambda}_1, \underline{\lambda}_2) = \sup_{\substack{\lambda_1 \geq 0 \\ \lambda_2 \geq 0}} \underline{L}_A^*(\lambda_1, \lambda_2) = \underline{I}_A, \tag{4.60}$$

where it is assumed that the maximizers $\underline{\lambda}_1$ and $\underline{\lambda}_2$ exist and are finite.

Remark: As in Section 4.3.3.2, the lower bound given by Equation (4.59) can be interpreted as the dual function which corresponds to a relaxed version of optimization problem (4.56). First, the scalar g in the cost function is replaced by $\underline{g}$ which leads to a smaller cost for each given state sequence $(\boldsymbol{x}_k : k \in \mathbb{N}_0)$ and control sequence $(\boldsymbol{u}_k : k \in \mathbb{N}_0)$. Second, the covariance matrix of the noise in the control channel which enters the dynamical system as additional process noise is scaled down from $g\boldsymbol{C_n}$ to $\underline{g}\boldsymbol{C_n}$. Moreover, the covariance matrix of the noise in the observation channel which acts like additional observation noise is scaled from $t\boldsymbol{C_q}$ down to $\underline{t}\boldsymbol{C_q}$ (cf. Figure 4.2). Finally, the power constraints are relaxed by $\underline{t}P_1 \leq \overline{t}P_{\mathrm{Tx},1}$ and $\underline{g}P_2 \leq \overline{g}P_{\mathrm{Tx},2}$, respectively, which corresponds to an increase of the available transmit power by a factor of $\underline{t}^{-1}\overline{t} \geq 1$ and $\underline{g}^{-1}\overline{g} \geq 1$. Consequently, the optimal value of the relaxed optimization problem is an underestimate of the solution of problem (4.56). Note that with the monotonicity and continuity properties of the functions L^* and $\underline{L}_A^*$ (cf. Equation 4.59) which are summed up in Appendix A9, the results presented there show that the lower bound $\underline{I}_A$ in Equation (4.60) is consistent, i. e., for $\underline{t} \to t$, $\overline{t} \to t$, $\underline{g} \to g$ and $\overline{g} \to g$, we have $\underline{I}_A \to I(t, g)$.[10]

If the optimization problem for the determination of $\underline{I}_A$ is not feasible, we obtain $\underline{I}_A = \infty$. Since this value is still a lower bound for $I(t, g)$, the inequality in Equation (4.60) also holds for this case

[10]For a discussion of the monotonicity and continuity of the stabilizing solutions of the DAREs which are involved in Equation (4.59), we refer again to [100], Appendix A4 and [104]. The continuity holds in our case under the assumption that $\boldsymbol{R}$ in Equation (4.14) and $\boldsymbol{C_q}$ or $\boldsymbol{C_v}$ in Equation (4.15) have full rank.

but the intermediate steps have to be omitted. In order to check the feasibility for a given interval A, the approach described in Section 4.3.1 can be used, where the relaxations mentioned in the remark above have to be taken into account, i. e., the replacement of $g\boldsymbol{C_n}$ and $t\boldsymbol{C_q}$ by $\underline{g}\boldsymbol{C_n}$ and $\underline{t}\boldsymbol{C_q}$, respectively, in Equation (4.25), and of the (scaled) transmit power constraints $tP_1 \leq tP_{\mathrm{Tx},1}$ and $gP_2 \leq gP_{\mathrm{Tx},2}$ by $\underline{t}P_1 \leq \overline{t}P_{\mathrm{Tx},1}$ and $\underline{g}P_2 \leq \overline{g}P_{\mathrm{Tx},2}$, respectively, in Equation (4.22).

A further modification of the monotonic optimization Algorithm 4.1 concerns the subdivision rule for certain set A in order to obtain lower bounds for subsets of A. For the one dimensional case discussed in Section 4.3.3, an interval has been divided into two subsets by simple bisection, i. e., the set has been split in the middle. A similar approach is used for the two dimensional case with $A \subset \mathbb{R}_+^2$. According to [102], if a set A is selected during the branch and bound procedure for a subdivision, the longer side of A is divided at its midpoint which provides two subsets A_1 and A_2 of A which have equal size and where $A_1 \cup A_2 = A$.

Finally, in order to deal with the fact that the optimization variables t and g are elements of the unbounded set $\mathbb{R}_+^2$, the pair (t, g) is mapped as in Section 4.3.3 to the set $B =]0, 1[^2$ by

$$(\pi, \gamma) = \left(\frac{t}{1+t}, \frac{g}{1+g}\right), \qquad (t, g) \in \mathbb{R}_+^2. \tag{4.61}$$

For the application of the branch and bound procedure, the boundary of B is included, i. e., the monotonic optimization algorithm is initialized with $B = [0, 1]^2$. The lower boundary points $\underline{\pi} = 0$ and $\underline{\gamma} = 0$ which correspond to $\underline{t} = 0$ and $\underline{g} = 0$ can be inserted directly in Equation (4.59) for the computation of the lower bound $\underline{I}_A$ (cf. Equation 4.60). The application of the upper boundary points with $\overline{\pi} = 1$ and $\overline{\gamma} = 1$ corresponds to an optimization problem where either the power constraint for the observation channel or the control channel is removed, which leads to an inactive power constraint with $\lambda_1 = 0$ or $\lambda_2 = 0$, respectively.[11] Thus, these points can be included in the preceding discussion. Note that Example 4.3.1 (see Figure 4.5) demonstrated that it is not obvious if a certain pair (t, g) of transmit and receive scaling factors leads to a feasible optimization problem. Additionally, even if the lower bound $\underline{I}_A$ corresponds to a relaxed problem, we have to take into account that it still may be infeasible. It is therefore necessary to perform a feasibility check for the computation of the value $I(t, g)$ with $(t, g) \in A$ as well as for the corresponding lower bound $\underline{I}_A$ (or the respective mapped variables (π, γ), see Equation 4.61).

Example 4.3.6 In Section 4.2.3, Example 4.2.1 illustrated some properties of the problem of joint controller and transceiver optimization. The respective values of t and g, i. e., the transmit and receive scaling factors, and of the resulting Lagrange multipliers λ_1 and λ_2 were given without an explanation how they have been obtained. At this point, Algorithm 4.2 is applied to optimization problem (4.56) to compute these values. The algorithm is initialized with the set $B = [0, 1]^2$ (cf. Equation 4.61) and with a desired relative accuracy of $\varepsilon = 10^{-2}$. Due to the repeated subdivision, subsets $A \subset B$ which lead to an infeasible optimization problem for the determination of the lower bound $\underline{I}_A$ are eventually discarded. The subdivision rule for the set B and its subsets in line 11 of Algorithm 4.2 is the bisection of the larger side of B as mentioned above.

As in Example 4.2.1, the SNRs of the observation and the control channel are constrained by

$$\log_{10}\left(\operatorname{tr}\left[\boldsymbol{C_q}\right]^{-1} P_{\mathrm{Tx},1}\right) = \log_{10}\left(\operatorname{tr}\left[\boldsymbol{C_n}\right]^{-1} P_{\mathrm{Tx},2}\right) = 31.1,$$

[11] This can also be interpreted as setting $\overline{t}P_{\mathrm{Tx},1} = \infty$ or $\overline{g}P_{\mathrm{Tx},2} = \infty$, respectively, and using the convention $0 \cdot \infty = 0$ in Equations (4.59) and (4.60).

Algorithm 4.2 Branch and bound approach for the joint optimization of controller and transceiver under power constraints (cf. [102, 103])

Select a desired relative accuracy $\varepsilon > 0$
Use as initial partition $\mathbb{S}_1 = \mathbb{P}_1 = B$
$k = 1$
$\underline{I}^* = 0$ and $\overline{I}^* = \infty$
while $1 - \frac{\underline{I}^*}{\overline{I}^*} \geq \varepsilon$ **do**
    Compute the lower bound $\underline{I}_A$ for each $A \in \mathbb{P}_k$
    Compute the upper bound $\overline{I}_A$ for each $A \in \mathbb{P}_k$
    Determine the smallest upper bound $\overline{I}^* = \min_{A \in \mathbb{S}_k} \overline{I}_A$
    Remove every $A \in \mathbb{S}_k$ with $\underline{I}_A \geq \overline{I}^*$ from $\mathbb{S}_k$ and let the set of remaining members of $\mathbb{S}_k$ be $\mathbb{R}_k$
    Determine $\underline{I}^* = \min_{A \in \mathbb{R}_k} \underline{I}_A$ as well as $B = \operatorname{argmin}_{A \in \mathbb{R}_k} \underline{I}_A$ (if more than one minimizer is present, choose one randomly)
    Determine the partition $\mathbb{P}_{k+1}$ of B according to the chosen subdivision method
    $\mathbb{S}_{k+1} = (\mathbb{R}_k \setminus B) \cup \mathbb{P}_{k+1}$
    $k \leftarrow k + 1$
end while

where the parameters of the dynamical system to be controlled and of the noise sequences are again taken from Example 3.1.1. With the described initialization, 4.2 provides the solution

$$t^* \approx 2.42618 \qquad \text{and} \qquad g^* \approx 2.00569.$$

For these values, the corresponding Lagrange multipliers are given by

$$\lambda_1^* \approx 9.40201 \qquad \text{and} \qquad \lambda_2^* \approx 23.9293.$$

The transmit power constraints are fulfilled with an absolute deviation of less than $2 \cdot 10^{-10}$, and the value of the LQG cost function for the solution of problem (4.56) is given by

$$J_\infty^* \approx 2580.25.$$

Note that the derivatives of L^* (see Equation 4.57) w.r.t. t and g (where the results of Appendix A4 are used to determine the derivatives) are not zero at t^*, g^*, λ_1^* and λ_2^*, respectively, implying that these values are not optimal. In order to determine a local optimum in the neighborhood of (t^*, g^*), a gradient descent which is initialized at this point is performed. The resulting values of the transmit and receive scaling factors as well as the associated Lagrange multipliers read as

$$\begin{aligned} t_\nabla &\approx 2.42593, & g_\nabla &\approx 2.00565, \\ \lambda_{1,\nabla} &\approx 9.44165, & \lambda_{2,\nabla} &\approx 24.0299. \end{aligned}$$

The norm of the gradient of L^* at this point is $\|\nabla_{t,g,\lambda_1,\lambda_2} L^*(t_\nabla, g_\nabla, \lambda_{1,\nabla}, \lambda_{2,\nabla})\|_2 \approx 1.75 \cdot 10^{-8}$. The absolute value of the difference between the corresponding LQG cost and the value of J_∞^* given above is less than $3 \cdot 10^{-4}$, i. e., even less than the desired relative accuracy.

For a comparison of the result provided by the monotonic optimization approach with the solution of the subsequent local optimization based on the gradient descent, Figure 4.14 shows a contour plot of the function I (see Equation 4.58) in the neighborhood of pairs (t^*, g^*) and (t_∇, g_∇). Thus, for the whole range of transmit and receive scaling factors shown there, the power constrained controller optimization is feasible and the power constraints are fulfilled. The smallest value of the resulting LQG cost function is attained using (t_∇, g_∇). This pair which has been obtained by the gradient descent mentioned above is depicted in Figure 4.14 as the point $\times$. The associated value of the cost function is given by $I(t_\nabla, g_\nabla) \approx 2580.25168$.

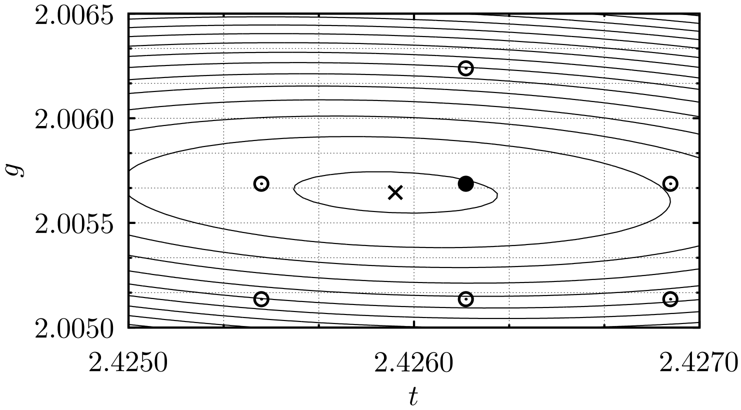


Figure 4.14: Values of the function $I(t, g)$ in the neighborhood of the result provided by Algorithm 4.2. Points which have been evaluated by Algorithm 4.2 are denoted by $\circ$, the minimizing point by $\times$.

The LQG cost at the point (t^*, g^*) (depicted as $\bullet$ in the figure above) which is the result of the monotonic optimization approach, is $I(t^*, g^*) \approx 2580.25192$, i. e., only slightly larger than $I(t_\nabla, g_\nabla)$. Feasible points which have also been evaluated by Algorithm 4.2 are marked by $\circ$. Note that the largest value of the cost function for all pairs of transmit and receive scaling factors shown in Figure 4.14 is approximately 2580.28058.

In the example above, we were able to determine feasible values of the transmit and receive scaling factors t and g as well as the Lagrange multipliers λ_1 and λ_2 such that the gradient of the function L^* vanishes. Thus, it has been shown that the actual accuracy of the result which is determined by the monotonic optimization approach is significantly better than ε, i. e., the relative accuracy which is guaranteed by Algorithm 4.2.

Unfortunately, despite the fact that it worked for Example 4.3.6, the gradient descent algorithm mentioned above is based on some assumptions which might not hold in general. Nevertheless, we give a short outline of the approach in the following. The gradient descent is actually performed using the function $I(t, g)$ (cf. Equation 4.58) which provides the optimal value of the power constrained controller optimization with fixed values of t and g. Since this function is the supremum w.r.t. the Lagrange multipliers λ_1 and λ_2, it is not obvious how to compute its derivatives w.r.t. t and g. Assume that $I(t, g)$ is actually a maximum and that it is attained at λ_1^* and λ_2^*, i. e.,

$$I(t, g) = \max_{\substack{\lambda_1 \geq 0 \\ \lambda_2 \geq 0}} L^*(t, g, \lambda_1, \lambda_2) = L^*(t, g, \lambda_1^*, \lambda_2^*), \tag{4.62}$$

and

$$(\lambda_1^*, \lambda_2^*) = \operatorname*{argmax}_{\substack{\lambda_1 \geq 0 \\ \lambda_2 \geq 0}} L^*(t, g, \lambda_1, \lambda_2). \tag{4.63}$$

With reference to Equation (4.63), the optimizing Lagrange multipliers λ_1^* and λ_2^* are functions of the transmit and receive scaling factors, i. e.,

$$\lambda_1^* = \lambda_1(t, g) \qquad \text{and} \qquad \lambda_2^* = \lambda_2(t, g). \tag{4.64}$$

Using the chain rule for the computation of the derivative of I w.r.t. t and g, we get

$$\frac{\partial I}{\partial t} = \frac{\partial L^*}{\partial t} + \frac{\partial L^*}{\partial \lambda_1}\frac{\partial \lambda_1}{\partial t} + \frac{\partial L^*}{\partial \lambda_2}\frac{\partial \lambda_2}{\partial t}, \tag{4.65}$$

where it is assumed that the respective derivatives exist (especially for the derivatives of the Lagrange multipliers w.r.t. t this is not obvious). The derivative $\frac{\partial I}{\partial g}$ is given by the analogous expression. If the transmit power constraints are active, i. e., $\lambda_1 > 0$ and $\lambda_2 > 0$, Equation (4.63) together with the assumption of differentiability implies that $\frac{\partial L^*}{\partial \lambda_1}(t, g, \lambda_1^*, \lambda_2^*) = 0$ as well as $\frac{\partial L^*}{\partial \lambda_2}(t, g, \lambda_1^*, \lambda_2^*) = 0$. Thus, it holds

$$\frac{\partial I}{\partial t}(t, g) = \frac{\partial L^*}{\partial t}(t, g, \lambda_1^*, \lambda_2^*), \tag{4.66}$$

with the analogous result for the derivative w.r.t. g. Consequently, the derivatives of I w.r.t. t and g are obtained by first solving the power constrained LQG control problem for fixed values of the transmit and receive scaling factors, which provides the Lagrange multipliers λ_1^* and λ_2^* (cf. Equations 4.62 and 4.63). Then, the derivative of L^* (see Equation 4.57) w.r.t. t and g, respectively, is computed for the given values λ_1^* and λ_2^*.

The assumption of strictly positive values of λ_1^* and λ_2^* at the minimum of I, i. e., the power constraints are active and thus hold with equality, can be verified with Equation (4.57). The KKT conditions (see, e. g., [95, Section 4.2.13]), which are necessary for a locally optimal point, must hold, i. e., the derivatives of L^* w.r.t. t and g must vanish at the optimum. Thus, we require that

$$\begin{aligned} &\frac{\partial L^*}{\partial t}(t, g, \lambda_1^*, \lambda_2^*) = \operatorname{tr}\left[\boldsymbol{P}\frac{\partial \boldsymbol{C}_{\tilde{\boldsymbol{x}}}}{\partial t}\right] - \lambda_1^* P_{\mathrm{Tx},1} = 0, \quad \text{and} \\ &\frac{\partial L^*}{\partial g}(t, g, \lambda_1^*, \lambda_2^*) = \operatorname{tr}\left[\boldsymbol{K}\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\right] + \operatorname{tr}\left[\boldsymbol{P}\frac{\partial \boldsymbol{C}_{\tilde{\boldsymbol{x}}}}{\partial g}\right] + \operatorname{tr}\left[\boldsymbol{R}\boldsymbol{C}_{\boldsymbol{n}}\right] - \lambda_2^* P_{\mathrm{Tx},2} = 0. \end{aligned} \tag{4.67}$$

With the results of Appendix A4 it can be shown that the derivative of $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}$ w.r.t. t and g, respectively, is a positive semidefinite matrix. Assuming that the derivative $\frac{\partial \boldsymbol{C}_{\tilde{\boldsymbol{x}}}}{\partial t}$ and the noise covariance matrix $\boldsymbol{C}_{\boldsymbol{n}}$ or the weighting matrix $\boldsymbol{R}$ are positive definite and noting that $P_{\mathrm{Tx},1} > 0$ and $P_{\mathrm{Tx},2} > 0$, it can be seen that $\lambda_1^* > 0$ and $\lambda_2^* > 0$.

It has to be mentioned again that the argumentation based on Equations (4.62) – (4.67) relies on a number of assumptions which do not hold in general. Nevertheless, the goal of the preceding discussion is not to provide a universally applicable local minimization algorithm for the determination of the optimal transmit and receive scaling factors. Instead, it demonstrates that the result provided by the monotonic optimization approach may be much better than the guaranteed accuracy suggests.

A final example demonstrates that the introduction of transmit and receive scaling factors is beneficial for the system performance even if no explicit transmit power constraints are given and the control loop which is closed over two communication channels could be optimized using an unconstrained LQG approach.

Example 4.3.7 Let $J^*_{\infty,\text{LQG}}$ be the value of the LQG cost function which is obtained by an optimization with no power constraints and with no transceivers, i. e., with $t = 1$ and $g = 1$, which means that $J^*_{\infty,\text{LQG}} = L^*(1, 1, 0, 0)$, see Equation (4.57). This case has been considered in Example 3.1.1 and provides the optimal value

$$J^*_{\infty,\text{LQG}} \approx 800.5.$$

The corresponding transmit powers have been computed there and are given by (cf. Equation 3.16) $P_{1,\text{LQG}} \approx 4189$ and $P_{2,\text{LQG}} \approx 1661$.

For the comparison with the LQG controller which uses the optimal scaling factors t and g, the transmit power constraints of problem (4.56) are set to the values of the unconstrained controller optimization, i. e., we require that $P_1 \leq P_{\text{Tx},1} = P_{1,\text{LQG}}$ and $P_2 \leq P_{\text{Tx},2} = P_{2,\text{LQG}}$. The application of Algorithm 4.2 to the system in Example 3.1.1 determines the solution of the joint optimization of controller and transceiver and provides the corresponding parameters

$$\begin{aligned} t^* &\approx 0.0790306, & g^* &\approx 0.0804537, \\ \lambda_1^* &\approx 0.313852, & \lambda_2^* &\approx 0.960286. \end{aligned}$$

Note that Algorithm 4.2 is initialized with $B = [0, 1]^2$, where the mapping to the actual optimization variables t and g is performed according to Equation (4.61), and a relative accuracy of $\varepsilon = 10^{-2}$ is chosen. Using the values given above, the power constraints of problem (4.56) are met with an absolute deviation of less than 10^{-9}, i. e., the power constrained LQG controller with the optimal choice of transmit and receive scaling factors uses the same amount of transmit power as the LQG controller without transceiver and power constraints. Nevertheless, the optimal value provided by Algorithm 4.2 is

$$J^*_\infty \approx 119.037.$$

This significant decrease of the LQG cost is due to the relatively small values of t and g which reduce the negative effect of the channel noise on the system performance. Because of the amplification of the transmit signals in the observation and control channel (recall that these signals are multiplied by $t^{-\frac{1}{2}}$ and $g^{-\frac{1}{2}}$, respectively), the Lagrange multipliers λ_1 and λ_2 have a positive value. Thus, a penalty based on the respective transmit power is added to the Lagrangian L^* (cf. Equation 4.11). This Lagrangian is the effective cost function for the determination of the optimal controller. Obviously, the fact that the controller now uses the weighting matrices $\boldsymbol{Q}+\lambda_1\boldsymbol{C}^{\text{T}}\boldsymbol{C}$ and $\boldsymbol{R}+\lambda_2\mathbf{I}_{N_u}$ (cf. Equation 4.14) as performance criterion instead of $\boldsymbol{Q}$ and $\boldsymbol{R}$ alone (cf. Equation 3.4) and thus does not simply focus in the smallest value of the original cost function, is more than compensated by the reduction of the channel noise which is fed into the closed loop control system.

4.4 Discussion

Having identified some of its shortcomings, the system model for the solution of power constrained LQG control problems from Chapter 3 has been extended in the present chapter by the introduction of scaling factors as the most simple instances of linear and memoryless transmitters and receivers at the input and the output of the system to be controlled. These additional degrees of freedom allow for an amplification or attenuation of the signals which are transmitted in the closed control loop and thus offer the possibility to either meet transmit power requirements which can not be fulfilled by the controller alone or to improve the system performance if the available transmit power is increased.

In Section 4.2.1, the unsurprising fact has been shown that the only impact of transmit and receive scaling factors effectively is a change of the variance of the channel noise sequences. Consequently, for fixed values of these factors, the power constrained controller optimization reduces to the case discussed in Chapter 3. A new quality is introduced by the joint optimization of the transmit and receive scaling factors and the controller in Sections 4.2.2 and 4.2.3. Compared to the optimization of the controller alone, the appealing property of convexity is lost and the jointly optimal solution of transmitters, receivers and controller can not be determined by a dual approach, which has been illustrated in Section 4.2.3. Even the set of feasible values of the transmit and receive scaling factors for given power constraints is not necessarily connected (see Example 4.3.1). Note that the non-convexity of distributed, i. e., structurally constrained, controller optimizations is a well known phenomenon, see, e. g., [39–41].

In order to get an impression about the set of feasible transmit power constraints, Pareto optimal values have been computed by the minimization of the weighted sum of the transmit powers of the observation and the control channel. In Section 4.3.2, it has been shown that this approach requires the determination of the optimal ratio of the transmit and receive scaling factors which is a demanding task due to the non-convexity of the considered optimization problem and the existence of local optima (see Example 4.3.2). Despite the fact that convexity can not be used to efficiently calculate Pareto optimal values of the transmit powers, the problem exhibits monotonicity properties which allow for the application of the monotonic optimization framework for global minimization [102, 103], see Section 4.3.3. This approach has been used to minimize the weighted sum of the transmit powers in Section 4.3.3.1, where it could also be observed that not all Pareto optimal values can be determined this way. In order to fill this gap, such transmit powers have been computed in Section 4.3.3.2 by minimizing the power of the observation channel with a constraint for the power of the control channel. Again, the monotonic optimization framework has been applied. Finally, a solution method for the original problem of the joint optimization of controller, transmitter and receiver based on the presented global optimization method has been introduced in Section 4.3.4.

Note that we did no consider problems as in [106] where the communication channel itself has a non-trivial transfer function, or tracking problems as in [35, 46]. The reason is that if the corresponding processes, represented by the transfer function of the communication channel or the power spectral density of the signal to be tracked, allow for a state space representation, they can be included in the description of the dynamical system to be controlled and the resulting optimization can be treated as a standard LQG control problem.

Despite its simplicity, the model of additive noise communication channel is used in a considerable amount of literature related to NCS, e. g., [29, 35, 59, 81–88, 99, 107]. Especially the more recent publications often focus on a scenario where only one communication channel, i. e., the observation or the control channel, is present in the control loop while the other channel is assumed to be ideal. For the further restriction to Single-Input Single-Output (SISO) systems, fundamental results have been derived like the smallest value of the channel SNR which allows for the stabilization of an unstable Linear Time-Invariant (LTI) system [29, 89, 106, 107]. Since this scenario is a special case of the general system model considered in Chapter 4, the proposed numerical methods are also applicable and allow for the minimization of transmit powers or the optimization of an LQG cost function subject to a transmit power or SNR constraint. The corresponding results for SISO systems with one communication channel in the control loop are presented in Chapter 6 where also some of the results of [29], which rely on a frequency domain approach, are reproduced in the framework of this thesis.

It is important to note that in [35, 46, 108] essentially the same problem as in this thesis, i. e., a power constrained controller optimization with the additional degree of freedom to choose the variance of the channel noise, is treated for the one channel SISO case. The main difference is that a frequency domain approach is applied and that the properties of the optimization w.r.t. the channel noise variance (in our case the transmit or receive scaling factors) are not discussed in detail. Especially the non-convexity of the problem which becomes apparent in the examples presented in Chapter 4 and which prohibits a straightforward solution is not considered.

Finally, we want to mention the early contribution [84] to the investigation of control loops which are closed over communication channels with fixed SNR, where the imperfect communication is due to computations with finite precision and thus rounding errors. The SNR constraints are not explicitly taken into account, but the covariance matrices of the channel noises are expressed as functions of the covariance matrix of the system state and thus implicitly define the constraints. The corresponding optimization problem is solved using a formulation based on Lyapunov(-like) and Riccati(-like) equations and with a Lagrangian approach. Additionally, an algorithm for the computation of a solution (if one exists) is provided. A major difference to the other approaches mentioned so far is that a static feedback controller which has only access to a noisy version of the system state is used, i. e., no state estimate is computed. Consequently, the available state information is not used optimally. The absence of the state estimation problem leads to the fact the optimization problem considered in [84] is convex and can be solved with an approach analogous to the one presented in Chapter 3, where the respective continuous-time formulation of the problem has to be used. This comment also applies to [109] where the authors assume that the controller has perfect knowledge about the system state and that the communication channel can be modeled as an additive noise channel. Despite the fact that the variance of the channel noise is considered as an optimization variable, the joint optimization together with the controller turns out to be a convex problem. It is expected that this property is lost for the output feedback case, i. e., if the choice of the channel noise has an impact on the error of the optimal state estimate.

5. Joint Optimization of Controller and Diagonal Transceivers

In Chapter 4, we considered the problem of the joint optimization of a Linear Quadratic Gaussian (LQG) controller together with *scalar* transceivers, i. e., the processing of signals at the input and the output of the dynamical system to be controlled was represented by a scaled identity matrix (see Figure 4.1). The non-convexity of the problem was the reason for the introduction of the monotonic optimization framework for the global minimization of the LQG cost function subject to transmit power constraints. Due to the generality of this optimization approach, the question arises now if it can be used to extend the simple scaling factors at the system input and output to a larger class of transceivers. In the following, it will be shown that this can be achieved in principle, but a straightforward generalization of the preceding results is only possible for the special case of diagonal covariance matrices of the channel noise. For the general case, a suboptimal approach based on the diagonalization of the covariance matrices is proposed in Section 5.2.1. Finally, the case of additive noise with general covariance matrix in addition to linear distortions of the channel input signal, represented by a channel matrix $\boldsymbol{H}$ (see Equation 1.2 and Figure 1.4), is discussed in Section 5.2.2.

5.1 Communication Channels With Diagonal Noise Covariance Matrices

With the assumption that both covariance matrices of the additive noise sequences in the observation and the control channel are diagonal, the methods for the analysis and the optimization of the power constrained control problem developed in Chapter 4 can be extended in a straightforward way to diagonal transmit and receive scaling matrices. We start with the description of the respective model of the communication channels, transmitters and receivers.

5.1.1 System Model

The system to be controlled is, as in the previous to chapters, the discrete-time linear dynamical system which has been introduced in Equation (1.1), with the assumption that it is time-invariant, i. e., the system parameters are constant. For the closed loop control, additive noise communication channels are used for the transmission of observations from the system output to the controller and of control signals from the controller back to the input of the system (see Section 1.6 and Figures 3.1 and 4.1). Assume now that the communication channels which are used to transmit the N_y-dimensional observations and the N_u-dimensional control signals consist of N_y and N_u independent scalar sub-channels, respectively. In this case, it is reasonable to model the corresponding additive noise in all scalar channels as mutually independent random sequences. Due to the independence assumption, the covariance matrices which describe the joint distribution of the scalar noise sequences are diagonal, i. e.,

$$\boldsymbol{C}_{\boldsymbol{q}} = \operatorname{diag}\left[c_{q,i}\right]_{i=1}^{N_y} \qquad \text{and} \qquad \boldsymbol{C}_{\boldsymbol{n}} = \operatorname{diag}\left[c_{n,i}\right]_{i=1}^{N_u}, \tag{5.1}$$

respectively. Thus, the variance of the channel noise in the i-th scalar channel for the observations and the control signal is given by $c_{q,i} > 0$, $i \in \{1, 2, \ldots, N_y\}$, and $c_{n,i} > 0$, $i \in \{1, 2, \ldots, N_u\}$, respectively.[1]

In Chapter 4, the system model which has been used in Chapter 3 for the solution of the LQG control problem under power constraints has been extended by scaling factors at the input and the output of the dynamical system to be controlled. These additional degrees of freedom allowed for a more efficient use of the available communication resources compared to the case when the controller alone has to satisfy all control and communication requirements. Regarding the model of independent sub-channels, represented by the diagonal noise covariance matrices shown in Equation (5.1), it is desirable to use a different scaling factor for each of these channels. The reason for this extension is that the quality of the channels, represented by the respective Signal to Noise Ratios (SNRs), is not equal in general, or it may be necessary to take into account different power constraints for different channels. Formally, this modification is represented by replacing the scaled identity matrices at the system input and output (see Figure 4.1) by diagonal matrices, i. e., by introducing the diagonal scaling

$$\boldsymbol{T} = \operatorname{diag}\left[t_i\right]_{i=1}^{N_y}, \qquad t_i > 0,\ i \in \{1, 2, \ldots, N_y\}, \tag{5.2}$$

for the observation channel and

$$\boldsymbol{G} = \operatorname{diag}\left[g_i\right]_{i=1}^{N_u}, \qquad g_i > 0,\ i \in \{1, 2, \ldots, N_u\}, \tag{5.3}$$

for the control channel. Note that the diagonal elements of $\boldsymbol{T}$ and $\boldsymbol{G}$ are assumed to be strictly positive. A negative sign has no influence on the closed control loop because the optimal LQG controller can invert the sign of the respective communication channel without changing the diagonal covariance matrix of the channel noise. Consequently, negative values can be excluded without loss of generality. This is not true for the restriction to non-zero values. In Chapter 4, the value of the scaling factors at the system input and output had to be non-zero. Otherwise the properties of stabilizability or detectability of the resulting system would have been lost. With the different scaling factors for the individual scalar sub-channels which are taken into account in this chapter, some of these channels can be switched off, i. e., multiplied by zero at the channel input or output, as long as the resulting open loop system remains stabilizable and observable. Regarding the joint optimization of the controller and the transceivers, it is not obvious that the optimizing values of the transmit and receive scaling factors for the individual sub-channels are non-zero because every scalar channel introduces additional noise to the closed loop system and requires additional transmit power. Thus, the optimal transceivers may switch off one or more scalar channels. Since the focus of this chapter lies on the investigation of diagonal instead of scalar transceivers, we do not consider the additional problem of how to determine the optimal subset of the available scalar channels for the transmission of control inputs and of observations obtained at the system output.

With the restriction to strictly positive scaling factors for each sub-channel (cf. Equations 5.2 and 5.3), we obtain the system model depicted in Figure 5.1 which allows for a notation analogous to the preceding chapter due to the invertibility of the scaling matrices $\boldsymbol{T}$ and $\boldsymbol{G}$. Note that the square root of $\boldsymbol{T}^{-1}$ and $\boldsymbol{G}$ is used in Figure 5.1. The reason for this step is analogous to the scenario with scalar transceivers: the effective variance of the noise sequence in each scalar communication channel is proportional to the respective diagonal element of $\boldsymbol{T}$ and $\boldsymbol{G}$, which will become clear in the following.

[1]The variances of all scalar sub-channels are assumed to be strictly positive. Otherwise, the optimization of the LQG controller subject to power constraints is of no interest because any constraint can be met in this case.

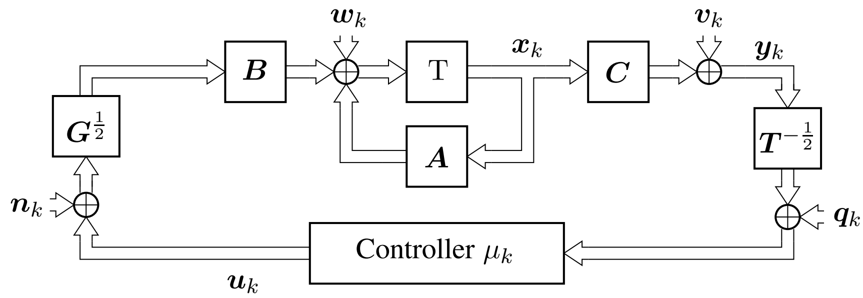


Figure 5.1: Model of the control loop which is closed over two channels with additive noise $\boldsymbol{q}_k$ and $\boldsymbol{n}_k$. Both noise sequences are assumed to have diagonal covariance matrices. At the system input and output, the diagonal receiver $\boldsymbol{G}^{\frac{1}{2}}$ and transmitter $\boldsymbol{T}^{-\frac{1}{2}}$, respectively, is introduced.

5.1.2 Optimization Problem and Solution

The optimization problem to be solved is, as in Chapters 3 and 4, an LQG control problem with power constraints for the observation and the control channel. The respective transmit powers which are limited by the available communication resources are given by expressions analogous to Equations (4.8) and (4.9), i. e.,

$$\begin{aligned} P_1 &= \lim_{N\to\infty} \frac{1}{N} \mathrm{E}\left[\sum_{n=0}^{N-1} \boldsymbol{x}_n^{\mathrm{T}} \boldsymbol{C}^{\mathrm{T}} \boldsymbol{T}^{-1} \boldsymbol{C} \boldsymbol{x}_n + \boldsymbol{v}_n^{\mathrm{T}} \boldsymbol{T}^{-1} \boldsymbol{v}_n\right] \qquad \text{and} \\ P_2 &= \lim_{N\to\infty} \frac{1}{N} \mathrm{E}\left[\sum_{n=0}^{N-1} \boldsymbol{u}_n^{\mathrm{T}} \boldsymbol{u}_n\right]. \end{aligned} \tag{5.4}$$

The values of P_1 and P_2 represent the sum of the powers of all scalar sub-channels for the transmission of observations and control signals, respectively. In the following, we consider constraints for these sum powers, but it is not a problem to take into account individual constraints for the scalar channels in the presented framework. This can be verified for fixed values of $\boldsymbol{T}$ and $\boldsymbol{G}$ using the problem formulation in Equation (3.22) where the sum-power constraints are expressed by the trace of the covariance matrices of the signals to be transmitted. Constraints for the diagonal elements of these covariance matrices (which represent the transmit powers of the individual scalar channels) are again inequalities containing linear functions of the optimization variables which renders the resulting optimization problem convex. The remaining minimization w.r.t. the diagonal elements of $\boldsymbol{T}$ and $\boldsymbol{G}$ has the monotonicity properties which are required for the application of the monotonic optimization framework introduced in Section 4.3.3 and which will be used for the subsequent derivations.

It has already been mentioned that we consider sum-power constraints (cf. Equation 5.4) in the following for the sake of simplicity. The cost function to be minimized is the infinite horizon average cost (see Appendix A6.3)

$$J_\infty = \lim_{N\to\infty} \frac{1}{N} \mathrm{E}\left[\boldsymbol{x}_N^{\mathrm{T}} \boldsymbol{Q}_N \boldsymbol{x}_N + \sum_{n=0}^{N-1} \begin{bmatrix} \boldsymbol{x}_n \\ \boldsymbol{G}^{\frac{1}{2}} \boldsymbol{u}_n \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{S} \\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{R} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_n \\ \boldsymbol{G}^{\frac{1}{2}} \boldsymbol{u}_n \end{bmatrix}\right] + \mathrm{tr}\left[\boldsymbol{R}\boldsymbol{G}^{\frac{1}{2}} \boldsymbol{C}_{\boldsymbol{n}} \boldsymbol{G}^{\frac{1}{2}}\right]. \tag{5.5}$$

Since the input which is actually applied to the dynamical system (see Figure 5.1) is $\boldsymbol{G}^{\frac{1}{2}}(\boldsymbol{u}_k + \boldsymbol{n}_k)$, $k \in \mathbb{N}_0$, the LQG cost in Equation (5.5) contains the additional term $\mathrm{tr}\left[\boldsymbol{R}\boldsymbol{G}^{\frac{1}{2}} \boldsymbol{C}_{\boldsymbol{n}} \boldsymbol{G}^{\frac{1}{2}}\right]$ due to the channel noise which is fed into the system. The control sequence $(\boldsymbol{u}_k : k \in \mathbb{N}_0)$ is multiplied by $\boldsymbol{G}^{\frac{1}{2}}$ at the system input which makes it necessary to take into account the control signal $\boldsymbol{G}^{\frac{1}{2}}\boldsymbol{u}_k$, $k \in \mathbb{N}_0$, in the cost function. The problem of the joint optimization of an LQG controller and the diagonal matrices $\boldsymbol{T}$ and $\boldsymbol{G}$ thus reads as

$$\begin{aligned}
&\underset{\boldsymbol{T},\boldsymbol{G},\mu_0,\mu_1,\mu_2\ldots}{\text{minimize}} \quad J_\infty \\
&\text{subject to} \quad \boldsymbol{x}_{k+1} = \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{B}\boldsymbol{G}^{\frac{1}{2}}(\boldsymbol{u}_k + \boldsymbol{n}_k) + \boldsymbol{w}_k, && k \in \mathbb{N}_0, \\
&\qquad \boldsymbol{y}_k = \boldsymbol{C}\boldsymbol{x}_k + \boldsymbol{v}_k, && k \in \mathbb{N}_0, \\
&\qquad \boldsymbol{u}_k = \mu_k(\mathcal{I}_k), && k \in \mathbb{N}_0, \\
&\qquad \mathcal{I}_k = \begin{cases} \{(\boldsymbol{T}^{-\frac{1}{2}}\boldsymbol{y}_0 + \boldsymbol{q}_0)\}, & k = 0, \\ \{(\boldsymbol{T}^{-\frac{1}{2}}\boldsymbol{y}_0 + \boldsymbol{q}_0), \ldots, (\boldsymbol{T}^{-\frac{1}{2}}\boldsymbol{y}_k + \boldsymbol{q}_k), \boldsymbol{u}_0, \ldots, \boldsymbol{u}_{k-1}\}, & k \in \mathbb{N}, \end{cases} \\
&\qquad \boldsymbol{T} = \operatorname{diag}\,[t_i]_{i=1}^{N_y} > \boldsymbol{0}_{N_y \times N_y}, \\
&\qquad \boldsymbol{G} = \operatorname{diag}\,[g_i]_{i=1}^{N_u} > \boldsymbol{0}_{N_u \times N_u}, \\
&\qquad P_1 \le P_{\mathrm{Tx},1}, \\
&\qquad P_2 \le P_{\mathrm{Tx},2}.
\end{aligned} \tag{5.6}$$

Note that the requirement of positive diagonal elements of $\boldsymbol{T}$ and $\boldsymbol{G}$ has been expressed as positive definiteness constraints above for notational compactness.

As in Section 4.3.4, we start the solution of optimization problem (5.6) by minimizing the LQG cost function w.r.t. the controller alone and for fixed values of $\boldsymbol{T}$ and $\boldsymbol{G}$. It has already been mentioned in Section 4.3.4 that the minimization of the Lagrangian associated with problem (5.6) i. e., of

$$L = J_\infty + \lambda_1 \left(P_1 - P_{\mathrm{Tx},1}\right) + \lambda_2 \left(P_2 - P_{\mathrm{Tx},2}\right), \tag{5.7}$$

w.r.t. to the controller μ_k, $k \in \mathbb{N}_0$, where $\lambda_1 \ge 0$ and $\lambda_2 \ge 0$ are the Lagrange multipliers which are associated with the constraints for P_1 and P_2, respectively, is again an LQG problem. This makes it easy to determine the minimum of L which is given by

$$\begin{aligned} L^*(\boldsymbol{T}, \boldsymbol{G}, \lambda_1, \lambda_2) = \operatorname{tr}\left[\boldsymbol{K}\big(\boldsymbol{C}_{\boldsymbol{w}} + \boldsymbol{B}\boldsymbol{G}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\big)\right] + \operatorname{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\right] + \lambda_1 \operatorname{tr}\left[\boldsymbol{T}^{-1}\boldsymbol{C}_{\boldsymbol{v}}\right] + \operatorname{tr}\left[\boldsymbol{R}\boldsymbol{G}\boldsymbol{C}_{\boldsymbol{n}}\right] \\ - \lambda_1 P_{\mathrm{Tx},1} - \lambda_2 P_{\mathrm{Tx},2}, \end{aligned} \tag{5.8}$$

where we used the fact that diagonal matrices commute and thus the square roots of $\boldsymbol{G}$ can be merged. Note that L^* in Equation (5.8) is explicitly denoted as a function of $\boldsymbol{T}$ and $\boldsymbol{G}$, and the dual variables λ_1 and λ_2. The matrices $\boldsymbol{K}$ and $\boldsymbol{P}$ are given by the stabilizing solution of the Discrete Algebraic Riccati Equation (DARE)

$$\begin{aligned} \boldsymbol{K} &= \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B}\boldsymbol{G}^{\frac{1}{2}} + \boldsymbol{S}\boldsymbol{G}^{\frac{1}{2}}\right)\left(\boldsymbol{G}^{\frac{1}{2}}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B}\boldsymbol{G}^{\frac{1}{2}} + \boldsymbol{G}^{\frac{1}{2}}\boldsymbol{R}\boldsymbol{G}^{\frac{1}{2}} + \lambda_2 \mathrm{I}_{N_u}\right)^{-1} \\ &\qquad \times \left(\boldsymbol{G}^{\frac{1}{2}}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} + \boldsymbol{G}^{\frac{1}{2}}\boldsymbol{S}^{\mathrm{T}}\right) + \boldsymbol{Q} + \lambda_1 \boldsymbol{C}^{\mathrm{T}}\boldsymbol{T}^{-1}\boldsymbol{C} \\ &= \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B} + \boldsymbol{S}\right)\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B} + \boldsymbol{R} + \lambda_2 \boldsymbol{G}^{-1}\right)^{-1}\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} + \boldsymbol{S}^{\mathrm{T}}\right) \\ &\qquad + \boldsymbol{Q} + \lambda_1 \boldsymbol{C}^{\mathrm{T}}\boldsymbol{T}^{-1}\boldsymbol{C}, \end{aligned} \tag{5.9}$$

with $\boldsymbol{P} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \boldsymbol{K} + \boldsymbol{Q} + \lambda_1 \boldsymbol{C}^{\mathrm{T}}\boldsymbol{T}^{-1}\boldsymbol{C}$. The covariance matrix $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}$ of the optimal estimate of the system state is determined by the stabilizing solution of the DARE

$$\begin{aligned}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} &= \boldsymbol{A}\bigg(\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{T}^{-\frac{1}{2}}\Big(\boldsymbol{T}^{-\frac{1}{2}}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{T}^{-\frac{1}{2}} + \boldsymbol{T}^{-\frac{1}{2}}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{T}^{-\frac{1}{2}} + \boldsymbol{C}_{\boldsymbol{q}}\Big)^{-1}\boldsymbol{T}^{-\frac{1}{2}}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\bigg)\boldsymbol{A}^{\mathrm{T}} \\ &\qquad + \boldsymbol{C}_{\boldsymbol{w}} + \boldsymbol{B}\boldsymbol{G}^{\frac{1}{2}}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{G}^{\frac{1}{2}}\boldsymbol{B}^{\mathrm{T}} \\ &= \boldsymbol{A}\left(\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{T}\boldsymbol{C}_{\boldsymbol{q}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\right)\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{w}} + \boldsymbol{B}\boldsymbol{G}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\end{aligned} \tag{5.10}$$

and reads as $\boldsymbol{C}_{\tilde{\boldsymbol{x}}} = \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{T}\boldsymbol{C}_{\boldsymbol{q}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}$. Due to the convexity of the optimization problem for given values of $\boldsymbol{T}$ and $\boldsymbol{G}$ and with the assumption that strong duality holds, the minimal value of the optimization problem (5.6) w.r.t. the controller alone is given by

$$I(\boldsymbol{T}, \boldsymbol{G}) = \sup_{\substack{\lambda_1 \geq 0 \\ \lambda_2 \geq 0}} L^*(\boldsymbol{T}, \boldsymbol{G}, \lambda_1, \lambda_2). \tag{5.11}$$

For the remaining minimization of the LQG cost function w.r.t. to the diagonal matrices $\boldsymbol{T}$ and $\boldsymbol{G}$, the branch and bound approach presented in Section 4.3.3 is applied. In order to see that the required monotonicity is also present for the model with diagonal transceivers, note that the scaled covariance matrices of the channel noise sequences which contribute to the Lagrangian in Equation (5.8) and the covariance matrix in Equation (5.10) read as

$$\boldsymbol{T}\boldsymbol{C}_{\boldsymbol{q}} = \operatorname{diag}\left[t_i c_{q,i}\right]_{i=1}^{N_y} \qquad \text{and} \qquad \boldsymbol{G}\boldsymbol{C}_{\boldsymbol{n}} = \operatorname{diag}\left[g_i c_{n,i}\right]_{i=1}^{N_u}. \tag{5.12}$$

Thus, if $\overline{\boldsymbol{T}} \geq \boldsymbol{T}$ and $\overline{\boldsymbol{G}} \geq \boldsymbol{G}$, the corresponding error covariance matrices have the property that $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}(\overline{\boldsymbol{T}}, \overline{\boldsymbol{G}}) \geq \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}(\boldsymbol{T}, \boldsymbol{G})$, which follows from [100] or the results from Appendix A4. Additionally, it is intuitively clear that an increased noise variance results in a larger variance of the state estimation error in any subspace of the state space (or at least not in a smaller variance). These arguments carry over to the covariance matrix $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}$.[2] A final application of the results provides the relation $\boldsymbol{K}(\overline{\boldsymbol{T}}, \overline{\boldsymbol{G}}) \leq \boldsymbol{K}(\boldsymbol{T}, \boldsymbol{G})$ since $\boldsymbol{K}$ depends on the inverse of the diagonal transmit and receive matrix, i. e., $\boldsymbol{C}^{\mathrm{T}}\overline{\boldsymbol{T}}^{-1}\boldsymbol{C} \leq \boldsymbol{C}^{\mathrm{T}}\boldsymbol{T}^{-1}\boldsymbol{C}$ and $\overline{\boldsymbol{G}}^{-1} \leq \boldsymbol{G}^{-1}$.

In order to apply the branch and bound approach for the minimization of I (see Equation 5.11) w.r.t. $\boldsymbol{T}$ and $\boldsymbol{G}$, the monotonicity results mentioned above are used. Recall that this approach sequentially partitions the set B of possible values of the optimization variables. For each subset A of such a partition, upper and lower bounds of the optimal value of the optimization problem under consideration are computed where the optimization variables are constrained to lie within the subset A. The availability of such bounds then allows for a systematic search of the global optimum within the original set B. We refer to Section 4.3.3 for a detailed description.

Assume that a set A of diagonal transmit and receive matrices is defined by

$$A = \left\{(\boldsymbol{T}, \boldsymbol{G}) \,\middle|\, \underline{\boldsymbol{T}} \leq \boldsymbol{T} \leq \overline{\boldsymbol{T}}, \underline{\boldsymbol{G}} \leq \boldsymbol{G} \leq \overline{\boldsymbol{G}},\ \boldsymbol{T} = \operatorname{diag}\left[t_i\right]_{i=1}^{N_y}, \boldsymbol{G} = \operatorname{diag}\left[g\right]_{i=1}^{N_u}\right\}, \tag{5.13}$$

where $\underline{\boldsymbol{T}}, \overline{\boldsymbol{T}}, \underline{\boldsymbol{G}}$ and $\overline{\boldsymbol{G}}$ are diagonal matrices with positive diagonal elements. A lower bound for the optimal value of the problem (5.6) when $\boldsymbol{T}$ and $\boldsymbol{G}$ are constrained to lie in A is obtained as

[2] This means that by increasing the noise variance, it holds $\boldsymbol{z}^{\mathrm{T}}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}(\overline{\boldsymbol{T}}, \overline{\boldsymbol{G}})\boldsymbol{z} \geq \boldsymbol{z}^{\mathrm{T}}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}(\boldsymbol{T}, \boldsymbol{G})\boldsymbol{z}$ for all $\boldsymbol{z} \in \mathbb{R}^{N_x}$ and thus $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}(\overline{\boldsymbol{T}}, \overline{\boldsymbol{G}}) \geq \boldsymbol{C}_{\tilde{\boldsymbol{x}}}(\boldsymbol{T}, \boldsymbol{G})$.

follows. As a first step, note that for all pairs of transmitters and receivers within this set, a lower bound for the Lagrangian from Equation (5.8) is given by

$$
\begin{aligned}
L^*(\boldsymbol{T}, \boldsymbol{G}, \lambda_1, \lambda_2) &= \operatorname{tr}\left[\boldsymbol{K}(\boldsymbol{T}, \boldsymbol{G})\left(\boldsymbol{C}_{\boldsymbol{w}} + \boldsymbol{B}\boldsymbol{G}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\right)\right] + \operatorname{tr}\left[\boldsymbol{P}(\boldsymbol{T}, \boldsymbol{G})\boldsymbol{C}_{\tilde{\boldsymbol{x}}}(\boldsymbol{T}, \boldsymbol{G})\right] \\
&\qquad + \lambda_1 \operatorname{tr}\left[\boldsymbol{T}^{-1}\boldsymbol{C}_{\boldsymbol{v}}\right] + \operatorname{tr}\left[\boldsymbol{R}\boldsymbol{G}\boldsymbol{C}_{\boldsymbol{n}}\right] - \lambda_1 P_{\mathrm{Tx},1} - \lambda_2 P_{\mathrm{Tx},2} \\
&\geq \operatorname{tr}\left[\boldsymbol{K}(\boldsymbol{T}, \boldsymbol{G})\left(\boldsymbol{C}_{\boldsymbol{w}} + \boldsymbol{B}\underline{\boldsymbol{G}}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\right)\right] + \operatorname{tr}\left[\boldsymbol{P}(\boldsymbol{T}, \boldsymbol{G})\boldsymbol{C}_{\tilde{\boldsymbol{x}}}(\underline{\boldsymbol{T}}, \underline{\boldsymbol{G}})\right] \\
&\qquad + \lambda_1 \operatorname{tr}\left[\overline{\boldsymbol{T}}^{-1}\boldsymbol{C}_{\boldsymbol{v}}\right] + \operatorname{tr}\left[\boldsymbol{R}\underline{\boldsymbol{G}}\boldsymbol{C}_{\boldsymbol{n}}\right] - \lambda_1 P_{\mathrm{Tx},1} - \lambda_2 P_{\mathrm{Tx},2} \\
&= \operatorname{tr}\left[\boldsymbol{K}(\boldsymbol{T}, \boldsymbol{G})\left(\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}(\underline{\boldsymbol{T}}, \underline{\boldsymbol{G}}) - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}(\underline{\boldsymbol{T}}, \underline{\boldsymbol{G}})\right)\right] \\
&\qquad + \operatorname{tr}\left[\left(\boldsymbol{Q} + \lambda_1 \boldsymbol{C}^{\mathrm{T}}\boldsymbol{T}^{-1}\boldsymbol{C}\right)\boldsymbol{C}_{\tilde{\boldsymbol{x}}}(\underline{\boldsymbol{T}}, \underline{\boldsymbol{G}})\right] \\
&\qquad + \lambda_1 \operatorname{tr}\left[\overline{\boldsymbol{T}}^{-1}\boldsymbol{C}_{\boldsymbol{v}}\right] + \operatorname{tr}\left[\boldsymbol{R}\underline{\boldsymbol{G}}\boldsymbol{C}_{\boldsymbol{n}}\right] - \lambda_1 P_{\mathrm{Tx},1} - \lambda_2 P_{\mathrm{Tx},2} \\
&\geq \operatorname{tr}\left[\boldsymbol{K}(\overline{\boldsymbol{T}}, \overline{\boldsymbol{G}})\left(\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}(\underline{\boldsymbol{T}}, \underline{\boldsymbol{G}}) - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}(\underline{\boldsymbol{T}}, \underline{\boldsymbol{G}})\right)\right] \\
&\qquad + \operatorname{tr}\left[\left(\boldsymbol{Q} + \lambda_1 \boldsymbol{C}^{\mathrm{T}}\overline{\boldsymbol{T}}^{-1}\boldsymbol{C}\right)\boldsymbol{C}_{\tilde{\boldsymbol{x}}}(\underline{\boldsymbol{T}}, \underline{\boldsymbol{G}})\right] \\
&\qquad + \lambda_1 \operatorname{tr}\left[\overline{\boldsymbol{T}}^{-1}\boldsymbol{C}_{\boldsymbol{v}}\right] + \operatorname{tr}\left[\boldsymbol{R}\underline{\boldsymbol{G}}\boldsymbol{C}_{\boldsymbol{n}}\right] - \lambda_1 P_{\mathrm{Tx},1} - \lambda_2 P_{\mathrm{Tx},2} \\
&= \underline{L}^*_A(\lambda_1, \lambda_2).
\end{aligned}
\tag{5.14}
$$

For the derivation, the matrices $\boldsymbol{K}$, $\boldsymbol{P}$, $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}$ and $\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}$ are denoted as functions of the diagonal transmit and receive matrices and the fact is used that $\boldsymbol{P} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \boldsymbol{K} + \boldsymbol{Q} + \lambda_1\boldsymbol{C}^{\mathrm{T}}\boldsymbol{T}^{-1}\boldsymbol{C}$. Note the similar expressions in Equations (4.36) and (4.59) which have been derived earlier. A reformulation of the lower bound $\underline{L}^*_A(\lambda_1, \lambda_2)$ as

$$
\begin{aligned}
\underline{L}^*_A(\lambda_1, \lambda_2) &= \operatorname{tr}\left[\boldsymbol{K}(\overline{\boldsymbol{T}}, \overline{\boldsymbol{G}})\left(\boldsymbol{C}_{\boldsymbol{w}} + \boldsymbol{B}\underline{\boldsymbol{G}}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\right)\right] + \operatorname{tr}\left[\boldsymbol{P}(\overline{\boldsymbol{T}}, \overline{\boldsymbol{G}})\boldsymbol{C}_{\tilde{\boldsymbol{x}}}(\underline{\boldsymbol{T}}, \underline{\boldsymbol{G}})\right] \\
&\qquad + \lambda_1 \operatorname{tr}\left[\overline{\boldsymbol{T}}^{-1}\boldsymbol{C}_{\boldsymbol{v}}\right] + \operatorname{tr}\left[\boldsymbol{R}\underline{\boldsymbol{G}}\boldsymbol{C}_{\boldsymbol{n}}\right] - \lambda_1 P_{\mathrm{Tx},1} - \lambda_2 P_{\mathrm{Tx},2}
\end{aligned}
\tag{5.15}
$$

shows that it is the dual function of a power constrained LQG problem where the controller gain (represented by the matrices $\boldsymbol{K}$ and $\boldsymbol{P}$) is determined using $\overline{\boldsymbol{T}}$ and $\overline{\boldsymbol{G}}$, whereas the estimator problem (represented by the error covariance matrices $\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}$ and $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}$) is solved using $\underline{\boldsymbol{T}}$ and $\underline{\boldsymbol{G}}$, i. e., by an underestimate of the variance of channel noise sequences. The fact that the controller gain is determined based on $\overline{\boldsymbol{T}}$ and $\overline{\boldsymbol{G}}$ results in the possibility to use a control signal with smaller variance compared to the case with $\boldsymbol{G}$ since the gain at the system input is increased. Thus, the power constraint on the control signal is less restrictive. The application of $\overline{\boldsymbol{T}}$ at the system output effectively relaxes the power constraint for the observation channel. Together with the underestimate of the variances of the channel noise sequences, the corresponding relaxed optimization problem leads to the lower bound $\underline{L}^*_A$ shown in Equation (5.15).

With the availability of a lower bound for L^*, the respective lower bound for the optimal value of the optimization problem (5.6) for given values of $\boldsymbol{T}$ and $\boldsymbol{G}$ within the set A (cf. Equation 5.13) is determined analogous to Equation (4.60):

$$
I(\boldsymbol{T}, \boldsymbol{G}) = \sup_{\substack{\lambda_1 \geq 0 \\ \lambda_2 \geq 0}} L^*(\boldsymbol{T}, \boldsymbol{G}, \lambda_1, \lambda_2) \geq L^*(\boldsymbol{T}, \boldsymbol{G}, \underline{\lambda}_1, \underline{\lambda}_2) \geq \underline{L}^*_A(\underline{\lambda}_1, \underline{\lambda}_2) = \sup_{\substack{\lambda_1 \geq 0 \\ \lambda_2 \geq 0}} \underline{L}^*_A(\lambda_1, \lambda_2) = \underline{I}_A, \tag{5.16}
$$

where it is again assumed that strong duality holds and the corresponding optimization problem is feasible. For the case that the problem which is used for the determination of the lower bound is not feasible, $\underline{I}_A$ as well as $I(\boldsymbol{T}, \boldsymbol{G})$ are unbounded from above. The consistency of the lower bound, i. e., the fact that the value of the lower bound converges to the actual value of I if the set A collapses to a single point, is shown in Appendix A9. This property depends on the continuity of the solutions $\boldsymbol{K}$ and $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{p}}$ of the respective DAREs w.r.t. $\boldsymbol{T}$, $\boldsymbol{G}$. We refer again to [104] for a discussion of the continuity properties and note that the stabilizing solutions of the DAREs are continuous w.r.t. $\boldsymbol{T}$ and $\boldsymbol{G}$ if $\boldsymbol{R}$ and either $\boldsymbol{C}_{\boldsymbol{v}}$ or $\boldsymbol{C}_{\boldsymbol{q}}$ have full rank.

For the application of the branch and bound approach to the determination of the global optimum of problem (5.6) w.r.t. $\boldsymbol{T}$ and $\boldsymbol{G}$, it remains to find an upper bound for the optimal value within a given set A (see Equation 5.13). Additionally, a subdivision rule for such sets which provides an exhaustive partitioning process has to be determined as well as a method to handle the unboundedness and positivity of the diagonal elements of $\boldsymbol{T}$ and $\boldsymbol{G}$. The upper bound is simply obtained by evaluating I at an arbitrary point from A, i. e.,

$$\overline{I}_A = I(\boldsymbol{T}, \boldsymbol{G}), \qquad (\boldsymbol{T}, \boldsymbol{G}) \in A. \tag{5.17}$$

The subdivision rule is chosen to be a bisection, i. e., a set A is partitioned in two smaller sets by dividing it at the midpoint of its longest side. A partition process performed this way is exhaustive, i. e., eventually collapses to a single point (see, e. g., [102, 103]). For the sake of simplicity, the initial set for the search of the optimal diagonal transmit and receive matrices is chosen to be

$$B = \left\{ (\boldsymbol{T}, \boldsymbol{G}) \left| \underline{t}\, \mathbf{I}_{N_y} \leq \boldsymbol{T} \leq \overline{t}\, \mathbf{I}_{N_y}, \underline{g}\, \mathbf{I}_{N_u} \leq \boldsymbol{G} \leq \overline{g}\, \mathbf{I}_{N_u} \right. \right\}, \tag{5.18}$$

where $\underline{t}$, $\overline{t}$, $\underline{g}$ and $\overline{g}$ have to be selected according to the desired search interval.[3] Note that this choice does not guarantee that the optimal matrices $\boldsymbol{T}$ and $\boldsymbol{G}$ are elements of B. In the following, it is assumed that this requirement holds by choosing $\underline{t}$ and $\underline{g}$ sufficiently small and $\overline{t}$ and $\overline{g}$ sufficiently large. With the lower bound $\underline{I}_A$ (cf. Equation 5.16), the upper bound $\overline{I}_A$ (cf. Equation 5.17), the selection of the initial set B (cf. Equation 5.18) and the bisection subdivision rule, Algorithm 4.2 from Section 4.3.4 can now be used for the determination of the optimal diagonal transmit and receive matrices $\boldsymbol{T}$ and $\boldsymbol{G}$ within the set B.

In order to demonstrate the benefits of the joint optimization of LQG controller and scalar as well as diagonal transceivers, the following example considers a scenario with diagonal covariance matrices of the channel noise sequences and a power constraint which can also be fulfilled by the LQG controller alone.

Example 5.1.1 Analogous to the examples in the preceding sections, the system and noise parameters given in Example 3.1.1 are used, but channel noise sequences $(\boldsymbol{q}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{n}_k : k \in \mathbb{N}_0)$ are assumed to have uncorrelated components with the same variance as before. Thus, their covariance matrices read as

$$\boldsymbol{C}_{\boldsymbol{q}} = \mathrm{diag}\,[1, 0.5] \quad \text{and} \quad \boldsymbol{C}_{\boldsymbol{n}} = \mathrm{diag}\,[0.7, 0.3]\,,$$

[3] The positive diagonal elements of $\boldsymbol{T}$ and $\boldsymbol{G}$ can be mapped to the interval $]0, 1[$ as in the preceding sections (see Equations 4.39, 4.55 and 4.61).The boundary points 0 and 1 can also be included for the computation of the upper and lower bounds $\overline{I}_A$ and $\underline{I}_A$, respectively. These points correspond either to the case that one or more components of the observation or the control signal are multiplied by zero (which effectively reduces the dimension of the observation or the control channel) or to the case that one or more components of transmitted or received signals are amplified by an unbounded scaling factor (which leads to an unbounded value of the respective upper or lower bound).

respectively. Using these matrices, the solution of the LQG control problem without transmitters, receivers and power constraints is computed as a reference. Taking into account a limited transmit power, solutions are determined according to the approaches presented in Chapter 3 (no transmitters and receivers but power constraints), Section 4.3.4 (scalar transceivers) and Section 5.1 (diagonal transceivers). Table 5.1 shows the results for different LQG controller optimizations.

	LQG cost	P_1	P_2	$\boldsymbol{T}$	$\boldsymbol{G}$	λ_1	λ_2
no constraints	954.7	5020	1972	$\mathbf{I}_2$	$\mathbf{I}_2$	—	—
no Tx/Rx	1513	2377	1585	$\mathbf{I}_2$	$\mathbf{I}_2$	1.318	1.556
scalar[4]	895.6	2377	1585	$0.8347 \cdot \mathbf{I}_2$	$0.5401 \cdot \mathbf{I}_2$	3.038	10.21
diagonal	71.36	2377	1585	$\operatorname{diag}[t_1, t_2]$ $t_1 = 0.09930$ $t_2 = 0.06334$	$\operatorname{diag}[g_1, g_2]$ $g_1 = 0.005894$ $g_2 = 0.1551$	0.02177	0.008706
ideal channels	28.84	∞	∞	—	—	—	—

Table 5.1: Comparison of power constrained control approaches with and without transmitters and receivers.

The first row contains the values for the optimal LQG controller without transceiver (equivalently, the transmitter and receiver is represented by the identity matrix) which does not take into account any power constraints and closes the control loop over the communication channels with additive noise. Without power constrains, there are also no dual variables λ_1 and λ_2. The values of the resulting transmit powers correspond to SNRs of $10 \log_{10}(\operatorname{tr}\left[\boldsymbol{C_q}\right]^{-1} P_1) \approx 35.25$ and $10 \log_{10}(\operatorname{tr}\left[\boldsymbol{C_n}\right]^{-1} P_2) \approx 32.95$, respectively.

