A Unified Framework for the Analysis and Design of Networked Control Systems

Eduardo I. Silva

Ingeniero Civil Electrónico
M.Sc. in Electronic Engineering

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

School of Electrical Engineering
and Computer Science

The University of Newcastle
Callaghan, NSW 2308, Australia

February, 2009

The thesis contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying subject to the provisions of the Copyright Act 1968.

I hereby certify that the work embodied in this thesis has been done in collaboration with other researchers, including my supervisors. I have included as part of the thesis a statement clearly outlining the extent of collaboration, with whom and under what auspices.

I hereby certify that the work embodied in this thesis contains published papers and scholarly work of which I am a joint author. I have included as part of the thesis a written statement, endorsed by my supervisor, attesting to my contribution to the joint publications and scholarly work.

Eduardo I. Silva, 18/02/2009

I hereby certify that parts of this thesis (in particular Chapter 3, and early versions of Chapters 4 and 7) are the result of collaboration with my supervisors Prof. G.C. Goodwin and Dr. D.E. Quevedo. This collaboration has been based on normal candidate-supervisors relations. Chapters 5 and 6 are the result of collaboration with Mr. M.S. Derpich and Dr. J. Østergaard (University of Aalborg, Denmark). Again, this collaboration has been based on a normal collaborative practice, where the bulk of work and intellectual effort has been undertaken by myself.

I hereby certify that parts of this thesis, and early versions of some of the results herein, have been previously published in peer reviewed conferences and journals. In such publications, my contributions have been either significant or fundamental.

Eduardo I. Silva, 18/02/2009

Acknowledgements

This thesis is the result of work carried out with the help and support of several individuals.

I would like to first acknowledge the unconditional love and support of my wife, Andrea, and my beautiful little princess, Isabel. I know that you have put your dreams on hold to be by my side while I realized mine.

Secondly, I would like to thank my supervisor, Prof. Graham C. Goodwin, for his bulletproof good disposition, and for giving me the freedom to work on what I saw fit. Giving too much freedom to a student is always double-edged. I hope that my work lived up to his expectations. I am also grateful to my co-supervisor, Dr. Daniel E. Quevedo. Even though his views of things are not always compatible with mine, his criticism and suggestions always helped to improve my work.

I cannot forget to mention that the person responsible for me ending up in Newcastle is Prof. Mario E. Salgado (Universidad Técnica Federico Santa María, Chile). His constant support has been an important factor in the realization of this thesis.

I must also acknowledge the help and support of Milan S. Derpich and Dr. Jan Østergaard (University of Aalborg, Denmark). Without the stamina of Milan, and the broad and deep knowledge on information theory of Jan, many of the results in this thesis would have never seen the light. Thank you for sharing with me your ideas and time. (I hope we will be able to keep working together in the future.)

During my stay in Newcastle, I had also the chance to interact with Prof. Vincent Wertz (catholic University of Louvain, Belgium). Working with him was stimulating. I had also the opportunity to interact (very shortly) with Prof. Arie Feuer (Technion, Israel). His interest in my work is certainly appreciated.

I must also express my deep gratitude to all the people at the CDSC that made my stay in Newcastle more enjoyable. I must thank Juan C. Aguero for his permanent good disposition and good mood, for sharing with me his good memory with references, and also for sparing me long walks to the library. I must also thank Cristian R. Rojas for being always willing to help with mathematical issues. I also thank Boris I. Godoy for his good disposition.

I must also acknowledge the help, infinite patience, and extremely good disposition of Dianne Piefke and Jayne Disney.

I am also grateful to Professor Stephen Boyd (Stanford, USA) and Professor Jie Chen (University of California, Riverside, USA) for acting as reviewers of this thesis. The comments of the third anonymous reviewer, which led to the correction of an early version of the proof of Theorem 5.3, are also appreciated.

The financial support received from the Centre for Complex Dynamic Systems & Control (CDSC), from The University of Newcastle, and from the Australian Government is also acknowledged.

A Andrea e Isabel.

Contents

Abstract

This thesis studies control systems with communication constraints. Such constraints arise due to the fact that practical control systems often use non-transparent communication links, i.e., links subject to data-rate constraints, random data-dropouts or random delays. Traditional control theory cannot deal with such constraints and the need for new tools and insights arises.

We study two problems: control with average data-rate constraints and control over analog erasure channels with i.i.d. dropout profiles.

When focusing on average data-rate constraints, it is natural to ask whether information theoretic ideas may assist the study of networked control systems. In this thesis we show that it is possible to use fundamental information theoretic concepts to arrive at a framework that allows one to tackle performance related control problems. In doing so, we show that there exists an *exact* link between control systems subject to average data-rate limits, and control systems which are closed over *additive i.i.d. noise channels subject to a signal-to-noise ratio constraint.*

On the other hand, in the case of control systems subject to i.i.d. data-dropouts, we show that there exists a *second-order moments equivalence* between a linear feedback system which is interconnected over an analog erasure channel, and the same system when it is interconnected over an *additive i.i.d. noise channel subject to a signal-to-noise ratio constraint.*

From the results foreshadowed above, it follows that the study of control systems closed over signal-to-noise ratio constrained additive i.i.d. noise channels is a task of relevance to many networked control problems. Moreover, the interplay between signal-to-noise ratio constraints and control objectives is an interesting issue in its own right. This thesis starts with such a study. Then, we use the resultant insights to address performance issues in control systems subject to either average data-rate constraints or i.i.d. data-dropouts. Our approach shows that, once key equivalences are exposed, standard control intuition and synthesis machinery can be used to tackle networked control problems in an *exact* manner. It also sheds light into fundamental results in the literature and gives (partial) answers to several previously open questions. We believe that the insights in this thesis are of fundamental importance and, to the best of the author's knowledge, novel.

Symbols and Acronyms

The following is a list of symbols and acronyms commonly used throughout this thesis. In some cases, the definitions are vague; proper definitions will be given where appropriate.

EC	entropy coder
ECDQ	entropy coded dithered quantizer
ED	entropy decoder
i.i.d.	independent and identically distributed
LTI	linear time invariant
LMI	linear matrix inequality
MIMO	multiple-input multiple-output
MP	minimum phase
MSS	mean square stable or mean square stability
NMP	non-minimum phase
PDF	probability density function
PSD	power spectral density
SISO	single-input single-output
wss	wide sense stationary
$\triangleq$	equal by definition
$\{\cdot\}$	a set
$\{\cdot\}_m$	a multi-set (see [11])
$\#$	cardinality (number of elements)
$\in$	set membership

$\subseteq$	contained in or equal to
$\subset$	contained but not equal to
$\cap$	intersection
$\cup$	union
$\setminus$	set exclusion
$\mathbb{N}_0$	$\{0, 1, 2, 3, \cdots\}$
$\mathbb{Z}$	$\{\cdots, -2, -1, 0, 1, 2, \cdots\}$
$\mathbb{R}$	the real numbers
$\mathbb{R}_0^+$	$\{x \in \mathbb{R} : 0 \leq x < \infty\}$
$\mathbb{R}^+$	$\{x \in \mathbb{R} : 0 < x < \infty\}$
$\mathbb{C}$	the complex plane
$\mathbb{R}^n$	$\mathbb{R} \times \cdots \times \mathbb{R}$ (n times)
$\mathbb{C}^n$	$\mathbb{C} \times \cdots \times \mathbb{C}$ (n times)
∞	infinity (note that $\infty \notin \mathbb{N} \cup \mathbb{Z} \cup \mathbb{R} \cup \mathbb{C}$; see [141].)
(a, b)	$\{x \in \mathbb{R} : a < x < b\}$
$(a, b]$	$\{x \in \mathbb{R} : a < x \leq b\}$
$[a, b)$	$\{x \in \mathbb{R} : a \leq x < b\}$
$[a, b]$	$\{x \in \mathbb{R} : a \leq x \leq b\}$
$\|x\|$	magnitude (absolute value) of $x \in \mathbb{C}$
$\bar{x}$	conjugate of $x \in \mathbb{C}$
$\mathrm{Re}\{x\}$	real part of $x \in \mathbb{C}$
$\mathrm{Im}\{x\}$	imaginary part of $x \in \mathbb{C}$
A^T	transpose of the matrix A
A^H	conjugate transpose of the matrix A (Hermitian operator)
$A \geq 0$	the matrix A is positive semi-definite
$A > 0$	the matrix A is positive definite
$\mathrm{diag}\{a_1, \cdots, a_n\}$	diagonal matrix containing $a_1, \cdots, a_n$ in its diagonal
ε_i	i^{th} vector of the canonical basis in $\mathbb{R}^n$, where n is understood from the context

z	argument of the Z-transform or forward shift operator, where the meaning is clear from the context
$A(z)$	transfer function in discrete time (scalar or multidimensional)
$\{A(z)\}\|_{z=0}$	$A(0)$
$A(\infty)$	$\lim_{z\to\infty} A(z)$
$A(z)^{\sim}$	shorthand for $A(z^{-1})^T$
$[A(z)]_{\mathcal{H}_2^\perp}$	strictly unstable part of $A(z) \in \mathcal{L}_2$
$[A(z)]_{\mathcal{H}_2}$	stable and strictly proper part of $A(z) \in \mathcal{L}_2$
$\|\|A(z)\|\|_2^2$	squared 2-norm of $A(z) \in \mathcal{L}_2$
$\otimes$	Kronecker product (see, e.g., [11, 19])
$\operatorname{vec}\{\cdot\}$	column stacking operator (see, e.g., [11, 19])
$\operatorname{vec}^{-1}\{\cdot\}$	inverse of the column stacking operator
$\mathcal{E}\{\cdot\}$	expectation operator
$\mathcal{E}\{\cdot\|\cdot\}$	conditional expectation
$\mu_x(k)$	mean at time k of the process x
$\sigma_x^2(k)$	variance at time k of the process x
μ_x	stationary mean of x, i.e., $\lim_{k\to\infty} \mu_x(k)$
σ_x^2	stationary variance of x, i.e., $\lim_{k\to\infty} \sigma_x^2(k)$
$\Omega_x(z)$	spectral factor of the process x
$S_x(z)$	power spectral density of the process x
x	shorthand for $\{x(k)\}_{k\in\mathbb{F}}$, with $\mathbb{F} = \mathbb{N}_0$ or $\mathbb{F} = \mathbb{Z}$
x^k	shorthand for $x(0), x(1), \cdots, x(k)$
x_i^k	shorthand for $x(i), x(i+1), \cdots, x(k)$
ln	natural logarithm
$\log_a$	logarithm in base a
inf	infimum
min	minimum
arg inf	argument that “infimizes” (see Section 2.2)
arg min	argument that minimizes

$\mathcal{R}$	set of all real rational transfer functions
$\mathcal{R}_p$	subset of $\mathcal{R}$ containing all proper transfer functions
$\mathcal{R}_{sp}$	subset of $\mathcal{R}$ containing all strictly proper transfer functions
$\mathcal{RH}_\infty$	subset of $\mathcal{R}_p$ containing all stable transfer functions
$\overline{\mathcal{RH}}_\infty$	subset of $\mathcal{R}_p$ containing all marginally stable transfer functions
$\mathcal{RH}_2$	subset of $\mathcal{R}_{sp}$ containing all stable transfer functions
$\mathcal{RH}_2^\perp$	subset of $\mathcal{R}$ containing all transfer functions that have only unstable poles (or no poles at all)
$\mathcal{U}_\infty$	subset of $\mathcal{RH}_\infty$ containing all square transfer functions with inverses in $\mathcal{RH}_\infty$
$\overline{\mathcal{U}}_\infty$	subset of $\mathcal{RH}_\infty$ containing all square transfer functions with inverses in $\overline{\mathcal{RH}}_\infty$
$\mathcal{L}_2$	subset of $\mathcal{R}$ containing all transfer functions whose frequency response is square integrable over the unit circle (see, e.g., [113, 182])

(The previous definitions are valid irrespective of the dimension of the transfer functions. If necessary, we will add an $n \times m$ superscript to refer to transfer functions of a specific size.)

$H(x)$	discrete entropy of (the discrete random variable) x
$H(x\|y)$	conditional discrete entropy of x given y
$h(x)$	differential entropy of (the continuous random variable) x
$h(x\|y)$	conditional differential entropy of x given y
$I(x;y)$	mutual information between x and y
$I(x;y\|z)$	conditional mutual information between x and y, given z
$D(x\|\|y)$	divergence between the distributions of x and y
$\bar{h}(x)$	entropy rate of the process x
$I_\infty(x;y)$	average scalar mutual information between the processes x and y
$I_\infty(x \to y)$	average directed mutual information between the processes x and y
$x \leftrightarrow y \leftrightarrow z$	x, y and z form a Markov chain (in that order)

Chapter 1

Introduction

1.1 Networked Control Systems

Control theory addresses the problem of how to feasibly manipulate a system so as to achieve a desired behavior [53]. A long standing assumption in standard control theory is that there exist transparent communication links between the *plant(s)*, i.e., the system(s) to be controlled, and the *controller(s)*, i.e., the system(s) which manipulate the plant(s) and are designed to achieve the desired behavior. This paradigm is suitable for applications where controllers and plants are physically close to one another, but is inappropriate when large scale systems are considered (see, e.g., [146]). For example, the high complexity and wide geographical distribution of power systems makes the assumption of (perfect) communication between all plants and controllers invalid (see, e.g., [66]). In these cases, *decentralized* architectures, i.e., control architectures where there exists a constrained communication pattern between the agents involved, are preferred over *full-MIMO* architectures where all agents are assumed to freely communicate with each other.[1]

As discussed above, communication constraints may manifest themselves in decentralized control architectures as the absence of communication links between the different agents in a control system. However, an issue that has been left aside from the standard control theoretic paradigm (including the standard decentralized control literature) is the fact that, even in the simplest SISO feedback loop, the communication channels may be far from transparent. Thus, communication constraints that have a nature that is essentially different from that of those

[1] It is worth pointing out that, in some cases, not only geographic distribution, but also reliability, cost, cabling and tuning issues make the implementation of full-MIMO control architectures impractical (see also [140, 144, 146, 184] and the many references therein).

arising in decentralized control theory arise. Such constraints originate in the characteristics of the *channels* and associated communication *protocols*, and may become *bottlenecks in the achievable performance*. Control systems where this happens are collectively referred to as *networked control systems* (see, e.g., [3, 4, 22]).

It is common to implement complex control systems over digital networks due to the many advantages that such technology provides (see, e.g., [96, 114, 124, 126]). In particular, standard communication technologies such as RS-485 or RS-232 (see [150,183]) and Ethernet (in its hard-wired and wireless versions; see, e.g., [89, 96, 126]) have been used in the context of networked control systems. Also, in different areas of control system applications, the necessity of standardization has led to several specialized protocols. For example, the CAN (controller area network) protocol emerged from the automotive industry, and Fieldbus has been developed for process control (see, e.g., [77, 96, 183]). The use of such communication systems to transmit control relevant information poses several challenges, as discussed below.

Irrespective of the actual nature of the physical links and communication protocols, if they are based on digital processing (as in most practical cases), then it becomes immediately clear that data needs to be *quantized* prior to transmission. That is, analog data from, for example, sensors needs to be converted into a digital word so that it can be processed and/or sent over a digital communication link. But the challenges do not stop here. Some communication protocols such as Ethernet, lack deterministic arbitration procedures to grant access to the network. This means that data may become *randomly delayed* or even *lost* (see, e.g., [96]). Data loss is also an issue in wireless communication systems. In those cases, environmental conditions may suddenly change and, as a consequence, the channel characteristics may change as well (see, e.g, [52]). This means that signals sent with, for example, a given power that was deemed satisfactory for a certain channel gain, may be too weak to be properly decoded at the receiver, due to the fact that the channel has changed since its gain was last estimated. In other cases, communication constraints arise due to the fact that the communication medium is shared amongst different users, i.e., the need of appropriate *scheduling policies* arise (see, e.g., [183]).

We thus conclude that the use of practical communication links leads inevitably to control problems where channel/network artifacts play a role. These effects include quantization, data-dropouts, random delays and scheduling issues. Key questions within this framework relate to the *interplay between the channel/network artifacts and the stability and performance of the associated control system.* A unifying framework for the analysis and design of general networked control systems does not yet exist. Nevertheless, there has been significant progress in recent years, as witnessed by the special issues [3, 4, 22] and by the handbook [77]. For example, quantization and data-rate constraints have been studied in, e.g., [21,34,42,44,50,71,81,116,118, 147,175,186]. The issue of data-dropout has been studied in, e.g., [41,78,82,85,86,100,148,154], whilst random time delays have been considered in, e.g., [95, 120, 121, 132, 172, 177]. Scheduling has been studied in, e.g., [54, 183, 187]. Good survey papers are [76, 118]. The former focuses on data-rate issues, and the latter on data-dropouts and delays.

Even though practical communication systems usually rely on digital signal processing, it is of theoretical importance to study networked situations where the communication channels (and the transmission technology) are assumed analog and the communication constraints arise, for example, as a limit on the transmission power (see, e.g., [18, 48, 55, 135, 159] and also [40]). This line of work is interesting because it enables one to grasp fundamental limitations and design issues in a simple framework. As such, this approach sheds light onto the explicit interplay between communication limits and control objectives. Very interesting work related to this framework has been documented in [105–107]. There, fundamental inequalities regarding information flow (in the standard information theoretic sense; see, e.g., [31, 108]) have been uncovered (see also [40]).

In the context of decentralized control of MIMO plants, networks can also play significant roles. Indeed, it is well known that constraining the controller architecture limits the achievable performance (see, e.g., [56, 91, 128, 140, 161, 169] and the references therein). This motivates the study of decentralized control architectures that are enriched with *additional, even though limited, communication links.* These additional communication channels may provide performance gains and help overcome the limitations that constraints on controller structure impose. For example, [80] showed that the set of plants that are stabilizable by decentralized architectures can be enlarged by means of appropriate communication resources usage. Also, [134] has provided evidence that, by exploiting the possibility of transmitting large packets (as made feasible by modern networks such as Ethernet), one can recover centralized performance in an architecture where different agents communicate over a network. Other relevant results can be found in [85, 109, 159, 189].

In this thesis we study two networked control problems: control with average data-rate constraints and control over analog erasure channels with i.i.d. dropout profiles. Instrumental to our studies is the consideration of additive noise channels with signal-to-noise ratio constraints. To put this thesis into context, we next give a brief overview of relevant literature on these subjects.

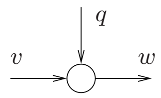

Figure 1.1: Additive i.i.d. noise channel.

1.1.1 Control over additive noise channels with signal-to-noise ratio constraints

The simplest model for an analog communication link is the additive white noise channel depicted in Figure 1.1. In that figure, q is an independent zero mean white noise sequence with a given distribution. Clearly, if the variance of q is fixed, then the effects of q can be made vanishingly small by increasing the variance of v to a sufficiently large value. Thus, in principle, an additive white noise channel presents no communications constraints. To properly model a physical transmission medium which is subject to communication constraints, one needs to add constraints on either q or v. It is customary to place a limit on the stationary variance of the signal to be transmitted, i.e., on the variance of v. This assumption, plus the additional assumption of q being Gaussian, yields the well-known additive white Gaussian noise channel model. This model is useful to model analog transmission media, where the power of the transmitted signal is constrained, e.g., due to technological or regulatory constraints (see, e.g, [31, 52, 174]).

The work [17, 18] considers a simple networked control situation where an additive white Gaussian noise channel is placed in the feedback path, signals are sent over the channel without any pre- or post-processing (i.e., no coding is in place), and the channel noise is the only external source in the feedback loop. For discrete time LTI SISO plant models,[2] and when considering a static state feedback law, [17,18] showed that a necessary and sufficient condition on the channel signal-to-noise ratio γ that allows one to find a controller that guarantees the stability of the resulting feedback loop is

$$\gamma > \left(\prod_{i=1}^{n_p} |p_i|^2 \right) - 1, \tag{1.1}$$

where $\{p_1, \cdots, p_{n_p}\}$ is the multi-set[3] of unstable plant poles. On the other hand, if dynamic LTI output feedback is employed, then the requirements on γ are strictly greater than those in (1.1) unless the plant is minimum-phase (MP) and has relative degree one [17, 18]. Thus,

[2] [17, 18] also consider continuous time models. We will focus only on the discrete time case in this thesis.
[3] Multi-sets are "sets" where multiplicity matters (see, e.g., [11]).

signal-to-noise ratio requirements for stability are, in general, more stringent in the dynamic output feedback case.[4] (We note, however, that we show in this thesis that if feedback from the channel output to the channel input is available, then no signal-to-noise ratio penalties arise in the dynamic feedback case. That is, the bound derived in [18] for the static state feedback case also applies to the dynamic output feedback case, when there is feedback around the channel.) The authors of [17, 18] also make a connection between their results and the results in [116]. In the next section we will further comment on this circle of ideas.

The work [49] considers the same setting as in [17, 18], but adds a process disturbance to the picture. It is shown in [49] that, if one considers general (possibly time varying and non-linear) coders and decoders, then condition (1.1) remains necessary for the mean square stability of the loop when state feedback is used. The results in [49] also give a lower bound on the plant state mean square norm. In turn, this shows that, due to the presence of process disturbances, the loop performance becomes arbitrarily poor if the signal-to-noise ratio approaches its minimal admissible value. This result is hardly surprising and is consistent with related results in, e.g., [116]. It is not clear, however, whether the performance bounds given in [49] are tight. For the case of noiseless MP plant models with relative degree one, it has also been shown in [49] that condition (1.1) is sufficient for a stabilizing LTI output feedback law that satisfies the channel power constraint to exist. These results leave open the question of how to construct stable feedback loops where the channel signal-to-noise ratio is arbitrarily close to (1.1), when disturbances are present. To answer that question, it becomes necessary to scale the signals sent through the channel, as discussed in [48] and in this thesis.[5] Also, structural issues that may influence the minimal signal-to-noise ratio for stability have not been explored yet in the literature.

In [17, 18], the authors use different techniques when dealing with either the static state feedback case, or the dynamic output feedback case. In [47], for a continuous time setting with no external noise sources besides the channel noise (which can be easily extended to the discrete time case; see [139]), the authors use LQG/LTR [171] ideas to unify both approaches. The motivation behind this is not only of technical interest. Indeed, it is easy to see that the sensitivity function associated with the static state feedback controller that achieves the minimal

---

[4]For the case of sampled data systems, [17, 18] showed that use of generalized hold functions allows one to circumvent the limitations that appear in the output feedback case. Nevertheless, this approach is not consistent with a purely discrete time framework.

[5]The work [48] was not available when most of the related results in this thesis were developed.

signal-to-noise ratio identified in [17, 18] is an all-pass function.[6] Therefore, the robustness and performance of the resulting closed loop are far from being satisfactory. Moreover, in the general case, it is cumbersome to arrive at closed form expressions for the optimal sensitivity functions (with respect to stabilization) using the techniques in [17,18]. This motivated the authors of [47] to modify the original problem of stabilization with a power constraint so as to take performance issues into account, and to be able to explicitly characterize the resulting sensitivity functions. To that end, the LQG/LTR framework is useful. Nevertheless, the results in [47] do not address the problem of what is the best performance achievable for a given signal-to-noise ratio constraint.

The work [137] (see also the associated thesis [135]) extends the results in [17,18] to the case of channels with colored noise and bandwidth limitations. It is proved that these constraints increase the signal-to-noise ratio required for stabilization. Another interesting extension has been presented in [136], where the interplay between minimal signal-to-noise ratio for stabilization and the out-of-band attenuation of the channel (i.e., the attenuation outside the assigned bandwidth for the channel) has been explicitly characterized for first and second order continuous time systems. As expected, stringent out-of-band attenuation requirements increase the signal-to-noise ratio requirements for stability. Other performance issues have been explored in [138] (see also Chapter 4 in [135]). In that work, the form of the optimal sensitivity function (with respect to stabilization) has been studied. Also, given any sensitivity function, the gap between the achieved signal-to-noise ratio and the minimal signal-to-noise ratio for stability is characterized as a means of studying additional signal-to-noise ratio requirements for performance. The problem with this approach is that it does not address fundamental performance questions such as what is the best sensitivity function (in some sense) for a given signal-to-noise ratio constraint. The results in [138] are then particularized to certain specific sensitivity functions that, for example, achieve a specified tracking error energy or frequency response when no communication constraints are present.

It follows from the previous discussion that the interplay between closed loop performance and signal-to-noise ratio constraints has not been fully explored. An exception to this is the recent work [48]. That work considers the setting discussed earlier, but assumes that there exist scaling factors at the input and output of the channel (see Figure 1.2). From both a theoretical and practical perspective, this is a very sensible assumption: the signal that is effectively transmitted over the channel may be a scaled version of the real signal of interest. (Note that, in this case, the power constraint is still on the channel input, i.e., on $\bar{v}$ in Figure 1.2.) In a discrete time setting

[6] The same applies to the output feedback case when the plant is MP and has relative degree one.

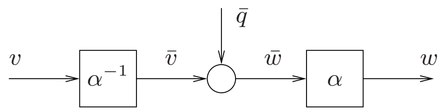

Figure 1.2: Additive i.i.d. noise channel with pre- and post-scaling.

where the output of a noisy plant is scaled and sent over the channel to be processed by an LTI controller collocated with the plant, [48] studies the best achievable performance (as measured by the plant output variance) and the corresponding optimal LTI controllers and scaling factors. To that end, standard LQG theory is employed. The problem becomes significantly simplified due to the fact that the power constraint is, basically, a constraint on the plant output variance, i.e., on the very signal whose variance is used to measure performance. (We will return to this point in Chapter 4 of this thesis.) It is worth mentioning that the results in [48] apply only to MP plants with relative degree one. It is also important to note that the study of quadratic optimal control problems with variance constraints is not a new subject (see, e.g., [13, 15, 103] and the many references therein). The new element here is the scaling factors at the channel input and output, which turn power constraints into signal-to-noise ratio constraints.

The results in [48] are interesting because they reveal the interplay between the scaling factor and the resulting best performance. Also, connections between best performance and channel capacity are established using the concept of entropy rate (see, e.g., [31]). It is also shown in [48] that the power constraint is active at the optimum, which is consistent with intuition. Even though the convexity of the associated optimization problems is implicitly exploited in the results of [48], its role is not as explicit as in, e.g., the approaches detailed in [15] for non-networked systems.

Form the previous discussion we can distill the following open problems:

(i) Is it possible, without leaving the realm of LTI architectures, to achieve stability at signal-to-noise ratios arbitrarily close to (1.1) for *arbitrary* plant models and *arbitrary* external noise sources? A positive answer to this question would imply, necessarily, a new element not considered in the work referred to above.

(ii) Is it possible to formulate a *general framework* to deal with performance issues and signal-to-noise ratio constraints? This should encompass arbitrary noise sources, any plant model and take architectural issues into account (e.g., coding and decoding schemes, feedback around channel, etc.).

(iii) How can one design coding schemes that mitigate the effects of signal-to-noise ratio constraints on closed loop performance, when a nominal controller design is already available? (Clearly, this is related to problem (ii) above.)

We have addressed some of the above questions in our work reported in [55, 159, 160, 162]. In Chapters 3 and 4 of this thesis, we present some of those results. We also and give explicit answers to all of the questions raised above.

1.1.2 Control with average data-rate limits

When dealing with practical digital networks, data needs to be quantized prior to transmission. Thus, quantization issues and data-rate limits arise as communication constraints. In some cases, the communication capacity of a given network may be sufficiently large so as to make quantization issues irrelevant (see, e.g., the work surveyed in [76, 148] and also [132]). However, there exist situations where the resources assigned to a particular control signal value or measurement are limited and, hence, quantization effects become important [118]. Also, from a purely theoretical perspective, quantization is a very interesting and challenging communication issue: it is highly non-linear, non-differentiable and, hence, hard to analyze [60].

In standard control theory, quantization has usually been treated as an undesirable effect that should be compensated for (see, e.g., [5, 110]). This stands in stark contrast to the perspective adopted by information theory, where quantization has always been considered as an integral part of systems (see, e.g., [31, 156]). In that area of research, quantization is seen as an *essential tool* for achieving digital communication over practical channels. Specifically, quantizers are seen as information encoders. This line of reasoning has recently been brought to the control theory arena (see, e.g., [116, 142, 143, 147, 175] and related work). It is worth noting that a different view of quantization, more closely tied to standard non-linear control theory than to information theory, has also been developed inside the control community (see, e.g., [21, 34, 42, 50, 88, 129] and many others).