For the power constrained controller optimizations, we require that the logarithmic values of the SNRs of the observation as well as of the control channel fulfill the inequalities

$$10 \log_{10}\left(\operatorname{tr}\left[\boldsymbol{C_q}\right]^{-1} P_1\right) \leq 32 \qquad \text{and} \qquad 10 \log_{10}\left(\operatorname{tr}\left[\boldsymbol{C_n}\right]^{-1} P_2\right) \leq 32,$$

respectively, which corresponds to the upper bounds $P_{\mathrm{Tx},1} \approx 2377$ and $P_{\mathrm{Tx},2} \approx 1585$ for the respective transmit powers P_1 and P_2. For these values, the power constrained LQG control problem without transmitters and receivers which has been discussed in Chapter 3 is feasible. The solution of the resulting convex optimization problem is shown in the second row of Table 5.1. Note that the value of the cost function is increased significantly compared to the solution of the unconstrained LQG problem. This is not surprising since the transmit powers of both the observation and the control channel are required to be smaller than the powers for the case without power constraints.

The third row of Table 5.1 contains the results when a simple scaling factor is introduced at the system input and output (i. e., the transmitter and the receiver is represented by a scaled identity matrix) and optimized jointly with the controller using the approach described in 4.3.4. The desired relative accuracy of the minimal value of the cost function is $\varepsilon = 10^{-2}$ and the values of the transmit and receive scaling factors are mapped to the set $[0, 1]^2$ as described in Section 4.3.4 (see Equation 4.61). The subdivision rule is a bisection in the image space of this mapping. Since the scaling factors for $\boldsymbol{T} = t\mathbf{I}_2$ and $\boldsymbol{G} = g\mathbf{I}_2$ are both smaller than one, the impact of the channel noise on the control performance is reduced, which results in a smaller value of the LQG cost

[4] Note that in Section 4.3, the transmit power constraints are multiplied by t and g, respectively. For the constrained optimization problem without this multiplication (e. g., when comparing to the case without transceivers), the Lagrange multipliers read as $\nu_1 = t\lambda_1 \approx 2.536$ and $\nu_2 = g\lambda_2 \approx 5.516$, respectively, see Example 4.2.1.

function even with a smaller amount of transmit power compared to the original and unconstrained LQG controller.

Note that the problems related to the non-convexity of the joint optimization problem which have been encountered in the preceding sections remain for the case of diagonal noise covariance matrices. For example, Figure 5.2 shows pairs (t, g) of transmit and receive scaling factors which lead to a feasible optimization problem for the given power constraints. Compared to Figure 4.5 which shows analogous results for non-diagonal covariance matrices, we observe a similar behavior, especially the disjoint regions of feasible pairs. Consequently, local optimization approaches do not guarantee to find the global optimum which requires the application of global approaches like the presented branch and bound algorithm.

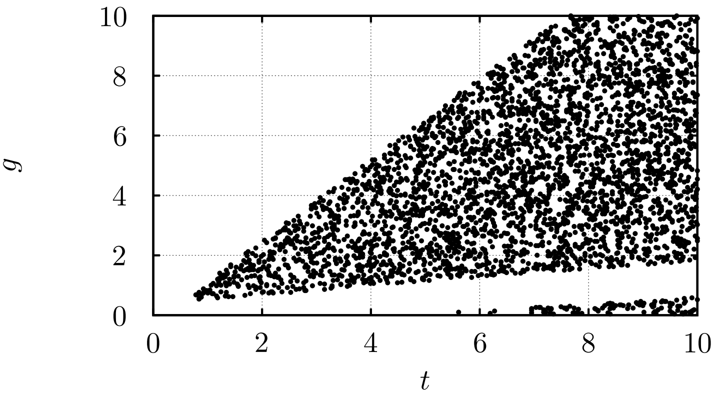


Figure 5.2: Feasibility of the optimization problem (4.10) (i. e., the joint optimization of controller and scalar transceivers) for $10\log_{10}(\mathrm{tr}\,[\boldsymbol{C_q}]^{-1}\,P_{\mathrm{Tx},1}) = 32$ and $10\log_{10}(\mathrm{tr}\,[\boldsymbol{C_n}]^{-1}\,P_{\mathrm{Tx},2}) = 32$ and given values of t and g. A feasible pair (t, g) is denoted by a dot.

Finally, the monotonic optimization approach is used to determine the optimal diagonal transmit and receive matrices $\boldsymbol{T} = \mathrm{diag}\,[t_1, t_2]$ and $\boldsymbol{G} = \mathrm{diag}\,[g_1, g_2]$ with $t_1, t_2, g_1, g_2 > 0$, i. e., the search for the global optimum is performed in a four dimensional set. Nevertheless, Algorithm 4.2 from Section 4.3.4 can readily be applied. The initial set for the optimization is chosen according to Equation (5.18) with $\underline{t} = \underline{g} = 10^{-10}$ and $\overline{t} = \overline{g} = 10^{7}$ and the desired relative accuracy of the result is $\varepsilon = 2 \cdot 10^{-2}$. Note that despite the fact that it is not necessary here, each scalar optimization variable is again mapped to the set $[0, 1]$ as in Chapter 4 (see Equations 4.39, 4.55 and 4.61). The subdivision using a bisection of sets selected by Algorithm 4.2 is then performed w.r.t. subsets of $[0, 1]^4$. With this setup, Algorithm 4.2 provides the solution shown in the fourth row of Table 5.1. Note that the value of the LQG cost is reduced by a factor of more than 12 compared to the case with scalar transceivers and a factor of more than 21 compared to the power constrained LQG controller without transmitter and receiver.

In order to get an impression about the control performance which could be achieved without the noisy communication channels, the last row provides the performance of an unconstrained LQG controller where it is allowed to use an infinitely large amount of transmit power for the transmission of observations and control signals or, equivalently, where no channel noise is present.

5.1.3 Impact of the Channel Assignment

Until now, the assignment of individual components of the input and output signals of the dynamical system to be controlled to the components of the communication channels, i. e., the scalar sub-channels, has been considered to be given and fixed. However, this assignment is only arbitrary and has no impact on the performance of the closed loop system if the variances of the noise sequences in all sub-channels are identical. Otherwise, the response of the dynamical system to the noise which is fed in at the system input, i. e., the channel noise of the control channel, and the impact of the noise in the observation channel on the state estimation error is generally different for each permutation of the available scalar sub-channels and the associated variances of the channel noise sequences.

Since the dynamical system to be controlled and the communication channels which are used for the information exchange between the system and the controller are distinct (physical or technical) entities, it is a degree of freedom for the system designer to decide which scalar component of the system output or the control signal is transmitted over which scalar sub-channel. This assignment of inputs and outputs of the dynamical system to inputs and outputs of the communication channels is formally described by a permutation matrix

$$\boldsymbol{F} = \sum_{k=1}^{K} \boldsymbol{e}_k^{(K)} \boldsymbol{e}_{i_k}^{(K),\mathrm{T}} \in \{0,1\}^{K \times K}, \tag{5.19}$$

where K is either equal to the dimension N_y of the system output or to N_u, i. e., the dimension of the system input. The permutation is determined by the numbers $i_k \in \{1, 2, \ldots, K\}$ for $k \in \{1, 2, \ldots, K\}$ with the property that $i_k \in \{1, 2, \ldots, K\} \setminus \{i_1, i_2, \ldots, i_{k-1}\}$. In the following we assume that different permutations can be chosen at the system input and output, which results in the following modification of the state and observation equation of the dynamical system to be controlled:

$$\begin{aligned} \boldsymbol{x}_{k+1} &= \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{B}\boldsymbol{F}_2\boldsymbol{G}^{\frac{1}{2}}\left(\boldsymbol{u}_k + \boldsymbol{n}_k\right) + \boldsymbol{w}_k, \\ \boldsymbol{y}_k &= \boldsymbol{F}_1\left(\boldsymbol{C}\boldsymbol{x}_k + \boldsymbol{v}_k\right), \end{aligned} \tag{5.20}$$

for $k \in \mathbb{N}_0$, and where $\boldsymbol{F}_1 \in \{0,1\}^{N_y \times N_y}$ and $\boldsymbol{F}_2 \in \{0,1\}^{N_u \times N_u}$ are permutation matrices (cf. Equation 5.19). Compared to the system model described in Section 5.1.1, we obtain the same scenario by replacing the system input matrix $\boldsymbol{B}$ with $\boldsymbol{B}\boldsymbol{F}_2$, the system output matrix $\boldsymbol{C}$ with $\boldsymbol{F}_1\boldsymbol{C}$ and the covariance matrix $\boldsymbol{C}_{\boldsymbol{v}}$ with $\boldsymbol{F}_1\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}_1^{\mathrm{T}}$. Additionally, due to the permutation of the system input, the weighting matrices $\boldsymbol{R}$ and $\boldsymbol{S}$ for the control signal in the LQG cost function (cf. Equation 5.5) have to be replaced with $\boldsymbol{F}_2^{\mathrm{T}}\boldsymbol{R}\boldsymbol{F}_2$ and $\boldsymbol{S}\boldsymbol{F}_2$, respectively. Thus, for fixed channel assignments, given by $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$, the optimal scalar or diagonal transceivers can be determined using the approach presented in Section 5.1.2.

In order to demonstrate the effect of changing the assignment of system inputs and outputs to the scalar sub-channels, Example 5.1.1 is revisited. In the following we will see that the comparison of the results presented there using no, scalar and diagonal transceivers should be interpreted carefully.

Example 5.1.2 The diagonal covariance matrices of the noise sequences in the observation and the control channel, i. e., $\boldsymbol{C}_{\boldsymbol{q}} = \operatorname{diag}[1, 0.5]$ and $\boldsymbol{C}_{\boldsymbol{n}} = \operatorname{diag}[0.7, 0.3]$, and the associated power constraints, i. e., transmit powers P_1 and P_2 which are limited according to

$10\log_{10}(\operatorname{tr}\left[\boldsymbol{C_q}\right]^{-1} P_1) \leq 32$ and $10\log_{10}\left(\operatorname{tr}\left[\boldsymbol{C_n}\right]^{-1} P_2\right) \leq 32$, from Example 5.1.1 are used for the solution of the optimization problem (5.6).[5] Since a permutation of the scalar components of the system input and output is taken into account, the state and observation equation for the constraints of problem (5.6) have to be replaced with Equation (5.20). Additionally, the weighting matrices $\boldsymbol{Q}$, $\boldsymbol{R}$ and $\boldsymbol{S}$ of the LQG cost function (see Equation 5.5) have to be replaced with $\boldsymbol{Q}$, i. e., no replacement, $\boldsymbol{F}_2^{\mathrm{T}}\boldsymbol{R}\boldsymbol{F}_2$ and $\boldsymbol{S}\boldsymbol{F}_2$, respectively. The values of $\boldsymbol{Q}$, $\boldsymbol{R}$ and $\boldsymbol{S}$ are provided by Example 3.1.1.

Since the dimensions of the input and the output signals of the dynamical system are $N_u = N_y = 2$, there exist only two permutations for the scalar components of these signals, i. e.,

$$\boldsymbol{F}_i \in \{\mathbf{I}_2, \boldsymbol{\Pi}\}, \tag{5.21}$$

for $i \in \{1, 2\}$, where the non-identity permutation matrix is given by

$$\boldsymbol{\Pi} = \begin{bmatrix} & 1 \\ 1 & \end{bmatrix}. \tag{5.22}$$

Table 5.2 shows the resulting values of the LQG cost and the associated transmit powers for the observation and the control channel using different permutations of the input and output signals of the dynamical system. Additionally, the values of the optimizing scalar and diagonal transmitters and receivers are provided. For the sake completeness, results which have already been presented in Table 5.1, i. e., for $\boldsymbol{F}_1 = \boldsymbol{F}_2 = \mathbf{I}_2$, are also shown in Table 5.2.

Note that the different channel assignments have a significant impact on the value of the cost function as well as on the transmit powers for the unconstrained, standard LQG controller without any transceiver (i. e., with $\boldsymbol{T} = \boldsymbol{G} = \mathbf{I}_2$), which are shown in the first row of Table 5.2. The qualitative results for this scenario carry over to the case when the transmit power constraints are included in the controller optimization, provided in the second row of Table 5.2. Since at least one power constraint is violated by the LQG controller for the unconstrained case irrespective of the channel assignment, the value of the cost function increases if it is required that both constraints are satisfied. Interestingly, the power constrained controller optimization becomes infeasible if the assignment for the observation channel is switched, i. e., if $\boldsymbol{F}_1 = \boldsymbol{\Pi}$ and $\boldsymbol{F}_2 = \mathbf{I}_2$. On the other hand, the power constraint for the control channel is inactive if only the assignment for the control channel is switched or if both $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ are equal to $\boldsymbol{\Pi}$. As in Chapter 3, we observe that the introduction of power constraints without the degrees of freedom which are offered by transmitters and receivers at the output and the input of the dynamical system to be controlled lead to unsatisfactory results.

The third row of Table 5.2 shows the optimal values of the LQG cost function, the associated transmit powers and the optimizing values of the scalar transceivers which are obtained using the approach presented in Chapter 4. For the determination of these results, Algorithm 4.2 has been applied with a relative desired accuracy of $\varepsilon = 10^{-2}$. The search interval for the transmit and receive scaling factors t and g has been set to $[10^{-10}, 10^7]$ and mapped into the interval $[0, 1]$ (see Section 4.3.4). The subdivision rule is again a bisection of sets in the image space of this mapping (cf. Equation 4.61). For the obtained results, all transmit power constraints hold with equality for each possible channel assignment, but the value of the cost function varies significantly, i. e.,

[5] The constraints for the SNRs of the individual communication channels translate to the power constraints $P_1 \leq 2377$ and $P_2 \leq 1585$, where the numerical values are rounded to four significant digits.

		$\boldsymbol{F}_1 = \boldsymbol{F}_2 = \mathbf{I}_2$	$\boldsymbol{F}_1 = \boldsymbol{\Pi}$, $\boldsymbol{F}_2 = \mathbf{I}_2$	$\boldsymbol{F}_1 = \mathbf{I}_2$, $\boldsymbol{F}_2 = \boldsymbol{\Pi}$	$\boldsymbol{F}_1 = \boldsymbol{F}_2 = \boldsymbol{\Pi}$
no constraints	LQG cost	954.7	1032	622.0	675.7
	P_1	5020	5407	3247	3505
	P_2	1972	2149	1317	1441
no Tx/Rx	LQG cost	1513	∞	656.2	736.9
	P_1	2377	—	2377	2377
	P_2	1585	—	1337	1459
scalar	LQG cost	895.6	2659	131.2	173.2
	P_1	2377	2377	2377	2377
	P_2	1585	1585	1585	1585
	$\boldsymbol{T}$	$0.8347 \cdot \mathbf{I}_2$	$4.802 \cdot \mathbf{I}_2$	$0.1216 \cdot \mathbf{I}_2$	$0.1454 \cdot \mathbf{I}_2$
	$\boldsymbol{G}$	$0.5401 \cdot \mathbf{I}_2$	$0.06583 \cdot \mathbf{I}_2$	$0.09930 \cdot \mathbf{I}_2$	$0.1281 \cdot \mathbf{I}_2$
diagonal	LQG cost	71.36	70.47	72.89	72.05
	P_1	2377	2377	2377	2377
	P_2	1585	1585	1585	1585
	$\boldsymbol{T}$	$t_1 = 0.09930$ $t_2 = 0.06334$	$t_1 = 0.04597$ $t_2 = 0.1416$	$t_1 = 0.09871$ $t_2 = 0.06334$	$t_1 = 0.04543$ $t_2 = 0.1416$
	$\boldsymbol{G}$	$g_1 = 0.005894$ $g_2 = 0.1551$	$g_1 = 0.005894$ $g_2 = 0.1532$	$g_1 = 0.1082$ $g_2 = 0.009364$	$g_1 = 0.1088$ $g_2 = 0.009116$

Table 5.2: Impact of different assignments of system inputs and outputs to scalar sub-channels on the LQG cost function and the transmit powers of the observation and control channel.

the maximal cost of 2659 is more than 20 times larger than the minimal value of 131.2 which is obtained for $\boldsymbol{F}_1 = \mathbf{I}_2$ and $\boldsymbol{F}_2 = \boldsymbol{\Pi}$.

Finally, the last row of Table 5.2 contains the results of the power constrained LQG problem where the optimal diagonal transmit and receive matrices $\boldsymbol{T} = \operatorname{diag}[t_1, t_2]$ and $\boldsymbol{G} = \operatorname{diag}[g_1, g_2]$ are applied. Due to the slower convergence compared to the scalar case (recall that 4 instead of 2 parameters have to be optimized), the relative desired accuracy is now chosen to be $\varepsilon = 2 \cdot 10^{-2}$ and the search interval for each scalar optimization variable is again $[10^{-10}, 10^{7}]$, together with the mapping to $[0, 1]$ and the bisection subdivision rule. As expected, the value of the cost function can be further decreased while the power constraints hold with equality.

In contrast to Example 5.1.1 which presented only the results of the first column of Table 5.2 above, the gains which can be achieved by the optimization of diagonal instead of scalar transceivers are less impressive. Comparing the best performance of each approach which takes into account the power constraints, the LQG cost without transceivers can be reduced by a factor of approximately 9.3 and the cost using scalar transceivers by a factor of approximately 1.8 if the optimal diagonal transmitters and receivers are applied. Nevertheless, for the optimal diagonal transceivers, a remarkable observation is that the minimal value of the cost function differs from the maximal one by less that 4%. This suggests that the diagonal transceivers are better suited for scalar sub-channels with unequal noise variances. It is however important to remember that not only the channel noise variances determine the behavior of the closed loop system but also the parameters of the dynamical system to be controlled. Consequently, it is not clear if the results presented using one specific example can be generalized to other scenarios.

The example above demonstrates that if more than one scalar sub-channel is available for the transmission of observations and control signals, respectively, in the closed control loop, it becomes important how the components of input and output signals of the dynamical system to be controlled are assigned to the available scalar sub-channels of the control and the observation channel, respectively.[6] The following example demonstrates that the optimal assignment does not only depend on the properties of the dynamical system and the communication channels but also on the operating point, i. e., the values of the available transmit powers $P_{\mathrm{Tx},1}$ and $P_{\mathrm{Tx},2}$.

Example 5.1.3 The same scenario as in Example 5.1.2 is considered here, i. e., the dynamical system to be controlled and the communication channels are described by the same parameters. In the following, we solve the problem of minimizing the weighted sum $\rho P_1 + (1-\rho)P_2$, $\rho \in [0,1]$, of the transmit powers P_1 and P_2 (cf. Equation 5.4) w.r.t. the LQG controller and scalar transceivers, i. e., using $\boldsymbol{T} = t\mathbf{I}_2$ and $\boldsymbol{G} = g\mathbf{I}_2$. This problem has already been investigated in Section 4.3.3.1. At this point, the goal is to demonstrate the impact of the channel assignment which is represented by the permutation matrices $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ (see Equations 5.20, 5.21, 5.22) on the resulting Pareto optimal transmit powers. For the computation of these powers, optimization problem (4.30) is solved for 1000 values of the weighting parameter ρ, where $\rho = (1+\theta)^{-1}$ and the values of θ are logarithmically spaced in the interval $[10^{-10}, 10^{10}]$ (see Example 4.3.3).

Figure 5.3 shows the resulting Pareto optimal values of the SNRs $\varphi_1 = \operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{q}}\right]^{-1} P_1$ and $\varphi_2 = \operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{n}}\right]^{-1} P_2$ for different choices of the permutation matrices $\boldsymbol{F}_i \in \{\mathbf{I}_2, \boldsymbol{\Pi}\}$, $i \in \{1,2\}$ (see Equation 5.21 and 5.22). For their computation, Algorithm 4.1 is initialized with the set $B = [0,1]$, where the optimization variable $\alpha \in \mathbb{R}_+$ is mapped to this set by the function given in Equation (4.39). For more details we refer to Section 4.3.3. The desired relative accuracy for the result of the optimization is chosen to be $\varepsilon = 10^{-3}$.

It can be observed that the assignment of inputs and outputs of the dynamical system to the available scalar sub-channels has a clear impact on the set of feasible transmit powers or SNRs, respectively, which is bounded by the Pareto optimal values. Especially for large values of φ_2, the associated minimal value of φ_1 can be reduced significantly. Note that this is not true if the roles of φ_1 and φ_2 are changed. In this case the channel assignment has a much smaller impact on the set of feasible SNRs. Another important observation is that there is no unique optimal channel assignment in the sense that the set of feasible SNRs is maximized. For example, over a large range of SNRs the choice $\boldsymbol{F}_1 = \boldsymbol{F}_2 = \boldsymbol{\Pi}$ (solid line) seems to be the best channel assignment. However, in a neighborhood of $10\log_{10}(\varphi_1) \approx 50$, there exist pairs of SNRs which are only feasible for $\boldsymbol{F}_1 = \mathbf{I}_2$ and $\boldsymbol{F}_2 = \boldsymbol{\Pi}$ (dash-dotted line).

Finally, recall the results presented in Table 5.2, i. e., the solutions of an LQG control problem with the constraints $10\log_{10}(\varphi_1) \leq 32$ and $10\log_{10}(\varphi_2) \leq 32$. Despite the fact that this operating point has the largest distance to the Pareto optimal SNRs which are obtained using the channel assignment $\boldsymbol{F}_1 = \boldsymbol{F}_2 = \boldsymbol{\Pi}$, the optimal value of the cost function is achieved with $\boldsymbol{F}_1 = \mathbf{I}_2$ and $\boldsymbol{F}_2 = \boldsymbol{\Pi}$. It can be seen that the optimal channel assignments which have been found for a certain design criterion do not necessarily carry over to optimization problems using a different criterion.

---

[6] At this point it becomes obvious that the concentration on diagonal transmitters and receivers is restrictive and that is in general possible to obtain better results for general transmit and receive filter matrices. Unfortunately, the monotonicity properties which have been used so far for the joint optimization of controller and transceivers are not present in the latter case.

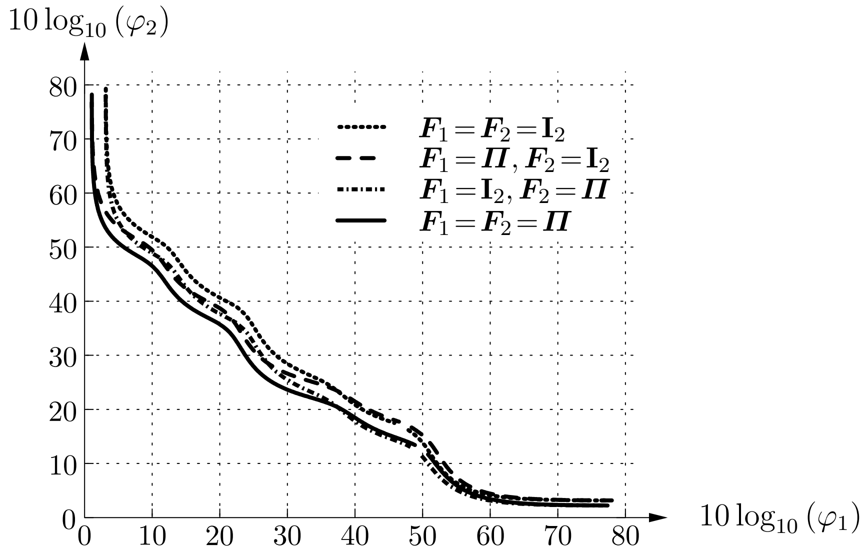


Figure 5.3: Pareto optimal values of SNRs φ_1 of the observation channel and φ_2 of the control channel obtained by the minimization of the weighted sum power $\rho P_1 + (1-\rho)P_2$, $\rho \in [0,1]$, using scalar transceivers.

The goal of this section is to demonstrate that the problem of the assignment of inputs and outputs of the dynamical system to be controlled to the available scalar sub-channels of the control and observation channel can not be neglected. Unfortunately, we are not able to provide a systematic approach to the determination of the optimal assignment. Consequently, the two examples which are shown in this section provide results where different assignments are simply fixed and the corresponding optimization problems are solved using the approaches presented in the preceding chapters and sections. Note further that the consideration of transmitters and receivers which consist of diagonal and permutation matrices is restrictive and it would be desirable to determine general optimal transceivers, i. e., matrices without structural constraints. However, a suitable approach to this problem is still missing.

5.2 Generalizations

The channel model with diagonal noise covariance matrices which has been introduced in Section 5.1.1 to describe a scenario with parallel and independent communication channels allowed for an extension of simple scaling factors at the input and output of these channels to the class of diagonal transmitters and receivers. Thus, the approach for the joint optimization of the LQG controller and the diagonal transceivers can be used for any communication system which is able to provide a set of independent scalar channels with mutually independent channel noise sequences. However, for the system model at hand, it is possible to consider some specific generalizations of the diagonal scenario investigated in Section 5.1.

5.2.1 Communication Channels With General Noise Covariance Matrices

In the following, an approach is proposed which reduces the case of communication channels with non-diagonal channel noise covariance matrices to the scenario presented in Section 5.1. In this context, it is not possible to optimize diagonal transceiver matrices based on the monotonicity properties of the associated optimization problem because these properties which are required by the branch and bound approach in the preceding sections are not present any more. The reason for this behavior can be found, e. g., in the DARE for the determination of the covariance matrix of the prediction (and estimation) error of the system state shown in Equation (5.10). If $\boldsymbol{C_q}$ and $\boldsymbol{C_n}$ are not diagonal, this equation becomes

$$\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} = \boldsymbol{A}\left(\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C_v} + \boldsymbol{T}^{\frac{1}{2}}\boldsymbol{C_q}\boldsymbol{T}^{\frac{1}{2}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\right)\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{C_w} + \boldsymbol{B}\boldsymbol{G}^{\frac{1}{2}}\boldsymbol{C_n}\boldsymbol{G}^{\frac{1}{2}}\boldsymbol{B}^{\mathrm{T}}. \tag{5.23}$$

For the diagonal matrices $\boldsymbol{T}$ and $\boldsymbol{G}$, it is generally not true that, e. g., for $\overline{\boldsymbol{T}}^{\frac{1}{2}} \geq \boldsymbol{T}^{\frac{1}{2}}$, it follows that $\overline{\boldsymbol{T}}^{\frac{1}{2}}\boldsymbol{C_q}\overline{\boldsymbol{T}}^{\frac{1}{2}} \geq \boldsymbol{T}^{\frac{1}{2}}\boldsymbol{C_q}\boldsymbol{T}^{\frac{1}{2}}$. A simple example illustrates this fact.

Example 5.2.1 Let $\boldsymbol{D} = \operatorname{diag}\left[d_k\right]_{k=1}^{2}$ be a diagonal matrix with $d_1, d_2 \in \left]0, 1\right[$ and

$$\boldsymbol{X} = \begin{bmatrix} 10.5 & -9.5 \\ -9.5 & 10.5 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & \\ & 20 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\frac{1}{\sqrt{2}}$$

be a positive definite matrix with the corresponding eigenvalue decomposition. Without loss of generality, we consider the definiteness of the matrix $\boldsymbol{X} - \boldsymbol{DXD}$ in order to investigate if $\overline{\boldsymbol{D}}\boldsymbol{X}\overline{\boldsymbol{D}} \geq \boldsymbol{DXD}$ for $\overline{\boldsymbol{D}} \geq \boldsymbol{D}$. With the specific choice of

$$\boldsymbol{D} = \begin{bmatrix} 0.75 & \\ & 0.25 \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{q} = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

we observe that

$$\boldsymbol{q}^{\mathrm{T}}\left(\boldsymbol{X} - \boldsymbol{DXD}\right)\boldsymbol{q} = -1,$$

which shows that $\boldsymbol{X} \not\geq \boldsymbol{DXD}$. In fact, the matrix $\boldsymbol{X} - \boldsymbol{DXD}$ is indefinite.

Since the partial ordering the of diagonal transmit and receive matrices $\boldsymbol{T}$ and $\boldsymbol{G}$ does not carry over to the diagonally scaled noise covariance matrices, i. e., $\boldsymbol{T}^{\frac{1}{2}}\boldsymbol{C_q}\boldsymbol{T}^{\frac{1}{2}}$ and $\boldsymbol{G}^{\frac{1}{2}}\boldsymbol{C_n}\boldsymbol{G}^{\frac{1}{2}}$, the monotonicity results from [100] can not be used to determine the bounds which are required by the monotonic optimization approach introduced earlier. Actually, it is not difficult to construct examples where the solutions of the corresponding DAREs do not inherit the partial ordering of $\boldsymbol{T}$ and $\boldsymbol{G}$. As a result, it is not clear how to jointly optimize an LQG controller together with diagonal transceivers and non-diagonal noise covariance matrices.

Having noticed that the applicability of the monotonic optimization approach is prevented by the fact that the covariance matrices of the channel noise sequences are not diagonal, and motivated by the results from, e. g., [110–112], we propose a heuristic design approach. Consider the system model shown in Figure 5.4 where it is *not* assumed that the covariance matrices of the channel noise sequences $(\boldsymbol{q}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{n}_k : k \in \mathbb{N}_0)$ are diagonal. For this case, it is proposed to chose a transmit and receive matrix, respectively, with the special structure

$$\tilde{\boldsymbol{T}} = \boldsymbol{U_q}\boldsymbol{T}^{-\frac{1}{2}} \qquad \text{and} \qquad \tilde{\boldsymbol{G}} = \boldsymbol{G}^{\frac{1}{2}}\boldsymbol{U}_{\boldsymbol{n}}^{\mathrm{T}}, \tag{5.24}$$

respectively, where the orthonormal matrices $\boldsymbol{U_q}$ and $\boldsymbol{U_n}$ are given by the eigenvalue decompositions of the noise covariance matrices

$$\boldsymbol{C_q} = \boldsymbol{U_q}\boldsymbol{\Lambda_q}\boldsymbol{U}_{\boldsymbol{q}}^{\mathrm{T}} \qquad \text{and} \qquad \boldsymbol{C_n} = \boldsymbol{U_n}\boldsymbol{\Lambda_n}\boldsymbol{U}_{\boldsymbol{n}}^{\mathrm{T}}, \tag{5.25}$$

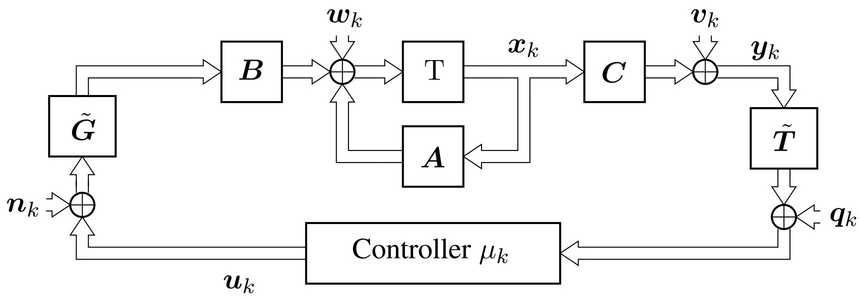


Figure 5.4: Model of the control loop which is closed over two channels with additive noise $\boldsymbol{q}_k$ and $\boldsymbol{n}_k$. At the system input and output, the receiver $\tilde{\boldsymbol{G}}$ and transmitter $\tilde{\boldsymbol{T}}$, respectively, is introduced.

with the diagonal matrices of eigenvalues $\boldsymbol{\Lambda_q}$ and $\boldsymbol{\Lambda_n}$. At this point, it is important to note that the eigenvalue decompositions are not unique, e. g., it is possible to introduce a permutation of the eigenvalues which are contained in $\boldsymbol{\Lambda_q}$ and $\boldsymbol{\Lambda_n}$ and the associated eigenvectors, i. e., the columns of $\boldsymbol{U_q}$ and $\boldsymbol{U_n}$. This permutation corresponds to the problem of the channel assignment discussed in Section 5.1.3.

Analogous to Section 5.1, $\boldsymbol{T}$ and $\boldsymbol{G}$ are diagonal matrices with positive diagonal elements, i. e.,

$$\boldsymbol{T} = \operatorname{diag}\,[t_i]_{i=1}^{N_y} \qquad \text{and} \qquad \boldsymbol{G} = \operatorname{diag}\,[g_i]_{i=1}^{N_u}\,, \tag{5.26}$$

with $t_i > 0$ and $g_i > 0$. Due to the invertibility of the transmitter and receiver matrices, the joint optimization problem of LQG controller and transceiver is almost identical to problem (5.6) and reads as (see also Equations 5.4 and 5.5)

$$\underset{\boldsymbol{T},\boldsymbol{G},\mu_0,\mu_1,\mu_2\ldots}{\text{minimize}} \lim_{N\to\infty}\frac{1}{N}\mathrm{E}\left[\boldsymbol{x}_N^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{x}_N+\sum_{n=0}^{N-1}\begin{bmatrix}\boldsymbol{x}_n\\ \tilde{\boldsymbol{G}}\boldsymbol{u}_n\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{Q} & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{R}\end{bmatrix}\begin{bmatrix}\boldsymbol{x}_n\\ \tilde{\boldsymbol{G}}\boldsymbol{u}_n\end{bmatrix}\right]+\operatorname{tr}\left[\boldsymbol{R}\tilde{\boldsymbol{G}}\boldsymbol{C_n}\tilde{\boldsymbol{G}}^{\mathrm{T}}\right] \tag{5.27}$$

$$\begin{aligned}
\text{subject to}\quad \boldsymbol{x}_{k+1} &= \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{B}\tilde{\boldsymbol{G}}(\boldsymbol{u}_k+\boldsymbol{n}_k)+\boldsymbol{w}_k, && k\in\mathbb{N}_0,\\
\boldsymbol{y}_k &= \boldsymbol{C}\boldsymbol{x}_k+\boldsymbol{v}_k, && k\in\mathbb{N}_0,\\
\boldsymbol{u}_k &= \mu_k(\mathcal{I}_k), && k\in\mathbb{N}_0,\\
\mathcal{I}_k &= \begin{cases}\{(\tilde{\boldsymbol{T}}\boldsymbol{y}_0+\boldsymbol{q}_0)\}, & k=0,\\ \{(\tilde{\boldsymbol{T}}\boldsymbol{y}_0+\boldsymbol{q}_0),\ldots,(\tilde{\boldsymbol{T}}\boldsymbol{y}_k+\boldsymbol{q}_k),\boldsymbol{u}_0,\ldots,\boldsymbol{u}_{k-1}\}, & k\in\mathbb{N},\end{cases}\\
\boldsymbol{U_q}^T\tilde{\boldsymbol{T}} &= \boldsymbol{T}^{-\frac{1}{2}} = \left(\operatorname{diag}\,[t_i]_{i=1}^{N_y}\right)^{-\frac{1}{2}} > \boldsymbol{0}_{N_y\times N_y},\\
\tilde{\boldsymbol{G}}\boldsymbol{U_n} &= \boldsymbol{G}^{\frac{1}{2}} = \left(\operatorname{diag}\,[g_i]_{i=1}^{N_u}\right)^{\frac{1}{2}} > \boldsymbol{0}_{N_u\times N_u},\\
\lim_{N\to\infty}\frac{1}{N}\,\mathrm{E}&\left[\sum_{n=0}^{N-1}\boldsymbol{x}_n^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}\tilde{\boldsymbol{T}}^{\mathrm{T}}\tilde{\boldsymbol{T}}\boldsymbol{C}\boldsymbol{x}_n+\boldsymbol{v}_n^{\mathrm{T}}\tilde{\boldsymbol{T}}^{\mathrm{T}}\tilde{\boldsymbol{T}}\boldsymbol{v}_n\right] \le P_{\mathrm{Tx},1},\\
\lim_{N\to\infty}\frac{1}{N}\,\mathrm{E}&\left[\sum_{n=0}^{N-1}\boldsymbol{u}_n^{\mathrm{T}}\boldsymbol{u}_n\right] \le P_{\mathrm{Tx},2}.
\end{aligned}$$

Paralleling the derivations in Section 5.1, we observe that for given matrices $\tilde{\boldsymbol{T}}$ and $\tilde{\boldsymbol{G}}$ the minimization of the Lagrangian associated with the problem above (where only the power constraints are dualized) w.r.t. the controller is an LQG problem and has the optimal value (cf. Equation 5.8)

$$\begin{aligned} L^*(\boldsymbol{T},\boldsymbol{G},\lambda_1,\lambda_2) &= \operatorname{tr}\left[\boldsymbol{K}\left(\boldsymbol{C}_{\boldsymbol{w}}+\boldsymbol{B}\tilde{\boldsymbol{G}}\boldsymbol{C}_{\boldsymbol{n}}\tilde{\boldsymbol{G}}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}}\right)\right]+\operatorname{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\right]+\lambda_1\operatorname{tr}\left[\tilde{\boldsymbol{T}}\boldsymbol{C}_{\boldsymbol{v}}\tilde{\boldsymbol{T}}^{\mathrm{T}}\right] \\ &\quad+\operatorname{tr}\left[\boldsymbol{R}\tilde{\boldsymbol{G}}\boldsymbol{C}_{\boldsymbol{n}}\tilde{\boldsymbol{G}}^{\mathrm{T}}\right]-\lambda_1P_{\mathrm{Tx},1}-\lambda_2P_{\mathrm{Tx},2} \\ &= \operatorname{tr}\left[\boldsymbol{K}\left(\boldsymbol{C}_{\boldsymbol{w}}+\boldsymbol{B}\boldsymbol{G}\boldsymbol{\Lambda}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\right)\right]+\operatorname{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\right]+\lambda_1\operatorname{tr}\left[\boldsymbol{T}^{-1}\boldsymbol{C}_{\boldsymbol{v}}\right] \\ &\quad+\operatorname{tr}\left[\boldsymbol{R}\boldsymbol{G}\boldsymbol{\Lambda}_{\boldsymbol{n}}\right]-\lambda_1P_{\mathrm{Tx},1}-\lambda_2P_{\mathrm{Tx},2}, \end{aligned} \tag{5.28}$$

where the fact has been used that $\tilde{\boldsymbol{G}}\boldsymbol{C}_{\boldsymbol{n}}\tilde{\boldsymbol{G}}^{\mathrm{T}}=\boldsymbol{G}^{\frac{1}{2}}\boldsymbol{U}_{\boldsymbol{n}}^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{U}_{\boldsymbol{n}}\boldsymbol{G}^{\frac{1}{2}}=\boldsymbol{G}\boldsymbol{\Lambda}_{\boldsymbol{n}}$ and that $\operatorname{tr}\left[\tilde{\boldsymbol{T}}\boldsymbol{C}_{\boldsymbol{v}}\tilde{\boldsymbol{T}}^{\mathrm{T}}\right]=\operatorname{tr}\left[\boldsymbol{U}_{\boldsymbol{n}}\boldsymbol{T}^{-\frac{1}{2}}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{T}^{-\frac{1}{2}}\boldsymbol{U}_{\boldsymbol{n}}^{\mathrm{T}}\right]=\operatorname{tr}\left[\boldsymbol{T}^{-1}\boldsymbol{C}_{\boldsymbol{v}}\right]$. Applying analogous identities, the matrices $\boldsymbol{K}$, $\boldsymbol{P}$ and $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}$ are determined by the DAREs (see Equations 5.9 and 5.10)

$$\begin{aligned} \boldsymbol{K} &= \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A}-\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B}+\boldsymbol{S}\right)\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B}+\boldsymbol{R}+\lambda_2\boldsymbol{G}^{-1}\right)^{-1}\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A}+\boldsymbol{S}^{\mathrm{T}}\right) \\ &\quad+\boldsymbol{Q}+\lambda_1\boldsymbol{C}^{\mathrm{T}}\boldsymbol{T}^{-1}\boldsymbol{C}, \end{aligned} \tag{5.29}$$

where $\boldsymbol{P}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A}-\boldsymbol{K}+\boldsymbol{Q}+\lambda_1\boldsymbol{C}^{\mathrm{T}}\boldsymbol{T}^{-1}\boldsymbol{C}$, and

$$\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}=\boldsymbol{A}\left(\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}-\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}+\boldsymbol{C}_{\boldsymbol{v}}+\boldsymbol{T}\boldsymbol{\Lambda}_{\boldsymbol{q}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\right)\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{C}_{\boldsymbol{w}}+\boldsymbol{B}\boldsymbol{G}\boldsymbol{\Lambda}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}, \tag{5.30}$$

which results in $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}=\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}-\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}+\boldsymbol{C}_{\boldsymbol{v}}+\boldsymbol{T}\boldsymbol{\Lambda}_{\boldsymbol{q}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}$. Comparing Equations (5.8), (5.9) and (5.10) from Section 5.1 with Equations (5.28), (5.29) and (5.30) and recalling that $\boldsymbol{T}$ and $\boldsymbol{G}$ are diagonal matrices with positive diagonal elements, it becomes clear that problem of finding the optimal values of $\boldsymbol{T}$ and $\boldsymbol{G}$ is identical to the problem considered in Section 5.1.

Remark: The basic idea for the specific choice of $\tilde{\boldsymbol{T}}$ and $\tilde{\boldsymbol{G}}$ in Equation (5.24) is to apply an orthonormal matrix to the transmit an received signal, respectively, which diagonalizes the covariance matrix of channel noise. The orthonormality ensures that the (sum) power of the transmit signals is not increased and that there is no amplification of the (sum) power of the channel noise at the receivers. The monotonic optimization framework can then be used to determine the optimal diagonal scaling matrices $\boldsymbol{T}$ and $\boldsymbol{G}$. Unfortunately, with the introduction of the orthonormal matrices for the joint optimization of the controller, the transmitter and the receiver, it is not possible any more to consider individual power constraints for the different scalar components of the transmitted vectors of observations and control signals. The reason is that even though these orthonormal transformations diagonalize the covariance matrices of the channel noise, this is not true for the summands in the DARE for the determination of the matrix $\boldsymbol{K}$ which are introduced by the consideration of individual power constraints. Only for the sum power constraints considered in the optimization problem (5.27), the special choice of $\tilde{\boldsymbol{T}}$ and $\tilde{\boldsymbol{G}}$ shown in Equation (5.24) preserves the monotonicity properties which are required by the global optimization approach.

The following example demonstrates that the proposed heuristic for the joint optimization of controller, transmitter and receiver has the potential to reduce the LQG cost function for the general case of non-diagonal covariance matrices of the channel noise. The reduction is in the same order of magnitude as it could be observed for the diagonal case in Example 5.1.1, but the impact of the channel assignment (cf. Section 5.1.2) is even more severe as in Example 5.1.2.

Example 5.2.2 Using the system and noise parameters from Example 3.1.1 (i. e., the covariance matrices of the channel noise sequences are *not* diagonal), solutions of the LQG control problem with different transceiver configurations are compared. Table 5.3 shows the corresponding results. The first row contains the values of the LQG cost function and the transmit powers of the observation and the control channel when no transmitter and receiver is used and the control loop is closed over the additive noise communication channels.

Example 5.1.2 demonstrated that a permutation of the scalar components of the input and output signals of the dynamical system to be controlled may have a significant impact on the performance of the closed loop control system (see Equation 5.20 for the corresponding modification of the system model using the permutation matrices $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$). Despite the fact that we do not assume here that the covariance matrices of the channel noise sequences $(\boldsymbol{q}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{n}_k : k \in \mathbb{N}_0)$ are diagonal, such a permutation is of course still possible. The second row of Table 5.3 thus shows the results for the optimal channel assignment which are obtained using $\boldsymbol{F}_1 = \mathbf{I}_2$ and $\boldsymbol{F}_2 = \boldsymbol{\Pi}$ (we refer again to Equations 5.20, 5.21 and 5.22 for a definition of these matrices).

<table>
<tr><th></th><th>LQG cost</th><th>P_1</th><th>P_2</th><th>$\boldsymbol{T}$</th><th>$\boldsymbol{G}$</th><th>λ_1</th><th>λ_2</th></tr>
<tr><td>no constraints</td><td>800.5</td><td>4189</td><td>1661</td><td>$\mathbf{I}_2$</td><td>$\mathbf{I}_2$</td><td>—</td><td>—</td></tr>
<tr><td>no constraints (opt. assign.)</td><td>413.6</td><td>2126</td><td>886.1</td><td>$\mathbf{I}_2$</td><td>$\mathbf{I}_2$</td><td>—</td><td>—</td></tr>
<tr><td>no Tx/Rx</td><td>∞</td><td>—</td><td>—</td><td>$\mathbf{I}_2$</td><td>$\mathbf{I}_2$</td><td>—</td><td>—</td></tr>
<tr><td>no Tx/Rx (opt. assign.)</td><td>415.4</td><td>1932</td><td>890.8</td><td>$\mathbf{I}_2$</td><td>$\mathbf{I}_2$</td><td>0.01981</td><td>0</td></tr>
<tr><td>scalar[7]</td><td>2580</td><td>1932</td><td>1288</td><td>$2.426 \cdot \mathbf{I}_2$</td><td>$2.006 \cdot \mathbf{I}_2$</td><td>9.402</td><td>23.92</td></tr>
<tr><td>scalar (opt. assign.)</td><td>115.5</td><td>1932</td><td>1288</td><td>$0.1228 \cdot \mathbf{I}_2$</td><td>$0.1185 \cdot \mathbf{I}_2$</td><td>0.4974</td><td>0.3243</td></tr>
<tr><td>diagonal</td><td>84.77</td><td>1932</td><td>1288</td><td>$\boldsymbol{T} = \mathrm{diag}\,[t_1, t_2]$
$t_1 = 0.1824$
$t_2 = 0.05785$</td><td>$\boldsymbol{G} = \mathrm{diag}\,[g_1, g_2]$
$g_1 = 0.02811$
$g_2 = 0.1155$</td><td>0.03812</td><td>0.01805</td></tr>
<tr><td>diagonal (opt. assign.)</td><td>79.33</td><td>1932</td><td>1288</td><td>$t_1 = 0.1852$
$t_2 = 0.06004$</td><td>$g_1 = 0.4066$
$g_2 = 0.005894$</td><td>0.03326</td><td>0.01058</td></tr>
<tr><td>ideal channels</td><td>28.84</td><td>∞</td><td>∞</td><td>—</td><td>—</td><td>—</td><td>—</td></tr>
</table>

Table 5.3: Comparison of power constrained control approaches with and without transmitters and receivers.

The next step is the addition of the power, or precisely SNR, constraints

$$10 \log_{10} \left(\mathrm{tr} \left[\boldsymbol{C}_{\boldsymbol{q}} \right]^{-1} P_1 \right) \leq 31.1 \qquad \text{and} \qquad 10 \log_{10} \left(\mathrm{tr} \left[\boldsymbol{C}_{\boldsymbol{n}} \right]^{-1} P_2 \right) \leq 31.1$$

to the problem of minimizing the LQG cost function. These constraints correspond to a limitation of the transmit powers to approximately $P_1 \leq 1932$ and $P_2 \leq 1288$, respectively. The addition of these constraints leads to an infeasible optimization problem when the degrees of freedom which are provided by a transmitter and receiver are missing and the communication channels are used

[7] In order to compare the values of the Lagrange multipliers, the scaled values $\nu_1 = t\lambda_1 \approx 22.81$ and $\nu_2 = g\lambda_2 \approx 47.99$ have to be considered.

as they are, i. e., with $\boldsymbol{F}_1 = \boldsymbol{F}_2 = \mathbf{I}_2$. The corresponding value of the cost function is denoted as ∞ in the third row of Table 5.3. With the optimal channel assignment, represented by $\boldsymbol{F}_1 = \mathbf{I}_2$ and $\boldsymbol{F}_2 = \boldsymbol{\Pi}$, the problem becomes feasible (see fourth row of Table 5.3), where the constraint for the transmit power P_2 remains inactive.

With the introduction of scaling factors at the input and the output of the dynamical system to be controlled, the power constrained joint optimization problem of the LQG controller and these scaling factors becomes feasible for $\boldsymbol{F}_1 = \boldsymbol{F}_2 = \mathbf{I}_2$ and is solved using the approach presented in Section 4.3.4. The relative accuracy for the branch and bound algorithm is chosen to be $\varepsilon = 10^{-2}$. The set for the search of the optimal scaling factors is $[10^{-10}, 10^7]^2$, which is again mapped to $[0, 1]^2$ to obtain the initial set B of Algorithm 4.2 (see Equation 4.61) and the bisection subdivision rule is applied. Note that with the optimal channel assignment using $\boldsymbol{F}_1 = \mathbf{I}_2$ and $\boldsymbol{F}_2 = \boldsymbol{\Pi}$, we observe a tremendous decrease of the LQG cost, which demonstrates that care must be taken for the assignment of the communication channels if scalar transceivers are applied to a system with multiple inputs and outputs.

Finally, with a relative accuracy of $\varepsilon = 2 \cdot 10^{-2}$, the transmit and receive matrices based on the diagonalization of the noise covariance matrices and diagonal transmitters and receivers $\boldsymbol{T}$ and $\boldsymbol{G}$ (see Equation 5.24) are optimized. The initial set B for the branch and bound algorithm is obtained by mapping $[10^{-10}, 10^5]^4$ to a subset of $[0, 1]^4$ analogous to Equations (4.39), (4.55) and (4.61), i. e., the minimal and maximal values for the search of the optimizing diagonal elements of $\boldsymbol{T}$ and $\boldsymbol{G}$ have been chosen to be 10^{-10} and 10^5. The eigenvalue decomposition of $\boldsymbol{C}_{\boldsymbol{q}}$ and $\boldsymbol{C}_{\boldsymbol{n}}$ which is computed for the actual transmitters $\tilde{\boldsymbol{T}}$ and $\tilde{\boldsymbol{G}}$ provides the diagonalized noise covariance matrices

$$\boldsymbol{\Lambda}_{\boldsymbol{q}} \approx \operatorname{diag}[0.4807, 1.019] \qquad \text{and} \qquad \boldsymbol{\Lambda}_{\boldsymbol{n}} \approx \operatorname{diag}[0.05279, 0.9472], \tag{5.31}$$

respectively. Note that the ascending order of the eigenvalues is chosen, but due to the non-uniqueness of the decomposition w.r.t. a permutation of the eigenvalues, a different order is possible. This degree of freedom is analogous to the channel assignment discussed in Section 5.1.3 for the case of diagonal covariance matrices of the channel noise sequences. Unfortunately, we are not able to provide a satisfactory answer to the question which permutation is optimal. Nevertheless, the ascending order is chosen arbitrarily at this point since we observed a relatively small dependence of the value of the LQG cost function on the channel assignment in Example 5.1.2, at least for the case of diagonal transceivers.

The seventh and eighth row of Table 5.3 provide a picture analogous to the results of Example 5.1.2: the LQG cost which can be achieved using scalar transceivers can be further decreased if diagonal transmit and receive matrices are applied, but the gain is moderate if the optimal channel assignments are used. Additionally, changing the order of the eigenvalues has a minor impact on the optimal value of the cost function which suggests that the diagonal transceivers have the potential to balance an unequal distribution of the eigenvalues of the noise covariance matrices even without the optimal choice of their order. Note that the optimal permutation of the eigenvalues is given by the matrices $\boldsymbol{F}_1 = \mathbf{I}_2$ and $\boldsymbol{F}_2 = \boldsymbol{\Pi}$, i. e., $\boldsymbol{\Lambda}_{\boldsymbol{q}}$ is given by Equation (5.31) whereas the order of the eigenvalues in $\boldsymbol{\Lambda}_{\boldsymbol{n}}$ is reversed, and that the associated cost of 79.33 is only determined with a relative accuracy of approximately $4.1 \cdot 10^{-2}$ which is obtained after 500000 iterations.

The last row shows the value of the LQG cost function which can be achieved if an infinitely large amount of transmit power is available or, equivalently, no channel noise is present.

As a final remark concerning the branch and bound approach for the determination of the optimal transmitters and receivers together with the optimal controller, note that a good solution candidate, represented by the upper bound which is determined by the monotonic optimization algorithm, can often be found relatively fast. Nevertheless, the certificate that it is actually a good solution, which is provided by the lower bound, needs a lot of computations, i. e., iterations of the algorithm. The following example illustrates this fact for the optimization of $\boldsymbol{T}$ and $\boldsymbol{G}$ from Example 5.2.2.

Example 5.2.3 In order to demonstrate the behavior of the monotonic optimization approach given by Algorithm 4.2, we revisit Example 5.2.2 and the problem of finding the optimal values of $\boldsymbol{T}$ and $\boldsymbol{G}$ for the transmitter and receiver proposed in Equation (5.24). The solution presented there resulted in the value 84.77 of the LQG cost function.

Figure 5.5 shows the evolution of the minimal value $\overline{I}^*$ of the upper bound (the cost function which is evaluated at specific values of $\boldsymbol{T}$ and $\boldsymbol{G}$) and the minimal value $\underline{I}^*$ of the lower bound (the solution of a relaxed optimization problem) with an increasing number of iterations of Algorithm 4.2. Note that the optimization is performed w.r.t. a four-dimensional set in this case. The set for the search of the optimizing matrices $\boldsymbol{T}$ and $\boldsymbol{G}$ is chosen according to Equation (5.18). The values of $\underline{t}$ and $\underline{g}$ and those of $\overline{t}$ and $\overline{g}$ are 10^{-10} and 10^5, respectively. The initial set B of Algorithm 4.2 is obtained by mapping this set to a subset of $[0, 1]^4$ analogous to Equations (4.39), (4.55) and (4.61).[8] One iteration corresponds to a bisection of the (mapped) set which provides the smallest value $\underline{I}^*$ of all currently available lower bounds, and the determination of the upper and lower bound for the cost function in the so obtained sets.

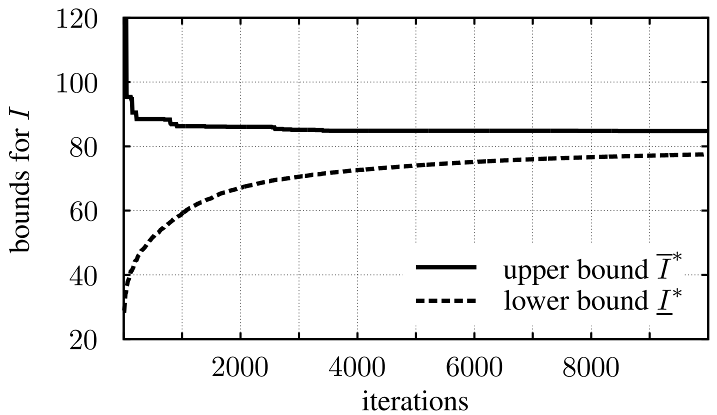


Figure 5.5: Smallest upper bound $\overline{I}^*$ and lower bound $\underline{I}^*$ over the number of iterations of Algorithm 4.2.

It can be observed that the value of the smallest upper bound $\overline{I}^*$ decreases relatively fast and provides a feasible result for controller, transmitter and receiver with a corresponding cost below 100 after 44 iterations and with a cost below 90 after 219 iterations.[9] Nevertheless, the confirmation that this is actually an acceptable value, which is given by the smallest lower bound $\underline{I}^*$, needs a significantly larger amount of iterations. The desired relative accuracy of $\varepsilon = 2 \cdot 10^{-2}$ and the

[8] Again, this mapping is not necessary at this point but is performed in order to be consistent with preceding examples.

[9] On a 2.2 GHz AMD Opteron CPU, 219 iterations can be executed in approximately one minute of CPU time using a Matlab implementation of the branch and bound algorithm.

associated value 84.77 of the LQG cost function is attained after 130650 iterations.[10] This number corresponds in the worst case[11] to 391950 evaluations of the optimization problem (5.27). Note that this number is equivalent to approximately 25 sampling points in each dimension (i. e., for t_1, t_2, g_1 and g_2) which are distributed over a search interval spanning 15 orders of magnitude. Since it is not obvious where these points should be placed by an approach which only samples the diagonal elements of $\boldsymbol{T}$ and $\boldsymbol{G}$ and does not use the information provided by the branch and bound approach, it is expected that the resulting solution will be generally worse compared to the monotonic optimization approach with the same computational complexity.

5.2.2 Communication Channels With Linear Distortions

For the second generalization of the optimization approach presented in Section 5.1, we take into account the general model of the communication channel which is described in Section 1.6 (see Equation 1.2 and Figure 1.4). Thus, besides non-diagonal covariance matrices of the channel noise sequences, linear distortions of the signals which are transmitted over the observation and the control channel are considered. These distortions are represented by the channel matrices $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$, respectively, where the dimension of these matrices is discussed below. Figure 5.6 depicts the model of the closed loop control system which extends the model used in Section 5.2.1 (see Figure 5.4) by the channel matrices.

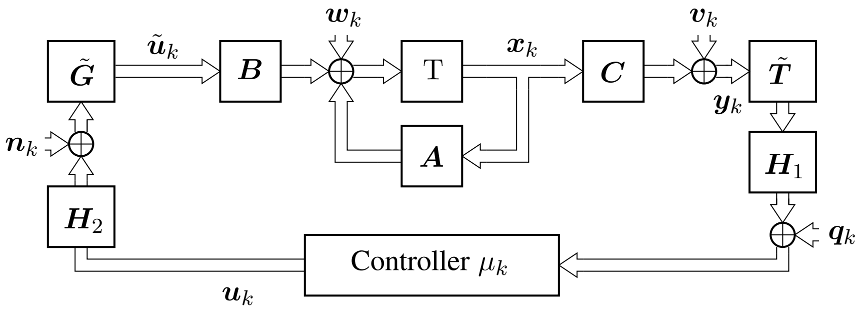


Figure 5.6: Model of the control loop which is closed over two channels with additive noise $\boldsymbol{q}_k$ and $\boldsymbol{n}_k$. At the system input and output, the receiver $\tilde{\boldsymbol{G}}$ and transmitter $\tilde{\boldsymbol{T}}$, respectively, is introduced. Besides the additive noise, the channels lead to linear distortions, represented by $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$.

All model assumptions which have been made in the preceding chapters, e. g., the independence of all noise sequences or constant parameters of the dynamical system, are used in the following and thus not restated at this point. Note that the channel matrices $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ are also assumed to be constant. Consequently, they fit into the LQG control framework which has been applied so far. Thus, the information at time index k which is available for the computation of control signals $\boldsymbol{u}_k$ reads as

$$\mathcal{I}_k = \begin{cases} \{(\boldsymbol{H}_1\tilde{\boldsymbol{T}}\boldsymbol{y}_0 + \boldsymbol{q}_0)\}, & k = 0, \\ \{(\boldsymbol{H}_1\tilde{\boldsymbol{T}}\boldsymbol{y}_0 + \boldsymbol{q}_0), \ldots, (\boldsymbol{H}_1\tilde{\boldsymbol{T}}\boldsymbol{y}_k + \boldsymbol{q}_k), \boldsymbol{u}_0, \ldots, \boldsymbol{u}_{k-1}\}, & k \in \mathbb{N}, \end{cases} \tag{5.32}$$

[10] The final value of the cost function is actually found after 46901 iterations, but the refinement of the lower bound for the optimal value needs the larger number of iterations.

[11] Assume that no subset can be excluded from the search. Then, due to the bisection in every iteration, it is necessary to compute one upper bound (the other one can be taken from the original, unpartitioned set) and two lower bounds for each of the two new sets. For the computation of each bound, the optimization problem (5.27) is solved one time.

and the actual input signal for the dynamical system to be controlled as

$$\tilde{\boldsymbol{u}}_k = \tilde{\boldsymbol{G}}\left(\boldsymbol{H}_2 \boldsymbol{u}_k + \boldsymbol{n}_k\right). \tag{5.33}$$

The goal is now to jointly optimize the controller, i. e., the mapping from the information $\mathcal{I}_k$ to the control signal $\boldsymbol{u}_k$, $k \in \mathbb{N}_0$, together with the transmit matrix $\tilde{\boldsymbol{T}}$ and the receive matrix $\tilde{\boldsymbol{G}}$. It has been mentioned earlier that this problem is hard to solve in general, but that the restriction to diagonal transmit and receive matrices allows for a systematic approach for channels without distortions, i. e., for $\boldsymbol{H}_1 = \mathbf{I}_{N_y}$ and $\boldsymbol{H}_2 = \mathbf{I}_{N_u}$, and diagonal channel noise covariance matrices. For non-diagonal covariance matrices, we proposed an approach in Section 5.2.1 which is based on the diagonalization of these matrices and thus reduces the more general case to the diagonal one. The same idea will be applied in the following, where we derive the necessary transformations at the transmitter and the receiver based on the solution of the optimal control problem for fixed matrices $\tilde{\boldsymbol{T}}$ and $\tilde{\boldsymbol{G}}$.

For the sake of simplicity, we assume in the following that $\boldsymbol{H}_1$ has more rows than columns, i. e., is tall, and that $\boldsymbol{H}_2$ has more columns than rows, i. e., is a wide matrix. This simplifies the subsequent derivations and is motivated by a scenario where the largest amount of computational complexity is represented by the controller and where the transmitter and receiver perform only memoryless linear operations. Thus, it is justified to assume that a larger dimension of the channel input and output signals can be handled at the controller and not at the input and the output of the dynamical system. Finally, we make the assumption that $\boldsymbol{H}_1$ has full column rank, $\boldsymbol{H}_2$ has full row rank, and that the number of columns of $\boldsymbol{H}_1$ is N_y whereas $\boldsymbol{H}_2$ has N_u rows in order to be compatible with the dimensions of the dynamical system.

With the introduction of the channel matrices $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$, the control problem (5.27) presented in Section 5.2.1 becomes

$$\begin{aligned}
&\underset{\boldsymbol{T},\boldsymbol{G},\mu_0,\mu_1,\ldots}{\text{minimize}} \lim_{N\to\infty} \frac{1}{N}\mathrm{E}\left[\boldsymbol{x}_N^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{x}_N + \sum_{n=0}^{N-1}\begin{bmatrix}\boldsymbol{x}_n\\ \tilde{\boldsymbol{G}}\boldsymbol{H}_2\boldsymbol{u}_n\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{Q} & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{R}\end{bmatrix}\begin{bmatrix}\boldsymbol{x}_n\\ \tilde{\boldsymbol{G}}\boldsymbol{H}_2\boldsymbol{u}_n\end{bmatrix}\right] + \mathrm{tr}\left[\boldsymbol{R}\tilde{\boldsymbol{G}}\boldsymbol{C}_{\boldsymbol{n}}\tilde{\boldsymbol{G}}^{\mathrm{T}}\right] \qquad (5.34)\\
&\text{subject to}\quad \boldsymbol{x}_{k+1} = \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{B}\tilde{\boldsymbol{G}}(\boldsymbol{H}_2\boldsymbol{u}_k + \boldsymbol{n}_k) + \boldsymbol{w}_k, \qquad k \in \mathbb{N}_0,\\
&\qquad \boldsymbol{y}_k = \boldsymbol{C}\boldsymbol{x}_k + \boldsymbol{v}_k, \qquad k \in \mathbb{N}_0,\\
&\qquad \boldsymbol{u}_k = \mu_k(\mathcal{I}_k), \qquad k \in \mathbb{N}_0,\\
&\qquad \mathcal{I}_k = \begin{cases}\{(\boldsymbol{H}_1\tilde{\boldsymbol{T}}\boldsymbol{y}_0 + \boldsymbol{q}_0)\}, & k = 0,\\ \{(\boldsymbol{H}_1\tilde{\boldsymbol{T}}\boldsymbol{y}_0 + \boldsymbol{q}_0), \ldots, (\boldsymbol{H}_1\tilde{\boldsymbol{T}}\boldsymbol{y}_k + \boldsymbol{q}_k), \boldsymbol{u}_0, \ldots, \boldsymbol{u}_{k-1}\}, & k \in \mathbb{N},\end{cases}\\
&\qquad \tilde{\boldsymbol{T}} = \boldsymbol{V}_{\boldsymbol{T}}\boldsymbol{T}^{-\frac{1}{2}}, \quad \boldsymbol{T} = \mathrm{diag}\,[t_i]_{i=1}^{N_y} > \mathbf{0}_{N_y\times N_y},\\
&\qquad \tilde{\boldsymbol{G}} = \boldsymbol{G}^{\frac{1}{2}}\boldsymbol{V}_{\boldsymbol{G}}, \quad \boldsymbol{G} = \mathrm{diag}\,[g_i]_{i=1}^{N_u} > \mathbf{0}_{N_u\times N_u},\\
&\qquad \lim_{N\to\infty}\frac{1}{N}\,\mathrm{E}\left[\sum_{n=0}^{N-1}\boldsymbol{x}_n^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}\tilde{\boldsymbol{T}}^{\mathrm{T}}\tilde{\boldsymbol{T}}\boldsymbol{C}\boldsymbol{x}_n + \boldsymbol{v}_n^{\mathrm{T}}\tilde{\boldsymbol{T}}^{\mathrm{T}}\tilde{\boldsymbol{T}}\boldsymbol{v}_n\right] \le P_{\mathrm{Tx},1},\\
&\qquad \lim_{N\to\infty}\frac{1}{N}\,\mathrm{E}\left[\sum_{n=0}^{N-1}\boldsymbol{u}_n^{\mathrm{T}}\boldsymbol{u}_n\right] \le P_{\mathrm{Tx},2}.
\end{aligned}$$

Note that the channel matrix $\boldsymbol{H}_2$ is included in the cost function of problem (5.34) since it affects the actual input signal of the dynamical system. Additionally, the state equation and the expression for the information $\mathcal{I}_k$ which is available to the controller at time index $k \in \mathbb{N}_0$ are modified according to Equations (5.33) and (5.32), respectively. Finally, the transmit matrix $\tilde{\boldsymbol{T}}$ and receive matrix $\tilde{\boldsymbol{G}}$ are chosen to have the special structure

$$\tilde{\boldsymbol{T}} = \boldsymbol{V}_{\boldsymbol{T}} \boldsymbol{T}^{-\frac{1}{2}} \qquad \text{and} \qquad \tilde{\boldsymbol{G}} = \boldsymbol{G}^{\frac{1}{2}} \boldsymbol{V}_{\boldsymbol{G}}, \tag{5.35}$$

respectively, where $\boldsymbol{T}$ and $\boldsymbol{G}$ are diagonal matrices with positive diagonal elements as in Equation (5.26). This choice is analogous to the transmitter and receiver given by Equation (5.24). However, the matrices $\boldsymbol{V}_{\boldsymbol{T}}$ and $\boldsymbol{V}_{\boldsymbol{G}}$ are not determined by the eigenvalue decompositions of the covariance matrices $\boldsymbol{C}_{\boldsymbol{q}}$ and $\boldsymbol{C}_{\boldsymbol{n}}$ alone as in Section 5.2.1. In the following, we propose a specific choice which effectively reduces the general channel model considered in this section to the one of Section 5.1.

For fixed transmit and receive matrices $\tilde{\boldsymbol{T}}$ and $\tilde{\boldsymbol{G}}$, i. e., without the optimization of the LQG cost function w.r.t. the associated diagonal matrices $\boldsymbol{T}$ and $\boldsymbol{G}$, the solution of optimization problem (5.34) can be determined as in the preceding sections. More precisely, the dual function of the problem (5.34) which has to be maximized w.r.t. the dual variables $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ reads as (cf. Equation 5.28)

$$\begin{aligned} L^*\left(\boldsymbol{T}, \boldsymbol{G}, \lambda_1, \lambda_2\right) = \operatorname{tr}\left[\boldsymbol{K}\left(\boldsymbol{C}_{\boldsymbol{w}} + \boldsymbol{B}\tilde{\boldsymbol{G}}\boldsymbol{C}_{\boldsymbol{n}}\tilde{\boldsymbol{G}}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}}\right)\right] + \operatorname{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\right] + \lambda_1 \operatorname{tr}\left[\tilde{\boldsymbol{T}}\boldsymbol{C}_{\boldsymbol{v}}\tilde{\boldsymbol{T}}^{\mathrm{T}}\right] \\ + \operatorname{tr}\left[\boldsymbol{R}\tilde{\boldsymbol{G}}\boldsymbol{C}_{\boldsymbol{n}}\tilde{\boldsymbol{G}}^{\mathrm{T}}\right] - \lambda_1 P_{\mathrm{Tx},1} - \lambda_2 P_{\mathrm{Tx},2}. \end{aligned} \tag{5.36}$$

As before, λ_1 and λ_2 are the dual variables associated with the power constraints $P_1 \leq P_{\mathrm{Tx},1}$ and $P_2 \leq P_{\mathrm{Tx},2}$, respectively. In contrast to Section 5.2.1, the DAREs for the determination of $\boldsymbol{K}$ and $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}$ are now given by[12]

$$\begin{aligned} \boldsymbol{K} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B}\tilde{\boldsymbol{G}}\boldsymbol{H}_2 + \boldsymbol{S}\tilde{\boldsymbol{G}}\boldsymbol{H}_2\right)\left(\boldsymbol{H}_2^{\mathrm{T}}\tilde{\boldsymbol{G}}^{\mathrm{T}}\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B} + \boldsymbol{R}\right)\tilde{\boldsymbol{G}}\boldsymbol{H}_2 + \lambda_2 \mathrm{I}_{N_H}\right)^{-1} \\ \times \left(\boldsymbol{H}_2^{\mathrm{T}}\tilde{\boldsymbol{G}}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} + \boldsymbol{H}_2^{\mathrm{T}}\tilde{\boldsymbol{G}}^{\mathrm{T}}\boldsymbol{S}^{\mathrm{T}}\right) + \boldsymbol{Q} + \lambda_1 \boldsymbol{C}^{\mathrm{T}}\tilde{\boldsymbol{T}}^{\mathrm{T}}\tilde{\boldsymbol{T}}\boldsymbol{C}, \end{aligned} \tag{5.37}$$

where $N_H \geq N_u$ is the number of columns of $\boldsymbol{H}_2 \in \mathbb{R}^{N_u \times N_H}$, and

$$\begin{aligned} \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} = \boldsymbol{A}\Big(\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\tilde{\boldsymbol{T}}^{\mathrm{T}}\boldsymbol{H}_1^{\mathrm{T}}\left(\boldsymbol{H}_1\tilde{\boldsymbol{T}}\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}}\right)\tilde{\boldsymbol{T}}^{\mathrm{T}}\boldsymbol{H}_1^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{q}}\right)^{-1}\boldsymbol{H}_1\tilde{\boldsymbol{T}}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\Big)\boldsymbol{A}^{\mathrm{T}} \\ + \boldsymbol{C}_{\boldsymbol{w}} + \boldsymbol{B}\tilde{\boldsymbol{G}}\boldsymbol{C}_{\boldsymbol{n}}\tilde{\boldsymbol{G}}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}}. \end{aligned} \tag{5.38}$$

Analogously, the matrix $\boldsymbol{P}$ which depends on the solution of Equation (5.37) reads as $\boldsymbol{P} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \boldsymbol{K} + \boldsymbol{Q} + \lambda_1 \boldsymbol{C}^{\mathrm{T}}\tilde{\boldsymbol{T}}^{\mathrm{T}}\tilde{\boldsymbol{T}}\boldsymbol{C}$, whereas the covariance matrix $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}$ of the state estimation error is given by $\boldsymbol{C}_{\tilde{\boldsymbol{x}}} = \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\tilde{\boldsymbol{T}}^{\mathrm{T}}\boldsymbol{H}_1^{\mathrm{T}}\left(\boldsymbol{H}_1\tilde{\boldsymbol{T}}\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}}\right)\tilde{\boldsymbol{T}}^{\mathrm{T}}\boldsymbol{H}_1^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{q}}\right)^{-1}\boldsymbol{H}_1\tilde{\boldsymbol{T}}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}$.

For the determination of the matrices $\boldsymbol{V}_{\boldsymbol{T}}$ and $\boldsymbol{V}_{\boldsymbol{G}}$ (see Equation 5.35) such that the general scenario described above is reduced to the diagonal case discussed in Section 5.1, Equations (5.37)

[12]Note that we assumed that $\boldsymbol{H}_2$ is a wide matrix. Thus, the inverse in Equation (5.37) exists only for $\lambda_2 > 0$, i. e., if the transmit power constraint of the control channel is active. This case can be assumed without loss of generality because otherwise the solution $\boldsymbol{K}$ (and consequently $\boldsymbol{P}$) would not depend on $\tilde{\boldsymbol{G}}$. In this case the value of the LQG cost function would be minimized for vanishing values of $\tilde{\boldsymbol{G}}$, i. e., by eliminating the effect of the control channel noise (cf. Equation 5.36).

and (5.38) are suitably rewritten. We begin with Equation (5.38), more precisely with the covariance matrix $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}$ of the state estimation error. With the assumption that $\boldsymbol{C}_{\boldsymbol{q}}$ is invertible[13], and by the application of the matrix inversion lemma (see, e. g., [113]), the equation which determines $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}$ also reads as

$$\begin{aligned}\boldsymbol{C}_{\tilde{\boldsymbol{x}}} &= \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\tilde{\boldsymbol{T}}^{\mathrm{T}}\boldsymbol{H}_1^{\mathrm{T}}\left(\boldsymbol{H}_1\tilde{\boldsymbol{T}}\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}+\boldsymbol{C}_{\boldsymbol{v}}\right)\tilde{\boldsymbol{T}}^{\mathrm{T}}\boldsymbol{H}_1^{\mathrm{T}}+\boldsymbol{C}_{\boldsymbol{q}}\right)^{-1}\boldsymbol{H}_1\tilde{\boldsymbol{T}}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} \\ &= \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\tilde{\boldsymbol{T}}^{\mathrm{T}}\boldsymbol{H}_1^{\mathrm{T}} \\ &\qquad\times\left(\boldsymbol{C}_{\boldsymbol{q}}^{-1}-\boldsymbol{C}_{\boldsymbol{q}}^{-1}\boldsymbol{H}_1\tilde{\boldsymbol{T}}\left(\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}+\boldsymbol{C}_{\boldsymbol{v}}\right)^{-1}+\tilde{\boldsymbol{T}}^{\mathrm{T}}\boldsymbol{H}_1^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{q}}^{-1}\boldsymbol{H}_1\tilde{\boldsymbol{T}}\right)^{-1}\tilde{\boldsymbol{T}}^{\mathrm{T}}\boldsymbol{H}_1^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{q}}^{-1}\right) \\ &\qquad\times\boldsymbol{H}_1\tilde{\boldsymbol{T}}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} \\ &= \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\Big(\tilde{\boldsymbol{T}}^{\mathrm{T}}\boldsymbol{H}_1^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{q}}^{-1}\boldsymbol{H}_1\tilde{\boldsymbol{T}}-\tilde{\boldsymbol{T}}^{\mathrm{T}}\boldsymbol{H}_1^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{q}}^{-1}\boldsymbol{H}_1\tilde{\boldsymbol{T}} \\ &\qquad\times\left(\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}+\boldsymbol{C}_{\boldsymbol{v}}\right)^{-1}+\tilde{\boldsymbol{T}}^{\mathrm{T}}\boldsymbol{H}_1^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{q}}^{-1}\boldsymbol{H}_1\tilde{\boldsymbol{T}}\right)^{-1}\tilde{\boldsymbol{T}}^{\mathrm{T}}\boldsymbol{H}_1^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{q}}^{-1}\boldsymbol{H}_1\tilde{\boldsymbol{T}}\Big)\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}.\end{aligned} \tag{5.39}$$

With a second application of the matrix inversion lemma and by inserting the result in Equation (5.38), we finally get

$$\begin{aligned}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} = \boldsymbol{A}\left(\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}-\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}+\boldsymbol{C}_{\boldsymbol{v}}+\left(\tilde{\boldsymbol{T}}^{\mathrm{T}}\boldsymbol{H}_1^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{q}}^{-1}\boldsymbol{H}_1\tilde{\boldsymbol{T}}\right)^{-1}\right)^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\right)\boldsymbol{A}^{\mathrm{T}} \\ +\boldsymbol{C}_{\boldsymbol{w}}+\boldsymbol{B}\tilde{\boldsymbol{G}}\boldsymbol{C}_{\boldsymbol{n}}\tilde{\boldsymbol{G}}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}},\end{aligned} \tag{5.40}$$

i. e., a DARE for the determination of the covariance matrix $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}$ where the effective channel noise has the covariance matrix $\left(\tilde{\boldsymbol{T}}^{\mathrm{T}}\boldsymbol{H}_1^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{q}}^{-1}\boldsymbol{H}_1\tilde{\boldsymbol{T}}\right)^{-1}$. Note that we assumed the invertibility of this matrix which must be ensured by the choice of $\tilde{\boldsymbol{T}}$. Thus, following the same approach as in Section 5.2.1, the monotonic optimization approach for the determination of the diagonal matrix $\boldsymbol{T}$ shown in Equation (5.35) can be applied if $\boldsymbol{V}_{\boldsymbol{T}}$ diagonalizes the matrix $\boldsymbol{H}_1^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{q}}^{-1}\boldsymbol{H}_1$.[14] Consequently, we use the orthonormal matrix

$$\boldsymbol{V}_{\boldsymbol{T}} = \boldsymbol{U}_{\boldsymbol{H}_1,\boldsymbol{q}}, \tag{5.41}$$

which is determined by the eigenvalue decomposition

$$\boldsymbol{H}_1^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{q}}^{-1}\boldsymbol{H}_1 = \boldsymbol{U}_{\boldsymbol{H}_1,\boldsymbol{q}}\boldsymbol{\Lambda}_{\boldsymbol{H}_1,\boldsymbol{q}}^{-1}\boldsymbol{U}_{\boldsymbol{H}_1,\boldsymbol{q}}^{\mathrm{T}}, \tag{5.42}$$

where the diagonal matrix $\boldsymbol{\Lambda}_{\boldsymbol{H}_1,\boldsymbol{q}}$ contains the *inverse* of the eigenvalues of $\boldsymbol{H}_1^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{q}}^{-1}\boldsymbol{H}_1$. Note that this notation has been chosen in order to be compatible with the results in Sections 5.1 and 5.2.1. Additionally, due to the non-uniqueness of the eigenvalue decomposition w.r.t., e. g., a permutation of the eigenvalues, the discussion from Section 5.1.3 concerning the channel assignment is also relevant at this point.

The situation for the determination of the matrix $\boldsymbol{V}_{\boldsymbol{G}}$ is more involved. The simple reason is that it has an influence on the LQG cost function through the state estimation error (see Equation

[13]This assumption has already been made earlier since a non-invertible covariance matrix would provide a noise free sub-channel such that an arbitrary (positive) power constraint could be fulfilled.

[14]Since the authors of [110–112] consider communication systems which minimize the mean square error of the transmitted data, they take into account the analogous matrix and diagonalize it by an eigenvalue decomposition.

5.40) and *additionally* due to the modified input matrix of the dynamical system which is not $\boldsymbol{B}$ but $\boldsymbol{B}\tilde{\boldsymbol{G}}$. This latter point is reflected by the DARE for the determination of the matrix $\boldsymbol{K}$ in Equation (5.37). Paralleling the derivations in Equations (5.39) and (5.40), this DARE can be rewritten as

$$\begin{aligned}\boldsymbol{K} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B}+\boldsymbol{S}\right)\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B}+\boldsymbol{R}+\left(\lambda_2^{-1}\tilde{\boldsymbol{G}}\boldsymbol{H}_2\boldsymbol{H}_2^{\mathrm{T}}\tilde{\boldsymbol{G}}^{\mathrm{T}}\right)^{-1}\right)^{-1}\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A}+\boldsymbol{S}^{\mathrm{T}}\right)\\ + \boldsymbol{Q} + \lambda_1\boldsymbol{C}^{\mathrm{T}}\tilde{\boldsymbol{T}}^{\mathrm{T}}\tilde{\boldsymbol{T}}\boldsymbol{C},\end{aligned} \tag{5.43}$$

where the assumption that $\lambda_2 > 0$ is used.

Now we are in the position to propose a matrix $\boldsymbol{V}_{\boldsymbol{G}}$ such that the monotonic optimization framework can be applied. The scenario which is considered in the present section is reduced to the one from Section 5.1 if $\boldsymbol{V}_{\boldsymbol{G}}$ jointly diagonalizes the channel noise covariance matrix $\boldsymbol{C}_{\boldsymbol{n}}$ and the product $\boldsymbol{H}_2\boldsymbol{H}_2^{\mathrm{T}}$ by congruence, i. e., such that both $\boldsymbol{V}_{\boldsymbol{G}}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{V}_{\boldsymbol{G}}^{\mathrm{T}}$ and $\boldsymbol{V}_{\boldsymbol{G}}\boldsymbol{H}_2\boldsymbol{H}_2^{\mathrm{T}}\boldsymbol{V}_{\boldsymbol{G}}^{\mathrm{T}}$ are diagonal. The conditions which ensure that this is possible can be found in, e. g., [114, p. 229] and hold in our case. The authors of [114] discuss this topic in detail and also provide a method for the diagonalization. However, we propose an alternative procedure which hopefully contributes to obtain some intuition about the relation between the properties of the channel matrix $\boldsymbol{H}_2$, the noise covariance matrix $\boldsymbol{C}_{\boldsymbol{n}}$, and the resulting matrix $\boldsymbol{V}_{\boldsymbol{G}}$. As a first step, the eigenvalue decomposition of $\boldsymbol{H}_2\boldsymbol{H}_2^{\mathrm{T}}$ is computed, i. e.,

$$\boldsymbol{H}_2\boldsymbol{H}_2^{\mathrm{T}} = \boldsymbol{U}_{\boldsymbol{H}_2}\boldsymbol{\Lambda}_{\boldsymbol{H}_2}\boldsymbol{U}_{\boldsymbol{H}_2}^{\mathrm{T}}, \tag{5.44}$$

where the diagonal matrix $\boldsymbol{\Lambda}_{\boldsymbol{H}_2}$ contains the eigenvalues of $\boldsymbol{H}_2\boldsymbol{H}_2^{\mathrm{T}}$ which are strictly positive by assumption. As a next step, the following eigenvalue decomposition is determined:

$$\boldsymbol{\Lambda}_{\boldsymbol{H}_2}^{-\frac{1}{2}}\boldsymbol{U}_{\boldsymbol{H}_2}^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{U}_{\boldsymbol{H}_2}\boldsymbol{\Lambda}_{\boldsymbol{H}_2}^{-\frac{1}{2}} = \boldsymbol{U}_{\boldsymbol{H}_2,\boldsymbol{n}}\boldsymbol{\Lambda}_{\boldsymbol{H}_2,\boldsymbol{n}}\boldsymbol{U}_{\boldsymbol{H}_2,\boldsymbol{n}}^{\mathrm{T}}, \tag{5.45}$$

where the diagonal matrix $\boldsymbol{\Lambda}_{\boldsymbol{H}_2,\boldsymbol{n}}$ again contains the strictly positive eigenvalues. Since $\boldsymbol{U}_{\boldsymbol{H}_2,\boldsymbol{n}}$ is orthonormal, it is easy to verify that the matrix

$$\boldsymbol{V}_{\boldsymbol{G}} = \boldsymbol{U}_{\boldsymbol{H}_2,\boldsymbol{n}}^{\mathrm{T}}\boldsymbol{\Lambda}_{\boldsymbol{H}_2}^{-\frac{1}{2}}\boldsymbol{U}_{\boldsymbol{H}_2}^{\mathrm{T}} \tag{5.46}$$

jointly diagonalizes $\boldsymbol{H}_2\boldsymbol{H}_2^{\mathrm{T}}$ and $\boldsymbol{C}_{\boldsymbol{n}}$ by congruence, i. e.,

$$\boldsymbol{V}_{\boldsymbol{G}}\boldsymbol{H}_2\boldsymbol{H}_2^{\mathrm{T}}\boldsymbol{V}_{\boldsymbol{G}}^{\mathrm{T}} = \mathbf{I}_{N_u} \qquad \text{and} \qquad \boldsymbol{V}_{\boldsymbol{G}}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{V}_{\boldsymbol{G}}^{\mathrm{T}} = \boldsymbol{\Lambda}_{\boldsymbol{H}_2,\boldsymbol{n}}. \tag{5.47}$$

With the proposed choice of $\boldsymbol{V}_{\boldsymbol{T}}$ (see Equations 5.41 and 5.42) and $\boldsymbol{V}_{\boldsymbol{G}}$ (see Equations 5.46, 5.44 and 5.45), and together with the matrices $\tilde{\boldsymbol{T}}$ and $\tilde{\boldsymbol{G}}$ given by Equation (5.35), the DAREs which have to be solved for the determination of the solution of optimization problem (5.34) read as

$$\begin{aligned}\boldsymbol{K} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B}+\boldsymbol{S}\right)\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B}+\boldsymbol{R}+\lambda_2\boldsymbol{G}^{-1}\right)^{-1}\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A}+\boldsymbol{S}^{\mathrm{T}}\right)\\ + \boldsymbol{Q} + \lambda_1\boldsymbol{C}^{\mathrm{T}}\boldsymbol{T}^{-1}\boldsymbol{C},\end{aligned} \tag{5.48}$$

and

$$\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} = \boldsymbol{A}\left(\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}+\boldsymbol{C}_{\boldsymbol{v}}+\boldsymbol{T}\boldsymbol{\Lambda}_{\boldsymbol{H}_1,\boldsymbol{q}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\right)\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{w}} + \boldsymbol{B}\boldsymbol{G}\boldsymbol{\Lambda}_{\boldsymbol{H}_2,\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}, \tag{5.49}$$

with the associated matrix $\boldsymbol{P} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \boldsymbol{K} + \boldsymbol{Q} + \lambda_1 \boldsymbol{C}^{\mathrm{T}}\boldsymbol{T}^{-1}\boldsymbol{C}$ and the error covariance matrix $\boldsymbol{C}_{\tilde{\boldsymbol{x}}} = \boldsymbol{C}^{\mathrm{p}}_{\tilde{\boldsymbol{x}}} - \boldsymbol{C}^{\mathrm{p}}_{\tilde{\boldsymbol{x}}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}^{\mathrm{p}}_{\tilde{\boldsymbol{x}}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{T}\boldsymbol{\Lambda}_{\boldsymbol{H}_1,\boldsymbol{q}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}^{\mathrm{p}}_{\tilde{\boldsymbol{x}}}$. With these results, the dual function from Equation (5.36) which is associated with problem (5.34) is finally given by

$$\begin{aligned} L^*(\boldsymbol{T}, \boldsymbol{G}, \lambda_1, \lambda_2) = \operatorname{tr}\left[\boldsymbol{K}\left(\boldsymbol{C}_{\boldsymbol{w}} + \boldsymbol{B}\boldsymbol{G}\boldsymbol{\Lambda}_{\boldsymbol{H}_2,\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\right)\right] + \operatorname{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\right] + \lambda_1 \operatorname{tr}\left[\boldsymbol{T}^{-1}\boldsymbol{C}_{\boldsymbol{v}}\right] \\ + \operatorname{tr}\left[\boldsymbol{R}\boldsymbol{G}\boldsymbol{\Lambda}_{\boldsymbol{H}_2,\boldsymbol{n}}\right] - \lambda_1 P_{\mathrm{Tx},1} - \lambda_2 P_{\mathrm{Tx},2}. \end{aligned} \tag{5.50}$$

The problem described above has the same properties as the results shown Sections 5.1 and 5.2.1, i. e., in Equations (5.8), (5.9) and (5.10) or (5.28), (5.29) and (5.30). Thus, the same approach as before can be used for the determination of the optimizing diagonal matrices $\boldsymbol{T}$ and $\boldsymbol{G}$.

Note that the structure of the transmit and receive matrices $\tilde{\boldsymbol{T}}$ and $\tilde{\boldsymbol{G}}$ given by Equation (5.35) and the choice of $\boldsymbol{V}_{\boldsymbol{T}}$ and $\boldsymbol{V}_{\boldsymbol{G}}$ shown in Equations (5.41) and (5.46) can not be motivated by some optimality criterion. The reason for these specific decisions is that they preserve the property of monotonicity which can be used by the monotonic optimization approach introduced in Chapter 4. We have no answer to the question how the optimal transmit and receive matrices look like in general and how much loss in performance has to be accepted due to the restricted structure of the transmitters and the receivers which are considered in this thesis.

5.3 Feasibility and Achievable Transmit Powers

We conclude this section with the consideration of the feasibility question for the power constrained controller design with diagonal transmitters and receivers. An associated problem is the determination of the set of feasible transmit power constraints. For the scalar case, this problem has been treated in Sections 4.3.1 – 4.3.3. With fixed transmit and receive matrices $\boldsymbol{T}$ and $\boldsymbol{G}$ (or $\tilde{\boldsymbol{T}}$ and $\tilde{\boldsymbol{G}}$, respectively), the approach presented in Section 4.3.1 can be applied in order to determine if a pair of power constraints, given by the values $P_{\mathrm{Tx},1}$ and $P_{\mathrm{Tx},2}$ of the available transmit powers, can be fulfilled. Note that a slightly different formulation of the power constraints has been used in Section 4.3.1, but the general approach stays the same. The remaining problem is to identify if a pair of transmit and receive matrices exists such that the power constrained controller optimization is feasible.

In Section 4.3.3, Pareto optimal values of the transmit powers for the observation and the control channel have been determined. These values separate feasible from infeasible problems since no transmit powers can be realized which are smaller than P_1 for the observation channel and at the same time smaller than P_2 for the control channel if (P_1, P_2) is a Pareto optimal pair. On the other hand, all inequality constraints which allow for a power larger than P_1 for the observation channel and at the same time for a power larger than P_2 for the control channel can be fulfilled. In order to determine the set of feasible power constraints for the case of diagonal transmitters and receivers, we adapt the approach presented in Section 4.3.3.2 for the computation of Pareto optimal pairs of transmit powers.[15] Thus, the transmit power of one communication channel is minimized while the power of the other channel is required to be smaller than a given value. To this end, recall the optimization problem (4.40) for the determination of the minimal value of P_1 subject to the constraint that $P_2 \leq P_{\mathrm{Tx},2}$, where $P_{\mathrm{Tx},2} > 0$ is the available transmit power for the control channel. For the problem at hand, the powers P_1 and P_2 are given by Equation (5.4). Consequently, the

[15] Since the heuristic approaches proposed in Section 5.2 reduce to the optimization of diagonal transceivers for channels with diagonal noise covariance matrices, only the latter case will be considered in the following.

optimization problem to be solved reads as

$$
\begin{aligned}
\underset{\boldsymbol{T},\boldsymbol{G},\mu_0,\mu_1,\mu_2\ldots}{\text{minimize}} \quad & P_1 && (5.51)\\
\text{subject to} \quad & \boldsymbol{x}_{k+1} = \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{B}\boldsymbol{G}^{\frac{1}{2}}(\boldsymbol{u}_k + \boldsymbol{n}_k) + \boldsymbol{w}_k, && k \in \mathbb{N}_0,\\
& \boldsymbol{y}_k = \boldsymbol{C}\boldsymbol{x}_k + \boldsymbol{v}_k, && k \in \mathbb{N}_0,\\
& \boldsymbol{u}_k = \mu_k(\mathcal{I}_k), && k \in \mathbb{N}_0,\\
& \mathcal{I}_k = \begin{cases} \{(\boldsymbol{T}^{-\frac{1}{2}}\boldsymbol{y}_0 + \boldsymbol{q}_0)\}, & k = 0,\\ \{(\boldsymbol{T}^{-\frac{1}{2}}\boldsymbol{y}_0 + \boldsymbol{q}_0), \ldots, (\boldsymbol{T}^{-\frac{1}{2}}\boldsymbol{y}_k + \boldsymbol{q}_k), \boldsymbol{u}_0, \ldots, \boldsymbol{u}_{k-1}\}, & k \in \mathbb{N}, \end{cases}\\
& \boldsymbol{T} = \operatorname{diag}\,[t_i]_{i=1}^{N_y} > \boldsymbol{0}_{N_y \times N_y},\\
& \boldsymbol{G} = \operatorname{diag}\,[g_i]_{i=1}^{N_u} > \boldsymbol{0}_{N_u \times N_u},\\
& P_2 \leq P_{\mathrm{Tx},2}.
\end{aligned}
$$

In Section 4.3.3.2, a reformulation of the system model has been used to demonstrate that if the problem is feasible, the infimum of the transmit power of the observation channel subject to the constrained power of the control channel is achieved if the transmit scaling factor t^{-1} goes to zero and the receive scaling g grows towards infinity with a fixed ratio $\alpha = g^{-1}t$ (see also Figure 4.6 and Equation 4.28). Simply speaking, if the signal at the system input is amplified with an unbounded scaling factor while the output signal is attenuated by a vanishing scalar, the channel noise sequences $(\boldsymbol{q}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{n}_k : k \in \mathbb{N}_0)$ become the dominant driving disturbances whereas the impact of the process and observation noise $(\boldsymbol{w}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{v}_k : k \in \mathbb{N}_0)$ on the transmit powers becomes negligible. Thus, the infimal transmit powers are determined using the same system model as for the optimization problem (5.51) but with vanishing process and observation noise.

The same line of argumentation also holds for the case of diagonal transmit and receiver matrices, i. e., the solution of problem (5.51) provides optimizing matrices $\boldsymbol{T}$ and $\boldsymbol{G}$ with unbounded diagonal elements. In order to see this, consider the optimization problem (5.6) and the dual function L^* for fixed values of $\boldsymbol{T}$ and $\boldsymbol{G}$ which is given by Equation (5.8). For the problem (5.51), we obtain the analogous function which is given by

$$
L^*(\boldsymbol{T},\boldsymbol{G},\lambda) = \operatorname{tr}\left[\boldsymbol{K}\left(\boldsymbol{C}_{\boldsymbol{w}} + \boldsymbol{B}\boldsymbol{G}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\right)\right] + \operatorname{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\right] + \operatorname{tr}\left[\boldsymbol{T}^{-1}\boldsymbol{C}_{\boldsymbol{v}}\right] - \lambda P_{\mathrm{Tx},2}, \tag{5.52}
$$

where $\lambda \geq 0$ is the dual variable associated with the transmit power constraint. The matrix $\boldsymbol{K}$ is the stabilizing solution of the DARE (cf. Equation 5.9)

$$
\boldsymbol{K} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B}\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B} + \lambda\boldsymbol{G}^{-1}\right)^{-1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} + \boldsymbol{C}^{\mathrm{T}}\boldsymbol{T}^{-1}\boldsymbol{C}. \tag{5.53}
$$

With this solution, the matrix $\boldsymbol{P}$ is given by $\boldsymbol{P} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \boldsymbol{K} + \boldsymbol{C}^{\mathrm{T}}\boldsymbol{T}^{-1}\boldsymbol{C}$. The covariance matrix of the state estimation error $\boldsymbol{C}_{\tilde{\boldsymbol{x}}} = \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{T}\boldsymbol{C}_{\boldsymbol{q}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}$ depends on the stabilizing solution of the DARE (cf. Equation 5.10)

$$
\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} = \boldsymbol{A}\left(\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{T}\boldsymbol{C}_{\boldsymbol{q}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}\right)\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{w}} + \boldsymbol{B}\boldsymbol{G}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}. \tag{5.54}
$$

Note that a common scaling of $\boldsymbol{T}$ and $\boldsymbol{G}$ by a factor $c > 0$ leads to the inverse scaling of the solution $\boldsymbol{K}$ from Equation (5.53) and the associated matrix $\boldsymbol{P}$, i. e., $\boldsymbol{K}(c\boldsymbol{T}, c\boldsymbol{G}) = c^{-1}\boldsymbol{K}(\boldsymbol{T},\boldsymbol{G})$ and $\boldsymbol{P}(c\boldsymbol{T}, c\boldsymbol{G}) = c^{-1}\boldsymbol{P}(\boldsymbol{T},\boldsymbol{G})$. Thus, the evaluation of L^* at $c\boldsymbol{T}$ and $c\boldsymbol{G}$ results in

$$
\begin{aligned}
L^*(c\boldsymbol{T}, c\boldsymbol{G}, \lambda) &= \operatorname{tr}\left[c^{-1}\boldsymbol{K}\left(\boldsymbol{C}_{\boldsymbol{w}} + \boldsymbol{B}c\boldsymbol{G}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\right)\right] + \operatorname{tr}\left[c^{-1}\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\right] + \operatorname{tr}\left[c^{-1}\boldsymbol{T}^{-1}\boldsymbol{C}_{\boldsymbol{v}}\right] - \lambda P_{\mathrm{Tx},2}\\
&= \operatorname{tr}\left[\boldsymbol{K}\left(c^{-1}\boldsymbol{C}_{\boldsymbol{w}} + \boldsymbol{B}\boldsymbol{G}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\right)\right] + \operatorname{tr}\left[\boldsymbol{P}\underline{\boldsymbol{C}_{\tilde{\boldsymbol{x}}}}\right] + \operatorname{tr}\left[\boldsymbol{T}^{-1}c^{-1}\boldsymbol{C}_{\boldsymbol{v}}\right] - \lambda P_{\mathrm{Tx},2},
\end{aligned} \tag{5.55}
$$

where $\underline{\boldsymbol{C}_{\tilde{\boldsymbol{x}}}} = c^{-1}\boldsymbol{C}_{\tilde{\boldsymbol{x}}} = \underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}} - \underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\boldsymbol{C}^{\mathrm{T}}\big(\boldsymbol{C}\underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\boldsymbol{C}^{\mathrm{T}} + c^{-1}\boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{T}\boldsymbol{C}_{\boldsymbol{q}}\big)^{-1}\boldsymbol{C}\underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}$ which depends on

$$\underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}} = \boldsymbol{A}\Big(\underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}} - \underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\boldsymbol{C}^{\mathrm{T}}\big(\boldsymbol{C}\underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\boldsymbol{C}^{\mathrm{T}} + c^{-1}\boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{T}\boldsymbol{C}_{\boldsymbol{q}}\big)^{-1}\boldsymbol{C}\underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\Big)\boldsymbol{A}^{\mathrm{T}} + c^{-1}\boldsymbol{C}_{\boldsymbol{w}} + \boldsymbol{B}\boldsymbol{G}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}. \tag{5.56}$$

Together with the fact that a smaller noise variance results in a smaller state estimation error in Equation (5.56), it can be seen that the value of the dual function in Equation (5.55) decreases monotonically with increasing values of c for each $\boldsymbol{T}$, $\boldsymbol{G}$ and λ. In the limit for $c \to \infty$, Equation (5.55) is the dual function for the optimization problem (5.51) with vanishing observation and process noise. The corresponding optimal value is smaller than the one that can be achieved in any scenario where these noise sequences are non-zero. Thus, the infimum of problem (5.51), if it is feasible, is achieved for transmit and receive matrices $\boldsymbol{T}$ and $\boldsymbol{G}$ with unbounded diagonal elements.

One result of the discussion above is the unboundedness of the optimizing transmitter $\boldsymbol{T}$ and receiver $\boldsymbol{G}$ when observation and process noise is present. A second one is the fact that if this limiting case is considered or, equivalently, if $\boldsymbol{C}_{\boldsymbol{w}} = \boldsymbol{0}_{N_x \times N_x}$ and $\boldsymbol{C}_{\boldsymbol{v}} = \boldsymbol{0}_{N_y \times N_y}$, the corresponding optimization problem becomes invariant w.r.t. a common scaling of $\boldsymbol{T}$ and $\boldsymbol{G}$. Thus, in order to determine unique optimizing transmitters and receivers, it is necessary to keep an arbitrary diagonal element of either $\boldsymbol{T}$ or $\boldsymbol{G}$ fixed, e. g., by setting it to one.[16] Taking into account these results, the infimal value of the transmit power in the observation channel subject to a power constraint for the control channel is determined by the following optimization problem:

$$\underset{\boldsymbol{T},\boldsymbol{G},\mu_0,\mu_1,\mu_2\ldots}{\text{minimize}} \quad \lim_{N\to\infty} \frac{1}{N}\,\mathrm{E}\left[\sum_{n=0}^{N-1} \boldsymbol{x}_n^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{T}^{-1}\boldsymbol{C}\boldsymbol{x}_n\right] \tag{5.57}$$

$$\begin{aligned}
\text{subject to} \quad \boldsymbol{x}_{k+1} &= \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{B}\boldsymbol{G}^{\frac{1}{2}}(\boldsymbol{u}_k + \boldsymbol{n}_k), && k \in \mathbb{N}_0,\\
\boldsymbol{y}_k &= \boldsymbol{C}\boldsymbol{x}_k, && k \in \mathbb{N}_0,\\
\boldsymbol{u}_k &= \mu_k(\mathcal{I}_k), && k \in \mathbb{N}_0,\\
\mathcal{I}_k &= \begin{cases} \{(\boldsymbol{T}^{-\frac{1}{2}}\boldsymbol{y}_0 + \boldsymbol{q}_0)\}, & k = 0,\\ \{(\boldsymbol{T}^{-\frac{1}{2}}\boldsymbol{y}_0 + \boldsymbol{q}_0), \ldots, (\boldsymbol{T}^{-\frac{1}{2}}\boldsymbol{y}_k + \boldsymbol{q}_k), \boldsymbol{u}_0, \ldots, \boldsymbol{u}_{k-1}\}, & k \in \mathbb{N}, \end{cases}\\
\boldsymbol{T} &= \operatorname{diag}\,[t_i]_{i=1}^{N_y} > \boldsymbol{0}_{N_y \times N_y},\\
\boldsymbol{G} &= \operatorname{diag}\,[g_i]_{i=1}^{N_u} > \boldsymbol{0}_{N_u \times N_u},\\
t_1 &= 1,\\
\lim_{N\to\infty} \frac{1}{N}\,\mathrm{E}&\left[\sum_{n=0}^{N-1} \boldsymbol{u}_n^{\mathrm{T}}\boldsymbol{u}_n\right] \leq P_{\mathrm{Tx},2}.
\end{aligned}$$

The choice of setting $t_1 = 1$ is arbitrary, any other diagonal element could have been selected.[17] Note that this additional constraint reduces the dimensionality of the optimization w.r.t. $\boldsymbol{T}$ and $\boldsymbol{G}$ by one. This fact has already been observed in Section 4.3.3 where the optimization w.r.t. to the transmit and receive scaling factors t and g reduced to the determination of the optimal ratio of these factors.

[16]Setting a diagonal element of $\boldsymbol{T}$ or $\boldsymbol{G}$ to one corresponds to a normalization of all other diagonal elements w.r.t. the selected one. The remaining optimization is then w.r.t. these normalized values. For example, in Section 4.3.2, the scaling factor g has been set to one and the normalized value of t has been called α, see Equation (4.28).

[17]This is of course only true due to the assumption that $\boldsymbol{T}$ and $\boldsymbol{G}$ are invertible. Without this restriction, setting a diagonal element of those matrices to one excludes a solution where the respective value is zero.

For fixed transmit and receive matrices, the dual function of optimization problem (5.57) which is obtained by the dualization of the transmit power constraint is (cf. Equation 5.52)

$$L^*(\boldsymbol{T}, \boldsymbol{G}, \lambda) = \operatorname{tr}\left[\boldsymbol{K}\boldsymbol{B}\boldsymbol{G}\boldsymbol{C}_{\boldsymbol{n}}\boldsymbol{B}^{\mathrm{T}}\right] + \operatorname{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\right] - \lambda P_{\mathrm{Tx},2}, \tag{5.58}$$

where $\boldsymbol{K}$ and $\boldsymbol{P}$ are determined using Equation (5.53) and $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}$ is obtained with the solution of Equation (5.54) where the covariance matrices $\boldsymbol{C}_{\boldsymbol{v}}$ and $\boldsymbol{C}_{\boldsymbol{w}}$ are set to zero. Note that one diagonal element of $\boldsymbol{T}$ or $\boldsymbol{G}$, e. g., the first diagonal element of $\boldsymbol{T}$ for optimization problem (5.57), has been chosen to be fixed and is thus not an optimization variable. With the assumption that strong duality holds, the solution of optimization problem (5.57) is determined by maximizing the concave dual function L^* w.r.t. $\lambda \geq 0$. For the remaining optimization w.r.t. the positive diagonal elements of $\boldsymbol{T}$ and $\boldsymbol{G}$, the monotonic optimization approach which has been used in Section 5.1 can be readily applied since the optimization problem (5.57) and the resulting dual function L^* in Equation (5.58) have the required monotonicity properties. Thus, the lower bound which has been derived in Equations (5.14)–(5.16) is also applicable for the determination of the optimizing transmit and receive matrices $\boldsymbol{T}$ and $\boldsymbol{G}$ using the monotonic optimization framework.

Analogous to the results shown in Figures 3.4, 4.8 and 4.13, the following example presents a numerical evaluation of the optimization problem (5.57), i. e., Pareto optimal values of the transmit powers for the case of diagonal transmit and receive matrices are determined.

Example 5.3.1 For the computation of Pareto optimal pairs of transmit powers, the system and noise parameters from Example 3.1.1 are used. Since the covariance matrices $\boldsymbol{C}_{\boldsymbol{q}}$ and $\boldsymbol{C}_{\boldsymbol{n}}$ of the channel noise sequences are not diagonal, the heuristic optimization approach presented in Section 5.2.1 is applied, i. e., the transmit and receive matrix is composed of an orthonormal matrix, which effectively diagonalizes the covariance matrix of the channel noise, and a diagonal matrix $\boldsymbol{T}$ and $\boldsymbol{G}$, respectively (see Equations 5.24 and 5.25). Besides the controller, these diagonal matrices which act on the so obtained channel with diagonal noise covariance matrices are the variables to be optimized.

Figure 5.7 shows the SNRs which are obtained by solving the optimization problem (5.57), where 11 values of the available transmit power $P_{\mathrm{Tx},2}$ are selected according to $10\log_{10}\left(\operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{n}}\right]^{-1} P_{\mathrm{Tx},2}\right) \in \{0, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50\}$. For the determination of the optimizing transmit and receive matrices, Algorithm 4.2 is initialized with the set B which is obtained by mapping the set $[10^{-5}, 10^5]^3$ to a subset of $[0, 1]^3$ analogous to Equations (4.39), (4.55) and (4.61), i. e., the diagonal elements of those matrices which are not fixed (cf. optimization problem 5.57) are constrained to be larger than 10^{-5} and smaller than 10^5. Despite the fact that a larger set would be desirable for the determination of the global optimum, the reason for this choice is the very slow convergence of the monotonic optimization algorithm for the problem at hand which renders larger search spaces problematic. The subdivision rule for Algorithm 4.2 is again a bisection (in the image space of the mapping described above) of the largest side of the subsets of B which are generated during the branching process. Finally, the desired relative accuracy of the result is chosen to be $\varepsilon = 10^{-2}$.

In order to compare the Pareto optimal transmit powers or SNRs, respectively, with the case where transmitter and receiver are scaled identity matrices (see Section 4.3), the results presented in Figures 4.8 and 4.13 are shown as a reference (solid line and shaded area). Since we observed a strong dependence of the system performance and transmit powers on the channel assignment (cf. Section 5.1.3) for the case of scalar transceivers, Pareto optimal pairs of transmit powers or SNRs, respectively, are computed for all of the four possible channel assignments. The corresponding SNRs are depicted in Figure 5.7 by dashed lines. Note that the original assignment is not optimal for almost all operating points and that significant gains are observed for a large range of SNRs.

The solutions of optimization problem (5.57) for diagonal transceivers are denoted by $\times$ in Figure 5.7. For a verification of these results, the power constrained power minimization is additionally solved where the roles of P_1 and P_2 have been changed, i. e., the transmit power P_2 is minimized subject to the constraint $P_1 \leq P_{\mathrm{Tx},1}$. The corresponding results, where $P_{\mathrm{Tx},1}$ is chosen according to $10\log_{10}\left(\operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{q}}\right]^{-1} P_{\mathrm{Tx},1}\right) \in \{0, 5, 10, 15, 20, 25, 27.5, 30, 35, 40, 45, 50, 55\}$, are denoted by $\bullet$ in Figure 5.7. It can be seen that the region of feasible power or SNR constraints becomes larger for the optimal choice of diagonal transceivers compared to the case of scalar transceivers, but the gains are relatively small for medium to large values of the SNR $\varphi_2 = \operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{n}}\right]^{-1} P_2$ of the control channel. Note that we do not consider a permutation of the eigenvalues in $\boldsymbol{\Lambda}_{\boldsymbol{q}}$ and $\boldsymbol{\Lambda}_{\boldsymbol{n}}$ given by Equation (5.31) (see also Equation 5.25) which is analogous to the problem of the optimal channel assignment.

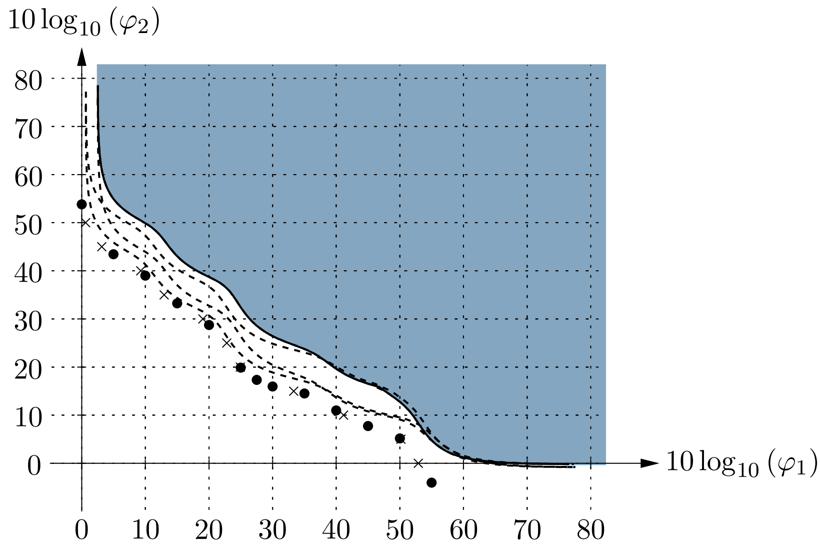


Figure 5.7: Pareto optimal values of SNRs determined by optimization problem (5.57) (denoted by $\times$). Results of an analogous optimization where the roles of P_1 and P_2 are exchanged are denoted by $\bullet$. The shaded region is taken from Figures 4.8 and 4.10. The SNRs φ_1 and φ_2 are defined in Equation (3.17).

For the computation of the optimizing diagonal matrices $\boldsymbol{T}$ and $\boldsymbol{G}$, problem (5.57) (or the analogous problem where the roles of P_1 and P_2 are swapped) is solved, i. e., the first diagonal element t_1 of $\boldsymbol{T}$ is set to the constant value one. Only for the minimization of P_1 subject to the constraint $10\log_{10}\left(\operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{n}}\right]^{-1} P_2\right) \leq 0$ the second diagonal element g_2 of $\boldsymbol{G}$ is set to the value one which results in a better relative accuracy after 500000 iterations of the branch and bound algorithm. Note that the desired relative accuracy of $\varepsilon = 10^{-2}$ is not achieved after 500000 iterations of Algorithm 4.2 for any result with $10\log_{10}(\varphi_1) \geq 45$, where $\varphi_1 = \operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{q}}\right]^{-1} P_1$. The other results are obtained with the required accuracy after 6653 to 246207 iterations. For the two solutions of optimization problem (5.57) with a power constraint of $10\log_{10}\left(\operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{n}}\right]^{-1} P_{\mathrm{Tx},2}\right) \in \{0, 5\}$, the relative accuracy after 500000 iterations is approximately 0.6, which means that the obtained solutions exceed the optimal value by a factor of 2.5 in the worst case. Unfortunately, although supported by these two solutions, the other mentioned results do not show an acceptable accuracy.

In Example 5.3.1 above, the set for the search of each diagonal element of $\boldsymbol{T}$ and $\boldsymbol{G}$ is chosen to be $[10^{-5}, 10^{5}]$. This limitation is due to slow and in some cases unsatisfactory convergence of the monotonic optimization algorithm, e. g., for large values of the transmit power of the observation channel. Nevertheless, the obtained results suggest that not all subspaces of the communication channels are used for the transmission of system outputs and control inputs in order to obtain Pareto optimal values of transmit powers. This observation is discussed in the following example.

Example 5.3.2 Consider the matrices $\boldsymbol{T}$ and $\boldsymbol{G}$ which are computed in Example 5.3.1 by Algorithm 4.2 for the optimization problem (5.57) with a power constraint according to $10\log_{10}\left(\operatorname{tr}\left[\boldsymbol{C_n}\right]^{-1} P_{\mathrm{Tx},2}\right) = 50$, the initial set $B = [10^{-5}, 10^{5}]^3$ and the required accuracy $\varepsilon = 10^{-2}$. Since the first diagonal element of $\boldsymbol{T}$ is set to the constant value one, the resulting matrices read as

$$\boldsymbol{T} \approx \operatorname{diag}\left[1, 10^{5}\right] \qquad \text{and} \qquad \boldsymbol{G} \approx \operatorname{diag}\left[1.722 \cdot 10^{-3},\, 4.547 \cdot 10^{-3}\right]. \tag{5.59}$$

The eigenvalue decompositions which are needed for the determination of the actual transmit and receive matrices $\tilde{\boldsymbol{T}}$ and $\tilde{\boldsymbol{G}}$ (see Equations 5.24 and 5.25) are again computed with an increasing order of the eigenvalues, i. e., the diagonalized noise covariance matrices read as

$$\boldsymbol{\Lambda_q} \approx \operatorname{diag}[0.4807, 1.019] \qquad \text{and} \qquad \boldsymbol{\Lambda_n} \approx \operatorname{diag}[0.05279, 0.9472]. \tag{5.60}$$

Noting that the transmitter uses the inverse of $\boldsymbol{T}^{\frac{1}{2}}$ for a diagonal scaling of transmit signals (see Equation 5.24), it can be seen that the power which corresponds to the larger eigenvalue of the covariance matrix $\boldsymbol{C_q}$ is attenuated significantly compared to the transmit power for the channel which is associated with the smaller eigenvalue. This is also true for the actual transmit powers since the (asymptotic) covariance matrix of the system output $\boldsymbol{y}_k$, $k \in \mathbb{N}_0$, of the optimization problem (5.57) is given by (note that the covariance matrix of the observation noise is $\boldsymbol{C_v} = \boldsymbol{0}_{2\times 2}$)

$$\boldsymbol{C_y} = \boldsymbol{C}\boldsymbol{C_x}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C_v} \approx \begin{bmatrix} 1.727 & 0.8041 \\ 0.8041 & 1.273 \end{bmatrix},$$

i. e., the variances of the scalar components of the system output, which are the signals to be transmitted, have the same order of magnitude.[18] Consequently, the covariance matrix of the signal which is transmitted over the effective channel with diagonalized noise covariance matrix reads as

$$\boldsymbol{T}^{-\frac{1}{2}}\boldsymbol{C_y}\boldsymbol{T}^{-\frac{1}{2}} \approx \begin{bmatrix} 1.727 & 0.002543 \\ 0.002543 & 1.273 \cdot 10^{-5} \end{bmatrix}.$$

This suggests that only one of the two dimensions which are provided by the observation channel is effectively used and that the second diagonal element of the transmit matrix $\boldsymbol{T}$ is only bounded[19] from above due to the restricted search interval of the monotonic optimization algorithm.

With the conjecture that the optimal transmitter only uses the subspace of the observation channel which is associated with the smaller eigenvalue of the noise covariance matrix $\boldsymbol{C_q}$, it is

---

[18]For the determination of the covariance matrix $\boldsymbol{C_y}$, the transmit and receiver matrices $\tilde{\boldsymbol{T}} = \boldsymbol{U_q}\boldsymbol{T}^{-\frac{1}{2}}$ and $\tilde{\boldsymbol{G}} = \boldsymbol{G}^{\frac{1}{2}}\boldsymbol{U}_{\boldsymbol{n}}^{\mathrm{T}}$ (cf. Equation 5.24), where $\boldsymbol{T}$ and $\boldsymbol{G}$ are given by Equation (5.59), of course have to be taken into account in the model of the closed loop system.

[19]Taking into account the normalization w.r.t. the first diagonal element of $\boldsymbol{T}$, the case that the second diagonal element of $\boldsymbol{T}$ goes to infinity means that the inverse scaling of the corresponding signal to be transmitted goes to zero, i. e., this component of the transmit signal is switched off.

straightforward to modify the system model used in optimization problem (5.57) accordingly. Due to the fixed value of the first diagonal element of $\boldsymbol{T}$, the transmitter $\tilde{\boldsymbol{T}}$ becomes (cf. Equation 5.24)

$$\tilde{\boldsymbol{T}} = \boldsymbol{U}_{\boldsymbol{q}} \boldsymbol{e}_1^{(2)} \boldsymbol{e}_1^{(2),\mathrm{T}}. \tag{5.61}$$

Note that this matrix is not invertible, which makes it necessary to formulate the optimization problem for the determination of Pareto optimal transmit powers analogous to problem (5.27). Since the transmitter in Equation (5.61) is constant, the solution of optimization problem (5.57) has only to be determined w.r.t. the controller and the diagonal matrix $\boldsymbol{G}$ at the receiver. To this end, Algorithm 4.2 is applied with the initial set B which is obtained by mapping the set $[10^{-10}, 10^{10}]^2$ to a subset of $[0,1]^2$ analogous to Equation (4.61) for the two variables of the diagonal elements of $\boldsymbol{G}$. The desired relative accuracy is $\varepsilon = 10^{-3}$ and the subdivision rule is again bisection. The resulting receiver matrix is obtained after 823 iterations and given by

$$\boldsymbol{G} \approx \operatorname{diag}\left[1.651 \cdot 10^{-3}, 4.599 \cdot 10^{-3}\right], \tag{5.62}$$

which is very close to the result determined by the full optimization of the transmitter and the receiver shown in Equation (5.59). Using the same power constraint, given by $10\log_{10}\left(\operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{n}}\right]^{-1} P_{\mathrm{Tx},2}\right) = 50$, the value of P_1 which is achieved by the transmit and receive matrices in Equation (5.59), i. e., with $\tilde{\boldsymbol{T}} = \boldsymbol{U}_{\boldsymbol{q}}\boldsymbol{T}^{-\frac{1}{2}}$ and $\tilde{\boldsymbol{G}} = \boldsymbol{G}^{\frac{1}{2}}\boldsymbol{U}_{\boldsymbol{n}}^{\mathrm{T}}$, is $P_1 \approx 1.7267$. The respective result for the fixed transmitter from Equation (5.61) and the receiver matrix $\boldsymbol{G}$ from Equation (5.62) is given by $P_1 \approx 1.7265$, where we have to keep in mind that the latter value is obtained using a larger set for the search of the optimal receiver and a desired relative accuracy of $\varepsilon = 10^{-3}$.

Having noticed that (almost) identical results are achieved for the computation of Pareto optimal transmit powers if one dimension of the observation channel is not used, the next step is to determine more of such transmit powers under the restriction that the transmitter $\tilde{\boldsymbol{T}}$ is given by Equation (5.61). To this end, 80 values of $P_{\mathrm{Tx},2}$ are sampled in the interval $10\log_{10}\left(\operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{n}}\right]^{-1} P_{\mathrm{Tx},2}\right) \in [14.098, 79]$ and the transmit powers P_1 of the observation channel for the corresponding power constraints are determined. The results are shown in Figure 5.8 by the dashed line. As before, the initial set for the branch and bound algorithm is chosen to be $[10^{-10}, 10^{10}]$, but the desired relative accuracy now is $\varepsilon = 10^{-2}$.

It can be observed that the results for the fixed transmitter which uses only one of the two dimensions of the observation channel (see Equation 5.61) coincide over a large range of SNRs with the results where this restriction is not imposed. Nevertheless, the SNR of the control channel can not be decreased below $10\log_{10}\left(\operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{n}}\right]^{-1} P_2\right) \approx 14.098$, which demonstrates that the chosen transmitter is not optimal for all possible transmit power constraints. However, an optimization problem analogous to the one investigated so far suggests similar properties of the transmitter and receiver for the case of smaller values of P_2, i. e., the power for the control channel.

Consider the problem of minimizing the transmit power P_2 with a constraint for the power P_1 w.r.t. the controller, the diagonal transmitter $\boldsymbol{T}$ and receiver $\boldsymbol{G}$.[20] The application of the monotonic optimization algorithm to this minimization problem with a power constraint according to $10\log_{10}\left(\operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{q}}\right]^{-1} P_{\mathrm{Tx},1}\right) = 40$ and where, in contrast to the case considered above, the first diagonal element of $\boldsymbol{G}$ is fixed to be one provides the value $P_1 \approx 12.543$. The associated diagonal

[20] For the minimization, an optimization problem analogous to problem (5.57) is formulated where the roles of P_1 and P_2 are exchanged and P_1 is constrained by the value of $P_{\mathrm{Tx},1}$. Additionally, since we consider the case of non-diagonal noise covariance matrices, the system model of Section 5.2.1 is used.

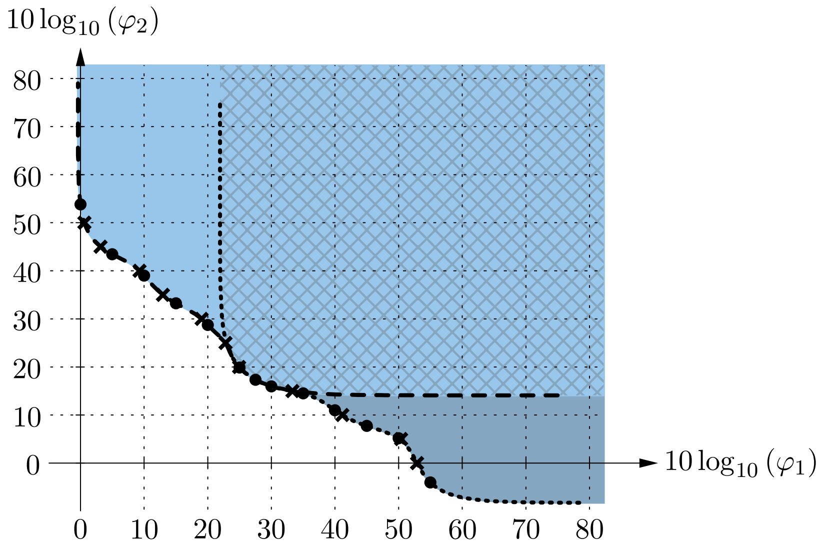


Figure 5.8: Pareto optimal values of SNRs φ_1 and φ_2 (see Equation 3.17) determined by switching off the second component of the transmit signal in the observation channel (dashed line) and the receive signal in the control channel (dotted line). For a comparison, the results shown in Figure 5.7 (no signal switched off, $\bullet$ and $\times$) are also shown.

transmit and receive matrices read as

$$\boldsymbol{T} \approx \operatorname{diag}\left[0.4884, 0.1730\right] \qquad \text{and} \qquad \boldsymbol{G} \approx \operatorname{diag}\left[1,\ 9.875 \cdot 10^{-4}\right].$$

The initial set for the monotonic optimization is chosen to be $B = [10^{-5}, 10^{5}]^3$ and the solution has the relative accuracy of $\varepsilon = 10^{-2}$. Increasing the desired accuracy to $\varepsilon = 10^{-3}$ slightly changes the result to

$$\boldsymbol{T} \approx \operatorname{diag}\left[0.4868, 0.1700\right] \qquad \text{and} \qquad \boldsymbol{G} \approx \operatorname{diag}\left[1,\ 1.321 \cdot 10^{-4}\right], \tag{5.63}$$

which leads to the transmit power $P_1 \approx 12.529$. This suggests that the second diagonal element of $\boldsymbol{G}$ is driven towards zero and thus only the subspace of the channel which is associated with the smaller eigenvalue of $\boldsymbol{C_n}$ is used for the transmission of control signals to the system input. In order to emphasize this suggestion, note that the covariance matrix of the signal which is applied to the input of the dynamical system to be controlled reads as

$$\boldsymbol{G}^{\frac{1}{2}}\boldsymbol{U}_{\boldsymbol{n}}^{\mathrm{T}}\left(\boldsymbol{L}\boldsymbol{C}_{\hat{\boldsymbol{x}}}\boldsymbol{L}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{n}}\right)\boldsymbol{U}_{\boldsymbol{n}}\boldsymbol{G}^{\frac{1}{2}} \approx \begin{bmatrix} 12.53 & 0.001171 \\ 0.001171 & 1.122 \cdot 10^{-7} \end{bmatrix} + \boldsymbol{G}^{\frac{1}{2}}\boldsymbol{\Lambda}_{\boldsymbol{n}}\boldsymbol{G}^{\frac{1}{2}},$$

where the transmitter and receiver with the diagonal matrices from Equation (5.63) are used and where $\boldsymbol{\Lambda}_{\boldsymbol{n}}$ is given in Equation (5.60). It can be seen that the controller puts much more weight on the first component of the signal which is fed into the system than on the second such that the latter one is dominated by the channel noise. Thus, it seems to be reasonable to switch off the second component of the system input by setting $\boldsymbol{G} = \operatorname{diag}\left[1, 0\right]$ which results in the fixed receiver (cf. Equation 5.24)

$$\tilde{\boldsymbol{G}} = \boldsymbol{e}_1^{(2)}\boldsymbol{e}_1^{(2),\mathrm{T}}\boldsymbol{U}_{\boldsymbol{n}}^{\mathrm{T}}. \tag{5.64}$$

The minimization of the transmit power P_2 subject to the constraint $P_1 \leq P_{\mathrm{Tx},1}$ now has to be performed w.r.t. to the controller and the diagonal matrix $\boldsymbol{T}$ at the transmitter. This optimization problem can again be solved using the branch and bound approach given by Algorithm 4.2. With the initial set $B = [10^{-10}, 10^{10}]^2$ and a desired relative accuracy of $\varepsilon = 10^{-3}$, the transmit power $P_1 \approx 12.527$ and the associated transmitter

$$\boldsymbol{T} \approx \operatorname{diag}[0.4873, 0.1700]$$

are obtained. This solution is almost identical to the respective transmit matrix in Equation (5.63) which has been computed without the assumption that the second diagonal element of $\boldsymbol{G}$ is zero.

For a comparison with previous results, the minimization of P_2 subject to the requirement that the transmit power P_1 does not exceed $P_{\text{Tx},1}$ is solved using the fixed transmitter given in Equation (5.64). For the power constraint, 95 values of $P_{\text{Tx},1}$ are sampled from the interval $10\log_{10}\left(\operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{q}}\right]^{-1} P_{\text{Tx},1}\right) \in [21.93, 79]$. The initial set of the branch and bound algorithm is again $B = [10^{-10}, 10^{10}]^2$ and the desired accuracy is chosen to be $\varepsilon = 10^{-2}$. The resulting transmit powers or SNRs, respectively, are depicted in Figure 5.8 by the dotted line. Note that SNRs below the value of $10\log_{10}\left(\operatorname{tr}\left[\boldsymbol{C}_{\boldsymbol{q}}\right]^{-1} P_{\text{Tx},1}\right) \approx 21.93$ are not feasible with the specific choice of the receiver from Equation (5.64).

Example 5.3.2 above demonstrates that the assumption of invertible transmit and receive matrices $\boldsymbol{T}$ and $\boldsymbol{G}$ (or $\tilde{\boldsymbol{T}}$ and $\tilde{\boldsymbol{G}}$, respectively) is too restrictive in general. It has already been mentioned in Section 5.1 that the additional power which is required to transmit, e. g., a two-dimensional instead of a scalar signal, and the additional channel noise which is fed into the closed control loop may lead to the fact that the optimal transceivers do not use all available degrees of freedom of the communication channels. In this case, it is optimal to allocate all available communication resources, i. e., transmit power, to a lower dimensional subspace of the channel. Unfortunately, a systematic answer to the question when this behavior can be observed is not available. The example above suggests that for very low transmit powers of the observation and the control channel, only the subspace which corresponds to the smallest eigenvalue of the respective noise covariance matrix is actually used for the transmission of observations or control signals. However, also the assignment of the scalar components of the control and observation channels to the individual inputs and outputs of the dynamical system to be controlled is of importance (see Section 5.1.3 and Example 5.2.2). This fact is not taken into account for any result presented in Examples 5.2.3, 5.3.1 and 5.3.2 where the assignment is determined by computing the eigenvalue decomposition of the covariance matrices of the channel noise sequences with an ascending order of the eigenvalues. The question if this is the optimal allocation remains open at this point.

5.4 Discussion

In Chapter 4, the transceivers have been chosen to be scaled identity matrices. The degrees of freedom provided by the scaling of transmit and receive signals have only impact on the transmit powers or, taking a different perspective, the variances of the channel noise sequences. While this is the only possibility to affect the transmit signals in scenarios with Single-Input Single-Output (SISO), i. e., scalar, communication channels, the Multiple-Input Multiple-Output (MIMO) channel models considered here allow for the application of general transmitter and receiver matrices at the output and the input of the dynamical system to be controlled. Consequently, within the class of linear memoryless transceiver schemes, the goal is to determine the optimal transmitter and receiver matrices and not only scaling factors.