We will now briefly expand on the last line of research mentioned above. Most of that work considers scenarios having no noise sources and where the only source of uncertainty is the system initial state. If one considers quantizers that are static and have a finite number of levels, then asymptotic stability can not be ensured in general [34]. The latter is a very important result and makes clear that the effects of quantization cannot be *rigourously* analyzed using simple additive noise models, as has been the usual practice in the signal processing literature (see, e.g, [83, 122]) and in early control literature focusing on the mitigation of quantization effects (see [32, 110, 111]). In these cases, one needs to abandon the notion of asymptotic convergence to a point and satisfy oneself with containability or so-called practical stability (see, e.g, [34, 42, 79, 186]). Nevertheless, if one (infinitely) increases the precision of the quantizer towards the origin, as in the case of logarithmic quantizers (see, e.g., [42, 50, 79]), then asymptotic stability can be ensured with static quantizers. This comes, of course, at the expense of having to use an infinite number of quantization levels. It has been proven in [42] that logarithmic quantizers are, indeed, the coarsest quantizers (in an appropriate sense) that allow one to achieve quadratic (asymptotic) stabilization of a single input LTI discrete time system. In order to design such a quantizer, the authors propose an LQR related problem and show that the resulting quantizer is a function of the unstable plant poles only. The case of finite logarithmic quantizers is also treated in [42]. An extension of the techniques employed in [42] to the MIMO case seems very difficult. This motivated [50] to explore alternative techniques to derive the results in [42]. Using the classic sector bound approach to model logarithmic quantizers, the authors of [50] were able to recover the results in [42] by means of an appropriately formulated $\mathcal{H}_\infty$ robust control problem. This indicates that, for SISO plant models, and when focusing on quadratic stabilizability, the sector bound approach is not conservative. The approach taken in [50] was then generalized to MIMO problems and performance related questions. In those cases, however, the notion of quantizer density becomes less clear and the sector bound approach seems to become conservative (see also [71]). An alternative viewpoint, where finite control input value constraints are explicitly imposed on a standard MPC formulation, is presented in [54, 129]. Within that framework, stability can be guaranteed only in certain cases and obtaining the control signal may become computationally expensive. Nevertheless, the latter approach is suitable for addressing performance related problems, and allows one to use the insights of standard MPC at the design stage. Indeed, except for the results in [50, 54, 129], useful performance issues have mostly been overlooked in the works described so far.

Many of the problems that appear when employing fixed level quantizers can be overcome if one uses dynamic and time-varying quantizers (see, e.g., [21, 115, 175, 186]). In those works, a finite number of levels in the quantizer is assumed, but the scaling of the quantizer is changed on-line so as to be able to properly accommodate the "size" of the values being quantized. With such a quantization scheme, it is possible to show that any asymptotically stabilizing static state feedback control law is still asymptotically stabilizing when the scaling policy of the quantizer is properly designed [21]. This scaling factor is usually referred to as a *zooming factor*. The results rely upon the fact that the zooming factor can be continuously varied. If the zooming factor is, instead, constrained to take some values in a finite set, then only practical stability is achieved [21].

The results surveyed so far, focus mainly on analyzing stability of control systems with quantized measurements or control values. They do not, however, explicitly study the minimal (average) data-rates (in, e.g., bits per sample) that allow one to achieve a given control objective. In the context of ideal digital channels, i.e., channels that transmit data without errors or delays at a constrained (average) rate, this question is closely related to so-called (noiseless) *rate distortion theory* (i.e., *source coding* problems; see, e.g., [10, 31, 157, 192]). The associated design problem is how to quantize, with the smallest (average) data-rate, whilst achieving a prescribed degree of performance, as measured by a suitable fidelity or distortion measure. A typical distortion measure is the mean square error, but it is also possible to define discrete measures. For example, the degree of fidelity may be defined to be zero when closed loop stability is achieved and very large otherwise (see, e.g., [174]). In this example, the design of schemes that minimize the fidelity measure subject to a data-rate constraint, amounts to finding strategies that stabilize with the given data-rate constraint. In this way, many *black-and-white* control questions can be possed as source coding problems. Of course, the problem is much more interesting if one focuses on performance questions.

In principle, one may be tempted to believe that standard information theory provides immediate *straight out of the box* answers to control problems with data-rate limitations. Unfortunately, this is not the case [118]. This is due to the fact that standard information theory (and in particular standard rate distortion theory) allows for coding using arbitrarily long sequences which incur arbitrarily long time delays. In addition, most of the general results on rate distortion theory do not take causality into account [10, 157]. It thus becomes clear that standard rate distortion theory is not useful to deal with control problems, where causality and delay issues are of fundamental importance. It is worth noting that some progress has been made towards a causal rate distortion theory, but the results are limited to special cases, or are valid only in the limit of high rates (see, e.g, [98, 119]). No general rate distortion theory for real time systems is currently available.

The discussion in the previous paragraph makes the work documented in [116, 175, 176, 186] specially relevant (even though the focus in that line of work lies on stability questions and not performance). The first results that pointed out that there exists a data-rate under which (memoryless) quantized control cannot keep the state of a noiseless plant bounded were presented in [7, 186]. The results were later refined in, e.g., [115, 175], where stronger notions of stability were considered. This was made possible thanks to the use of encoders and decoders with memory and zooming techniques, such as those mentioned before. A very important extension was published in [116], where the authors focus on noisy plant models subject to mild conditions on the noise sources. Within the latter setting, [116] shows that it is possible to find causal coders, decoders and controllers such that the resulting closed loop system is mean square stable (MSS) if and only if the *average* data-rate (in bits per sample), say $\bar{R}$, satisfies

$$\bar{R} > \sum_{i=1}^{n_p} \log_2 |p_i| , \tag{1.2}$$

where $\{p_1, \cdots, p_{n_p}\}$ denotes the multi-set of unstable plant poles. The above result holds for every coder, decoder and controller in the class of (time varying and non-linear) causal systems and, therefore, condition (1.2) establishes a fundamental separation line between what is achievable in networked control systems over digital channels and what is not (when the problem of interest is MSS). We note that bounds similar to (1.2) arise as solutions to a variety of different problems (e.g., observability, deterministic stability, etc.) and under different assumptions on the channels and coding schemes (see, e.g., [43, 75, 115, 116, 174, 175]). Indeed, the quantity on the right hand side of (1.2) is a fundamental measure of the difficulty of stabilizing a system, as explored in [117, 147]. This approach has been extended to performance questions in the case of noiseless plant models, as studied in [147].

Showing that the rate at which a stable control system is transmitting data must satisfy (1.2) (i.e., *necessity*) is fairly simple and employs more or less standard tools (see [118, 175] and also [40, 49, 105]). On the other hand, designing an actual coding scheme that achieves stability at any rate above the absolute limit (i.e., the proof of *sufficiency*) is much more involved (see [116,175]). All constructions known to date that actually achieve average data rates close to (1.2) use complex coding schemes that, in principle, have infinite memory. The consideration of coding schemes with limited (or no memory) is much more involved [118] and no simple explicit solutions are currently available (see, e.g., Section VI in [175]).

Almost all the works referred to in the previous two paragraphs, focus solely on stability issues. There exist, however, bounds on the mean square norm of the plant state that make explicit the fact that, as the data rate approaches the absolute minimum for stability in (1.2), the performance becomes arbitrarily poor when disturbances are present, no matter how the coder, decoder and controller are chosen (see [116, 118] and also [49]). It is, unfortunately, unclear whether these bounds are (in general) tight or not.

A more general performance-oriented approach has been pursued in [118, 176], where conditions for separation and certainty equivalence [8] were investigated in the context of quadratic stochastic problems with data-rate constraints. It has been proven in [118] that, if the encoder has a specific recursive structure, then certainty equivalence and a *quasi*-separation principle hold. That is, the encoder can be first designed for the uncontrolled process and, in a second stage, the controller can be designed using explicit knowledge of the encoder. (Thus, the controller depends on the encoder choice and hence quasi-separation and not strict separation holds.) This result is interesting, but it does not give an explicit characterization of the optimal encoding policy. This is also a drawback of the results reported in [176], where the trade-off between data-rate and performance is studied. Unfortunately, the results in [176] are expressed in terms of the, so-called, sequential rate distortion function, which is difficult to compute (except, perhaps, in the high-rate regime).

There exists a wealth of work dealing with the general problem of the interplay between channel capacity and closed loop control (see, e.g., [40, 49, 105, 174] and related work). When dealing with ideal digital channels, one essentially assumes that the channel is "simply a pipeline of bits". This, in turn, implies that, somehow, the problem of *channel coding* (which is complementary to that of source coding; see, e.g., [31, 51, 156]) has been solved. Of course, results such as that in (1.2) imply that the capacity of the underlying channel that is required to stabilize a system must be greater than the left hand side of (1.2). However, the catch in these conclusions is that the bound in (1.2) has been derived assuming error free transmission. In order to guarantee arbitrarily low probability of errors when transmitting digital data over a physical medium, one requires, necessarily, long delays or a very high capacity (see, e.g, [31]). This makes the bound in (1.2) too optimistic when interpreted as a capacity requirement. To summarize, if one focuses on ideal digital channels, then this is equivalent to focusing on the source coding problem only (i.e., quantization problems only). Channel coding is not an issue. This is, of course, a simplification which, nonetheless, yields interesting results that can be used as benchmarks in practical situations.

From the results surveyed above we conclude that the following questions remain open:

(i) It has been pointed out in [118] that optimal control problems that involve data-rate constraints are hard to formulate and solve. Moreover, it is clear that solutions to such problems involve the optimal design of quantizers and, in the most general case, would require the solution of the *causal rate-distortion problem*, which is an open problem [98,119]. It is thus relevant to explore contact points between control theory and information theory (in particular, rate distortion theory) so as to develop a framework that allows one to address control problems beyond those of stabilization when data-rate limits are present.

(ii) The results surveyed above usually rely on complex coding schemes with arbitrarily large memory. This raises the question as to whether it is possible to obtain explicit stability and performance related results when memoryless coding and decoding schemes are employed.

(iii) The work [18] has made an interesting connection between the minimal signal-to-noise ratio needed to stabilize a SISO LTI plant model over an additive white Gaussian noise channel, and the capacity of such a channel. It turns out that the minimal capacity for MSS equals the right hand side in (1.2). However, the settings used in [116] and [18] are different and the relationship between both results seems to be only incidental. (Recall the discussion regarding channel and source coding made above.) Hence, it would be interesting to investigate if, in a setup that considers error- and delay-free digital channels, there exists an actual relationship between signal-to-noise ratio related results[7] and (average) data-rate related results.

The above questions are addressed in Chapters 5 and 6 of this thesis.

1.1.3 Control with data-dropouts

A final networked control system issue considered in this thesis is that of data-dropouts. As pointed out before, there exist situations where the main communication constraint of importance is the existence of random data loss. In the simplest model of such a channel, data is either received uncorrupted at the receiving end with a given probability, or is completely lost. This leads to the notion of *erasure* channel (see, e.g., [76] and the many references therein). Moreover, it is often assumed that the channel is analog, i.e., that there is no need to quantize the data.[8] If the dropout probability can be assumed constant over time, then it is possible to model data dropouts as an independent sequence of i.i.d. Bernoulli random variables (see, e.g., [41, 148, 154]). Extensions to more general cases, where data dropouts are assumed to obey a Markov chain, can be found in, e.g., [99]. Other models for data dropouts (sometimes referred to as *deterministic models* [76]) are explored in, e.g., [73, 193].

Early work in this area includes [70, 120]. In [120], the effect of holding previously received data when data-dropout occurs, or estimating missing data based on previously received values

[7] This requires an appropriate definition of what is the signal-to-noise ratio we are interested in. It is not the (underlying physical) channel signal-to-noise ratio.

[8] Hence, the usual control-related erasure channels are *not* erasure channels in the usual information theoretic sense described in, e.g., [31].

is explored. On the other hand, [70] studies the effect of replacing missing data by zeros. Of course, one can do much better if properly designed dropout compensators are considered. For example, the work [100] (see also the thesis [99]) derives a closed form expression for the power spectral density (PSD) of the output of a simple closed loop when the feedback loop comprises an analog i.i.d. erasure channel and a dropout compensator. These expressions are then employed to design an LTI system in the dropout compensator so as to minimize the output variance. Key to this procedure is the observation that the resulting switched system (actually a Markov jump linear system (MJLS); see, e.g., [30]) is equivalent to a linear system with an external noise source that has a variance that depends on the variance of a signal within the loop. This allows one to pose the design problem as a quasi-convex optimization problem.

In some cases, the performance gains arising from mere data-dropout compensation may not be sufficient. In these cases, a complete controller re-design is needed. The work [154] uses ideas from MJLS theory to synthesize controllers that minimize an $\mathcal{H}_\infty$ functional in a networked situation where an i.i.d. erasure channel exists between sensors and controller. To that end, a bounded real lemma for MJLS (see, e.g., [33, 153, 158]) is employed to write the associated optimization problem as a semi-definite program (see also [15]). In this approach, the controller is assumed to have instantaneous and exact knowledge of the dropout process. A similar approach has been used in [82] in a more general context, where several channels are available for communication.

In [41], the author considers a general setup where multiple LTI controllers and LTI plants are connected over a *fading network* (which generalizes the idea of analog erasure channel). In that network model, the randomness of the channel is concentrated in a structured uncertainty block and, hence, the resulting feedback loop can be analyzed and designed using standard robust control theory (see, e.g., [167, 194]). The main focus in [41] is on a situation having no external noise sources and where, accordingly, stabilization is the main concern. It is proven that the (robust) MSS of the resulting feedback loop is equivalent to the feasibility of a linear matrix inequality (LMI). In a second step, the problem of designing controllers that maximize the robustness margin of the feedback loop is possed as a non-convex optimization problem that can be tackled using the well-known D-K iteration procedure (see, e.g., [167, 194]). Finally, [41] particularize their results to the case of a single analog i.i.d. erasure channel. In this case, the computation of the stability margin reduces to a standard $\mathcal{H}_2$ problem and it is proven that, in the perfect measured state feedback case,[9] and when the plant has one input, it is possible to

[9] Recall that no external sources are considered in [41].

find an LTI controller that achieves MSS if and only if the probability of successful transmission p satisfies

$$p > 1 - \frac{1}{\prod_{i=1}^{n_p} |p_i|^2}, \tag{1.3}$$

where, as before, $\{p_1, \cdots, p_{n_p}\}$ denotes the multi-set of unstable plant poles. Simple manipulation of this result suggests that, once an appropriate capacity-like notion for analog erasure channels is defined, the problem of stabilization over additive white Gaussian noise channels is equivalent to that of stabilization over analog i.i.d. erasure channels (see Section 8.2 in [41]).

It is also possible to address the problem of control over unreliable channels from the perspective of classic LQG theory (see, e.g., [6, 94]). This approach has been adopted in [78]. The authors of [78] consider a setup where independent analog i.i.d. erasure channels are present in both the link between sensors and controllers, and the link between controllers and actuators. No dropout compensation is assumed. Neither is any pre-processing of the sent signals utilized, i.e., state measurements and control values are sent directly over the channels and, if they are dropped, then lost data is replaced by zeros. The authors of [78] were one of the first to point out that there exist[10] problems where fundamental differences arise depending on whether there exist acknowledgements that testify successful transmissions or not. Protocols where such acknowledgements are available are referred to as *TCP-like* protocols, and when no acknowledgements are available the corresponding protocols are called *UDP-like* protocols. Within this setup, [78] shows that in the finite horizon noiseless LQ control problem, and when TCP-like protocols are employed, separation between estimation and control exists, and a closed form for the optimal estimators and control laws is given. For the infinite horizon LQG case, necessary and sufficient conditions that guarantee MSS are developed. These conditions reduce to a simple test on the spectral radius of the system structural matrix ("A" matrix) when the plant model has a square and invertible input to state matrix ("B" matrix). In the case of UDP-like protocols, the situation is much more involved since separation does not hold. Nevertheless, by assuming separation, [78] proposes a suboptimal controller. Sufficient conditions that guarantee MSS when the proposed scheme is used and the input to state matrix is invertible are given. However, these results involve coupled Riccati equations. Thus, the results do not give insights into the stabilization problem except in trivial cases.

The line of work in [78] has been extended in the work surveyed in [148] (see also related work by the same authors in, e.g., [164,165]). The setup in [148] is almost identical to that of [78], but

[10] Given the results in [68] surveyed below, this claim should be understood as valid for certain classes of problems only.

the former assumes that only noisy output measurements are available. The emphasis in [148] lies in making clear that separation holds only for the TCP-like protocol case, and that there exists a region in the plane of successful transmission probabilities within which stability can be guaranteed. These regions are protocol dependent and, as expected, the region associated with the TCP-like case contains that of UDP-like protocols. (It is also worth noting that these regions are consistent with the bound in (1.3) and the results in [78].) The results in [148] also make clear that LTI policies such as those in [41] are, in general, suboptimal. This is due to the fact that, even in the TCP-like case, the estimation gain is an explicit function of the dropout process and, as such, is time varying (and remains time varying for all times). In the UDP-like case, optimal control laws are shown to be non-linear in general. This motivates the study of optimal LTI LQG controllers as pursued in [166], where an iterative method for the computation of such controllers is proposed.

The work [68] extends the work detailed in [148] in several ways (but focuses on the case of communication constraints in the feedback path only). Firstly, the problem of interest is possed as an LQG problem where an encoder is present at the sending side of the channel and where, at the receiving end, the controller is preceded by a decoder. This gives additional flexibility. Secondly, a control (and coding-decoding) policy is proposed whose optimality holds *irrespective* of the actual channel dropout profile. Another important result in [68] relates to the fact that, for the setup used, separation between the estimation and control components of the cost function holds. Thus, it suffices to apply a standard LQ control law to the best mean square estimate of the plant state to arrive at the optimal control and estimation policy. These state estimates are constructed using a standard Kalman filter at the sending end (the encoder). At the receiving end, the decoder, which is assumed to know whether data was lost or not, uses a model of the plant to replace missing state measurements. It is proven that this simple policy actually achieves the best estimate of the state, thus providing a very simple solution to the original problem. For the case of Markov-chain governed data dropouts, [68] characterizes the stability of the resulting feedback loop in terms of the spectral radius of a specific matrix (much in the spirit of [120]). The results in [68] do not address, however, the question of convergence to a steady state solution, nor do they give an explicit characterization of the requirements on the dropout profiles that guarantee stabilization using the optimal control policy. The results in [68] are extended in [67] to the case of two channels in the feedback path.

In this thesis we focus on the simplest model for data dropouts. Namely, we focus on analog i.i.d. erasure channels. Within this context, and given the previous literature review, we have identified the following open questions:

(i) Is it possible to extend the results in [100], which relate spectral properties of a specific feedback architecture closed over an analog i.i.d. erasure channel to those of a simple LTI architecture, to more general situations? Indeed, the relationship between stabilization over additive white Gaussian noise channels and analog i.i.d. erasure channels pointed out in [41] suggests that the approach in [100] might be useful to give explicit conditions that guarantee MSS in an LTI framework. This leads naturally to the following additional question:

(ii) In what sense (if any, besides stabilization) are additive white Gaussian noise channels equivalent to analog i.i.d. erasure channels?

(iii) As pointed out in [148], optimal LQG policies are time varying and, in the UDP-like protocol case, non-linear. Is it possible to use *standard* tools to synthesize LTI control schemes that are optimal within a restricted class of control laws? Whilst the results in [166] certainly point in that direction, simpler characterizations of the solutions might be possible in some cases.

(iv) It is also interesting to investigate whether the conditions for MSS in the restricted LTI setup mentioned in (iii) above are more stringent than those in the general time varying case. The results in [41] imply that this is not the case in the noiseless state feedback case, but the general case has not been studied yet.

We address the questions raised above in Chapter 7 of this thesis.

1.2 Thesis Framework

In the previous section, we reviewed literature related to the material in this thesis and pointed out several open questions. We next describe the general setup and the main assumptions in this thesis. Specifically, we present the networked control architecture of interest and briefly comment on the channel models and the control and coding architectures to be used, and on the control problems to be addressed.

1.2.1 The networked control architecture

In this thesis we mainly focus on the control of discrete time LTI SISO plant models. Unlike standard non-networked control systems (see, e.g., [53, 167, 194]), we assume that the feedback

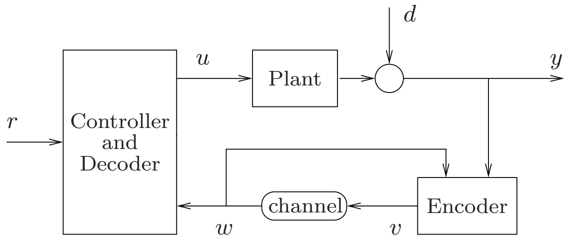

Figure 1.3: The networked architecture considered in this thesis.

path in the control system comprises a *non-transparent channel* and a *coding scheme*.[11] The coding scheme, which is a key and novel aspect of networked control, has two parts: the *encoder* and the *decoder*. The encoder is in charge of appropriately processing the signal to be sent over the channel so as to mitigate channel effects and/or to translate the measurements into symbols that the channel can understand. (This is the case of, e.g., digital channels where the symbols are binary words.) On the other hand, the decoder is in charge of translating the channel symbols back into the standard signal domain. Once the decoding operation is completed, the resulting signal can be used by the *controller* to generate the control signal. The setup described above is illustrated in Figure 1.3. We note that, in Figure 1.3, the controller and decoder have been merged into a single block labeled *controller-and-decoder*. (Indeed, it is possible to argue that a separation between the controller and the decoder is artificial and may not always yield the most general architecture.) In Figure 1.3, we also show a reference signal r, and a plant output disturbance d. As will become apparent as we proceed, the inclusion of additional disturbance signals and measurement noise does not pose any additional difficulties to our framework. It is also worth mentioning that we will assume that there exists (one-step delayed) *feedback from the output of the channel to its input*. This feature gives additional flexibility to accommodate control objectives, and will be shown to be of fundamental importance in some cases. We will also examine cases where such feedback is not available.[12]

Throughout this thesis we will (mostly) assume that the following holds:[13]

Assumption 1.1 (Plant, initial states and signals)

[11] In practice, it is usually the case that both the link between sensors and controller, and between controller and actuators, are non-transparent. Nevertheless, the study of a simple architecture where only one channel is limited allows one to unveil fundamental limitations and issues in networked control systems. (This assumption is common in the literature; see, e.g., [18, 76, 118].)

[12] When feedback around the channel is available, then the encoder structure in Figure 1.3 is akin to that of a two-degree-of-freedom controller (also called two parameter controller) (see, e.g., [182]).

[13] To avoid unnecessary page-turning, we will repeat (or slightly modify) these assumptions where appropriate.

(a) The plant is a discrete time SISO LTI system, having a strictly proper and non-zero transfer function, and a detectable and stabilizable underlying realization.

(b) The signals r and d are second order[14] *mutually uncorrelated wss scalar processes with rational power spectral densities (PSDs). At least one signal, r or d, is non-zero. If r or d are non-zero, then their PSDs are non-zero for all frequencies.*

(c) The initial state of the plant and of other dynamic systems in Figure 1.3 are jointly second order random variables.

We assume the plant to be SISO and LTI for simplicity. In principle, it would be possible to extend many of our results to the MIMO case at the expense of additional technicalities.[15] The remaining assumptions on the plant are mild and are satisfied by most plants of practical significance (see, e.g., [53, 167, 194]). Most of the assumptions regarding the external signals r and d are standard in the classical stochastic control framework (see, e.g., [6, 61, 94, 168]). One assumption that needs further comment refers to the fact that we do not allow both r and d to be simultaneously zero sequences. This is done to simplify the analysis at some points, but is not an essential assumption. Nevertheless, we believe that some external signal will always be present in practice and, therefore, assuming either r or d non-zero is appropriate. The assumption of r and d having a PSD that, if not identically zero, is non-zero at all frequencies will be justified later (see Remark 2.3 in Chapter 2). Finally, the assumptions regarding the initial states are mild and standard.

1.2.2 The channels

There exist many useful channel models available in the literature (see, e.g., [31,148,174,179]). As is true for all models, they tend to highlight only some particular characteristics and constraints of the physical communication medium and communication protocol. In this thesis we will focus on the following three models (which will be formally defined in the corresponding chapters):

(i) **Additive white (or i.i.d.) noise channel with signal-to-noise ratio constraint:** This model comprises an additive white noise channel where the noise variance is fixed, a power constraint on the channel input is in place, and where pre- and post-scaling factors are used (see Figure 1.2). It is easy to see that this model amounts to the situation in

[14] Nomenclature associated with random processes is presented in Chapter 2.

[15] A simple MIMO extension of part of our results is provided in Chapter 3.

Figure 1.1, if q is a zero mean white noise sequence whose *variance is a decision variable*,[16] and where the stationary variance constraint on $\bar{v}$ translates into a constraint on the ratio between the stationary variance of v and the variance of q, i.e., where communication constraints arise as stationary *signal-to-noise ratio constraints*. In some cases, we will require stronger assumptions on q. In particular, we will sometimes assume that q is not only white, but a sequence of i.i.d. random variables.

(ii) **Ideal digital channel with average data-rate constraint:** In this channel model, data is transmitted instantly, and with no errors, from the sending end to the receiving end. The communication constraints arise due to the fact that the channel has a restricted alphabet (i.e., can transmit only symbols from a countable set), and that there exists a bound on the *average* length (in bits or nats[17]) of the transmitted symbols.[18] When dealing with practical channels, bandwidths are limited and thus bit-rates need to be instantaneously limited. Accordingly, and as mentioned before, the consideration of average data-rate limited channels is only a simplification that, nevertheless, allows one to establish bounds on what can be achieved in practice and what cannot.

(iii) **Analog i.i.d. erasure channel:** In this model, data is either received uncorrupted with a given probability (that is assumed to be constant over time), or is completely lost. It is important to emphasize that this channel model is analog in that it is able to transmit any real number with no need to quantize the data.

1.2.3 The control and coding architecture

Having described the architecture and channel models considered, we now turn to a brief review of the control and coding architectures that will be presented in the forthcoming chapters. Unlike most of the literature, our interest lies primarily in simple schemes that use (in most cases) only LTI filters. It is clear that this choice may constrain the performance of the resulting networked control systems. Nonetheless, it opens the door to use standard linear control theory to analyze and design networked control systems in an exact manner (i.e., without approximations). Also, surprisingly, our results show that the constraints we impose on the control and coding architecture are not always a limiting factor.

[16]In Figure 1.2 the variance of $q \triangleq w - v$ depends on the choice of the scaling factor α and on the (fixed) variance of the underlying noise source $\bar{q}$. Thus, changing α amounts to changing the variance of q.

[17]Nats arise when one uses basis e instead of basis 2 to represent digital words ($1[\text{bit}] = \ln 2[\text{nat}]$).

[18]In such a channel, the instantaneous length of the transmitted symbols may be unbounded.

1.2.4 The control problems

In this thesis, we not only address the problem of networked control system stabilization, but we also tackle performance related questions. In particular, we adopt the notion of mean square stability, as defined in, e.g., [29, 30, 84]. When dealing with performance questions, we focus on the tracking error e, defined via (see Figure 1.3)

$$e \triangleq r - y, \tag{1.4}$$

and use its stationary variance as performance measure. (As the reader will certainly notice, our results can be straightforwardly extended to other weighted quadratic performance measures.)

1.3 Overall Aim of the Thesis

This thesis focuses on studying the *interplay between communication constraints and the performance and stability* of the networked control situation of Figure 1.3, when the channel is one of the channels described in Section 1.2.2. To that end, we aim at *establishing links* between networked situations that use ideal digital channels with average data-rate constraints, or analog i.i.d. erasure channels, and networked situations that employ additive noise channels with signal-to-noise ratio constraints.

In this thesis we also focus on *exploring to what extent simple LTI control and coding schemes provide a means to approach (or achieve) some of the known fundamental limits in networked control systems*. As already mentioned in Section 1.1, the latter usually rely on complex[19] control and coding architectures. In doing so, we *bring together ideas from control theory and information theory*, and we also *adapt standard control theory* methodologies to deal with the challenges that networked control systems pose.

1.4 Overview of Thesis Contents and Contributions

This thesis is organized in eight chapters (including the present one) and three appendices. The most important contribution of this thesis is to propose *a framework to study control problems with either signal-to-noise ratio constraints, or average data-rate constraints, or i.i.d. data-dropouts, in an unified manner*. Our framework is not only useful for the analysis of such

[19] At least time varying.

control systems, but also yields design procedures that use standard control system design tools. In particular, the content and contributions of each chapter can be summarized as follows:[20]

- **Chapter 2:** This is a review chapter where the notation and conventions used throughout the thesis are presented. We also present known results regarding the stability of feedback systems and MSS, which are used repeatedly in the remaining chapters.

- **Chapter 3:** This chapter deals with a basic networked control system setup, where a pre-designed controller has to be implemented using a signal-to-noise ratio constrained additive white noise channel. To mitigate communication effects, we propose a feedback-based perfect reconstruction coding architecture, and analyze the interplay between a channel signal-to-noise ratio constraint and both MSS and performance of the resulting networked control system. When focusing on MSS, our results show that the use of feedback allows one to reduce (in some cases, drastically) signal-to-noise ratio requirements as compared to an architecture that does not use feedback based coding. Regarding performance questions, we provide a methodology to design optimal coding schemes in an iterative fashion.

  This chapter also extends the above setting to the control of MIMO plant models. In that case, we focus on situations where non-transparent communication channels are used as additional channels which are aimed at counteracting the performance limitations arising from decentralized architectures. This leads to a control architecture that we call *partly networked*.

  The main contributions of this chapter are the following:

  (i) An explicit and analytic characterization of the minimal signal-to-noise ratio needed to achieve MSS for the given architecture is obtained. This characterization is given in terms of unstable plant and controller poles only. The role of feedback in the coding scheme is made explicit.

  (ii) An iterative procedure to design the filters in the proposed coding scheme is developed. This method is based on convex optimization and gives optimal performance for a given signal-to-noise ratio constraint.

  (iii) Regarding the MIMO case, we provide a methodology to design a partly networked control architecture having performance in mind. As a by-product, we show that, for the given architecture, it is sometimes not advisable to use the additional communication channels, unless they are extremely reliable. In other words, we provide

[20] Chapter 2 and the appendices present known results and, accordingly, no contributions.

evidence that suggests that, in some cases, *no information is better than corrupted information.*

- **Chapter 4:** In this chapter we extend the results established in the first part of Chapter 3 to general control and coding architectures for SISO plant models. We consider the most general architecture that is possible to build around a signal-to-noise ratio constrained additive white noise channel, where feedback is available from its output. Within this setting, we study signal-to-noise ratio requirements for MSS and the interplay between performance and signal-to-noise ratio constraints. Our results make explicit the role that design degrees of freedom play in networked control systems. The results in this chapter serve as a backbone for the rest of the thesis.

  The main contributions of this chapter are as follows:

  (i) An explicit and analytic characterization is given for the minimal signal-to-noise ratio needed to achieve MSS in the proposed architecture. This characterization is stated in terms of unstable plant poles only. If structural constraints are imposed on the architecture (e.g., no channel feedback is available, etc.), then other plant features also play a role and we provide an exact characterization of their effects.

  (ii) A general framework is developed to study the interplay between performance and signal-to-noise ratio constraints.

  (iii) An iterative procedure is given, based on convex optimization, to design the filters in the proposed control and coding scheme. This procedure allows one to optimize performance for a given signal-to-noise ratio constraint, and also enables one to find the minimal signal-to-noise ratio needed to attain a specified level of performance.

- **Chapter 5:** This chapter is intended to provide a link between information theory and control theory. Our aim is to show that it is possible to use fundamental concepts such as differential entropy and (directed) mutual information to arrive at a framework that is suitable for analyzing and designing control systems where the channels have an average data-rate constraint. To that end, we begin by providing bounds on average data-rates that are valid in a general setup. We then specialize these results to the case of *i.i.d. source coding schemes* (a notion defined in this thesis). Within this setup, it is shown that it is possible to bound, from above, the average data-rate across the channel by the signal-to-noise ratio of an associated additive i.i.d. noise channel. This reveals the important fact that the tools and insights developed in Chapter 4 are actually useful to deal, in an exact manner, with average data-rate constraints. Finally, we discuss what an entropy coded dithered quantizer is. Chapter 5 prepares the ground for the results of Chapter 6.

The main contributions of this chapter are the following:

(i) We establish lower bounds on average data-rates. These bounds depend explicitly on the information (i.e., the data) available at the encoder and decoder.

(ii) For i.i.d. source coding schemes, we show that the lower bounds on average data-rates established in (i) can be bounded by (and sometimes written in terms of) the stationary spectral characteristics of the input and output of the coding scheme. This result makes it possible to establish a simple link between information theory (in particular, rate distortion theory) and standard control theory.

(iii) We show that there exists a formal relationship between analog additive i.i.d. channels and channels with average data-rate limits. This allows us to explain the relationship between the result in [17, 18] and those in [116].

(iv) We slightly extend known results regarding entropy coded dithered quantizers to the case where there is feedback around the quantizer.

- **Chapter 6:** This chapter uses the results in Chapter 5 to deal with control systems which are closed over ideal digital channels with average data-rate constraints. We first consider MSS and show that by using control and coding architectures formed by LTI filters and i.i.d. source coding schemes, it is possible to achieve MSS at average data-rates that are close to the lowest bounds provided in [116]; see (1.2). Our results make explicit the role that degrees of freedom have in the design process, and also the role played by the memory of the i.i.d. source coding scheme. We end this chapter by providing guaranteed upper bounds on the minimal average data-rate that allow one to attain a given performance level. In all our results, entropy coded dithered quantizers play an important role as a means to get close (and in some cases achieve) the average data-rates that our theory promises.

  The main contributions of this chapter can be summarized as follows:

  (i) We characterize lower bounds on the average data-rate needed to achieve MSS when an i.i.d. source coding scheme is employed and suitable Gaussianity assumptions are imposed. These bounds reveal the role played by the information available at the encoder and decoder, and also the interplay between that information and the degrees of freedom provided by the LTI filters in the proposed architecture.

(ii) Using an entropy coded dithered quantizer as a building block, we provide guaranteed upper bounds on the average data-rate needed to achieve MSS with the proposed control and coding architecture. We show that the gap between our bounds and the bounds in [116] (see (1.2)) is composed of two terms: the first arises due to the inefficiency of practical entropy coders, and the second arises due to the fact that the quantization noise generated by entropy coded dithered quantizers is not Gaussian, but uniform.

(iii) We give an upper bound on the average data-rate needed to achieve a certain level of performance with the proposed control and coding architecture.

- **Chapter 7:** In this chapter we show that there exists a *second-order moments equivalence* between a linear system interconnected over an analog i.i.d. erasure channel, and the same linear system interconnected over a signal-to-noise ratio constrained additive i.i.d. noise channel. In the study of control systems with i.i.d. data-dropouts, this equivalence plays a role that is akin to the role that the results of Chapter 5 played in Chapter 6, i.e., it provides a *bridge between standard control theory and networked control theory.* Once this key equivalence is revealed, we use it to analyze and design control systems where the feedback path uses an analog i.i.d. erasure channel. In particular, we characterize bounds on the channel dropout probability that ensure MSS. Although we employ only LTI coding and control schemes, we recover the bounds found in the literature for which linear time varying schemes have been proposed. Finally, we present a convex-optimization-based methodology to optimize performance in networked control systems closed over analog i.i.d. erasure channels.

  The main contributions of this chapter are the following:

  (i) We show that the feedback interconnection of an LTI system over an analog i.i.d. erasure channel is MSS if and only if the same LTI system, when interconnected over a signal-to-noise ratio constrained additive i.i.d. noise channel, is asymptotically stable and an inequality constraint on the successful transmission probability is satisfied. This result, as well as the one detailed in (ii) below, extends the results in [100].

  (ii) We show that the signals in a general feedback loop, where LTI systems are interconnected over an analog i.i.d. erasure channel, exhibit the same PSDs as those of the corresponding signals in the feedback loop that arises when the erasure channel is replaced by a signal-to-noise ratio constrained additive i.i.d. noise channel, plus a gain.

(iii) For a general control and coding architecture that uses only LTI filters, we establish a bound on the successful transmission probability that allows one to guarantee MSS. When TCP-like protocols are used, this bound is given in terms of the unstable plant poles only. When TCP-like protocols cannot be used, then our results show that additional plant features play a role.

(iv) We propose a procedure based on convex optimization to design a simplified control and coding architecture so as to optimize performance for a given successful transmission probability.

- **Chapter 8:** This chapter wraps up and summarizes the thesis. We also discuss directions for future research.
- **Appendices:** Appendix A presents background on analytic $\mathcal{H}_2$ optimization techniques. These results are used to characterize explicit solutions to the optimization problems that are encountered throughout the thesis. Appendix B discusses the standard additive noise model for uniform quantization, which is employed in some of the simulation studies in Chapters 3 and 4. Finally, Appendix C provides a succinct review of basic definitions in information theory. These facts and definitions are employed in Chapter 5.

1.5 Associated Publications

The following is a list of publications by the author (or with significant contributions by the author):

- **Refereed journals**

  (i) Directly related to the content of this thesis:

  1. 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, January 2008.
  2. **E.I. Silva**, G.C. Goodwin, and D.E. Quevedo. A framework for the design of controllers and coders in SNR constrained networked systems. Submitted to Automatica, November 2008.

  (ii) Indirectly related to the content of this thesis:

3. D.E. Quevedo, **E.I. Silva** and G.C. Goodwin: "Control over Unreliable Networks having Packet Loss and Variable Transmission Delay". IEEE Journal on Selected Areas in Communications, Vol. 26, No.4, pp. 672-685, May 2008.
4. D.E. Quevedo, **E.I. Silva** and G.C. Goodwin: "Subband Coding for Networked Control Systems". To appear in the International Journal of Robust and Nonlinear Control, 2008.

(iii) Other publications by the author having no relationship with this thesis:

5. **E.I. Silva** and M.E. Salgado: "Performance bounds for feedback control of nonminimum-phase MIMO systems with arbitrary delay structure". IEE Proceedings - Control Theory and Applications - Vol. 152, No. 2, pp. 211-219, March 2005.
6. G.C. Goodwin, M.E. Salgado and **E.I. Silva**: "Time-domain performance limitations arising from decentralized architectures and their relationship to the RGA". International Journal of Control, Vol. 78, No. 13, pp. 1045-1062, September 2005.
7. M.E. Salgado and **E.I. Silva**: "Robustness issues in $\mathcal{H}_2$ optimal control of unstable plants". Systems and Control Letters, Vol. 55, No. 2, pp. 124-131, February 2006.
8. M.E. Salgado and **E.I. Silva**: "Achievable performance limitations for SISO plants with pole location constraints". International Journal of Control, Vol. 79, No. 3, pp. 263277, March 2006.
9. **E.I. Silva**, D.A. Oyarzún and M.E. Salgado: "On structurally constrained $\mathcal{H}_2$ performance bounds for stable MIMO plant models". IET Control Theory & Applications, Vol. 1, No. 4, pp. 1033–1045, July 2007.
10. M.E. Salgado, D.A. Oyarzún and **E.I. Silva**: "$\mathcal{H}_2$ Optimal Ripple-Free Deadbeat Controller Design". Automatica, Vol. 43, No. 11, pp. 1961–1967, November 2007.
11. M.S. Derpich, **E.I. Silva**, D.E. Quevedo and G.C. Goodwin: "On Optimal Perfect Reconstruction Feedback Quantizers". IEEE Transaction on Signal Processing, Vol. 56, No. 8 (part 2), pp. 3871-3890, August 2008.
12. A.A. Peters, M.E. Salgado and **E.I. Silva**: "Performance Bounds in Linear Control of Unstable MIMO Systems with Pole Location Constraint". Systems and Control Letters, Vol. 57, No. 5, pp. 392–399, May 2008.

- **Refereed conferences**

  (i) Directly related to the content of this thesis:

  1. G.C. Goodwin, D.E. Quevedo and **E.I. Silva**: "An Introduction to Networked Control Systems". Proceedings of the 6th Asian Control Conference (ASCC), Bali, Indonesia, 2006.
  2. **E.I. Silva**, G.C. Goodwin, D.E. Quevedo and M.S. Derpich: "Optimal Noise Shaping for Networked Control Systems". Proceedings of the European Control Conference (ECC), Kos, Greece, 2007.
  3. **E.I. Silva**, D.E. Quevedo and G.C. Goodwin: "Optimal Coding for Bit-Rate Limited Networked Control Systems in the Presence of Data Loss". Proceedings of the 46th IEEE Conference on Decision and Control (CDC), New Orleans, USA, 2007.
  4. G.C. Goodwin, **E.I. Silva** and D.E. Quevedo: "A Brief Introduction to the Analysis and Design of Networked Control Systems". Plenary address at the Chinese Control and Decision Conference, China, 2008.
  5. **E.I. Silva**, D.E. Quevedo and G.C. Goodwin: "Optimal Controller Design for Networked Control Systems". Proceedings of the 17th IFAC World Congress, Seoul, Korea, 2008.
  6. **E.I. Silva**, G.C. Goodwin and D.E. Quevedo: "On Networked Control Architectures for MIMO Plants". Proceedings of the 17th IFAC World Congress, Seoul, Korea, 2008.
  7. **E.I. Silva**, M.S. Derpich, J. Østergaard and D.E. Quevedo: "Simple Coding for Achieving Mean Square Stability over Bit-rate Limited Channels". Accepted for presentation at the 47th IEEE Conference on Decision and Control (CDC), Cancún, México, 2008.

  (ii) Indirectly related to the content of this thesis:

  8. G.C. Goodwin, D.E. Quevedo and **E.I. Silva**: "Filter Banks in Networked Control". Proceedings of the 17th International Symposium on Mathematical Theory of Networks and Systems (MTNS), Kyoto, Japan, 2006.
  9. D.E. Quevedo, **E.I. Silva** and G.C. Goodwin: "Packetized Predictive Control over Erasure Channels". Proceedings of the 26th American Control Conference (ACC), New York, USA, 2007.

10. D.E. Quevedo, **E.I. Silva** and G.C. Goodwin: "Performance Limits in Multi-Channel Networked Control System Architectures". Proceedings of the 3rd IFAC Symposium On System, Structure And Control (SSSC), Brazil, 2007.
11. D.E. Quevedo, **E.I. Silva** and D. Nešić: "Design of Multiple Actuator-Link Control Systems with Packet Dropouts". Proceedings of the 17th IFAC World Congress, Seoul, Korea, 2008.

(iii) Other publications by the author having no relationship with this thesis:

12. D.A. Oyarzún and **E.I. Silva**: "Effect of downstream feedback on the achievable performance of feedback control loops for serial processes". Proceedings of the European Control Conference (ECC), Kos, Greece, 2007.
13. **E.I. Silva**, B.P. McGrath, D.E. Quevedo and G.C. Goodwin: "Predictive Control of a Flying Capacitor Converter". Proceedings of the 26th American Control Conference (ACC), New York, USA, 2007.
14. **E.I. Silva**, D.E. Quevedo and G.C. Goodwin: "Optimal Multibit Digital to Analog Conversion". Proceedings of the 7th International Symposium on Communications and Information Technologies, Sydney, Australia, 2007.
15. V. Wertz, **E.I. Silva**, G.C. Goodwin and B. Codrons: "Performance Limitations arising in the Control of Power Plants". Proceedings of the 17th IFAC World Congress, Seoul, Korea, 2008.

- **Internal Reports**

1. **E.I. Silva**, D.E. Quevedo, M.S. Derpich and G.C. Goodwin: "Design of Feedback Quantizers for Networked Control Systems". Unpublished report available at `http://altair.elo.utfsm.cl/public_elo/uploads/pdf/Silva_FBQ_TR.pdf`

Chapter 2

Notation and Preliminaries

2.1 Introduction

The purpose of this chapter is to present definitions, conventions, and basic results that are used repeatedly throughout the thesis. Definitions and known results whose use is restricted to proofs or only to certain chapters, are compiled as separate appendices at the end of the thesis and are referred to where appropriate. An abbreviated list of symbols and brief standard definitions is included at the beginning of the thesis for ease of reference.

2.2 Transfer Functions

In this thesis we will consider discrete-time linear time invariant (LTI) systems with real rational transfer functions.[1] We denote the set of all real rational transfer functions by $\mathcal{R}$. We will use z as the forward shift operator or as the argument of the Z-transform, where the meaning is clear from the context. We define (finite and infinite) poles and zeros of transfer functions as usual (see, e.g., [87, 194]).

We say that $A(z) \in \mathcal{R}$ is proper (i.e., causal) if and only if $\lim_{z\to\infty} A(z)$ exists and is finite; we denote the subset of $\mathcal{R}$ that contains all proper transfer functions by $\mathcal{R}_p$. We say that $A(z) \in \mathcal{R}$ is strictly proper (i.e., strictly causal) if and only if $\lim_{z\to\infty} A(z) = 0$; the subset of $\mathcal{R}$ containing all strictly proper transfer functions is denoted by $\mathcal{R}_{sp}$. We say that $A(z) \in \mathcal{R}$ is biproper if and only if $\lim_{z\to\infty} A(z)$ is finite and non-singular.

[1] The term transfer function also applies to transfer matrices, i.e., multidimensional transfer functions.

A transfer function $A(z) \in \mathcal{R}_p$ is said to be *unstable* if and only if it has poles outside the unit disk or on the unit circle (i.e., in $\{z \in \mathbb{C} : |z| \geq 1\}$); it is said to be *stable* if and only if all its poles lie strictly inside the unit disk (i.e., in $\{z \in \mathbb{C} : |z| < 1\}$); it is said to be *marginally-stable* if and only if all its poles are inside the unit disk or on the unit circle (i.e., in $\{z \in \mathbb{C} : |z| \leq 1\}$). By extension, a pole p is stable if and only if $|p| < 1$, unstable if and only if $|p| \geq 1$ (or $p = \infty$), marginally-stable if and only if $|p| = 1$, and strictly unstable if and only if $|p| > 1$ (or $p = \infty$).

A transfer function $A(z) \in \mathcal{R}_p$ is said to be *non-minimum phase* (NMP) if and only if it has finite zeros outside the unit disk or on the unit circle; it is said to be *minimum phase* (MP) if and only if all its finite zeros lie strictly inside the unit disk; it is said to be *marginally-MP* if and only if all its finite zeros are inside the unit disk or on the unit circle. By extension, a zero c is MP if and only if $|c| < 1$, NMP if and only if $|c| \geq 1$ (or $c = \infty$), marginally-MP if and only if $|c| = 1$, and strictly-NMP if and only if $|c| > 1$ (or $c = \infty$).

For future reference, we define the following sets:

$$
\begin{aligned}
\mathcal{RH}_2 &\triangleq \{A(z) \in \mathcal{R}_{sp} : A(z) \text{ is stable}\}, \\
\mathcal{RH}_2^\perp &\triangleq \{A(z) \in \mathcal{R} : A(z) \text{ is constant or has only strictly unstable poles}\}, \\
\mathcal{RH}_\infty &\triangleq \{A(z) \in \mathcal{R}_p : A(z) \text{ is stable}\}, \\
\overline{\mathcal{RH}}_\infty &\triangleq \{A(z) \in \mathcal{R}_p : A(z) \text{ is marginally stable}\}, \\
\mathcal{U}_\infty &\triangleq \big\{A(z) \in \mathcal{RH}_\infty : A(z) \text{ is square, non-singular almost everywhere,}^2 \\
&\qquad\qquad\qquad\qquad\qquad\qquad \text{and } A(z)^{-1} \in \mathcal{RH}_\infty\big\}, \\
\overline{\mathcal{U}}_\infty &\triangleq \{A(z) \in \mathcal{RH}_\infty : A(z) \text{ is square, non-singular almost everywhere,} \\
&\qquad\qquad\qquad\qquad\qquad\qquad \text{and } A(z)^{-1} \in \overline{\mathcal{RH}}_\infty\big\}.
\end{aligned}
$$

If the dimension of a transfer function is important, then we will add a $n \times m$ subscript to any of the previously defined symbols.

Any $A(z) \in \mathcal{R}$ with no marginally stable poles belongs to $\mathcal{L}_2$,[3] i.e., is such that

$$
||A(z)||_2^2 \triangleq \text{trace}\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} A(e^{j\omega})A(e^{j\omega})^H d\omega\right\} < \infty. \tag{2.1}
$$

$||\cdot||_2$ corresponds to the norm in $\mathcal{L}_2$ (also called $2-$norm; see [46, 113, 182, 194] and also Appendix A). We also note that any $A(z) \in \mathcal{R}$ with no marginally stable poles can be written

[2] i.e., non-singular for almost every $z \in \mathbb{C}$.

[3] This is an abuse of notation. Formally, the function that arises when restricting attention to $z = e^{j\omega}$ is the one that belongs to $\mathcal{L}_2$.

as

$$A(z) = [A(z)]_{\mathcal{H}_2^\perp} + [A(z)]_{\mathcal{H}_2}\,, \tag{2.2}$$

where $[A(z)]_{\mathcal{H}_2^\perp} \in \mathcal{RH}_2^\perp$ contains the strictly unstable, constant and non-causal parts of $A(z)$, and where $[A(z)]_{\mathcal{H}_2} \in \mathcal{RH}_2$ corresponds to the stable and strictly proper part of $A(z)$.

Throughout this thesis, plant transfer functions are always denoted by $G(z)$ (except in Section 2.4). We denote the multi-set[4] of strictly unstable plant poles via

$$\left\{p_1, \cdots, p_{n_p}\right\}_m, \tag{2.3}$$

and the multi-set of marginally stable poles via

$$\left\{p_{n_p+1}, \cdots, p_{n_p+\bar{n}_p}\right\}_m. \tag{2.4}$$

(if $\left\{p_1, \cdots, p_{n_p}\right\}_m$ is empty [resp. $\left\{p_{n_p+1}, \cdots, p_{n_p+\bar{n}_p}\right\}_m$ is empty], then $n_p = 0$ [resp. $\bar{n}_p = 0$]). The number of occurrences of a pole p_i in any of the above defined multi-sets equals the algebraic multiplicity of the pole (see, e.g., [87]). If the plant $G(z)$ is SISO,[5] we define the products (see also, e.g., [155]):[6]

$$\xi_p(z) \triangleq \prod_{i=1}^{n_p} \frac{1 - z p_i}{z - p_i}, \quad \bar{\xi}_p(z) \triangleq \prod_{i=n_p+1}^{n_p+\bar{n}_p} \frac{z}{z - p_i}, \quad \bar{\xi}_p^{\varepsilon}(z) \triangleq \prod_{i=n_p+1}^{n_p+\bar{n}_p} \frac{z}{z - p_i(1-\varepsilon)}, \tag{2.5}$$

where $\varepsilon \in (0,1]$. Clearly, $\xi_p(z)$ and $\bar{\xi}_p(z)$ are biproper and unstable with stable inverses, $\bar{\xi}_p^{\varepsilon}(z) \in \mathcal{U}_\infty$, and

$$\lim_{\varepsilon \to 0^+} \bar{\xi}_p(z)^{-1} \bar{\xi}_p^{\varepsilon}(z) = 1. \tag{2.6}$$

In an analogous fashion we define the multi-set of strictly-NMP plant zeros via

$$\{c_1, \cdots, c_{n_c}\}_m, \tag{2.7}$$

the multi-set of marginally MP zeros via

$$\{c_{n_c+1}, \cdots, c_{n_c+\bar{n}_c}\}_m, \tag{2.8}$$

[4] Recall that a set has, by definition, only distinct elements. Multi-sets are an extension of the notion of set, where elements can appear more than once (see, e.g., [11]). Thus, the multi-set of unstable poles of $A(z)$ is the "set" of unstable poles of $A(z)$, counting multiplicities.

[5] In this thesis we mostly focus on SISO plant models. Analogous definitions can be given in the MIMO case (see, e.g., [74, 163, 178]).

[6] We adopt the convention that $\prod_{i=n_1}^{n_2} x_i = 1$ whenever $n_1 > n_2$.

and the products

$$\xi_c(z) \triangleq \prod_{i=1}^{n_c} \frac{1 - zc_i}{z - c_i}, \quad \bar{\xi}_c(z) \triangleq \prod_{i=n_c+1}^{n_c+\bar{n}_c} \frac{z}{z - c_i}, \quad \bar{\xi}_c^{\varepsilon}(z) \triangleq \prod_{i=n_c+1}^{n_c+\bar{n}_c} \frac{z}{z - c_i(1-\varepsilon)}, \tag{2.9}$$

where $\varepsilon \in (0, 1]$. We note that $\xi_c(z)$ and $\bar{\xi}_c(z)$ are biproper and unstable with stable inverses, $\bar{\xi}_c^{\varepsilon}(z) \in \mathcal{U}_\infty$, and

$$\lim_{\varepsilon \to 0^+} \bar{\xi}_c(z)^{-1} \bar{\xi}_c^{\varepsilon}(z) = 1. \tag{2.10}$$

The previous definitions apply to the plant transfer function $G(z)$. Of course, analogous definitions can be given for any SISO transfer function $X(z)$. To denote any given multi-set of poles or zeros of $X(z)$ we will add an X superscript to the symbols used in the case of $G(z)$. We will use the same convention with any of the previously defined products related to poles or zeros. (For example, $\{p_1^X, \cdots, p_{n_p^X}^X\}_m$ denotes the multi-set of strictly unstable poles of $X(z)$, and $\xi_p^X(z) \triangleq \prod_{i=1}^{n_p^X} \frac{1 - zp_i^X}{z - p_i^X}$ denotes the corresponding product.)

For any $X(z) \in \overline{\mathcal{U}}_\infty$ and parameter $\varepsilon \in (0, 1]$, we define a (non unique) transfer function $\hat{X}_\varepsilon(z)$ that belongs to $\mathcal{U}_\infty$ for any $\varepsilon \in (0, 1]$ and that is such that

$$\lim_{\varepsilon \to 0^+} \left\| X(z) - \hat{X}_\varepsilon(z) \right\|_2^2 = 0. \tag{2.11}$$

We define the operator $\mathcal{K}_\varepsilon\{\cdot\} : \overline{\mathcal{U}}_\infty \to \mathcal{U}_\infty$ via

$$\mathcal{K}_\varepsilon\{X(z)\} = \begin{cases} X(z) & \text{if } X(z) \in \mathcal{U}_\infty \ , \\ \hat{X}_\varepsilon(z) & \text{if } X(z) \in \overline{\mathcal{U}}_\infty. \end{cases} \tag{2.12}$$

In the scalar case, i.e., when $X(z) \in \overline{\mathcal{U}}_\infty^{1\times 1}$, it is easy to see that one choice for $\hat{X}_\varepsilon(z)$ is given by

$$\hat{X}_\varepsilon(z) \triangleq \bar{\xi}_c^{\varepsilon,X}(z)^{-1} \bar{\xi}_c^X(z) X(z). \tag{2.13}$$

If $X(z) \in \overline{\mathcal{U}}_\infty^{n\times n}$, then one can construct $\hat{X}_\varepsilon(z)$ using the multivariate analogues of $\xi_c^X(z)$, $\bar{\xi}_c^X(z)$ and $\bar{\xi}_D^{\varepsilon,X}(z)$ (see, e.g., [163, 178]). We will say that the $\mathcal{K}_\varepsilon\{\cdot\}$ operator is *redundant* if and only if its argument belongs to $\mathcal{U}_\infty$.

On a few occasions, we will need to use the notion of *strongly stabilizable* plant, i.e., plants that admit stable one degree-of-freedom (dof) controllers (see, e.g., [38, 182]). In those cases, it is useful to define the following:

$$\mathcal{C}_+ \triangleq \{c \in \mathbb{R} : 1 \le c < \infty, c \text{ is a plant zero}\} \triangleq \{c_1^+, \cdots, c_{n_c^+}^+\}, \tag{2.14}$$

$$\mathcal{C}_- \triangleq \{c \in \mathbb{R} : -\infty < c \le -1, c \text{ is a plant zero}\} \triangleq \{c_1^-, \cdots, c_{n_c^-}^-\}, \tag{2.15}$$

where $-\infty < c^-_{n_c^-} < \cdots < c^-_1 \leq -1$ and $1 \leq c^+_1 < \cdots c^+_{n_c^+} < \infty$. If $\mathcal{C}_+$ (resp. $\mathcal{C}_-$) is empty, then $n_c^+ = 0$ (resp. $n_c^- = 0$). If the plant $G(z)$ is strictly proper and $\mathcal{C}_+$ is non empty, then we define, $\forall i \in \{1, \cdots, n_c^+\}$,

$$\mathcal{I}_i^+ \triangleq \begin{cases} (c_i^+, c_{i+1}^+) & \text{if } i < n_c^+ \\ (c_i^+, \infty) & \text{if } i = n_c^+ \end{cases} \tag{2.16}$$

and

$$n_i^+ \triangleq \# \left\{ p \in \mathbb{R} : p \text{ is a pole of } G(z),\, p \in \mathcal{I}_i^+ \right\}_m, \quad \mathcal{I}^+ \triangleq \{ c_i \in \mathcal{C}_+ : n_i^+ \text{ is odd} \}. \tag{2.17}$$

Each set (interval) $\mathcal{I}_i^+$ corresponds to the portion of the positive real line between any two real NMP plant zeros, n_i^+ is the number of unstable plant poles (counting multiplicities) in $\mathcal{I}_i^+$, and $\mathcal{I}^+$ is the set of real positive NMP plant zeros at the beginning of intervals $\mathcal{I}_i^+$ that contain an odd number of unstable plant poles. If $\mathcal{C}_+$ is empty, then $\mathcal{I}^+$ is empty. In an analogous fashion, if the plant is strictly proper and $\mathcal{C}_-$ is not empty, then we define, $\forall i \in \{1, \cdots, n_c^-\}$,

$$\mathcal{I}_i^- \triangleq \begin{cases} (c_{i+1}^-, c_i^-) & \text{if } i < n_c^- \\ (-\infty, c_i^-) & \text{if } i = n_c^- \end{cases} \tag{2.18}$$

and

$$n_i^- \triangleq \# \left\{ p \in \mathbb{R} : p \text{ is a pole of } G(z),\, p \in \mathcal{I}_i^- \right\}_m, \quad \mathcal{I}^- \triangleq \{ c_i \in \mathcal{C}_- : n_i^- \text{ is odd} \}. \tag{2.19}$$

If $\mathcal{C}_-$ is empty, then $\mathcal{I}^-$ is empty. Finally, we define

$$\mathcal{I} = \mathcal{I}^+ \cup \mathcal{I}^- \tag{2.20}$$

and note that $\mathcal{I}$ is empty if and only if $G(z)$ is strongly stabilizable (see, e.g., [38, 182]).

2.3 Stochastic Processes

In this thesis we use the simplest notion of random process as defined in, e.g., [6, 168].[7] That is, we understand the random process $\{x(k)\}_{k\in\mathbb{N}_0}$ as a sequence of (possibly vector valued) *real* random variables that have well defined (joint) PDFs. For simplicity, we will use the shorthand notation x to refer to $\{x(k)\}_{k\in\mathbb{N}_0}$. In exceptional circumstances, we will consider bi-lateral random processes, i.e., we will consider processes of the form $\{x(k)\}_{k\in\mathbb{Z}}$. If that is the case, then

[7]We will use the names *process*, *random process*, *random signal*, *stochastic process*, etc., as synonyms.

we will make this fact explicit. If, in the sequel, we define a notion for random processes where $k \in \mathbb{N}_0$, then the same notion applies with the obvious changes to bi-lateral processes, and also to random variables (if it makes sense).