A first step towards this direction has been taken in the present chapter with the optimization of diagonal transmitters and receivers when the channel noise covariance matrices are diagonal. For the general case of non-diagonal covariance matrices, a heuristic has been proposed which is based on the diagonalization of these matrices and motivated by solutions developed in the area of signal processing and communication, e. g., [110–112], for a pure communication setting without a closed control loop. Unfortunately, those approaches can not be adapted in a straightforward way to the closed loop control scenario. For a pure communication setting, the transmitter and the receiver are designed to minimize a given cost function, e. g., the mean square error, subject to a transmit power constraint. For the LQG control setting, the design goal for, e. g., a transmitter at the system output, is to minimize the state estimation error, which is one contribution to the LQG cost function (see, e. g., Appendix A6.3, Summary A6.2), subject to a power constraint. A first problem is that the power constraint is associated with the covariance matrix of the system state whereas the state estimation error is associated with the covariance matrix of the innovation sequence computed by the Kalman filter. These two matrices are related to each other by the solution of a DARE and a discrete Lyapunov equation (cf. Equations 3.6, 3.7 and 3.13 – 3.15), which makes the consideration of this relation challenging. The second problem is the common scenario that the dimension of the system state is larger than the dimension of the communication channel, which violates the assumptions of [110, 111]. Finally, the solutions presented there use transceivers which are based on a diagonalization of the covariance matrix of the transmit signal. However, in a closed control loop, the covariance matrix of the system state depends on the actual transmit and receive matrix and thus can not be diagonalized independently to the transceiver design.[21]

For the case of diagonal covariance matrices of the observation and control channel noise sequences, the approach for the joint optimization of controller and transceivers presented in Chapter 4 could be extended to diagonal transmit and receive matrices in a straightforward way. The reason is that the monotonicity properties required by the monotonic optimization algorithm still hold for diagonal transceivers in combination with diagonal noise covariance matrices. However, it has been demonstrated in Section 5.1.3 that the assumption of unequal variances of the scalar components of the channel noise sequences leads to the unsurprising fact that the assignment of the components of input and output signals of the dynamical system to be controlled to the scalar components of the communication channels has a significant impact on the closed loop behavior, e. g., the transmit powers or the LQG cost function. This problem is not covered by diagonal transmit and receive matrices. Unfortunately, an approach for the determination of the optimal permutation of the scalar components of the communication channels is not available. Nevertheless, Example 5.1.2 showed that the impact of the channel assignment on the closed loop performance can be compensated up to a certain degree by the optimal choice of the diagonal transmitter and receiver, which is not the case for the optimal scalar transceivers.

In Section 5.2, two generalizations of the optimization of diagonal transceivers to the case of non-diagonal channel noise covariances and to communication channels with linear distortions have been proposed. To summarize, both approaches aim at a diagonalization of the communication channels and the associated noise covariance matrix. Once this goal is achieved, the problem of the joint optimization of controller and diagonal transceivers can be solved using the approach presented in Section 5.1. This means that the joint design of the control and the communication system is separated at this point. As a first step, a set of independent scalar communication channels

[21] Note further that the sequences which are transmitted in the closed control loop are temporally correlated in general, which can not be handled by memoryless transceivers. If not the control but the estimation problem alone is considered, the results of [115] can be applied which are also derived in [116] using a different approach.

is provided by the diagonalization of the available, non-diagonal channels. The joint optimization of diagonal transceivers and the controller subject to power constraints is performed subsequently. However, we do not claim that this separation is optimal. Finally, note that the authors of [117] investigate the case of parallel Gaussian communication channels, i. e., the scenario considered in Section 5.1, where only the control channel is modeled as an additive noise channel and where the state of the dynamical system to be controlled is available for the design of the controller, or encoder and decoder, respectively. They derive bounds for the stabilizability of the system under power constraints and investigate a more general class of controller, transmitter and receiver than in this thesis. However, this requires the knowledge of the system state and no results beyond stabilizability are available.

The dependence of properties of the transmitted signals on the transceivers which are used for their transmission makes it hard to find solutions which are jointly optimal together with the controller. Even the simple case with scalar transceivers is not easy to solve. Note that the authors of [50] analyzed the impact of a matrix transmitter for packet drop channels (cf. Section 2.1.2). But despite the fact that the state estimation problem alone is considered, i. e., no closed loop system is investigated, no systematic design procedure could be derived due to the problem structure. Thus, the joint optimal design of controller and transceivers remains a challenging task.

6. Results for SISO Systems With One Communication Channel

Throughout this thesis, a general system model with Multiple-Input Multiple-Output (MIMO) Linear Time-Invariant (LTI) systems and Additive White Gaussian Noise (AWGN) communication channels has been investigated. Additionally, both links which connect the dynamical system to be controlled with the controller have been modeled as communication channels with transmit power or, effectively, Signal to Noise Ratio (SNR) constraints, which models a spatial separation of controller and dynamical system. The resulting optimization problems allowed for numerical solution approaches, but their properties prevented an analytical investigation, e. g., the region of feasible transmit powers could only be determined numerically. Nevertheless, the related literature on Networked Control Systems (NCSs), e. g., [29, 35, 46, 53, 59, 85–87, 106, 107, 118], provides analytical results, like the relation of minimal transmit power (or SNR) to the unstable eigenvalues of the system matrix $\boldsymbol{A}$.

All references mentioned above have in common that they consider Single-Input Single-Output (SISO) systems, i. e., dynamical systems with scalar input and output signals, and only one communication channel in the control loop. This allows for a concentration on the impact of the channel on the control system without the coupling of the two different channels.[1] A second advantage of the SISO model is that the only parameter which describes the properties of a channel noise sequence is its variance and not a covariance matrix. Thus, the channel noise can not have a different impact on different subspaces of the transmit signals since all these signals are of dimension one. Nevertheless, the problem formulation and the insights from the preceding chapters provide valuable tools for a unified treatment of the scenarios with SISO systems and one communication channel. Using these tools, existing results are reproduced in a different context in the following.

6.1 Ideal Observation Channel

An ideal observation channel describes the situation where the observations which are obtained at the system output can be transmitted to the controller without any error, i. e., where the input and the output signal of the observation channel are identical. In the limit for $t \to 0$, this situation is a special case of the system model which has been used so far because the channel noise is effectively scaled down to zero (cf. Figure 4.2). Consequently, all preceding results from Chapter 4 can be applied to the scenario where the observations obtained at the output of the dynamical system are transmitted over an ideal communication channel to the controller, provided that there is no power constraint for the corresponding channel or that the available transmit power $P_{\mathrm{Tx},1}$ is set to infinity.

Figure 6.1 depicts the scenario with an ideal observation channel, where the special case of a SISO system is considered, i. e., where the dimension of the system input and output is equal to

[1]For example, a decrease of the available transmit power for the control signal typically leads to an increased variance of the system state and thus an increased demand for communication resources for the transmission of the system output signal to the controller.

one and thus y_k and u_k, $k \in \mathbb{N}_0$, are scalars. Consequently, the system input vector $\boldsymbol{b}$ and output vector $\boldsymbol{c}$ are elements of $\mathbb{R}^{N_x}$. A similar situation for the MIMO case has already been investigated in Section 4.3.3.2 and is depicted in Figure 4.11. By comparison, the model which is considered in Figure 6.1 is a special case with $\boldsymbol{C_q} = \boldsymbol{0}_{N_y \times N_y}$, $N_y = N_u = 1$ and $t = 1$, which leads to the equivalence $\alpha^{-1} = g$ (cf. Equation 4.28).

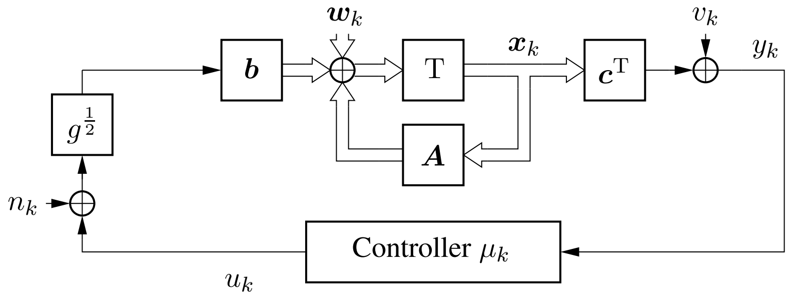


Figure 6.1: Model of the control loop which is closed over one channel with additive noise n_k. At the system input, the receive scaling $g^{\frac{1}{2}}$ is introduced.

Since it is has been shown in the preceding chapters how to jointly optimize the controller and the transmitter with a power constrained Linear Quadratic Gaussian (LQG) approach, which is also possible for the special case with $t = 1$ and $\boldsymbol{C_q} = \boldsymbol{0}_{N_y \times N_y}$ when $N_y = 1$, we address again the feasibility problem. More precisely, we want to determine the smallest value of the available transmit power $P_{\text{Tx},2}$ such that optimization problem (4.10) is feasible for the special case considered here. To this end, the feasibility test described in Section 4.3.1 is applied. Note that a power constraint for the ideal observation channel does not make sense at this point because it is either implicitly assumed that an infinite amount of power is available or, because no channel noise is present, any positive value of transmit power can be achieved by an appropriate scaling of the transmit signal y_k, $k \in \mathbb{N}_0$. Thus, the Lagrange multiplier λ_1 (cf. Equation 4.22) which is associated with the power constraint for the observation channel is set to zero. Additionally, in order to simplify the derivations, the transmit power constraints are not multiplied with the corresponding transmit and receive scaling factor, which provides the Lagrangian for the feasibility problem (cf. Equation 4.22)

$$L_{\text{feas}} = \lambda_2 (P_2 - P_{\text{Tx},2}), \tag{6.1}$$

where the transmit power P_2 is given by Equation (4.9) and we set $\lambda_1 = 0$. Consequently, the dual function (cf. Equation 4.23) for the feasibility test reads as

$$L^*_{\text{feas}}(\lambda_2) = \operatorname{tr}\left[\boldsymbol{K}\left(\boldsymbol{C_w} + g\boldsymbol{b}c_n\boldsymbol{b}^{\mathrm{T}}\right)\right] + \operatorname{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\right] - \lambda_2 P_{\text{Tx},2}, \tag{6.2}$$

where

$$\begin{aligned}\boldsymbol{K} &= \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - g^{\frac{1}{2}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{b}\left(g\boldsymbol{b}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{b} + \lambda_2\right)^{-1} g^{\frac{1}{2}}\boldsymbol{b}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} \\ &= \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{b}\left(\boldsymbol{b}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{b} + g^{-1}\lambda_2\right)^{-1}\boldsymbol{b}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A}\end{aligned} \tag{6.3}$$

and

$$\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}} = \boldsymbol{A}\left(\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}} - \boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\boldsymbol{c}\left(\boldsymbol{c}^{\mathrm{T}}\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\boldsymbol{c} + c_v\right)^{-1}\boldsymbol{c}^{\mathrm{T}}\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\right)\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{C_w} + g\boldsymbol{b}c_n\boldsymbol{b}^{\mathrm{T}}. \tag{6.4}$$

Additionally, we have $\boldsymbol{P} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \boldsymbol{K}$ and $\boldsymbol{C}_{\tilde{\boldsymbol{x}}} = \boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}} - \boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\boldsymbol{c}\left(\boldsymbol{c}^{\mathrm{T}}\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\boldsymbol{c} + c_v\right)^{-1}\boldsymbol{c}^{\mathrm{T}}\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}$. For the feasibility test in Section 4.3.1, the Lagrangian is maximized w.r.t. $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ in order to determine if the power constraints can be satisfied, where additionally the sum of the Lagrange multipliers is required to be equal to one. The latter constraint ensures that the dual function does not grow towards infinity (or zero) by a positive scaling of λ_1 and λ_2. Since in the present scenario we have $\lambda_1 = 0$, this constraint determines the value of λ_2 to be one and a maximization of the dual function (cf. optimization problem 4.26) is not necessary. Thus we are interested in the sign of $L^*_{\text{feas}}(1)$ which tells us if the minimal value of the difference $P_2 - P_{\text{Tx},2}$ can be negative and thus the power constraints can be fulfilled.

In order to analyze $L^*_{\text{feas}}(1)$, note that the solution of the Discrete Algebraic Riccati Equation (DARE) in Equation (6.3) is a scaled version of

$$\overline{\boldsymbol{K}} = \boldsymbol{A}^{\mathrm{T}}\overline{\boldsymbol{K}}\boldsymbol{A} - \boldsymbol{A}^{\mathrm{T}}\overline{\boldsymbol{K}}\boldsymbol{b}\left(\boldsymbol{b}^{\mathrm{T}}\overline{\boldsymbol{K}}\boldsymbol{b} + 1\right)^{-1}\boldsymbol{b}^{\mathrm{T}}\overline{\boldsymbol{K}}\boldsymbol{A}, \tag{6.5}$$

were $\boldsymbol{K} = g^{-1}\lambda_2\overline{\boldsymbol{K}}$. Using this identity, we get

$$L^*_{\text{feas}}(1) = \operatorname{tr}\left[\overline{\boldsymbol{K}}\left(g^{-1}\boldsymbol{C}_{\boldsymbol{w}} + \boldsymbol{b}c_n\boldsymbol{b}^{\mathrm{T}}\right)\right] + \operatorname{tr}\left[\overline{\boldsymbol{P}}\,\overline{\boldsymbol{C}_{\tilde{\boldsymbol{x}}}}\right] - P_{\text{Tx},2}, \tag{6.6}$$

where $\overline{\boldsymbol{P}} = g\lambda_2^{-1}\boldsymbol{P}$ and $\overline{\boldsymbol{C}_{\tilde{\boldsymbol{x}}}} = \overline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}} - \overline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\boldsymbol{c}\left(\boldsymbol{c}^{\mathrm{T}}\overline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\boldsymbol{c} + g^{-1}c_v\right)^{-1}\boldsymbol{c}^{\mathrm{T}}\overline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}$ with

$$\overline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}} = \boldsymbol{A}\left(\overline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}} - \overline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\boldsymbol{c}\left(\boldsymbol{c}^{\mathrm{T}}\overline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\boldsymbol{c} + g^{-1}c_v\right)^{-1}\boldsymbol{c}^{\mathrm{T}}\overline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\right)\boldsymbol{A}^{\mathrm{T}} + g^{-1}\boldsymbol{C}_{\boldsymbol{w}} + \boldsymbol{b}c_n\boldsymbol{b}^{\mathrm{T}}. \tag{6.7}$$

Since the first two summands of $L^*_{\text{feas}}(1)$ represent the actual transmit power P_2, the available transmit power $P_{\text{Tx},2}$ has to be larger than this sum, otherwise there exists no LQG controller which fulfills the power constraint.

Note that the preceding discussion is valid for a given value of g, where this scaling factor is still a degree of freedom which is used to determine a feasible solution. Thus, if we are interested in the smallest possible value of $P_{\text{Tx},2}$ such that $L^*_{\text{feas}}(1)$ (cf. Equation 6.6) is non-positive, we have to find the value of g which minimizes the sum $\operatorname{tr}\left[\overline{\boldsymbol{K}}\left(g^{-1}\boldsymbol{C}_{\boldsymbol{w}} + \boldsymbol{b}c_n\boldsymbol{b}^{\mathrm{T}}\right)\right] + \operatorname{tr}\left[\overline{\boldsymbol{P}}\,\overline{\boldsymbol{C}_{\tilde{\boldsymbol{x}}}}\right]$, which is identical to the actual transmit power P_2. Fortunately, it is easy to verify that $L^*_{\text{feas}}(1)$ is a monotonically decreasing function of g because an increase of g corresponds to a decrease of the variance of the driving noise process, given by the scaled covariance matrix $g^{-1}\boldsymbol{C}_{\boldsymbol{w}}$, and of the observation noise, given by $g^{-1}c_v$ (see Equations 6.6 and 6.7). Consequently, the infimal value of P_2 which can be approached arbitrarily close for $g \to \infty$ determines the lower bound for $P_{\text{Tx},2}$, i. e., in order to ensure feasibility of the power constrained LQG controller optimization, for $\lambda_2 = 1$ it must hold that (cf. Equation 6.6)

$$P_{\text{Tx},2} > \operatorname{tr}\left[\overline{\boldsymbol{K}}\boldsymbol{b}c_n\boldsymbol{b}^{\mathrm{T}}\right] + \operatorname{tr}\left[\overline{\boldsymbol{P}}\,\underline{\boldsymbol{C}_{\tilde{\boldsymbol{x}}}}\right], \tag{6.8}$$

with the asymptotic estimation error covariance matrix $\underline{\boldsymbol{C}_{\tilde{\boldsymbol{x}}}} = \underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}} - \underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\boldsymbol{c}\left(\boldsymbol{c}^{\mathrm{T}}\underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\boldsymbol{c}\right)^{-1}\boldsymbol{c}^{\mathrm{T}}\underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}$ and the corresponding DARE

$$\underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}} = \boldsymbol{A}\left(\underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}} - \underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\boldsymbol{c}\left(\boldsymbol{c}^{\mathrm{T}}\underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\boldsymbol{c}\right)^{-1}\boldsymbol{c}^{\mathrm{T}}\underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\right)\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{b}c_n\boldsymbol{b}^{\mathrm{T}}. \tag{6.9}$$

The dependence of the first summand of the right hand side of Equation (6.8) on the eigenvalues of the system matrix $\boldsymbol{A}$ can now be made explicit. It has been shown in [62] and [29] that for the DARE in Equation (6.5) it holds

$$\boldsymbol{b}^{\mathrm{T}}\overline{\boldsymbol{K}}\boldsymbol{b} + 1 = \prod\nolimits_{|\sigma_i|>1}|\sigma_i|^2, \tag{6.10}$$

where σ_i, $i \in \{1, 2, \ldots, N_x\}$, are the eigenvalues of the system matrix $\boldsymbol{A}$. Inserting Equation (6.10) in (6.8) results in the inequality

$$P_{\mathrm{Tx},2} > \left(\left(\prod\nolimits_{|\sigma_i|>1} |\sigma_i|^2 \right) - 1 \right) c_n + \operatorname{tr}\left[\overline{\boldsymbol{P}} \, \underline{\boldsymbol{C}_{\tilde{\boldsymbol{x}}}} \right]. \tag{6.11}$$

The first summand of the right hand side of Equation (6.11) is a well known result which provides the infimal value of the transmit power (or SNR when divided by the variance c_n of the channel noise) which has to be provided to stabilize an open loop unstable dynamical system over an additive noise channel.[2] The second summand which increases this value is also known, but in a different representation and with a different interpretation. In [29], this summand is absent if the controller has perfect information about the system state (the state feedback case) whereas it is a function of the relative degree and the non-minimum phase zeros of the open loop transfer function of the system to be stabilized (given by $\boldsymbol{A}$, $\boldsymbol{b}$ and $\boldsymbol{c}$) in the output feedback case, i. e., if the controller has only knowledge about y_k, $k \in \mathbb{N}_0$.

The absence of the second summand of the right hand side of Equation (6.11) is obvious for the state feedback case because the estimation error of the system state is zero if the state is known. For the output feedback case, the relative degree and the non-minimum phase zeros of the dynamical system have to be interpreted in terms of the resulting state estimation error. At this point, we only provide a qualitative discussion, the quantitative description can be found in [29]. Assume that the dynamical system to be controlled is minimum phase and has relative degree one, i. e., all zeros of its corresponding transfer function have a magnitude less than one and the degree of its denominator is greater than the degree of numerator by one. Using Equation (6.9), it is easy to verify that

$$\underline{\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}} = \boldsymbol{b} c_n \boldsymbol{b}^{\mathrm{T}} \tag{6.12}$$

is a positive semidefinite solution of the corresponding DARE if $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{b} \neq 0$, which holds because the relative degree of the system is assumed to be one (see, e. g., [32, p. 93]). With the minimum phase assumption, it is a stabilizing solution, i. e., the magnitude of all eigenvalues of

$$\boldsymbol{A}^{\mathrm{T}} - \boldsymbol{c}\left(\boldsymbol{c}^{\mathrm{T}} \underline{\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}} \boldsymbol{c}\right)^{-1} \boldsymbol{c}^{\mathrm{T}} \underline{\boldsymbol{C}_{\tilde{\boldsymbol{x}}}^{\mathrm{P}}} \boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}^{\mathrm{T}} - \boldsymbol{c}\left(\boldsymbol{c}^{\mathrm{T}}\boldsymbol{b}\right)^{-1} \boldsymbol{b}^{\mathrm{T}} \boldsymbol{A}^{\mathrm{T}} \tag{6.13}$$

is less than one. It is shown in [119] that the eigenvalues of the matrix in Equation (6.13) are the zeros of the open loop transfer function of the system under consideration, which have a magnitude less than one by assumption.[3] Having shown that the covariance matrix in Equation (6.12) is the positive semidefinite stabilizing solution of the DARE in Equation (6.9), it can be seen that the resulting covariance matrix of the state estimation error is zero, i. e., $\underline{\boldsymbol{C}_{\tilde{\boldsymbol{x}}}} = \boldsymbol{0}_{N_x \times N_x}$, and the second summand of the right hand size of Equation (6.11) is also zero.

The result above allows for a more intuitive interpretation. If the dynamical system to be controlled is minimum phase, it can be inverted using a stable system. Consequently, with the ideal

[2] The authors of [29] provide a remark for the case when $\boldsymbol{A}$ has eigenvalues with magnitude one. These eigenvalues obviously do not contribute to Equation (6.10) but still have to be stabilized. Nevertheless, this can be established with an arbitrarily small increase of transmit power.

[3] The eigenvalues of the matrix in Equation (6.13) correspond to the zeros of the transfer function for the system described by $\boldsymbol{A}^{\mathrm{T}}$, the system input vector $\boldsymbol{c}$ and the output vector $\boldsymbol{b}$. Since this system has the same transfer function as the original one, the above statement is correct. Additionally, it is shown in [120] that a stabilizing controller gain for this situation is given by $\boldsymbol{l}^{\mathrm{T}} = -\left(\boldsymbol{b}^{\mathrm{T}}\boldsymbol{c}\right)^{-1} \boldsymbol{b}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$, which provides the stable closed loop matrix $\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{c}\boldsymbol{l}^{\mathrm{T}} = \boldsymbol{A}^{\mathrm{T}} - \boldsymbol{c}\left(\boldsymbol{c}^{\mathrm{T}}\boldsymbol{b}\right)^{-1} \boldsymbol{b}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$.

knowledge of the system output sequence $(y_k : k \in \mathbb{N}_0)$, its input sequence which is the scaled sum of $(u_k : k \in \mathbb{N}_0)$ and $(n_k : k \in \mathbb{N}_0)$ can be reconstructed after the transient phase. Due to the relative degree of one, a one step delay has to be introduced for a causal inversion. Now, with the knowledge of the system input[4], the decay of the influence of the initial state (the closed loop system is stable) and the absence of process noise (since we are considering the asymptotic case for $g \to \infty$), the system state can be estimated asymptotically with no error. This interpretation also shows why the estimation error grows with the relative degree and for non-minimum phase systems. In the minimum phase case, the increase of the relative degree results in an additional delay for the state estimate, which increases the estimation error. Non-minimum phase systems can not be stably inverted and thus their input sequence (precisely, the channel noise which is fed into the system) can not be perfectly reconstructed, which also leads to an increase of the state estimation error.

6.2 Ideal Control Channel

For the second scenario which takes into account only one communication channel in the closed control loop, it is assumed that the control signal is transmitted to the input of the dynamical system to be controlled over an ideal channel, e. g., when the controller is spatially very close to the system input. As in Section 6.1, this is a special case of the general scenario considered in Chapter 4 if no power constraint is taken into account for the control channel, i. e., if $P_{\mathrm{Tx},2}$ is arbitrarily large. With no power constraint, the receive scaling factor g in the control channel can be chosen arbitrarily small which effectively results in a vanishing effect of the corresponding channel noise on the closed loop system. A second point of view is the case where the channel noise is zero, i. e., $\boldsymbol{C_n} = \mathbf{0}_{N_u \times N_u}$. In this case, a transmit power constraint is meaningless since it can be set to any positive value by the appropriate choice of g. Note that this is a special case of the scenario considered in Section 4.3.2 (cf. Figure 4.6). Using $\boldsymbol{C_n} = \mathbf{0}_{N_u \times N_u}$ with $N_y = N_u = 1$ there, the fixed value of $g = 1$ and the resulting identity $\alpha = t$ (cf. Equation 4.28) lead to the scenario depicted in Figure 6.2. Consequently, the results which have been derived earlier also apply to this special case.

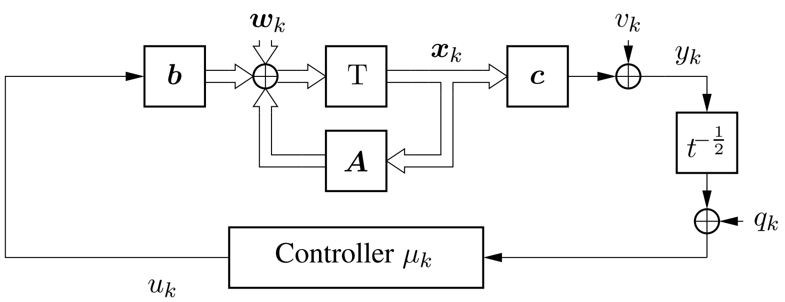


Figure 6.2: Model of the control loop which is closed over one channel with additive noise q_k. At the system output, the transmit scaling $t^{-\frac{1}{2}}$ is introduced.

For the investigation of the feasibility problem with the ideal control channel, i. e., the determination of the smallest value of the available transmit power $P_{\mathrm{Tx},1}$ which has to be provided such that a power constrained LQG controller exists, we parallel Section 6.1 to a large extent. Using

[4] Due to the one step delay, the system input $g^{\frac{1}{2}}(u_{k-1}+n_{k-1})$ can be reconstructed at time index k. In the noiseless case, this corresponds to the knowledge of the system state $\boldsymbol{x}_k = \boldsymbol{A}\boldsymbol{x}_{k-1} + \boldsymbol{b}g^{\frac{1}{2}}(u_{k-1}+n_{k-1})$.

again the approach presented in Section 4.3.1, the Lagrangian for the feasibility test reads as (cf. Equation 6.1):

$$L_{\text{feas}} = \lambda_1 \left(P_1 - P_{\text{Tx},1}\right), \tag{6.14}$$

where the actual transmit power P_1 of the observation channel is given by Equation (4.8). For a fixed value of the scaling factor t, the corresponding dual function reads as (cf. Equation 6.2)

$$L^*_{\text{feas}}(\lambda_1) = \operatorname{tr}\left[\boldsymbol{K}\boldsymbol{C}_{\boldsymbol{w}}\right] + \operatorname{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\right] - \lambda_1 P_{\text{Tx},1}, \tag{6.15}$$

where

$$\boldsymbol{K} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{b}\left(\boldsymbol{b}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{b}\right)^{-1}\boldsymbol{b}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} + t^{-1}\lambda_1\boldsymbol{c}\boldsymbol{c}^{\mathrm{T}} \tag{6.16}$$

and

$$\begin{aligned}\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}} &= \boldsymbol{A}\left(\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}} - t^{-\frac{1}{2}}\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\boldsymbol{c}\left(t^{-1}\boldsymbol{c}^{\mathrm{T}}\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\boldsymbol{c} + t^{-1}c_v + c_q\right)^{-1}t^{-\frac{1}{2}}\boldsymbol{c}^{\mathrm{T}}\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\right)\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{w}}\\ &= \boldsymbol{A}\left(\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}} - \boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\boldsymbol{c}\left(\boldsymbol{c}^{\mathrm{T}}\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\boldsymbol{c} + c_v + tc_q\right)^{-1}\boldsymbol{c}^{\mathrm{T}}\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\right)\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{w}}.\end{aligned} \tag{6.17}$$

The matrices $\boldsymbol{P}$ and $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}$ in Equation (6.15) are given by $\boldsymbol{P} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \boldsymbol{K} + t^{-1}\lambda_1\boldsymbol{c}\boldsymbol{c}^{\mathrm{T}}$ and $\boldsymbol{C}_{\tilde{\boldsymbol{x}}} = \boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}} - \boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\boldsymbol{c}\left(\boldsymbol{c}^{\mathrm{T}}\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\boldsymbol{c} + c_v + tc_q\right)^{-1}\boldsymbol{c}^{\mathrm{T}}\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}$. Since only one power constraint, represented by $P_{\text{Tx},1}$, is present, the feasibility test described in Section 4.3.1 reduces to the evaluation of $L^*_{\text{feas}}(\lambda_1)$ at $\lambda_1 = 1$ and to verify if its sign is not positive, i. e., the minimal value of the actual transmit power P_1 does not exceed the available power $P_{\text{Tx},1}$. As in Section 6.1, the degree of freedom which is provided by the transmit scaling factor t is used to determine the smallest lower bound for $P_{\text{Tx},1}$ such that the power constrained LQG controller design is feasible.

Note that the solution of the DARE in Equation (6.16) can also been obtained by solving

$$\overline{\boldsymbol{K}} = \boldsymbol{A}^{\mathrm{T}}\overline{\boldsymbol{K}}\boldsymbol{A} - \boldsymbol{A}^{\mathrm{T}}\overline{\boldsymbol{K}}\boldsymbol{b}\left(\boldsymbol{b}^{\mathrm{T}}\overline{\boldsymbol{K}}\boldsymbol{b}\right)^{-1}\boldsymbol{b}^{\mathrm{T}}\overline{\boldsymbol{K}}\boldsymbol{A} + \boldsymbol{c}\boldsymbol{c}^{\mathrm{T}} \tag{6.18}$$

and using the identity $\boldsymbol{K} = t^{-1}\lambda_1\overline{\boldsymbol{K}}$, i. e., both solutions only differ by a scalar factor. Together with the matrices $\overline{\boldsymbol{P}} = t\lambda_1^{-1}\boldsymbol{P}$ and $\overline{\boldsymbol{C}_{\tilde{\boldsymbol{x}}}} = \overline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}} - \overline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\boldsymbol{c}\left(\boldsymbol{c}^{\mathrm{T}}\overline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\boldsymbol{c} + t^{-1}c_v + c_q\right)^{-1}\boldsymbol{c}^{\mathrm{T}}\overline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}$, where

$$\overline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}} = \boldsymbol{A}\left(\overline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}} - \overline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\boldsymbol{c}\left(\boldsymbol{c}^{\mathrm{T}}\overline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\boldsymbol{c} + t^{-1}c_v + c_q\right)^{-1}\boldsymbol{c}^{\mathrm{T}}\overline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\right)\boldsymbol{A}^{\mathrm{T}} + t^{-1}\boldsymbol{C}_{\boldsymbol{w}}, \tag{6.19}$$

the value $L^*_{\text{feas}}(1)$ reads as

$$L^*_{\text{feas}}(1) = \operatorname{tr}\left[t^{-1}\overline{\boldsymbol{K}}\boldsymbol{C}_{\boldsymbol{w}}\right] + \operatorname{tr}\left[\overline{\boldsymbol{P}}\,\overline{\boldsymbol{C}_{\tilde{\boldsymbol{x}}}}\right] - P_{\text{Tx},1}. \tag{6.20}$$

Since this value is non-positive for a feasible optimization problem, the first two summands of the right hand side of Equation (6.20) should be as small as possible. This is achieved for $t \to \infty$, which is obvious for the first summand and intuitively clear for the second because the effective variance $t^{-1}c_v$ of the observation noise and the covariance matrix $t^{-1}\boldsymbol{C}_{\boldsymbol{w}}$ of the process noise are scaled down towards zero in this case which results in a monotonic decrease of the estimation error of the (scaled) system state. Consequently, the available transmit power must fulfill the inequality

$$P_{\text{Tx},1} > \operatorname{tr}\left[\overline{\boldsymbol{P}}\,\underline{\boldsymbol{C}_{\tilde{\boldsymbol{x}}}}\right], \tag{6.21}$$

where $\underline{\boldsymbol{C}_{\tilde{\boldsymbol{x}}}} = \underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}} - \underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\boldsymbol{c}\left(\boldsymbol{c}^{\mathrm{T}}\underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\boldsymbol{c} + c_q\right)^{-1}\boldsymbol{c}^{\mathrm{T}}\underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}$ with (cf. Equation 6.19)

$$\underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}} = \boldsymbol{A}\left(\underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}} - \underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\boldsymbol{c}\left(\boldsymbol{c}^{\mathrm{T}}\underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\boldsymbol{c} + c_q\right)^{-1}\boldsymbol{c}^{\mathrm{T}}\underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\right)\boldsymbol{A}^{\mathrm{T}}. \tag{6.22}$$

In order to obtain results analogous to Section 6.1, it is advantageous to rewrite the right hand side of Equation (6.21). This expression is nothing but the infimal value of the transmit power P_1 for the observation channel which can be achieved by an LQG controller in the limit for $t \to \infty$. For a finite value of t, this power is also given by (cf. Equations 3.8 and 4.8)

$$P_1 = t^{-1}\left(\boldsymbol{c}^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{x}}\boldsymbol{c} + c_v\right), \tag{6.23}$$

where the covariance matrix $\boldsymbol{C}_{\boldsymbol{x}}$ of the system state is determined according to Equations (3.13) – (3.15) and reads in our case as

$$\boldsymbol{C}_{\boldsymbol{x}} = (\boldsymbol{A} + \boldsymbol{b}\boldsymbol{l}^{\mathrm{T}})\boldsymbol{C}_{\hat{\boldsymbol{x}}}(\boldsymbol{A} + \boldsymbol{b}\boldsymbol{l}^{\mathrm{T}})^{\mathrm{T}} + \boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}, \tag{6.24}$$

with the optimal controller gain

$$\boldsymbol{l}^{\mathrm{T}} = -\left(\boldsymbol{b}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{b}\right)^{-1}\boldsymbol{b}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} = -\left(\boldsymbol{b}^{\mathrm{T}}\overline{\boldsymbol{K}}\boldsymbol{b}\right)^{-1}\boldsymbol{b}^{\mathrm{T}}\overline{\boldsymbol{K}}\boldsymbol{A}. \tag{6.25}$$

Using Equation (6.24) which is multiplied by t^{-1} for the determination of the transmit power (cf. Equation 6.23), the right hand side of Equation (6.21), which is the limit for $t \to \infty$, can now also be expressed as

$$\mathrm{tr}\left[\overline{\boldsymbol{P}}\underline{\boldsymbol{C}_{\tilde{\boldsymbol{x}}}}\right] = \boldsymbol{c}^{\mathrm{T}}\left((\boldsymbol{A} + \boldsymbol{b}\boldsymbol{l}^{\mathrm{T}})\underline{\boldsymbol{C}_{\hat{\boldsymbol{x}}}}(\boldsymbol{A} + \boldsymbol{b}\boldsymbol{l}^{\mathrm{T}})^{\mathrm{T}} + \underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\right)\boldsymbol{c}. \tag{6.26}$$

The (scaled) covariance matrix $\underline{\boldsymbol{C}_{\hat{\boldsymbol{x}}}}$ is the solution of a discrete Lyapunov equation analogous to Equation (3.15) where the error covariance matrices $\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}$ and $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}$ are replaced by their (scaled) counterparts $\underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}$ and $\underline{\boldsymbol{C}_{\tilde{\boldsymbol{x}}}}$, respectively. Thus, the inequality from Equation (6.21) becomes

$$P_{\mathrm{Tx},1} > \boldsymbol{c}^{\mathrm{T}}\underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\boldsymbol{c} + \boldsymbol{c}^{\mathrm{T}}(\boldsymbol{A} + \boldsymbol{b}\boldsymbol{l}^{\mathrm{T}})\underline{\boldsymbol{C}_{\hat{\boldsymbol{x}}}}(\boldsymbol{A} + \boldsymbol{b}\boldsymbol{l}^{\mathrm{T}})^{\mathrm{T}}\boldsymbol{c}. \tag{6.27}$$

We are now in the position to provide a formulation of this inequality analogous to Equation (6.11). Noting that the DAREs for the determination of $\overline{\boldsymbol{K}}$ (cf. Equation 6.5) in the preceding section and $\underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}$ (cf. Equation 6.22) in this section are structurally identical and obtained from each other when replacing $\boldsymbol{A}$ by $\boldsymbol{A}^{\mathrm{T}}$, $\boldsymbol{b}$ by $\boldsymbol{c}$, and a multiplication by c_q, the result of Equation (6.10) can be used and we obtain

$$\boldsymbol{c}^{\mathrm{T}}\underline{\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}}\boldsymbol{c} + c_q = \left(\prod\nolimits_{|\sigma_i|>1}|\sigma_i|^2\right)c_q, \tag{6.28}$$

where σ_i, $i \in \{1, 2, \ldots, N_x\}$, are the eigenvalues of $\boldsymbol{A}^{\mathrm{T}}$ and thus also of $\boldsymbol{A}$. This provides the final representation of the inequality which must hold in order to obtain a feasible power constrained LQG optimization problem:

$$P_{\mathrm{Tx},1} > \left(\left(\prod\nolimits_{|\sigma_i|>1}|\sigma_i|^2\right) - 1\right)c_q + \boldsymbol{c}^{\mathrm{T}}(\boldsymbol{A} + \boldsymbol{b}\boldsymbol{l}^{\mathrm{T}})\underline{\boldsymbol{C}_{\hat{\boldsymbol{x}}}}(\boldsymbol{A} + \boldsymbol{b}\boldsymbol{l}^{\mathrm{T}})^{\mathrm{T}}\boldsymbol{c}. \tag{6.29}$$

The first summand of the right hand side of Equation (6.29) is, analogous to Equation (6.11), the well known infimal value of the transmit power (or the SNR when divided by the channel noise variance c_q) which is necessary to stabilize an open loop unstable dynamical system over an additive noise communication channel. Since it depends on the error covariance matrix $\underline{\boldsymbol{C}}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}$ and the system output vector $\boldsymbol{c}$ alone, the controller has no influence on this part of the transmit power. Recalling that the optimal LQG controller uses the Kalman filter (see Appendix A7) to determine the optimal estimate of the system state, it seems reasonable that the available transmit power must be at least as large as $\boldsymbol{c}^{\mathrm{T}}\underline{\boldsymbol{C}}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\boldsymbol{c}$ since this is the variance of the *innovation* sequence (subtracted by the variance of the channel noise) which is computed by the Kalman filter to decorrelate the channel output sequence for the sequential estimate of the system state (see Equations A192, A193 and A196). Thus, the variance of the innovations, which represent the actual information processed by the Kalman filter, determines the infimal value of the transmit power $P_{\mathrm{Tx},1}$.

The second summand of the right hand side of Equation (6.29) can be interpreted as in Section 6.1 in terms of the relative degree of the system to be controlled and the zeros of its transfer function. Assume that the system to be controlled is minimum phase and has relative degree one. Repeating the argumentation of Section 6.1, it can be shown that

$$\overline{\boldsymbol{K}} = \boldsymbol{c}\boldsymbol{c}^{\mathrm{T}} \tag{6.30}$$

is a positive semidefinite stabilizing solution of the DARE in Equation (6.18), which results in the optimal controller gain (cf. Equation 6.25)

$$\boldsymbol{l}^{\mathrm{T}} = -(\boldsymbol{b}^{\mathrm{T}}\boldsymbol{c})^{-1}\boldsymbol{c}^{\mathrm{T}}\boldsymbol{A}. \tag{6.31}$$

It is now easy to verify that the product $\boldsymbol{c}^{\mathrm{T}}\left(\boldsymbol{A} + \boldsymbol{b}\boldsymbol{l}^{\mathrm{T}}\right)$ is equal to $\boldsymbol{0}^{\mathrm{T}}_{N_x}$ and thus the second summand on the right hand side of Equation (6.29) is zero. This behavior has been reported in [86] and explained by the fact that the optimal controller inverts the system to be controlled, which is possible using a stable system due to the minimum phase assumption, and subtracts the optimal prediction of the system output from the actual output. A transmit power of zero would be achieved if the system output could be subtracted from itself, but since the controller can only determine an (imperfect) state estimate, the signal that must be subtracted from the system output in order to minimize the variance of the difference is its optimally predicted value. Consequently, the variance of the remaining transmit signal is the variance of the prediction error, given by $\boldsymbol{c}^{\mathrm{T}}\underline{\boldsymbol{C}}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\boldsymbol{c}$. The prediction and not estimate is subtracted due to the one step delay of the closed control loop which is a result of the relative degree, assumed to be one, of the dynamical system to be controlled.

From this point of view it becomes clear that the summand $\boldsymbol{c}^{\mathrm{T}}(\boldsymbol{A} + \boldsymbol{b}\boldsymbol{l}^{\mathrm{T}})\underline{\boldsymbol{C}}_{\hat{\boldsymbol{x}}}(\boldsymbol{A} + \boldsymbol{b}\boldsymbol{l}^{\mathrm{T}})^{\mathrm{T}}\boldsymbol{c}$ in Equation (6.29) is strictly larger than zero for the non-minimum phase case or for systems with a relative degree larger than one. In this case, additional delay is present in the control loop and thus the variance of the transmit signal increases because the one step prediction of the system output can not be subtracted at the transmitter any more. With a non-minimum phase system, it is not possible to stably invert the dynamical system to be controlled and thus it is not possible for a controller to subtract an arbitrary signal from the system output.

6.3 Discussion

Due to the general treatment of the joint optimization of LQG controller and scalar transceivers, the results presented in Chapter 4 can easily be applied to the case of SISO systems with only one communication channel in the control loop. In order to compare the presented approaches with known results, Chapter 6 considered this special case since it received considerable attention in the existing literature on NCSs. The solutions obtained in Chapter 4 are confirmed by the reproduction of well known results for the SISO one channel case, i. e., the feasible transmit power which is lower bounded by a function of the unstable eigenvalues of the open loop system. Additionally, the fact that this lower bound is increased for non-minimum phase systems with relative degree larger than one has also been verified.

For the SISO case where only the control signals are transmitted over an imperfect, i. e., noisy, communication channel, the authors of [29] carry out an extensive analysis of control loops which are closed over an additive noise channel using frequency domain formulations. They consider continuous-time as well as discrete-time systems with state and output feedback and provide expressions for the minimal SNRs which is necessary to stabilize open loop unstable systems in terms of the unstable poles, non-minimum phase zeros and the relative degree of the system. The optimization of transmit and receive scaling factors is not considered because these degrees of freedom are not necessary in their system model in the absence of process and observation noise. Additionally, system performance is not considered since the authors investigate the relation of SNR and stabilizability.

The authors of [86] describe the observation channel by an additive noise model while the control channel is assumed to be ideal and extend the system model of [29] by introducing process noise and an inverse scaling of the transmit and received signal of the communication channel. The cost function to be minimized is the variance of the observations at the system output, i. e., the transmit signal of the observation channel. With the assumption of a minimum phase system with relative degree equal to one, the structure of the optimal controller is derived and it is shown that the system output is (except for the channel noise) identical with the innovation sequence of the optimal estimator. It is also demonstrated how the transmit scaling factor can be used for a trade-off between the value of the cost function and the available transmit power. Additionally, information theoretic tools are applied to analyze the dependence of the transmit power requirements on the process noise sequence. In [87], information theoretic quantities are used for the investigation of the transmit power requirements for general non-linear and time varying transmit and receive strategies.

A different approach to the consideration of a transmit scaling factor at the output of a dynamical system to be controlled is proposed in [85], where again the observation channel is assumed to be non-ideal. The authors extend the classical LQG cost function by an additional summand containing the squared value of the scaling factor. The resulting optimization problem is analyzed in terms of (quasi-)convexity for a scalar dynamical system where the transmitter has access to the scalar system state. Instead of frequency domain formulations, the analysis relies on the DAREs for the optimal regulator and state estimator. Using the modified LQG cost function, the transmit power is bounded by preventing the scaling factor at the transmitter from going to infinity, but there is no explicit control over the actual transmit power, e. g., by a power constraint for the optimization of the LQG controller.

A constrained optimization approach based on transfer functions is used in [46] and [35] for the design of an LQG controller subject to a limited SNR of the communication channel. The variance of the channel noise becomes a design variable, which renders the considered problem equivalent to one with a power constraint and scaling factors at transmitter and receiver. This fact is pointed out in [35]. Besides the investigation of minimal SNR requirements, also performance issues are addressed and a joint optimization of the controller and the channel noise variance is proposed. Unfortunately, no statement about the (non-)convexity of the problem is provided. Since it has been demonstrated that the general problem considered in Chapter 4 is not convex, it is not obvious that the algorithm proposed in [46] is capable to determine the global optimum. Nevertheless, many questions treated in Chapter 4 are motivated by [46] and [35].

7. Summary and Conclusions

In the context of optimal control, it is natural to consider constraints in the optimization problems which are solved for the determination of the optimal controller. This is reflected by the monographs [78] and [121] which provide a lot of examples for constrained controller designs. Networked Control Systems (NCSs) are a special class of such optimization problems, where the channels which are used for the exchange of information in a control system are explicitly taken into account. Consequently, the constraints to be considered are not due to the required performance of the control system but due to the limitations which are inherent to communication systems. Ultimately, these limitations are the result of finite physical resources for the transmission of information, e. g., bandwidth or power. This leads to effects like received signals which are corrupted by noise, data loss because of collisions of users which have to share the same communication channel or finite data rates due to channels with limited Shannon capacity, which are observed for different types of communication infrastructures.

In order to take into account the communication system for the design of the controller, the optimization problems which are formulated for the determination of a controller have to be modified depending on the actual model of the communication channels which are used for the transmission of information. Section 2.1 provided an overview of the variety of models which are commonly used in the literature on NCSs and references which give an example of the various approaches to the joint treatment of communications and control. In this thesis, we decided for a simple additive noise channel model. This implies that the limited communication resource is the transmit power because otherwise an unbounded amplification of the transmit signal could be used to eliminate the negative impact of the additive disturbance.

For the investigation of a control loop which is closed over power constrained channels with additive noise, we applied the Linear Quadratic Gaussian (LQG) framework. There are several reasons for this choice. First of all, it is possible to build on a well understood and well developed basis for the design of an optimal controller which supports the investigation of the aspects related to the consideration of communication channels. Second, and even more important, this framework allows for a straightforward embedding of the communication system, i. e., the channel model and the associated power constraints, in the overall system model of the NCS. Consequently, the analysis and the design of the control and the communication system can be performed jointly and on the same level of abstraction. This is an interesting property compared to, e. g., packet drop channels where the complete communication system is reduced to the probability of a data loss and thus has to be designed to minimize this probability irrespective of the control application. Nevertheless, such channel models are of great importance if an existing communication infrastructure like the Internet shall be used for a control task.

We started the consideration of additive noise communication channels and the associated power constraints in Chapter 3, where both links which connect the dynamical system to be controlled with the controller, i. e., for the transmission of observations and control signals, are modeled as noisy channels. This represents a remote control scenario where the dynamical system and the controller are spatially separated. For the situation investigated in Chapter 3, the controller is the only degree of freedom to fulfill the transmit power constraints of both communication channels. In this case, the constrained controller optimization is a convex problem which can be solved relatively easy. Solution approaches date back to the 1980s [81, 96, 122] and are systematically investigated in the context of convex optimization in [78]. Thus, the presented results are not new and serve as the introduction of the basic problem to be solved. Note that we have chosen a reformulation of the standard infinite horizon LQG control problem with average cost in terms of the stationary covariance matrices of the system state and the control signal. Such a formulation has recently been used by [88] and allows for a direct application of tools and software packages for convex optimization like [92] and [93].

The results presented in Chapter 3 also demonstrated that it is problematic to use the controller as the only degree of freedom to optimize a control specific performance criterion and at the same time to fulfill the transmit power constraints. The reason is that these quantities, i. e., the LQG cost and the transmit powers, are linked by the properties of the closed loop system and it is thus not possible in general to use an increased amount of transmit power to improve the control performance, i. e., decrease the value of the cost function.

In contrast to a considerable amount of the literature on NCSs which investigates the model of additive noise channels, we restricted the available amount of transmit power while keeping the variance of the channel noise fixed instead of limiting the Signal to Noise Ratio (SNR) of the communication channels, i. e., the ratio of these two quantities. This emphasizes the point of view that the channel noise is a physical quantity, e. g., thermal noise, which can not be changed. However, in addition to the controller, there is the degree of freedom to design a transmitter and a receiver at the input and the output of the communication channel in order to satisfy the power constraints. This point of view has been made explicit in Chapter 4 with the introduction of scaling factors at the input and the output of the dynamical system to be controlled as the most simple instances of transmitters and receivers. Note that this extension of the system model represents no loss of generality w.r.t. to comparable approaches, e. g., [35, 46], where no transmit or receive scaling is taken into account but the variance of the channel noise is considered as a degree of freedom. Additionally, the respective SNR constraints can be reformulated as power constraints.

The joint optimization of the LQG cost function w.r.t. the separated elements controller, transmitter and receiver can be interpreted as a controller optimization with structural constraints. In more detail, the controller is required to be split up into two gain factors which directly act on the input and output signals of the dynamical system, and a general function which maps the output of one communication channel to the input of the other channel. Thus, the investigation of the joint optimization problem in Section 4.2.2 revealed a well known property of distributed controller design: the non-convexity of the associated optimization problem [39–41]. For example, it has been shown that the set of pairs of the transmit and receive scaling factors which lead to a feasible power constrained optimization problem is not necessarily connected. Even the unconstrained optimization problem for the determination of Pareto optimal values of the transmit powers turned out to be non-convex and to have several local optima. Consequently, such problems are not tractable using standard methods for convex optimization.

Since fixed values of transmit and receive scaling factors result in a convex controller design problem, only a small number of variables has to be determined in the non-convex setting. It is thus possible to sample these variables. However, this approach allows for no statement about the global optimality of an obtained result and the quality of a possibly suboptimal solution. Fortunately, although non-convex, the considered optimization problems exhibit monotonicity properties which can be used for the application of a branch and bound method in order to determine the global optimum [102, 103]. The derivation of lower bounds for the minimal value of the optimization problems under consideration and the application of the branch and bound approach to the determination of the optimizing transmit and receive scaling factors are main contributions of this thesis.

Note that the branch and bound method which exploits the monotonicity properties of the optimization problem can also be interpreted as a systematic way to sample its "non-convex" variables. The required lower bounds are obtained by solving a relaxed convex LQG control problem with fixed values of the transmit and receive scaling factors. The upper bounds are simply calculated by solving the original convex LQG control problem for fixed transmitters and receivers. Since the determination of these bounds thus has the same computational complexity as the evaluation of the optimization problem in a conventional sampling of the "non-convex" variables, the application of the branch and bound approach is not prohibitively complex, at least for a small number of variables.

In Chapter 5, the joint optimization of the LQG controller and the scalar transmit and receive scaling factors has been extended to transceivers which are represented by diagonal matrices. This extension is straightforward for communication channels with diagonal noise covariance matrices. For the general case of non-diagonal covariance matrices (Section 5.2.1) and linear channel disturbances (Section 5.2.2), we proposed a modification of the transmitter and receiver such that the resulting equivalent communication channels are reduced to the diagonal case. Consequently, this step allows for the application of the branch and bound approach for the determination of optimizing diagonal transceiver matrices. However, the diagonalization of the communication channels is performed independently of the controller design and thus can not be claimed to be optimal.

Finally, the results in Chapter 6 have been presented in order to apply the framework for the analysis and design of power constrained control systems to scenarios which are commonly used in the literature on NCSs. Some of the fundamental results which have been derived in the past for Single-Input Single-Output (SISO) systems with only one communication channel in the control loop have been reproduced.

The investigation of NCSs involves a variety of concepts which have been developed independently for the analysis and the design of communication and control systems. Especially the models of the communication channels which are used for the information exchange in a control loop and the associated constraints determine to a large extent the approach to the joint design of the controller and the components of the considered communication system. To name a few examples, a structural separation of transceivers and the controller can be treated in the context of distributed control systems, packet drop channels can be considered within the framework of Markov Jump Linear Systems (see, e. g., [109] and references therein), and additive noise communication channels with power constraints fit in the LQG framework. However, there are several problems which prevent a direct application of results from the communication and information theory to NCSs and render the joint treatment of control and communications a challenging task.

The first problem is that sophisticated coding and transmission schemes which can be applied to a pure communication task often require the processing of data in large blocks and thus introduce a considerable amount of delay which can not be tolerated in a control loop in general. The reduction of this delay typically results in more distortion or transmission errors which degrades the quality of the information which is available to the controller. Additionally, such distortions are also fed into the dynamical system to be controlled which makes it even harder to determine a controller which achieves a satisfactory performance. In the worst case, the controller is not able to combat the distortions and errors with the limited amount of available information which renders the problem of the joint design of the control and the communication system infeasible.

Such considerations are of course not relevant for the investigation of additive noise channels in the LQG framework with the additional restriction to linear transmitters and receivers. However, the second problem which is encountered in optimization problems for NCSs is still present. In a pure communication setting, the information source can be treated independently of the transmitters and receivers which are designed for the transmission of information over a communication channel. For example, a correlated source can be whitened in order to remove redundancy or scaled appropriately such that it has a desired variance. This separation is not possible for the transmit signals in a closed control loop where the sources of information are the dynamical system (for the transmission of system outputs to the controller) and the controller (for the transmission of control signals to the system input). Since they are interconnected in a feedback loop, the choice of a transmitter or a receiver has a direct impact on the properties of the signals to be transmitted or received. Consequently, the communication system and the information which is exchanged using this system are not independent any more. This fact has been observed for the minimization of the LQG cost function under power constraints in Chapter 4 where the covariance matrices of the signals to be transmitted are functions of the transmit and receive scaling factors. This is the main reason for the difficulties which are encountered for the joint optimization of the LQG controller, the transmitter and the receiver even for the case of a simple scaling of transmit and receive signals.

In this thesis, we proposed a numerical solution for the problem of a joint optimization of the controller and the transmit and receive scaling factors in the LQG framework, where the control loop is closed over power constrained additive noise channels. This approach has also been generalized to the optimization of diagonal transmitters and receivers. However, there remain open questions. For the system model under consideration, we observed that the assignment of the individual scalar components of the system input and output signals to the components (or subspaces) of the available communication channels has a significant impact on the performance of the closed loop control system. We have no satisfactory answer to the question how to choose this assignment, i. e., how to jointly optimize (diagonal) transmitters and receivers and the respective permutation of the input and output signals of the dynamical system to be controlled. A further question is of more interest but also even more difficult to answer. What are the optimal linear transmitters and receivers? For a communication (or estimation) setting, the answer to this question is known for memoryless [110–112] and temporally correlated [115, 116] sources. However, due to the coupling of the sources of information and the transmitters and receivers which are used for the information exchange, it is not obvious how these approaches can be extended to a closed loop scenario. For a SISO system with one control channel in the control loop, the authors of [123] propose a joint transceiver and controller design. They provide a reformulation of the problem which is not convex but quasi-convex and thus a systematic approach for the determination of the global optimum is possible.

As a final remark, the non-convexity of optimization problems for the joint design of control and communication systems complicates the determination of their globally optimal solution but, at least for some cases, does not prevent the possibility to determine it. An example is the joint design of transmitters and receivers for data transmission or estimation, respectively [115, 116]. This non-convex optimization problem can be reformulated as a problem which allows for an efficient determination of the solution. Even if such a reformulation is not available, the problem may still have properties which can be exploited by global optimization algorithms. The problems which have been investigated in this thesis have such a property, i. e., monotonicity in the “non-convex” optimization variables, which is the basis for a branch and bound method for the determination of the global optimum. This approach can be interpreted as a systematic sampling procedure which additionally provides information about the accuracy of the obtained results. Consequently, at least for low-dimensional problems, such non-convex optimization problems are feasible.

Appendix

A1. Sampling of Continuous-Time Linear Systems

Although physical (stochastic) processes are generally functions of a continuous time parameter, estimators and controllers which are used for their observation and manipulation are often implemented using digital computers. Consequently, those devices are represented by discrete-time systems and thus it is necessary to sample the continuous-time processes to be estimated or controlled. Note that in this thesis only linear systems are used which is also the focus of this section.

We start with a simple deterministic example. In the following, it will be extended towards the case which is relevant for the derivation of the stochastic difference equations that are used to describe discrete-time linear dynamical systems. Consider the following system of homogeneous linear differential equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{x}_t = \boldsymbol{A}_t \boldsymbol{x}_t, \qquad t \in \mathbb{R}_{+,0}, \tag{A1}$$

where the coefficient matrix $\boldsymbol{A}_t \in \mathbb{R}^{N\times N}$ is also a function of t. The initial value of the equation above is $\boldsymbol{x}_0 \in \mathbb{R}^N$. The solution of Equation (A1) can be expressed as (cf. [4, p. 11], [124, p. 20])

$$\boldsymbol{x}_t = \boldsymbol{\Phi}_{t,0}\boldsymbol{x}_0, \qquad t \in \mathbb{R}_{+,0}, \tag{A2}$$

where $\boldsymbol{\Phi}_{t,s} \in \mathbb{R}^{N\times N}$ with $s, t \in \mathbb{R}_{+,0}$ is the so-called *transition matrix* from the time s to t. It is the solution of the matrix differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\Phi}_{t,0} = \boldsymbol{A}_t \boldsymbol{\Phi}_{t,0}, \qquad t \in \mathbb{R}_{+,0}, \tag{A3}$$

with initial value $\boldsymbol{\Phi}_{0,0} = \mathbf{I}_N$. Unfortunately, the transition matrix does not allow for a closed form expression in general but can be determined by an infinite series (see, e. g., [125] and [124, pp. 20-21]). Nevertheless, it has the properties (cf. [4, p. 12])

$$\boldsymbol{\Phi}_{t_1,t_0} = \boldsymbol{\Phi}_{t_1,s}\boldsymbol{\Phi}_{s,t_0}, \qquad t_0,\, t_1,\, s \in \mathbb{R}_{+,0}, \tag{A4}$$

$$\boldsymbol{\Phi}_{t_1,t_0} = \boldsymbol{\Phi}_{t_0,t_1}^{-1}, \qquad t_0,\, t_1 \in \mathbb{R}_{+,0}, \tag{A5}$$

where the inverse always exists since $\boldsymbol{\Phi}_{t_1,t_0}$ is non-singular [126, pp. 171.172], [124, p. 28]. For the case of a constant coefficient matrix, i. e., when $\boldsymbol{A}_t = \boldsymbol{A}$ for all $t \geq 0$, the transition matrix can be expressed using the matrix exponential (see, e. g., [4, p. 13]) and reads as

$$\boldsymbol{\Phi}_{t_1,t_0} = \mathrm{e}^{\boldsymbol{A}(t_1-t_0)}, \qquad t_1, t_0 \in \mathbb{R}_{+,0}, \tag{A6}$$

where the formal notation of a square matrix $\boldsymbol{X}$ in the exponent is defined as

$$\mathrm{e}^{\boldsymbol{X}} = \sum_{k=0}^{\infty} \frac{1}{k!}\boldsymbol{X}^k. \tag{A7}$$

The properties of the matrix exponential are exactly those of a transition matrix [126, p. 170].

Having established the results for homogeneous systems, we now turn to inhomogeneous ones (cf. [4, p. 12], [126, p. 173], [124, Section 1.6]),

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{x}_t = \boldsymbol{A}_t\boldsymbol{x}_t + \boldsymbol{B}_t\boldsymbol{u}_t, \qquad t \in \mathbb{R}_{+,0}, \tag{A8}$$

with the initial value $\boldsymbol{x}_0 \in \mathbb{R}^N$. The equation is denoted analogous to linear systems using the input matrix $\boldsymbol{B}_t$ which is in general a function of t. Note that $\boldsymbol{u}_t$ is assumed to be deterministic at this point. With the transition matrix, the solution of Equation (A8) is given by

$$\boldsymbol{x}_t = \boldsymbol{\Phi}_{t,0}\boldsymbol{x}_0 + \int_0^t \boldsymbol{\Phi}_{t,\tau}\boldsymbol{B}_\tau\boldsymbol{u}_\tau \,\mathrm{d}\tau, \qquad t \in \mathbb{R}_{+,0}, \tag{A9}$$

which can be verified by inserting the result in the differential equation from Equation (A8). It can also be derived by computing

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{\Phi}_{t,0}^{-1}\boldsymbol{x}_t) &= -\boldsymbol{\Phi}_{t,0}^{-1}\left(\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\Phi}_{t,0}\right)\boldsymbol{\Phi}_{t,0}^{-1}\boldsymbol{x}_t + \boldsymbol{\Phi}_{t,0}^{-1}\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{x}_t \\ &= -\boldsymbol{\Phi}_{t,0}^{-1}\boldsymbol{A}_t\boldsymbol{\Phi}_{t,0}\boldsymbol{\Phi}_{t,0}^{-1}\boldsymbol{x}_t + \boldsymbol{\Phi}_{t,0}^{-1}\left(\boldsymbol{A}_t\boldsymbol{x}_t + \boldsymbol{B}_t\boldsymbol{u}_t\right) \\ &= \boldsymbol{\Phi}_{t,0}^{-1}\boldsymbol{B}_t\boldsymbol{u}_t,\end{aligned} \tag{A10}$$

where Equation (A3) is used for the derivative of $\boldsymbol{\Phi}_{t,0}$. Integrating both sides from 0 to t results in

$$\boldsymbol{\Phi}_{t,0}^{-1}\boldsymbol{x}_t - \boldsymbol{\Phi}_{0,0}^{-1}\boldsymbol{x}_0 = \int_0^t \boldsymbol{\Phi}_{\tau,0}^{-1}\boldsymbol{B}_\tau\boldsymbol{u}_\tau \,\mathrm{d}\tau. \tag{A11}$$

Since $\boldsymbol{\Phi}_{0,0} = \mathbf{I}_N$ (cf. Equation A3) and due to the properties given in Equations (A4) and (A5), Equation (A11) is equivalent to Equation (A9).

When dealing with stochastic systems, one could simply consider the model above with a stochastic disturbance, i. e.,

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{x}_t = \boldsymbol{A}_t\boldsymbol{x}_t + \boldsymbol{w}_t, \qquad t \in \mathbb{R}_{+,0}, \tag{A12}$$

where $(\boldsymbol{w}_t : t \in \mathbb{R}_{+,0})$ is a white noise stochastic process with $\mathrm{E}_{\boldsymbol{w}_t}[\boldsymbol{w}_t] = \mathbf{0}_N$, $t \geq 0$, and $\mathrm{E}_{\boldsymbol{w}_s,\boldsymbol{w}_t}\left[\boldsymbol{w}_s\boldsymbol{w}_t^{\mathrm{T}}\right] = \boldsymbol{C}_{\boldsymbol{w}_t}\delta(s-t)$, where δ denotes the Dirac distribution. In this case, the solution of Equation (A12) is (cf. Equation A9)

$$\boldsymbol{x}_t = \boldsymbol{\Phi}_{t,0}\boldsymbol{x}_0 + \int_0^t \boldsymbol{\Phi}_{t,\tau}\boldsymbol{w}_\tau \,\mathrm{d}\tau, \qquad t \in \mathbb{R}_{+,0}. \tag{A13}$$

The expected value of the second summand of this equation is zero due to the linearity of the integration and its covariance matrix is given by[1]

$$\begin{aligned}\mathrm{E}_{\boldsymbol{w}}\left[\int_0^t \boldsymbol{\Phi}_{t,\tau_1}\boldsymbol{w}_{\tau_1}\,\mathrm{d}\tau_1\left(\int_0^t \boldsymbol{\Phi}_{t,\tau_2}\boldsymbol{w}_{\tau_2}\,\mathrm{d}\tau_2\right)^{\mathrm{T}}\right] &= \int_0^t\int_0^t \boldsymbol{\Phi}_{t,\tau_1}\,\mathrm{E}_{\boldsymbol{w}_{\tau_1,\tau_2}}\left[\boldsymbol{w}_{\tau_1}\boldsymbol{w}_{\tau_2}^{\mathrm{T}}\right]\boldsymbol{\Phi}_{t,\tau_2}^{\mathrm{T}}\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2 \\ &= \int_0^t\int_0^t \boldsymbol{\Phi}_{t,\tau_1}\boldsymbol{C}_{\boldsymbol{w}_{\tau_1}}\delta(\tau_1-\tau_2)\boldsymbol{\Phi}_{t,\tau_2}^{\mathrm{T}}\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2 \\ &= \int_0^t \boldsymbol{\Phi}_{t,\tau}\boldsymbol{C}_{\boldsymbol{w}_\tau}\boldsymbol{\Phi}_{t,\tau}^{\mathrm{T}}\,\mathrm{d}\tau.\end{aligned} \tag{A14}$$

[1]The expected value in Equation (A14) is denoted with index $\boldsymbol{w}$ and not $\boldsymbol{w}_\tau$ with $\tau \in [0,t]$ since we deal with an uncountable number of random variables at this point.

Using the abbreviation

$$\boldsymbol{W}_t = \int_0^t \boldsymbol{\Phi}_{t,\tau} \boldsymbol{C}_{\boldsymbol{w}_\tau} \boldsymbol{\Phi}_{t,\tau}^{\mathrm{T}} \,\mathrm{d}\tau, \tag{A15}$$

the result of the equation above is determined by the matrix differential equation (cf. [124, pp. 60-61] and [3, p. 66])

$$\frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{W}_t = \boldsymbol{A}_t \boldsymbol{W}_t + \boldsymbol{W}_t \boldsymbol{A}_t^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{w}_t} \tag{A16}$$

with initial value $\boldsymbol{W}_0 = \mathbf{0}_{N \times N}$, which can be verified by inserting Equation (A15) in Equation (A16).[2]

Although the result is correct (cf. [4, p. 470]), it is derived without taking into account the stochastic nature of $(\boldsymbol{w}_t : t \in \mathbb{R}_{+,0})$, which makes it necessary to introduce stochastic differential equations and the concept of calculus for stochastic processes. The corresponding theory is out of the scope of this section. A detailed description of the rigorous approach using stochastic differential equations in the form of

$$\mathrm{d}\boldsymbol{x}_t = \boldsymbol{A}_t \boldsymbol{x}_t \,\mathrm{d}t + \mathrm{d}\boldsymbol{\eta}_t, \tag{A17}$$

where $(\boldsymbol{\eta}_t : t \in \mathbb{R}_{+,0})$ is a Wiener process with differential covariance matrix $\boldsymbol{C}_{\boldsymbol{\eta}_t} \,\mathrm{d}t$ can be found in, e. g., [3, Chapter 3].