The following definitions are standard and can be found in, e.g., [6, 168].

Definition 2.1 (Mean, Covariance, etc.) *Consider two random processes x and y. Then:*

1. *The* mean *of x, denoted by $\mu_x(k)$, is defined via $\mu_x(k) \triangleq \mathcal{E}\{x(k)\}$, where $\mathcal{E}\{\cdot\}$ denotes the expectation operator.*

2. *The* cross-covariance function *between x and y, denoted by $R_{xy}(k+\tau, k)$, is defined via*

$$R_{xy}(k+\tau, k) \triangleq \mathcal{E}\left\{(x(k+\tau) - \mu_x(k+\tau))\,(y(k) - \mu_y(k))^T\right\} \tag{2.21}$$

   for all k and τ such that $x(k+\tau)$ and $y(k)$ are defined.[8] *If $x = y$, then we define*

$$R_x(k+\tau, k) \triangleq R_{xx}(k+\tau, k) \tag{2.22}$$

   as the covariance *function of x.*

3. *The* variance matrix *of x, denoted by $P_x(k)$, is defined via $P_x(k) \triangleq R_x(k,k)$. The* variance *of x is defined via $\sigma_x^2(k) \triangleq \mathrm{trace}\{P_x(k)\}$.*

4. *The* second order moments matrix *of x, denoted by $M_x(k)$, is defined via $M_x(k) \triangleq \mathcal{E}\left\{x(k)x(k)^T\right\}$.*

□□

Recall that two random variables α and β are said to be *uncorrelated* if and only if

$$\mathcal{E}\left\{(\alpha - \mu_\alpha)\,(\beta - \mu_\beta)^T\right\} = 0. \tag{2.23}$$

By extension, two processes x and y are said to be *uncorrelated* if and only if

$$\mathcal{E}\left\{(x(i) - \mu_x(i))\,(y(j) - \mu_y(j))^T\right\} = 0, \quad \forall i, j, \tag{2.24}$$

and the process x is said to be uncorrelated with the random variable α if and only if

$$\mathcal{E}\left\{(x(i) - \mu_x(i))\,(\alpha - \mu_\alpha)^T\right\} = 0, \quad \forall i. \tag{2.25}$$

In this thesis we will encounter the following special random processes (see also [6, 168]):

[8] If x and y are bi-lateral, then $x(k+\tau)$ and $y(k)$ are defined for every (k, τ); if, for example, x is not-bilateral, the $x(k+\tau)$ is defined only for $k + \tau \geq 0$.

Definition 2.2 (Types of random processes) *Consider a random process x. Then:*

1. *x is said to be* white *if and only if it is a sequence of uncorrelated random variables with the same mean and variance matrix.*

2. *x is an* i.i.d. sequence *if and only if it is a sequence of identically distributed and independent (i.i.d.) random variables.*

3. *x is an* independent *(resp.* uncorrelated*) process, without specifying with respect to what, if and only if x is independent of (resp. uncorrelated with)* any *other random variable or process. (By extension, we say that a random variable is an independent (resp. uncorrelated) random variable if and only if it is independent of (resp. uncorrelated with)* any *other random variable or process.)*

4. *x is a* second order *process if and only if its mean and second order moments matrix exist and are finite for every $k \in \mathbb{N}_0$ and, moreover, remain finite when $k \to \infty$. (By extension, we say that a random variable is a second order one if and only if it has well defined and finite mean and variance matrix.)*

5. *x is an* asymptotically wss *process if and only if its stationary mean and its stationary covariance function, denoted by μ_x and $R_x(\tau)$ and defined via*[9]

$$\mu_x \triangleq \lim_{k\to\infty} \mu_x(k), \quad R_x(\tau) \triangleq \lim_{k\to\infty} R_x(k+\tau, k), \quad \tau \in \mathbb{N}_0, \tag{2.26}$$

respectively, exist and are finite. In these cases, we also define the stationary variance of x via

$$\sigma_x^2 \triangleq \lim_{k\to\infty} \sigma_x^2(k). \tag{2.27}$$

6. *If x is a bi-lateral process (and only in that case), then we will say that x is* wide sense stationary *(wss) if and only if its mean $\mu_x(k)$ is constant for every $k \in \mathbb{Z}$, and its covariance function is a function of $\tau \in \mathbb{Z}$ only. In these cases we have*

$$\mu_x = \mu_x(k), \quad R_x(\tau) = R_x(k+\tau, k), \quad \sigma_x^2 = \sigma_x^2(k), \quad \forall k \in \mathbb{Z}. \tag{2.28}$$

[9]Note that the definitions in (2.26) and (2.27) are consistent with the definitions in (2.28). Also note the abuse of notation: when μ_x (σ_x^2) has k as an argument, then it refers to the mean (variance) of x at time k, whereas with no arguments refers to the stationary mean (variance) of x (if it exists), or to the mean (variance) of x when x is wss. On the other hand, when R_x has two arguments, it refers to the covariance function of x, and when it has only one argument it refers to the stationary covariance function of x (if it exists) or to the covariance of x when x is wss.

□□

Remark 2.1 (Jointly) *For any given two random processes x and y, define $z \triangleq [x\ y]^T$. If x, y and z are second order processes, then we say that x and y are* jointly second order *processes. The same logic applies to all the notions defined in Definition 2.2.* □□

Remark 2.2 *If x and y are jointly wss, then their cross-covariance function is a function of τ only. In these cases we define*

$$R_{xy}(\tau) \triangleq R_{xy}(k+\tau, k). \tag{2.29}$$

Analogously, if x and y are jointly asymptotically wss, then $R_{xy}(\tau)$ refers to the stationary cross-covariance function between x and y (i.e., $\lim_{k\to\infty} R_{xy}(k+\tau, k)$). □□

When dealing with wss processes, it is useful to define the following:

Definition 2.3 (Power spectral density) *Consider two jointly wss processes x and y. Then:*

1. *The* power spectral density *(PSD) of x, denoted by $S_x(z)$, is defined via*

$$S_x(z) \triangleq \sum_{\tau=-\infty}^{\infty} R_x(\tau) z^{-\tau} \quad \Leftrightarrow \quad R_x(\tau) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_x(e^{j\omega}) e^{j\omega\tau} d\omega. \tag{2.30}$$

   (Note that $S_x(z)$ is guaranteed to be well defined at least on the unit circle, i.e., for $z = e^{j\omega}$; see, e.g., Section 3.4 in [168].)

2. *The* cross-PSD *between x and y, denoted by $S_{xy}(z)$, is defined via*

$$S_{xy}(z) \triangleq \sum_{\tau=-\infty}^{\infty} R_{xy}(\tau) z^{-\tau}. \tag{2.31}$$

□□

When dealing with cross-PSDs, it is sometimes convenient to use the following notation (and its analogue in the case of PSDs):

$$S_{xy}^{+}(z) \triangleq \sum_{\tau=1}^{\infty} R_{xy}(\tau) z^{-\tau}, \quad S_{xy}^{-}(z) \triangleq \sum_{\tau=-\infty}^{-1} R_{xy}(\tau) z^{-\tau}. \tag{2.32}$$

Thus,

$$S_{xy}(z) = S_{xy}^{+}(z) + S_{xy}^{-}(z) + R_{xy}(0) \tag{2.33}$$

and, clearly, $S_{xy}^{+}(z)^{\sim} = S_{yx}^{-}(z)$ and $S_{xy}^{-}(z)^{\sim} = S_{yx}^{+}(z)$, where $A(z)^{\sim}$ is a shorthand for $A(z^{-1})^T$.

We end this section recalling the well-known spectral factorization theorem:

Theorem 2.1 (Spectral factorization (see, e.g., [2, 6, 168])) *Consider a wss process x with rational PSD $S_x(z)$. If $S_x(z)$ has full normal rank,*[10] *then there exists $\Omega_x(z) \in \overline{\mathcal{U}}_\infty$ such that*

$$S_x(z) = \Omega_x(z)\Omega_x(z)^\sim. \tag{2.34}$$

$\Omega_x(z)$ will be referred to as a spectral factor *of the process x.* □□□

Remark 2.3 (Spectral factors in $\mathcal{U}_\infty$) *If $S_x(z)$ has full rank for every $z = e^{j\omega}$ (equivalently, $S_x(z)$ has no zeros on the unit circle), then $\Omega_x(z)$ can be chosen in $\mathcal{U}_\infty$. We also note that, if $S_x(z)$ has zeros on the unit circle, then one can always pick $\hat{\Omega}_x(z) \in \mathcal{U}_\infty$ such that it approximates $\Omega_x(z)$ arbitrarily well in, e.g., $2-$norm sense. Thus, we will often assume that the processes we encounter admit spectral factors in $\mathcal{U}_\infty$.* □□

From the definition of the $2-$norm, it immediately follows that the variance of a wss process x can be written in terms of its spectral factor as (see (2.1) and (2.30))

$$\sigma_x^2 = \text{trace}\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \Omega_x(e^{j\omega})\Omega_x(e^{j\omega})^H d\omega\right\} = ||\Omega_x(z)||_2^2 . \tag{2.35}$$

Remark 2.4 (Asymptotically wss processes and stationary PSDs) *Consider an asymptotically wss process x and a bi-lateral process $\bar{x}$ such that, $\forall\, k \in \mathbb{N}_0$,*

$$\mu_{\bar{x}}(k) = \mu_x, \quad R_{\bar{x}}(k+\tau, k) = \begin{cases} R_x(\tau) & \text{if } \tau \geq 0 \\ R_x(-\tau) & \text{if } \tau < 0 \end{cases}. \tag{2.36}$$

By construction, $\bar{x}$ is a wss process and, hence, admits a PSD. We define the stationary PSD *of x, which, by a slight abuse of notation, will be denoted by $S_x(z)$, as the PSD of $\bar{x}$, i.e.,*

$$S_x(z) \triangleq S_{\bar{x}}(z). \tag{2.37}$$

We note that, as far as spectral and second order properties are concerned, $\bar{x}$ corresponds to the process that one "would observe" when studying x for sufficiently large time instants (i.e., when $k \to \infty$).

Using the previous definition, we can define stationary spectral factors *for asymptotically wss processes as the spectral factors of the corresponding stationary PSDs. Also,* (2.35) *applies to asymptotically wss processes. However, in that case, $\Omega_x(z)$ corresponds to a stationary spectral factor of x, and σ_x^2 corresponds to the stationary variance of x.* □□

[10] Equivalently, $S_x(z)$ is non-singular almost everywhere.

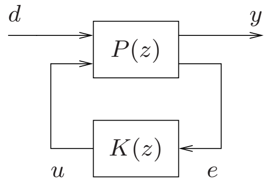

Figure 2.1: General feedback interconnection of systems $P(z)$ and $K(z)$.

2.4 Internal Stability

When dealing with feedback systems, the fundamental question of stability arises. This section summarizes basic stability results for linear systems (see, e.g., [87, 94, 194]).

Consider the feedback system of Figure 2.1, where $P(z) \in \mathcal{R}_p$ and $K(z) \in \mathcal{R}_p$ are the transfer functions of discrete-time LTI systems. Partition $P(z)$ in a way such that

$$\begin{bmatrix} y \\ e \end{bmatrix} = \begin{bmatrix} P_{11}(z) & P_{12}(z) \\ P_{21}(z) & P_{22}(z) \end{bmatrix} \begin{bmatrix} d \\ u \end{bmatrix}, \tag{2.38}$$

and assume that the corresponding underlying realization is given by

$$\left(A, \begin{bmatrix} B_1 & B_2 \end{bmatrix}, \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}, \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} \right), \tag{2.39}$$

where the state space representation matrices have been partitioned in a way consistent with the partition of $P(z)$ in (2.38).

Definition 2.4 (Internal stability [194]) *The system of Figure 2.1 is said to be* internally stable *if and only if the state of the interconnected system converges to zero for all finite initial states when* $d = 0$. □□

A necessary condition for internal stability is that the underlying realizations of both $P(z)$ and $K(z)$ be detectable and stabilizable (see, e.g, [87, 94, 194]).

Definition 2.5 (Well-posed feedback loop [194]) *The system of Figure 2.1 is said to be* well-posed *if and only if* $\det(I - P_{22}(\infty)K(\infty)) \neq 0$. □□

It is easy to see that a well-posed feedback system is such that the transfer functions between any two points are well-defined and proper (see, e.g., [194]).

Definition 2.6 (Stabilizable plant) *Consider a plant with transfer function* $P(z) \in \mathcal{R}_p$ *and realization given by* (2.39)*. The plant is said to be* stabilizable *if and only if there exists* $K(z) \in \mathcal{R}_p$ *such that the feedback loop in Figure 2.1 is internally stable and well-posed.* □□

Plants that are stabilizable can be easily characterized as follows:

Lemma 2.1 (Stabilizable plant (see Lemma 12.1 in [194])) *A plant with transfer function $P(z) \in \mathcal{R}_p$ and realization given by (2.39) is stabilizable if and only if the realization $(A, B_2, C_2, 0)$ is stabilizable and detectable.*

Clearly, $(A, B_2, C_2, 0)$ being stabilizable and detectable is sufficient for the underlying realization of the plant to be detectable and stabilizable (but it is by no means necessary).

For any plant with transfer function $P(z) \in \mathcal{R}_p$ and realization given by (2.39) we define the set

$$\mathcal{S} \triangleq \{K(z) \in \mathcal{R}_p : \text{feedback loop of Figure 2.1 is internally stable and well-posed}\}. \tag{2.40}$$

Any controller $K(z) \in \mathcal{S}$ will be referred to as *admissible* for the plant or, equivalently, admissible for $P(z)$. If $T_{dy}(z)$ denotes the transfer function between d and y in the feedback system of Figure 2.1, then we define

$$\mathcal{T} \triangleq \{T_{dy}(z) \in \mathcal{R}_p : K(z) \in \mathcal{S}\}. \tag{2.41}$$

We note that both $\mathcal{S}$ and $\mathcal{T}$ are not empty if and only if the plant (in short, $P(z)$) is stabilizable. If that is the case, then the next theorem provides a characterization of both $\mathcal{S}$ and $\mathcal{T}$:

Theorem 2.2 (Youla parametrization (see, e.g., [46, 113, 182, 194]) *Consider the feedback system in Figure 2.1 and assume that $P(z)$ is stabilizable. Consider a doubly coprime factorization of $P_{22}(z)$ over $\mathcal{RH}_\infty$, i.e., consider $X_i(z)$, $Y_i(z)$, $X_d(z)$, $Y_d(z)$, $N_i(z)$, $D_i(z)$, $N_d(z)$, $D_d(z) \in \mathcal{RH}_\infty$, with $X_i(z)$, $X_d(z)$, $D_i(z)$, $D_d(z)$ biproper, such that*

$$P_{22}(z) = N_d(z)D_d(z)^{-1} = D_i(z)^{-1}N_i(z) \tag{2.42}$$

and

$$\begin{bmatrix} X_i(z) & -Y_i(z) \\ -N_i(z) & D_i(z) \end{bmatrix} \begin{bmatrix} D_d(z) & Y_d(z) \\ N_d(z) & X_d(z) \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}. \tag{2.43}$$

Then:

1. *$K(z) \in \mathcal{S}$ if and only if*

$$K(z) = (X_i(z) - Q(z)N_i(z))^{-1}(Y_i(z) - Q(z)D_i(z)), \tag{2.44}$$

where $Q(z) \in \mathcal{RH}_\infty$ is such that $\det(X_i(\infty) - Q(\infty)N_i(\infty)) \neq 0$.

2. $T_{dy}(z) \in \mathcal{T}$ *if and only if*

$$T_{dy}(z) = T^o_{dy}(z) - P_{12}(z)D_d(z)Q(z)D_i(z)P_{21}(z), \tag{2.45}$$

where $Q(z)$ *is as in Part 1 above, and*

$$T^o_{dy}(z) \triangleq P_{11}(z) + P_{12}(z)D_d(z)Y_i(z)P_{21}(z). \tag{2.46}$$

3. In (2.45), *the transfer functions* $T^o_{dy}(z), P_{12}(z)D_d(z)$ *and* $D_i(z)P_{21}(z)$ *belong to* $\mathcal{RH}_\infty$.

□□

The importance of Theorem 2.2 is that it gives an *affine* characterization of all transfer functions from d to y that are achievable in the feedback loop of Figure 2.1 with an admissible controller $K(z)$. As such, Theorem 2.2 is the starting point for most optimization based controller synthesis procedures (see also [13] for a detailed discussion of this point).

2.5 Mean Square Stability

In this thesis we study dynamic systems subject to random inputs. In this framework, deterministic notions of stability do not make sense and need to be extended accordingly (see, e.g., [93]). In particular, we will focus on the concept of mean square stability (defined below).

Consider a discrete-time LTI system described via

$$x(k+1) = Ax(k) + Bu(k), \quad k \in \mathbb{N}_0, \quad x(0) = x_0, \tag{2.47}$$

where, for every $k \in \mathbb{N}_0$, $x(k) \in \mathbb{R}^{n_x}$ denotes the system state, $u(k) \in \mathbb{R}^{n_u}$ denotes the system input, and A and B are real matrices of appropriate dimensions. We will assume that the following holds:

Assumption 2.1 (Inputs and initial states)

(a) The initial state x_0 *is a second order random variable having mean* μ_0 *and variance matrix* $P_0 \geq 0$.

(b) The input u *is a second order white sequence uncorrelated with* x_0*, having mean* μ_u *and variance matrix* $P_u \geq 0$.

□□

We are now ready to define mean square stability:

Definition 2.7 (Mean square stability [29, 30, 84]) *The LTI system in* (2.47) *is* mean square stable *(MSS[11]) if and only if, for every x_0 and every u that satisfy Assumption 2.1, there exist $\mu_x \in \mathbb{R}^{n_x}$ and $M_x \in \mathbb{R}^{n_x \times n_x}$, $M_x \geq 0$, both not depending on x_0, such that*[12]

$$\lim_{k\to\infty} \mathcal{E}\left\{x(k)\right\} = \mu_x, \quad \lim_{k\to\infty} \mathcal{E}\left\{x(k)x(k)^T\right\} = M_x. \tag{2.48}$$

□□

In words, Definition 2.7 states that system (2.47) is MSS if and only if its state x has well defined and finite stationary mean and stationary second order moments matrix.

In order to characterize the systems of the form (2.47) that are MSS, we start by noting that the following holds:

Lemma 2.2 (See Section 4.2.2 in [168] or Section 3.3 in [6])) *Consider the LTI system in* (2.47) *and suppose that Assumption 2.1 holds. Then, for every k and τ in $\mathbb{N}_0$,*

$$\mu_x(k+1) = A\mu_x(k) + B\mu_u, \tag{2.49}$$

$$P_x(k+1) = AP_x(k)A^T + BP_uB^T, \tag{2.50}$$

$$R_x(k+\tau, k) = A^\tau P_x(k), \tag{2.51}$$

where $\mu_x(0) = \mu_0$ and $P_x(0) = P_0$. □□□

Lemma 2.2 allows one to prove the following:

Theorem 2.3 (MSS) *The LTI system in* (2.47) *is MSS if and only if A is Hurwitz (i.e., has all its eigenvalues strictly inside the unit circle).*

Proof:

- ($\Rightarrow$) If the system (2.47) is MSS, then, for any x_0 and u satisfying Assumption 2.1, the mean of x converges to a well defined, finite and unique stationary value μ_x that does not depend on the initial state x_0. From Lemma 2.2 we have that, if this is true, then A must be Hurwitz (otherwise, $\mu_x(k)$ does not converge or its stationary value depends on the mean of x_0).

[11] Will also use MSS to refer to mean square stability, where the meaning is clear from the context.

[12] We note that the definition of limit implies that these quantities, if they exist, need to be unique.

- ($\Leftarrow$) If A is Hurwitz, then Lemma 2.2 immediately implies that, for every x_0 and u satisfying Assumption 2.1, $\mu_x(k)$ is well defined and finite for every k and, moreover, converges to $(I-A)^{-1}B\mu_u$ (which does not depend on x_0) as $k \to \infty$. It also follows immediately that $P_x(k)$ is well defined and finite for every $k \in \mathbb{N}_0$ and any x_0 and u satisfying Assumption 2.1. Since $P_u \geq 0$ and A is Hurwitz, basic properties of Lyapunov equations (see, e.g., Section 21.1 in [194]) allow one to conclude that the equation

$$X = AXA^T + BP_uB^T \tag{2.52}$$

admits an unique (well defined and finite) solution, say $X = X_o$, that is, in addition, positive semidefinite and (of course) not dependent on x_0. Thus, use of the proof of Corollary 2 in Section 4.2.2 in [168] allows one to conclude that $P_x(k) \to X_o$ and the result follows.

□□□

Corollary 2.1 (Asymptotic wss) *If the LTI system in* (2.47) *is MSS and Assumption 2.1 holds, then its state x is a second order asymptotically wss random process with stationary PSD*

$$S_x(z) = H_x(z)P_uH_x(z)^\sim, \tag{2.53}$$

where $H_x(z) \triangleq (zI-A)^{-1}B$.

Proof: This result follows immediately from the proof of Theorem 2.3 and Corollary 2 in Section 4.2.2 in [168]. □□□

Corollary 2.2 (Properties of linear combinations of the state) *Consider the LTI system in* (2.47) *and define (here, $\ell \in \mathbb{N}_0$ is just an index)*

$$y_\ell(k) \triangleq C_\ell x(k) + D_\ell u(k), \tag{2.54}$$

where $y_\ell(k) \in \mathbb{R}^{n_{y_\ell}}$, and C_ℓ and D_ℓ are constant matrices of appropriate dimensions. If system (2.47) *is MSS and Assumption 2.1 holds, then y_ℓ is a second order asymptotically wss process with stationary PSD*

$$S_{y_\ell}(z) = H_\ell(z)P_uH_\ell(z)^\sim, \tag{2.55}$$

where $H_\ell(z) \triangleq C_\ell(zI-A)^{-1}B + D_\ell$. Any two processes y_i and y_j of the form (2.54) *are jointly second order asymptotically wss processes.*

Proof: Immediate. □□□

Remark 2.5 (The case of u being wss) *If Assumption 2.1(b) is replaced by u being a second order wss process, uncorrelated with x_0, and having rational PSD $S_u(z) = \Omega_u(z)\Omega_u(z)^\sim$ with $\Omega_u(z) \in \mathcal{RH}_\infty$, then the definition of MSS still applies. Moreover, standard results from* representation *of stochastic processes (see, e.g., Chapter 4 in [6]) allow one to easily see that:*

- *A modified version of Lemma 2.2 holds, where the matrices A and B are replaced by augmented matrices that include the original system matrices and a state space realization of $\Omega_u(z)$.*
- *Theorem 2.3 and Corollaries 2.1 and 2.2 still hold but, in the present case, P_u in equations* (2.53) *and* (2.55) *must be replaced by $S_u(z)$.*

□□

Remark 2.6 (Correlation between u and x_0) *If the correlation between u and x_0 is finite for every $k \in \mathbb{N}_0$ (and remains finite as $k \to \infty$), then the previous characterizations of MSS and of the stationary PSDs of the system state and outputs still apply.* □□

The key message of the previous results is that, when dealing with LTI systems and second order wss processes, the characterization of second-order-moment-related properties of the system state and output is straightforward. In particular, we have made explicit that MSS is equivalent to the asymptotic stability of the underlying LTI system. These observations have importance when put into the context of control system design. For example, it follows from the results in this section that, if, in the feedback system of Figure 2.1, the external signal d is a second order wss process and the initial states of both plant and controller are second order random variables, then the closed loop will be MSS if and only if the controller $K(z)$ is an admissible controller for the plant $P(z)$ (i.e., if and only if $K(z)$ stabilizes $P(z)$ and the feedback loop is well-posed).

2.6 Optimization Problems

Most of the results in this thesis are obtained as solutions to optimization problems. To streamline our notation, we will adopt the following convention regarding optimal values. Consider a set $\mathcal{X}$ and a function J defined on $\mathcal{X}$. If

$$J_o \triangleq \inf_{x \in \mathcal{X}} J(x) \tag{2.56}$$

exists (and is not $-\infty$) and $\inf_{x\in\mathcal{X}} J(x)$ is achievable in $\mathcal{X}$, i.e., there exists $\mathcal{X}_o \subseteq \mathcal{X}$ such that $J_o = J(x)$ for every $x \in \mathcal{X}_o$, then x_o defined via

$$x_o \triangleq \arg \inf_{x \in \mathcal{X}} J(x) \tag{2.57}$$

refers to any $x \in \mathcal{X}_o$. In those cases we often replace inf by min. If, on the other hand, J_o exists, but is not achievable with any $x \in \mathcal{X}$, then there exists a sequence $\{x_n\}_{n\in\mathbb{N}_0}$, $x_n \in \mathcal{X}\ \forall n \in \mathbb{N}_0$, such that

$$\lim_{n\to\infty} J(x_n) = \inf_{x\in\mathcal{X}} J(x) = J_o. \tag{2.58}$$

(In other words, J_o cannot be achieved, but can be approximated to any degree of accuracy with $x \in \mathcal{X}$ in the sense described in (2.58).) In these cases, x_o defined as in (2.57) corresponds to any $x \in \hat{\mathcal{X}}$ such that $J_o = J(x)$, where $\hat{\mathcal{X}} \supset \mathcal{X}$ is an appropriate set.

2.7 Summary

This chapter has provided basic definitions and results that will play a relevant role in the remainder of this thesis. Other known results of less widespread use can be found in the appendices. We would like to warn the reader that most definitions and the notation introduced in this chapter and the appendices are sometimes used without further comment in the forthcoming chapters. If she/he happens to find an unfamiliar symbol, then she/he should be able to find its definition in this chapter, in the appendices or in the list of symbols at the beginning of this thesis.

Chapter 3

Optimal Coding System Design

3.1 Introduction

This chapter deals with networked control systems closed over signal-to-noise ratio constrained additive i.i.d noise channels. We adopt a simplified viewpoint, where a controller is assumed to have already been designed[1] under the assumption of transparent communication links. Within this framework, we propose to use a feedback based coding scheme so as to reduce the impact of the signal-to-noise ratio constraint on closed loop performance. We also study the MSS of the resulting networked control setup. In particular, we obtain a simple characterization of the minimal channel signal-to-noise ratio that allows one to design a feedback coding system that guarantees closed loop stability. This result is expressed in terms of the plant and controller unstable poles only. For stable controllers, and regardless of the plant zeros or relative degree, our results show that the minimal admissible signal-to-noise ratio equals the minimal signal-to-noise ratio identified in [17, 18] for the less demanding static state feedback case.

The idea of designing coding schemes to embellish given controller designs is not new. For example, our previous work documented in [55] considers a coding scheme that turns out to be a special case of the one considered here. On the other hand, [160] considers the same coding architecture as the one studied in this chapter, but the design procedure in [160] assumes that the channel signal-to-noise ratio is *sufficiently large*. The methodology used in this chapter does not require that assumption. Also, the stability analysis included in the current chapter goes beyond the results of [55, 160]. Other work (loosely) related to the results in this chapter

[1] More general scenarios will be considered later.

are [24, 25, 101, 112]. We will comment on that work in Section 3.6, later in this chapter.

In the second part of this chapter we extend our framework to the MIMO case. We consider the control of two-input two-output MIMO LTI systems for which a decentralized controller has been successfully designed. We investigate the possible benefits of enriching the decentralized control structure with additional communication channels. These channels allow one to implement full MIMO (i.e., centralized) controllers where the off-diagonal terms communicate through non transparent channels. Within this setup, we show, for a given full-MIMO controller, how to design coding schemes that optimize overall performance. Unsurprisingly, for high signal-to-noise ratio channels, the networked MIMO architecture outperforms the decentralized one. However, an interesting finding is that, in some situations, the networked architecture will perform better than the decentralized one, only if the channels have very large signal-to-noise ratios.

The remainder of this chapter is organized as follows: Section 3.2 presents the architecture of interest and our working assumptions. Section 3.3 focuses on MSS, whilst Section 3.4 presents an iterative procedure for designing the proposed coding scheme. Our approach is then extended to the MIMO case in Section 3.5. Section 3.6 presents some numerical examples that, among other things, illustrate how to use our results to design feedback quantizers that mitigate the impact of quantization on closed loop performance, when a simplified additive noise model for quantization is adopted (see Appendix B). Section 3.7 draws conclusions.

3.2 Problem Definition

In this chapter we will consider the networked control architecture depicted in Figure 3.1. In that figure, r is the reference signal, d models output disturbances and the channel is an additive i.i.d. noise channel with a signal-to-noise ratio constraint. To avoid ambiguities, we introduce the following definition:

Definition 3.1 (Additive i.i.d. noise channel with a signal-to-noise ratio constraint) *Consider the channel in Figure 3.2, where v is the input, w the output, q models channel imperfections, and*

$$w(k) \triangleq v(k) + q(k), \quad \forall k \in \mathbb{N}_0, \quad \forall v(k) \in \mathbb{R}. \tag{3.1}$$

We say that the channel is a signal-to-noise ratio constrained additive i.i.d. noise channel *if and only if:*

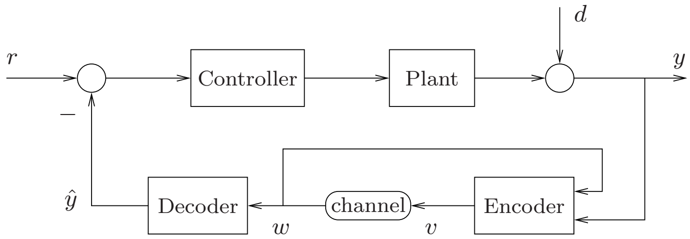

Figure 3.1: Considered networked control architecture.

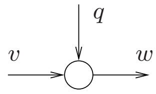

Figure 3.2: Additive i.i.d. noise channel.

- *q is an independent second order zero-mean i.i.d. sequence with variance σ_q^2. The latter is a quantity to be designed, and*
- *the ratio between the stationary variance of v and the variance of q is upper bounded, i.e., there exists $\Gamma \in \mathbb{R}_0^+$ such that*

$$\gamma \triangleq \frac{\sigma_v^2}{\sigma_q^2} \leq \Gamma. \tag{3.2}$$

(Sometimes, we will use the wording additive i.i.d. noise channel with signal-to-noise ratio constraint Γ *to emphasize the fact that the channel has a signal-to-noise ratio upper bounded by Γ.)* □□

Remark 3.1 *In many situations one can relax the independence requirements in Definition 3.1 to just uncorrelatedness. If one does so, then the channel noise becomes just white (not i.i.d.). In this thesis, however, we consider settings where independence is actually required.* □□

Remark 3.2 (Design of σ_q^2) *As discussed in Section 1.2.2, additive i.i.d. noise channels with signal-to-noise ratio constraints arise when scaling factors are used at the input and output of an additive i.i.d. noise channel that has a power constraint. Thus, the design of such scaling factors amounts to the design of the equivalent noise variance σ_q^2. As another illustration of what designing the variance σ_q^2 means, note that, as we will show in Chapter 5, in the case of control systems subject to average data-rate constraints, the choice of the source coding scheme determines the statistics of the resulting quantization noise. Accordingly, designing the source coding scheme amounts to choosing the equivalent noise statistics.* □□

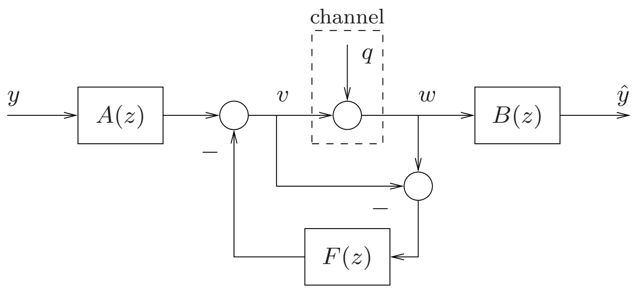

Figure 3.3: Feedback coding architecture (including channel model).

We will focus on a situation where the controller has *already been designed* assuming a transparent communication link[2] and, thus, only the coding scheme has to be designed. The control loop formed by the plant and controller, when transparent communication links are in place (i.e., when $\hat{y} = y$ in Figure 3.1) will be referred to as the *nominal loop* (or nominal design).

We will consider LTI coding systems and, consistent with the setup introduced in Chapter 1, we will assume that the output of the channel w is known at the sending end with one step delay. The use of any additional communication channel not explicitly shown in Figure 3.1 is, however, impossible. Inspired by standard feedback coding ideas (see, e.g., [37, 83, 151]), we propose to use the coding architecture in Figure 3.3 (where we have made the channel model explicit). In that figure, $A(z), B(z)$ and $F(z)$ are filters in $\mathcal{R}_p$ that need to be designed. Since w is available with one step delay, $F(z)$ must be strictly proper.

As already mentioned above, we are interested in designing coding systems for pre-specified nominal designs. In this setting, it is natural to employ coding systems that, in the absence of channel artifacts, have unit transfer function. These coding architectures are usually referred to as *perfect reconstruction* ones (see, e.g., [37]). In our case, it is easy to see from Figure 3.3 that

$$\hat{y} = B(z)A(z)y + B(z)(1 - F(z))q. \tag{3.3}$$

Therefore, perfect reconstruction is equivalent to $B(z) = A(z)^{-1}$.

Consistent with the previous discussion, we will focus on feedback coding schemes that satisfy the following:

Constraint 3.1 (Structural constraints on the feedback coder) *The feedback coder filters are such that $B(z) = A(z)^{-1}$ and $F(z) \in \mathcal{R}_{sp}$.* □□

[2]This can, of course, be carried out using any standard design tool (see, e.g., [53, 167]).

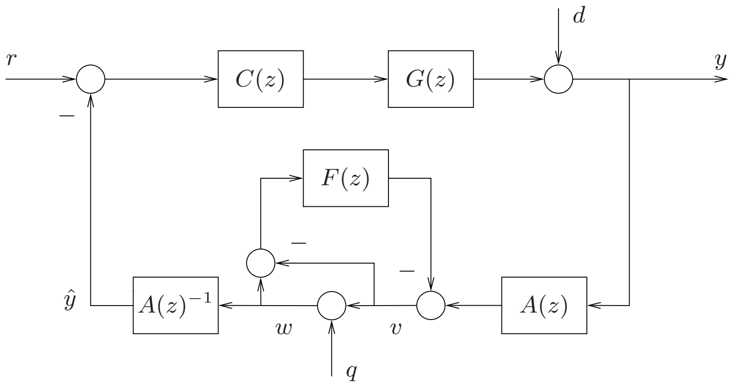

Figure 3.4: Networked control system with feedback coding.

The networked control system that arises when the proposed coding architecture is used in the networked control system of Figure 3.1 is shown in Figure 3.4. In that figure, $G(z)$ is the plant transfer function, $C(z)$ is the controller transfer function, and we have made explicit the perfect reconstruction constraint. In the sequel, we will assume that the following holds:

Assumption 3.1 (Plant, nominal design, signals and initial states)

(a) The plant transfer function $G(z)$ belongs to $\mathcal{R}_{sp}^{1\times 1}$, is non-zero, and has an underlying realization that is detectable and stabilizable.

(b) The controller $C(z)$ belongs to $\mathcal{R}_{p}^{1\times 1}$, and is a fixed non-zero one-degree-of-freedom (one-dof) admissible controller for $G(z)$.[3]

(c) The signals r and d are second order mutually uncorrelated wss scalar processes with rational PSDs. At least one signal, r or d, is non-zero. If r or d are non-zero, then they admit spectral factors in $\mathcal{U}_\infty$.

(d) The initial states of all filters in Figure 3.4 (including plant and controller) are jointly second order random variables.

□□

The assumptions on $G(z)$, r, d, and the initial states are as in Chapter 1 and were justified there. The assumption of $C(z)$ being non-zero discards non interesting situations, where the nominal loop is such that $G(z)$ is left in open loop (i.e., uncontrolled). Assuming that the

[3] i.e., the standard one-dof control loop formed by $G(z)$ and $C(z)$ (see, e.g., [53]) is internally stable and well posed.

controller is such that the nominal loop is internally stable and well-posed is, of course, sensible in our context where the coding system is to be designed after the controller has already been chosen.

The main focus of this chapter lies in studying the interplay between stationary signal-to-noise ratio constraints, i.e., constraints on

$$\gamma \triangleq \frac{\sigma_v^2}{\sigma_q^2}, \tag{3.4}$$

and the performance of the networked control architecture of Figure 3.4. In particular, we will first focus on signal-to-noise ratio requirements for MSS. In a second stage, we will show how to design the coding scheme in Figure 3.3 having closed loop performance in mind.

For future reference, we define the nominal sensitivity function $S(z)$ and the nominal complementary sensitivity function $T(z)$ via (see, e.g., [53])

$$S(z) \triangleq \frac{1}{1+G(z)C(z)}, \quad T(z) \triangleq 1 - S(z). \tag{3.5}$$

(We also suggest that the reader familiarize him/herself with the notation introduced in Section 2.2, if not already done so.)

3.3 Mean Square Stability

In this section we study stability properties of the networked control system in Figure 3.4.

In our setting, one has the freedom to choose $A(z)$, $F(z)$ and the channel noise variance σ_q^2 (recall Definition 3.1). As long as Assumption 3.1 holds, the feedback system in Figure 3.4 is MSS if an only if $A(z) \in \mathcal{U}_\infty$, $F(z) \in \mathcal{RH}_2$ and $\sigma_q^2 \in \mathbb{R}_0^+$ (i.e., σ_q^2 is finite).[4] The following results holds:

Theorem 3.1 (Minimal signal-to-noise ratio for MSS) *Consider the feedback system in Figure 3.4, where q is the noise in an additive i.i.d. noise channel, and suppose that Assumption 3.1 holds. Then,*

$$\gamma_{\text{inf}} \triangleq \inf_{\substack{A(z)\in\mathcal{U}_\infty \\ F(z)\in\mathcal{RH}_2 \\ \sigma_q^2\in\mathbb{R}_0^+}} \gamma = \left(\prod_{i=1}^{n_p} |p_i|^2\right)\left(\prod_{i=1}^{n_p^C} |p_i^C|^2\right) - 1, \tag{3.6}$$

where $\{p_1^C, \cdots, p_{n_p^C}^C\}_m$ denotes the multi-set of strictly unstable poles of the controller $C(z)$. The infimum γ_{inf} is not achievable, but γ can be made arbitrarily close to γ_{inf} upon choosing a

[4] Recall that $\mathbb{R}_0^+ \triangleq \{x \in \mathbb{R} : 0 \le x < \infty\}$.

sufficiently large (but finite) $\sigma_q^2 \in \mathbb{R}_0^+$*, any* $A(z) \in \mathcal{U}_\infty$ *and*[5]

$$F(z) = 1 - \mathcal{K}_\varepsilon \left\{ \left(\xi_p(z) \xi_p^C(z) S(z) \right) \right\}^{-1} \xi_p(\infty) \xi_p^C(\infty), \tag{3.7}$$

where ε *is a sufficiently small (but non-zero) real value in* $(0, 1]$*. Moreover, the* $\mathcal{K}_\varepsilon\{\cdot\}$ *operator in* (3.7) *is redundant if and only if both* $G(z)$ *and* $C(z)$ *have no poles on the unit circle.*

Proof: Straightforward analysis reveals that, if Assumption 3.1 holds, $A(z) \in \mathcal{U}_\infty$, $F(z) \in \mathcal{RH}_2$ and $\sigma_q^2 \in \mathbb{R}_0^+$, then σ_v^2 exists, is finite and we can write

$$\gamma = \frac{\sigma_v^2}{\sigma_q^2} = ||T(z) + S(z)F(z)||_2^2 + \frac{||A(z) \left[T(z)\Omega_r(z) \;\; S(z)\Omega_d(z) \right]||_2^2}{\sigma_q^2}, \tag{3.8}$$

where $T(z)$ is the nominal complementary sensitivity function, $S(z)$ is the nominal sensitivity function, and $\Omega_r(z) \in \mathcal{RH}_\infty$ (resp. $\Omega_d(z) \in \mathcal{RH}_\infty$) is a spectral factor of r (resp. d). Defining

$$\xi(z) \triangleq \xi_p(z) \xi_p^C(z), \tag{3.9}$$

we have that, for any $A(z) \in \mathcal{U}_\infty$, $F(z) \in \mathcal{RH}_2$ and $\sigma_q^2 \in \mathbb{R}_0^+$,

$$\begin{aligned}
\gamma &\geq ||T(z) + S(z)F(z)||_2^2 \\
&\overset{(a)}{=} ||\xi(z) - \xi(z)S(z) + \xi(z)S(z)F(z)||_2^2 \\
&\overset{(b)}{=} ||\xi(z) - \xi(0)||_2^2 + ||\xi(0) - \xi(\infty) + (\xi(\infty) - \xi(z)S(z)) + \xi(z)S(z)F(z)||_2^2 \\
&\overset{(c)}{=} ||\xi(z) - \xi(0)||_2^2 + ||\xi(0) - \xi(\infty)||_2^2 + ||(\xi(\infty) - \xi(z)S(z)) + \xi(z)S(z)F(z)||_2^2 \\
&\overset{(d)}{\geq} \left(\prod_{i=1}^{n_p} |p_i|^2 \right) \left(\prod_{i=1}^{n_p^C} \left|p_i^C\right|^2 \right) - 1,
\end{aligned} \tag{3.10}$$

where (a) follows from the definition of $T(z)$ and $S(z)$, the fact that $\xi_p(z)$ and $\xi_p^C(z)$ are unitary (hence, $\xi(z)$ is unitary; see Definition A.1 in Appendix A) and properties of the 2-norm, (b) follows from Part 4 in Fact A.1 in Appendix A and the fact that $\xi(z) \in \mathcal{RH}_2^\perp$ and $\xi(z)S(z) \in \mathcal{RH}_\infty$ (this follows from the definition of $\xi(z)$ and the fact that $S(z)$ is the sensitivity function of a stable and well-posed control loop [Assumption 3.1]), (c) follows from Part 3 in Fact A.1 and the fact that $\xi(0) - \xi(\infty) \in \mathcal{RH}_2^\perp$ and $(\xi(\infty) - \xi(z)S(z)) + \xi(z)S(z)F(z) \in \mathcal{RH}_2$ (recall that $F(z) \in \mathcal{RH}_2$), and (d) follows from the Residue Theorem (see, e.g., Appendix A in [155] and also the proof of Theorem 4.1 in Chapter 4).

From the previous discussion, it follows that

$$\gamma \overset{(a)}{\geq} ||T(z) + S(z)F(z)||_2^2 \overset{(b)}{\geq} \left(\prod_{i=1}^{n_p} |p_i|^2 \right) \left(\prod_{i=1}^{n_p^C} \left|p_i^C\right|^2 \right) - 1 = \gamma_{\text{inf}}. \tag{3.11}$$

[5]Recall the definition of $\mathcal{K}_\varepsilon\{\cdot\}$ in Section 2.2.

Equality in (a) holds for $\sigma_q^2 \in \mathbb{R}_0^+$ if and only if $A(z)\,[\,T(z)\Omega_r(z) \;\; S(z)\Omega_d(z)\,] = 0$ which, given Assumption 3.1 and the fact that $A(z) \in \mathcal{U}_\infty$, cannot hold. Nevertheless, if σ_q^2 is sufficiently large (but finite), then the gap between both sides in (a) can be made arbitrarily small. On the other hand, equality in (b) holds if and only if (recall (3.9))

$$F(z) = 1 - \left(\xi_p(z)\xi_p^C(z)S(z)\right)^{-1} \xi_p(\infty)\xi_p^C(\infty). \tag{3.12}$$

The right hand side in (3.12) belongs to $\mathcal{RH}_2$ if and only if $\left(\xi_p(z)\xi_p^C(z)S(z)\right)^{-1} \in \mathcal{RH}_\infty$, i.e., if and only if neither the plant nor the controller have marginally stable poles. If this is not the case, then use of the $\mathcal{K}_\varepsilon\{\cdot\}$ operator allows one to build a function in $\mathcal{RH}_2$ that achieves a gap in (b) as small as desired (see (3.7)). (This follows directly form the definition of $\mathcal{K}_\varepsilon\{\cdot\}$ introduced in Section 2.2.) The proof is completed upon noting that $A(z)$ plays no role in our previous derivations (so, any $A(z) \in \mathcal{U}_\infty$ is useful). □□□

Theorem 3.1 gives an explicit characterization of the minimal channel signal-to-noise ratio that is compatible with MSS. This minimal signal-to-noise ratio depends only on the unstable poles of the plant $G(z)$ and controller $C(z)$. We note that, if the plant is strongly stabilizable (i.e., if it admits stable one-dof admissible controllers; see, e.g., [38, 182]) and the nominal controller is marginally stable, then γ_{inf} is a function of the unstable plant poles only; namely,

$$\gamma_{\text{inf}} = \left(\prod_{i=1}^{n_p} |p_i|^2\right) - 1. \tag{3.13}$$

From Theorem 3.1 we see that any design for $F(z), A(z)$ and σ_q^2 in Figure 3.4 that guarantees MSS is such that the corresponding stationary signal-to-noise ratio satisfies $\gamma > \gamma_{\text{inf}}$. Conversely, for any $\gamma > \gamma_{\text{inf}}$, it is possible to choose $F(z), A(z)$ and σ_q^2 so as to guarantee MSS and achieve a signal-to-noise ratio equal to γ. Such a choice can be readily found using the results in Theorem 3.1. As a consequence, we have:

Corollary 3.1 (Conditions on Γ) *Consider the networked control situation in Figure 3.1, where the coding architecture is as in Figure 3.3, the channel is an additive i.i.d. noise channel with signal-to-noise ratio constraint Γ, and Assumption 3.1 holds. $\Gamma > \gamma_{\text{inf}}$ is necessary and sufficient to be able to find $A(z)$, $F(z)$ and σ_q^2 such that MSS is guaranteed and the signal-to-noise ratio constraint $\gamma \leq \Gamma$ satisfied.*

Proof: Immediate from Theorem 3.1 and Definition 3.1. □□□

Remark 3.3 (Other architectures) *It is easy to see that the following holds:*

- *If* $A(z)$ *is not a decision variable, but a fixed transfer function in* $\mathcal{U}_\infty$, *then the results of both Theorem 3.1 and Corollary 3.1 still apply.*
- *If* $F(z) \in \mathcal{RH}_2$ *is fixed, then* $\sigma_q^2 \in \mathbb{R}_0^+$ *if and only if*

$$\Gamma \geq \gamma > ||T(z) + S(z)F(z)||_2^2 . \tag{3.14}$$

Thus, if $F(z) \in \mathcal{RH}_2$ *is fixed, then there exist* $A(z)$ *and* σ_q^2 *such that MSS is guaranteed, and the signal-to-noise ratio constraint is satisfied, if and only if* (3.14) *holds.*

□□

Remark 3.4 (Previous work) *If we fix* $F(z) = 0$ *and it were possible to redesign the controller (which is* not *the case here), then Theorem III.2 in [18] allows one to conclude that the minimal stationary channel signal-to-noise ratio compatible with MSS obeys*

$$\gamma_{\text{inf}} = \left(\prod_{i=1}^{n_p} |p_i|^2 \right) - 1 + \Delta_G, \tag{3.15}$$

where $\Delta_G \geq 0$ *depends on the strictly-NMP zeros and relative degree of the plant model* ($\Delta_G = 0$ *if and only if* $G(z)$ *has no strictly-NMP zeros and has relative degree one). If* $G(z)$ *is strongly stabilizable, and one chooses a stable nominal controller, then it is immediate to see from* (3.15) *and* (3.13) *that the signal-to-noise ratio requirements for MSS in the case that one employs no feedback in the coder (i.e., when* $F(z) = 0$*) are, in general, more demanding than in the proposed architecture.*[6] *(The only case in which the signal-to-noise ratio constraints are the same arises if* $G(z)$ *has no strictly-NMP zeros and relative degree one.) This is an important indication of the benefits that feedback coding brings to networked control, and is consistent with well-known results that show that additional degrees of freedom are beneficial for control (see, e.g., [26, 28, 182]). This issue is explored in greater detail in Chapter 4.* □□

3.4 Design for Performance

In this section we go beyond stability and focus on how to actually design the proposed coding architecture so as to minimize the impact of signal-to-noise ratio constraints on closed loop performance. To that end, we define the tracking error e via

$$e \triangleq r - y, \tag{3.16}$$

[6] We note that the gap may be very significative if, e.g., the plant has high relative degree.

and use its stationary variance σ_e^2 as performance measure. The problem of interest can be stated as follows:

Problem 3.1 (Best performance with signal-to-noise ratio constraint) *Consider the networked control situation in Figure 3.1, where the coding architecture is as in Figure 3.3, the channel is an additive i.i.d. noise channel with signal-to-noise ratio constraint* Γ*, and Assumption 3.1 holds. Find (or prove the problem to be infeasible)*

$$\left[\sigma_e^2\right]_\Gamma \triangleq \inf_{\substack{A(z)\in\mathcal{U}_\infty \\ F(z)\in\mathcal{RH}_2 \\ \sigma_q^2\in\mathbb{R}_0^+ \\ \gamma\leq\Gamma}} \sigma_e^2 \tag{3.17}$$

and find the filters $F(z)$ *and* $A(z)$ *that achieve* $\left[\sigma_e^2\right]_\Gamma$ *(or, at least, approximate* $\left[\sigma_e^2\right]_\Gamma$ *with arbitrary precision).*[7]

Problem 3.1 has been tackled in the literature using two assumptions. In [55] the authors[8] constrained themselves to a case where $F(z) = 0$ and only $A(z)$ is to be designed. On the other hand, [160] gives an analytic and explicit approximation to the solution of Problem 3.1 under the assumption that the signal-to-noise ratio constraint Γ is *high enough*. In this chapter we tackle the general case, and make no assumptions regarding the size of Γ. A similar (and in principle simpler) problem has been recently solved in [37] in the context of feedback quantization for analog-to-digital conversion (see, e.g., [83, 151]). Unfortunately, the technique employed in [37] does not seem to yield an immediate explicit characterization of the solution in the present case. Instead of pursuing that line of reasoning here, we will present below an iterative approach that is guaranteed to yield performance that is close to optimal.

Before describing the proposed design procedure, we note that the following holds:

Fact 3.1 (Asymptotic behavior of $\left[\sigma_e^2\right]_\Gamma$) *Consider the networked control situation in Figure 3.1, where the coding architecture is as in Figure 3.3, the channel is an additive i.i.d. noise channel with signal-to-noise ratio constraint* Γ*, and Assumption 3.1 holds. Then:*

1. *If* $\Gamma \to \gamma_{\inf}$*, then* $\left[\sigma_e^2\right]_\Gamma \to \infty$*.*
2. *If* $\Gamma \to \infty$*, then* $\left[\sigma_e^2\right]_\Gamma \to \sigma_N^2$*, where* $\sigma_N^2 \triangleq ||S(z)\left[\Omega_r(z)\ \ \Omega_d(z)\right]||_2^2$ *is the tracking error variance in the nominal loop.*

[7] We note that, under Assumption 3.1, the constraints on $F(z)$, $A(z)$ and σ_q^2 are necessary and sufficient to achieve MSS.

[8] Note that the current author is a co-author of [37, 55, 160].

Proof: It is easy to see that, provided Assumption 3.1 holds, $A(z) \in \mathcal{U}_\infty$, $F(z) \in \mathcal{RH}_2$ and $\sigma_q^2 \in \mathbb{R}_0^+$ (as required for MSS), then σ_e^2 exists, is finite and obeys

$$\sigma_e^2 = ||S(z)\left[\Omega_r(z) \;\; \Omega_d(z)\right]||_2^2 + \sigma_q^2 \left|\left|T(z)A(z)^{-1}(1-F(z))\right|\right|_2^2, \tag{3.18}$$

where all symbols are as in the proof of Theorem 3.1.

1. If $\gamma \to \gamma_{\text{inf}}$, then the proof of Theorem 3.1 (see (3.11)) allows one to conclude that $\sigma_q^2 \to \infty$. Therefore, $\sigma_e^2 \to \infty$ unless $T(z)A(z)^{-1}(1-F(z)) = 0$ which, since $F(z) \in \mathcal{RH}_2$, $A(z) \in \mathcal{U}_\infty$ and Assumption 3.1 holds, is impossible. The result thus follows.

2. It follows from (3.8) that, for any $A(z) \in \mathcal{U}_\infty$ and $F(z) \in \mathcal{RH}_2$, $\gamma \to \infty$ if and only if $\sigma_q^2 \to 0$ (recall that Assumption 3.1 holds and, hence, the numerator in the second term in the right hand side of (3.8) is finite and non-zero). The result is now immediate from (3.18).

□□□

From Fact 3.1 we see that the requirements on the stationary channel signal-to-noise ratio for MSS are inadequate from the perspective of performance. This is, indeed, the motivation behind Problem 3.1. Since Fact 3.1 gives a characterization of $\left[\sigma_e^2\right]_\Gamma$ when $\Gamma \to \infty$ or $\Gamma \to \gamma_{\text{inf}}$, we will omit these cases in our subsequent discussion.

3.4.1 Choosing the pre- and post-filter $A(z)$

We begin by showing how to choose $A(z)$, when $F(z) \in \mathcal{RH}_2$ is given. To that end we define, for any $F(z) \in \mathcal{RH}_2$,

$$J_F \triangleq \inf_{\substack{A(z)\in\mathcal{U}_\infty \\ \sigma_q^2\in\mathbb{R}_0^+ \\ \gamma\leq\Gamma}} \sigma_e^2, \tag{3.19}$$

and

$$(A_F(z), \sigma_F^2) \triangleq \arg \inf_{\substack{A(z)\in\mathcal{U}_\infty \\ \sigma_q^2\in\mathbb{R}_0^+ \\ \gamma\leq\Gamma}} \sigma_e^2. \tag{3.20}$$

It follows from Remark 3.3 that, since $F(z)$ is fixed, Γ must satisfy (3.14) in order for the optimization problem in (3.19) to be feasible. Having this in mind, the next theorem characterizes J_F and the corresponding optimal quantization noise variance σ_F^2 and the optimal filter $A_F(z)$:

Theorem 3.2 (Optimal $A(z)$ for a given $F(z)$) *Consider the networked control situation in Figure 3.1, where the coding architecture is as in Figure 3.3, the channel is an additive i.i.d. noise channel with signal-to-noise ratio constraint Γ, and Assumption 3.1 holds. If $F(z)$ is any fixed transfer function in $\mathcal{RH}_2$ and $||T(z)+S(z)F(z)||_2^2 < \Gamma < \infty$, then:*

1. *The minimal tracking error variance that satisfies the signal-to-noise ratio constraint, i.e., J_F, is given by*

$$J_F = \sigma_N^2 + \frac{\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \left|T(e^{j\omega})(1-F(e^{j\omega}))\right| \sqrt{\left|T(e^{j\omega})\Omega_r(e^{j\omega})\right|^2 + \left|S(e^{j\omega})\Omega_d(e^{j\omega})\right|^2}\, d\omega\right)^2}{\Gamma - ||T(z)+S(z)F(z)||_2^2}, \tag{3.21}$$

where σ_N^2 is the tracking error variance in the nominal loop (see Fact 3.1).

2. *J_F can be achieved upon choosing $A_F(z)$ such that*

$$\left|A_F(e^{j\omega})\right|^4 = \alpha \frac{\left|T(e^{j\omega})(1-F(e^{j\omega}))\right|^2}{\left|T(e^{j\omega})\Omega_r(e^{j\omega})\right|^2 + \left|S(e^{j\omega})\Omega_d(e^{j\omega})\right|^2}, \quad \forall \omega \in [-\pi, \pi], \tag{3.22}$$

where α is any arbitrary positive real number, and

$$\sigma_F^2 = \frac{||A_F(z)\,[\,T(z)\Omega_r(z) \;\; S(z)\Omega_d(z)\,]||_2^2}{\Gamma - ||T(z)+S(z)F(z)||_2^2}. \tag{3.23}$$

Proof: Since $\gamma_{\inf} \leq ||T(z)+S(z)F(z)||_2^2 < \Gamma < \infty$ (see also (3.11)), we know that the optimal noise variance σ_F^2 belongs to $\mathbb{R}^+$ (see proof of Fact 3.1). We will thus focus on $\sigma_q^2 \in \mathbb{R}^+$. Since Assumption 3.1 holds, we have that for any $A(z) \in \mathcal{U}_\infty$, $F(z) \in \mathcal{RH}_2$ and $\sigma_q^2 \in \mathbb{R}^+$, both σ_e^2 and γ exists, are finite and obey

$$\sigma_e^2 = ||S(z)\,[\,\Omega_r(z) \;\; \Omega_d(z)\,]||_2^2 + \sigma_q^2 \left|\left|T(z)A(z)^{-1}(1-F(z))\right|\right|_2^2, \tag{3.24}$$

$$\gamma = ||T(z)+S(z)F(z)||_2^2 + \frac{||A(z)\,[\,T(z)\Omega_r(z) \;\; S(z)\Omega_d(z)\,]||_2^2}{\sigma_q^2}, \tag{3.25}$$

where all symbols are as in the proof of Theorem 3.1.