Finally, we are in the position to derive the sampled version of the stochastic differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{x}_t = \boldsymbol{A}_t \boldsymbol{x}_t + \boldsymbol{B}_t \boldsymbol{u}_t + \boldsymbol{w}_t, \qquad t \in \mathbb{R}_{+,0}, \tag{A18}$$

with initial condition[3] $\boldsymbol{x}_0 \in \mathbb{R}^N$ and where $(\boldsymbol{w}_t : t \in \mathbb{R}_{+,0})$ is zero mean[4] white Gaussian noise, i. e., $\boldsymbol{w}_t \sim \mathcal{N}\left(\mathbf{0}_N, \boldsymbol{C}_{\boldsymbol{w}_t}\right)$ and $\mathrm{E}_{\boldsymbol{w}_s,\boldsymbol{w}_t}\left[\boldsymbol{w}_s \boldsymbol{w}_t^{\mathrm{T}}\right] = \boldsymbol{C}_{\boldsymbol{w}_t} \delta(s-t)$. We assume that Equation (A18) is sampled at equidistant time indices $t = kT$, $k \in \mathbb{N}_0$, where $T > 0$ is the sampling period, and use the notation

$$\boldsymbol{x}_k = \boldsymbol{x}_{kT}, \qquad k \in \mathbb{N}_0, \tag{A19}$$

where the ambiguity that $(\boldsymbol{x}_k : k \in \mathbb{N}_0)$ is a random sequence and $(\boldsymbol{x}_t : t \in \mathbb{R}_{+,0})$ is a random process of a continuous parameter should be clear from the context. Additionally, it is assumed that $\boldsymbol{u}_t$, $t \in \mathbb{R}_{+,0}$, which is interpreted as the control input of a dynamical system, is constant within one sampling period, i. e.,

$$\boldsymbol{u}_t = \boldsymbol{u}_k, \qquad t \in [kT, (k+1)T[\quad \text{with} \quad k \in \mathbb{N}_0. \tag{A20}$$

This is one possibility to perform the "discrete-to-continuous" conversion which is the counterpart of the sampling and is commonly referred to as *zero order hold*. Using the results of Equations

[2]This result is of particular interest for constant matrices $\boldsymbol{A}_t = \boldsymbol{A}$ and $\boldsymbol{C}_{\boldsymbol{w}_t} = \boldsymbol{C}_{\boldsymbol{w}}$ for all $t \in \mathbb{R}_{+,0}$. In this case, the differential equation has a close relation to the investigation of the stability of linear systems. For the equilibrium solution with $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{W}_t = \mathbf{0}_{N \times N}$, Equation (A16) is called the continuous-time Lyapunov equation.

[3]The initial value can also be a random vector. This case can be incorporated in the derivations above without difficulties, see, e. g., [3, Section 3.6].

[4]The assumption that the stochastic process has zero mean is not restrictive since a different expected value can be considered by simply adding a constant vector.

(A9) and (A13), we get for the time index $(k+1)T$ with $k \in \mathbb{N}_0$

$$\begin{aligned}
\boldsymbol{x}_{k+1} &= \boldsymbol{\Phi}_{(k+1)T,0}\boldsymbol{x}_0 + \int_0^{(k+1)T} \boldsymbol{\Phi}_{(k+1)T,\tau}\boldsymbol{B}_\tau \boldsymbol{u}_\tau \,\mathrm{d}\tau + \int_0^{(k+1)T} \boldsymbol{\Phi}_{(k+1)T,\tau}\boldsymbol{w}_\tau \,\mathrm{d}\tau \\
&= \boldsymbol{\Phi}_{(k+1)T,kT}\left(\boldsymbol{\Phi}_{kT,0}\boldsymbol{x}_0 + \int_0^{kT} \boldsymbol{\Phi}_{kT,\tau}\boldsymbol{B}_\tau \boldsymbol{u}_\tau \,\mathrm{d}\tau + \int_0^{kT} \boldsymbol{\Phi}_{kT,\tau}\boldsymbol{w}_\tau \,\mathrm{d}\tau\right) \\
&\quad + \int_{kT}^{(k+1)T} \boldsymbol{\Phi}_{(k+1)T,\tau}\boldsymbol{B}_\tau \boldsymbol{u}_\tau \,\mathrm{d}\tau + \int_{kT}^{(k+1)T} \boldsymbol{\Phi}_{(k+1)T,\tau}\boldsymbol{w}_\tau \,\mathrm{d}\tau \\
&= \boldsymbol{\Phi}_{(k+1)T,kT}\boldsymbol{x}_k + \int_{kT}^{(k+1)T} \boldsymbol{\Phi}_{(k+1)T,\tau}\boldsymbol{B}_\tau \,\mathrm{d}\tau \boldsymbol{u}_k + \int_{kT}^{(k+1)T} \boldsymbol{\Phi}_{(k+1)T,\tau}\boldsymbol{w}_\tau \,\mathrm{d}\tau \\
&= \hat{\boldsymbol{A}}_k \boldsymbol{x}_k + \hat{\boldsymbol{B}}_k \boldsymbol{u}_k + \hat{\boldsymbol{w}}_k,
\end{aligned} \tag{A21}$$

where we also applied the property of the transition matrix from Equation (A4). In the last line of Equation (A21), one can identify the parameters of the discrete-time system which is obtained by sampling the corresponding differential equation as

$$\begin{aligned}
\hat{\boldsymbol{A}}_k &= \boldsymbol{\Phi}_{(k+1)T,kT}, \\
\hat{\boldsymbol{B}}_k &= \int_{kT}^{(k+1)T} \boldsymbol{\Phi}_{(k+1)T,\tau}\boldsymbol{B}_\tau \,\mathrm{d}\tau, \qquad k \in \mathbb{N}_0,
\end{aligned} \tag{A22}$$

while the parameters of the discrete-time noise sequence $(\hat{\boldsymbol{w}}_k : k \in \mathbb{N}_0)$ are

$$\hat{\boldsymbol{w}}_k \sim \mathcal{N}\left(\boldsymbol{0}_N, \boldsymbol{C}_{\hat{\boldsymbol{w}}_k}\right), \tag{A23}$$

$$\mathrm{E}_{\hat{\boldsymbol{w}}_m,\hat{\boldsymbol{w}}_n}\left[\hat{\boldsymbol{w}}_m \hat{\boldsymbol{w}}_n^{\mathrm{T}}\right] = \begin{cases} \boldsymbol{C}_{\hat{\boldsymbol{w}}_n} & m = n, \\ \boldsymbol{0}_{N\times N} & \text{otherwise}, \end{cases} \tag{A24}$$

$$\boldsymbol{C}_{\hat{\boldsymbol{w}}_k} = \int_{kT}^{(k+1)T} \boldsymbol{\Phi}_{(k+1)T,\tau}\boldsymbol{C}_{\boldsymbol{w}_\tau}\boldsymbol{\Phi}_{(k+1)T,\tau}^{\mathrm{T}} \,\mathrm{d}\tau. \tag{A25}$$

These properties follow directly from the fact that $(\boldsymbol{w}_t : t \in \mathbb{R}_{+,0})$ is a Gaussian white noise process. The linear operation of integration does not change the type of distribution and the zero mean (Equation A23). The intervals for the integration of the continuous-time noise which results in $\hat{\boldsymbol{w}}_k = \int_{kT}^{(k+1)T} \boldsymbol{\Phi}_{(k+1)T,\tau}\boldsymbol{w}_\tau \,\mathrm{d}\tau$ are disjoint for different values of $k \in \mathbb{N}_0$. Consequently, the discrete-time noise is an uncorrelated sequence of random variables (cf. Equation A24). Finally, the covariance of the noise sequence is given by Equation (A25), which results from the derivations in Equations (A12)-(A14).

A2. The Discrete Lyapunov Equation

Let $\boldsymbol{A} \in \mathbb{R}^{N\times N}$ and $\boldsymbol{Q} \in \mathbb{R}^{N\times N}$ be constant matrices and $\boldsymbol{X} \in \mathbb{R}^{N\times N}$ be a matrix variable. The discrete Lyapunov equation is defined as

$$\boldsymbol{X} = \boldsymbol{A}\boldsymbol{X}\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{Q}. \tag{A26}$$

The main points associated with this equation are the conditions for the existence of a unique solution $\boldsymbol{X}$ and, if it exists, its properties. A detailed discussion about the discrete Lyapunov equation can be found in, e. g., [31, Appendix D]. There are several approaches for the determination of a solution of the equation, cf. [127]. For the *direct method*, the Lyapunov equation (A26) is rewritten as a system of linear equations, i. e.,

$$\begin{aligned} \operatorname{vec}[\boldsymbol{X}] &= (\boldsymbol{A}\otimes\boldsymbol{A})\operatorname{vec}[\boldsymbol{X}] + \operatorname{vec}[\boldsymbol{Q}] \\ \Leftrightarrow (\mathbf{I}_{N^2} - (\boldsymbol{A}\otimes\boldsymbol{A}))\operatorname{vec}[\boldsymbol{X}] &= \operatorname{vec}[\boldsymbol{Q}], \end{aligned} \tag{A27}$$

where vec is the operator which stacks the columns of $\boldsymbol{X}$ in a column vector of dimension N^2, and $\otimes$ denotes the Kronecker product (cf. [128]). Since the eigenvalues of $\boldsymbol{A}\otimes\boldsymbol{A}$ are $\lambda_i\lambda_j$, $i,j\in\{1,2,\ldots,N\}$, where λ_i, $i\in\{1,2,\ldots,N\}$, are the eigenvalues of $\boldsymbol{A}$ (cf. [126, pp. 234-235]), we see that a unique solution only exists if $\lambda_i\lambda_j \neq 1$, $\forall i,j$. For the case when $\lambda_i\lambda_j = 1$ for some i,j, there exists no solution if $\operatorname{vec}[\boldsymbol{Q}]$ is not an element of the space spanned by the columns of $(\mathbf{I}_{N^2} - (\boldsymbol{A}\otimes\boldsymbol{A}))$, otherwise there exist infinitely many solutions.

The *iterative* method can be used if all eigenvalues of $\boldsymbol{A}$ have magnitude less than one. Note that in this case, the solution is always unique since $|\lambda_i\lambda_j| < 1$, $i,j\in\{1,2,\ldots,N\}$. Then, the solution of Equation (A26) is obtained by the iteration

$$\boldsymbol{X}_{k+1} = \boldsymbol{A}\boldsymbol{X}_k\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{Q}, \tag{A28}$$

which converges to (cf. [5, pp. 63-64])

$$\boldsymbol{X} = \sum_{k=0}^{\infty} \boldsymbol{A}^k\boldsymbol{Q}\boldsymbol{A}^{\mathrm{T},k}. \tag{A29}$$

From the result in Equation (A29) it can be seen that the unique solution $\boldsymbol{X}$ is symmetric if $\boldsymbol{Q}$ is symmetric and that $\boldsymbol{X}$ is positive (semi)definite if $\boldsymbol{Q}$ is positive (semi)definite.[5]

Finally, although not the best approach from a numerical perspective, the unique solution of the Lyapunov equation (A26) when all eigenvalues of $\boldsymbol{A}$ have magnitude less than one and $\boldsymbol{Q}$ is positive semidefinite can be determined by the convex program

$$\underset{\boldsymbol{Y}}{\operatorname{minimize}}\ \operatorname{tr}[\boldsymbol{Y}] \qquad \text{subject to} \qquad \begin{bmatrix} \boldsymbol{Y}-\boldsymbol{Q} & \boldsymbol{A}\boldsymbol{Y} \\ \boldsymbol{Y}\boldsymbol{A}^{\mathrm{T}} & \boldsymbol{Y} \end{bmatrix} \geq \mathbf{0}_{2N\times 2N}, \tag{A30}$$

where the fact is used that in this case the linear matrix inequality[6]

$$\boldsymbol{Y} \geq \boldsymbol{A}\boldsymbol{Y}\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{Q} \tag{A31}$$

can be rewritten as a positive semidefinite Schur complement[7] which is equivalent to the constraint of the above optimization problem. Additionally, any positive semidefinite matrix $\boldsymbol{Y}$ which fulfills the inequality (A31) has the property[8] $\boldsymbol{Y}\geq\boldsymbol{X}$, where $\boldsymbol{X}$ is the solution of Equation (A26).

[5]More generally, the solution $\boldsymbol{X}$ is positive definite if the pair $(\boldsymbol{A},\boldsymbol{Q}^{\frac{1}{2}})$ is controllable, see [31, Appendix D].

[6]For a definition and detailed treatment of linear matrix inequalities, see, e. g., [129]. An important feature is that linear matrix inequalities define convex sets which can be handled by numerical algorithms for the solution of convex optimization problems.

[7]For properties of the Schur complement in terms of definiteness, see, e. g., [94, Appendix A.5.5].

[8]If $\boldsymbol{Y}\geq\boldsymbol{A}\boldsymbol{Y}\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{Q}$, it holds $\boldsymbol{Y}=\boldsymbol{A}\boldsymbol{Y}\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{Q}+\boldsymbol{\Phi}$, where $\boldsymbol{\Phi}$ is a positive semidefinite matrix of appropriate dimensions. Thus, $\boldsymbol{Y}=\boldsymbol{Y}_1+\boldsymbol{Y}_2$, where $\boldsymbol{Y}_1=\boldsymbol{A}\boldsymbol{Y}_1\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{Q}$ and $\boldsymbol{Y}_2=\boldsymbol{A}\boldsymbol{Y}_2\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{\Phi}$ are the positive semidefinite solutions of two Lyapunov equations. Since the solutions are unique, we have $\boldsymbol{Y}_1=\boldsymbol{X}$, which shows that $\boldsymbol{Y}\geq\boldsymbol{X}$.

A3. The Discrete Algebraic Riccati Equation

Let $\boldsymbol{A} \in \mathbb{R}^{N\times N}$, $\boldsymbol{B} \in \mathbb{R}^{N\times M}$, $\boldsymbol{Q} \in \mathbb{R}^{N\times N}$, $\boldsymbol{R} \in \mathbb{R}^{M\times M}$ and $\boldsymbol{S} \in \mathbb{R}^{N\times M}$ be constant matrices with

$$\boldsymbol{Q} \geq \boldsymbol{0}_{N\times N}, \qquad \boldsymbol{R} > \boldsymbol{0}_{M\times M}, \qquad \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{S} \\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{R} \end{bmatrix} \geq \boldsymbol{0}_{(N+M)\times(N+M)}. \tag{A32}$$

Then, the associated Discrete Algebraic Riccati Equation (DARE) is given by

$$\boldsymbol{X} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{A} - \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{B} + \boldsymbol{S}\right)\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{B} + \boldsymbol{R}\right)^{-1}\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{A} + \boldsymbol{S}^{\mathrm{T}}\right) + \boldsymbol{Q}, \tag{A33}$$

with the matrix variable $\boldsymbol{X} \in \mathbb{R}^{N\times N}$. Note that with the substitution

$$\tilde{\boldsymbol{A}} = \boldsymbol{A} - \boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{S}^{\mathrm{T}}, \tag{A34}$$

the equation above can be rewritten as (cf. [130])

$$\boldsymbol{X} = \tilde{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{X}\tilde{\boldsymbol{A}} - \tilde{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{B}\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{B} + \boldsymbol{R}\right)^{-1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X}\tilde{\boldsymbol{A}} + \boldsymbol{Q} - \boldsymbol{S}\boldsymbol{R}^{-1}\boldsymbol{S}^{\mathrm{T}}. \tag{A35}$$

Unlike the discrete Lyapunov equation (see Section A2), the DARE is a non-linear matrix equation and it is more involved to discuss the conditions for the existence of a solution and of its properties.

Especially for the purposes of estimation and control, we are interested in positive semidefinite and stabilizing solutions, i. e., positive semidefinite matrices $\boldsymbol{X}$ which fulfill Equation (A33) such that the matrix

$$\boldsymbol{A}_{\mathrm{cl}} = \boldsymbol{A} + \boldsymbol{B}\boldsymbol{L}, \tag{A36}$$

where

$$\boldsymbol{L} = -\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{B} + \boldsymbol{R}\right)^{-1}\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{A} + \boldsymbol{S}^{\mathrm{T}}\right), \tag{A37}$$

has all eigenvalues inside the complex unit disk, i. e., with magnitude less than one. An important fact is that if a stabilizing solution exists, it is positive semidefinite and *unique* [31, p. 775]. There are several results which state the existence of stabilizing solutions, positive definite solutions and the uniqueness of these solutions. We will not discuss these results and the corresponding proofs at this point but refer to the detailed treatment of the topic in [31, Appendix E and Chapter 14] and also [6, Chapter 4].

In many cases, the following result is of most interest. Recall the substitution given in Equation (A34) and additionally define

$$\tilde{\boldsymbol{Q}} = \boldsymbol{Q} - \boldsymbol{S}\boldsymbol{R}^{-1}\boldsymbol{S}^{\mathrm{T}}, \tag{A38}$$

where $\tilde{\boldsymbol{Q}} \geq \boldsymbol{0}_{N\times N}$ due to the assumptions of Equation (A32). The DARE (A33) has a unique positive semidefinite solution if and only if the pair $(\boldsymbol{A}, \boldsymbol{B})$ is stabilizable and the pair $(\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{Q}}^{\frac{1}{2}})$ is detectable [31, Appendix E], in which case this solution is also stabilizing. This unique solution can be obtained in several ways. The simplest one is using the iteration

$$\boldsymbol{X}_{k+1} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{X}_k\boldsymbol{A} - \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{X}_k\boldsymbol{B} + \boldsymbol{S}\right)\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X}_k\boldsymbol{B} + \boldsymbol{R}\right)^{-1}\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X}_k\boldsymbol{A} + \boldsymbol{S}^{\mathrm{T}}\right) + \boldsymbol{Q} \tag{A39}$$

with an arbitrary positive semidefinite initialization $\boldsymbol{X}_0 \geq \boldsymbol{0}_{N\times N}$. Under the conditions mentioned above, this iteration converges to the positive semidefinite stabilizing solution

$$\boldsymbol{X} = \lim_{k\to\infty} \boldsymbol{X}_k, \tag{A40}$$

irrespective of the (positive semidefinite) initialization (cf. [6, Section 4.1] and [31, Chapter 14]). A more efficient and numerically stable method for the computation of the stabilizing solution of the DARE is the so-called Schur or invariant subspace method (cf. [130] and [31, Appendix E.7]). Finally, the solution of Equation (A33) can also be determined by the convex program

$$\underset{\boldsymbol{Y}}{\operatorname{maximize}}\ \operatorname{tr}[\boldsymbol{Y}] \quad \text{subject to} \quad \begin{bmatrix} \boldsymbol{A}^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{A}-\boldsymbol{Y}+\boldsymbol{Q} & \boldsymbol{A}^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{B}+\boldsymbol{S} \\ \boldsymbol{B}^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{A}+\boldsymbol{S}^{\mathrm{T}} & \boldsymbol{B}^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{B}+\boldsymbol{R} \end{bmatrix} \geq \mathbf{0}_{(N+M)\times(N+M)}, \quad \boldsymbol{Y} \geq \mathbf{0}_{N\times N}, \tag{A41}$$

because, under the stabilizability and detectability assumption, the stabilizing solution of the DARE is also the maximal[9] element (cf. [31, Appendix E.3] and [100]) of the set of positive semidefinite matrices $\boldsymbol{X}$ which satisfy the Riccati inequality

$$\boldsymbol{X} \leq \boldsymbol{A}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{A} - \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{B}+\boldsymbol{S}\right)\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{B}+\boldsymbol{R}\right)^{-1}\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{A}+\boldsymbol{S}^{\mathrm{T}}\right)+\boldsymbol{Q}. \tag{A42}$$

Note that the condition $\boldsymbol{R} > \mathbf{0}_{M\times M}$ (cf. Equation A32) is not necessary for the existence of a positive semidefinite and stabilizing solution of the DARE and can be relaxed to $\boldsymbol{R} \geq \mathbf{0}_{M\times M}$. The only requirement to be fulfilled is that the matrix $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{B}+\boldsymbol{R}$ is invertible (cf. Equation A33). In this case, the optimization-based solution approach given in Equation (A41) can still be applied, and the Schur method can be modified to deal with singular matrices $\boldsymbol{R}$. However, in this case the stabilizability of $(\boldsymbol{A},\boldsymbol{B})$ and, e. g., for $\boldsymbol{S}=\mathbf{0}_{N\times M}$, detectability of $(\boldsymbol{A},\boldsymbol{Q}^{\frac{1}{2}})$ do not guarantee a unique positive definite solution which is also stabilizing. Consider the simple counterexample with $M = 1$, $\boldsymbol{Q}=\boldsymbol{c}\boldsymbol{c}^{\mathrm{T}}$ and $r = 0$ (the matrix $\boldsymbol{R}$ reduces to a scalar here), where $(\boldsymbol{A},\boldsymbol{b})$ is stabilizable and $(\boldsymbol{A},\boldsymbol{c})$ is detectable and with $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{b} \neq 0$. The corresponding DARE reads as

$$\boldsymbol{X} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{A} - \boldsymbol{A}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{b}(\boldsymbol{b}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{b})^{-1}\boldsymbol{b}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{A}+\boldsymbol{c}\boldsymbol{c}^{\mathrm{T}}, \tag{A43}$$

and it is easy to verify that $\boldsymbol{X} = \boldsymbol{c}\boldsymbol{c}^{\mathrm{T}}$ is a positive semidefinite solution. Nevertheless, the closed loop matrix (cf. Equations A36 and A37)

$$\boldsymbol{A}_{\mathrm{cl}} = \boldsymbol{A} - (\boldsymbol{b}^{\mathrm{T}}\boldsymbol{c})^{-1}\boldsymbol{b}\boldsymbol{c}^{\mathrm{T}}\boldsymbol{A} \tag{A44}$$

has in general eigenvalues with magnitude larger than one.[10] A more detailed treatment of the case when $\boldsymbol{R}$ is singular can be found in, e. g., [131]. One result is that if the corresponding DARE has a stabilizing solution, it is unique. The authors of [131] also provide a necessary and sufficient condition for the existence which is more involved than the stabilizability and detectability conditions for the non-singular case and which is checked in the process of computing the stabilizing solution. A further discussion of this topic and the analysis of more general Riccati equations and inequalities is provided by [132].

As a final remark, note that for the special case of $\boldsymbol{Q} = \mathbf{0}_{N\times N}$ and $\boldsymbol{S} = \mathbf{0}_{N\times M}$, the requirement of detectability of $(\boldsymbol{A},\tilde{\boldsymbol{Q}})$ is obviously not fulfilled. Consequently, the existence of a unique positive semidefinite stabilizing solution of the DARE is not guaranteed. Nevertheless, as long as $(\boldsymbol{A},\boldsymbol{B})$ is stabilizable, there exists at least one positive semidefinite solution which is semi-stabilizing, i. e., the corresponding matrix $\boldsymbol{A}_{\mathrm{cl}}$ given by Equation (A36) has no eigenvalues with magnitude larger than one (see [31, p. 775]).

[9] The maximal element $\boldsymbol{X}_+$ which satisfies inequality (A42) has the property that $\boldsymbol{X}_+ \geq \boldsymbol{X}$ for all positive semidefinite matrices $\boldsymbol{X}$ which satisfy inequality (A42).

[10] The eigenvalues of $\boldsymbol{A}_{\mathrm{cl}}$ which is given in Equation (A44) can be identified as the zeros of the transfer function of a linear dynamical system described by $\boldsymbol{A}$, $\boldsymbol{b}$ and $\boldsymbol{c}$, cf. [119].

A4. Derivative of the Discrete Algebraic Riccati Equation

Let $\boldsymbol{X}$ be the stabilizing positive semidefinite solution of the DARE in Equation (A33) and let the matrix $\boldsymbol{L}$ be given according to Equation (A37), i. e., the magnitude of all eigenvalues of $\boldsymbol{A}_{\mathrm{cl}} = \boldsymbol{A} + \boldsymbol{B}\boldsymbol{L}$ is less than one. Additionally, let the parameters $\boldsymbol{Q}$ and $\boldsymbol{R}$ be functions of a scalar variable β. Then, the derivative of $\boldsymbol{X}$ w.r.t. β is given by

$$\begin{aligned}\frac{\partial \boldsymbol{X}}{\partial \beta} &= \boldsymbol{A}^{\mathrm{T}}\frac{\partial \boldsymbol{X}}{\partial \beta}\boldsymbol{A} + \boldsymbol{A}^{\mathrm{T}}\frac{\partial \boldsymbol{X}}{\partial \beta}\boldsymbol{B}\boldsymbol{L} + \boldsymbol{L}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}}\frac{\partial \boldsymbol{X}}{\partial \beta}\boldsymbol{A} + \boldsymbol{L}^{\mathrm{T}}\left(\boldsymbol{B}^{\mathrm{T}}\frac{\partial \boldsymbol{X}}{\partial \beta}\boldsymbol{B} + \frac{\partial \boldsymbol{R}}{\partial \beta}\right)\boldsymbol{L} + \frac{\partial \boldsymbol{Q}}{\partial \beta} \\ &= \boldsymbol{A}_{\mathrm{cl}}^{\mathrm{T}}\frac{\partial \boldsymbol{X}}{\partial \beta}\boldsymbol{A}_{\mathrm{cl}} + \boldsymbol{L}^{\mathrm{T}}\frac{\partial \boldsymbol{R}}{\partial \beta}\boldsymbol{L} + \frac{\partial \boldsymbol{Q}}{\partial \beta},\end{aligned} \tag{A45}$$

which is a discrete Lyapunov equation (see Appendix A2). Since the magnitude of all eigenvalues of $\boldsymbol{A}_{\mathrm{cl}}$ is less than one, this equation has a unique solution. Note that the solution is symmetric if $\boldsymbol{L}^{\mathrm{T}}\frac{\partial \boldsymbol{R}}{\partial \beta}\boldsymbol{L} + \frac{\partial \boldsymbol{Q}}{\partial \beta}$ is a symmetric matrix.

Before computing the second derivative of $\boldsymbol{X}$ w.r.t. β, we determine the derivative

$$\begin{aligned}\frac{\partial \boldsymbol{L}}{\partial \beta} &= -\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{B} + \boldsymbol{R}\right)^{-1}\left(\boldsymbol{B}^{\mathrm{T}}\frac{\partial \boldsymbol{X}}{\partial \beta}\boldsymbol{B} + \frac{\partial \boldsymbol{R}}{\partial \beta}\right)\boldsymbol{L} - \left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{B} + \boldsymbol{R}\right)^{-1}\boldsymbol{B}^{T}\frac{\partial \boldsymbol{X}}{\partial \beta}\boldsymbol{A} \\ &= -\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{B} + \boldsymbol{R}\right)^{-1}\left(\boldsymbol{B}^{\mathrm{T}}\frac{\partial \boldsymbol{X}}{\partial \beta}\boldsymbol{A}_{\mathrm{cl}} + \frac{\partial \boldsymbol{R}}{\partial \beta}\boldsymbol{L}\right).\end{aligned} \tag{A46}$$

The second derivative of $\boldsymbol{X}$ w.r.t. β reads as

$$\begin{aligned}\frac{\partial^2 \boldsymbol{X}}{\partial \beta^2} &= \boldsymbol{A}_{\mathrm{cl}}^{\mathrm{T}}\frac{\partial^2 \boldsymbol{X}}{\partial \beta^2}\boldsymbol{A}_{\mathrm{cl}} + \frac{\partial \boldsymbol{L}^{\mathrm{T}}}{\partial \beta}\boldsymbol{B}\frac{\partial \boldsymbol{X}}{\partial \beta}\boldsymbol{A}_{\mathrm{cl}} + \boldsymbol{A}_{\mathrm{cl}}^{\mathrm{T}}\frac{\partial \boldsymbol{X}}{\partial \beta}\boldsymbol{B}\frac{\partial \boldsymbol{L}}{\partial \beta} \\ &\qquad + \frac{\partial \boldsymbol{L}^{\mathrm{T}}}{\partial \beta}\frac{\partial \boldsymbol{R}}{\partial \beta}\boldsymbol{L} + \boldsymbol{L}^{\mathrm{T}}\frac{\partial \boldsymbol{R}}{\partial \beta}\frac{\partial \boldsymbol{L}}{\partial \beta} + \boldsymbol{L}^{\mathrm{T}}\frac{\partial^2 \boldsymbol{R}}{\partial \beta^2}\boldsymbol{L} + \frac{\partial^2 \boldsymbol{Q}}{\partial \beta^2} \\ &= \boldsymbol{A}_{\mathrm{cl}}^{\mathrm{T}}\frac{\partial^2 \boldsymbol{X}}{\partial \beta^2}\boldsymbol{A}_{\mathrm{cl}} - 2\left(\boldsymbol{A}_{\mathrm{cl}}^{\mathrm{T}}\frac{\partial \boldsymbol{X}}{\partial \beta}\boldsymbol{B} + \boldsymbol{L}^{\mathrm{T}}\frac{\partial \boldsymbol{R}}{\partial \beta}\right)\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{B} + \boldsymbol{R}\right)^{-1}\left(\boldsymbol{B}^{\mathrm{T}}\frac{\partial \boldsymbol{X}}{\partial \beta}\boldsymbol{A}_{\mathrm{cl}} + \frac{\partial \boldsymbol{R}}{\partial \beta}\boldsymbol{L}\right) \\ &\qquad + \boldsymbol{L}^{\mathrm{T}}\frac{\partial^2 \boldsymbol{R}}{\partial \beta^2}\boldsymbol{L} + \frac{\partial^2 \boldsymbol{Q}}{\partial \beta^2},\end{aligned} \tag{A47}$$

where Equation (A46) has been used for the last line of the equation above. Like Equation (A45), (A47) has a unique solution.

Note that if $\boldsymbol{Q}$ and $\boldsymbol{R}$ are affine functions of β, i. e., $\boldsymbol{Q} = \boldsymbol{Q}_1 + \beta\boldsymbol{Q}_2$ and $\boldsymbol{R} = \boldsymbol{R}_1 + \beta\boldsymbol{R}_2$ where $\boldsymbol{Q}_1$, $\boldsymbol{Q}_2$, $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ are positive semidefinite and $\beta \geq 0$, $\frac{\partial \boldsymbol{X}}{\partial \beta}$ is positive semidefinite whereas $\frac{\partial^2 \boldsymbol{X}}{\partial \beta^2}$ is negative semidefinite. This is easy to verify because $\boldsymbol{L}^{\mathrm{T}}\frac{\partial \boldsymbol{R}}{\partial \beta}\boldsymbol{L} + \frac{\partial \boldsymbol{Q}}{\partial \beta} = \boldsymbol{L}^{\mathrm{T}}\boldsymbol{R}_2\boldsymbol{L} + \boldsymbol{Q}_2 \geq \mathbf{0}_{N\times N}$, which implies the positive semidefiniteness of the first derivative. For the second derivative, note that $\boldsymbol{L}^{\mathrm{T}}\frac{\partial^2 \boldsymbol{R}}{\partial \beta^2}\boldsymbol{L} + \frac{\partial^2 \boldsymbol{Q}}{\partial \beta^2} = \mathbf{0}_{N\times N}$ and that $\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{B} + \boldsymbol{R}\right) > \mathbf{0}_{M\times M}$, where N and M correspond to the dimensions of $\boldsymbol{Q}$ and $\boldsymbol{R}$ shown in Equation (A32). Due to the negative sign, the second derivative of $\boldsymbol{X}$ w.r.t. β is negative semidefinite.

A5. Properties of Expected Values

A5.1 Expected Value of a Function of Random Variables

Let $\boldsymbol{x}$ be a random vector, g a measurable function and $\boldsymbol{y} = g(\boldsymbol{x})$ be the random vector which is obtained by applying g to $\boldsymbol{x}$. Then,

$$\mathrm{E}_{\boldsymbol{y}}\left[\boldsymbol{y}\right] = \mathrm{E}_{\boldsymbol{x}}\left[g(\boldsymbol{x})\right]. \tag{A48}$$

Proof. This is a standard result and can be found, e. g., in [34, pp. 142-143]. □

A5.2 Conditional Expected Value and Conditional Expectation

Although using the same notation, we distinguish between the conditional *expected value* and the conditional *expectation* of a random variable. Let $\boldsymbol{x}$, $\boldsymbol{y}$ be random vectors with a joint distribution. Then, the conditional expected value of $\boldsymbol{x}$ given the *event*[11] $\{\boldsymbol{y} = \boldsymbol{\eta}\}$, where $\boldsymbol{\eta}$ is a constant vector, is a deterministic vector given by the expression (cf. [34, p. 231])

$$\mathrm{E}_{\boldsymbol{x}|\boldsymbol{y}}\left[\boldsymbol{x}|\, \boldsymbol{\eta}\right] = \int \boldsymbol{x} f_{\boldsymbol{x}|\boldsymbol{y}}\left(\boldsymbol{x}|\boldsymbol{\eta}\right) \mathrm{d}\boldsymbol{x}. \tag{A49}$$

Thus, the conditional expected value can be interpreted as the value of a function

$$\varphi(\boldsymbol{\eta}) = \mathrm{E}_{\boldsymbol{x}|\boldsymbol{y}}\left[\boldsymbol{x}|\, \boldsymbol{\eta}\right] \tag{A50}$$

evaluated at $\boldsymbol{\eta}$.

In contrast, the conditional expectation is itself a random variable. A straightforward definition of this random variable is obtained by replacing the realization $\boldsymbol{\eta}$ in Equation (A50) by the random vector $\boldsymbol{y}$, i. e., the random vector $\boldsymbol{z}$ which is given by

$$\boldsymbol{z} = \mathrm{E}_{\boldsymbol{x}|\boldsymbol{y}}\left[\boldsymbol{x}|\, \boldsymbol{y}\right] = \varphi(\boldsymbol{y}) \tag{A51}$$

is the conditional expectation of $\boldsymbol{x}$ given $\boldsymbol{y}$.

The rigorous definition (cf., e. g., [33, p. 347]) of the conditional expectation is as follows: Let $(\Omega, \mathbb{F}, P)$ be a probability space and x a random variable[12] on this space. Additionally, let $\mathbb{G} \subset \mathbb{F}$ be a sub σ-algebra of $\mathbb{F}$. Then, a random variable z which is measurable w.r.t. $\mathbb{G}$ is called conditional expectation of x given $\mathbb{G}$ if

$$\int_B x(\omega)\, \mathrm{d}P(\omega) = \int_B z(\omega)\, \mathrm{d}P(\omega), \qquad \forall B \in \mathbb{G}. \tag{A52}$$

Using the indicator function[13] $\mathbb{I}_B(\omega)$, this equation can be rewritten in terms of expected values[14]:

$$\mathrm{E}\left[x\, \mathbb{I}_B\right] = \mathrm{E}\left[z\, \mathbb{I}_B\right], \qquad \forall B \in \mathbb{G}. \tag{A53}$$

[11] The distinction between $\boldsymbol{y}$ and $\boldsymbol{\eta}$ should emphasize the fact that $\boldsymbol{y}$ is a random vector while $\boldsymbol{\eta}$ is a realization.

[12] The definition of the conditional expectation can be extended to random vectors if the components are treated separately and the σ-algebra which represents the condition is generated jointly by all random variables that are the components of the given random vector.

[13] $\mathbb{I}_B(\omega) = 1$ if $\omega \in B$ and 0 otherwise.

[14] The subscripts of the expectation operators have been dropped to emphasize that x and z are defined on the same probability space and that the expected value is determined using the same probability measure P, cf. Equation (A52).

The random variable z is denoted by $z = \mathrm{E}_{x|\mathbb{G}}\left[x|\,\mathbb{G}\right]$. In order to be consistent with the notation introduced in Equation (A51), let $\mathbb{G}$ be the σ-algebra generated by the random variable y which maps $(\Omega, \mathbb{F})$ to $(\Omega', \mathbb{F}')$, i. e.,

$$\mathbb{G} = \{B|B = y^{-1}(A), A \in \mathbb{F}'\}. \tag{A54}$$

Keeping this in mind, we obtain

$$\mathrm{E}_{x|y}\left[x|\,y\right] = \mathrm{E}_{x|\mathbb{G}}\left[x|\,\mathbb{G}\right]. \tag{A55}$$

An important property of the conditional expectation is the following: let $\boldsymbol{x}$ and $\boldsymbol{y}$ be two random vectors with a joint distribution. Then,

$$\mathrm{E}_{\boldsymbol{y}}\left[\mathrm{E}_{\boldsymbol{x}|\boldsymbol{y}}\left[\boldsymbol{x}|\,\boldsymbol{y}\right]\right] = \mathrm{E}_{\boldsymbol{x}}\left[\boldsymbol{x}\right]. \tag{A56}$$

Proof. Equation (A56) is a direct result of Equation (A53) by setting $B = \Omega$. Another proof uses the definition of the conditional expectation as $\varphi(\boldsymbol{y})$ (cf. Equations A49-A51). Thus, we obtain

$$\begin{aligned}
\mathrm{E}_{\boldsymbol{y}}\left[\mathrm{E}_{\boldsymbol{x}|\boldsymbol{y}}\left[\boldsymbol{x}|\,\boldsymbol{y}\right]\right] &= \int \mathrm{E}_{\boldsymbol{x}|\boldsymbol{y}}\left[\boldsymbol{x}|\,\boldsymbol{\eta}\right] f_{\boldsymbol{y}}\left(\boldsymbol{\eta}\right) \mathrm{d}\boldsymbol{\eta} = \iint \boldsymbol{x} f_{\boldsymbol{x}|\boldsymbol{y}}\left(\boldsymbol{x}|\boldsymbol{\eta}\right) \mathrm{d}\boldsymbol{x} f_{\boldsymbol{y}}\left(\boldsymbol{\eta}\right) \mathrm{d}\boldsymbol{\eta} \\
&= \iint \boldsymbol{x} f_{\boldsymbol{x},\boldsymbol{y}}\left(\boldsymbol{x}, \boldsymbol{\eta}\right) \mathrm{d}\boldsymbol{x}\, \mathrm{d}\boldsymbol{\eta} = \int \boldsymbol{x} f_{\boldsymbol{x}}\left(\boldsymbol{x}\right) \mathrm{d}\boldsymbol{x} \\
&= \mathrm{E}_{\boldsymbol{x}}\left[\boldsymbol{x}\right].
\end{aligned} \tag{A57}$$

□

A5.3 Pointwise Minimization

Let $\boldsymbol{x} : \Omega \rightarrow \mathbb{R}^{N_x}$ and $\boldsymbol{y} : \Omega \rightarrow \mathbb{R}^{N_y}$ be random vectors with a joint distribution, $h : \mathbb{R}^{N_y} \rightarrow \mathbb{R}^M$ be a measurable function of $\boldsymbol{y}$ and $g : \mathbb{R}^{N_x} \times \mathbb{R}^M \rightarrow \mathbb{R}_{+,0}$ a non-negative, measurable function of $\boldsymbol{x}$ and $h(\boldsymbol{y})$. Assuming its existence, the determination of the function h^* which minimizes $\mathrm{E}_{\boldsymbol{x},\boldsymbol{y}}\left[g(\boldsymbol{x}, h(\boldsymbol{y}))\right]$ can be determined pointwise, i. e.,

$$h^*(\boldsymbol{\eta}) = \operatorname*{argmin}_{\boldsymbol{z} \in \mathbb{R}^M} \mathrm{E}_{\boldsymbol{x}|\boldsymbol{y}}\left[g(\boldsymbol{x}, \boldsymbol{z})|\,\boldsymbol{\eta}\right], \qquad \boldsymbol{\eta} \in \mathbb{R}^{N_y}. \tag{A58}$$

Proof.

$$\begin{aligned}
\min_h \mathrm{E}_{\boldsymbol{x},\boldsymbol{y}}\left[g(\boldsymbol{x}, h(\boldsymbol{y}))\right] &= \min_h \iint f_{\boldsymbol{x},\boldsymbol{y}}\left(\boldsymbol{\xi}, \boldsymbol{\eta}\right) g(\boldsymbol{\xi}, h(\boldsymbol{\eta}))\, \mathrm{d}\boldsymbol{\xi}\, \mathrm{d}\boldsymbol{\eta} \\
&= \min_h \int f_{\boldsymbol{y}}\left(\boldsymbol{\eta}\right) \int f_{\boldsymbol{x}|\boldsymbol{y}}\left(\boldsymbol{\xi}|\boldsymbol{\eta}\right) g(\boldsymbol{\xi}, h(\boldsymbol{\eta}))\, \mathrm{d}\boldsymbol{\xi}\, \mathrm{d}\boldsymbol{\eta} \\
&= \min_h \int f_{\boldsymbol{y}}\left(\boldsymbol{\eta}\right) \mathrm{E}_{\boldsymbol{x}|\boldsymbol{y}}\left[g(\boldsymbol{x}, h(\boldsymbol{\eta}))|\,\boldsymbol{\eta}\right] \mathrm{d}\boldsymbol{\eta}.
\end{aligned} \tag{A59}$$

Note that all Probability Density Functions (PDFs) and the function g are non-negative. Thus, the minimum of the expression above is attained by using the function h which, evaluated at each given value $\boldsymbol{\eta}$, minimizes $\mathrm{E}_{\boldsymbol{x}|\boldsymbol{y}}\left[g(\boldsymbol{x}, h(\boldsymbol{\eta}))|\,\boldsymbol{\eta}\right]$, i. e., the problem is to determine the value $\boldsymbol{z} = h(\boldsymbol{\eta}) \in \mathbb{R}^M$ which minimizes the conditional expected value at a given point $\boldsymbol{\eta}$. Consequently,

$$\min_h \mathrm{E}_{\boldsymbol{x},\boldsymbol{y}}\left[g(\boldsymbol{x}, h(\boldsymbol{y}))\right] = \int f_{\boldsymbol{y}}\left(\boldsymbol{\eta}\right) \min_{\boldsymbol{z} \in \mathbb{R}^M} \mathrm{E}_{\boldsymbol{x}|\boldsymbol{y}}\left[g(\boldsymbol{x}, \boldsymbol{z})|\,\boldsymbol{\eta}\right] \mathrm{d}\boldsymbol{\eta}, \tag{A60}$$

where the minimizer is $\boldsymbol{z}^* = h^*(\boldsymbol{\eta})$, i. e., is the value of the optimal function h^* evaluated at $\boldsymbol{\eta}$. □

Remark: A different approach of proving the above result can be found in, e. g., [3, pp. 260-261]. There, it is argued that

$$\min_{\boldsymbol{z}\in\mathbb{R}^M} \mathrm{E}_{\boldsymbol{x}|\boldsymbol{y}}\left[g(\boldsymbol{x},\boldsymbol{z})|\,\boldsymbol{\eta}\right] \leq \mathrm{E}_{\boldsymbol{x}|\boldsymbol{y}}\left[g(\boldsymbol{x},h(\boldsymbol{\eta}))|\,\boldsymbol{\eta}\right] \tag{A61}$$

for every function h. Consequently, taking the expected value w.r.t. $\boldsymbol{y}$ on both sides, we obtain

$$\mathrm{E}_{\boldsymbol{y}}\left[\min_{\boldsymbol{z}\in\mathbb{R}^M} \mathrm{E}_{\boldsymbol{x}|\boldsymbol{y}}\left[g(\boldsymbol{x},\boldsymbol{z})|\,\boldsymbol{y}\right]\right] \leq \mathrm{E}_{\boldsymbol{x},\boldsymbol{y}}\left[g(\boldsymbol{x},h(\boldsymbol{y}))\right]. \tag{A62}$$

Since this inequality holds for every function h, it is also true that

$$\mathrm{E}_{\boldsymbol{y}}\left[\min_{\boldsymbol{z}\in\mathbb{R}^M} \mathrm{E}_{\boldsymbol{x}|\boldsymbol{y}}\left[g(\boldsymbol{x},\boldsymbol{z})|\,\boldsymbol{y}\right]\right] \leq \min_{h} \mathrm{E}_{\boldsymbol{x},\boldsymbol{y}}\left[g(\boldsymbol{x},h(\boldsymbol{y}))\right]. \tag{A63}$$

On the other hand, the minimum of the expected value w.r.t. h must be smaller than its value using an arbitrary function, e. g., h^* which is given in Equation (A58). This results in

$$\min_{h} \mathrm{E}_{\boldsymbol{x},\boldsymbol{y}}\left[g(\boldsymbol{x},h(\boldsymbol{y}))\right] \leq \mathrm{E}_{\boldsymbol{x},\boldsymbol{y}}\left[g(\boldsymbol{x},h^*(\boldsymbol{y}))\right] = \mathrm{E}_{\boldsymbol{y}}\left[\min_{\boldsymbol{z}\in\mathbb{R}^M} \mathrm{E}_{\boldsymbol{x}|\boldsymbol{y}}\left[g(\boldsymbol{x},\boldsymbol{z})|\,\boldsymbol{y}\right]\right]. \tag{A64}$$

Putting both inequalities (A63) and (A64) together leads finally to the result

$$\min_{h} \mathrm{E}_{\boldsymbol{x},\boldsymbol{y}}\left[g(\boldsymbol{x},h(\boldsymbol{y}))\right] = \mathrm{E}_{\boldsymbol{y}}\left[\min_{\boldsymbol{z}\in\mathbb{R}^M} \mathrm{E}_{\boldsymbol{x}|\boldsymbol{y}}\left[g(\boldsymbol{x},\boldsymbol{z})|\,\boldsymbol{y}\right]\right]. \tag{A65}$$

A6. LQG Control

The term Linear Quadratic Gaussian (LQG) describes scenarios where the associated control problems deal with *linear* dynamical systems, *quadratic* cost functions and a *Gaussian* model for the stochastic driving processes. For the dynamical system to be controlled, we adopt the system model introduced in Section 1.5, i. e.,

$$\begin{aligned} \boldsymbol{x}_{k+1} &= \boldsymbol{A}_k\boldsymbol{x}_k + \boldsymbol{B}_k\boldsymbol{u}_k + \boldsymbol{w}_k, \\ \boldsymbol{y}_k &= \boldsymbol{C}_k\boldsymbol{x}_k + \boldsymbol{v}_k, \qquad\qquad k \in \{0,1,\ldots,N-1\}, \end{aligned} \tag{A66}$$

where $\boldsymbol{x}_k \in \mathbb{R}^{N_x}$ is the system state at time index k, $\boldsymbol{u}_k \in \mathbb{R}^{N_u}$ is the control input and $\boldsymbol{y}_k \in \mathbb{R}^{N_y}$ is the output of the system. The system matrix at time index k is given by $\boldsymbol{A}_k$, the system input matrix is $\boldsymbol{B}_k$ and the output matrix is $\boldsymbol{C}_k$, where all dimensions are according to the state, input and output vectors, respectively. The driving noise sequence $(\boldsymbol{w}_k : k \in \{0,1,\ldots,N-1\})$ is assumed to be independent but not necessarily identically Gaussian distributed, i. e., $\boldsymbol{w}_k \sim \mathcal{N}\left(\boldsymbol{0}_{N_x}, \boldsymbol{C}_{\boldsymbol{w}_k}\right)$. The same holds for the measurement noise sequence $(\boldsymbol{v}_k : k \in \{0,1,\ldots,N-1\})$, i. e., $\boldsymbol{v}_k \sim \mathcal{N}\left(\boldsymbol{0}_{N_y}, \boldsymbol{C}_{\boldsymbol{v}_k}\right)$. The initial state is $\boldsymbol{x}_0 \sim \mathcal{N}\left(\boldsymbol{\mu}_{\boldsymbol{x}_0}, \boldsymbol{C}_{\boldsymbol{x}_0}\right)$. Additionally, the process noise sequence, the observation noise sequence and the initial state are assumed to be mutually independent.[15] The number $N \in \mathbb{N}$ is the so-called horizon, i. e., it determines the number of steps which are considered for the LQG problem.

[15]The independence assumption can be relaxed if $(\boldsymbol{w}_k : k \in \{0,1,\ldots,N-1\})$ is a Gauss-Markov sequence, i. e., admits for a state space representation which is driven by white noise. Thus, by augmenting the state equation (A66) (cf. [38, p. 38]), the problem of correlated noise can be reduced to the white (independent) noise case. The same holds for the observation noise $(\boldsymbol{v}_k : k \in \{0,1,\ldots,N-1\})$ and if the different noise sequences are mutually correlated (cf. [37, Section 11.2] and [133, Sections 5.9 and 5.10]).

The last part for the description of the LQG control problem is the quadratic cost function which is used to determine the control sequence $(\boldsymbol{u}_k : k \in \{0, \ldots, N-1\})$. It reflects the goal to keep the deviation of the system state from the origin small without an excessively large control effort. The criterion for the quantification of the terms "small" and "large" is the expected value of the squared and weighted Euclidean norm of the state and the control input vector, respectively. The expected value is applied because the state as well as the control are in general random variables due to the process and the measurement noise. Since the dynamical system is considered over a horizon of N steps, the overall cost function is chosen to be the sum of the cost of the individual steps. Thus, the cost function of the LQG control problem reads as

$$J = \mathrm{E}_{\boldsymbol{x}_0,\boldsymbol{w}_0,\ldots,\boldsymbol{w}_{N-1},\boldsymbol{v}_0,\ldots,\boldsymbol{v}_{N-1}} \left[\boldsymbol{x}_N^{\mathrm{T}} \boldsymbol{Q}_N \boldsymbol{x}_N + \sum_{n=0}^{N-1} \begin{bmatrix} \boldsymbol{x}_n \\ \boldsymbol{u}_n \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \boldsymbol{Q}_n & \boldsymbol{S}_n \\ \boldsymbol{S}_n^{\mathrm{T}} & \boldsymbol{R}_n \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_n \\ \boldsymbol{u}_n \end{bmatrix} \right], \tag{A67}$$

where the weight matrices $\boldsymbol{Q}_n \in \mathbb{R}^{N_x \times N_x}$, $\boldsymbol{R}_n \in \mathbb{R}^{N_u \times N_u}$ and $\boldsymbol{S}_n \in \mathbb{R}^{N_x \times N_u}$ fulfill the conditions

$$\boldsymbol{Q}_n \geq \mathbf{0}_{N_x \times N_x}, \qquad \boldsymbol{R}_n > \mathbf{0}_{N_u \times N_u}, \qquad \begin{bmatrix} \boldsymbol{Q}_n & \boldsymbol{S}_n \\ \boldsymbol{S}_n^{\mathrm{T}} & \boldsymbol{R}_n \end{bmatrix} \geq \mathbf{0}_{(N_x+N_u)\times(N_x+N_u)}, \tag{A68}$$

for $n \in \{0, 1, \ldots, N\}$ or $n \in \{0, 1, \ldots, N-1\}$, respectively, in order to obtain a convex cost function which is bounded below by zero.[16] These weight matrices are not given a priori but have to be chosen by the system designer to reflect the objectives of the specific control problem. As an example, for $\boldsymbol{S}_n = \mathbf{0}_{N_x \times N_u}$, $n \in \{0, 1, \ldots, N-1\}$, a scaling of $\boldsymbol{Q}_n$ while keeping $\boldsymbol{R}_n$ fixed can be used to trade the goal of keeping the system state small against the goal of a small control effort.[17] Additionally, the weight matrices can be used to emphasize subspaces of the state and control space, respectively, where large deviations are more severe than in other subspaces. Note that the expected value in Equation (A67) is taken w.r.t. all random variables which are involved in the description of the dynamical system to be controlled and the measurements of the system output (cf. Equation A66) over the horizon N.[18]

For the determination of the optimal control input for the dynamical system given by Equation (A66), it is typically assumed in the LQG framework that at time index k the controller has access to all measurements of the system output up to this time and all preceding (self-generated) control inputs. The information state at time index k thus reads as the set

$$\mathcal{I}_k = \begin{cases} \{\boldsymbol{y}_0\}, & k = 0, \\ \{\boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_k, \boldsymbol{u}_0, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_{k-1}\}, & k \in \{1, 2, \ldots, N-1\}. \end{cases} \tag{A69}$$

Note that this assumption results in a recursive description of the information state, i. e., $\mathcal{I}_{k+1} = \{\mathcal{I}_k, \boldsymbol{y}_{k+1}, \boldsymbol{u}_k\}$, $k \in \{0, 1, \ldots, N-2\}$, with the initial set $\mathcal{I}_0 = \{\boldsymbol{y}_0\}$.

[16] The condition $\boldsymbol{R}_n > 0$ can be relaxed to $\boldsymbol{R}_n \geq 0$ (cf. Section A3), but care must be taken w.r.t. the stability of the optimal control (cf. [101]). Note that the minimization of the cost function given in Equation (A67) may still be a well-posed problem for indefinite weight matrices, see, e. g., [97, 98].

[17] This is a typical case for a multicriterion optimization problem, see, e. g., [94, Section 4.7].

[18] The expected value in Equation (A67) represents some abuse of notation because $\boldsymbol{x}_n$, $n \in \{0, 1, \ldots, N\}$, and $\boldsymbol{u}_n$, $n \in \{0, 1, \ldots, N-1\}$, are random vectors which are taken into account for the computation of the expected value. Nevertheless, since they can be represented as functions of the fundamental random vectors $\boldsymbol{x}_0, \boldsymbol{w}_0, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_{N-1}$, $\boldsymbol{v}_0, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_{N-1}$, they are not explicitly listed in the index of the expectation operator.

Now we are in the position to formulate the LQG control problem. The goal is to determine a sequence of measurable [19] functions μ_k, $k \in \{0, 1, \ldots, N-1\}$, of the information state $\mathcal{I}_k$, i. e.,

$$\boldsymbol{u}_k = \mu_k(\mathcal{I}_k), \qquad k \in \{0, 1, \ldots, N-1\}, \tag{A70}$$

such that the cost function given by Equation (A67) is minimized for the dynamical system given by Equation (A66). Formally, the control problem reads as

$$\begin{aligned} \underset{\mu_0, \mu_1, \ldots, \mu_{N-1}}{\text{minimize}} \; & \mathrm{E}_{\boldsymbol{x}_0, \boldsymbol{w}_0, \ldots, \boldsymbol{w}_{N-1}, \boldsymbol{v}_0, \ldots, \boldsymbol{v}_{N-1}} \left[\boldsymbol{x}_N^{\mathrm{T}} \boldsymbol{Q}_N \boldsymbol{x}_N + \sum_{n=0}^{N-1} \begin{bmatrix} \boldsymbol{x}_n \\ \boldsymbol{u}_n \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \boldsymbol{Q}_n & \boldsymbol{S}_n \\ \boldsymbol{S}_n^{\mathrm{T}} & \boldsymbol{R}_n \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_n \\ \boldsymbol{u}_n \end{bmatrix} \right], \\ \text{subject to} \quad & \boldsymbol{x}_{k+1} = \boldsymbol{A}_k \boldsymbol{x}_k + \boldsymbol{B}_k \boldsymbol{u}_k + \boldsymbol{w}_k, \quad k \in \{0, 1, \ldots, N-1\}, \\ & \boldsymbol{y}_k = \boldsymbol{C}_k \boldsymbol{x}_k + \boldsymbol{v}_k, \qquad\qquad\quad k \in \{0, 1, \ldots, N-1\}, \\ & \boldsymbol{u}_k = \mu_k(\mathcal{I}_k), \qquad\qquad\qquad k \in \{0, 1, \ldots, N-1\}, \end{aligned} \tag{A71}$$

where $\mathcal{I}_k$ is the information set given by Equation (A69).

Note that the above optimization problem can not only be used to express the desire to keep the system state close to the origin. It is also possible to minimize the deviation from a given non-zero point or a trajectory which can be deterministic or stochastic. In order to explain this modification, assume that the control objective is to follow a trajectory generated by the dynamical system

$$\boldsymbol{\xi}_{k+1} = \boldsymbol{A}_k^{(R)} \boldsymbol{\xi}_k + \boldsymbol{\omega}_k, \qquad k \in \{0, 1, \ldots, N-1\}, \tag{A72}$$

where $\boldsymbol{\xi}_k \in \mathbb{R}^M$ is the state of the reference system and $(\boldsymbol{\omega}_k : k \in \{0, 1, \ldots, N-1\})$ with $\boldsymbol{\omega}_k \sim \mathcal{N}\left(\mathbf{0}_M, \boldsymbol{C}_{\boldsymbol{\omega}_k}\right)$, $k \in \{0, 1, \ldots, N-1\}$, is the independent driving noise process which is also assumed to be independent of all other random variables. The behavior of the reference system is determined by the system matrix $\boldsymbol{A}_k^{(R)} \in \mathbb{R}^{M \times M}$ and the parameters of noise distribution. For the sake of simplicity, let $N_x = M$. Using this assumption, the squared Euclidean norm of the deviation of the system state from the state of the reference can be expressed as

$$\|\boldsymbol{x}_k - \boldsymbol{\xi}_k\|_2^2 = \begin{bmatrix} \boldsymbol{x}_k \\ \boldsymbol{\xi}_k \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \mathbf{I}_{N_x} & -\mathbf{I}_{N_x} \\ -\mathbf{I}_{N_x} & \mathbf{I}_{N_x} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_k \\ \boldsymbol{\xi}_k \end{bmatrix}, \tag{A73}$$

where the vector $\boldsymbol{z}_k^{\mathrm{T}} = \left[\boldsymbol{x}_k^{\mathrm{T}}, \boldsymbol{\xi}_k^{\mathrm{T}}\right]^{\mathrm{T}}$ can be interpreted as the state of a dynamical system described by the difference equation

$$\boldsymbol{z}_{k+1} = \begin{bmatrix} \boldsymbol{A}_k & \\ & \boldsymbol{A}_k^{(R)} \end{bmatrix} \boldsymbol{z}_k + \begin{bmatrix} \boldsymbol{B}_k \\ \mathbf{0}_{M \times N_u} \end{bmatrix} \boldsymbol{u}_k + \begin{bmatrix} \boldsymbol{w}_k \\ \boldsymbol{\omega}_k \end{bmatrix}, \quad k \in \{0, 1, \ldots, N-1\}, \tag{A74}$$

Thus, by choosing the weight matrices $\boldsymbol{Q}_n$ according to Equation (A73), the tracking problem can be recast to fit in the formulation of the optimization problem given by Equation (A71).

A6.1 Dynamic Programming

The LQG control problem which has been introduced above fits into the more general framework of dynamic programming. It considers the problem of the optimal choice of a sequence of decisions,

[19] Precisely, the function μ_k at time index k is measurable w.r.t. the σ-algebra generated by the random variables $\boldsymbol{x}_0, \boldsymbol{w}_0, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_{k-1}, \boldsymbol{v}_0, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$.

where the decision in one step changes the basis for the decisions of subsequent steps. Examples for such situations are shortest path problems, inventory control or games like chess [6]. A key property of these problems is the necessity to trade the minimal cost of a single, local decision against the possibility that this step will lead to a higher cost of future decisions.

According to [6, pp. 2-3], two features characterize the considered model. The first one is a discrete-time dynamical system

$$\boldsymbol{x}_{k+1} = f_k(\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{w}_k), \qquad k \in \{0, 1, \ldots, N-1\}, \tag{A75}$$

where $\boldsymbol{x}_k$ is the system state, $\boldsymbol{u}_k$ is the control input and $\boldsymbol{w}_k$ is a random disturbance with properties given below. The initial state $\boldsymbol{x}_0$ is assumed to be a random variable. Additionally, the function f_k maps the system state, control input and disturbance at time index k to the state at the subsequent time index. The number of considered steps is given by the horizon N. The sets from which all variables are taken and the underlying probability space are not explicitly given at this point because they do not provide much insight into the description of the generic problem.

The goal of dynamic programming is to chose a sequence $(\boldsymbol{u}_k : k \in \{0, 1, \ldots, N-1\})$ of control inputs such that the dynamical system given by Equation (A75) behaves in an optimal way, where the optimality criterion is defined below. Thus, for the determination of $\boldsymbol{u}_k$, information about the system state at time index k is necessary. Since we adopt the model of [6, pp. 218-219], it is assumed that this information is provided by observations (or measurements) $\boldsymbol{y}_k$ according to

$$\boldsymbol{y}_k = \begin{cases} h_0(\boldsymbol{x}_0, \boldsymbol{v}_0), & k = 0, \\ h_k(\boldsymbol{x}_k, \boldsymbol{u}_{k-1}, \boldsymbol{v}_k), & k \in \{1, 2, \ldots, N-1\}, \end{cases} \tag{A76}$$

where $\boldsymbol{v}_k$ represents a random measurement disturbance and h_k are functions which map the system state, the control input and the measurement disturbance to the actual observations. For the sake of simplicity, we assume that $(\boldsymbol{w}_k : k \in \{0, 1, \ldots, N-1\})$ and $(\boldsymbol{v}_k : k \in \{0, 1, \ldots, N-1\})$ are independent random sequences which are also mutually independent and do not depend on the initial state $\boldsymbol{x}_0$.[20]

The second feature of the model for dynamic programming concerns the cost which is incurred at each step and which provides the optimality criterion for the determination of the control input. It is given by the functions g_k, $k \in \{0, 1, \ldots, N-1\}$, which depend on the state, the disturbance and the control variable $\boldsymbol{u}_k$ to be determined, i. e., the decision to be made, and since these quantities are not deterministic due to the system and observation model, the expected value of the cost is considered. An important assumption is that the overall cost over the horizon N is additive, i. e.,

$$J_N = \mathrm{E}_{\boldsymbol{x}_0, \boldsymbol{w}_0, \ldots, \boldsymbol{w}_{N-1}, \boldsymbol{v}_0, \ldots, \boldsymbol{v}_{N-1}} \left[g_N(\boldsymbol{x}_N) + \sum_{n=0}^{N-1} g_n(\boldsymbol{x}_n, \boldsymbol{u}_n, \boldsymbol{w}_n) \right], \tag{A77}$$

where $g_N(\boldsymbol{x}_N)$ is the cost incurred by the terminal state $\boldsymbol{x}_N$ alone. Finally, it is assumed that for the determination of the control input at time index k, all observations up to time k and all preceding inputs can be used. Thus, the available information set is given by

$$\mathcal{I}_k = \begin{cases} \{\boldsymbol{y}_0\}, & k = 0, \\ \{\boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_k, \boldsymbol{u}_0, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_{k-1}\}, & k \in \{1, 2, \ldots, N-1\}. \end{cases} \tag{A78}$$

[20] This assumption can be relaxed, e. g., if the dependencies of the noise sequences can be described by a state space model which is driven by independent noise. In that case, the original and the noise model can be combined to an augmented state space model with independent driving noise. Additionally, dependencies of the noise distributions on the state and control inputs may be considered, cf. [6, Chapter 5.1].

In order to minimize the cumulative cost of the individual steps over the horizon N (cf. Equation A77), functions have to be determined which map the available information about the system state (cf. Equation A78) to an appropriate input for the dynamical system (cf. Equation A75), i. e.,

$$\boldsymbol{u}_k = \mu_k(\mathcal{I}_k), \qquad k \in \{0, 1, \ldots, N-1\}, \tag{A79}$$

where μ_k is a measurable function of $\mathcal{I}_k$. The optimization problem to be solved thus reads as

$$\begin{aligned} \underset{\mu_0,\mu_1,\ldots,\mu_{N-1}}{\text{minimize}} \quad & \mathrm{E}_{\boldsymbol{x}_0,\boldsymbol{w}_0,\ldots,\boldsymbol{w}_{N-1},\boldsymbol{v}_0,\ldots,\boldsymbol{v}_{N-1}} \left[g_N(\boldsymbol{x}_N) + \sum_{n=0}^{N-1} g_n(\boldsymbol{x}_n, \boldsymbol{u}_n, \boldsymbol{w}_n) \right], \\ \text{subject to} \quad & \boldsymbol{x}_{k+1} = f_k(\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{w}_k), \qquad k \in \{0, 1, \ldots, N-1\}, \\ & \boldsymbol{y}_k = \begin{cases} h_0(\boldsymbol{x}_0, \boldsymbol{v}_0), & k = 0, \\ h_k(\boldsymbol{x}_k, \boldsymbol{u}_{k-1}, \boldsymbol{v}_k), & k \in \{1, 2, \ldots, N-1\}, \end{cases} \\ & \boldsymbol{u}_k = \mu_k(\mathcal{I}_k), \qquad k \in \{0, 1, \ldots, N-1\}, \end{aligned} \tag{A80}$$

where $\mathcal{I}_k$ is given by Equation (A78).