Since the nominal controller is fixed, the first term in (3.24) is constant. We will thus focus on the second term. Using (3.25), the constraint $\gamma \leq \Gamma$ can be written in the equivalent form

$$\sigma_q^2 \geq \frac{||A(z)\,[\,T(z)\Omega_r(z) \;\; S(z)\Omega_d(z)\,]||_2^2}{\Gamma - ||T(z)+S(z)F(z)||_2^2}. \tag{3.26}$$

Thus,

$$
\begin{aligned}
&\sigma_q^2 \left|\left|T(z)A(z)^{-1}(1-F(z))\right|\right|_2^2 \\
&\quad \overset{(a)}{\geq} \frac{||A(z)\left[T(z)\Omega_r(z) \;\; S(z)\Omega_d(z)\right]||_2^2 \left|\left|T(z)A(z)^{-1}(1-F(z))\right|\right|_2^2}{\Gamma - ||T(z)+S(z)F(z)||_2^2} \\
&\quad \overset{(b)}{\geq} \frac{\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|T(e^{j\omega})(1-F(e^{j\omega}))\right|\sqrt{\left|T(e^{j\omega})\Omega_r(e^{j\omega})\right|^2+\left|S(e^{j\omega})\Omega_d(e^{j\omega})\right|^2}\,d\omega\right)^2}{\Gamma - ||T(z)+S(z)F(z)||_2^2}, \quad (3.27)
\end{aligned}
$$

where the last inequality follows from the Cauchy-Schwartz inequality and Fact A.1 in Appendix A. Equality in (a) holds if and only if σ_q^2 equals the right hand size of (3.26). On the other hand, equality in (b) holds if and only if $A(z)$ is such that its magnitude raised to the fourth power equals the right hand side of (3.22) (which, given Assumption 3.1, is finite for all $\omega \in [-\pi, \pi]$). The result follows noting that $A_F(z)$ can always be approximated, to any desired degree of accuracy, by a rational filter in $\mathcal{U}_\infty$ (see, e.g., [1, 72, 141]). □□□

The characterization of $A_F(z)$ given by Theorem 3.2, although explicit, is generally not satisfied by any transfer function in $\mathcal{U}_\infty$. This is due to the fact that the fourth root of the right hand side in (3.22) is, usually, irrational. Nevertheless, as mentioned in the proof of Theorem 3.2, it is always possible to find a filter in $\mathcal{U}_\infty$ that achieves a performance that is as close as desired to J_F.[9] In practice, it is normally sufficient to consider reasonably low order filters to approximate $A_F(z)$ (see also [55]).

3.4.2 Choosing the feedback filter $F(z)$

We next address the problem of choosing $F(z)$ when $A(z) \in \mathcal{U}_\infty$ is given. Consistent with the notation introduced before, we define

$$
J_A \triangleq \inf_{\substack{F(z)\in\mathcal{RH}_2 \\ \sigma_q^2\in\mathbb{R}_0^+ \\ \gamma\leq\Gamma}} \sigma_e^2, \tag{3.28}
$$

and

$$
\left(F_A(z), \sigma_A^2\right) \triangleq \arg \inf_{\substack{F(z)\in\mathcal{RH}_2 \\ \sigma_q^2\in\mathbb{R}_0^+ \\ \gamma\leq\Gamma}} \sigma_e^2. \tag{3.29}
$$

[9] This is, of course, consistent with the definition of $A_{opt}^{F(z)}$ and the definition of arg inf (see Section 2.6).

Our aim in this section is to provide a one parameter characterization for J_A, the corresponding optimal filter $F_A(z)$, and the optimal quantization noise variance σ_A^2. Towards that goal, we begin by studying an auxiliary problem:

Lemma 3.1 (Auxiliary problem) *Define the functional*

$$L_\epsilon(F(z)) \triangleq \epsilon J(F(z)) + (1-\epsilon)R(F(z)), \tag{3.30}$$

where $\epsilon \in [0,1]$,

$$J(F(z)) \triangleq \left|\left|T(z)A(z)^{-1}(1-F(z))\right|\right|_2^2, \quad R(F(z)) \triangleq ||T(z)+S(z)F(z)||_2^2, \tag{3.31}$$

$A(z) \in \mathcal{U}_\infty$, *and* $T(z)$ *and* $S(z)$ *are defined in* (3.5) *with* $G(z)$ *and* $C(z)$ *satisfying Assumption 3.1. Then,*

$$F_\epsilon(z) \triangleq \arg \inf_{F(z)\in\mathcal{RH}_2} L_\epsilon(F(z)) = 1 - P_o(\infty)P_o(z)^{-1}, \tag{3.32}$$

where $P_o(z)$ *is an outer factor (see Appendix A) of* [10]

$$P_\epsilon(z) \triangleq \begin{bmatrix} \sqrt{\epsilon}\ \xi_c(z)\xi_c^C(z)T(z)A(z)^{-1} \\ \sqrt{1-\epsilon}\ \xi_p(z)\xi_p^C(z)S(z) \end{bmatrix}. \tag{3.33}$$

Moreover, $F_\epsilon(z) \in \mathcal{RH}_2$ *if and only if* $\epsilon \in (0,1)$, *or* $\epsilon = 0$ *and* $G(z)C(z)$ *has no marginally stable poles, or* $\epsilon = 1$ *and* $G(z)C(z)$ *has no marginally-MP zeros.*

Proof: (Our proof technique follows the ideas described, e.g., in [26].) Using the proof of Theorem 3.1 we can write L_ϵ as

$$L_\epsilon(F(z)) = (1-\epsilon)\left(\gamma_{\inf} + ||(\xi(z)S(z)-\xi(\infty)) - \xi(z)S(z)F(z)||_2^2\right) + \\ \epsilon\left|\left|T(z)A(z)^{-1} - T(z)A(z)^{-1}F(z)\right|\right|_2^2, \tag{3.34}$$

where the symbols are as in the proof of Theorem 3.1. Defining

$$H(z) \triangleq \xi_c(z)\xi_c^C(z)T(z)A(z)^{-1}, \tag{3.35}$$

it is possible to write

$$\begin{aligned} \left|\left|T(z)A(z)^{-1} - T(z)A(z)^{-1}F(z)\right|\right|_2^2 &\overset{(a)}{=} ||H(z)-H(z)F(z)||_2^2 \\ &= ||H(\infty) + (H(z)-H(\infty)) - H(z)F(z)||_2^2 \\ &\overset{(c)}{=} |H(\infty)|^2 + ||(H(z)-H(\infty)) - H(z)F(z)||_2^2, \end{aligned} \tag{3.36}$$

[10] Recall the notation introduced in Section 2.2.

where (a) follows from the fact that $\xi_c(z)\xi_c^C(z)$ is unitary, and (b) follows using Part 3 in Fact A.1 in Appendix A and the fact that $H(\infty) \in \mathcal{RH}_2^\perp$ and $(H(z) - H(\infty)) - H(z)F(z) \in \mathcal{RH}_2$. (This follows from the definition of $\xi_c(z)\xi_c^C(z)$, the fact that $A(z) \in \mathcal{U}_\infty$, $F(z) \in \mathcal{RH}_2$, and that $T(z)$ is the complementary sensitivity of an internally stable and well-posed control loop [Assumption 3.1].) From (3.34) and (3.36) it follows that

$$L_\epsilon(F(z)) = (1-\epsilon)\gamma_{\inf} + \epsilon\,|H(\infty)|^2 + \hat{L}_\epsilon(F(z)), \tag{3.37}$$

where

$$\hat{L}_\epsilon(F(z)) \triangleq ||W(z) - P_\epsilon(z)F(z)||_2^2\,, \tag{3.38}$$

$P_\epsilon(z)$ (defined in (3.33)) belongs to $\mathcal{RH}_\infty$, is biproper, and

$$W(z) \triangleq \begin{bmatrix} \sqrt{\epsilon}\ \ (H(z) - H(\infty)) \\ \sqrt{1-\epsilon}\ \ (\xi(z)S(z) - \xi(\infty)) \end{bmatrix} \in \mathcal{RH}_2. \tag{3.39}$$

Since Assumption 3.1 holds, it is clear that $P_\epsilon(z)$ has full normal column rank. Therefore, $P_o(z) \in \overline{\mathcal{U}}_\infty$ (see Theorem A.1). Moreover, since Assumption 3.1 holds, $T(z)$ and $S(z)$ share no zeros on the unit circle. Since $A(z) \in \mathcal{U}_\infty$, it follows that $P_o(z) \in \mathcal{U}_\infty\ \forall\, \epsilon \in (0,1)$. If $\epsilon \in \{0,1\}$, then $P_o(z)$ has as marginally-MP zeros the marginally-MP zeros of either $S(z)$ or $T(z)$ (i.e., the marginally unstable poles of $G(z)C(z)$ or marginally-MP zeros of $G(z)C(z)$).

Consider the unitary matrix

$$\phi(z) \triangleq \begin{bmatrix} P_i(z)^\sim \\ I - P_i(z)P_i(z)^\sim \end{bmatrix}, \tag{3.40}$$

where $P_i(z)$ is the inner factor of $P_\epsilon(z)$ in the inner-outer factorization $P_\epsilon(z) = P_i(z)P_o(z)$. Since $\phi(z)$ is unitary and $P_i(z)^\sim P_i(z) = I$, we have from (3.38) that

$$\begin{aligned} \hat{L}_\epsilon(F(z)) &= ||\phi(z)\,(W(z) - P_\epsilon(z)F(z))||_2^2 \\ &= ||(I - P_i(z)P_i(z)^\sim)\,W(z)||_2^2 + ||P_i(z)^\sim W(z) - P_o(z)F(z)||_2^2 \\ &\overset{(a)}{=} ||(I - P_i(z)P_i(z)^\sim)\,W(z)||_2^2 + ||P_o(z) - P_i(z)^\sim P_\epsilon(\infty) - P_o(z)F(z)||_2^2 \\ &\overset{(b)}{=} ||(I - P_i(z)P_i(z)^\sim)\,W(z)||_2^2 + ||P_o(\infty) - P_i(z)^\sim P_\epsilon(\infty)||_2^2 + \\ &\qquad ||(P_o(z) - P_o(\infty)) - P_o(z)F(z)||_2^2\,, \end{aligned} \tag{3.41}$$

where (a) follows from the definition of $W(z)$, $P_\epsilon(z)$, $P_i(z)$ and $P_o(z)$, and (b) follows from Part 3 in Fact A.1 and the fact that $P_o(\infty) - P_i(z)^\sim P_\epsilon(\infty) \in \mathcal{RH}_2^\perp$ and $(P_o(z) - P_o(\infty)) - P_o(z)F(z) \in$

$\mathcal{RH}_2$. Since $P_o(z) \in \overline{\mathcal{U}}_\infty$, the first part of the result follows. Our remaining claim follows from the properties of $P_o(z)$ discussed previously. □□□

The characterization of $F_\epsilon(z)$ given in Lemma 3.1 plays an essential role in our subsequent discussion. It is worth mentioning that the only critical step when calculating $F_\epsilon(z)$ is the inner-outer factorization of $P(z)$. This can be done using the factorization techniques in, e.g., [46,123].

The next theorem provides a characterization of the optimal filter $F_A(z)$ in terms of $F_\epsilon(z)$.

Theorem 3.3 (Optimal $F(z)$ for a fixed $A(z)$) *Consider the networked control situation in Figure 3.1, where the coding architecture is as in Figure 3.3, the channel is an additive i.i.d. noise channel with signal-to-noise ratio constraint Γ, and Assumption 3.1 holds. Assume that $A(z)$ is any fixed function in $\mathcal{U}_\infty$ and that $\gamma_{\inf} < \Gamma < \infty$. Define the set*

$$\Sigma_\Gamma \triangleq \left\{ \sigma_q^2 \in \mathbb{R}_0^+ : \sigma_q^2 \geq \frac{||A(z)\left[T(z)\Omega_r(z) \;\; S(z)\Omega_d(z)\right]||_2^2}{\Gamma - \gamma_{\inf}} \right\} \tag{3.42}$$

and also the function $f : \Sigma_\Gamma \to [0,1]$ defined via

$$f(\sigma_q^2) \triangleq \begin{cases} 1 & \text{if } g(\sigma_q^2) > R(F_1(z)), \\ \epsilon & \text{if } \epsilon \in [0,1] \text{ is such that } g(\sigma_q^2) = R(F_\epsilon(z)), \end{cases} \tag{3.43}$$

where

$$g(\sigma_q^2) \triangleq \Gamma - \frac{||A(z)\left[T(z)\Omega_r(z) \;\; S(z)\Omega_d(z)\right]||_2^2}{\sigma_q^2}, \tag{3.44}$$

and R and $F_\epsilon(z)$ defined as in (3.31) and (3.32), respectively. Then:

1. *The optimal noise variance σ_A^2 and the optimal filter $F_A(z)$ are given by*

$$\sigma_A^2 = \arg \min_{\sigma_q^2 \in \Sigma_\Gamma} \sigma_q^2 \left|\left| T(z)A(z)^{-1}(1 - F_{f(\sigma_q^2)}(z)) \right|\right|_2^2, \quad F_A(z) = F_{f(\sigma_A^2)}(z). \tag{3.45}$$

   Thus, the minimal tracking error variance J_A satisfies

$$J_A = \sigma_N^2 + \sigma_A^2 \left|\left| T(z)A(z)^{-1}(1 - F_A(z)) \right|\right|_2^2, \tag{3.46}$$

   where σ_N^2 is the tracking error variance in the nominal loop (see Fact 3.1).

2. *If $g(\sigma_A^2) < R(F_1(z))$, or $g(\sigma_A^2) \geq R(F_1(z))$ and $G(z)C(z)$ has no marginal-MP zeros, then J_A is achievable (i.e., $\sigma_A^2 \in \mathbb{R}_0^+$ and $F_A(z) \in \mathcal{RH}_2$). If none of the previous conditions*

hold, then it is possible to make σ_e^2 arbitrarily close to J_A, while violating the signal-to-noise ratio constraint as little as desired, upon choosing $\sigma_q^2 = \sigma_A^2$ and

$$F(z) = 1 - P_A(\infty)\mathcal{K}_\varepsilon \left\{P_A(z)\right\}^{-1}, \tag{3.47}$$

where $P_A(z) \in \overline{\mathcal{U}}_\infty$ is such that $F_A(z) = 1 - P_A(\infty)P_A(z)^{-1}$, and ε is a sufficiently small (but non-zero) real in $(0, 1]$.

Proof:

1. Provided Assumption 3.1 holds, $A(z) \in \mathcal{U}_\infty$, $F(z) \in \mathcal{RH}_2$ and $\sigma_q^2 \in \mathbb{R}_0^+$, then σ_e^2 exists, is finite and obeys (see (3.18) and recall the definition of σ_N^2 in Fact 3.1)

$$\sigma_e^2 = \sigma_N^2 + \sigma_q^2 \left|\left|T(z)A(z)^{-1}(1 - F(z))\right|\right|_2^2, \tag{3.48}$$

where all symbols are as in the proof of Theorem 3.1. Therefore, $J_A = \sigma_N^2 + \hat{J}_A$, where

$$\hat{J}_A = \inf_{\sigma_q^2 \in \mathbb{R}_0^+} \inf_{\substack{F(z) \in \mathcal{RH}_2 \\ \gamma \leq \Gamma}} J(F(z)), \tag{3.49}$$

and J was defined in (3.31). Since $\gamma_{\text{inf}} < \Gamma$, it follows from the definition of γ_{inf} that the problem of finding J_A (equivalently, of finding $\hat{J}_A$) is feasible. Also, we know from the proof of Fact 3.1, that $\gamma_{\text{inf}} < \Gamma < \infty$ also implies that $\sigma_A^2 \in \mathbb{R}^+$. Therefore, one can replace the inf operator by the min operator in the outer optimization problem in (3.49).

The constraint $\gamma \leq \Gamma$ is equivalent to

$$R(F(z)) \leq \Gamma - \frac{\left|\left|A(z)\left[T(z)\Omega_r(z) \;\; S(z)\Omega_d(z)\right]\right|\right|_2^2}{\sigma_q^2} = g(\sigma_q^2), \tag{3.50}$$

where R was defined in (3.31). By definition of γ_{inf} and the proof of Theorem 3.1, it follows that $R(F(z)) \geq \gamma_{\text{inf}}$. Therefore, for any $\sigma_q^2 \in \mathbb{R}^+$, the inner optimization problem in (3.49) is feasible if and only if $g(\sigma_q^2) \geq \gamma_{\text{inf}} \Leftrightarrow \sigma_q^2 \in \Sigma_\Gamma$.

From the previous discussion, it follows that

$$\hat{J}_A = \min_{\sigma_q^2 \in \Sigma_\Gamma} \inf_{\substack{F(z) \in \mathcal{RH}_2 \\ R(F(z)) \leq g(\sigma_q^2)}} J(F(z)). \tag{3.51}$$

For any fixed $\sigma_q^2 \in \Sigma_\Gamma$, we define

$$\bar{F}_{\sigma_q^2}(z) \triangleq \arg \inf_{\substack{F(z) \in \mathcal{RH}_2 \\ R(F(z)) \leq g(\sigma_q^2)}} J(F(z)). \tag{3.52}$$

The well known KKT optimality conditions (see, e.g., [16,102]) state that $\bar{F}_{\sigma_q^2}(z)$ must be a stationary point of

$$L(F(z)) \triangleq \lambda_1 J(F(z)) + \lambda_2 R(F(z)), \tag{3.53}$$

for some non-negative λ_1 and λ_2, satisfying $\lambda_1 + \lambda_2 > 0$ and $\lambda_2 \left(R(\bar{F}_{\sigma_q^2}(z)) - g(\sigma_q^2) \right) = 0$. Since $\lambda_1 + \lambda_2 > 0$, we can define $\epsilon \triangleq \lambda_1(\lambda_1 + \lambda_2)^{-1}$ (which is clearly in $[0,1]$). We thus conclude that $\bar{F}_{\sigma_q^2}(z)$ must be a stationary point of L_ϵ for some $\epsilon \in [0,1]$ (see (3.30)). Moreover, since $A(z) \in \mathcal{U}_\infty$ and, by Assumption 3.1, neither $S(z)$ nor $T(z)$ are zero, it follows from Lemma A.2 in Appendix A that both J and R are strictly convex functionals of $F(z)$. Thus, L_ϵ is also strictly convex in $F(z)$ and, as such, has only one stationary point at $F_\epsilon(z)$ (see (3.32)). As a consequence, $\bar{F}_{\sigma_q^2}(z) = F_\epsilon(z)$ for some $\epsilon \in [0,1]$ such that $(1-\epsilon)\left(R(\bar{F}_{\sigma_q^2}(z)) - g(\sigma_q^2) \right) = 0$.

Consider any given $\sigma_q^2 \in \Sigma_\Gamma$. If $R(F_1(z)) \geq g(\sigma_q^2)$, then, by convexity, there exists $\epsilon \in [0,1]$ such that $R(F_\epsilon(z)) = g(\sigma_q^2)$. In these cases, it is immediate to see that the inequality constraint in (3.52) is active at the optimum. As a consequence, we have that, if $R(F_1(z)) \geq g(\sigma_q^2)$, then (recall the definition of f)

$$\bar{F}_{\sigma_q^2}(z) = F_{f(\sigma_q^2)}(z). \tag{3.54}$$

If, on the other hand, $R(F_1(z)) < g(\sigma_q^2)$, then, by convexity again, $R(F_\epsilon(z)) < g(\sigma_q^2)$ for every $\epsilon \in [0,1]$ and the constraint on R is redundant. In those cases, one can choose $\bar{F}_{\sigma_q^2}(z) = F_1(z) = F_{f(\sigma_q^2)}(z)$, where we have used the definition of f. (We note that in this case there are infinitely many $F(z)$ that satisfy (3.52).)

The previous analysis provides an expression for the parameter $F(z)$ that solves the inner problem in (3.51) for every $\sigma_q^2 \in \Sigma_\Gamma$. To solve (3.51), it suffices to pick the value of σ_q^2 that minimizes $\sigma_q^{-2}\, J(\bar{F}_{\sigma_q^2}(z)) = \sigma_q^{-2}\, J(F_{f(\sigma_q^2)}(z))$. This is achieved by solving the optimization problem in (3.45) which always admits a solution since $\sigma_A^2 \in \Sigma_\Gamma$. The result is now immediate.

2. To prove this part, it suffices to analyze the cases where $F_A(z) \notin \mathcal{RH}_2$. By Lemma 3.1, this can be reduced to the study of the cases where $f(\sigma_A^2) \in \{0,1\}$ and $G(z)C(z)$ has either zeros or poles on the unit circle.

   Since $\Gamma > \gamma_{\inf}$ and the curve of all points in the (J,R) plane that are achievable with $F(z) = F_\epsilon(z)$, $\epsilon \in [0,1]$, is strictly convex, we conclude that, for every $\sigma_q^2 \in \Sigma_\Gamma$, we have

$f(\sigma_q) > 0$. On the other hand, using convexity again one has that, if $g(\sigma_q^2) < R(F_1(z))$, then necessarily $f(\sigma_A^2) < 1$. Therefore, Lemma 3.1 allows one to conclude that the only case where $F_A(z) \notin \mathcal{RH}_2$ arises when $g(\sigma_q^2) \geq R(F_1(z))$ (hence, $f(\sigma_q^2) = 1$) and $G(z)C(z)$ has marginally-MP zeros. The result now follows from the definition of $\mathcal{K}_\varepsilon\{\cdot\}$ and (3.32).

□□□

Remark 3.5 *From the proof of Theorem 3.3 it follows that, since $\gamma_{\text{inf}} < \Gamma < \infty$, one of the conditions in the definition of f is always satisfied (see* (3.43)*).* □□

Theorem 3.3 provides a one parameter characterization of the minimal tracking error variance that is compatible with the signal-to-noise constraint, when $A(z) \in \mathcal{U}_\infty$ is fixed. The scalar parameter, namely the optimal channel noise variance σ_A^2, can be found using any standard line search procedure and, as such, its calculation embodies no additional difficulties. Our result also gives a characterization of the optimal channel noise variance σ_A^2 and the optimal filter $F_A(z)$ that allows one to achieve (or approximate arbitrarily well) J_A at the expense of having a signal-to-noise ratio no greater than Γ (or arbitrarily close to Γ).

We end this section by noting that the characterization of $F_A(z)$ simplifies considerably, if certain conditions are met:

Corollary 3.2 (Optimal $F(z)$ for a fixed $A(z)$ (simplified)) *Consider the networked control situation in Figure 3.1, where the coding architecture is as in Figure 3.3, the channel is an additive i.i.d. noise channel with signal-to-noise ratio constraint Γ, and Assumption 3.1 holds. Assume that $A(z)$ is any fixed function in $\mathcal{U}_\infty$ and that $\gamma_{\text{inf}} < \Gamma < \infty$. If, for every $\sigma_q^2 \in \Sigma_\Gamma$,*

$$\Gamma - \frac{||A(z)\left[T(z)\Omega_r(z)\ \ S(z)\Omega_d(z)\right]||_2^2}{\sigma_q^2} \leq R(F_1(z)), \tag{3.55}$$

then

$$\sigma_A^2 = \frac{||A(z)\left[T(z)\Omega_r(z)\ \ S(z)\Omega_d(z)\right]||_2^2}{\Gamma - ||T(z) + S(z)F_{\epsilon^*}(z)||_2^2}, \quad F_A(z) = F_{\epsilon^*}(z), \tag{3.56}$$

where

$$\epsilon^* = \arg\min_{\epsilon\in[0,1]} \frac{||A(z)\left[T(z)\Omega_r(z)\ \ S(z)\Omega_d(z)\right]||_2^2 \left|\left|T(z)A(z)^{-1}(1 - F_\epsilon(z))\right|\right|_2^2}{\Gamma - ||T(z) + S(z)F_\epsilon(z)||_2^2}. \tag{3.57}$$

Proof: Immediate from the proof of Theorem 3.3. □□□

Remark 3.6 *We note that, for the result of Corollary 3.2 to hold, it is sufficient that* $\Gamma - \sigma_A^{-2} \left|\left| A(z) \left[T(z)\Omega_r(z) \;\; S(z)\Omega_d(z) \right] \right|\right|_2^2 \leq R(F_1(z))$. *Usefulness of this condition is limited since* σ_A^2 *is not known* a-priori. □□

3.4.3 Design procedure and final remarks

In this section we show how to use the results in Sections 3.4.1 and 3.4.2 to iteratively design the proposed coding architecture. It should be clear that one can always choose to fix one of the coder filters (trivial choices are $A(z) = 1$ or $F(z) = 0$) and then use Theorem 3.2 or 3.3 to design the free parameter and the corresponding quantization noise variance σ_q^2. This choice will, clearly, limit the achievable performance. To exploit the full potential of the proposed coding architecture, we suggest that one uses the following algorithm:

Algorithm 3.1 (Iterative design procedure)

- Step 1: *Pick a tolerance value* $\rho > 0$ *and set* $A(z) = A_o(z)$, *where* $A_o(z)$ *is any function in* $\mathcal{U}_\infty$. *Set* $i = 0$ *and* $\sigma_i^2 = \infty$.
- Step 2: *Set* $i = i + 1$ *and fix* $A(z)$ *at its current value. Use Theorem 3.3 to obtain* $F_A(z)$ *and* σ_A^2. *Set* $\sigma_i^2 = J_A$ *and* $F(z) = F_A(z)$.
- Step 3: *Set* $i = i + 1$ *and fix* $F(z)$ *at its current value. Use Theorem 3.2 to obtain* $A_F(z)$ *and* σ_F^2. *Set* $\sigma_i^2 = J_F$ *and* $A(z) = A_F(z)$.
- Step 4: *If* $\left|\sigma_i^2 - \sigma_{i-1}^2\right| \sigma_i^{-2} < \rho$, *then* Stop. *Else, return to* Step 2.

□□

It should be clear that it is not certain that Algorithm 3.1 will converge to the global minimum of the problem of minimizing σ_e^2 subject to the considered signal-to-noise ratio constraint. Nevertheless, it is easy to see that, by construction, Algorithm 3.1 reduces the value of σ_e^2 at each iteration and, therefore, it necessarily converges to a local minimum. To find the global minimum, we suggest to use multiple starting points and/or to try the trivial variant of Algorithm 3.1 that starts by choosing an initial value for $F(z)$ instead of $A(z)$.

In general, $A_F(z) \neq 1$ and $F_A(z) \neq 0$. Thus, fixing $A(z)$ or $F(z)$ and optimally choosing the other filter will obviously provide a coding architecture that enhances closed loop performance when compared with a situation where no pre- or post-processing is available. It is also clear that the use of Algorithm 3.1 allows one to design coding architectures that will always outperform those designed using the guidelines in our previous work described in [55]. This is a consequence of the fact that [55] constrains $F(z)$ to be identically zero.

A more interesting discussion arises if one compares the results in this chapter with the results in another of our early contributions, namely [160]. In that work, it is assumed that γ is *sufficiently high* and guidelines for the design of $F(z)$ and $A(z)$ are provided. An obvious problem with that approach is that deciding, *a priori*, which γs are high enough seems to be impossible. Our current proposal makes no approximations. Therefore, it must necessarily outperform the performance achieved by the filter choices in [160]. (Of course, the filters suggested in [160] may provide a good starting point for Algorithm 3.1.)

Simulation results that verify our conclusions are included in Section 3.6.

3.5 A Simple Extension MIMO Plants

In this section we briefly extend some of the ideas presented above to the case of two-by-two MIMO systems. We will consider a situation where a given plant model $G(z) \in \mathcal{R}_{sp}^{2\times 2}$ is currently controlled using a completely decentralized (i.e., diagonal) controller that uses two transparent communication links. Our aim is to explore the possible benefits of replacing this control architecture by a networked full-MIMO controller that makes use of the existing transparent communication links and of two additional non-transparent communication channels.

We assume that a full-MIMO controller has already been designed assuming transparent communication between plant and controller. We denote this controller by $C(z) \in \mathcal{R}_p^{2\times 2}$. The diagonal terms of this controller are implemented without communication constraints (i.e., using the existing links). However, the off-diagonal terms communicate with the plant using non transparent communication links (i.e., using the additional links). We will refer to such a networked architecture as a *partly networked* control architecture.

Each non transparent communication link comprises a perfect reconstruction coder-decoder pair and a signal-to-noise constrained additive i.i.d. noise channel as depicted in Figure 3.5 (recall Definition 3.1). In that figure, $F_i(z) \in \mathcal{R}_p^{1\times 1}$ is the i-th ($i \in \{1, 2\}$) coder transfer function, v_i is the i-th channel input and w_i is the i-th channel output. These signals are related via

$$w_i = v_i + q_i, \tag{3.58}$$

where q_i is the i-th channel noise. Consistent with Definition 3.1 we will assume the following:

Assumption 3.2 (Channel model)

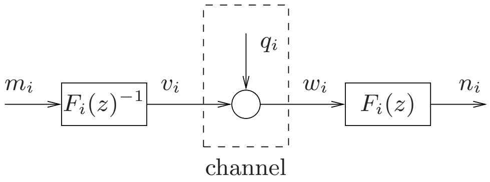

Figure 3.5: i-th communication link.

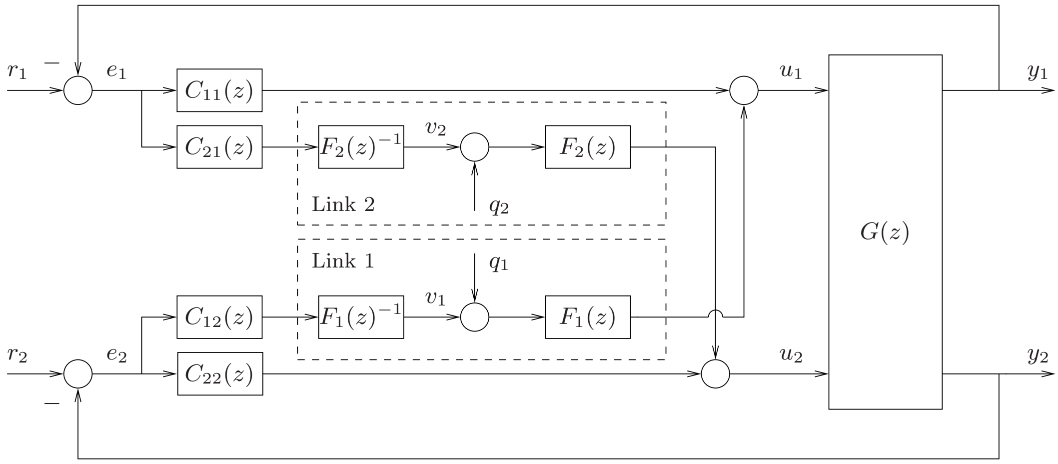

Figure 3.6: Partly networked MIMO control architecture.

(a) The noise signals q_1 *and* q_2 *are independent second order zero-mean i.i.d. sequences, having variances* σ_1^2 *and* σ_2^2 *which are design parameters.*

(b) There exists a constraint on the stationary channel signal-to-noise ratios, i.e.,

$$\gamma_i \triangleq \frac{\sigma_{v_i}^2}{\sigma_i^2} \leq \Gamma_i, \quad i \in \{1, 2\}, \tag{3.59}$$

where Γ_1 *and* Γ_2 *are given constants in* $\mathbb{R}_0^+$.

□□

The networked control system which results from employing the links described above to implement the off-diagonal terms of $C(z)$ is shown in Figure 3.6. In that figure, $u \triangleq [u_1\ u_2]^T$ is the plant input, $y \triangleq [y_1\ y_2]^T$ is the plant output, $r \triangleq [r_1\ r_2]^T$ is a reference sequence, and $e \triangleq [e_1\ e_2]^T$ denotes the tracking error, i.e.,

$$e \triangleq r - y. \tag{3.60}$$

In the remainder of this chapter we will show how to choose $F_1(z)$ and $F_2(z)$ so as to mitigate the effects of the communication constraints on the stationary variance of the tracking error σ_e^2.

To that end, we will make the following assumption whose validity can be justified using the same arguments employed to justify Assumption 3.1:

Assumption 3.3 (Plant, controller, reference signal and initial states)

(a) The plant transfer function $G(z)$ belongs to $\mathcal{R}_{sp}^{2\times 2}$, is non-zero, and has an underlying realization that is detectable and stabilizable.

(b) The controller $C(z)$ belongs to $\mathcal{R}_{p}^{2\times 2}$, and is a fixed non-zero one-dof-admissible controller for $G(z)$.

(c) The reference signal r is a second order wss process that admits a spectral factor in $\mathcal{U}_\infty$.

(d) The initial states of all the filters in Figure 3.6 (including plant and controller) are jointly second order random variables.

□□

For future reference we define (compare with (3.5))

$$S(z) \triangleq (I + G(z)C(z))^{-1}, \quad T(z) \triangleq I - S(z). \tag{3.61}$$

3.5.1 Mean square stability

In this section we provide conditions that guarantee the MSS of the feedback system of Figure 3.6. As a byproduct, we will show that, in some cases of interest, the advantages of a full-MIMO controller design can be void due to the communication constraints that appear in the partly networked implementation.

It is immediate to see that, provided Assumptions 3.2(a) and 3.3 hold, the feedback system in Figure 3.6 is MSS if and only if $F_1(z), F_2(z) \in \mathcal{U}_\infty$ and $\sigma_1^2, \sigma_2^2 \in \mathbb{R}_0^+$. Therefore, we have the following result:

Theorem 3.4 (MSS) *Consider the feedback system in Figure 3.6 and suppose that Assumptions 3.2(a) and 3.3 holds. Define*

$$F(z) \triangleq \operatorname{diag}\{F_1(z), F_2(z)\}, \tag{3.62}$$

and

$$A_1(z) \triangleq F_1(z)^{-1} C_{12}(z) \varepsilon_2^T S(z), \quad A_2(z) \triangleq F_2(z)^{-1} C_{21}(z) \varepsilon_1^T S(z). \tag{3.63}$$

Then, for any $F_1(z), F_2(z) \in \mathcal{U}_\infty$, it is possible to choose σ_i^2 so that the feedback system in Figure 3.6 is MSS and the constraint (3.59) in Assumption 3.2(b) holds if and only if

$$\Gamma_1 > B_1 \triangleq \left|\left|C_{12}(z)\varepsilon_2^T S(z)G(z)\varepsilon_1\right|\right|_2^2, \tag{3.64}$$

$$\Gamma_2 > B_2 \triangleq \left|\left|C_{21}(z)\varepsilon_2^T S(z)G(z)\varepsilon_2\right|\right|_2^2, \tag{3.65}$$

and

$$(\Gamma_1 - B_1)(\Gamma_2 - B_2) > ||A_1(z)G(z)F(z)\varepsilon_2||_2^2 \, ||A_2(z)G(z)F(z)\varepsilon_1||_2^2 . \tag{3.66}$$

Proof: Straightforward analysis of Figure 3.6 shows that, provided Assumptions 3.2(a) and 3.3 hold and $F_1(z), F_2(z) \in \mathcal{U}_\infty$, $\sigma_{v_i}^2$, $i \in \{1,2\}$, exist, are finite and obey

$$\sigma_{v_i}^2 = ||A_i(z)\Omega_r(z)||_2^2 + \sum_{j=1}^{2} \sigma_j^2 \, ||A_i(z)G(z)F(z)\varepsilon_j||_2^2 . \tag{3.67}$$

It thus makes sense to write

$$\gamma_i = \frac{\sigma_{v_i}^2}{\sigma_i^2} = \frac{||A_i(z)\Omega_r(z)||_2^2}{\sigma_i^2} + \sum_{j=1}^{2} \frac{\sigma_j^2}{\sigma_i^2} \, ||A_i(z)G(z)F(z)\varepsilon_j||_2^2 . \tag{3.68}$$

Using the definition of $A_i(z)$, one obtains from (3.68)

$$\gamma_i - B_i = \frac{||A_i(z)\Omega_r(z)||_2^2}{\sigma_i^2} + \sum_{\substack{j=1 \\ j\neq i}}^{2} \frac{\sigma_j^2 \, ||A_i(z)G(z)F(z)\varepsilon_j||_2^2}{\sigma_i^2}. \tag{3.69}$$

Solving (3.69) for σ_i^2, and using the definition of $A_i(z)$ again, it follows that

$$\sigma_i^2 = \Delta^{-1}\Big((\gamma_j - B_j) \, ||A_i(z)\Omega_r(z)||_2^2 + ||A_j(z)\Omega_r(z)||_2^2 \, ||A_i(z)G(z)F(z)\varepsilon_j||_2^2 \Big), \tag{3.70}$$

where $j \in \{1,2\}$, $j \neq i$, and

$$\Delta = (\gamma_1 - B_1)(\gamma_2 - B_2) - ||A_1(z)G(z)F(z)\varepsilon_2||_2^2 \, ||A_2(z)G(z)F(z)\varepsilon_1||_2^2 . \tag{3.71}$$

It easily follows from (3.69)-(3.71) and the definition of Γ_i, that it is possible to choose $\sigma_i^2 \in \mathbb{R}_0^+$ (hence, MSS is achieved) and satisfy $\gamma_i \leq \Gamma_i$ if and only if Γ_i satisfies (3.64)-(3.66). The proof is now complete. □□□

It is illustrative to note that, as long as Assumptions 3.2(a) and 3.3 hold, and $F_1(z), F_2(z) \in \mathcal{U}_\infty$, condition (3.64) (resp. (3.65)) is necessary and sufficient for the existence of $\sigma_1^2 \in \mathbb{R}_0^+$ (resp. $\sigma_2^2 \in \mathbb{R}_0^+$) that guarantees the satisfaction of Assumption 3.2(b) when $q_2 = 0$, i.e., when

$\Gamma_2 \to \infty$ (resp. when $\Gamma_1 \to \infty$). In other words, condition (3.64) (resp. (3.65)) is necessary and sufficient for MSS when only the channel associated to Link 1 (resp. Link 2) has communication constraints. On the other hand, for finite Γ_is (and hence for finite γ_is) both channels have communication constraints. Hence, (3.64)-(3.65) are no longer sufficient for MSS, and (3.66) is required. It is also interesting to note that, if both channels have communication constraints and $C(z)$ and $G(z)$ are anti-diagonal transfer matrices, then (3.66) is trivially satisfied when (3.64)-(3.65) hold.

An immediate consequence of Theorem 3.4 is the following:

Corollary 3.3 (Arbitrarily poor performance) *Consider the feedback system in Figure 3.6 and suppose that Assumptions 3.2 and 3.3 hold. Assume that* (3.64)-(3.66) *are satisfied and define*

$$\mathcal{G} \triangleq \left\{(\Gamma_1, \Gamma_2) \in \mathbb{R}_0^2 : \ \Gamma_1 \textit{ and } \Gamma_2 \textit{ achieve equality in } (3.66)\right\}. \tag{3.72}$$

Then, for any fixed $F_1(z), F_2(z) \in \mathcal{U}_\infty$ *and any* $(\bar{\Gamma}_1, \bar{\Gamma}_2) \in \mathcal{G}$

$$\lim_{(\Gamma_1,\Gamma_2)\to(\bar{\Gamma}_1,\bar{\Gamma}_2)} \inf_{\substack{\sigma_i^2 \in \mathbb{R}_0^+ \\ \gamma_i \le \Gamma_i \\ i\in\{1,2\}}} \sigma_e^2 = \infty. \tag{3.73}$$

Proof: Provided Assumptions 3.2 and 3.3 hold, we have that, for any $F_i(z) \in \mathcal{U}_\infty$ and $\sigma_i^2 \in \mathbb{R}_0^+$, σ_e^2 exists, is finite and obeys

$$\sigma_e^2 = ||S(z)\Omega_r(z)||_2^2 + \sum_{i=1}^{2} \sigma_i^2 \, ||S(z)G(z)\varepsilon_i F_i(z)||_2^2 \, . \tag{3.74}$$

By definition of $\mathcal{G}$ and the proof of Theorem 3.4 we know that, if $(\Gamma_1, \Gamma_2) \to (\bar{\Gamma}_1, \bar{\Gamma}_2)$, the the only way to satisfy $\gamma_i \le \Gamma_i$ is by making $\sigma_i^2 \to \infty$. Since Assumption 3.3 holds and $F_i(z) \in \mathcal{U}_\infty$, it follows that $S(z)G(z)\varepsilon_i F_i(z) \neq 0$ and, hence, $\sigma_e^2 \to \infty$. This completes the proof. □□□

Although very simple, Corollary 3.3 has an interesting implication: for any given full-MIMO controller, and no matter how the coding system is chosen, there exist sufficiently poor channels which render the performance of the resulting partly networked closed loop arbitrarily bad. In these cases, *any* stabilizing decentralized controller (that does not use the non-transparent communication channels) will provide better performance. As will become clear in the examples of Section 3.6, the channel signal-to-noise ratios do not need to be artificially low for the partly MIMO control architecture to perform poorly. Indeed, depending on plant and controller features, the left hand side of (3.66) may be large, thus requiring high signal-to-noise ratios to be satisfied.

Remark 3.7 *We emphasize that our conclusions are tied to the fact that the controller $C(z)$ is assumed fixed. If, on the contrary, $C(z)$ is also taken as a design parameters, then the resulting partly networked control architecture will certainly outperform any given decentralized design.*

3.5.2 Optimal coding system design

In this section we show how to design the coders $F_i(z)$ and the channel noise variances σ_i^2, so as to optimize performance subject to constraints on the channel signal-to-noise ratios γ_i. To do so, we will introduce a simplifying assumption. From Theorem 3.4 we know that the constraints on γ_i are compatible with MSS if and only if (3.64)-(3.66) are satisfied. The first two conditions establish restrictions on Γ_i that depend only on the plant model and the full-MIMO controller design. As such, those conditions can be readily imposed on Γ_i. However, the third condition depends on $F_i(z)$, which is a design choice. Thus, it is complicated to explicitly characterize the set of all Γ_is (and $F_i(z)$s) that satisfy (3.66) for a given plant and full-MIMO controller design. We will thus assume that the following holds:

Assumption 3.4 (Simplifying assumption) *The upper bounds Γ_i on γ_i, $i \in \{1, 2\}$, are high enough so that, for every $F_i(z) \in \mathcal{U}_\infty$, $i \in \{1, 2\}$, the satisfaction of* (3.66) *is guaranteed when* (3.64) *and* (3.65) *hold.* □□

Provided Assumptions 3.2(a) and 3.3 hold, and $F_1(z), F_2(z) \in \mathcal{U}_\infty$, we have (see the proof of Theorem 3.4)

$$\gamma_i - B_i = \frac{||A_i(z)\Omega_r(z)||_2^2}{\sigma_i^2} + \sum_{\substack{j=1\\ j\neq i}}^{2} \frac{\sigma_j^2 \, ||A_i(z)G(z)F(z)\varepsilon_j||_2^2}{\sigma_i^2} \geq \frac{||A_i(z)\Omega_r(z)||_2^2}{\sigma_i^2}. \tag{3.75}$$

Thus,

$$\sigma_i^2 \geq \frac{||A_i(z)\Omega_r(z)||_2^2}{\gamma_i - B_i} \tag{3.76}$$

and, as a consequence (recall (3.74)),

$$\sigma_e^2 \geq ||S(z)\Omega_r(z)||_2^2 + \sum_{i=1}^{2} \frac{||A_i(z)\Omega_r(z)||_2^2 \, ||S(z)G(z)\varepsilon_i F_i(z)||_2^2}{\gamma_i - B_i}. \tag{3.77}$$

It is clear that, if the signal-to-noise ratios γ_1 and γ_2 are high enough, then the bound in (3.77) will be tight. If this is the case, then the coders and channel noise variances that minimize σ_e^2 also minimize

$$J \triangleq \sum_{i=1}^{2} \frac{||A_i(z)\Omega_r(z)||_2^2 \, ||S(z)G(z)\varepsilon_i F_i(z)||_2^2}{\gamma_i - B_i}. \tag{3.78}$$

We have the following result:

Theorem 3.5 (Suboptimal coders) *Consider the functional J defined in (3.78) and suppose that Assumption 3.3(a-c) hold. If γ_1 and γ_2 are fixed at values such that $\gamma_i > B_i$, $i \in \{1,2\}$, then the filters $F_1^o(z)$ and $F_2^o(z)$, defined via*

$$(F_1^o(z), F_2^o(z)) \triangleq \arg \inf_{\substack{F_1(z)\in\mathcal{U}_\infty \\ F_2(z)\in\mathcal{U}_\infty}} J, \tag{3.79}$$

are such that

$$\left|F_i^o(e^{j\omega})\right|^4 = \alpha_i \frac{M_i(e^{j\omega})M_i(e^{j\omega})^H}{\left(S(e^{j\omega})G(e^{j\omega})\varepsilon_i\right)^H S(e^{j\omega})G(e^{j\omega})\varepsilon_i}, \quad i \in \{1,2\}, \tag{3.80}$$

where

$$M_1(z) \triangleq C_{12}(z)\varepsilon_2^T S(z)\Omega_r(z), \quad M_2(z) \triangleq C_{21}(z)\varepsilon_1^T S(z)\Omega_r(z), \tag{3.81}$$

and α_i is an arbitrary positive constant. With this choice,

$$J = \sum_{i=1}^{2} \frac{\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \sqrt{M_i(e^{j\omega})M_i(e^{j\omega})^H \left(S(e^{j\omega})G(e^{j\omega})\varepsilon_i\right)^H S(e^{j\omega})G(e^{j\omega})\varepsilon_i}\, d\omega\right)^2}{\gamma_i - B_i}. \tag{3.82}$$

Proof: Fact A.1 and the Cauchy Schwartz inequality imply that, for $i \in \{1,2\}$,

$$\begin{aligned}
&||A_i(z)\Omega_r(z)||_2^2 \, ||S(z)G(z)\varepsilon_i F_i(z)||_2^2 \\
&= \left|\left|\sqrt{A_i(z)\Omega_r(z)\left(A_i(z^{-1})\Omega_r(z^{-1})\right)^T}\right|\right|_2^2 \left|\left|\sqrt{F_i(z^{-1})^T \left(S(z)G(z^{-1})\varepsilon_i\right)^T S(z)G(z)\varepsilon_i F_i(z)}\right|\right|_2^2 \\
&\geq \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \sqrt{M_i(e^{j\omega})M_i(e^{j\omega})^H \left(S(e^{j\omega})G(e^{j\omega})\varepsilon_i\right)^H S(e^{j\omega})G(e^{j\omega})\varepsilon_i(e^{j\omega})}\, d\omega\right)^2
\end{aligned} \tag{3.83}$$

where we have used the definitions of $M_i(z)$ and $A_i(z)$ (see (3.63) and (3.81)). In (3.83), equality holds if and only if there exists $\alpha_i \in \mathbb{R}^+$ such that

$$\sqrt{A_i(e^{j\omega})\Omega_r(e^{j\omega})\left(A_i(e^{j\omega})\Omega_r(e^{j\omega})\right)^H} = \frac{1}{\sqrt{\alpha_i}}\sqrt{F_i(e^{j\omega})^H \left(S(e^{j\omega})G(e^{j\omega})\varepsilon_i\right)^H S(e^{j\omega})G(e^{j\omega})\varepsilon_i F_i(e^{j\omega})},$$

for every $\omega \in [-\pi, \pi]$. Since B_i does not depend on $F_1(z)$ or $F_2(z)$, the theorem follows. □□□

Theorem 3.5 shows how to choose filters $F_1(z)$ and $F_2(z)$ in $\mathcal{U}_\infty$ so as to minimize J when γ_1 and γ_2 are given. Since each term in J is a decreasing function of either γ_1 or γ_2, it follows that, in order to minimize J, it suffices to make γ_i equal to the largest admissible value, i.e.,

$\gamma_i = \Gamma_i$. If $\Gamma_i > B_i$ and Assumption 3.4 holds, then this choice for γ_i guarantees that the channel noise variances σ_i^2 (calculated via (3.70)) belong to $\mathbb{R}_0^+$. Accordingly, the suggested choices for $F_i(z)$ and γ_i are such that, provided $\Gamma_i > B_i$ and Assumption 3.4 holds, the MSS of the partly networked control architecture is guaranteed and a lower bound on σ_e^2, namely J, is minimized subjext to $\gamma_i \leq \Gamma_i$. If Γ_i are high enough, then one can expect the proposed coder and channel noise variance choices to be close to the actual optimal choices.

The procedure outlined above allows one to design coding systems having the impact of communication constraints on overall closed loop performance in mind. An interesting feature of the proposed filters is that they do not depend on the channel signal-to-noise ratios. This allows one to conjecture that the optimal filters will perform well for a large class of communication channels.

3.6 Examples

We next present examples that illustrate the main results in this chapter. We will employ signal-to-noise ratio constrained additive i.i.d. noise channels *to model* digital channels that are able to transmit b bits per sample and that are preceded by a finite uniform quantizer (see, e.g., [83,181] and also [55, 188]). Within this model, which is described in Appendix B, the signal-to-noise ratio γ of the additive channel can be related to the number of bits b in the quantizer via

$$\gamma = \frac{3}{\alpha^2}\left(2^b - 1\right)^2, \tag{3.84}$$

where α is the quantizer loading factor (see [83]).

We will start by illustrating the results in Section 3.4.

Example 3.1 (SISO case) *Consider the feedback system of Figure 3.4 with plant and (nominal) controller given by*

$$G(z) = \frac{1}{z - 0.8}, \qquad C(z) = \frac{z - 0.8}{z - 1}. \tag{3.85}$$

Assume that the output disturbance d is zero and that the reference r has a spectral factor given by

$$\Omega_r(z) = \frac{0.02z}{z - 0.9}. \tag{3.86}$$

Suppose that the signal-to-noise ratio constrained additive i.i.d. noise channel models a uniform quantizer having a fixed loading factor of $\alpha = 4$, and $b \in \{1, \cdots, 8\}$ bits.

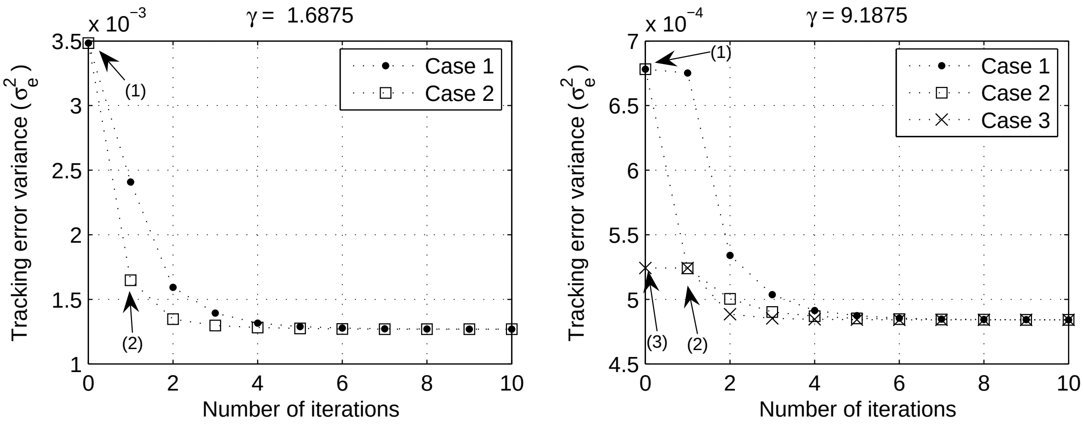

Figure 3.7: Tracking error as function of the number of iterations in Algorithm 3.1 (see Example 3.1 for details).

Figure 3.7 shows the stationary tracking error variance σ_e^2 *as a function of the number of iterations in Algorithm 3.1 for two representative values of the quantizer signal-to-noise ratio:* $\gamma = 1.6875$ *and* $\gamma = 9.1875$*, which correspond to* $b = 2$ *and* $b = 3$*, respectively (see* (3.84)*). Cases 1 and 2 refer to iterations that start with* $A(z) = 1$ *and* $F(z) = 0$*. In Case 1 we initially fixed* $A(z)$*, whereas in Case 2 we start by fixing* $F(z)$*. Case 3 refers to iterations that start with the choice for the filters suggested in [160]. We note that the filter* $F(z)$ *proposed in [160] satisfies the conditions of Theorem 3.2 if and only if* $\gamma > 4.42$*. (Accordingly, we omitted Case 3 in Figure 3.7 when* $\gamma = 1.6875$*.) It can be seen that rapid convergence of Algorithm 3.1 occurs and, more interestingly, that the limiting performance does not depend on the order in which the filters are calculated, or on the initial values chosen. Thus, local minima related issues do not seem to play a role in this example.*

In Figure 3.7 we have identified three points. The first of these (point (1)) refers to the performance achieved without coding ($F(z) = 0$ *and* $A(z) = 1$*). The second (point (2)) refers to the performance achieved when employing the optimal coding system proposed in [55]. The third (point (3)) refers to the performance achieved using the filters suggested in [160].*

The results show that coding is, indeed, necessary to achieve the best possible loop performance. (Compare point (1) with, e.g., the value of σ_e^2 *after 10 iterations.) Furthermore, Algorithm 3.1 yields coding systems that perform better than our previous proposals in [55, 160]. (Compare points (2) and (3) with the limiting value for* σ_e^2*.) It is also interesting to mention that, for* $b > 4$*, the performance provided by the filters in [160] is substantially closer to the*

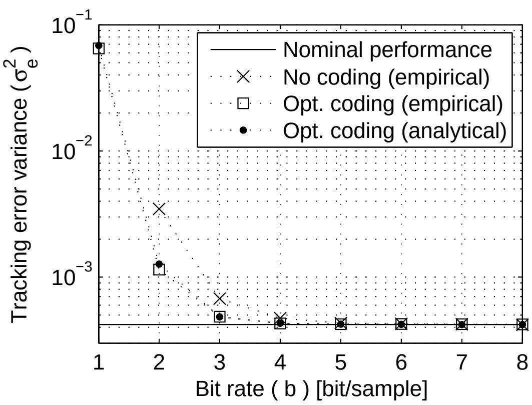

Figure 3.8: Tracking error as function of the channel bit rate for Example 3.1.

limiting value of σ_e^2 *than in the case shown in Figure 3.7. This suggest, as mentioned before, that the filters in [160] may provide good starting points for the iterative procedure proposed here.*

We end this example by studying the behavior of the stationary tracking error variance as a function of the channel bit rate b. The results are presented in Figure 3.8, where "Nominal performance" refers to the performance achieved by the nominal loop (without quantization), "No coding (empirical)" refers to simulated results[11] *when no coding is employed (i.e., when* $A(z) = 1$ *and* $F(z) = 0$*), "Opt. coding (empirical)" refers to simulated results obtained with the filters suggested by Algorithm 3.1 (after* 10 *iterations), and "Opt. coding (analytical)" refers to the corresponding predictions made using the simplified noise model for quantization described in Appendix B. One can see that, as expected, the effects of quantization vanish as* $b \to \infty$*. Interestingly, the predictions made using the model turn out to be very accurate for every bit rate: indeed, for* $b \geq 3$ *the relative errors are of less than* 1% *and, for* $b \in \{1, 2\}$*, the relative errors are around* 8%*. (We note that* $F(z) = 0$ *does not satisfy* (3.14) *for* $b = 1$*. Accordingly, we have omitted the non coded simulation results for* $b = 1$*.)* □□

The results presented above suggest that one can design feedback quantizers using the simplified model for quantization described in Appendix B, together with the tools developed in this chapter. Of course, such a design is heuristic in that the model for quantization is only an approximation. In particular, stability guarantees can not be provided (see also Section III.B in [118]). Nonetheless, from a practical point of view, our results go beyond those documented

[11] All simulations use an actual (undithered) uniform quantizer with $L = 2^b - 1$ levels. For each b, the results correspond to the average of 200 simulations (each one 10^5 samples long and using a different reference realization).

in, e.g., [24, 25, 101]. In that work, a precise deterministic stability analysis is presented when the coding system is constrained to be a delta-modulator (or variations thereof;[12] see, e.g., [83]), but no performance issues are addressed. Another recent publication that deals with the design of feedback quantizers for control systems is [112]. In that work, the authors propose a coding architecture similar to the one proposed in this chapter, but restrict the quantizers to have infinitely many levels and a prespecified quantization step. Interestingly, the optimal coder in [112] (which focuses on minimizing a time domain functional) turns out to have a structure that is a special case of the architecture considered here.

We next illustrate the results of Section 3.5 by two examples. We will assume for simplicity that both channels have the same bit-rates $b = b_1 = b_2$.[13]

Example 3.2 (Simple MIMO case) *Consider the feedback system of Figure 3.6 with plant*

$$G(z) = \begin{bmatrix} \dfrac{0.6}{(z-0.8)} & \dfrac{0.4}{(z-0.8)} \\ \dfrac{1}{(z-0.5)} & \dfrac{1}{(z-0.5)} \end{bmatrix} \tag{3.87}$$

and controller

$$C(z) = \begin{bmatrix} \dfrac{5(z-0.8)}{z-1} & \dfrac{-2(z-0.5)}{z-1} \\ \dfrac{-5(z-0.8)}{z-1} & \dfrac{3(z-0.5)}{z-1} \end{bmatrix}. \tag{3.88}$$

Assume that the reference r has a spectral factor given by

$$\Omega_r(z) = \frac{0.0049627(z+0.9934)}{(z^2-1.97z+0.9802)} I, \tag{3.89}$$

and that the signal-to-noise ratio constrained additive i.i.d. noise channels model uniform quantizers having a fixed loading factor of $\alpha = 4$ and an equal number of bits b. For comparison purposes, also consider the following stabilizing decentralized controller for $G(z)$:

$$C_d(z) = \begin{bmatrix} \dfrac{1.3333(z-0.8)}{(z+0.8)(z-1)} & 0 \\ 0 & \dfrac{0.8(z-0.5)}{(z+0.8)(z-1)} \end{bmatrix}. \tag{3.90}$$

[12] i.e., unlike our proposal, the general multi-bit case is not treated in [24, 25, 101].

[13] We note that, since the proposed coding filters are independent of the channel signal-to-noise ratios, our results opens the possibility of investigating optimal communication resource allocation schemes, as discussed in, e.g., [82, 131, 188]. In particular, one can envisage alternative bit allocation schemes: for example, given a total budget of b_T bits, one can choose b_1 and b_2 so as to minimize J, while respecting $b_T = b_1 + b_2$ (we refer the reader to [83] for details).

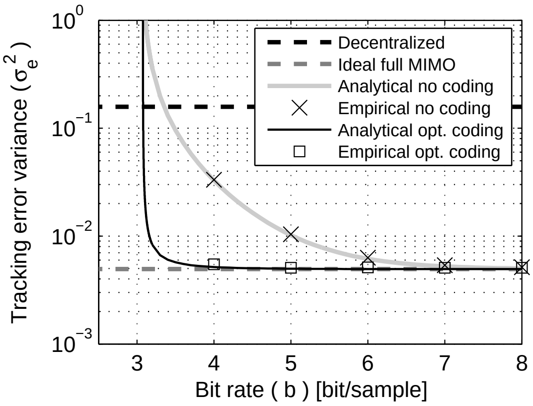

Figure 3.9: Tracking error variance as a function of the per-channel bit rate (Example 3.2).