The main idea for the solution of the optimization problem (A80) is the *principle of optimality* (cf. [134, p. 83] or [6, p. 18]). In order to understand this principle, consider the following "sub-problem" of (A80):

$$\begin{aligned} \underset{\mu_m,\mu_{m+1},\ldots,\mu_{N-1}}{\text{minimize}} \quad & \mathrm{E}_{\boldsymbol{x}_m,\boldsymbol{w}_m,\ldots,\boldsymbol{w}_{N-1},\boldsymbol{v}_m,\ldots,\boldsymbol{v}_{N-1}|\mathcal{I}_m} \left[g_N(\boldsymbol{x}_N) + \sum_{n=m}^{N-1} g_n(\boldsymbol{x}_n, \boldsymbol{u}_n, \boldsymbol{w}_n) \middle| \mathcal{I}_m \right], \\ \text{subject to} \quad & \boldsymbol{x}_{k+1} = f_k(\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{w}_k), \quad k \in \{m, m+1, \ldots, N-1\}, \\ & \boldsymbol{y}_k = h_k(\boldsymbol{x}_k, \boldsymbol{u}_{k-1}, \boldsymbol{v}_k), \quad k \in \{m, m+1, \ldots, N-1\}, \\ & \boldsymbol{u}_k = \mu_k(\mathcal{I}_k), \quad k \in \{m, m+1, \ldots, N-1\}, \end{aligned} \tag{A81}$$

with $m > 0$. This problem considers only the *tail* of the optimization problem given by (A80). The first m steps are assumed to be given and described by the information set $\mathcal{I}_m$ which contains the measurements up to time index m and the control inputs which are functions of these measurements. The objective of (A81) is to determine the optimal control input for the *remaining* steps when the first m steps have already been made and are fixed. Loosely speaking, if we came to some point after an arbitrary choice for the first m steps, the question to be answered is what is the optimal way to finish the optimization problem to the terminal point at time index N. This is the reason why the cost function of the optimization problem (A81) is called the *cost-to-go*. The solution of the minimization of this cost subject to the state and observation equations results in the optimal sequence of control inputs for the remaining part of the horizon which now depends on the choice of control inputs for the first m steps and the resulting measurements, i. e., on $\mathcal{I}_m$.

The principle of optimality states the following: assume that an optimal sequence of functions μ_k^*, $k \in \{0, 1, \ldots, N-1\}$, for the computation of the control input has been determined for the original problem (A80) and is used for the first m steps of the control problem, which results in the information set $\mathcal{I}_m$ at time index m. Then, given this information set, the remaining part of the optimal functions μ_k^*, $k \in \{m, m+1, \ldots, N-1\}$, also minimizes the cost-to-go, i. e., is the solution of the optimization problem (A81). Thus, the solution of the optimization over the complete horizon is also the solution of all tail problems.

The solution of the optimization problem (A80) can now be obtained by recursively solving the tail problems described by (A81). In order to derive the recursive representation of the problem, let J^* be the optimal value of the problem (A80), i. e.,

$$J^* = \min_{\mu_0,\mu_1,\dots,\mu_{N-1}} \mathrm{E}_{\boldsymbol{x}_0,\boldsymbol{w}_0,\dots,\boldsymbol{w}_{N-1},\boldsymbol{v}_0,\dots,\boldsymbol{v}_{N-1}} \left[g_N(\boldsymbol{x}_N) + \sum_{n=0}^{N-1} g_n(\boldsymbol{x}_n, \mu_n(\mathcal{I}_n), \boldsymbol{w}_n), \right] \tag{A82}$$

subject to the corresponding constraints. Since the control input $\boldsymbol{u}_0$ must be determined as a function μ_0 of $\mathcal{I}_0$, recall Section A5.3 and use the result of Section A5.2 to introduce the conditional expected value given $\mathcal{I}_0$. Additionally, due to the fact that the cost function is additive, J^* can be rewritten as

$$\begin{aligned} J^* = \min_{\mu_0,\mu_1,\dots,\mu_{N-1}} \mathrm{E}_{\boldsymbol{x}_0,\boldsymbol{v}_0} \Bigg[\mathrm{E}_{\boldsymbol{x}_0,\boldsymbol{w}_0,\dots,\boldsymbol{w}_{N-1},\boldsymbol{v}_0,\dots,\boldsymbol{v}_{N-1}|\mathcal{I}_0} \Bigg[& g_0(\boldsymbol{x}_0, \mu_0(\mathcal{I}_0), \boldsymbol{w}_0) \\ & + g_N(\boldsymbol{x}_N) + \sum_{n=1}^{N-1} g_n(\boldsymbol{x}_n, \mu_n(\mathcal{I}_n), \boldsymbol{w}_n) \Bigg| \mathcal{I}_0 \Bigg] \Bigg], \end{aligned} \tag{A83}$$

where $\mathcal{I}_0 = \{\boldsymbol{y}_0\}$ (cf. Equation A78) and the distribution of $\boldsymbol{y}_0$ is completely described by the joint distribution of $\boldsymbol{x}_0$ and $\boldsymbol{v}_0$ (cf. Equation A76). Following the arguments of Section A5.3, the functions μ_k, $k \in \{0, 1, \dots, N-1\}$, which minimize J can be determined pointwise for each given value of $\mathcal{I}_0$. Thus, the minimization can be carried out inside the outer expected value and we get[21]

$$\begin{aligned} J^* = \mathrm{E}_{\boldsymbol{x}_0,\boldsymbol{v}_0} \Bigg[\min_{\boldsymbol{u}_0} \Bigg(& \mathrm{E}_{\boldsymbol{x}_0,\boldsymbol{w}_0|\mathcal{I}_0} \left[g_0(\boldsymbol{x}_0, \boldsymbol{u}_0, \boldsymbol{w}_0) | \, \mathcal{I}_0 \right] \\ & + \min_{\mu_1,\dots,\mu_{N-1}} \mathrm{E}_{\boldsymbol{x}_1,\boldsymbol{w}_1,\dots,\boldsymbol{w}_{N-1},\boldsymbol{v}_1,\dots,\boldsymbol{v}_{N-1}|\mathcal{I}_0} \Bigg[g_N(\boldsymbol{x}_N) + \sum_{n=1}^{N-1} g_n(\boldsymbol{x}_n, \mu_n(\mathcal{I}_n), \boldsymbol{w}_n) \Bigg| \mathcal{I}_0 \Bigg] \Bigg) \Bigg], \end{aligned} \tag{A84}$$

where the random variables that are considered for the individual expected values have been adapted to the respective arguments. The principle of optimality ensures that the functions μ_k, $k \in \{1, 2, \dots, N-1\}$, which are obtained by the minimization of the second summand of Equation (A84) alone are also optimal for the overall problem of minimizing J^*.[22] Comparing this minimization problem with Equation (A82), we identify an essentially equivalent problem which starts at $k = 1$ and is conditioned on the information $\mathcal{I}_0$. Consequently, the same steps which led to Equation (A84) can be repeated to obtain

$$\begin{aligned} & \min_{\mu_1,\dots,\mu_{N-1}} \mathrm{E}_{\boldsymbol{x}_1,\boldsymbol{w}_1,\dots,\boldsymbol{w}_{N-1},\boldsymbol{v}_1,\dots,\boldsymbol{v}_{N-1}|\mathcal{I}_0} \Bigg[g_N(\boldsymbol{x}_N) + \sum_{n=1}^{N-1} g_n(\boldsymbol{x}_n, \mu_n(\mathcal{I}_n), \boldsymbol{w}_n) \Bigg| \mathcal{I}_0 \Bigg] \\ & = \mathrm{E}_{\boldsymbol{x}_1,\boldsymbol{v}_1|\mathcal{I}_0} \Bigg[\min_{\mu_1,\dots,\mu_{N-1}} \mathrm{E}_{\boldsymbol{x}_1,\boldsymbol{w}_1,\dots,\boldsymbol{w}_{N-1},\boldsymbol{v}_1,\dots,\boldsymbol{v}_{N-1}|\mathcal{I}_1} \Bigg[g_N(\boldsymbol{x}_N) + \sum_{n=1}^{N-1} g_n(\boldsymbol{x}_n, \mu_n(\mathcal{I}_n), \boldsymbol{w}_n) \Bigg| \mathcal{I}_0, \boldsymbol{y}_1, \boldsymbol{u}_0 \Bigg] \Bigg| \mathcal{I}_0 \Bigg], \end{aligned} \tag{A85}$$

[21] The notation of Equation (A84) is potentially misleading because the inner conditional expectation is a random variable due to the fact that the condition $\mathcal{I}_0$ is a random variable. Thus, the minimization w.r.t. $\boldsymbol{u}_0$ and μ_k, $k \in \{1, 2, \dots, N-1\}$, has to be understood in the sense of Section A5.3, i. e., pointwise.

[22] A detailed derivation of this argument can be found in [6, Chapter 1.5].

where the equivalence $\mathcal{I}_1 = \{\mathcal{I}_0, \boldsymbol{y}_1, \boldsymbol{u}_0\}$ (cf. Equation A78) and the fact that $\boldsymbol{u}_0$ is a deterministic function of $\mathcal{I}_0$ have been used. Denote the solution of the optimization problem (A81) as $J_m(\mathcal{I}_m)$, i. e.,

$$J_m(\mathcal{I}_m) = \min_{\mu_m,\ldots,\mu_{N-1}} \mathrm{E}_{\boldsymbol{x}_m,\boldsymbol{w}_m,\ldots,\boldsymbol{w}_{N-1},\boldsymbol{v}_m,\ldots,\boldsymbol{v}_{N-1}|\mathcal{I}_m} \left[g_N(\boldsymbol{x}_N) + \sum_{n=m}^{N-1} g_n(\boldsymbol{x}_n, \mu_n(\mathcal{I}_n), \boldsymbol{w}_n) \middle| \mathcal{I}_m \right], \quad \text{(A86)}$$

subject to the corresponding constraints. Using this notation, Equation (A84) can be rewritten as

$$J^* = \mathrm{E}_{\boldsymbol{x}_0,\boldsymbol{v}_0} \left[J_0(\mathcal{I}_0) \right], \quad \text{(A87)}$$

and the expression inside the expected value becomes (cf. Equations A84 and A85)

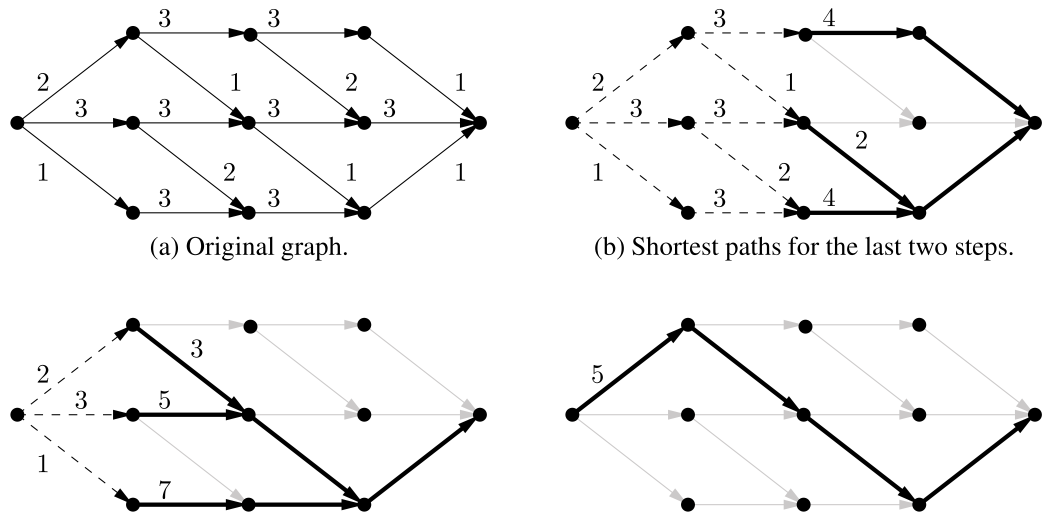
$$\begin{aligned} J_0(\mathcal{I}_0) &= \min_{\boldsymbol{u}_0} \left(\mathrm{E}_{\boldsymbol{x}_0,\boldsymbol{w}_0|\mathcal{I}_0} \left[g_0(\boldsymbol{x}_0, \boldsymbol{u}_0, \boldsymbol{w}_0) | \, \mathcal{I}_0 \right] + \mathrm{E}_{\boldsymbol{x}_1,\boldsymbol{v}_1|\mathcal{I}_0} \left[J_1(\mathcal{I}_1) | \, \mathcal{I}_0 \right] \right) \\ &= \min_{\boldsymbol{u}_0} \mathrm{E}_{\boldsymbol{x}_0,\boldsymbol{w}_0,\boldsymbol{v}_1|\mathcal{I}_0} \left[g_0(\boldsymbol{x}_0, \boldsymbol{u}_0, \boldsymbol{w}_0) + J_1(\mathcal{I}_0, \boldsymbol{y}_1, \boldsymbol{u}_0) | \, \mathcal{I}_0 \right]. \end{aligned} \quad \text{(A88)}$$

A further repetition of the above steps finally results in the recursive determination of the optimal control inputs and the associated costs:

$$J_m(\mathcal{I}_m) = \begin{cases} \min\limits_{\boldsymbol{u}_m} \mathrm{E}_{\boldsymbol{x}_m,\boldsymbol{w}_m,\boldsymbol{v}_{m+1}|\mathcal{I}_m} \left[g_m(\boldsymbol{x}_m, \boldsymbol{u}_m, \boldsymbol{w}_m) + J_{m+1}(\mathcal{I}_m, \boldsymbol{y}_{m+1}, \boldsymbol{u}_m) | \, \mathcal{I}_m \right], \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad m \in \{0, \ldots, N-2\}, \\ \min\limits_{\boldsymbol{u}_{N-1}} \mathrm{E}_{\boldsymbol{x}_{N-1},\boldsymbol{w}_{N-1}|\mathcal{I}_{N-1}} [g_{N-1}(\boldsymbol{x}_{N-1}, \boldsymbol{u}_{N-1}, \boldsymbol{w}_{N-1}) \\ \qquad\qquad + g_N(f_{N-1}(\boldsymbol{x}_{N-1}, \boldsymbol{u}_{N-1}, \boldsymbol{w}_{N-1})) | \mathcal{I}_{N-1}], \qquad m = N-1. \end{cases} \quad \text{(A89)}$$

This result provides the following interpretation of the optimization problem: formally, the minimization of the cost function J can be split in two parts. First, assuming that the inputs for the first m steps of the problem which lead to the information $\mathcal{I}_m$ are given, optimize the remaining steps based on this history, i. e., minimize $J_m(\mathcal{I}_m)$. Second, with the knowledge about the minimal cost-to-go, determine the first m control inputs such that the sum of the cost which is associated with these steps and the optimal cost-to-go is minimal. Equation (A89) shows that this can be performed one step after another going backwards in time and starting with $m = N-1$. At this time index, the terminal stage of the optimization problem is considered and the cost-to-go is simply determined by g_N and g_{N-1}. Subsequently, a sequence of optimization problems is solved by going one step back and minimizing the sum of the cost associated with this step and the previously determined minimal cost-to-go. Depending on the actual structure of the dynamic system and the cost function, this *dynamic programming algorithm* can lead to an efficient solution method.

As an example for the application of the dynamic programming algorithm, consider the problem of finding the shortest path through a directed graph which is depicted in Figure A1a. The goal is to traverse the graph from the left to the right. The nodes of the graph correspond to the states of the system and a state transition to the move from one node to a neighboring node along the edge which connects both. The cost of such a transition is denoted by the weight of the edge and the cost of a path from the left to the right is the sum of the weights of the edges which make up the path. Note that the control input corresponds to the choice of the specific neighboring node to move to. Although this is an informal description of the problem, it should be clear that it fits in the framework of dynamic programming, i. e., the propagation of the system state can be described according to Equation (A75) and the overall cost accumulates the cost of the individual steps of the problem (cf. Equation A77). In order to keep things simple, the problem is assumed to be deterministic, i. e., the state transitions as well as the associated costs are deterministic, and it is assumed that the system state is known perfectly.



(a) Original graph. (b) Shortest paths for the last two steps.

(c) Shortest paths for the last three steps. (d) Shortest path through the graph.

Figure A1: Shortest path problem solved by dynamic programming.

Since exactly four state transitions are necessary to move from the left to the right of the graph and it is not allowed to stay at a specific node, we have $N = 4$.[23] The determination of the minimal cost-to-go for the last transition is trivial (there is only one edge connecting the right node with each preceding one). Thus, we start with the determination of the minimal cost-to-go for $m = N - 2 = 2$. The result is shown in Figure A1b. Starting from the top node, the shortest path to the destination has weight 4, from the center note 2 and from the bottom note we obtain the cost 4. Now the optimal cost of the tail of the path (the optimal cost-to-go) is known depending on the state this tail is starting from and irrespective of the path that led to this state. Thus, for the determination of the optimal cost-to-go for $m = N - 3 = 1$, it is only necessary to determine the minimal sum of the cost of a state transition from any state to the neighboring top, center or bottom node and the remaining cost-to-go from that node which has been optimized one step earlier. Figure A1c shows the result which gives a cost of 3 from the top node, 5 from the center node and 7 from the bottom node to the destination. For the calculation of the minimal overall cost, i. e., the shortest path from the left to the right, this step has to be repeated which is again simple due to the fact that the starting node has only one edge which connects it to each of the neighboring nodes. The shortest path, which has a weight of 5, is shown in Figure A1d.

Having a closer look on the above example, the reader with a background in communication and information theory might recognize the Viterbi algorithm [135]. This algorithm is an instance of the dynamic programming principle which is commonly executed forward in time. This is not a contradiction to the description that has been chosen here since for deterministic shortest path problems, a path from the starting point to the destination has the same length as from the destination to the starting point. Thus, the graph can be equivalently traversed backward or forward.

[23] An equivalent problem is to allow to stay at a node and associate an arbitrary positive cost with the stay. In that case, the optimal choice is to directly move to the next node because the cost of any path with a stay is always larger than without one.

A6.2 Solution of the LQG Control Problem

The LQG control problem can be solved by applying the dynamic programming approach which has been introduced in the last section. To this end, we identify that the function which describes the evolution of the system state (cf. Equation A75) is given by Equation (A66), i, e.,

$$f_k(\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{w}_k) = \boldsymbol{A}_k\boldsymbol{x}_k + \boldsymbol{B}_k\boldsymbol{u}_k + \boldsymbol{w}_k, \qquad k \in \{0, 1, \ldots, N-1\}, \tag{A90}$$

and the observation equation (cf. Equation A76)[24] is

$$h_k(\boldsymbol{x}_k, \boldsymbol{v}_k) = \boldsymbol{C}_k\boldsymbol{x}_k + \boldsymbol{v}_k, \qquad k \in \{0, 1, \ldots, N-1\}. \tag{A91}$$

Finally, comparing the cost functions of Equation (A77) and Equation (A67), we observe that

$$g_N(\boldsymbol{x}_N) = \boldsymbol{x}_N^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{x}_N \tag{A92}$$

and

$$g_k(\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{w}_k) = \boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{Q}_k\boldsymbol{x}_k + \boldsymbol{u}_k^{\mathrm{T}}\boldsymbol{R}_k\boldsymbol{u}_k + 2\boldsymbol{u}_k^{\mathrm{T}}\boldsymbol{S}_k^{\mathrm{T}}\boldsymbol{x}_k, \qquad k \in \{0, 1, \ldots, N-1\}. \tag{A93}$$

Using these functions, the optimization problem (A80) leads to the LQG problem (A71). It is solved by recursively minimizing backwards in time the cost-to-go, i. e., the cost for the remaining steps of the optimization problem when the first m steps have already been made and provided the information (cf. Equation A69)

$$\mathcal{I}_m = \begin{cases} \{\boldsymbol{y}_0\}, & m = 0, \\ \{\boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_m, \boldsymbol{u}_0, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_{m-1}\}, & m \in \{1, 2, \ldots, N-1\}, \end{cases} \tag{A94}$$

starting with $m = N-1$. Thus, it is assumed that the optimal control inputs $\boldsymbol{u}_0, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_{N-2}$ have already been applied and the observations $\boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{N-1}$ are known. The minimal cost-to-go for the last step of the control problem is then given by (cf. Equation A89)

$$\begin{aligned} J_{N-1}(\mathcal{I}_{N-1}) &= \min_{\boldsymbol{u}_{N-1}} \mathrm{E}_{\boldsymbol{x}_{N-1},\boldsymbol{w}_{N-1}|\mathcal{I}_{N-1}} \Big[\boldsymbol{x}_{N-1}^{\mathrm{T}}\boldsymbol{Q}_{N-1}\boldsymbol{x}_{N-1} + \boldsymbol{u}_{N-1}^{\mathrm{T}}\boldsymbol{R}_{N-1}\boldsymbol{u}_{N-1} + 2\boldsymbol{u}_{N-1}^{\mathrm{T}}\boldsymbol{S}_{N-1}^{\mathrm{T}}\boldsymbol{x}_{N-1} \\ &\quad + (\boldsymbol{A}_{N-1}\boldsymbol{x}_{N-1} + \boldsymbol{B}_{N-1}\boldsymbol{u}_{N-1} + \boldsymbol{w}_{N-1})^{\mathrm{T}}\boldsymbol{Q}_N(\boldsymbol{A}_{N-1}\boldsymbol{x}_{N-1} + \boldsymbol{B}_{N-1}\boldsymbol{u}_{N-1} + \boldsymbol{w}_{N-1})\Big|\mathcal{I}_{N-1}\Big] \\ &= \min_{\boldsymbol{u}_{N-1}} \mathrm{E}_{\boldsymbol{x}_{N-1},\boldsymbol{w}_{N-1}|\mathcal{I}_{N-1}} \Big[\boldsymbol{x}_{N-1}^{\mathrm{T}}\left(\boldsymbol{A}_{N-1}^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{A}_{N-1} + \boldsymbol{Q}_{N-1}\right)\boldsymbol{x}_{N-1} \\ &\quad + \boldsymbol{u}_{N-1}^{\mathrm{T}}\left(\boldsymbol{B}_{N-1}^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{B}_{N-1} + \boldsymbol{R}_{N-1}\right)\boldsymbol{u}_{N-1} \\ &\quad + 2\boldsymbol{u}_{N-1}^{\mathrm{T}}\left(\boldsymbol{B}_{N-1}^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{A}_{N-1} + \boldsymbol{S}_{N-1}^{\mathrm{T}}\right)\boldsymbol{x}_{N-1} + \boldsymbol{w}_{N-1}^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{w}_{N-1}\Big|\mathcal{I}_{N-1}\Big]. \end{aligned} \tag{A95}$$

Note that due to the assumption that the process noise $\boldsymbol{w}_{N-1}$ has zero mean and is independent of all other random variables, i. e., independent of $\mathcal{I}_{N-1}$, the terms of Equation (A95) which depend linearly on $\boldsymbol{w}_{N-1}$ are zero and have been dropped. In order to determine the minimal cost-to-go

[24] Note that Equation (A76) allows for $\boldsymbol{y}_k$ to be a function of $\boldsymbol{u}_{k-1}$. Since the linear system given in Equation (A66) is assumed to provide observations which are not directly influenced by the control input, the argument $\boldsymbol{u}_{k-1}$ of h_k is omitted.

$J_{N-1}(\mathcal{I}_{N-1})$, the derivative w.r.t. $\boldsymbol{u}_{N-1}$ of the function to be minimized is determined and set to zero. This leads to the equation

$$\begin{aligned}2\left(\boldsymbol{B}_{N-1}^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{B}_{N-1}+\boldsymbol{R}_{N-1}\right)\boldsymbol{u}_{N-1}&\\ +2\left(\boldsymbol{B}_{N-1}^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{A}_{N-1}+\boldsymbol{S}_{N-1}^{\mathrm{T}}\right)\mathrm{E}_{\boldsymbol{x}_{N-1}|\mathcal{I}_{N-1}}\left[\boldsymbol{x}_{N-1}\middle|\mathcal{I}_{N-1}\right]&=\boldsymbol{0}_{N_u}.\end{aligned} \tag{A96}$$

Thus, the optimal control input which minimizes the cost-to-go for the last step of the control problem reads as

$$\boldsymbol{u}_{N-1}=-\left(\boldsymbol{B}_{N-1}^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{B}_{N-1}+\boldsymbol{R}_{N-1}\right)^{-1}\left(\boldsymbol{B}_{N-1}^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{A}_{N-1}+\boldsymbol{S}_{N-1}^{\mathrm{T}}\right)\mathrm{E}_{\boldsymbol{x}_{N-1}|\mathcal{I}_{N-1}}\left[\boldsymbol{x}_{N-1}\middle|\mathcal{I}_{N-1}\right], \tag{A97}$$

where the inverse always exists by the assumption $\boldsymbol{R}_k > \boldsymbol{0}_{N_u\times N_u}$, $k\in\{0,1,\dots,N-1\}$, and $\boldsymbol{Q}_k \geq \boldsymbol{0}_{N_x\times N_x}$, $k\in\{0,1,\dots,N-1\}$ (cf. Equation A68). In order to simplify the notation, we use the abbreviation

$$\hat{\boldsymbol{x}}_k=\mathrm{E}_{\boldsymbol{x}_k|\mathcal{I}_k}\left[\boldsymbol{x}_k\middle|\mathcal{I}_k\right],\quad k\in\{0,1,\dots,N-1\}, \tag{A98}$$

for the conditional mean of the system state $\boldsymbol{x}_k$ given the information $\mathcal{I}_k$ at time index k. Additionally, the deviation of the system state from the conditional mean is denoted by

$$\tilde{\boldsymbol{x}}_k=\boldsymbol{x}_k-\hat{\boldsymbol{x}}_k,\quad k\in\{0,1,\dots,N-1\}. \tag{A99}$$

Inserting the result from Equation (A97) in Equation (A95) and using the notation from Equations (A98) and (A99), the minimal cost-to-go reads as

$$\begin{aligned}J_{N-1}(\mathcal{I}_{N-1})=\mathrm{E}_{\boldsymbol{x}_{N-1},\boldsymbol{w}_{N-1}|\mathcal{I}_{N-1}}\Big[&\boldsymbol{x}_{N-1}^{\mathrm{T}}\left(\boldsymbol{A}_{N-1}^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{A}_{N-1}+\boldsymbol{Q}_{N-1}\right)\boldsymbol{x}_{N-1}\\ &-\hat{\boldsymbol{x}}_{N-1}\left(\boldsymbol{A}_{N-1}^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{B}_{N-1}+\boldsymbol{S}_{N-1}\right)\left(\boldsymbol{B}_{N-1}^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{B}_{N-1}+\boldsymbol{R}_{N-1}\right)^{-1}\\ &\times\left(\boldsymbol{B}_{N-1}^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{A}_{N-1}+\boldsymbol{S}_{N-1}^{\mathrm{T}}\right)\hat{\boldsymbol{x}}_{N-1}+\boldsymbol{w}_{N-1}^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{w}_{N-1}\Big|\mathcal{I}_{N-1}\Big],\end{aligned} \tag{A100}$$

where it has also been used that

$$\mathrm{E}_{\tilde{\boldsymbol{x}}_k|\mathcal{I}_k}\left[\tilde{\boldsymbol{x}}_k\middle|\mathcal{I}_k\right]=\mathrm{E}_{\boldsymbol{x}_k|\mathcal{I}_k}\left[\boldsymbol{x}_k-\hat{\boldsymbol{x}}_k\middle|\mathcal{I}_k\right]=\boldsymbol{0}_{N_x},\quad k\in\{0,1,\dots,N-1\}. \tag{A101}$$

Finally, due to this property, Equation (A100) can be rewritten in terms of $\boldsymbol{x}_{N-1}$ and $\tilde{\boldsymbol{x}}_{N-1}$:

$$J_{N-1}(\mathcal{I}_{N-1})=\mathrm{E}_{\boldsymbol{x}_{N-1},\boldsymbol{w}_{N-1}|\mathcal{I}_{N-1}}\left[\boldsymbol{x}_{N-1}^{\mathrm{T}}\boldsymbol{K}_{N-1}\boldsymbol{x}_{N-1}+\tilde{\boldsymbol{x}}_{N-1}^{\mathrm{T}}\boldsymbol{P}_{N-1}\tilde{\boldsymbol{x}}_{N-1}+\boldsymbol{w}_{N-1}^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{w}_{N-1}\middle|\mathcal{I}_{N-1}\right], \tag{A102}$$

where

$$\boldsymbol{P}_{N-1}=\left(\boldsymbol{A}_{N-1}^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{B}_{N-1}+\boldsymbol{S}_{N-1}\right)\left(\boldsymbol{B}_{N-1}^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{B}_{N-1}+\boldsymbol{R}_{N-1}\right)^{-1}\left(\boldsymbol{B}_{N-1}^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{A}_{N-1}+\boldsymbol{S}_{N-1}^{\mathrm{T}}\right) \tag{A103}$$

and

$$\boldsymbol{K}_{N-1}=\boldsymbol{A}_{N-1}^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{A}_{N-1}+\boldsymbol{Q}_{N-1}-\boldsymbol{P}_{N-1}. \tag{A104}$$

Before we proceed with the next step of the minimization of the overall cost function given by Equation (A67), note that the last term of $J_{N-1}(\mathcal{I}_{N-1})$ (cf. Equation A102) does not depend on any control input or observation due to the independence assumption. Thus,

$$\mathrm{E}_{\boldsymbol{x}_{N-1},\boldsymbol{w}_{N-1}|\mathcal{I}_{N-1}}\left[\boldsymbol{w}_{N-1}^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{w}_{N-1}\middle|\mathcal{I}_{N-1}\right]=\mathrm{E}_{\boldsymbol{w}_{N-1}}\left[\boldsymbol{w}_{N-1}^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{w}_{N-1}\right], \tag{A105}$$

which is a constant. The same holds for the second term, i. e., the quadratic form in $\tilde{\boldsymbol{x}}_{N-1}$, w.r.t. the control inputs, which is not as obvious and proved in [6, pp. 231-232]. In order to show this independence, note that the state equation (cf. Equation A66) is a linear function of the system state, the control input and the driving noise. Thus, the system state at time index $k+1$ can be expressed as a linear function of the initial state $\boldsymbol{x}_0$, the driving noise sequence and the control input sequence up to time index k. Formally, this means that

$$\boldsymbol{x}_{k+1} = \boldsymbol{\Phi}_k \boldsymbol{\xi}_k + \boldsymbol{\Psi}_k \boldsymbol{\gamma}_k, \quad k \in \{0, 1, \ldots, N-1\}, \tag{A106}$$

where $\boldsymbol{\Phi}_k \in \mathbb{R}^{N_x \times (k+2)N_x}$ and $\boldsymbol{\Psi}_k \in \mathbb{R}^{N_x \times (k+1)N_u}$ are constructed from the system matrices $\boldsymbol{A}_n$ and the input matrices $\boldsymbol{B}_n$, $n \in \{0, 1, \ldots, k\}$, and

$$\begin{aligned} \boldsymbol{\xi}_k &= \left[\boldsymbol{x}_0^{\mathrm{T}}, \boldsymbol{w}_0^{\mathrm{T}}, \boldsymbol{w}_1^{\mathrm{T}}, \ldots, \boldsymbol{w}_k^{\mathrm{T}}\right]^{\mathrm{T}}, \\ \boldsymbol{\gamma}_k &= \left[\boldsymbol{u}_0^{\mathrm{T}}, \boldsymbol{u}_1^{\mathrm{T}}, \ldots, \boldsymbol{u}_k^{\mathrm{T}}\right]^{\mathrm{T}}, \qquad k \in \{0, 1, \ldots, N-1\}. \end{aligned} \tag{A107}$$

This means that the system state can be separated in an uncontrolled part which exclusively depends on the initial state and the process noise, and a controlled part which is a function of the control inputs alone. Note that at time index k, the vector $[\boldsymbol{u}_0^{\mathrm{T}}, \boldsymbol{u}_1^{\mathrm{T}}, \ldots, \boldsymbol{u}_{k-1}^{\mathrm{T}}]^{\mathrm{T}}$ is contained in the information set $\mathcal{I}_k$, see Equation (A69). It follows that

$$\begin{aligned} \tilde{\boldsymbol{x}}_k &= \boldsymbol{x}_k - \hat{\boldsymbol{x}}_k = \boldsymbol{x}_k - \mathrm{E}_{\boldsymbol{x}_k|\mathcal{I}_k}\left[\boldsymbol{x}_k | \mathcal{I}_k\right] \\ &= \boldsymbol{\Phi}_{k-1}\left(\boldsymbol{\xi}_{k-1} - \mathrm{E}_{\boldsymbol{\xi}_{k-1}|\mathcal{I}_k}\left[\boldsymbol{\xi}_{k-1} | \mathcal{I}_k\right]\right) + \boldsymbol{\Psi}_{k-1}\boldsymbol{\gamma}_{k-1} - \boldsymbol{\Psi}_{k-1}\boldsymbol{\gamma}_{k-1}, \end{aligned} \tag{A108}$$

i. e., the summand of $\tilde{\boldsymbol{x}}_k$ containing the vector $[\boldsymbol{u}_0^{\mathrm{T}}, \boldsymbol{u}_1^{\mathrm{T}}, \ldots, \boldsymbol{u}_{k-1}^{\mathrm{T}}]^{\mathrm{T}}$ cancels out. Thus, a potential dependence of this expression on the control input is only due to the information $\mathcal{I}_k$ when computing the conditional expected value of $\boldsymbol{\xi}_{k-1}$. In order to show that this is not the case, it is convenient to express $\mathcal{I}_k$ in terms of $\boldsymbol{\xi}_{k-1}$ and $\boldsymbol{\gamma}_{k-1}$. Using Equations (A66) and (A106), we get

$$\boldsymbol{y}_k = \begin{cases} \boldsymbol{C}_0 \boldsymbol{x}_0 + \boldsymbol{v}_0, & k = 0, \\ \boldsymbol{C}_k \boldsymbol{\Phi}_{k-1} \boldsymbol{\xi}_{k-1} + \boldsymbol{C}_k \boldsymbol{\Psi}_{k-1} \boldsymbol{\gamma}_{k-1} + \boldsymbol{v}_k, & k \in \{1, 2, \ldots, N-1\}. \end{cases} \tag{A109}$$

Since the vector $\boldsymbol{\gamma}_{k-1}$, represented by the sequence $(\boldsymbol{u}_n : n \in \{0, 1, \ldots, k-1\})$, is part of $\mathcal{I}_k$, the information provided by the set $\mathcal{I}_k$ and the set

$$\bar{\mathcal{I}}_k = \begin{cases} \{\boldsymbol{y}_0\}, & k = 0, \\ \{\boldsymbol{y}_0, \bar{\boldsymbol{y}}_1, \ldots, \bar{\boldsymbol{y}}_k, \boldsymbol{u}_0, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_{k-1}\}, & k \in \{1, 2, \ldots, N-1\}, \end{cases} \tag{A110}$$

where

$$\bar{\boldsymbol{y}}_k = \boldsymbol{y}_k - \boldsymbol{C}_k \boldsymbol{\Psi}_{k-1} \boldsymbol{\gamma}_{k-1}, \quad k \in \{1, 2, \ldots, N-1\}, \tag{A111}$$

is equivalent since $\mathcal{I}_k$ and $\bar{\mathcal{I}}_k$ are mapped to each other by an invertible and deterministic function. Finally, a reasonable assumption is that $\boldsymbol{\xi}_{k-1}$, which determines the uncontrolled part of the system state $\boldsymbol{x}_k$, is conditionally independent of the control input given $\boldsymbol{y}_0, \bar{\boldsymbol{y}}_1, \ldots, \bar{\boldsymbol{y}}_k$. This assumption is justified if it is not known how the control inputs are computed and only the numerical values are given. In that case one can model the control inputs as stochastically independent of $\boldsymbol{\xi}_{k-1}$. It is also justified for the setting of the LQG problem where the control inputs $\boldsymbol{u}_k$ are computed as deterministic functions of the available information $\mathcal{I}_k$, $k \in \{0, 1, \ldots, N-1\}$. Here, it is easy to show that $\boldsymbol{u}_k$ can be expressed as a deterministic function of $\boldsymbol{y}_0, \bar{\boldsymbol{y}}_1, \ldots, \bar{\boldsymbol{y}}_k$. Thus, the knowledge of the control inputs does not provide additional information about $\boldsymbol{\xi}_{k-1}$, compared to the observations of the uncontrolled part of the system state (cf. Equation A111). Putting everything together, we observe that

$$\begin{aligned}\tilde{\boldsymbol{x}}_k &= \boldsymbol{\Phi}_{k-1}\left(\boldsymbol{\xi}_{k-1} - \mathrm{E}_{\boldsymbol{\xi}_{k-1}|\mathcal{I}_k}\left[\boldsymbol{\xi}_{k-1} \middle| \mathcal{I}_k\right]\right) \\ &= \boldsymbol{\Phi}_{k-1}\left(\boldsymbol{\xi}_{k-1} - \mathrm{E}_{\boldsymbol{\xi}_{k-1}|\boldsymbol{y}_0,\bar{\boldsymbol{y}}_1,\ldots,\bar{\boldsymbol{y}}_k}\left[\boldsymbol{\xi}_{k-1} \middle| \boldsymbol{y}_0, \bar{\boldsymbol{y}}_1, \ldots, \bar{\boldsymbol{y}}_k\right]\right),\end{aligned} \tag{A112}$$

i. e., $\tilde{\boldsymbol{x}}_k$ is not a function of the control inputs.

Having shown that the choice of the control input has no effect on $\tilde{\boldsymbol{x}}_{N-1}$, the minimal cost-to-go at time index $N-1$ can be written as

$$\begin{aligned}J_{N-1}(\mathcal{I}_{N-1}) = {}& \mathrm{E}_{\boldsymbol{x}_{N-1}|\mathcal{I}_{N-1}}\left[\boldsymbol{x}_{N-1}^{\mathrm{T}}\boldsymbol{K}_{N-1}\boldsymbol{x}_{N-1} \middle| \mathcal{I}_{N-1}\right] \\ &+ \mathrm{E}_{\tilde{\boldsymbol{x}}_{N-1}|\mathcal{I}_{N-1}}\left[\tilde{\boldsymbol{x}}_{N-1}^{\mathrm{T}}\boldsymbol{P}_{N-1}\tilde{\boldsymbol{x}}_{N-1} \middle| \mathcal{I}_{N-1}\right] + \mathrm{E}_{\boldsymbol{w}_{N-1}}\left[\boldsymbol{w}_{N-1}^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{w}_{N-1}\right],\end{aligned} \tag{A113}$$

where the last two summands do not depend on the control input. Following the dynamic programming approach (cf. Equation A89), the minimal cost-to-go for the time index $N-2$ is determined by minimizing the sum of the cost associated with the isolated step at this time index and the remaining cost-to-go, given by $J_{N-1}(\mathcal{I}_{N-1})$, i. e.,

$$\begin{aligned}J_{N-2}(\mathcal{I}_{N-2}) = {}& \min_{\boldsymbol{u}_{N-2}} \mathrm{E}_{\boldsymbol{x}_{N-2},\boldsymbol{w}_{N-2},\boldsymbol{v}_{N-1}|\mathcal{I}_{N-2}}\Big[\boldsymbol{x}_{N-2}^{\mathrm{T}}\boldsymbol{Q}_{N-2}\boldsymbol{x}_{N-2} \\ &+ \boldsymbol{u}_{N-2}^{\mathrm{T}}\boldsymbol{R}_{N-2}\boldsymbol{u}_{N-2} + 2\boldsymbol{u}_{N-2}^{\mathrm{T}}\boldsymbol{S}_{N-2}^{\mathrm{T}}\boldsymbol{x}_{N-2} \\ &+ \mathrm{E}_{\boldsymbol{x}_{N-1}|\mathcal{I}_{N-1}}\left[\boldsymbol{x}_{N-1}^{\mathrm{T}}\boldsymbol{K}_{N-1}\boldsymbol{x}_{N-1} \middle| \mathcal{I}_{N-1}\right] \\ &+ \mathrm{E}_{\tilde{\boldsymbol{x}}_{N-1}|\mathcal{I}_{N-1}}\left[\tilde{\boldsymbol{x}}_{N-1}^{\mathrm{T}}\boldsymbol{P}_{N-1}\tilde{\boldsymbol{x}}_{N-1} \middle| \mathcal{I}_{N-1}\right] + \mathrm{E}_{\boldsymbol{w}_{N-1}}\left[\boldsymbol{w}_{N-1}^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{w}_{N-1}\right] \Big| \mathcal{I}_{N-2}\Big] \\ = {}& \min_{\boldsymbol{u}_{N-2}} \mathrm{E}_{\boldsymbol{x}_{N-2},\boldsymbol{w}_{N-2}|\mathcal{I}_{N-2}}\Big[\boldsymbol{x}_{N-2}^{\mathrm{T}}\boldsymbol{Q}_{N-2}\boldsymbol{x}_{N-2} \\ &+ \boldsymbol{u}_{N-2}^{\mathrm{T}}\boldsymbol{R}_{N-2}\boldsymbol{u}_{N-2} + 2\boldsymbol{u}_{N-2}^{\mathrm{T}}\boldsymbol{S}_{N-2}^{\mathrm{T}}\boldsymbol{x}_{N-2} \\ &+ (\boldsymbol{A}_{N-2}\boldsymbol{x}_{N-2}+\boldsymbol{B}_{N-2}\boldsymbol{u}_{N-2}+\boldsymbol{w}_{N-2})^{\mathrm{T}}\boldsymbol{K}_{N-1}(\boldsymbol{A}_{N-2}\boldsymbol{x}_{N-2}+\boldsymbol{B}_{N-2}\boldsymbol{u}_{N-2}+\boldsymbol{w}_{N-2}) \Big| \mathcal{I}_{N-2}\Big] \\ &+ \mathrm{E}_{\tilde{\boldsymbol{x}}_{N-1}|\mathcal{I}_{N-2}}\left[\tilde{\boldsymbol{x}}_{N-1}^{\mathrm{T}}\boldsymbol{P}_{N-1}\tilde{\boldsymbol{x}}_{N-1} \middle| \mathcal{I}_{N-2}\right] + \mathrm{E}_{\boldsymbol{w}_{N-1}}\left[\boldsymbol{w}_{N-1}^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{w}_{N-1}\right],\end{aligned} \tag{A114}$$

where the minimization is not carried out for the last two terms because they do not depend on $\boldsymbol{u}_{N-2}$. In order to obtain this expression for $J_{N-2}(\mathcal{I}_{N-2})$, note that

$$\begin{aligned}\mathcal{I}_{N-1} &= \{\mathcal{I}_{N-2}, \boldsymbol{y}_{N-1}, \boldsymbol{u}_{N-2}\} \\ &= \{\mathcal{I}_{N-2}, \boldsymbol{C}_{N-1}\left(\boldsymbol{A}_{N-2}\boldsymbol{x}_{N-2} + \boldsymbol{B}_{N-2}\boldsymbol{u}_{N-2} + \boldsymbol{w}_{N-2}\right) + \boldsymbol{v}_{N-1}, \boldsymbol{u}_{N-2}\}.\end{aligned} \tag{A115}$$

Using this equivalence, we get

$$\begin{aligned}&\mathrm{E}_{\boldsymbol{x}_{N-2},\boldsymbol{w}_{N-2},\boldsymbol{v}_{N-1}|\mathcal{I}_{N-2}}\left[\mathrm{E}_{\boldsymbol{x}_{N-1}|\mathcal{I}_{N-1}}\left[\boldsymbol{x}_{N-1}^{\mathrm{T}}\boldsymbol{K}_{N-1}\boldsymbol{x}_{N-1} \middle| \mathcal{I}_{N-1}\right] \middle| \mathcal{I}_{N-2}\right] \\ &= \mathrm{E}_{\boldsymbol{x}_{N-2},\boldsymbol{w}_{N-2},\boldsymbol{y}_{N-1}|\mathcal{I}_{N-2}}\Big[\mathrm{E}_{\boldsymbol{x}_{N-2},\boldsymbol{w}_{N-2}|\mathcal{I}_{N-1}}\Big[(\boldsymbol{A}_{N-2}\boldsymbol{x}_{N-2}+\boldsymbol{B}_{N-2}\boldsymbol{u}_{N-2}+\boldsymbol{w}_{N-2})^{\mathrm{T}}\boldsymbol{K}_{N-1} \\ &\qquad \times (\boldsymbol{A}_{N-2}\boldsymbol{x}_{N-2}+\boldsymbol{B}_{N-2}\boldsymbol{u}_{N-2}+\boldsymbol{w}_{N-2}) \Big| \mathcal{I}_{N-2}, \boldsymbol{y}_{N-1}, \boldsymbol{u}_{N-2}\Big]\Big| \mathcal{I}_{N-2}\Big] \\ &= \mathrm{E}_{\boldsymbol{x}_{N-2},\boldsymbol{w}_{N-2}|\mathcal{I}_{N-2}}\Big[\left(\boldsymbol{A}_{N-2}\boldsymbol{x}_{N-2} + \boldsymbol{B}_{N-2}\boldsymbol{u}_{N-2} + \boldsymbol{w}_{N-2}\right)^{\mathrm{T}}\boldsymbol{K}_{N-1} \\ &\qquad \times \left(\boldsymbol{A}_{N-2}\boldsymbol{x}_{N-2} + \boldsymbol{B}_{N-2}\boldsymbol{u}_{N-2} + \boldsymbol{w}_{N-2}\right) \Big| \mathcal{I}_{N-2}\Big],\end{aligned} \tag{A116}$$

since $\boldsymbol{u}_{N-2}$ is not considered to be a random vector due to the minimization (cf. Equation A114), and

$$\begin{aligned}
&\mathrm{E}_{\boldsymbol{x}_{N-2},\boldsymbol{w}_{N-2},\boldsymbol{v}_{N-1}|\mathcal{I}_{N-2}}\left[\mathrm{E}_{\tilde{\boldsymbol{x}}_{N-1}|\mathcal{I}_{N-1}}\left[\tilde{\boldsymbol{x}}_{N-1}^{\mathrm{T}}\boldsymbol{P}_{N-1}\tilde{\boldsymbol{x}}_{N-1}\middle|\,\mathcal{I}_{N-1}\right]\middle|\,\mathcal{I}_{N-2}\right] \\
&\quad= \mathrm{E}_{\boldsymbol{x}_{N-1},\boldsymbol{y}_{N-1}|\mathcal{I}_{N-2}}\left[\mathrm{E}_{\boldsymbol{x}_{N-1}|\mathcal{I}_{N-1}}\left[(\boldsymbol{x}_{N-1}-\hat{\boldsymbol{x}}_{N-1})^{\mathrm{T}}\boldsymbol{P}_{N-1}(\boldsymbol{x}_{N-1}-\hat{\boldsymbol{x}}_{N-1})\middle|\,\mathcal{I}_{N-2},\boldsymbol{y}_{N-1},\boldsymbol{u}_{N-2}\right]\middle|\,\mathcal{I}_{N-2}\right] \\
&\quad= \mathrm{E}_{\tilde{\boldsymbol{x}}_{N-1}|\mathcal{I}_{N-2}}\left[\tilde{\boldsymbol{x}}_{N-1}^{\mathrm{T}}\boldsymbol{P}_{N-1}\tilde{\boldsymbol{x}}_{N-1}\middle|\,\mathcal{I}_{N-2}\right],
\end{aligned} \tag{A117}$$

where we used the result from Section A5.2 and the fact that $\tilde{\boldsymbol{x}}_k$, $k \in \{0, 1, \ldots, N\}$, does not depend on the control input.

The solution of the minimization problem given by Equation (A114) becomes obvious by comparing it with Equation (A95). The two optimization problems are principally identical except for the different time index, weighting matrices and additive constants, which has no effect on the structure of the solution. Thus, we conclude that (cf. Equation A97)

$$\boldsymbol{u}_{N-2} = -\left(\boldsymbol{B}_{N-2}^{\mathrm{T}}\boldsymbol{K}_{N-1}\boldsymbol{B}_{N-2}+\boldsymbol{R}_{N-2}\right)^{-1}\left(\boldsymbol{B}_{N-2}^{\mathrm{T}}\boldsymbol{K}_{N-1}\boldsymbol{A}_{N-2}+\boldsymbol{S}_{N-2}^{\mathrm{T}}\right)\mathrm{E}_{\boldsymbol{x}_{N-2}|\mathcal{I}_{N-2}}\left[\boldsymbol{x}_{N-2}\middle|\,\mathcal{I}_{N-2}\right], \tag{A118}$$

and that the optimal cost-to-go $J_{N-2}(\mathcal{I}_{N-2})$ can be expressed by a quadratic form in $\boldsymbol{x}_{N-2}$ and additional constants described by $\tilde{\boldsymbol{x}}_{N-2}$ and $\boldsymbol{w}_{N-2}$ (cf. Equation A113). Using these results, we can proceed with the minimization of the overall cost function going backwards in time and get

$$\boldsymbol{u}_k = -\left(\boldsymbol{B}_k^{\mathrm{T}}\boldsymbol{K}_{k+1}\boldsymbol{B}_k + \boldsymbol{R}_k\right)^{-1}\left(\boldsymbol{B}_k^{\mathrm{T}}\boldsymbol{K}_{k+1}\boldsymbol{A}_k + \boldsymbol{S}_k^{\mathrm{T}}\right)\mathrm{E}_{\boldsymbol{x}_k|\mathcal{I}_k}\left[\boldsymbol{x}_k\middle|\,\mathcal{I}_k\right], \tag{A119}$$

where[25]

$$\boldsymbol{K}_k = \boldsymbol{A}_k^{\mathrm{T}}\boldsymbol{K}_{k+1}\boldsymbol{A}_k + \boldsymbol{Q}_k - \boldsymbol{P}_k, \tag{A120}$$

$$\boldsymbol{P}_k = \left(\boldsymbol{A}_k^{\mathrm{T}}\boldsymbol{K}_{k+1}\boldsymbol{B}_k + \boldsymbol{S}_k\right)\left(\boldsymbol{B}_k^{\mathrm{T}}\boldsymbol{K}_{k+1}\boldsymbol{B}_k + \boldsymbol{R}_k\right)^{-1}\left(\boldsymbol{B}_k^{\mathrm{T}}\boldsymbol{K}_{k+1}\boldsymbol{A}_k + \boldsymbol{S}_k^{\mathrm{T}}\right), \tag{A121}$$

for $k \in \{0, 1, \ldots, N-1\}$ and $\boldsymbol{K}_N = \boldsymbol{Q}_N$, and

$$J_k(\mathcal{I}_k) = \mathrm{E}_{\boldsymbol{x}_k|\mathcal{I}_k}\left[\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{K}_k\boldsymbol{x}_k\middle|\,\mathcal{I}_k\right] + \sum_{i=k}^{N-1}\mathrm{E}_{\tilde{\boldsymbol{x}}_i|\mathcal{I}_k}\left[\tilde{\boldsymbol{x}}_i^{\mathrm{T}}\boldsymbol{P}_i\tilde{\boldsymbol{x}}_i\middle|\,\mathcal{I}_k\right] + \sum_{i=k}^{N-1}\mathrm{E}_{\boldsymbol{w}_i}\left[\boldsymbol{w}_i^{\mathrm{T}}\boldsymbol{K}_{i+1}\boldsymbol{w}_i\right]. \tag{A122}$$

Finally, noting that $J^* = \mathrm{E}_{\boldsymbol{x}_0,\boldsymbol{v}_0}\left[J_0(\mathcal{I}_0)\right]$ (cf. Equation A87) and that $\mathcal{I}_0 = \{\boldsymbol{y}_0\} = \{\boldsymbol{C}_0\boldsymbol{x}_0 + \boldsymbol{v}_0\}$, we get the optimal cost of the LQG control problem:

$$\begin{aligned}
J^* &= \mathrm{E}_{\boldsymbol{x}_0}\left[\boldsymbol{x}_0^{\mathrm{T}}\boldsymbol{K}_0\boldsymbol{x}_0\right] + \sum_{k=0}^{N-1}\mathrm{E}_{\tilde{\boldsymbol{x}}_k}\left[\tilde{\boldsymbol{x}}_k^{\mathrm{T}}\boldsymbol{P}_k\tilde{\boldsymbol{x}}_k\right] + \sum_{k=0}^{N-1}\mathrm{E}_{\boldsymbol{w}_k}\left[\boldsymbol{w}_k^{\mathrm{T}}\boldsymbol{K}_{k+1}\boldsymbol{w}_k\right] \\
&= \boldsymbol{\mu}_{\boldsymbol{x}_0}^{\mathrm{T}}\boldsymbol{K}_0\boldsymbol{\mu}_{\boldsymbol{x}_0} + \mathrm{tr}\left[\boldsymbol{K}_0\boldsymbol{C}_{\boldsymbol{x}_0}\right] + \sum_{k=0}^{N-1}\mathrm{tr}\left[\boldsymbol{P}_k\boldsymbol{C}_{\tilde{\boldsymbol{x}}_k}\right] + \sum_{k=0}^{N-1}\mathrm{tr}\left[\boldsymbol{K}_{k+1}\boldsymbol{C}_{\boldsymbol{w}_k}\right].
\end{aligned} \tag{A123}$$

It can be seen that the optimal cost consists of three contributions. The first one is due to the initial state and is given by its mean value and the associated uncertainty which is described by its

[25]It is shown in [6, pp. 149-150] that $\boldsymbol{K}_k$, and consequently $\boldsymbol{P}_k$, are positive semidefinite matrices for $k \in \{0, 1, \ldots, N-1\}$.

covariance matrix. The second contribution is due to the fact that the system state is not known exactly but has to be estimated by computing its conditional mean given the available information $\mathcal{I}_k$ at time index k, i. e., $\hat{\boldsymbol{x}}_k = \mathrm{E}_{\boldsymbol{x}_k|\mathcal{I}_k}[\boldsymbol{x}_k|\mathcal{I}_k]$. The associated cost is given by the covariance matrix of the estimation error $\tilde{\boldsymbol{x}}_k = \boldsymbol{x}_k - \hat{\boldsymbol{x}}_k$. The third contribution is due to the noise process $(\boldsymbol{w}_k : k \in \{0, 1, \ldots, N-1\})$ which drives the system state away from zero.

What is still missing is the description of the computation of the conditional mean $\mathrm{E}_{\boldsymbol{x}_k|\mathcal{I}_k}[\boldsymbol{x}_k|\mathcal{I}_k]$, $k \in \{0, 1, \ldots, N-1\}$, which is not simple in general if arbitrary distributions of the contributing random variables are considered. Nevertheless, due to the assumption of Gaussian random variables and a linear system, which leads to the fact that the evolution of the system state and the associated observations are described by a Gauss-Markov process, the computationally efficient Kalman filter algorithm can be used to determine the conditional expectations and the associated error covariance matrices. This algorithm is introduced in Appendix A7.

The following summary presents the results of this section in compact form. The weighting matrices are assumed to fulfill Equation (A68) and the sequences $(\boldsymbol{w}_k : k \in \{0, 1, \ldots, N-1\})$ and $(\boldsymbol{v}_k : k \in \{0, 1, \ldots, N-1\})$ of independent random vectors have zero mean, are mutually independent and additionally independent of the initial state $\boldsymbol{x}_0$.

Summary A6.1 The optimization problem

$$\underset{\mu_0,\mu_1,\ldots,\mu_{N-1}}{\text{minimize}} \ \mathrm{E}_{\boldsymbol{x}_0,\boldsymbol{w}_0,\ldots,\boldsymbol{w}_{N-1},\boldsymbol{v}_0,\ldots,\boldsymbol{v}_{N-1}}\left[\boldsymbol{x}_N^{\mathrm{T}}\boldsymbol{Q}_N\boldsymbol{x}_N + \sum_{n=0}^{N-1}\begin{bmatrix}\boldsymbol{x}_n\\ \boldsymbol{u}_n\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{Q}_n & \boldsymbol{S}_n\\ \boldsymbol{S}_n^{\mathrm{T}} & \boldsymbol{R}_n\end{bmatrix}\begin{bmatrix}\boldsymbol{x}_n\\ \boldsymbol{u}_n\end{bmatrix}\right],$$

$$\begin{aligned}\text{subject to} \quad \boldsymbol{x}_{k+1} &= \boldsymbol{A}_k\boldsymbol{x}_k + \boldsymbol{B}_k\boldsymbol{u}_k + \boldsymbol{w}_k, && k \in \{0, 1, \ldots, N-1\},\\ \boldsymbol{y}_k &= \boldsymbol{C}_k\boldsymbol{x}_k + \boldsymbol{v}_k, && k \in \{0, 1, \ldots, N-1\},\\ \boldsymbol{u}_k &= \mu_k(\mathcal{I}_k), && k \in \{0, 1, \ldots, N-1\},\end{aligned}$$

where

$$\mathcal{I}_k = \begin{cases}\{\boldsymbol{y}_0\}, & k = 0,\\ \{\boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_k, \boldsymbol{u}_0, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_{k-1}\}, & k \in \{1, 2, \ldots, N-1\},\end{cases}$$

is solved by

$$\boldsymbol{u}_k = -\left(\boldsymbol{B}_k^{\mathrm{T}}\boldsymbol{K}_{k+1}\boldsymbol{B}_k + \boldsymbol{R}_k\right)^{-1}\left(\boldsymbol{B}_k^{\mathrm{T}}\boldsymbol{K}_{k+1}\boldsymbol{A}_k + \boldsymbol{S}_k^{\mathrm{T}}\right)\mathrm{E}_{\boldsymbol{x}_k|\mathcal{I}_k}[\boldsymbol{x}_k|\mathcal{I}_k],$$

where

$$\begin{aligned}\boldsymbol{K}_k &= \boldsymbol{A}_k^{\mathrm{T}}\boldsymbol{K}_{k+1}\boldsymbol{A}_k + \boldsymbol{Q}_k - \boldsymbol{P}_k,\\ \boldsymbol{P}_k &= \left(\boldsymbol{A}_k^{\mathrm{T}}\boldsymbol{K}_{k+1}\boldsymbol{B}_k + \boldsymbol{S}_k\right)\left(\boldsymbol{B}_k^{\mathrm{T}}\boldsymbol{K}_{k+1}\boldsymbol{B}_k + \boldsymbol{R}_k\right)^{-1}\left(\boldsymbol{B}_k^{\mathrm{T}}\boldsymbol{K}_{k+1}\boldsymbol{A}_k + \boldsymbol{S}_k^{\mathrm{T}}\right),\end{aligned}$$

for $k \in \{0, 1, \ldots, N-1\}$, and $\boldsymbol{K}_N = \boldsymbol{Q}_N$. The optimal value of the cost function is

$$J^* = \mathrm{E}_{\boldsymbol{x}_0}\left[\boldsymbol{x}_0^{\mathrm{T}}\boldsymbol{K}_0\boldsymbol{x}_0\right] + \sum_{k=0}^{N-1}\mathrm{E}_{\tilde{\boldsymbol{x}}_k}\left[\tilde{\boldsymbol{x}}_k^{\mathrm{T}}\boldsymbol{P}_k\tilde{\boldsymbol{x}}_k\right] + \sum_{k=0}^{N-1}\mathrm{E}_{\boldsymbol{w}_k}\left[\boldsymbol{w}_k^{\mathrm{T}}\boldsymbol{K}_{k+1}\boldsymbol{w}_k\right].$$

A6.3 Infinite Horizon LQG Control with Average Cost

In practice, control systems can operate over a very long time. For the LQG control problem, this means that the horizon N becomes large and, in the limit, grows to infinity. Having a look at Equation (A123), it becomes clear that in this case the optimal cost J^* grows without a bound and thus does not allow for an evaluation of the actual control performance. One approach to the solution of this problem is not to consider J^*, but the *average cost per stage* of the control problem, i. e.,

$$J_\infty = \lim_{N\to\infty} \frac{1}{N} J, \tag{A124}$$

where J is the LQG cost function given by Equation (A67). Note that while J is always finite for finite values of N if $\boldsymbol{x}_k$ and $\boldsymbol{u}_k$ have bounded second order moments, this is not necessarily the case for J_∞.

Of special interest is the investigation of time-invariant systems with stationary process and observation noise sequences and constant weighting matrices for the LQG cost function. The reason is that for this scenario, it turns out that the optimal controller is also time-invariant and can be determined with low computational and memory requirements. For the stationary and time-invariant scenario, assume that

$$\begin{aligned} \boldsymbol{x}_{k+1} &= \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{B}\boldsymbol{u}_k + \boldsymbol{w}_k, \\ \boldsymbol{y}_k &= \boldsymbol{C}\boldsymbol{x}_k + \boldsymbol{v}_k, \qquad\qquad k \in \mathbb{N}_0, \end{aligned} \tag{A125}$$

i. e., a Linear Time-Invariant (LTI) system is considered. Additionally, $(\boldsymbol{w}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{v}_k : k \in \mathbb{N}_0)$ are identically and independently distributed (i.i.d.) sequences which are additionally mutually independent and independent of the initial system state $\boldsymbol{x}_0 \sim \mathcal{N}\left(\boldsymbol{\mu}_{\boldsymbol{x}_0}, \boldsymbol{C}_{\boldsymbol{x}_0}\right)$. Since the noise sequences are assumed to be i.i.d., i. e., they are also stationary, and Gaussian, we have $\boldsymbol{w}_k \sim \mathcal{N}\left(\boldsymbol{0}_{N_x}, \boldsymbol{C}_{\boldsymbol{w}}\right)$ and $\boldsymbol{v} \sim \mathcal{N}\left(\boldsymbol{0}_{N_y}, \boldsymbol{C}_{\boldsymbol{v}}\right)$, $k \in \mathbb{N}_0$. Finally, the cost function is given by (cf. Equations A67 and A68)

$$J = \mathrm{E}_{\boldsymbol{x}_0,\boldsymbol{w}_0,\ldots,\boldsymbol{w}_{N-1},\boldsymbol{v}_0,\ldots,\boldsymbol{v}_{N-1}}\left[\boldsymbol{x}_N^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}_N + \sum_{n=0}^{N-1}\begin{bmatrix}\boldsymbol{x}_n\\ \boldsymbol{u}_n\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{Q} & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{R}\end{bmatrix}\begin{bmatrix}\boldsymbol{x}_n\\ \boldsymbol{u}_n\end{bmatrix}\right], \tag{A126}$$

with

$$\boldsymbol{Q} \geq \boldsymbol{0}_{N_x\times N_x}, \qquad \boldsymbol{R} > \boldsymbol{0}_{N_u\times N_u}, \qquad \begin{bmatrix}\boldsymbol{Q} & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{R}\end{bmatrix} \geq \boldsymbol{0}_{(N_x+N_u)\times(N_x+N_u)}. \tag{A127}$$

For this scenario, the average optimal cost per stage can be obtained using the results from Section A6.2 and is given by (cf. Equation A123)

$$\frac{1}{N}J^* = \frac{1}{N}\left(\boldsymbol{\mu}_{\boldsymbol{x}_0}^{\mathrm{T}}\boldsymbol{K}_0\boldsymbol{\mu}_{\boldsymbol{x}_0} + \operatorname{tr}\left[\boldsymbol{K}_0\boldsymbol{C}_{\boldsymbol{x}_0}\right]\right) + \frac{1}{N}\sum_{k=0}^{N-1}\left(\operatorname{tr}\left[\boldsymbol{P}_k\boldsymbol{C}_{\tilde{\boldsymbol{x}}_k}\right] + \operatorname{tr}\left[\boldsymbol{K}_{k+1}\boldsymbol{C}_{\boldsymbol{w}}\right]\right), \tag{A128}$$

where $\boldsymbol{K}_k$ and $\boldsymbol{P}_k$, $k \in \{0, 1, \ldots, N-1\}$, are determined by the iteration (cf. Equations A120 and A121)

$$\boldsymbol{K}_k = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}_{k+1}\boldsymbol{A} - \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}_{k+1}\boldsymbol{B} + \boldsymbol{S}\right)\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}_{k+1}\boldsymbol{B} + \boldsymbol{R}\right)^{-1}\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}_{k+1}\boldsymbol{A} + \boldsymbol{S}^{\mathrm{T}}\right) + \boldsymbol{Q}, \tag{A129}$$

$$\boldsymbol{P}_k = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}_{k+1}\boldsymbol{A} - \boldsymbol{K}_k + \boldsymbol{Q}, \tag{A130}$$

with initial condition $\boldsymbol{K}_N = \boldsymbol{Q}$. The covariance matrix $\boldsymbol{C}_{\tilde{\boldsymbol{x}}_k}$, $k \in \{0, 1, \ldots, N-1\}$, of the state estimation error is obtained by the Kalman filter iterations (cf. Equations A182 and A198) and reads as

$$\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}_{k+1}} = \boldsymbol{A}\left(\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}_k} - \boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}_k}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}_k}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}_k}\right)\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{w}}, \tag{A131}$$

$$\boldsymbol{C}_{\tilde{\boldsymbol{x}}_k} = \boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}_k} - \boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}_k}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}_k}, \tag{A132}$$

with initial condition $\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}_0} = \boldsymbol{C}_{\boldsymbol{x}_0}$. Before the minimization of J_∞ is discussed, we analyze the behavior of the average optimal cost, i. e., $\frac{1}{N}J^*$, in the limit for $N \to \infty$. In this case and under the assumptions from Section A3, the iterations in Equations (A129) and (A131) converge to their respective unique positive semidefinite steady state solutions given by the DAREs

$$\boldsymbol{K} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B} + \boldsymbol{S}\right)\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B} + \boldsymbol{R}\right)^{-1}\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} + \boldsymbol{S}^{\mathrm{T}}\right) + \boldsymbol{Q} \tag{A133}$$

and

$$\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}} = \boldsymbol{A}\left(\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}} - \boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\right)\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{w}}, \tag{A134}$$

irrespective of the initial conditions. Consequently, the matrix $\boldsymbol{P}_k$ converges to $\boldsymbol{P}$ and the error covariance matrix $\boldsymbol{C}_{\tilde{\boldsymbol{x}}_k}$ to $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}$. Using this result for Equation (A128), it can be seen that the constant term depending on $\boldsymbol{\mu}_{\boldsymbol{x}_0}$ and $\boldsymbol{C}_{\boldsymbol{x}_0}$ vanishes in the limit for $N \to \infty$, whereas the remaining terms converge to constant values. Thus, we obtain

$$\lim_{N\to\infty}\frac{1}{N}J^* = \operatorname{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\right] + \operatorname{tr}\left[\boldsymbol{K}\boldsymbol{C}_{\boldsymbol{w}}\right]. \tag{A135}$$

The optimal controller becomes the LTI system

$$\boldsymbol{u}_k = \boldsymbol{L}\hat{\boldsymbol{x}}_k, \qquad k \in \mathbb{N}_0, \tag{A136}$$

where

$$\boldsymbol{L} = -\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B} + \boldsymbol{R}\right)^{-1}\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} + \boldsymbol{S}^{\mathrm{T}}\right), \tag{A137}$$

and the state estimate $\hat{\boldsymbol{x}}_k$ determined by the steady state Kalman filter (see Section A7), i. e.,

$$\begin{aligned} \hat{\boldsymbol{x}}^{\mathrm{P}}_{k+1} &= \boldsymbol{A}\hat{\boldsymbol{x}}_k + \boldsymbol{B}\boldsymbol{u}_k, \\ \hat{\boldsymbol{x}}_k &= \hat{\boldsymbol{x}}^{\mathrm{P}}_k + \boldsymbol{G}\left(\boldsymbol{y}_k - \boldsymbol{C}\hat{\boldsymbol{x}}^{\mathrm{P}}_k\right), \qquad k \in \mathbb{N}_0, \end{aligned} \tag{A138}$$

with

$$\boldsymbol{G} = \boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}}\right)^{-1}. \tag{A139}$$

It is an appealing property of this solution that it can be computed offline and implemented as an LTI filter. Additionally, it can be used as an approximate solution for the finite horizon control problem when N is sufficiently large.

Now we turn to the minimization of J_∞ given by Equation (A124) because so far the limit of the optimal LQG cost has been considered and not the minimum of the infinite horizon average cost. In general, the treatment of infinite horizon average cost problems, especially for problems with infinite state or control spaces, needs advanced mathematical tools which will not be discussed here. An introduction to these problems and associated solution methods can be found in, e. g., [6, Chapter 7] while a more detailed discussion is presented in [136, Chapter 4]. The methods introduced there are applied to the infinite horizon LQG problem with average cost in [136, Section 4.6.5] when the system state is perfectly known, i. e., when $\boldsymbol{y}_k = \boldsymbol{x}_k$, $k \in \mathbb{N}_0$, and it is shown that the solution according to Equations (A135) and (A136) (with $\hat{\boldsymbol{x}}_k$ replaced by $\boldsymbol{x}_k$ and $\boldsymbol{C}_{\tilde{\boldsymbol{x}}}$ by $\mathbf{0}_{N_x \times N_x}$) is indeed optimal.

A different approach to the solution of the infinite horizon LQG problem with average cost is the so-called $\mathcal{H}_2$ minimization [137–139]. The basis for this approach is the fact that the second moment of a stationary random sequence can be calculated using its power spectral density, see, e. g., [4, pp. 468-469], [31, pp. 194-195 and 203-204] or [34, pp. 420-421]. Thus, assuming that $(\boldsymbol{\xi}_k : k \in \mathbb{N}_0)$ is a K-dimensional stationary[26] random sequence, we get

$$\mathrm{E}_{\boldsymbol{\xi}_k}\left[\|\boldsymbol{\xi}_k\|_2^2\right] = \mathrm{tr}\left[\frac{1}{2\pi}\int_{-\pi}^{\pi} \boldsymbol{S}_{\boldsymbol{\xi}}(\mathrm{e}^{\mathrm{j}\omega})\,\mathrm{d}\omega\right], \qquad k \in \mathbb{N}_0, \tag{A140}$$

where $\boldsymbol{S}_{\boldsymbol{\xi}} \colon \mathbb{C} \to \mathbb{C}^{K \times K}$, is the element-wise z-transform of the correlation (matrix) sequence

$$\boldsymbol{R}_{\boldsymbol{\xi}}(k) = \mathrm{E}_{\boldsymbol{\xi}_n, \boldsymbol{\xi}_{n+k}}\left[\boldsymbol{\xi}_n \boldsymbol{\xi}_{n+k}^{\mathrm{T}}\right], \qquad n \in \mathbb{N}_0, k+n \in \mathbb{N}_0, \tag{A141}$$

cf. [4, p. 468]. Assume further that $(\boldsymbol{\xi}_k : k \in \mathbb{N}_0)$ is the output of an LTI system with transfer matrix $\boldsymbol{A} \colon \mathbb{C} \to \mathbb{C}^{K \times L}$, $L \in \mathbb{N}$, which is driven by a white noise sequence $(\boldsymbol{n}_k : k \in \mathbb{N}_0)$ with $\boldsymbol{S}_{\boldsymbol{n}}(\mathrm{e}^{\mathrm{j}\omega}) = \sigma_n^2 \mathbf{I}_{L \times L}$, $\omega \in \{-\pi, \pi\}$. Then, the second moment of the output sequence is given by (cf. [4, p. 469])

$$\mathrm{E}_{\boldsymbol{\xi}_k}\left[\|\boldsymbol{\xi}_k\|_2^2\right] = \mathrm{tr}\left[\frac{1}{2\pi}\int_{-\pi}^{\pi} \boldsymbol{A}(\mathrm{e}^{\mathrm{j}\omega})\boldsymbol{A}^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\omega})\,\mathrm{d}\omega\right]\sigma_n^2 = \|\boldsymbol{A}\|_2^2 \sigma_n^2, \tag{A142}$$

i. e., by the squared $\mathcal{H}_2$ norm of the system transfer matrix (see, e. g., [138, p. 85] or [139, p. 23]). Consequently, the minimization of $\mathrm{E}_{\boldsymbol{\xi}_k}\left[\|\boldsymbol{\xi}_k\|_2^2\right]$ is equivalent to the minimization of the $\mathcal{H}_2$ norm of the corresponding system transfer matrix.

The $\mathcal{H}_2$ minimization problem can be identified as an LQG control problem with infinite horizon and average cost. First, we have to assume that a time-invariant controller exists which leads to asymptotically and jointly stationary state and control sequences $(\boldsymbol{x}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{u}_k : k \in \mathbb{N}_0)$, respectively. Second, define

$$\boldsymbol{\xi}_k = \boldsymbol{E}\boldsymbol{x}_k + \boldsymbol{F}\boldsymbol{u}_k, \qquad k \in \mathbb{N}_0, \tag{A143}$$

where $\boldsymbol{E}$ and $\boldsymbol{F}$ are constant matrices of appropriate dimensions. Since the state and control sequences are asymptotically stationary, we consider

$$\lim_{k\to\infty} \mathrm{E}_{\boldsymbol{\xi}_k}\left[\|\boldsymbol{\xi}_k\|_2^2\right] = \lim_{k\to\infty} \mathrm{E}_{\boldsymbol{x}_k, \boldsymbol{u}_k}\left[\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{E}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{x}_k + \boldsymbol{u}_k^{\mathrm{T}}\boldsymbol{F}^{\mathrm{T}}\boldsymbol{F}\boldsymbol{u}_k + 2\boldsymbol{u}_k^{\mathrm{T}}\boldsymbol{F}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{x}_k\right], \tag{A144}$$

which can be identified to be the cost function J_∞ (cf. Equation A124) by setting $\boldsymbol{Q} = \boldsymbol{E}^{\mathrm{T}}\boldsymbol{E}$, $\boldsymbol{R} = \boldsymbol{F}^{\mathrm{T}}\boldsymbol{F}$ and $\boldsymbol{S} = \boldsymbol{E}^{\mathrm{T}}\boldsymbol{F}$ and using the fact that due to the convergence of the second order moments of the state and control sequence to their stationary values, the limit of the average in Equation (A124) and the limit in Equation (A144) have the same value. Finally, [137] and [138, Chapter 6.5] show that the solutions of the $\mathcal{H}_2$ optimization and of the stationary LQG control problem are the same, i. e., they are obtained by the solution of the two DAREs given in Equations (A133) and (A134) which are used to compute a linear function of the optimal state estimate. The

[26]This implies that the components of $\boldsymbol{\xi}_k$ are jointly stationary.

optimal value of the cost function is the sum of the contributions due to the process noise and the state estimation error (cf. Equation A135). We conclude that the solution given in Equations (A136)-(A139) is the optimal linear time-invariant controller which minimizes J_∞.

Summary A6.2 For the LTI system given in Equation (A125), the optimal linear time-invariant controller which minimizes the infinite horizon average cost function

$$J_\infty = \lim_{N\to\infty} \frac{1}{N} J,$$

where J is given by Equation (A126), reads as

$$\boldsymbol{u}_k = -\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B} + \boldsymbol{R}\right)^{-1}\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} + \boldsymbol{S}^{\mathrm{T}}\right)\hat{\boldsymbol{x}}_k, \qquad k \in \mathbb{N}_0,$$

where the optimal state estimate $\hat{\boldsymbol{x}}_k = \mathrm{E}_{\boldsymbol{x}_k|\mathcal{I}_k}\left[\boldsymbol{x}_k|\mathcal{I}_k\right]$, with $\mathcal{I}_k$ given by Equation (A94), is determined by the steady state Kalman filter (cf. Section A7), i. e.,

$$\begin{aligned}\hat{\boldsymbol{x}}^{\mathrm{P}}_{k+1} &= \boldsymbol{A}\hat{\boldsymbol{x}}_k + \boldsymbol{B}\boldsymbol{u}_k,\\ \hat{\boldsymbol{x}}_k &= \hat{\boldsymbol{x}}^{\mathrm{P}}_k + \boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}}\right)^{-1}\left(\boldsymbol{y}_k - \boldsymbol{C}\hat{\boldsymbol{x}}^{\mathrm{P}}_k\right), \qquad k \in \mathbb{N}_0.\end{aligned}$$

The result requires the stabilizing solutions of the two DAREs

$$\boldsymbol{K} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B} + \boldsymbol{S}\right)\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{B} + \boldsymbol{R}\right)^{-1}\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} + \boldsymbol{S}^{\mathrm{T}}\right) + \boldsymbol{Q}$$

and

$$\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}} = \boldsymbol{A}\left(\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}} - \boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\right)\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{w}}.$$

The optimal value of the cost function is given by

$$J^*_\infty = \mathrm{tr}\left[\boldsymbol{P}\boldsymbol{C}_{\tilde{\boldsymbol{x}}}\right] + \mathrm{tr}\left[\boldsymbol{K}\boldsymbol{C}_{\boldsymbol{w}}\right],$$

with

$$\boldsymbol{P} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{A} - \boldsymbol{K} + \boldsymbol{Q}$$

and

$$\boldsymbol{C}_{\tilde{\boldsymbol{x}}} = \boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}} - \boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}}\right)^{-1}\boldsymbol{C}\boldsymbol{C}^{\mathrm{P}}_{\tilde{\boldsymbol{x}}}.$$

A7. The Kalman Filter

The Kalman filter [140] is a very popular method for the optimal estimation of a random sequence which is described by a Markov model using incomplete and noisy observations. For the case of Gauss-Markov sequences, it is the optimal estimator in the mean square sense, without the Gaussianity the best linear estimator in that sense [37, Section 3.2]. By introducing a state space model for the sequence to be estimated, it allows for an efficient sequential estimation procedure which can be performed on-line when new observations become available. In the following, we present the basic idea of sequential estimation for general Markov sequences, while in Section A7.2 the Kalman filter algorithm is derived using a state space description of the Gauss-Markov sequence to be estimated.

Let $(\boldsymbol{x}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{y}_k : k \in \mathbb{N}_0)$ be two random sequences which are described by their joint distribution functions. The problem is to infer from the *observation sequence* $(\boldsymbol{y}_k : k \in \mathbb{N}_0)$ on the sequence of interest, $(\boldsymbol{x}_k : k \in \mathbb{N}_0)$, which cannot be observed directly.

Given a set of observed random vectors $\boldsymbol{y}_k$, $k \in \{i_1, i_2, \ldots, i_N\} \subset \mathbb{N}_0$, all information about a set of random vectors $\boldsymbol{x}_\ell$, $\ell \in \{j_1, j_2, \ldots, j_M\} \subset \mathbb{N}_0$, is provided by the conditional PDF

$$f_{\boldsymbol{x}_{j_1},\ldots,\boldsymbol{x}_{j_M}|\boldsymbol{y}_{i_1},\ldots,\boldsymbol{y}_{i_N}}\left(\boldsymbol{x}_{j_1},\ldots,\boldsymbol{x}_{j_M}|\boldsymbol{y}_{i_1},\ldots,\boldsymbol{y}_{i_N}\right), \tag{A145}$$

where we assume that the PDF exists and is well defined. Of special interest is the *filtering* problem, i. e., the determination of the conditional PDFs

$$f_{\boldsymbol{x}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_k}\left(\boldsymbol{x}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_k\right), \quad k \in \mathbb{N}_0. \tag{A146}$$

This can be interpreted as making inference about an unknown quantity, described by the random vector $\boldsymbol{x}_k$, at a time index k from measurements or observations $\boldsymbol{y}_i$, $i \in \{0, 1, \ldots, k\}$, which are obtained sequentially from an initial time 0 up to k. For a certain class of stochastic models, the determination of the conditional PDF can be performed sequentially when a new observation becomes available, which allows for the development of on-line estimation algorithms that rely on the knowledge of this PDF.

A7.1 Sequential Estimation for Markov Models

Assume that the joint PDFs of $(\boldsymbol{x}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{y}_k : k \in \mathbb{N}_0)$ can be written as

$$\begin{aligned} &f_{\boldsymbol{x}_0,\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_N}\left(\boldsymbol{x}_0,\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_N\right) \\ &\qquad = f_{\boldsymbol{x}_0}\left(\boldsymbol{x}_0\right)\prod_{k=1}^{N} f_{\boldsymbol{x}_k|\boldsymbol{x}_{k-1}}\left(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}\right)\prod_{k=0}^{N} f_{\boldsymbol{y}_k|\boldsymbol{x}_k}\left(\boldsymbol{y}_k|\boldsymbol{x}_k\right), \end{aligned} \tag{A147}$$

for all $N \in \mathbb{N}_0$.[27] This is possible if $(\boldsymbol{x}_k : k \in \mathbb{N}_0)$ is a Markov sequence, i. e.,

$$f_{\boldsymbol{x}_0,\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N}\left(\boldsymbol{x}_0,\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N\right) = f_{\boldsymbol{x}_0}\left(\boldsymbol{x}_0\right)\prod_{k=1}^{N} f_{\boldsymbol{x}_k|\boldsymbol{x}_{k-1}}\left(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}\right) \tag{A148}$$

for all $N \in \mathbb{N}_0$, where $f_{\boldsymbol{x}_0}(\boldsymbol{x}_0)$ is the PDF of the initial random vector and $f_{\boldsymbol{x}_{k+1}|\boldsymbol{x}_k}(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k)$, $k \in \mathbb{N}_0$, is the transition density of the sequence. Additionally, Equation (A147) implies that the

[27] We assume that the PDF exists and is well defined, which will also be the case in the remainder of this section.

conditional PDF of the observation sequence given the state sequence factorizes as

$$f_{\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_N|\boldsymbol{x}_0,\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N}\left(\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_N|\boldsymbol{x}_0,\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N\right)=\prod_{k=0}^{N} f_{\boldsymbol{y}_k|\boldsymbol{x}_k}\left(\boldsymbol{y}_k|\boldsymbol{x}_k\right), \tag{A149}$$

for all $N \in \mathbb{N}_0$. Loosely speaking, this requires that the observation $\boldsymbol{y}_k$, $k \in \mathbb{N}_0$, only depends on the state $\boldsymbol{x}_k$ at time index k. Formally, Equation (A149) holds if $\boldsymbol{y}_k$ is conditionally independent of all states $(\boldsymbol{x}_i : i \in \mathbb{N}_0, i \neq k)$ and observations $(\boldsymbol{y}_i : i \in \mathbb{N}_0, i \neq k)$ given the state $\boldsymbol{x}_k$ for all $k \in \mathbb{N}_0$. Consequently, the stochastic model of the state and observation sequence is fully determined by the transition densities $f_{\boldsymbol{x}_{k+1}|\boldsymbol{x}_k}\left(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k\right)$ and the observation densities $f_{\boldsymbol{y}_k|\boldsymbol{x}_k}\left(\boldsymbol{y}_k|\boldsymbol{x}_k\right)$ for $k \in \mathbb{N}_0$ as well as the PDF of the initial state $f_{\boldsymbol{x}_0}\left(\boldsymbol{x}_0\right)$.

In the following, it will be useful to note that the model assumptions which lead to Equation (A147) determine the conditional independence of $\boldsymbol{y}_k$ and $(\boldsymbol{y}_i : i \in \{0,1,\ldots,k-1\})$ given $\boldsymbol{x}_k$, i. e.,

$$f_{\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1},\boldsymbol{y}_k|\boldsymbol{x}_k}\left(\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1},\boldsymbol{y}_k|\boldsymbol{x}_k\right)=f_{\boldsymbol{y}_k|\boldsymbol{x}_k}\left(\boldsymbol{y}_k|\boldsymbol{x}_k\right) f_{\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}|\boldsymbol{x}_k}\left(\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}|\boldsymbol{x}_k\right), \tag{A150}$$

which follows directly from the assumption of the conditional independence above. Additionally, the state $\boldsymbol{x}_{k+1}$ is conditionally independent of $(\boldsymbol{y}_i : i \in \{0,\ldots,k\})$ given $\boldsymbol{x}_k$, which results from

$$\begin{aligned}
&f_{\boldsymbol{x}_{k+1},\boldsymbol{y}_0,\ldots,\boldsymbol{y}_k|\boldsymbol{x}_k}\left(\boldsymbol{x}_{k+1},\boldsymbol{y}_0,\ldots,\boldsymbol{y}_k|\boldsymbol{x}_k\right)=\frac{1}{f_{\boldsymbol{x}_k}\left(\boldsymbol{x}_k\right)} f_{\boldsymbol{x}_k,\boldsymbol{x}_{k+1},\boldsymbol{y}_0,\ldots,\boldsymbol{y}_k}\left(\boldsymbol{x}_k,\boldsymbol{x}_{k+1},\boldsymbol{y}_0,\ldots,\boldsymbol{y}_k\right)\\
&=\frac{1}{f_{\boldsymbol{x}_k}\left(\boldsymbol{x}_k\right)}\int\cdots\iint f_{\boldsymbol{x}_0}\left(\boldsymbol{x}_0\right)\prod_{i=1}^{k+1} f_{\boldsymbol{x}_i|\boldsymbol{x}_{i-1}}\left(\boldsymbol{x}_i|\boldsymbol{x}_{i-1}\right)\prod_{i=0}^{k+1} f_{\boldsymbol{y}_i|\boldsymbol{x}_i}\left(\boldsymbol{y}_i|\boldsymbol{x}_i\right)\mathrm{d}\boldsymbol{x}_0\ldots\mathrm{d}\boldsymbol{x}_{k-1}\,\mathrm{d}\boldsymbol{y}_{k+1}\\
&=\frac{f_{\boldsymbol{x}_{k+1}|\boldsymbol{x}_k}\left(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k\right)}{f_{\boldsymbol{x}_k}\left(\boldsymbol{x}_k\right)}\int\cdots\int f_{\boldsymbol{x}_0}\left(\boldsymbol{x}_0\right)\prod_{i=1}^{k} f_{\boldsymbol{x}_i|\boldsymbol{x}_{i-1}}\left(\boldsymbol{x}_i|\boldsymbol{x}_{i-1}\right)\prod_{i=0}^{k} f_{\boldsymbol{y}_i|\boldsymbol{x}_i}\left(\boldsymbol{y}_i|\boldsymbol{x}_i\right)\mathrm{d}\boldsymbol{x}_0\ldots\mathrm{d}\boldsymbol{x}_{k-1}\\
&=\frac{f_{\boldsymbol{x}_{k+1}|\boldsymbol{x}_k}\left(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k\right)}{f_{\boldsymbol{x}_k}\left(\boldsymbol{x}_k\right)} f_{\boldsymbol{x}_k,\boldsymbol{y}_0,\ldots\boldsymbol{y}_k}\left(\boldsymbol{x}_k,\boldsymbol{y}_0,\ldots\boldsymbol{y}_k\right)\\
&=f_{\boldsymbol{x}_{k+1}|\boldsymbol{x}_k}\left(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k\right) f_{\boldsymbol{y}_0,\ldots\boldsymbol{y}_k|\boldsymbol{x}_k}\left(\boldsymbol{y}_0,\ldots\boldsymbol{y}_k|\boldsymbol{x}_k\right).
\end{aligned} \tag{A151}$$

In the following, it will be shown how to determine the PDFs given in Equation (A146) sequentially. The result relies on the the application of Bayes' theorem for PDFs [34, p. 224]. We start by noting that

$$f_{\boldsymbol{x}_0|\boldsymbol{y}_0}\left(\boldsymbol{x}_0|\boldsymbol{y}_0\right)=\frac{f_{\boldsymbol{y}_0|\boldsymbol{x}_0}\left(\boldsymbol{y}_0|\boldsymbol{x}_0\right) f_{\boldsymbol{x}_0}\left(\boldsymbol{x}_0\right)}{f_{\boldsymbol{y}_0}\left(\boldsymbol{y}_0\right)}, \tag{A152}$$

where $f_{\boldsymbol{y}_0}\left(\boldsymbol{y}_0\right)=\int f_{\boldsymbol{y}_0|\boldsymbol{x}_0}\left(\boldsymbol{y}_0|\boldsymbol{x}_0\right) f_{\boldsymbol{x}_0}\left(\boldsymbol{x}_0\right)\mathrm{d}\boldsymbol{x}_0$. This expression is determined by the given stochastic model. Assume now that the PDF $f_{\boldsymbol{x}_{k-1}|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}}\left(\boldsymbol{x}_{k-1}|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}\right)$ is known for the index $(k-1) \in \mathbb{N}_0$. Then, the marginalization w.r.t. $\boldsymbol{x}_{k-1}$ and the application of the model assumption (cf. Equation A147) provides the density

$$\begin{aligned}
&f_{\boldsymbol{x}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}}\left(\boldsymbol{x}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}\right)=\int f_{\boldsymbol{x}_k,\boldsymbol{x}_{k-1}|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}}\left(\boldsymbol{x}_k,\boldsymbol{x}_{k-1}|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}\right)\mathrm{d}\boldsymbol{x}_{k-1}\\
&\quad=\int f_{\boldsymbol{x}_k|\boldsymbol{x}_{k-1},\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}}\left(\boldsymbol{x}_k|\boldsymbol{x}_{k-1},\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}\right) f_{\boldsymbol{x}_{k-1}|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}}\left(\boldsymbol{x}_{k-1}|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}\right)\mathrm{d}\boldsymbol{x}_{k-1}\\
&\quad\overset{(*)}{=}\int f_{\boldsymbol{x}_k|\boldsymbol{x}_{k-1}}\left(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}\right) f_{\boldsymbol{x}_{k-1}|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}}\left(\boldsymbol{x}_{k-1}|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}\right)\mathrm{d}\boldsymbol{x}_{k-1}.
\end{aligned} \tag{A153}$$

Note that the equation $(*)$ holds due to the conditional independence of $(\boldsymbol{x}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{y}_k : k \in \mathbb{N}_0)$. Analogously, we get

$$\begin{aligned} f_{\boldsymbol{y}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}}\left(\boldsymbol{y}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}\right) &= \int f_{\boldsymbol{y}_k,\boldsymbol{x}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}}\left(\boldsymbol{y}_k,\boldsymbol{x}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}\right)\mathrm{d}\boldsymbol{x}_k \\ &= \int f_{\boldsymbol{y}_k|\boldsymbol{x}_k,\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}}\left(\boldsymbol{y}_k|\boldsymbol{x}_k,\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}\right) f_{\boldsymbol{x}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}}\left(\boldsymbol{x}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}\right)\mathrm{d}\boldsymbol{x}_k \\ &= \int f_{\boldsymbol{y}_k|\boldsymbol{x}_k}\left(\boldsymbol{y}_k|\boldsymbol{x}_k\right) f_{\boldsymbol{x}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}}\left(\boldsymbol{x}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}\right)\mathrm{d}\boldsymbol{x}_k. \end{aligned} \tag{A154}$$

Having computed the two PDFs in Equations (A153) and (A154), the same reasoning without the marginalization step is applied to get the desired density (cf. Equation A146) at time index $k \in \mathbb{N}_0$:

$$\begin{aligned} f_{\boldsymbol{x}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_k}\left(\boldsymbol{x}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_k\right) &= \frac{f_{\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1},\boldsymbol{y}_k|\boldsymbol{x}_k}\left(\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1},\boldsymbol{y}_k|\boldsymbol{x}_k\right) f_{\boldsymbol{x}_k}\left(\boldsymbol{x}_k\right)}{f_{\boldsymbol{y}_0,\ldots,\boldsymbol{y}_k}\left(\boldsymbol{y}_0,\ldots,\boldsymbol{y}_k\right)} \\ &= \frac{f_{\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}|\boldsymbol{x}_k}\left(\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}|\boldsymbol{x}_k\right) f_{\boldsymbol{x}_k}\left(\boldsymbol{x}_k\right) f_{\boldsymbol{y}_k|\boldsymbol{x}_k}\left(\boldsymbol{y}_k|\boldsymbol{x}_k\right)}{f_{\boldsymbol{y}_0,\ldots,\boldsymbol{y}_k}\left(\boldsymbol{y}_0,\ldots,\boldsymbol{y}_k\right)} \\ &= \frac{f_{\boldsymbol{x}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}}\left(\boldsymbol{x}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}\right) f_{\boldsymbol{y}_k|\boldsymbol{x}_k}\left(\boldsymbol{y}_k|\boldsymbol{x}_k\right)}{f_{\boldsymbol{y}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}}\left(\boldsymbol{y}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}\right)}. \end{aligned} \tag{A155}$$

Thus, using Equations (A153), (A154) and (A155) together with the initial condition given by Equation (A152), the conditional PDFs for the filtering problem can be calculated sequentially.

Equations (A153) and (A154) are often called *prediction* and Equation (A155) *update* (or correction) steps of the solution of the filtering problem. This follows from the fact that Equations (A153) and (A154) describe the distribution of the quantity of interest, $\boldsymbol{x}_k$, or the observation $\boldsymbol{y}_k$, respectively, conditioned only on past observations up to time index $k-1$, i. e., the distribution of quantities that lie in the future w.r.t. the available observations. When a new observation $\boldsymbol{y}_k$ becomes available, this information can be used to update the conditional PDF of $\boldsymbol{x}_k$ given all available observations up to time k (see Equation A155).

Summary A7.1 Let $(\boldsymbol{x}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{y}_k : k \in \mathbb{N}_0)$ be two random sequences with a joint distribution described by Equation (A147). Then, given the conditional PDF

$$f_{\boldsymbol{x}_{k-1}|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}}\left(\boldsymbol{x}_{k-1}|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}\right), \quad k \in \mathbb{N},$$

the corresponding PDF for time index k can be computed sequentially by

$$f_{\boldsymbol{x}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_k}\left(\boldsymbol{x}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_k\right) = \frac{f_{\boldsymbol{x}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}}\left(\boldsymbol{x}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}\right) f_{\boldsymbol{y}_k|\boldsymbol{x}_k}\left(\boldsymbol{y}_k|\boldsymbol{x}_k\right)}{f_{\boldsymbol{y}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}}\left(\boldsymbol{y}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}\right)},$$

where

$$f_{\boldsymbol{x}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}}\left(\boldsymbol{x}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}\right) = \int f_{\boldsymbol{x}_k|\boldsymbol{x}_{k-1}}\left(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}\right) f_{\boldsymbol{x}_{k-1}|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}}\left(\boldsymbol{x}_{k-1}|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}\right)\mathrm{d}\boldsymbol{x}_{k-1}$$

and

$$f_{\boldsymbol{y}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}}\left(\boldsymbol{y}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}\right) = \int f_{\boldsymbol{y}_k|\boldsymbol{x}_k}\left(\boldsymbol{y}_k|\boldsymbol{x}_k\right) f_{\boldsymbol{x}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}}\left(\boldsymbol{x}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}\right)\mathrm{d}\boldsymbol{x}_k.$$

A7.2 Sequential Estimation for Gauss-Markov Models

Although the results of the preceding section are very general and can be applied to any stochastic model which satisfies Equation (A147), it may be difficult or even not possible to evaluate the integrals in Equations (A153) and (A154). On the other hand, the determination of the conditional PDFs for the filtering problem can be tremendously simplified for the case that $(\boldsymbol{x}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{y}_k : k \in \mathbb{N}_0)$ are jointly Gaussian random sequences. For such sequences, it is well known that all subsets of random vectors are jointly [34, p. 386], [37, p. 322] as well as conditionally [31, pp. 116-117] Gaussian distributed and thus the corresponding distributions are completely determined by the mean vector and the covariance matrix. Consequently, for the sequential determination of the conditional PDFs, it suffices to determine the corresponding means and covariances.

For the remainder of this section, the following result will be useful:

Theorem A7.1 Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be two jointly Gaussian distributed random vectors. Then, the mean vector and the covariance matrix of the conditional distribution of $\boldsymbol{x}$ given $\boldsymbol{y}$ are

$$\boldsymbol{\mu}_{\boldsymbol{x}|\boldsymbol{y}} = \boldsymbol{\mu}_{\boldsymbol{x}} + \boldsymbol{C}_{\boldsymbol{x},\boldsymbol{y}} \boldsymbol{C}_{\boldsymbol{y}}^{-1} (\boldsymbol{y} - \boldsymbol{\mu}_{\boldsymbol{y}}) \qquad \text{and} \tag{A156}$$

$$\boldsymbol{C}_{\boldsymbol{x}|\boldsymbol{y}} = \boldsymbol{C}_{\boldsymbol{x}} - \boldsymbol{C}_{\boldsymbol{x},\boldsymbol{y}} \boldsymbol{C}_{\boldsymbol{y}}^{-1} \boldsymbol{C}_{\boldsymbol{x},\boldsymbol{y}}^{\mathrm{T}}, \tag{A157}$$

respectively, where it is assumed that the inverse exists.

Proof. This is a standard result and can be found, e. g., in [31, pp. 116-117] or [70, p. 300]. □

Thus, for the Markov model introduced by Equation (A147), the restriction to Gauss-Markov sequences results in the conditional PDFs

$$f_{\boldsymbol{x}_{k+1}|\boldsymbol{x}_k}(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k) = \mathcal{N}\left(\boldsymbol{\mu}_{\boldsymbol{x}_{k+1}|\boldsymbol{x}_k}, \boldsymbol{C}_{\boldsymbol{x}_{k+1}|\boldsymbol{x}_k}\right), \qquad k \in \mathbb{N}_0, \tag{A158}$$

with (cf. Theorem A7.1)

$$\begin{aligned} \boldsymbol{\mu}_{\boldsymbol{x}_{k+1}|\boldsymbol{x}_k} &= \boldsymbol{\mu}_{\boldsymbol{x}_{k+1}} + \boldsymbol{C}_{\boldsymbol{x}_{k+1},\boldsymbol{x}_k} \boldsymbol{C}_{\boldsymbol{x}_k}^{-1} (\boldsymbol{x}_k - \boldsymbol{\mu}_{\boldsymbol{x}_k}) \qquad \text{and} \\ \boldsymbol{C}_{\boldsymbol{x}_{k+1}|\boldsymbol{x}_k} &= \boldsymbol{C}_{\boldsymbol{x}_{k+1}} - \boldsymbol{C}_{\boldsymbol{x}_{k+1},\boldsymbol{x}_k} \boldsymbol{C}_{\boldsymbol{x}_k}^{-1} \boldsymbol{C}_{\boldsymbol{x}_{k+1},\boldsymbol{x}_k}^{\mathrm{T}}. \end{aligned} \tag{A159}$$

Analogously,

$$f_{\boldsymbol{y}_k|\boldsymbol{x}_k}(\boldsymbol{y}_k|\boldsymbol{x}_k) = \mathcal{N}\left(\boldsymbol{\mu}_{\boldsymbol{y}_k|\boldsymbol{x}_k}, \boldsymbol{C}_{\boldsymbol{y}_k|\boldsymbol{x}_k}\right), \qquad k \in \mathbb{N}_0 \tag{A160}$$

with

$$\begin{aligned} \boldsymbol{\mu}_{\boldsymbol{y}_k|\boldsymbol{x}_k} &= \boldsymbol{\mu}_{\boldsymbol{y}_k} + \boldsymbol{C}_{\boldsymbol{y}_k,\boldsymbol{x}_k} \boldsymbol{C}_{\boldsymbol{x}_k}^{-1} (\boldsymbol{x}_k - \boldsymbol{\mu}_{\boldsymbol{x}_k}) \qquad \text{and} \\ \boldsymbol{C}_{\boldsymbol{y}_k|\boldsymbol{x}_k} &= \boldsymbol{C}_{\boldsymbol{y}_k} - \boldsymbol{C}_{\boldsymbol{y}_k,\boldsymbol{x}_k} \boldsymbol{C}_{\boldsymbol{x}_k}^{-1} \boldsymbol{C}_{\boldsymbol{y}_k,\boldsymbol{x}_k}^{\mathrm{T}}. \end{aligned} \tag{A161}$$

For the derivation of the Kalman filter, it is convenient to use a state space model for the sequences $(\boldsymbol{x}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{y}_k : k \in \mathbb{N}_0)$. This represents no loss of generality since Gauss-Markov sequences with the properties from Equation (A147) can always[28] be constructed with an appropriate linear state space model.[29]

[28] Of course we assume that all inverses which are used for the description of the model parameters exist. A state space model covering the singular case can also be constructed but will not be discussed here.

[29] For the case of two jointly Gaussian distributed random vectors, see, e. g., [70, Section 7.5].

Theorem A7.2 Let $(\boldsymbol{x}_k : k \in \mathbb{N}_0)$ be a Gauss-Markov sequence which, together with $(\boldsymbol{y}_k : k \in \mathbb{N}_0)$, satisfies Equation (A147). Then, the random sequences $(\boldsymbol{x}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{y}_k : k \in \mathbb{N}_0)$ with given conditional PDFs (cf. Equations A158 and A160) can be constructed using the state space model

$$\begin{aligned} \boldsymbol{x}_{k+1} &= \boldsymbol{A}_k \boldsymbol{x}_k + \boldsymbol{w}_k, \\ \boldsymbol{y}_k &= \boldsymbol{C}_k \boldsymbol{x}_k + \boldsymbol{v}_k, \qquad k \in \mathbb{N}_0, \end{aligned} \tag{A162}$$

where $(\boldsymbol{w}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{v}_k : k \in \mathbb{N}_0)$ are mutually independent sequences of independently distributed random vectors with $\boldsymbol{w}_k \sim \mathcal{N}\left(\boldsymbol{\mu}_{\boldsymbol{w}_k}, \boldsymbol{C}_{\boldsymbol{w}_k}\right)$ and $\boldsymbol{v}_k \sim \mathcal{N}\left(\boldsymbol{\mu}_{\boldsymbol{v}_k}, \boldsymbol{C}_{\boldsymbol{v}_k}\right)$, $k \in \mathbb{N}_0$. Additionally, the initial value $\boldsymbol{x}_0 \sim \mathcal{N}\left(\boldsymbol{\mu}_{\boldsymbol{x}_0}, \boldsymbol{C}_{\boldsymbol{x}_0}\right)$ is independent of $(\boldsymbol{w}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{v}_k : k \in \mathbb{N}_0)$. The parameters of the state space model are

$$\boldsymbol{A}_k = \boldsymbol{C}_{\boldsymbol{x}_{k+1},\boldsymbol{x}_k} \boldsymbol{C}_{\boldsymbol{x}_k}^{-1}, \qquad \boldsymbol{C}_k = \boldsymbol{C}_{\boldsymbol{y}_k,\boldsymbol{x}_k} \boldsymbol{C}_{\boldsymbol{x}_k}^{-1}, \tag{A163}$$

$$\boldsymbol{\mu}_{\boldsymbol{w}_k} = \boldsymbol{\mu}_{\boldsymbol{x}_{k+1}} - \boldsymbol{C}_{\boldsymbol{x}_{k+1},\boldsymbol{x}_k} \boldsymbol{C}_{\boldsymbol{x}_k}^{-1} \boldsymbol{\mu}_{\boldsymbol{x}_k}, \qquad \boldsymbol{\mu}_{\boldsymbol{v}_k} = \boldsymbol{\mu}_{\boldsymbol{y}_k} - \boldsymbol{C}_{\boldsymbol{y}_k,\boldsymbol{x}_k} \boldsymbol{C}_{\boldsymbol{x}_k}^{-1} \boldsymbol{\mu}_{\boldsymbol{x}_k}, \tag{A164}$$

$$\boldsymbol{C}_{\boldsymbol{w}_k} = \boldsymbol{C}_{\boldsymbol{x}_{k+1}|\boldsymbol{x}_k}, \qquad \boldsymbol{C}_{\boldsymbol{v}_k} = \boldsymbol{C}_{\boldsymbol{y}_k|\boldsymbol{x}_k}, \tag{A165}$$

which are completely determined by the Gaussian transition densities $f_{\boldsymbol{x}_{k+1}|\boldsymbol{x}_k}(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k)$ and $f_{\boldsymbol{y}_k|\boldsymbol{x}_k}(\boldsymbol{y}_k|\boldsymbol{x}_k)$, $k \in \mathbb{N}_0$, and $f_{\boldsymbol{x}_0}(\boldsymbol{x}_0)$.

Proof. The stochastic description of the Gauss-Markov sequence requires the knowledge of the transition densities $f_{\boldsymbol{x}_{k+1}|\boldsymbol{x}_k}(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k)$ and $f_{\boldsymbol{y}_k|\boldsymbol{x}_k}(\boldsymbol{y}_k|\boldsymbol{x}_k)$, $k \in \mathbb{N}_0$, as well as the distribution of the initial random vector $f_{\boldsymbol{x}_0}(\boldsymbol{x}_0)$. Since all random vectors are jointly Gaussian distributed, it can be seen with the results of Theorem A7.1 that this knowledge is equivalent to the knowledge of a sequence of affine functions which describe the conditional expected values of the distributions and a sequence of constant covariance matrices which describe the conditional covariances (cf. Equations A159 and A161). These sequences can be constructed using the state space model in Equation (A162). First, consider the conditional expected value of $\boldsymbol{x}_{k+1}$ given $\boldsymbol{x}_k$,

$$\begin{aligned} \boldsymbol{\mu}_{\boldsymbol{x}_{k+1}|\boldsymbol{x}_k} &= \mathrm{E}_{\boldsymbol{x}_{k+1}|\boldsymbol{x}_k}\left[\boldsymbol{x}_{k+1} | \, \boldsymbol{x}_k\right] \\ &= \mathrm{E}_{\boldsymbol{x}_k,\boldsymbol{w}_k|\boldsymbol{x}_k}\left[\boldsymbol{A}_k \boldsymbol{x}_k + \boldsymbol{w}_k | \, \boldsymbol{x}_k\right] = \boldsymbol{A}_k \boldsymbol{x}_k + \boldsymbol{\mu}_{\boldsymbol{w}_k}, \qquad k \in \mathbb{N}_0, \end{aligned} \tag{A166}$$

where the fact has been used that $(\boldsymbol{w}_k : k \in \mathbb{N}_0)$ is a sequence of independently distributed random vectors and thus $\boldsymbol{w}_k$ is independent of $\boldsymbol{x}_k$. Comparing this function which is affine in $\boldsymbol{x}_k$ with Equation (A159), we find that

$$\boldsymbol{A}_k = \boldsymbol{C}_{\boldsymbol{x}_{k+1},\boldsymbol{x}_k} \boldsymbol{C}_{\boldsymbol{x}_k}^{-1}, \qquad k \in \mathbb{N}_0, \tag{A167}$$

and

$$\begin{aligned} \boldsymbol{\mu}_{\boldsymbol{w}_k} &= \boldsymbol{\mu}_{\boldsymbol{x}_{k+1}|\boldsymbol{x}_k} - \boldsymbol{A}_k \boldsymbol{x}_k \\ &= \boldsymbol{\mu}_{\boldsymbol{x}_{k+1}} - \boldsymbol{C}_{\boldsymbol{x}_{k+1},\boldsymbol{x}_k} \boldsymbol{C}_{\boldsymbol{x}_k}^{-1} \boldsymbol{\mu}_{\boldsymbol{x}_k}, \qquad k \in \mathbb{N}_0. \end{aligned} \tag{A168}$$

With the independence of $\boldsymbol{x}_k$ and $\boldsymbol{w}_k$, the state space model generates the sequence of covariance matrices

$$\boldsymbol{C}_{\boldsymbol{x}_{k+1}} = \boldsymbol{A}_k \boldsymbol{C}_{\boldsymbol{x}_k} \boldsymbol{A}_k^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{w}_k}, \qquad k \in \mathbb{N}_0, \tag{A169}$$

with initial value $\boldsymbol{C}_{\boldsymbol{x}_0}$ as well as the sequence of cross-covariance matrices

$$\boldsymbol{C}_{\boldsymbol{x}_{k+1},\boldsymbol{x}_k} = \boldsymbol{A}_k \boldsymbol{C}_{\boldsymbol{x}_k}, \qquad k \in \mathbb{N}_0. \tag{A170}$$

This result provides the correct choice of the covariance matrix of $\boldsymbol{w}_k$ which is (cf. Equation A159)

$$\begin{aligned} \boldsymbol{C}_{\boldsymbol{x}_{k+1}|\boldsymbol{x}_k} &= \boldsymbol{A}_k \boldsymbol{C}_{\boldsymbol{x}_k} \boldsymbol{A}_k^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{w}_k} - \boldsymbol{A}_k \boldsymbol{C}_{\boldsymbol{x}_k} \boldsymbol{C}_{\boldsymbol{x}_k}^{-1} \boldsymbol{C}_{\boldsymbol{x}_k} \boldsymbol{A}_k^{\mathrm{T}} \\ &= \boldsymbol{C}_{\boldsymbol{w}_k}, \end{aligned} \tag{A171}$$

for $k \in \mathbb{N}_0$. The presented steps can be repeated for the sequence of observations $(\boldsymbol{y}_k : k \in \mathbb{N}_0)$ to obtain the results for the parameters $\boldsymbol{C}_k$, $\boldsymbol{\mu}_{\boldsymbol{v}_k}$ and $\boldsymbol{C}_{\boldsymbol{v}_k}$, $k \in \mathbb{N}_0$. Finally, due to the independence assumptions of the noise sequences $(\boldsymbol{w}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{v}_k : k \in \mathbb{N}_0)$, it is easy to verify that the conditional independence of the state and the observation sequences which lead to Equation (A147) holds for the linear state space model. □

In the following, we will use the state space model of Equation (A162) to derive the Kalman filter. At this point, we change the introduced notation to one which is more common in the existing literature on optimal estimation. As already mentioned, the conditional PDF of $\boldsymbol{x}_k$ given the available observations, i. e., $f_{\boldsymbol{x}_k|\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_k}(\boldsymbol{x}_k|\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_k)$, is fully described by (cf. Theorem A7.1) the conditional mean

$$\hat{\boldsymbol{x}}_k = \mathrm{E}_{\boldsymbol{x}_k|\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_k}\left[\boldsymbol{x}_k|\,\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_k\right] = \boldsymbol{\mu}_{\boldsymbol{x}_k|\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_k} \tag{A172}$$

and the conditional covariance matrix

$$\boldsymbol{C}_{\tilde{\boldsymbol{x}}_k} = \mathrm{E}_{\boldsymbol{x}_k|\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_k}\left[\left(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k\right)\left(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k\right)^{\mathrm{T}}\middle|\,\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_k\right] = \boldsymbol{C}_{\boldsymbol{x}_k|\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_k} \tag{A173}$$

for all $k \in \mathbb{N}_0$. Note that we introduce the vector

$$\tilde{\boldsymbol{x}}_k = \boldsymbol{x}_k - \hat{\boldsymbol{x}}_k, \qquad k \in \mathbb{N}_0, \tag{A174}$$

which represents the estimation error between the quantity of interest, $\boldsymbol{x}_k$, and its conditional mean estimate, $\hat{\boldsymbol{x}}_k = \mathrm{E}_{\boldsymbol{x}_k|\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_k}\left[\boldsymbol{x}_k|\,\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_k\right]$. As we have already seen in Section A7.1 (cf. Equation A153), the procedure of sequential estimation includes a *prediction* step, represented by the determination of the PDF $f_{\boldsymbol{x}_k|\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{k-1}}(\boldsymbol{x}_k|\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{k-1})$. The parameters of this density are given by

$$\hat{\boldsymbol{x}}_k^{\mathrm{P}} = \mathrm{E}_{\boldsymbol{x}_k|\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{k-1}}\left[\boldsymbol{x}_k|\,\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{k-1}\right] = \boldsymbol{\mu}_{\boldsymbol{x}_k|\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{k-1}} \tag{A175}$$

and

$$\boldsymbol{C}_{\tilde{\boldsymbol{x}}_k}^{\mathrm{P}} = \mathrm{E}_{\boldsymbol{x}_k|\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{k-1}}\left[\left(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k^{\mathrm{P}}\right)\left(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k^{\mathrm{P}}\right)^{\mathrm{T}}\middle|\,\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{k-1}\right] = \boldsymbol{C}_{\boldsymbol{x}_k|\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{k-1}} \tag{A176}$$

for $k \in \mathbb{N}$, where

$$\tilde{\boldsymbol{x}}_k^{\mathrm{P}} = \boldsymbol{x}_k - \hat{\boldsymbol{x}}_k^{\mathrm{P}}, \qquad k \in \mathbb{N}_0. \tag{A177}$$

The superscript P indicates the prediction of $\boldsymbol{x}_k$ at time index k given all observations up to time index $k-1$. Note that for $k = 0$, we have $\tilde{\boldsymbol{x}}_0^{\mathrm{P}} = \boldsymbol{\mu}_{\boldsymbol{x}_0}$ and consequently $\boldsymbol{C}_{\tilde{\boldsymbol{x}}_0}^{\mathrm{P}} = \boldsymbol{C}_{\boldsymbol{x}_0}$.

An important property of the conditional mean $\hat{\boldsymbol{x}}_k$ and the deviation $\tilde{\boldsymbol{x}}_k$ of the random vector $\boldsymbol{x}_k$ from its conditional mean is that these random vectors are uncorrelated [70, pp. 300-301], i. e.,[30]

$$\mathrm{E}_{\tilde{\boldsymbol{x}}_k,\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_k}\left[(\hat{\boldsymbol{x}}_k-\boldsymbol{\mu}_{\boldsymbol{x}_k})\tilde{\boldsymbol{x}}_k^{\mathrm{T}}\right]=\mathbf{0}_{N_x\times N_x},\qquad k\in\mathbb{N}_0. \tag{A178}$$

Additionally, $\tilde{\boldsymbol{x}}_k$ is uncorrelated with the random vectors it is conditioned on, i. e.,

$$\mathrm{E}_{\tilde{\boldsymbol{x}}_k,\boldsymbol{y}_i}\left[(\boldsymbol{y}_i-\boldsymbol{\mu}_{\boldsymbol{y}_i})\,\tilde{\boldsymbol{x}}_k^{\mathrm{T}}\right]=\mathbf{0}_{N_y\times N_x},\qquad i\in\{0,1,\ldots,k\}. \tag{A179}$$

These results hold analogously for the random vectors $\hat{\boldsymbol{x}}_k^{\mathrm{P}}$ and $\tilde{\boldsymbol{x}}_k^{\mathrm{P}}$.

Now we are in the position to derive the well known Kalman filter equations. To this end, we use the state space model introduced in Equation (A162), i. e.,

$$\begin{aligned}\boldsymbol{x}_{k+1}&=\boldsymbol{A}_k\boldsymbol{x}_k+\boldsymbol{w}_k\\ \boldsymbol{y}_k&=\boldsymbol{C}_k\boldsymbol{x}_k+\boldsymbol{v}_k,\qquad k\in\mathbb{N}_0,\end{aligned} \tag{A180}$$

where the initial state $\boldsymbol{x}_0$ and the random sequences $(\boldsymbol{w}_k:k\in\mathbb{N}_0)$ and $(\boldsymbol{v}_k:k\in\mathbb{N}_0)$ are independently (not necessary identically) jointly Gaussian distributed and mutually independent. Analogously to the prediction and correction steps shown in Section A7.1, the parameters of the corresponding PDFs, i. e., the conditional mean and covariance, will be determined sequentially in the following. Assume that these parameters $\hat{\boldsymbol{x}}_{k-1}$ and $\boldsymbol{C}_{\tilde{\boldsymbol{x}}_{k-1}}$ are known for some $(k-1)\in\mathbb{N}_0$. Then, for the prediction step, we get[31]

$$\begin{aligned}\hat{\boldsymbol{x}}_k^{\mathrm{P}}&=\mathrm{E}_{\boldsymbol{x}_k|\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{k-1}}\left[\boldsymbol{x}_k|\,\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{k-1}\right]\\ &=\mathrm{E}_{\boldsymbol{x}_{k-1},\boldsymbol{w}_{k-1}|\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{k-1}}\left[\boldsymbol{A}_{k-1}\boldsymbol{x}_{k-1}+\boldsymbol{w}_{k-1}|\,\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{k-1}\right]\\ &\overset{(*)}{=}\boldsymbol{A}_{k-1}\,\mathrm{E}_{\boldsymbol{x}_{k-1}|\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{k-1}}\left[\boldsymbol{x}_{k-1}|\,\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{k-1}\right]+\mathrm{E}_{\boldsymbol{w}_{k-1}}\left[\boldsymbol{w}_{k-1}\right]\\ &=\boldsymbol{A}_{k-1}\hat{\boldsymbol{x}}_{k-1}+\boldsymbol{\mu}_{\boldsymbol{w}_{k-1}}.\end{aligned} \tag{A181}$$

Equation $(*)$ holds because, due to the model assumptions, $\boldsymbol{w}_{k-1}$ is independent of the observations $\boldsymbol{y}_i$, $i\in\{0,1,\ldots,k-1\}$. For the corresponding covariance matrix we get

$$\begin{aligned}\boldsymbol{C}_{\tilde{\boldsymbol{x}}_k}^{\mathrm{P}}&=\mathrm{E}_{\boldsymbol{x}_k|\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{k-1}}\left[\left(\boldsymbol{x}_k-\hat{\boldsymbol{x}}_k^{\mathrm{P}}\right)\left(\boldsymbol{x}_k-\hat{\boldsymbol{x}}_k^{\mathrm{P}}\right)^{\mathrm{T}}\middle|\,\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{k-1}\right]\\ &=\mathrm{E}_{\boldsymbol{x}_{k-1},\boldsymbol{w}_{k-1}|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}}\Big[\left(\boldsymbol{A}_{k-1}\left(\boldsymbol{x}_{k-1}-\hat{\boldsymbol{x}}_{k-1}\right)+\left(\boldsymbol{w}_{k-1}-\boldsymbol{\mu}_{\boldsymbol{w}_{k-1}}\right)\right)\\ &\qquad\times\left(\boldsymbol{A}_{k-1}\left(\boldsymbol{x}_{k-1}-\hat{\boldsymbol{x}}_{k-1}\right)+\left(\boldsymbol{w}_{k-1}-\boldsymbol{\mu}_{\boldsymbol{w}_{k-1}}\right)\right)^{\mathrm{T}}\Big|\,\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}\Big]\\ &=\boldsymbol{A}_{k-1}\,\mathrm{E}_{\boldsymbol{x}_{k-1}|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}}\left[\left(\boldsymbol{x}_{k-1}-\hat{\boldsymbol{x}}_{k-1}\right)\left(\boldsymbol{x}_{k-1}-\hat{\boldsymbol{x}}_{k-1}\right)^{\mathrm{T}}\middle|\,\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}\right]\boldsymbol{A}_{k-1}^{\mathrm{T}}\\ &\qquad+\mathrm{E}_{\boldsymbol{w}_{k-1}}\left[\left(\boldsymbol{w}_{k-1}-\boldsymbol{\mu}_{\boldsymbol{w}_{k-1}}\right)\left(\boldsymbol{w}_{k-1}-\boldsymbol{\mu}_{\boldsymbol{w}_{k-1}}\right)^{\mathrm{T}}\right]\\ &=\boldsymbol{A}_{k-1}\boldsymbol{C}_{\tilde{\boldsymbol{x}}_{k-1}}\boldsymbol{A}_{k-1}^{\mathrm{T}}+\boldsymbol{C}_{\boldsymbol{w}_{k-1}}.\end{aligned} \tag{A182}$$

[30]The subtraction of $\boldsymbol{\mu}_{\boldsymbol{x}_k}$ can be omitted without changing the result because $\tilde{\boldsymbol{x}}_k$ is a zero mean random vector.

[31]Equation (A181) shows how to deal with linear dynamical systems which are additionally driven by a known input sequence, e. g., a controlled system. In this case, the input sequence can be interpreted as a different value of the mean vector $\boldsymbol{\mu}_{\boldsymbol{w}_{k-1}}$ of the process noise and thus, for the case of linear systems, can simply be added to Equation (A181). This is of importance for the LQG control approach, see Section A6.

The derivation of the correction step needs more work but offers a different view on the representation of random sequences. For a compact notation, define the vector $\boldsymbol{\gamma}_k$ which collects all observations up to time index k, i. e.,

$$\boldsymbol{\gamma}_k = \begin{bmatrix} \boldsymbol{y}_0 \\ \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_k \end{bmatrix}, \qquad k \in \mathbb{N}_0. \tag{A183}$$

Since the conditional distribution of $\boldsymbol{x}_k$ given $\boldsymbol{\gamma}_k$ is Gaussian, its mean value for $k \in \mathbb{N}$ is given by (cf. Theorem A7.1)

$$\begin{aligned} \hat{\boldsymbol{x}}_k &= \mathrm{E}_{\boldsymbol{x}_k|\boldsymbol{\gamma}_k}\left[\boldsymbol{x}_k \middle| \boldsymbol{\gamma}_k\right] = \boldsymbol{\mu}_{\boldsymbol{x}_k} + \boldsymbol{C}_{\boldsymbol{x}_k,\boldsymbol{\gamma}_k}\boldsymbol{C}_{\boldsymbol{\gamma}_k}^{-1}\left(\boldsymbol{\gamma}_k - \boldsymbol{\mu}_{\boldsymbol{\gamma}_k}\right) \\ &= \boldsymbol{\mu}_{\boldsymbol{x}_k} + \begin{bmatrix} \boldsymbol{C}_{\boldsymbol{x}_k,\boldsymbol{\gamma}_{k-1}} \boldsymbol{C}_{\boldsymbol{x}_k,\boldsymbol{y}_k} \end{bmatrix} \begin{bmatrix} \boldsymbol{C}_{\boldsymbol{\gamma}_{k-1}} & \boldsymbol{C}_{\boldsymbol{\gamma}_{k-1},\boldsymbol{y}_k} \\ \boldsymbol{C}_{\boldsymbol{\gamma}_{k-1},\boldsymbol{y}_k}^{\mathrm{T}} & \boldsymbol{C}_{\boldsymbol{y}_k} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{\gamma}_{k-1} - \boldsymbol{\mu}_{\boldsymbol{\gamma}_{k-1}} \\ \boldsymbol{y}_k - \boldsymbol{\mu}_{\boldsymbol{y}_k} \end{bmatrix}. \end{aligned} \tag{A184}$$

For a sequential computation of $\hat{\boldsymbol{x}}_k$ it would be comfortable to have uncorrelated observations, i. e., that $\boldsymbol{C}_{\boldsymbol{\gamma}_{k-1},\boldsymbol{y}_k} = \mathbf{0}_{kN_y \times N_y}$, $k \in \mathbb{N}$. In this case, the inverse in Equation (A184) would be block diagonal and the conditional mean could be written as a sum (cf. [6, p. 490]) where one summand is already known (cf. Equation A181):

$$\hat{\boldsymbol{x}}_k^{\mathrm{P}} = \mathrm{E}_{\boldsymbol{x}_k|\boldsymbol{\gamma}_{k-1}}\left[\boldsymbol{x}_k \middle| \boldsymbol{\gamma}_{k-1}\right] = \boldsymbol{\mu}_{\boldsymbol{x}_k} + \boldsymbol{C}_{\boldsymbol{x}_k,\boldsymbol{\gamma}_{k-1}}\boldsymbol{C}_{\boldsymbol{\gamma}_{k-1}}^{-1}\left(\boldsymbol{\gamma}_{k-1} - \boldsymbol{\mu}_{\boldsymbol{\gamma}_{k-1}}\right). \tag{A185}$$

Of course the desired cross covariance matrix is not zero because the observation sequence $(\boldsymbol{y}_k : k \in \mathbb{N}_0)$ is in general not a sequence of uncorrelated random vectors. But we have the degree of freedom to apply an invertible affine function to the collection of observations $\boldsymbol{\gamma}_k$ which does not change the mean vector and covariance matrix of the conditional distribution of $\boldsymbol{x}_k$ given the (transformed) observations. Define

$$\boldsymbol{\eta}_k = \boldsymbol{T}_k\boldsymbol{\gamma}_k + \boldsymbol{s}_k, \qquad k \in \mathbb{N}_0, \tag{A186}$$

where $\boldsymbol{T}_k^{-1}$ exists and $\boldsymbol{s}_k$ is a constant vector of appropriate dimension. It is easy to verify that

$$\begin{aligned} \mathrm{E}_{\boldsymbol{x}_k|\boldsymbol{\eta}_k}\left[\boldsymbol{x}_k \middle| \boldsymbol{\eta}_k\right] &= \boldsymbol{\mu}_{\boldsymbol{x}_k} + \boldsymbol{C}_{\boldsymbol{x}_k,\boldsymbol{\eta}_k}\boldsymbol{C}_{\boldsymbol{\eta}_k}^{-1}\left(\boldsymbol{\eta}_k - \boldsymbol{\mu}_{\boldsymbol{\eta}_k}\right) \\ &= \boldsymbol{\mu}_{\boldsymbol{x}_k} + \boldsymbol{C}_{\boldsymbol{x}_k,\boldsymbol{\gamma}_k}\boldsymbol{T}_k^{\mathrm{T}}\boldsymbol{T}_k^{-1,\mathrm{T}}\boldsymbol{C}_{\boldsymbol{\gamma}_k}^{-1}\boldsymbol{T}_k^{-1}\left(\boldsymbol{T}_k\boldsymbol{\gamma}_k + \boldsymbol{s}_k - \boldsymbol{T}_k\boldsymbol{\mu}_{\boldsymbol{\gamma}_k} - \boldsymbol{s}_k\right) \\ &= \mathrm{E}_{\boldsymbol{x}_k|\boldsymbol{\gamma}_k}\left[\boldsymbol{x}_k \middle| \boldsymbol{\gamma}_k\right]. \end{aligned} \tag{A187}$$

Analogously, the equality of the conditional covariance matrices can be shown. For the Kalman filter, we use an affine function with

$$\boldsymbol{T}_k = \begin{bmatrix} \mathbf{I}_{kN_y} & \mathbf{0}_{kN_y \times N_y} \\ -\boldsymbol{C}_{\boldsymbol{y}_k,\boldsymbol{\gamma}_{k-1}}\boldsymbol{C}_{\boldsymbol{\gamma}_{k-1}}^{-1} & \mathbf{I}_{N_y} \end{bmatrix} \tag{A188}$$

and

$$\boldsymbol{s}_k = \begin{bmatrix} \mathbf{0}_{kN_y} \\ -\boldsymbol{\mu}_{\boldsymbol{y}_k} + \boldsymbol{C}_{\boldsymbol{y}_k,\boldsymbol{\gamma}_{k-1}}\boldsymbol{C}_{\boldsymbol{\gamma}_{k-1}}^{-1}\boldsymbol{\mu}_{\boldsymbol{\gamma}_{k-1}} \end{bmatrix} \tag{A189}$$

for $k \in \mathbb{N}$, while for $k = 0$ the mapping with $\boldsymbol{T}_0 = \mathbf{I}_{N_y}$ and $\boldsymbol{s}_0 = -\boldsymbol{\mu}_{\boldsymbol{y}_0}$ is applied. The matrix $\boldsymbol{T}_k$ is clearly invertible since it is lower triangular with ones on the main diagonal. Additionally, note that $\boldsymbol{T}_k$ and $\boldsymbol{s}_k$ only depend on mean vectors and covariance matrices up to time index k. Having a closer look on $\boldsymbol{T}_k$ and $\boldsymbol{s}_k$, it can be seen that the vector $\boldsymbol{\eta}_k$ of the transformed observations reads as

$$\boldsymbol{\eta}_k = \begin{bmatrix} \gamma_{k-1} \\ \boldsymbol{e}_k \end{bmatrix}, \qquad k \in \mathbb{N}, \tag{A190}$$

and $\boldsymbol{\eta}_0 = \boldsymbol{e}_0$, where we introduced the vector $\boldsymbol{e}_k$, $k \in \mathbb{N}_0$. Using Equations (A183), (A186), (A188) and (A189), we identify

$$\begin{aligned} \boldsymbol{e}_k &= \boldsymbol{y}_k - \boldsymbol{\mu}_{\boldsymbol{y}_k} - \boldsymbol{C}_{\boldsymbol{y}_k,\gamma_{k-1}} \boldsymbol{C}_{\gamma_{k-1}}^{-1} \left(\gamma_{k-1} - \boldsymbol{\mu}_{\gamma_{k-1}}\right) \\ &= \boldsymbol{y}_k - \mathrm{E}_{\boldsymbol{y}_k|\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{k-1}} \left[\boldsymbol{y}_k | \, \boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{k-1}\right], \qquad k \in \mathbb{N}, \end{aligned} \tag{A191}$$

and $\boldsymbol{e}_0 = \boldsymbol{y}_0 - \boldsymbol{\mu}_{\boldsymbol{y}_0}$. It can be seen that $\boldsymbol{e}_k$ is the deviation of the observation $\boldsymbol{y}_k$ from its conditional mean given the past observations $\boldsymbol{y}_i$, $i \in \{0, 1, \ldots, k-1\}$, and thus it is uncorrelated with these observations (cf. Equation A179). Consequently, using the transformed observations $\boldsymbol{\eta}_k$, we get with Equations (A184) and (A187) the conditional mean

$$\begin{aligned} \hat{\boldsymbol{x}}_k &= \boldsymbol{\mu}_{\boldsymbol{x}_k} + \left[\boldsymbol{C}_{\boldsymbol{x}_k,\gamma_{k-1}} \boldsymbol{C}_{\boldsymbol{x}_k,\boldsymbol{e}_k}\right] \begin{bmatrix} \boldsymbol{C}_{\gamma_{k-1}} & \boldsymbol{0}_{kN_y \times N_y} \\ \boldsymbol{0}_{N_y \times kN_y} & \boldsymbol{C}_{\boldsymbol{e}_k} \end{bmatrix}^{-1} \begin{bmatrix} \gamma_{k-1} - \boldsymbol{\mu}_{\gamma_{k-1}} \\ \boldsymbol{e}_k - \boldsymbol{\mu}_{\boldsymbol{e}_k} \end{bmatrix} \\ &= \boldsymbol{\mu}_{\boldsymbol{x}_k} + \boldsymbol{C}_{\boldsymbol{x}_k,\gamma_{k-1}} \boldsymbol{C}_{\gamma_{k-1}}^{-1} \left(\gamma_{k-1} - \boldsymbol{\mu}_{\gamma_{k-1}}\right) + \boldsymbol{C}_{\boldsymbol{x}_k,\boldsymbol{e}_k} \boldsymbol{C}_{\boldsymbol{e}_k}^{-1} \left(\boldsymbol{e}_k - \boldsymbol{\mu}_{\boldsymbol{e}_k}\right) \\ &= \hat{\boldsymbol{x}}_k^{\mathrm{P}} + \boldsymbol{C}_{\boldsymbol{x}_k,\boldsymbol{e}_k} \boldsymbol{C}_{\boldsymbol{e}_k}^{-1} \boldsymbol{e}_k \end{aligned} \tag{A192}$$

for $k \in \mathbb{N}_0$. The last step is valid due to Equation (A185) and the fact that $\boldsymbol{e}_k$ has zero mean.

It remains to determine the covariance matrices $\boldsymbol{C}_{\boldsymbol{x}_k,\boldsymbol{e}_k}$ and $\boldsymbol{C}_{\boldsymbol{e}_k}$ for $k \in \mathbb{N}_0$. We start with the latter one and get by simple calculations[32]

$$\begin{aligned} \boldsymbol{C}_{\boldsymbol{e}_k} &= \mathrm{E}_{\gamma_k} \left[(\boldsymbol{y}_k - \mathrm{E}_{\boldsymbol{y}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}} \left[\boldsymbol{y}_k | \, \boldsymbol{y}_0, \ldots, \boldsymbol{y}_{k-1}\right])(\boldsymbol{y}_k - \mathrm{E}_{\boldsymbol{y}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{k-1}} \left[\boldsymbol{y}_k | \, \boldsymbol{y}_0, \ldots, \boldsymbol{y}_{k-1}\right])^{\mathrm{T}}\right] \\ &= \mathrm{E}_{\boldsymbol{x}_k,\boldsymbol{v}_k,\gamma_{k-1}} \Big[\left(\boldsymbol{C}_k \left(\boldsymbol{x}_k - \mathrm{E}_{\boldsymbol{x}_k|\gamma_{k-1}} \left[\boldsymbol{x}_k | \, \gamma_{k-1}\right]\right) + \left(\boldsymbol{v}_k - \boldsymbol{\mu}_{\boldsymbol{v}_k}\right)\right) \\ &\qquad\qquad \times \left(\boldsymbol{C}_k \left(\boldsymbol{x}_k - \mathrm{E}_{\boldsymbol{x}_k|\gamma_{k-1}} \left[\boldsymbol{x}_k | \, \gamma_{k-1}\right]\right) + \left(\boldsymbol{v}_k - \boldsymbol{\mu}_{\boldsymbol{v}_k}\right)\right)^{\mathrm{T}} \Big] \\ &= \mathrm{E}_{\boldsymbol{x}_k,\boldsymbol{v}_k,\gamma_{k-1}} \left[\left(\boldsymbol{C}_k \left(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k^{\mathrm{P}}\right) + \left(\boldsymbol{v}_k - \boldsymbol{\mu}_{\boldsymbol{v}_k}\right)\right) \left(\boldsymbol{C}_k \left(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k^{\mathrm{P}}\right) + \left(\boldsymbol{v}_k - \boldsymbol{\mu}_{\boldsymbol{v}_k}\right)\right)^{\mathrm{T}}\right] \\ &= \boldsymbol{C}_k \boldsymbol{C}_{\tilde{\boldsymbol{x}}_k}^{\mathrm{P}} \boldsymbol{C}_k^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}_k}, \end{aligned} \tag{A193}$$

where we used Equations (A162), (A176), (A183) and the fact that $\boldsymbol{v}_k$ is independent of $\boldsymbol{x}_k$ and γ_{k-1}. For the cross covariance matrix, together with Equation (A177) and the fact that $\hat{\boldsymbol{x}}_k^{\mathrm{P}}$ and $\tilde{\boldsymbol{x}}_k^{\mathrm{P}}$ are uncorrelated (cf. Equation A178), we get

$$\begin{aligned} \boldsymbol{C}_{\boldsymbol{x}_k,\boldsymbol{e}_k} &= \mathrm{E}_{\boldsymbol{x}_k,\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_k} \left[\boldsymbol{x}_k (\boldsymbol{y}_k - \mathrm{E}_{\boldsymbol{y}_k|\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{k-1}} \left[\boldsymbol{y}_k | \, \boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{k-1}\right])^{\mathrm{T}}\right] \\ &= \mathrm{E}_{\boldsymbol{x}_k,\boldsymbol{v}_k,\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{k-1}} \left[\left(\hat{\boldsymbol{x}}_k^{\mathrm{P}} + \tilde{\boldsymbol{x}}_k^{\mathrm{P}}\right) \left(\boldsymbol{C}_k \left(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k^{\mathrm{P}}\right) + \left(\boldsymbol{v}_k - \boldsymbol{\mu}_{\boldsymbol{v}_k}\right)\right)^{\mathrm{T}}\right] \\ &= \boldsymbol{C}_{\tilde{\boldsymbol{x}}_k}^{\mathrm{P}} \boldsymbol{C}_k^{\mathrm{T}}. \end{aligned} \tag{A194}$$

Finally, inserting Equation (A194), Equation (A193) and Equation (A191) in Equation (A192) and using the identity

$$\begin{aligned} \mathrm{E}_{\boldsymbol{y}_k|\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{k-1}} \left[\boldsymbol{y}_k | \, \boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{k-1}\right] &= \mathrm{E}_{\boldsymbol{x}_k,\boldsymbol{v}_k|\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{k-1}} \left[\boldsymbol{C}_k \boldsymbol{x}_k + \boldsymbol{v}_k | \, \boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{k-1}\right] \\ &= \boldsymbol{C}_k \hat{\boldsymbol{x}}_k^{\mathrm{P}} + \boldsymbol{\mu}_{\boldsymbol{v}_k}, \end{aligned} \tag{A195}$$

[32]For $k = 0$, we have $\boldsymbol{e}_0 = \boldsymbol{y}_0 - \boldsymbol{\mu}_{\boldsymbol{y}_0}$ as well as $\boldsymbol{C}_{\tilde{\boldsymbol{x}}_0}^{\mathrm{P}} = \boldsymbol{C}_{\boldsymbol{x}_0}$, which directly provides the result of Equations (A193) and (A194) using the state space model (cf. Equation A180).

we get the desired result

$$\hat{\boldsymbol{x}}_k = \hat{\boldsymbol{x}}_k^{\mathrm{P}} + \boldsymbol{C}_{\tilde{\boldsymbol{x}}_k}^{\mathrm{P}} \boldsymbol{C}_k^{\mathrm{T}} \left(\boldsymbol{C}_k \boldsymbol{C}_{\tilde{\boldsymbol{x}}_k}^{\mathrm{P}} \boldsymbol{C}_k^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}_k} \right)^{-1} \left(\boldsymbol{y}_k - \boldsymbol{C}_k \hat{\boldsymbol{x}}_k^{\mathrm{P}} - \boldsymbol{\mu}_{\boldsymbol{v}_k} \right), \tag{A196}$$

which can be computed from the known quantities $\hat{\boldsymbol{x}}_{k-1}$ and $\boldsymbol{C}_{\tilde{\boldsymbol{x}}_{k-1}}$ for $k \in \mathbb{N}$.

Finally, we determine the covariance matrix $\boldsymbol{C}_{\tilde{\boldsymbol{x}}_k}$ of the distribution in the correction step. To this end, note that (cf. Equations A191, A195 and A162),

$$\begin{aligned} \boldsymbol{e}_k &= \boldsymbol{y}_k - \boldsymbol{C}_k \hat{\boldsymbol{x}}_k^{\mathrm{P}} - \boldsymbol{\mu}_{\boldsymbol{v}_k} \\ &= \boldsymbol{C}_k \tilde{\boldsymbol{x}}_k^{\mathrm{P}} + (\boldsymbol{v}_k - \boldsymbol{\mu}_{\boldsymbol{v}_k}). \end{aligned} \tag{A197}$$

Keeping that in mind and using Equations (A192), (A193) and (A194), the desired covariance matrix reads as[33]

$$\begin{aligned} \boldsymbol{C}_{\tilde{\boldsymbol{x}}_k} &= \mathrm{E}_{\boldsymbol{x}_k, \boldsymbol{y}_0, \ldots, \boldsymbol{y}_k} \left[(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k)(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k)^{\mathrm{T}} \right] \\ &= \mathrm{E}_{\boldsymbol{x}_k, \boldsymbol{y}_0, \ldots, \boldsymbol{y}_k} \left[\left(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k^{\mathrm{P}} - \boldsymbol{C}_{\boldsymbol{x}_k, \boldsymbol{e}_k} \boldsymbol{C}_{\boldsymbol{e}_k}^{-1} \boldsymbol{e}_k \right) \left(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k^{\mathrm{P}} - \boldsymbol{C}_{\boldsymbol{x}_k, \boldsymbol{e}_k} \boldsymbol{C}_{\boldsymbol{e}_k}^{-1} \boldsymbol{e}_k \right)^{\mathrm{T}} \right] \\ &= \mathrm{E}_{\boldsymbol{x}_k, \boldsymbol{y}_0, \ldots, \boldsymbol{y}_k} \left[\tilde{\boldsymbol{x}}_k^{\mathrm{P}} \tilde{\boldsymbol{x}}_k^{\mathrm{P,T}} \right] - \mathrm{E}_{\boldsymbol{x}_k, \boldsymbol{y}_0, \ldots, \boldsymbol{y}_k} \left[\tilde{\boldsymbol{x}}_k^{\mathrm{P}} \boldsymbol{e}_k^{\mathrm{T}} \right] \boldsymbol{C}_{\boldsymbol{e}_k}^{-1} \boldsymbol{C}_{\boldsymbol{x}_k, \boldsymbol{e}_k}^{\mathrm{T}} \\ &\quad - \boldsymbol{C}_{\boldsymbol{x}_k, \boldsymbol{e}_k} \boldsymbol{C}_{\boldsymbol{e}_k}^{-1} \mathrm{E}_{\boldsymbol{x}_k, \boldsymbol{y}_0, \ldots, \boldsymbol{y}_k} \left[\boldsymbol{e}_k \tilde{\boldsymbol{x}}_k^{\mathrm{P,T}} \right] + \boldsymbol{C}_{\boldsymbol{x}_k, \boldsymbol{e}_k} \boldsymbol{C}_{\boldsymbol{e}_k}^{-1} \mathrm{E}_{\boldsymbol{y}_0, \ldots, \boldsymbol{y}_k} \left[\boldsymbol{e}_k \boldsymbol{e}_k^{\mathrm{T}} \right] \boldsymbol{C}_{\boldsymbol{e}_k}^{-1} \boldsymbol{C}_{\boldsymbol{x}_k, \boldsymbol{e}_k}^{\mathrm{T}} \\ &= \boldsymbol{C}_{\tilde{\boldsymbol{x}}_k}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}_k}^{\mathrm{P}} \boldsymbol{C}_k^{\mathrm{T}} \left(\boldsymbol{C}_k \boldsymbol{C}_{\tilde{\boldsymbol{x}}_k}^{\mathrm{P}} \boldsymbol{C}_k^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}_k} \right)^{-1} \boldsymbol{C}_k \boldsymbol{C}_{\tilde{\boldsymbol{x}}_k}^{\mathrm{P}}. \end{aligned} \tag{A198}$$

An important step for the derivation of the Kalman filter is the mapping of the observations using an affine function:

$$\begin{bmatrix} \boldsymbol{y}_0 \\ \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_{k-1} \\ \boldsymbol{y}_k \end{bmatrix} \xrightarrow{\boldsymbol{T}_k(\bullet) + \boldsymbol{s}_k} \begin{bmatrix} \boldsymbol{y}_0 \\ \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_{k-1} \\ \boldsymbol{y}_k - \mathrm{E}_{\boldsymbol{y}_k | \boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{k-1}} \left[\boldsymbol{y}_k | \, \boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{k-1} \right] \end{bmatrix} = \begin{bmatrix} \boldsymbol{\gamma}_{k-1} \\ \\ \boldsymbol{e}_k \end{bmatrix}. \tag{A199}$$

This step decomposes the available observations into $\boldsymbol{\gamma}_{k-1}$, which is used to compute $\hat{\boldsymbol{x}}_{k-1}$, and $\boldsymbol{e}_k$, which represents all information about the new observation $\boldsymbol{y}_k$ *that is not contained in* $\boldsymbol{\gamma}_{k-1}$. This is the reason why $\boldsymbol{e}_k$ is commonly called *innovation*. In other words, $\boldsymbol{e}_k$ is the part of $\boldsymbol{y}_k$ which lies in the space of random variables which are orthogonal to the space of random variables spanned by the observations $\boldsymbol{y}_i$, $i \in \{0, 1, \ldots, k-1\}$. Thus, $\boldsymbol{\gamma}_{k-1}$ and $\boldsymbol{e}_k$ are orthogonal (and uncorrelated since $\boldsymbol{e}_k$ is a zero mean random vector).

In Equation (A199), an invertible affine function has been used to decorrelate the observation $\boldsymbol{y}_k$ and the set of earlier observations, given by the vector $\boldsymbol{\gamma}_{k-1}$. Of course, such a function can be applied for every $k \in \mathbb{N}_0$, i. e.,

$$\begin{aligned} \boldsymbol{e}_k &= \boldsymbol{y}_k - \mathrm{E}_{\boldsymbol{y}_k | \boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{k-1}} \left[\boldsymbol{y}_k | \, \boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{k-1} \right] \\ &= \boldsymbol{y}_k - \boldsymbol{C}_k \hat{\boldsymbol{x}}_k^{\mathrm{P}} - \boldsymbol{\mu}_{\boldsymbol{v}_k} \end{aligned} \tag{A200}$$

[33] Note that the covariance matrix $\boldsymbol{C}_{\tilde{\boldsymbol{x}}_k}$ has been defined differently in Equation (A173), i. e., using the conditional distribution of $\boldsymbol{x}_k$ given the observations $\boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_k$, while Equation (A198) uses their joint distribution. This is not a contradiction since the estimation error $\tilde{\boldsymbol{x}}_k$ is uncorrelated with the observations (cf. Equation A179) and thus also stochastically independent due to Gaussianity. Consequently, $\boldsymbol{C}_{\tilde{\boldsymbol{x}}_k} = \boldsymbol{C}_{\boldsymbol{x}_k | \boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_k} = \mathrm{E}_{\boldsymbol{x}_k | \boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_k} \left[(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k)(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k)^{\mathrm{T}} \middle| \, \boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_k \right] = \mathrm{E}_{\tilde{\boldsymbol{x}}_k | \boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_k} \left[\tilde{\boldsymbol{x}}_k \tilde{\boldsymbol{x}}_k^{\mathrm{T}} \middle| \, \boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_k \right] = \mathrm{E}_{\tilde{\boldsymbol{x}}_k} \left[\tilde{\boldsymbol{x}}_k \tilde{\boldsymbol{x}}_k^{\mathrm{T}} \right] = \mathrm{E}_{\boldsymbol{x}_k, \boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_k} \left[(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k)(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k)^{\mathrm{T}} \right]$.

for $k \in \mathbb{N}$, and $\boldsymbol{e}_0 = \boldsymbol{y}_0 - \boldsymbol{\mu}_{\boldsymbol{y}_0}$. In this case, the sequence $(\boldsymbol{e}_k : k \in \mathbb{N}_0)$ is called the *innovation sequence*. It has the property that it is an orthogonal (and uncorrelated since it is zero mean) sequence. Consequently, the covariance matrices which describe the joint densities of all subsets of random vectors from the sequence $(\boldsymbol{e}_k : k \in \mathbb{N}_0)$ are block diagonal. This allows for a simple sequential computation of the parameters of the conditional densities $f_{\boldsymbol{x}_k|\boldsymbol{e}_0,\ldots,\boldsymbol{e}_k}(\boldsymbol{x}_k|\boldsymbol{e}_0,\ldots,\boldsymbol{e}_k)$ which is equivalent (not identical) to the density $f_{\boldsymbol{x}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_k}(\boldsymbol{x}_k|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_k)$ because observations are mapped to innovations by invertible affine functions.

The idea of a decorrelation of observations enables a sequential computation of the mean vector and covariance matrix of the distribution of the state vector $\boldsymbol{x}_k$ given the observations $\boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_k$. It is worth noting that it is not necessary to store all these observations, i. e., the Kalman filter algorithm has finite memory requirements. Equation (A200) shows that the computation of the innovation at time index k just needs the current observation $\boldsymbol{y}_k$ and the predicted system state $\hat{\boldsymbol{x}}_k^{\mathrm{P}}$. Since the latter one can be calculated sequentially (cf. Equations A181, A196, A182 and A198), it is only necessary to keep the current state estimate, estimation error covariance matrix and observation in the memory.

The innovation sequence is an equivalent representation of the original observation sequence since the mapping from the latter to the former is invertible. The motivation for this mapping is the fact that the elements of the innovation sequence are uncorrelated and thus orthogonal because they have zero mean. This orthogonalization is performed sequentially, i. e., each new observation $\boldsymbol{y}_{k+1}$, $k \in \mathbb{N}_0$, is orthogonalized against all previous observations $(\boldsymbol{y}_i : i \in \{0, \ldots, k\})$ or, equivalently, innovations $(\boldsymbol{e}_i : i \in \{0, \ldots, k\})$. In the context of sequential orthogonalization, the transformation of the observations to the innovation sequence is also called the Gram-Schmidt method (see, e. g., [34, pp. 270-272].

Summary A7.2 Let $(\boldsymbol{x}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{y}_k : k \in \mathbb{N}_0)$ be two random sequences generated by the state space model

$$\begin{aligned} \boldsymbol{x}_{k+1} &= \boldsymbol{A}_k \boldsymbol{x}_k + \boldsymbol{w}_k, \\ \boldsymbol{y}_k &= \boldsymbol{C}_k \boldsymbol{x}_k + \boldsymbol{v}_k, \qquad k \in \mathbb{N}_0, \end{aligned}$$

where $(\boldsymbol{w}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{v}_k : k \in \mathbb{N}_0)$ are independently (not necessarily identically) distributed random sequences with $\boldsymbol{w}_k \sim \mathcal{N}\left(\boldsymbol{\mu}_{\boldsymbol{w}_k}, \boldsymbol{C}_{\boldsymbol{w}_k}\right)$ and $\boldsymbol{v}_k \sim \mathcal{N}\left(\boldsymbol{\mu}_{\boldsymbol{v}_k}, \boldsymbol{C}_{\boldsymbol{v}_k}\right)$, $k \in \mathbb{N}_0$. Additionally, the initial value $\boldsymbol{x}_0 \sim \mathcal{N}\left(\boldsymbol{\mu}_{\boldsymbol{x}_0}, \boldsymbol{C}_{\boldsymbol{x}_0}\right)$, the sequence $(\boldsymbol{w}_k : k \in \mathbb{N}_0)$ and $(\boldsymbol{v}_k : k \in \mathbb{N}_0)$ are mutually independent. Then, the parameters of the PDF $f_{\boldsymbol{x}_k|\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_k}(\boldsymbol{x}_k|\boldsymbol{y}_0,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_k)$ can be calculated sequentially by

$$\begin{aligned} \hat{\boldsymbol{x}}_k &= \hat{\boldsymbol{x}}_k^{\mathrm{P}} + \boldsymbol{C}_{\tilde{\boldsymbol{x}}_k}^{\mathrm{P}} \boldsymbol{C}_k^{\mathrm{T}} \left(\boldsymbol{C}_k \boldsymbol{C}_{\tilde{\boldsymbol{x}}_k}^{\mathrm{P}} \boldsymbol{C}_k^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}_k}\right)^{-1} \left(\boldsymbol{y}_k - \boldsymbol{C}_k \hat{\boldsymbol{x}}_k^{\mathrm{P}} - \boldsymbol{\mu}_{\boldsymbol{v}_k}\right), \\ \boldsymbol{C}_{\tilde{\boldsymbol{x}}_k} &= \boldsymbol{C}_{\tilde{\boldsymbol{x}}_k}^{\mathrm{P}} - \boldsymbol{C}_{\tilde{\boldsymbol{x}}_k}^{\mathrm{P}} \boldsymbol{C}_k^{\mathrm{T}} \left(\boldsymbol{C}_k \boldsymbol{C}_{\tilde{\boldsymbol{x}}_k}^{\mathrm{P}} \boldsymbol{C}_k^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{v}_k}\right)^{-1} \boldsymbol{C}_k \boldsymbol{C}_{\tilde{\boldsymbol{x}}_k}^{\mathrm{P}}, \qquad k \in \mathbb{N}_0, \end{aligned}$$

and

$$\begin{aligned} \hat{\boldsymbol{x}}_k^{\mathrm{P}} &= \boldsymbol{A}_{k-1} \hat{\boldsymbol{x}}_{k-1} + \boldsymbol{\mu}_{\boldsymbol{w}_{k-1}}, \\ \boldsymbol{C}_{\tilde{\boldsymbol{x}}_k}^{\mathrm{P}} &= \boldsymbol{A}_{k-1} \boldsymbol{C}_{\tilde{\boldsymbol{x}}_{k-1}} \boldsymbol{A}_{k-1}^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{w}_{k-1}}, \qquad k \in \mathbb{N}. \end{aligned}$$

The initial values are $\hat{\boldsymbol{x}}_0^{\mathrm{P}} = \boldsymbol{\mu}_{\boldsymbol{x}_0}$ and $\boldsymbol{C}_{\tilde{\boldsymbol{x}}_0}^{\mathrm{P}} = \boldsymbol{C}_{\boldsymbol{x}_0}$.

A8. Convexity of the Trade-off Curve of Pareto Optimal Values

Let $f_1 : \mathbb{R}^N \to \mathbb{R}_{+,0}$ and $f_2 : \mathbb{R}^N \to \mathbb{R}_{+,0}$ be two convex functions and consider the joint minimization of f_1 and f_2 w.r.t. $\boldsymbol{x} \in \mathbb{R}^N$. By convexity, it holds (cf., e. g., [94, Section 3.1.1])

$$\begin{aligned} f_1(\alpha\boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2) &\leq \alpha f_1(\boldsymbol{x}_1) + (1-\alpha) f_1(\boldsymbol{x}_2), \qquad \text{and} \\ f_2(\alpha\boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2) &\leq \alpha f_2(\boldsymbol{x}_1) + (1-\alpha) f_2(\boldsymbol{x}_2), \end{aligned} \tag{A201}$$

for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^N$ and $\alpha \in [0,1]$. This is equivalent to

$$\begin{bmatrix} f_1(\alpha\boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2) \\ f_2(\alpha\boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2) \end{bmatrix} \leq \begin{bmatrix} \alpha f_1(\boldsymbol{x}_1) + (1-\alpha) f_1(\boldsymbol{x}_2) \\ \alpha f_2(\boldsymbol{x}_1) + (1-\alpha) f_2(\boldsymbol{x}_2) \end{bmatrix} = \alpha \begin{bmatrix} f_1(\boldsymbol{x}_1) \\ f_2(\boldsymbol{x}_1) \end{bmatrix} + (1-\alpha) \begin{bmatrix} f_1(\boldsymbol{x}_2) \\ f_2(\boldsymbol{x}_2) \end{bmatrix}, \tag{A202}$$

where the inequality has to be understood component-wise.

Now, let $\boldsymbol{z}_1^{*,\mathrm{T}} = [f_1(\boldsymbol{x}_1^*),\ f_2(\boldsymbol{x}_1^*)]^{\mathrm{T}}$ and $\boldsymbol{z}_2^{*,\mathrm{T}} = [f_1(\boldsymbol{x}_2^*),\ f_2(\boldsymbol{x}_2^*)]^{\mathrm{T}}$ be two Pareto optimal values, i. e., $\boldsymbol{x}_1^*$ and $\boldsymbol{x}_2^*$ are Pareto optimal points (for a definition of Pareto optimal values and points see, e. g., [94, Section 4.7.3]). Then, Equation (A202) tells us that the curve in $\mathbb{R}^2_{+,0}$ which connects $\boldsymbol{z}_1^*$ and $\boldsymbol{z}_2^*$ when f_1 and f_2 are evaluated at $\boldsymbol{x} = \alpha\boldsymbol{x}_1^* + (1-\alpha)\boldsymbol{x}_2^*$, $\alpha \in [0,1]$, lies below the straight line which connects $\boldsymbol{z}_1^*$ and $\boldsymbol{z}_2^*$. Thus, there must exist at least one more Pareto optimal value along this curve. The same argumentation can be repeated with an arbitrary pair of Pareto optimal values. This implies that the set of all these values, which describes the trade-off curve for the two objectives represented by f_1 and f_2, is convex. Otherwise, we obtain a contradiction because for a non-convex trade-off curve there exist triples of Pareto optimal values such that one of them lies above the straight line which connects the other two.

The argumentation above is not a rigorous proof but provides the basic ideas to understand the convexity of the optimal trade-off curve which is obtained by the determination of Pareto optimal values of two convex functions.

A9. Consistency of a Lower Bound for Convex-Monotonic Optimization Problems

With reference to [103], a convex-monotonic optimization problem exhibits convexity w.r.t. to a subset of optimization variables for fixed values of all other variables and monotonicity in those other variables for fixed values of the "convex" ones. In Chapters 4 and 5, several convex-monotonic minimization problems are solved using a branch and bound approach to determine the optimizing values of the "non-convex" variables.

With the assumption that strong duality holds, the solution of the convex part of the constrained minimization problem is obtained by maximizing the corresponding dual function (see, e. g., [94, Chapter 5] or [95, Chapter 6]). In order to generalize the notation presented in Chapters 4 and 5, i. e., Equations (4.52), (4.59) and (5.14), let the dual function be $L^*(\boldsymbol{\alpha}, \lambda_1, \lambda_2)$, where the vector $\boldsymbol{\alpha}$ contains the positive valued, "non-convex" variables, and $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ are the dual variables for two dualized inequality constraints.[34] Thus, the solution of the considered optimization problem is given by

$$I(\boldsymbol{\alpha}) = \sup_{\substack{\lambda_1 \geq 0 \\ \lambda_2 \geq 0}} L^*(\boldsymbol{\alpha}, \lambda_1, \lambda_2). \tag{A203}$$

[34] Comparing with Chapters 4 and 5, the components of $\boldsymbol{\alpha}$ represent the (ratio of the) transmit and receive scaling factors t and g, respectively, or the diagonal entries of the diagonal matrices $\boldsymbol{T}$ and $\boldsymbol{G}$, respectively. If only one inequality (power) constraint is considered, λ_1 or λ_2, respectively, has to be set to zero.

For the cases considered in Chapters 4 and 5, i. e., for channel noise covariances and weighting matrices $\boldsymbol{R}$ for the control input with full rank, the dual function L^* is continuous w.r.t. $\boldsymbol{\alpha}$ (see [104] for a discussion of the continuity of stabilizing solutions of a DARE). It is also continuous[35] and concave w.r.t. λ_1 and λ_2. For continuity, we refer again to [104], whereas the concavity always holds for dual functions [94, Section 5.1.2].

In order to obtain a lower bound for the minimal value of $I(\boldsymbol{\alpha})$ w.r.t. $\boldsymbol{\alpha}$ when these parameters can be chosen from a set A, the monotonicity of the dual function L^* has been used in Chapters 4 and 5 to obtain a lower bound

$$\underline{L}_A^*(\lambda_1, \lambda_2) \leq L^*(\boldsymbol{\alpha}, \lambda_1, \lambda_2), \qquad \boldsymbol{\alpha} \in A, \tag{A204}$$

which holds for all $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$. Consequently, the inequality also holds for the supremum, i. e.,

$$\underline{I}_A = \sup_{\substack{\lambda_1 \geq 0 \\ \lambda_2 \geq 0}} \underline{L}_A^*(\lambda_1, \lambda_2) \leq \sup_{\substack{\lambda_1 \geq 0 \\ \lambda_2 \geq 0}} L^*(\boldsymbol{\alpha}, \lambda_1, \lambda_2) = I(\boldsymbol{\alpha}), \qquad \boldsymbol{\alpha} \in A. \tag{A205}$$

The lower bound for the dual function shown in Equation (A204) is itself the dual function of a relaxed version of the original minimization problem, see Equations (4.52), (4.59) and (5.14). Thus, it is also concave in λ_1 and λ_2 and its supremum $\underline{I}_A$ is the optimal value of the relaxed problem.

For the guarantee that the branch and bound algorithm described in [103] finds the global minimum of $I(\boldsymbol{\alpha})$ over all possible parameter vectors $\boldsymbol{\alpha}$, the bound derived in Equation (A205) must be consistent, i. e., it has to be shown that if the set A collapses to a single point $\boldsymbol{\alpha}$, the value of the lower bound converges to the optimal value $I(\boldsymbol{\alpha})$ of the original optimization problem. To this end, let the set A be parametrized by the two vectors $\underline{\boldsymbol{\alpha}}$ and $\overline{\boldsymbol{\alpha}}$ with positive entries according to

$$A = \{\boldsymbol{\alpha} \,|\underline{\boldsymbol{\alpha}} \leq \boldsymbol{\alpha} \leq \overline{\boldsymbol{\alpha}}\}, \tag{A206}$$

where the inequalities have to be understood component-wise. Using such sets, the lower bound for the dual function given by Equation (A204) has the property that if $B \subset A$, it follows that $\underline{L}_A^*(\lambda_1, \lambda_2) \leq \underline{L}_B^*(\lambda_1, \lambda_2)$ for all $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$, which is due to the monotonicity of the original optimization problem w.r.t. the parameter vector $\boldsymbol{\alpha}$. Thus, $\underline{L}_A^*(\lambda_1, \lambda_2)$ is monotonically increasing w.r.t. $\underline{\boldsymbol{\alpha}}$ and decreasing w.r.t. $\overline{\boldsymbol{\alpha}}$. Note further that for the specific problems investigated in Chapters 4 and 5, $\underline{L}_A^*(\lambda_1, \lambda_2)$ is also continuous w.r.t. $\underline{\boldsymbol{\alpha}}$ and $\overline{\boldsymbol{\alpha}}$, which is a result of the continuity of the stabilizing solutions of the respective DAREs w.r.t. to the parameter vectors (see [104]). Additionally, taking into account Equation (A206), it holds

$$\underline{L}_A^*(\lambda_1, \lambda_2) = L^*(\boldsymbol{\alpha}, \lambda_1, \lambda_2) \qquad \text{for} \qquad \underline{\boldsymbol{\alpha}} = \overline{\boldsymbol{\alpha}} = \boldsymbol{\alpha}, \tag{A207}$$

i. e., the lower bound for the dual function is tight if the set A collapses to a single point. Having stated this property, which can be verified with Equations (4.52), (4.59), and (5.14), it will be shown that this is also true for the supremum. To this end, we discuss the feasible and infeasible case separately.

[35]Regarding the results of [104], the solution of the DARE in Equation (4.47) might not be continuous at $\lambda = 0$. Nevertheless, since the dual function is concave w.r.t. the dual variable, $\lambda = 0$ can not be the maximizer of the dual function when it is not continuous at this point.

Case 1: The optimization problem is feasible using the parameters $\boldsymbol{\alpha}$, i. e., $I(\boldsymbol{\alpha})$ is finite.

Since $I(\boldsymbol{\alpha})$ is the supremum of the dual function w.r.t. λ_1 and λ_2, there exist $\lambda_1' \geq 0$ and $\lambda_2' \geq 0$ such that

$$I(\boldsymbol{\alpha}) - L^*(\boldsymbol{\alpha}, \lambda_1', \lambda_2') < \varepsilon_1 \tag{A208}$$

for each $\varepsilon_1 > 0$. Additionally, due to the continuity and monotonicity of the lower bound $\underline{L}_A^*(\lambda_1, \lambda_2)$ and with Equation (A207), there exist $\underline{\boldsymbol{\alpha}} \leq \boldsymbol{\alpha}$ and $\overline{\boldsymbol{\alpha}} \geq \boldsymbol{\alpha}$ such that

$$L^*(\boldsymbol{\alpha}, \lambda_1', \lambda_2') - \underline{L}_A^*(\lambda_1', \lambda_2') < \varepsilon_2 \tag{A209}$$

for each $\varepsilon_2 > 0$. Note that this inequality holds for every set $B \subset A$ which is parametrized according to Equation (A206). Adding Equation (A208) and (A209), we get the inequality

$$I(\boldsymbol{\alpha}) - \underline{L}_A^*(\lambda_1', \lambda_2') < \varepsilon_1 + \varepsilon_2 \tag{A210}$$

for each $\varepsilon_1 + \varepsilon_2 > 0$. Since the supremum of $\underline{L}_A^*$ is larger than or equal to $\underline{L}_A^*(\lambda_1', \lambda_2')$, it can be seen that for a set A which is sufficiently small, it holds

$$I(\boldsymbol{\alpha}) - \underline{I}_A < \varepsilon \tag{A211}$$

for each $\varepsilon = \varepsilon_1 + \varepsilon_2 > 0$, i. e., the optimal value $\underline{I}_A$ of the lower bound converges to the optimal value $I(\boldsymbol{\alpha})$ of the original optimization problem if the set A collapses to the point $\boldsymbol{\alpha}$.

Case 2: The optimization problem is infeasible using the parameters $\boldsymbol{\alpha}$, i. e., $I(\boldsymbol{\alpha}) = \infty$.

It will be shown by contradiction that in this case also the lower bound $\underline{I}_A$ for the optimal value $I(\boldsymbol{\alpha})$ is not bounded from above and thus $\underline{I}_A \to \infty$ when the set A collapses to the point $\boldsymbol{\alpha}$. Assume that $\underline{I}_A$ converges to the finite value $c > 0$ for $\underline{\boldsymbol{\alpha}} \to \boldsymbol{\alpha}$ and $\overline{\boldsymbol{\alpha}} \to \boldsymbol{\alpha}$. Since $L^*(\boldsymbol{\alpha}, \lambda_1, \lambda_2)$ is unbounded from above, there exist λ_1' and λ_2' such that

$$L^*(\boldsymbol{\alpha}, \lambda_1', \lambda_2') - c > \varepsilon_1 \tag{A212}$$

for each $\varepsilon_1 > 0$. Now, let $\varepsilon_2 > 0$ and $\delta > 0$ with $\varepsilon_1 = \varepsilon_2 + \delta$. Then, with Equation (A209), there exist $\underline{\boldsymbol{\alpha}} \leq \boldsymbol{\alpha}$ and $\overline{\boldsymbol{\alpha}} \geq \boldsymbol{\alpha}$ such that

$$\underline{L}_A^*(\lambda_1', \lambda_2') - L^*(\boldsymbol{\alpha}, \lambda_1', \lambda_2') > \delta - \varepsilon_1. \tag{A213}$$

Adding Equation (A212) and (A213), we get

$$\underline{L}_A^*(\lambda_1', \lambda_2') - c > \delta. \tag{A214}$$

Since $\underline{I}_A$ is the supremum of $\underline{L}_A^*$ and thus larger than or equal to $\underline{L}_A^*(\lambda_1', \lambda_2')$, it can be seen that $\underline{I}_A$ does not converge to the finite value c and thus growths without an upper bound.

The two cases show that under the assumption that strong duality holds for the considered optimization problems, the lower bound $\underline{I}_A$ for the optimal value $I(\boldsymbol{\alpha})$ is consistent.

Bibliography

[1] C. E. Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, July and Oct. 1948, reprinted with corrections. [Online]. Available: http://cm.bell-labs.com/cm/ms/what/shannonday/paper.html

[2] J. Proakis, *Digital Communications*, 4th ed., ser. McGraw-Hill Series in Electrical and Computer Engineering. Boston, MA, USA: McGraw-Hill, 2001, iternational ed.

[3] K. J. Åström, *Introduction to Stochastic Control Theory*. Mineola, NY, USA: Dover Publications, Inc., 2006, Reprint. Originally published by Academic Press, Inc., New York, in 1970 as Volume 70 in the Mathematics in Science and Engineering series.

[4] H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*. New York, NY, USA: John Wiley & Sons, Inc., 1972.

[5] T. Söderström, *Discrete-Time Stochastic Systems, Estimation and Control*, ser. Prentice Hall International Series in Systems and Control Engineering. New York, NY, USA: Prentice Hall, 1994.

[6] D. P. Bertsekas, *Dynamic Programming and Optimal Control, Volume 1*, 3rd ed. Belmont, MA, USA: Athena Scientific, 2005.

[7] T. Kailath, *Linear Systems*, ser. Prentice-Hall Information and System Sciences. Englewood Cliffs, NJ, USA: Prentice-Hall, Inc., 1980.

[8] P. Antsaklis and J. Baillieul, "Guest Editorial Special Issue on Networked Control Systems," *IEEE Trans. Automat. Contr.*, vol. 49, no. 9, pp. 1421–1423, Sept. 2004.

[9] B. Zhu, B. Sinopoli, K. Poolla, and S. Sastry, "Estimation over Wireless Sensor Networks," in *Proceedings of the 2007 American Control Conference*, pp. 2732–2737, July 2007.

[10] B. Sinopoli, C. Sharp, L. Schenato, S. Schaffert, and S. S. Sastry, "Distributed Control Applications Within Sensor Networks," *Proc. IEEE*, vol. 91, no. 8, pp. 1235–1246, Aug. 2003.

[11] M. Epstein, L. Shi, and R. M. Murray, "Estimation Schemes for Networked Control Systems Using UDP-Like Communication," in *Proceedings of the 46th IEEE Conference on Decision and Control*, pp. 3945–3951, Dec. 2007.

[12] E. Garone, B. Sinopoli, and A. Casavola, "LQG Control Over Lossy TCP-like Networks With Probabilistic Packet Acknowledgements," in *Proceedings of the 47th IEEE Conference on Decision and Control*, pp. 2686–2691, Dec. 2008.

[13] R. C. Luo and T. Fukuda, Eds., "Special Issue on Networked Intelligent Robots Through the Internet," *Proc. IEEE*, vol. 91, no. 3, Mar. 2003.

[14] W. Zhang, M. S. Branicky, and S. M. Phillips, "Stability of Networked Control Systems," *IEEE Control Syst. Mag.*, vol. 21, no. 1, pp. 84–99, Feb. 2001.

[15] X. Liu and A. Goldsmith, "Wireless Medium Access Control in Networked Control Systems," in *Proceedings of the 2004 American Control Conference*, vol. 4, pp. 3605–3610, June 2004.

[16] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, "A Survey of Recent Results in Networked Control Systems," *Proc. IEEE*, vol. 95, no. 1, pp. 138–162, Jan. 2007, Special Issue on Technology of Networked Control Systems.

[17] B. Johansson, A. Speranzon, M. Johansson, and K. H. Johansson, "Distributed Model Predictive Consensus," in *Proceedings of the 17th International Symposium on Mathematical Theory of Networks and Systems*, pp. 2438–2444, July 2006.

[18] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and Cooperation in Networked Multi-Agent Systems," *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007, Special Issue on Technology of Networked Control Systems.

[19] L. G. Bushnell, "Networks and Control [Guest Editorial]," *IEEE Control Syst. Mag.*, vol. 21, no. 1, pp. 22–23, Feb. 2001.

[20] P. Antsaklis and J. Baillieul, "Special Issue on Technology of Networked Control Systems," *Proc. IEEE*, vol. 95, no. 1, pp. 5–8, Jan. 2007.

[21] S. H. Low, F. Paganini, and J. C. Doyle, "Internet Congestion Control," *IEEE Control Syst. Mag.*, vol. 22, no. 1, pp. 28–43, Feb. 2002.

[22] W. P. Blair, Jr. and D. D. Sworder, "Feedback control of a class of linear discrete systems with jump parameters and quadratic cost criteria," *International Journal of Control*, vol. 21, no. 5, pp. 833–841, 1975.

[23] B. E. Griffiths and K. A. Loparo, "Optimal control of jump-linear gaussian systems," *International Journal of Control*, vol. 42, no. 4, pp. 791–819, 1985.

[24] M. G. Safonov and M. Athans, "Gain and Phase Margin for Multiloop LQG Regulators," *IEEE Trans. Automat. Contr.*, vol. AC-22, no. 2, pp. 173–179, Apr. 1977.

[25] M. Athans, R. Ku, and S. B. Gershwin, "The Uncertainty Threshold Principle: Some Fundamental Limitations of Optimal Decision Making under Dynamic Uncertainty," *IEEE Trans. Automat. Contr.*, vol. 22, no. 3, pp. 491–495, June 1977.

[26] M. G. Safonov and M. K. H. Fan, "Editorial," *International Journal of Robust and Nonlinear Control*, vol. 7, no. 2, pp. 97–103, Feb. 1997.

[27] D. S. Bernstein and A. N. Michel, "A chronological bibliography on saturating actuators," *International Journal of Robust and Nonlinear Control*, vol. 5, no. 5, pp. 375–380, 1995.

[28] G. N. Nair and R. J. Evans, "Stabilization with data-rate-limited feedback: tightest attainable bounds," *Systems & Control Letters*, vol. 41, no. 1, pp. 49–56, Sept. 2000.

[29] J. H. Braslavsky, R. H. Middleton, and J. S. Freudenberg, "Feedback Stabilization Over Signal-to-Noise Ratio Constrained Channels," *IEEE Trans. Automat. Contr.*, vol. 52, no. 8, pp. 1391–1403, Aug. 2007.

[30] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman Filtering With Intermittent Observations," *IEEE Trans. Automat. Contr.*, vol. 49, no. 9, pp. 1453–1464, Sept. 2004, Special Issue on Networked Control Systems.

[31] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*, ser. Prentice Hall Information and System Sciences Series. Upper Saddle River, NJ, USA: Prentice Hall, Inc., 2000.

[32] G. E. Dullerud and F. Paganini, *A Course in Robust Control Theory, A Convex Approach*, ser. Texts in Applied Mathematics. New York, NY, USA: Springer, 2000, vol. 36.

[33] G. Grimmett and D. Stirzaker, *Probability and Random Processes*, 3rd ed. Oxford, UK: Oxford University Press, 2001.

[34] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*, 4th ed., ser. McGraw-Hill Series in Electrical and Computer Engineering. Boston, MA, USA: McGraw-Hill, 2001.

[35] 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, Feb. 2010.

[36] G. N. Nair and R. J. Evans, "Stabilizability of Stochastic Linear Systems with Finite Feedback Data Rates," *SIAM Journal on Control and Optimization*, vol. 43, no. 2, pp. 413–436, 2004.

[37] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, ser. Prentice-Hall Information and System Sciences Series. Englewood Cliffs, NJ, USA: Prentice-Hall, Inc., 1979.

[38] M. Aoki, *Optimization of Stochastic Systems; Topics in Discrete-Time Systems*, ser. Mathematics in Science and Engineering. New York, NY, USA: Academic Press, Inc., 1967.

[39] H. S. Witsenhausen, "A Counterexample in Stochastic Optimum Control," *SIAM Journal on Control*, vol. 6, no. 1, pp. 131–147, 1968.

[40] S. K. Mitter and A. Sahai, "Information and control: Witsenhausen revisited," in *Learning, Control and Hybrid Systems*, ser. Lecture Notes in Control and Information Sciences, Y. Yamamoto and S. Hara, Eds. Berlin, Germany: Springer-Verlag, 1999, vol. 241, pp. 281–293.

[41] M. Rotkowitz and S. Lall, "A Characterization of Convex Problems in Decentralized Control*," *IEEE Trans. Automat. Contr.*, vol. 51, no. 2, pp. 274–286, Feb. 2006.

[42] A. Sahai and S. Mitter, "The Necessity and Sufficiency of Anytime Capacity for Stabilization of a Linear System over a Noisy Communication Link – Part 1: Scalar Systems," *IEEE Trans. Inform. Theory*, vol. 52, no. 8, pp. 3369–3395, Aug. 2006.

[43] R. G. Gallager, *Information Theory and Reliable Communication*. New York, NY, USA: John Wiley & Sons, Inc., 1968.

[44] L. Xiao, M. Johansson, H. Hindi, S. Boyd, and A. Goldsmith, "Joint Optimization of Wireless Communication and Networked Control Systems," in *Switching and Learning in Feedback Systems*, ser. Lecture Notes in Computer Science, R. Murray-Smith and R. Shorten, Eds. Berlin, Germany: Springer-Verlag, 2005, vol. 3355, pp. 248–272.

[45] V. Gupta, B. Hassibi, and R. M. Murray, "Optimal LQG control across packet-dropping links," *Systems & Control Letters*, vol. 56, no. 6, pp. 439–446, June 2007.

[46] E. I. Silva, D. E. Quevedo, and G. C. Goodwin, "Optimal Controller Design for Networked Control Systems," in *Proceedings of the 17th IFAC World Congress*, pp. 5167–5172, July 2008.

[47] D. Liberzon and A. S. Morse, "Basic Problems in Stability and Design of Switched Systems," *IEEE Control Syst. Mag.*, vol. 19, no. 5, pp. 59–70, Oct. 1999.

[48] J.-W. Lee and G. E. Dullerud, "Uniform stabilization of discrete-time switched and Markovian jump linear systems," *Automatica*, vol. 42, no. 2, pp. 205–218, Feb. 2006.

[49] M. Huang and S. Dey, "Stability of Kalman filtering with Markovian packet losses," *Automatica*, vol. 43, no. 4, pp. 598–607, Apr. 2007.

[50] R. Blind, S. Uhlich, B. Yang, and F. Allgöwer, "Robustification and Optimization of a Kalman Filter with Measurement Loss using Linear Precoding," in *Proceedings of the 2009 American Control Conference*, pp. 2222–2227, June 2009.

[51] L. Shi, M. Epstein, and R. M. Murray, "Kalman Filtering Over a Packet-Dropping Network: A Probabilistic Perspective," *IEEE Trans. Automat. Contr.*, vol. 55, no. 3, pp. 594–604, Mar. 2010.

[52] N. Elia and J. N. Eisenbeis, "Limitations of Linear Remote Control over Packet Drop Networks," in *Proceedings of the 43rd IEEE Conference on Decision and Control*, vol. 5, pp. 5152–5157, Dec. 2004.

[53] Q. Ling and M. D. Lemmon, "Power Spectral Analysis of Networked Control Systems With Data Dropouts," *IEEE Trans. Automat. Contr.*, vol. 49, no. 6, pp. 955–959, June 2004.

[54] O. C. Imer, S. Yüksel, and T. Başar, "Optimal control of LTI systems over unreliable communication links," *Automatica*, vol. 42, no. 9, pp. 1429–1439, Sept. 2006.

[55] C. L. Robinson and P. R. Kumar, "Sending the Most Recent Observation is not Optimal in Networked Control: Linear Temporal Coding and Towards the Design of a Control Specific Transport Protocol," in *Proceedings of the 46th IEEE Conference on Decision and Control*, pp. 334–339, Dec. 2007.

[56] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. S. Sastry, "Foundations of Control and Estimation Over Lossy Networks," *Proc. IEEE*, vol. 95, no. 1, pp. 163–187, Jan. 2007, Special Issue on Technology of Networked Control Systems.

[57] C. L. Robinson and P. R. Kumar, "Optimizing Controller Location in Networked Control Systems with Packet Drops," *IEEE J. Select. Areas Commun.*, vol. 26, no. 4, pp. 661–671, May 2008.

[58] E. Silva, G. Goodwin, and D. Quevedo, "On the Design of Control Systems over Unreliable Channels," in *Proceedings of the European Control Conference 2009*, pp. 377–382, Aug. 2009.

[59] G. C. Goodwin, D. E. Quevedo, and E. I. Silva, "Architectures and coder design for networked control systems," *Automatica*, vol. 44, no. 1, pp. 248–257, Jan. 2008.

[60] D. F. Delchamps, "Stabilizing a Linear System with Quantized State Feedback," *IEEE Trans. Automat. Contr.*, vol. 35, no. 8, pp. 916–924, Aug. 1990.

[61] R. W. Brockett and D. Liberzon, "Quantized Feedback Stabilization of Linear Systems," *IEEE Trans. Automat. Contr.*, vol. 45, no. 7, pp. 1279–1289, July 2000.

[62] N. Elia and S. K. Mitter, "Stabilization of Linear Systems With Limited Information," *IEEE Trans. Automat. Contr.*, vol. 46, no. 9, pp. 1384–1400, Sept. 2001.

[63] H. Ishii and B. A. Francis, *Limited Data Rate in Control Systems with Networks*, ser. Lecture Notes in Control and Information Sciences. Berlin, Germany: Springer-Verlag, 2002, vol. 275.

[64] S. Tatikonda and S. Mitter, "Control Under Communication Constraints," *IEEE Trans. Automat. Contr.*, vol. 49, no. 7, pp. 1056–1068, July 2004.

[65] G. N. Nair, F. Fagnani, S. Zampieri, and R. J. Evans, "Feedback Control under Data Rate Constraints: An Overview," *Proc. IEEE*, vol. 95, no. 1, pp. 108–137, Jan. 2007, Special Issue on Technology of Networked Control Systems.

[66] S. Yüksel and T. Başar, "Control over Noisy Forward and Reverse Channels," *IEEE Trans. Automat. Contr.*, vol. 56, no. 5, pp. 1014–1029, May 2011.

[67] S. P. Meyn and R. L. Tweedy, *Markov Chains and Stochastic Stability*. London, UK: Springer-Verlag, 1993. [Online]. Available: http://probability.ca/MT/

[68] T. Söderström and P. Stoica, *System Identification*, ser. Prentice Hall International Series in Systems and Control Engineering. New York, NY, USA: Prentice Hall, 1989. [Online]. Available: http://user.it.uu.se/~ts/sysidbook.pdf

[69] J. K. Tugnait, L. Tong, and Z. Ding, "Single-User Channel Estimation and Equalization," *IEEE Signal Processing Mag.*, vol. 17, no. 3, pp. 16–28, May 2000.

[70] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*, ser. Addison-Wesley Series in Electrical and Computer Engineering: Digital Signal Processing. Reading, MA, USA: Addison-Wesley, 1991.

[71] R. E. Kalman, "Control of Randomly Varying Linear Dynamical Systems," in *Proceedings of Symposia in Applied Mathematics*, G. Birkhoff, R. Bellman, and C. C. Lin, Eds., vol. 13: Hydrodynamic Instability, pp. 287–298. American Mathematical Society, 1962.

[72] N. E. Nahi, "Optimal Recursive Estimation With Uncertain Observation," *IEEE Trans. Inform. Theory*, vol. IT-15, no. 4, pp. 457–462, July 1969.

[73] P. K. Rajasekaran, N. Satyanarayana, and M. Srinath, "Optimum Linear Estimation of Stochastic Signals in the Presence of Multiplicative Noise," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-7, no. 3, pp. 462–468, May 1971.

[74] J. K. Tugnait, "Stability of Optimum Linear Estimators of Stochastic Signals in White Multiplicative Noise," *IEEE Trans. Automat. Contr.*, vol. AC-26, no. 3, pp. 757–761, June 1981.

[75] W. L. De Koning, "Infinite Horizon Optimal Control of Linear Discrete Time Systems with Stochastic Parameters," *Automatica*, vol. 18, no. 4, pp. 443–453, July 1982.

[76] B. S. Chow and W. P. Birkemeier, "A New Recursive Filter for Systems with Multiplicative Noise," *IEEE Trans. Inform. Theory*, vol. 36, no. 6, pp. 1430–1435, Nov. 1990.

[77] N. Elia, "Remote stabilization over fading channels," *Systems & Control Letters*, vol. 54, no. 3, pp. 237–249, Mar. 2005.

[78] S. P. Boyd and C. H. Barratt, *Linear Controller Design: Limits of Performance*, ser. Prentice Hall Information and System Sciences Series. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1991.

[79] K. Zhou, J. C. Doyle, and K. Glover, *Robust and optimal control*. Upper Saddle River, NJ, USA: Prentice Hall, Inc., 1996.

[80] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, ser. Wiley Series in Telecommunications. New York, NY, USA: John Wiley & Sons, Inc., 1991.

[81] P. Mäkilä, T. Westerlund, and H. Toivonen, "Constrained Linear Quadratic Gaussian Control with Process Applications," *Automatica*, vol. 20, no. 1, pp. 15–29, Jan. 1984.

[82] 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.

[83] G. Zhu, M. A. Rotea, and R. Skelton, "A Convergent Algorithm for the Output Covariance Constraint Control Problem," *SIAM Journal on Control and Optimization*, vol. 35, no. 1, pp. 341–361, Jan. 1997.

[84] G. Shi, R. E. Skelton, and K. M. Grigoriadis, "Minimum Output Variance Control for FSN Models: Continuous-Time Case," *Mathematical Problems in Engineering*, vol. 6, no. 2-3, pp. 171–188, 2000.

[85] C.-K. Ko, X. Gao, S. Prajna, and L. J. Schulman, "On Scalar LQG Control with Communication Cost," in *Proceedings of the 44th IEEE Conference on Decision and Control, and European Control Conference*, pp. 2805–2810, Dec. 2005.

[86] J. S. Freudenberg, R. H. Middleton, and J. H. Braslavsky, "Stabilization with Disturbance Attenuation over a Gaussian Channel," in *Proceedings of the 46th IEEE Conference on Decision and Control*, pp. 3958–3963, Dec. 2007.

[87] J. S. Freudenberg, R. H. Middleton, and V. Solo, "Stabilization and Disturbance Attenuation Over a Gaussian Communication Channel," *IEEE Trans. Automat. Contr.*, vol. 55, no. 3, pp. 795–799, Mar. 2010.

[88] A. Gattami, "Generalized Linear Quadratic Control," *IEEE Trans. Automat. Contr.*, vol. 55, no. 1, pp. 131–136, Jan. 2010.

[89] E. I. Silva, M. S. Derpich, J. Østergaard, and D. E. Quevedo, "Simple Coding for Achieving Mean Square Stability over Bit-rate Limited Channels," in *Proceedings of the 47th IEEE Conference on Decision and Control*, pp. 2698–2703, Dec. 2008.

[90] V. Gupta, A. F. Dana, J. P. Hespanha, R. M. Murray, and B. Hassibi, "Data Transmission Over Networks for Estimation and Control," *IEEE Trans. Automat. Contr.*, vol. 54, no. 8, pp. 1807–1819, Aug. 2009.

[91] V. Gupta, "On an Estimation Oriented Routing Protocol," in *Proceedings of the 2010 American Control Conference*, pp. 580–585, June 2010.

[92] J. F. Sturm, "Using SeDuMi 1.02, a Matlab Toolbox for Optimization Over Symmetric Cones," *Optimization Methods and Software*, vol. 11, no. 1-4, pp. 625–653, 1999, (latest version 1.3). [Online]. Available: http://sedumi.ie.lehigh.edu

[93] J. Löfberg, "YALMIP : A Toolbox for Modeling and Optimization in MATLAB," in *Proceedings of the 2004 IEEE International Symposium on Computer Aided Control Systems Design,*, pp. 284–289, Sept. 2004. [Online]. Available: http://users.isy.liu.se/johanl/yalmip

[94] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.

[95] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming, Theory and Algorithms*, 3rd ed. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2006.

[96] H. T. Toivonen and P. M. Mäkilä, "Computer-aided design procedure for multiobjective LQG control problems," *International Journal of Control*, vol. 49, no. 2, pp. 655–666, Feb. 1989.

[97] A. C. M. Ran and H. L. Trentelman, "Linear Quadratic Problems With Indefinite Cost for Discrete Time Systems," *SIAM Journal on Matrix Analysis and Applications*, vol. 14, no. 3, pp. 776–797, July 1993.

[98] M. A. Rami, X. Chen, and X. Y. Zhou, "Discrete-time Indefinite LQ Control with State and Control Dependent Noises," *Journal of Global Optimization*, vol. 23, no. 3-4, p. 245–265, Aug. 2002.

[99] G. C. Goodwin, E. I. Silva, and D. E. Quevedo, "Analysis and Design of Networked Control Systems Using the Additive Noise Model Methodology," *Asian Journal of Control*, vol. 12, no. 4, p. 443–459, July 2010.

[100] A. C. M. Ran and R. Vreugdenhil, "Existence and Comparison Theorems for Algebraic Riccati Equations for Continuous- and Discrete-Time Systems," *Linear Algebra and its Applications*, vol. 99, pp. 63–83, Feb. 1988.

[101] B. Priel and U. Shaked, " 'Cheap' optimal control of discrete single input single output systems," *International Journal of Control*, vol. 38, no. 6, pp. 1087–1113, 1983.

[102] H. Tuy, F. Al-Khayyal, and P. Thach, "Monotonic Optimization: Branch and Cut Methods," in *Essays and Surveys in Global Optimization*, C. Audet, P. Hansen, and G. Savard, Eds. New York, NY, USA: Springer US, 2005, pp. 39–78.

[103] H. Tuy, "Partly Convex and Convex-Monotonic Optimization Problems," in *Modeling, Simulation and Optimization of Complex Processes*, H. G. Bock, H. X. Phu, E. Kostina, and R. Rannacher, Eds. Berlin, Germany: Springer, 2005, pp. 485–508.

[104] A. A. Stoorvogel and A. Saberi, "Continuity properties of solutions to H_2 and H_∞ Riccati equations," *Systems & Control Letters*, vol. 27, no. 4, pp. 209–222, Apr. 1996.

[105] M. B. Estrada, M. F. Garcia, M. Malabre, and J. C. M. García, "Left invertibility and duality for linear systems," *Linear Algebra and its Applications*, vol. 425, no. 2-3, pp. 345–373, Sept. 2007, Special Issue in honor of Paul Fuhrmann.

[106] A. J. Rojas, J. H. Braslavsky, and R. H. Middleton, "Fundamental limitations in control over a communication channel," *Automatica*, vol. 44, no. 12, pp. 3147–3151, Dec. 2008.

[107] N. Elia, "When Bode Meets Shannon: Control-Oriented Feedback Communication Schemes," *IEEE Trans. Automat. Contr.*, vol. 49, no. 9, pp. 1477–1488, Sept. 2004.

[108] E. I. Silva, "A Unified Framework for the Analysis and Design of Networked Control Systems," Ph.D. dissertation, School of Electrical Engineering and Computer Science, The University of Newcastle, Callaghan, Australia, Feb. 2009. [Online]. Available: http://hdl.handle.net/1959.13/34254

[109] S. A. Pulgar, E. I. Silva, and M. E. Selgado, "Optimal State-Feedback Design for MIMO systems subject to multiple SNR constraints," in *Proceedings of the 18th IFAC World Congress*, pp. 14 410–14 415, Aug. 2011.

[110] H. Sampath, P. Stoica, and A. Paulraj, "Generalized Linear Precoder and Decoder Design for MIMO Channels Using the Weighted MMSE Criterion," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2198–2206, Dec. 2001.

[111] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal Designs for Space-Time Linear Precoders and Decoders," *IEEE Trans. Signal Processing*, vol. 50, no. 5, pp. 1051–1064, May 2002.

[112] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx Beamforming Design for Multicarrier MIMO Channels: A Unified Framework for Convex Optimization," *IEEE Trans. Signal Processing*, vol. 51, no. 9, pp. 2381–2401, Sept. 2003.

[113] H. V. Henderson and S. R. Searle, "On Deriving the Inverse of a Sum of Matrices," *SIAM Review*, vol. 23, no. 1, pp. 53–60, Jan. 1981.

[114] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, UK: Cambridge University Press, 1991.

[115] J. Yang and S. Roy, "On Joint Transmitter and Receiver Optimization for Multiple-Input-Multiple-Output (MIMO) Transmission Systems," *IEEE Trans. Commun.*, vol. 42, no. 12, pp. 3221–3231, Dec. 1994.

[116] E. Johannesson, A. Ghulchak, A. Rantzer, and B. Bernhardsson, "MIMO Encoder and Decoder Design for Signal Estimation," in *Proceedings of the 19th International Symposium on Mathematical Theory of Networks and Systems*, pp. 2019–2025, July 2010.

[117] Z. Shu and R. H. Middleton, "Stabilization Over Power-Constrained Parallel Gaussian Channels," *IEEE Trans. Automat. Contr.*, vol. 56, no. 7, pp. 1718–1724, July 2011.

[118] A. J. Rojas, "Signal-to-noise ratio performance limitations for input disturbance rejection in output feedback control," *Systems & Control Letters*, vol. 58, no. 5, pp. 353–358, May 2009.

[119] R. W. Brockett, "Poles, Zeros, and Feedback: State Space Interpretation," *IEEE Trans. Automat. Contr.*, vol. 10, no. 2, pp. 129–135, Apr. 1965.

[120] Z. Zhang and J. S. Freudenberg, "Discrete-time Loop Transfer Recovery for Systems with Nonminimum Phase Zeros and Time Delays," *Automatica*, vol. 29, no. 2, pp. 351–363, Mar. 1993.

[121] G. C. Goodwin, M. M. Seron, and J. De Doná, *Constrained Control and Estimation , An Optimisation Approach*, ser. Communications and Control Engineering Series. London, UK: Springer-Verlag, 2005.

[122] P. Mäkilä, T. Westerlund, and H. Toivonen, "Constrained Linear Quadratic Gaussian Control," in *Proceedings of the 21st IEEE Conference on Decision and Control*, pp. 312–317, Dec. 1982.

[123] E. Johannesson, A. Rantzer, and B. Bernhardsson, "Optimal Linear Control for Channels with Signal-to-Noise Ratio Constraints," in *Proceedings of the 2011 American Control Conference*, pp. 521–526, June 2011.

[124] R. W. Brockett, *Finite Dimensional Linear Systems*, ser. Series in Decision and Control. New York, NY, USA: John Wiley & Sons, Inc., 1970.

[125] W. Magnus, "On the Exponential Solution of Differential Equations for a Linear Operator," *Communications on Pure and Applied Mathematics*, vol. 7, no. 4, pp. 649–673, Nov. 1954.

[126] R. Bellman, *Introduction to Matrix Analysis*, 2nd ed., ser. Classics In Applied Mathematics. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 1997, Reprint. Originally published by McGraw-Hill, New York, in 1970.

[127] G. Kitagawa, "An algorithm for solving the matrix equation $\boldsymbol{X} = \boldsymbol{F}\boldsymbol{X}\boldsymbol{F}^{\mathrm{T}} + \boldsymbol{S}$," *International Journal of Control*, vol. 25, no. 5, pp. 745–753, 1977.

[128] J. W. Brewer, "Kronecker Products and Matrix Calculus in System Theory," *IEEE Trans. Circuits Syst.*, vol. CAS-25, no. 9, pp. 772–781, Sept. 1978.

[129] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, ser. SIAM Studies in Applied Mathematics. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 1994, vol. 15.

[130] W. F. Arnold, III and A. J. Laub, "Generalized Eigenproblem Algorithms and Software for Algebraic Riccati Equations," *Proc. IEEE*, vol. 72, no. 12, pp. 1746–1754, Dec. 1984.

[131] V. Ionescu and M. Weiss, "On Computing the Stabilizing Solution of the Discrete-Time Riccati Equation," *Linear Algebra and its Applications*, vol. 174, pp. 229–238, Sept. 1992.

[132] A. A. Stoorvogel and A. Saberi, "The Discrete Algebraic Riccati Equation and Linear Matrix Inequality," *Linear Algebra and its Applications*, vol. 274, no. 1-3, pp. 317–365, Apr. 1998.

[133] P. S. Maybeck, *Stochastic Models, Estimation and Control, Volume 1*, ser. Mathematics in Science and Engineering. New York, NY, USA: Academic Press, Inc., 1979, vol. 141.

[134] R. E. Bellman, *Dynamic Programming*. Princeton, NJ, USA: Princeton University Press, 1957.

[135] G. D. Forney, Jr., "The Viterbi Algorithm," *Proc. IEEE*, vol. 61, no. 3, pp. 268–278, Mar. 1973.

[136] D. P. Bertsekas, *Dynamic Programming and Optimal Control, Volume 2*, 3rd ed. Belmont, MA, USA: Athena Scientific, 2005.

[137] T. Chen and B. A. Francis, "State-Space Solutions to Discrete-Time and Sampled-Data $\mathcal{H}_2$ Control Problems," in *Proceedings of the 31st IEEE Conference on Decision and Control*, vol. 1, pp. 1111–1116, Dec. 1992.

[138] ——, *Optimal Sampled Data Control Systems*, ser. Communications and Control Engineering Series. London, UK: Springer-Verlag, 1995. [Online]. Available: http://www.control.utoronto.ca/people/profs/francis/sd_book.pdf

[139] A. Saberi, P. Sannuti, and B. M. Chen, H_2 *Optimal Control*, ser. Prentice Hall International Series in Systems and Control Engineering. London, UK: Prentice Hall, 1995.

[140] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Transactions of the ASME–Journal of Basic Engineering*, vol. 82, Series D, pp. 35–45, 1960.