Figure 3.9 shows the tracking error variance in the partly networked MIMO architecture studied as a function of the per-channel bit-rate b in several situations: "Analytical no coding" refers to the performance predicted by (3.74) *and* (3.70) *when no coding is employed; "Empirical no coding" refers to simulated*[14] *performance when no coding is considered; "Analytical opt. coding" and "Empirical opt. coding" refer to analytical and simulated performance when the coders suggested by Theorem 3.5 are employed. For comparison purposes, Figure 3.9 also shows the non networked full MIMO performance ("Ideal full MIMO") and the performance achieved when using* $C_{d1}(z)$ *("Decentralized"). The results allow one to conclude that, in this case, the benefits of coding are significative. Indeed, for* $b = 4$ *and no coding, the performance achieved by the partly networked MIMO controller is more than* 5 *times worse than that achieved by the ideal full-MIMO control scheme. On the other hand, if optimal coding is employed, then the performance deterioration with respect to the non-networked situation is only* 3.5%. *We also note that the match between the predictions made using the simplified model for quantization in Appendix B and the results of non idealized simulations is remarkably good.*

As expected, non-networked full MIMO performance is recovered as $b \to \infty$*, irrespective of the coders used. It is also apparent that the non coded networked full MIMO architecture should be preferred to the decentralized one for* $b > 3.36$*. On the other hand, the optimally coded networked MIMO architecture provides significant improvement in performance for* $b > 3.20$*,*

[14]The analytical results are given for fractional values of b (this simply amounts to choosing non integer values for b in (3.84)). Of course, simulation results are presented for integer values of b only. All simulations use actual b-bit uniform quantizers in both channels. For each b, the results correspond to an average of 200 simulations (each one 10^5 samples long and using a different reference realization).

when compared to the decentralized architecture. It is also interesting to note that, if $b \to 3.07$*, then the performance becomes arbitrary poor no matter what the coding is. This is in agreement with Corollary 3.3 as a straightforward calculation based on* (3.84) *reveals.* □□

We will next examine a second MIMO example, where the plant is much harder to control.

Example 3.3 (Hard MIMO case) *We consider a classical* 2×2 *model of a distillation process, as described in [167]. Assuming properly scaled variables and a zero order hold at the plant input, it is possible to derive the following discrete time model:*

$$G(z) = \frac{1}{z - 0.9868} \begin{bmatrix} 1.1629 & -1.1444 \\ 1.4331 & -1.4516 \end{bmatrix}. \tag{3.91}$$

Although simple in appearance, this plant is extremely difficult to control. This is due to the fact that it is almost singular [167].

Consider the full MIMO controller

$$C(z) = \frac{z - 0.9868}{z - 1} \begin{bmatrix} 30.1564 & -23.7729 \\ 29.7712 & -24.1582 \end{bmatrix}, \tag{3.92}$$

and the decentralized controller

$$C_d(z) = \begin{bmatrix} \frac{0.261(z-0.734)}{(z-1)} & 0 \\ 0 & \frac{-0.375(z-0.6979)}{(z-1)} \end{bmatrix}. \tag{3.93}$$

We note that $C_d(z)$ *is a discretized version of a carefully tuned MIMO PI controller for* $G_2(z)$*, as proposed in [167]. Also assume that the reference spectral density is given by* $\Omega_r(z) = 0.1(z - 0.9)^{-1}I$*, and that the quantizers loading factors are given by* $\alpha = 5$*.*[15]

Figure 3.10 shows the tracking error variance provided by the partly networked MIMO architecture as a function of the per-channel bit-rate, in the same cases as those in Figure 3.9. The same qualitative behavior as in Example 3.2 is observed. However, the performance gains arising from coding are negligible in the present situation. This is due to the fact that the optimal coders have an almost flat frequency response.

An interesting feature of this example is that the performance achieved by the partly networked full-MIMO controller may become significatively worse than that achieved by the decentralized controller $C_d(z)$*, even for relatively high bit rates (namely, for* $b \to 7.15$*). Again, this is consistent with Corollary 3.3 and* (3.84). □□

[15] For this plant, saturation may compromise loop stability even when transparent channels are employed. To avoid this, we extended the quantizer dynamic range so that quantizer overflow (i.e., saturation) probability becomes completely negligible.

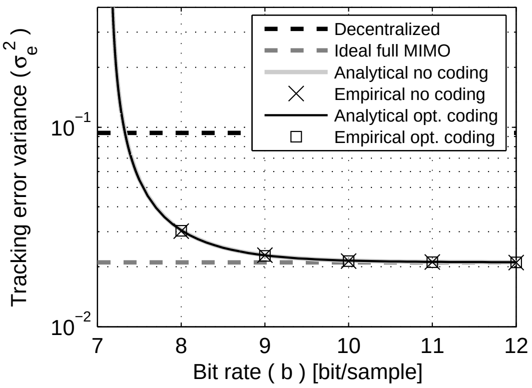

Figure 3.10: Tracking error variance as a function of the per-channel bit rate (Example 3.3).

3.7 Summary

This chapter has presented a methodology to design feedback based coding schemes that minimize the impact of signal-to-noise ratio constraints on closed loop performance. Our results show that feedback coding schemes are beneficial when compared to simpler schemes documented in the literature. We also characterized the smallest channel signal-to-noise ratio compatible with MSS. That study revealed that, for a given channel signal-to-noise ratio, the class of plants that are stabilizable when feedback coding is employed is significatively larger than the class of plants that are stabilizable when no feedback in the coder is employed. This result opens the door to investigate other LTI control and feedback coding architectures and the associated signal-to-noise ratio requirements for MSS. This is done in Chapter 4.

We also extended our framework to deal with two-by-two MIMO control problems. In that setting, we investigated how additional channels may improve the performance of standard decentralized control architectures. An interesting byproduct of our analysis is that, in some cases, the additional channels need to be of high quality if the (partly) networked full-MIMO architecture is to outperform a given decentralized design.[16]

An interesting extension of the results in this chapter lies in addressing the joint control and coding architecture design problem (see Chapter 4). Also, the general n-by-n MIMO case, and the problem of how to actually design partly networked full-MIMO controllers (and not just the filters in a coding scheme) are interesting subjects for future study.

[16]Of course, if one optimally designs a partly networked full-MIMO controller, then the resulting architecture will always outperform any decentralized design, since the latter scheme is a special case of the former.

Chapter 4

Optimal Control and Coding System Design

4.1 Introduction

This chapter extends part of the the results in Chapter 3 to a general control and coding architecture that employs a signal-to-noise ratio constrained additive i.i.d. noise channel in the feedback path. We start by studying signal-to-noise ratio requirements for MSS. In particular, we examine the role played by the degrees of freedom of the proposed architecture in the minimal channel signal-to-noise ratio compatible with MSS. Consistent with the results of Chapter 3, our results show that use of feedback in the coding architecture is essential to reduce the signal-to-noise ratio requirements for MSS. We also study the interplay between the performance of the proposed architecture and the channel signal-to-noise ratio constraint. This study yields an answer to the question of what is the best achievable performance in a general LTI networked control architecture when a signal-to-noise ratio constraint is present in the feedback path. The converse question, i.e., what are the requirements on the channel signal-to-noise ratio that guarantee a given level of performance, is also addressed. Our results are given in terms of the solutions to convex optimization problems, and can be readily used for numerical computations based on standard software (see, e.g., [57]).

As already discussed in Section 1.1, this chapter extends the results available in the literature in several ways: (i) we consider a general architecture, with no structural constraints besides those imposed by the communication constraints (compare with the results in Chapter 3 and also with the results in [48, 162]), (ii) our results can accommodate (almost) any SISO plant model (compare with [48]), and (iii) our framework can accommodate cases where arbitrarily many noise sources are present (compare with [48, 162]).

The rest of this chapter is organized as follows: Section 4.2 presents the architecture of interest and our working assumptions. Section 4.3 studies conditions on the channel signal-to-noise ratio that are necessary and sufficient for the MSS of the considered networked control system. In Section 4.4 we consider a general control problem where a signal-to-noise ratio constraint is present, and we study the interplay between that constraint and the variance of a specific output in the loop. These results are then used in Section 4.5 to study the control and coding architecture proposed in Section 4.2, as well as a relevant special case. A simple numerical example is presented in Section 4.6. Section 4.7 draws conclusions.

4.2 Problem Definition

In this chapter we will consider the general networked control architecture depicted in Figure 4.1. In that figure, d is a disturbance signal, r is the reference signal, and q is the noise in an additive i.i.d. noise channel subject to a signal-to-noise ratio constraint Γ (see Definition 3.1).

As was the case of Chapter 3, we will focus on control and coding architectures that use only LTI filters and no additional communication channels other than those explicitly shown in Figure 4.1. With these constraints in mind, we propose the control and coding architecture depicted in Figure 4.2. In that figure, $M(z)$ and $F(z)$ are LTI systems in $\mathcal{R}_p^{1\times 2}$ that are to be designed and

$$v = F(z)\begin{bmatrix} w \\ y \end{bmatrix} \triangleq F_1(z)w + F_2(z)y, \tag{4.1}$$

and

$$u = M(z)\begin{bmatrix} r \\ w \end{bmatrix} \triangleq M_1(z)r + M_2(z)w. \tag{4.2}$$

For future reference we also define

$$K(z) \triangleq \begin{bmatrix} F(z) & M(z) \end{bmatrix}. \tag{4.3}$$

The proposed control and coding architecture generalizes the coding architecture presented in Chapter 3 and, as the reader can easily verify, constitutes the most general architecture

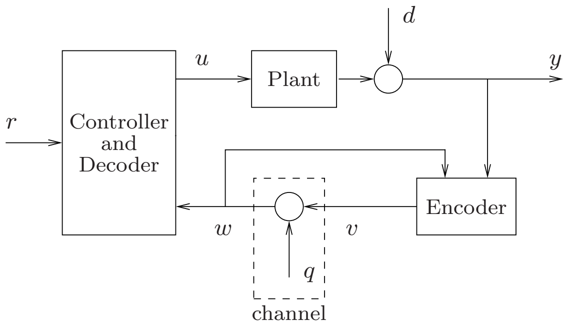

Figure 4.1: General networked control system.

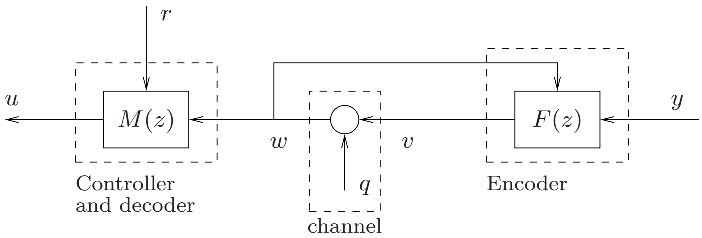

Figure 4.2: Proposed control and coding scheme.

that uses only LTI filters, feedback around the additive i.i.d. noise channel, and no additional communication channels. We would like to emphasize that, as the results in the forthcoming sections will make explicit, feedback from the output of the channel to the input of the filter $F(z)$ is a key feature of the proposed coding scheme. We also recall that, in our setup, the signal w is available with one step delay at the encoder side. Thus, $F_1(z) \in \mathcal{R}_{sp}$; this constraint will be enforced throughout the rest of this chapter.

The networked control system that arises when the proposed control and coding scheme is used in the loop of Figure 4.1 can be written as shown in Figure 4.3, where $G(z)$ is the plant transfer function. In the remainder of this chapter, we will work under the following assumption that, as discussed in Chapter 1, encompasses many cases of interest:

Assumption 4.1 (Plant, signals and initial states)

(a) The plant transfer function $G(z)$ belongs to $\mathcal{R}_{sp}^{1\times 1}$, is non-zero, and has a stabilizable and detectable underlying realization $(A_G, B_G, C_G, 0)$.

(b) The signals r and d are second order mutually uncorrelated wss scalar processes with rational PSDs. At least one signal, r or d, is non-zero. If r or d are non-zero, then they admit spectral factors in $\mathcal{U}_\infty$.

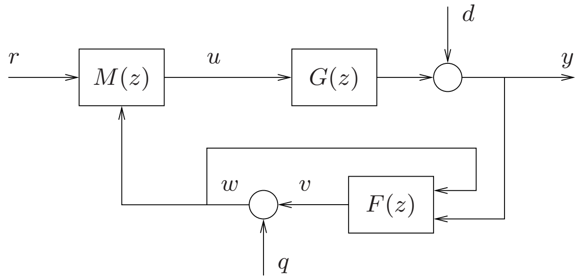

Figure 4.3: Networked control system that arises when using the control and coding architecture of Figure 4.2 in the scheme of Figure 4.1.

(c) The initial states of all the filters in Figure 4.3 (including plant model) are jointly second order random variables.

□□

The remainder of this chapter focuses on studying the interplay between stationary channel signal-to-noise ratio constraints, i.e., constraints on

$$\gamma \triangleq \frac{\sigma_v^2}{\sigma_q^2}, \tag{4.4}$$

and the performance of the control architecture depicted in Figure 4.3. As was the case for the architecture studied in Chapter 3, we first study signal-to-noise ratio requirements for MSS. Then, we consider the problem of designing the proposed control and coding architecture having performance in mind.

4.3 Mean Square Stability

Given our framework, the designer has the freedom to choose both $K(z)$ and the channel noise variance σ_q^2 which, of course, has to be finite. On the other hand, as long as Assumption 4.1 holds and $\sigma_q^2 \in \mathbb{R}_0^+$, the networked control system in Figure 4.3 is MSS if and only if $K(z)$ is such that the feedback loop in Figure 4.3 is internally stable and well-posed (see Chapter 2). Accordingly, if we define[1]

$$\mathcal{S} \triangleq \{K(z) \in \mathcal{R}_{sp} \times \mathcal{R}_p \times \mathcal{R}_p \times \mathcal{R}_p : \text{the feedback loop in Figure 4.3}$$
$$\text{is internally stable and well-posed}\}, \tag{4.5}$$

[1]In this definition, we make the constraint $F_1(z) \in \mathcal{R}_{sp}$ explicit.

then we should restrict our attention to $K(z) \in \mathcal{S}$ and $\sigma_q^2 \in \mathbb{R}_0^+$.

As mentioned before, feedback from w to $F(z)$ plays a central role in the performance of the coding scheme of Figure 4.2. To illustrate this, we will sometimes consider a special case of the proposed control and coding scheme where $F_1(z) = 0$. This case will be referred to as a control and coding architecture *without feedback*. For future reference we define

$$\mathcal{S}_C \triangleq \left\{ K(z) \in \mathcal{S} : F(z) = \begin{bmatrix} 0 & F_2(z) \end{bmatrix} \right\}. \tag{4.6}$$

We now give characterizations of $\mathcal{S}$ and $\mathcal{S}_C$.

Corollary 4.1 (Characterization of $\mathcal{S}$ and $\mathcal{S}_C$) *Consider the feedback system of Figure 4.3 and suppose that Assumption 4.1(a) holds. Then:*

1. *$K(z) \in \mathcal{S}$ if and only if $M(z)$ has a detectable and stabilizable underlying realization (A_M, B_M, C_M, D_M) such that*[2]

$$\left(\begin{bmatrix} A_G & B_G C_M & 0 \\ 0 & A_M & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} B_G D_M \varepsilon_2 \\ B_M \varepsilon_2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ C_G & 0 & 0 \end{bmatrix}, 0 \right) \tag{4.7}$$

is detectable and stabilizable, and

$$F(z) = (X_i(z) - Q(z)N_i(z))^{-1}(Y_i(z) - Q(z)D_i(z)) \begin{bmatrix} z^{-1} & 0 \\ 0 & 1 \end{bmatrix}, \tag{4.8}$$

where $Q(z)$ is a free parameter in $\mathcal{RH}_\infty$ and $X_i(z), N_i(z), Y_i(z)$ and $D_i(z)$ are as in Theorem 2.2 with

$$P_{22}(z) = \begin{bmatrix} z^{-1} \\ G(z)M_2(z) \end{bmatrix}. \tag{4.9}$$

2. *$K(z) \in \mathcal{S}_C$ if and only if $F_2(z)$ has a detectable and stabilizable underlying realization (A_2, B_2, C_2, D_2) such that*

$$\left(\begin{bmatrix} A_G & 0 \\ B_2 C_G & A_2 \end{bmatrix}, \begin{bmatrix} B_G \\ 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ D_2 C_G & C_2 \end{bmatrix}, 0 \right) \tag{4.10}$$

is detectable and stabilizable, and

$$M(z) = (X_i(z) - Q(z)N_i(z))^{-1}(Y_i(z) - Q(z)D_i(z)), \tag{4.11}$$

[2]When constructing B_M and D_M, recall that $M(z)$ obeys (4.2). In (4.7), ε_i refers to the i-th vector of the canonical basis of $\mathbb{R}^2$.

where $Q(z)$ is a free parameter in $\mathcal{RH}_\infty$ and $X_i(z), N_i(z), Y_i(z)$ and $D_i(z)$ are as in Theorem 2.2 with

$$P_{22}(z) = \begin{bmatrix} 0 \\ F_2(z)G(z) \end{bmatrix}. \tag{4.12}$$

Proof:

1. Since $F_1(z)$ must belong to $\mathcal{R}_{sp}$, we define

$$\bar{F}(z) \triangleq F(z) \begin{bmatrix} z & 0 \\ 0 & 1 \end{bmatrix} \in \mathcal{R}_p^{1\times 2}. \tag{4.13}$$

With this notation, the result follows from Theorem 2.2 and Lemma 2.1, where $\bar{F}(z)$ plays the role of $K(z)$ and $P_{22}(z)$ is as in (4.9). We also note that, since $G(z) \in \mathcal{R}_{sp}$, we have that $\det(X_i(\infty) - Q(\infty)N_i(\infty)) = \det(X_i(\infty)) \neq 0$ for every $Q(z) \in \mathcal{RH}_\infty$, as required (recall that $X_i(z)$ is biproper).

2. In this case, the result follows from Theorem 2.2 and Lemma 2.1, with $M(z)$ playing the role of $K(z)$ and $P_{22}(z)$ as in (4.12). As above, $G(z) \in \mathcal{R}_{sp}$ guarantees $\det(X_i(\infty)$ $-Q(\infty)N_i(\infty)) \neq 0$ for every $Q(z) \in \mathcal{RH}_\infty$.

□□□

We are now ready to present the first main result in this chapter. Namely, a characterization of the minimal channel signal-to-noise ratio that guarantees the MSS of control and coding architectures with no constraints on $K(z)$. The case of architectures without feedback is treated later in this section.

Theorem 4.1 (Minimal signal-to-noise ratio for MSS) *Consider the feedback system in Figure 4.3, where q is the noise in an additive i.i.d. noise channel, and suppose that Assumption 4.1 holds. If $G(z)$ is unstable, then*

$$\gamma_{\inf} \triangleq \inf_{\substack{K(z)\in\mathcal{S} \\ \sigma_q^2\in\mathbb{R}_0^+}} \gamma = \left(\prod_{i=1}^{n_p} |p_i|^2\right) - 1. \tag{4.14}$$

Unless $d = 0$ and $G(z)$ has no poles on the unit circle, $\gamma_{\inf}$ is not achievable, but can be approached with arbitrary precision with a sufficiently large (but finite) $\sigma_q^2 \in \mathbb{R}_0^+$ and $K(z)$ as in Part 1 of Corollary 4.1 with $M(z) \in \mathcal{RH}_\infty$ and (recall the definition of $\mathcal{K}_\varepsilon\{\cdot\}$ in (2.12))

$$Q(z) = -z\mathcal{K}_\varepsilon\{\xi_p(z)D(z)\}^{-1}(\xi_p(\infty) - X_i(z)\xi_p(z)D(z))\begin{bmatrix} 1 & 0 \end{bmatrix}, \tag{4.15}$$

where $D(z) \in \mathcal{RH}_\infty$ is any denominator of $G(z)M_2(z)$ in a coprime factorization of $G(z)M_2(z)$ over $\mathcal{RH}_\infty$, and ε is a sufficiently small (but non-zero) real in $(0,1]$. If the plant has no poles on the unit circle, then $\mathcal{K}_\varepsilon\{\cdot\}$ in (4.15) is redundant and γ_{inf} is achievable if and only if $d = 0$. To do so, it suffices to choose $K(z)$ as above, with $M_1(z)\Omega_r(z) = 0$ and any $\sigma_q^2 \in \mathbb{R}^+$.

Proof: Consider $M(z) \in \mathcal{R}_p$ with stabilizable and detectable underlying realization and such that (4.7) is detectable and stabilizable. Also consider $N(z), D(z) \in \mathcal{RH}_\infty$, coprime in $\mathcal{RH}_\infty$ and with $D(z)$ biproper, such that $G(z)M_2(z) = N(z)D(z)^{-1}$. Then,

$$N_d(z) \triangleq \begin{bmatrix} z^{-1}D(z) \\ N(z) \end{bmatrix}, \quad D_d(z) \triangleq D(z), \quad N_i(z) \triangleq \begin{bmatrix} z^{-1} \\ N(z) \end{bmatrix}, \quad D_i(z) \triangleq \begin{bmatrix} 1 & 0 \\ 0 & D(z) \end{bmatrix} \tag{4.16}$$

are such that $D_d(z)$ and $D_i(z)$ are biproper, (2.42) is satisfied with $P_{22}(z)$ as in (4.9), and there exist $X_i(z)$, $Y_i(z)$, $X_d(z)$, $Y_d(z) \in \mathcal{RH}_\infty$ satisfying (2.43). (This follows from the definition of coprimeness; see, e.g., [46, 182].) Thus, we can use Theorem 2.2 to conclude that all transfer functions from q to v in Figure 4.3, say $T_{qv}(z)$, that are achievable with $K(z) \in \mathcal{S}$ are given by

$$T_{qv}(z) = Y_i(z)N_d(z) - D(z)Q(z)N_i(z) = X_i(z)D(z) - 1 - D(z)Q(z)N_i(z), \tag{4.17}$$

where $Q(z) \in \mathcal{RH}_\infty$ and the second equality follows from (2.43). We also note that (2.43) implies that, since $G(z) \in \mathcal{R}_{sp}$ and $Y_i(z) \in \mathcal{RH}_\infty$,

$$X_i(\infty)D(\infty) - 1 = Y_i(\infty)N_d(\infty) = 0. \tag{4.18}$$

Straightforward analysis of Figure 4.3 reveals that, if Assumption 4.1 holds, and $K(z) \in \mathcal{S}$ and $\sigma_q^2 \in \mathbb{R}_0^+$ (as required for MSS), then σ_v^2 exists, is finite, and we can write

$$\gamma = \frac{\sigma_v^2}{\sigma_q^2} = ||T_{qv}(z)||_2^2 + \frac{||T_{drv}(z)\Omega_{dr}(z)||_2^2}{\sigma_q^2}, \tag{4.19}$$

where $T_{drv}(z)$ is the transfer function from $[d\ r]^T$ to v (and equals $T_{drw}(z)$ in (6.7)) and $\Omega_{dr}(z)$ is a spectral factor of $[d\ r]^T$. Thus, we have that, for any $K(z) \in \mathcal{S}$ and $\sigma_q^2 \in \mathbb{R}_0^+$,

$$\gamma \geq ||T_{qv}(z)||_2^2 = ||1 - X_i(z)D(z) + D(z)Q(z)N_i(z)||_2^2 . \tag{4.20}$$

We next use standard $\mathcal{H}_2$ optimization techniques (see, e.g., [18, 27, 145] and also Appendix A) to rewrite the right hand side in (4.20) in a more convenient way:[3]

$$
\begin{aligned}
&||1 - X_i(z)D(z) + D(z)Q(z)N_i(z)||_2^2 \\
&\stackrel{(a)}{=} \left|\left| \xi_p(z) - \xi_p(z)X_i(z)D(z) + \xi_p(z)D(z)Q(z)\begin{bmatrix} z^{-1} \\ N(z)\end{bmatrix}\right|\right|_2^2 \\
&\stackrel{(b)}{=} ||\xi_p(z) - \xi_p(0)||_2^2 + \left|\left| \xi_p(0) - \xi_p(\infty) + (\xi_p(\infty) - X_i(z)\xi_p(z)D(z)) + \xi_p(z)D(z)Q(z)\begin{bmatrix} z^{-1} \\ N(z)\end{bmatrix}\right|\right|_2^2 \\
&\stackrel{(c)}{=} ||\xi_p(z) - \xi_p(0)||_2^2 + ||\xi_p(0) - \xi_p(\infty)||_2^2 + \left|\left| (\xi_p(\infty) - X_i(z)\xi_p(z)D(z)) + \xi_p(z)D(z)Q(z)\begin{bmatrix} z^{-1} \\ N(z)\end{bmatrix}\right|\right|_2^2 \\
&\stackrel{(d)}{=} \left(\prod_{i=1}^{n_p} |p_i|^2\right) - 1 + \left|\left| z\left(\xi_p(\infty) - X_i(z)\xi_p(z)D(z)\right) + \xi_p(z)D(z)Q(z)\begin{bmatrix} 1 \\ zN(z)\end{bmatrix}\right|\right|_2^2
\end{aligned}
\tag{4.21}
$$

where (a) follows from the definition of $N_i(z)$, the fact that $\xi_p(z)$ is unitary (see definition A.1) and properties of the 2−norm, (b) follows from Part 4 in Fact A.1 (see Appendix A) and the fact that $\xi_p(z) \in \mathcal{RH}_2^\perp$, $\xi_p(z)D(z) \in \mathcal{RH}_\infty$ and

$$
-\xi_p(z)X_i(z)D(z) + \xi_p(z)D(z)Q(z)\begin{bmatrix} z^{-1} \\ N(z)\end{bmatrix} \in \mathcal{RH}_\infty, \tag{4.22}
$$

and (c) follows from Part 3 in Fact A.1 and the fact that $\xi_p(0) - \xi_p(\infty) \in \mathcal{RH}_2^\perp$ and

$$
(\xi_p(\infty) - X_i(z)\xi_p(z)D(z)) + \xi_p(z)D(z)Q(z)\begin{bmatrix} z^{-1} \\ N(z)\end{bmatrix} \in \mathcal{RH}_2 \tag{4.23}
$$

(recall that (4.18) holds and, since $G(z) \in \mathcal{R}_{sp}$, $N(z) \in \mathcal{RH}_2$). Finally, (d) follows from the fact that z is unitary and the following: by definition of 2−norm,

$$
\begin{aligned}
&||\xi_p(z) - \xi_p(0)||_2^2 + ||\xi_p(0) - \xi_p(\infty)||_2^2 = \\
&\qquad 1 + |\xi_p(0)|^2 - 2\,\mathrm{Re}\left\{\xi_p(0)^H \frac{1}{2\pi j}\oint_{\mathcal{C}_u} \xi_p(z)\frac{dz}{z}\right\} + |\xi_p(0) - \xi_p(\infty)|^2,
\end{aligned}
\tag{4.24}
$$

where $\mathcal{C}_u$ denotes the counter-clockwise oriented unit circle. Since $\xi_p(z)z^{-1}$ has only one pole inside the unit circle, it follows from the Residue Theorem (see, e.g., Appendix A in [155]) that

$$
||\xi_p(z) - \xi_p(0)||_2^2 + ||\xi_p(0) - \xi_p(\infty)||_2^2 = 1 - |\xi_p(0)|^2 + |\xi_p(0) - \xi_p(\infty)|^2 = \left(\prod_{i=1}^{n_p} |p_i|^2\right) - 1. \tag{4.25}
$$

[3]Recall the notation introduced in Section 2.2.

It now immediately follows from (4.20) and (4.21) that

$$\gamma \overset{(a)}{\geq} ||T_{qv}(z)||_2^2 \overset{(b)}{\geq} \left(\prod_{i=1}^{n_p} |p_i|^2\right) - 1, \tag{4.26}$$

for every $K(z) \in \mathcal{S}$ and every $\sigma_q^2 \in \mathbb{R}_0^+$. It remains to prove that γ can be made arbitrarily close to $\gamma_{\inf}$ by means of an appropriate choice for $K(z) \in \mathcal{S}$ and $\sigma_q^2 \in \mathbb{R}_0^+$.

We will start by establishing conditions under which equality in both (a) and (b) in (4.26) holds. Equality in both (a) and (b) holds if and only if $||T_{drv}(z)\Omega_{dr}(z)||_2^2\,\sigma_q^{-2} = 0$ and the second term in (4.21) equals zero. Since $G(z)$ is unstable and r and d are uncorrelated, we have that $K(z) \in \mathcal{S}$ and $\sigma_q^2 \in \mathbb{R}_0^+$ imply $||T_{drv}(z)\Omega_{dr}(z)||_2^2\,\sigma_q^{-2} = 0$ if and only if $r = d = 0$, or $d = 0$ and $M_1(z) = 0$. On the other hand, there exists $Q(z) \in \mathcal{RH}_\infty$ such that the last term in (4.21) equals zero[4] if and only if $\xi_p(z)D(z) \in \mathcal{U}_\infty$ which is equivalent to having $M_2(z) \in \mathcal{RH}_\infty$ and $G(z)$ with no poles on the unit circle. If that is the case, then one of such $Q(z)$ is given by

$$Q(z) = -z\left(\xi_p(z)D(z)\right)^{-1}\left(\xi_p(\infty) - X_i(z)\xi_p(z)D(z)\right)\begin{bmatrix}1 & 0\end{bmatrix} \in \mathcal{RH}_\infty. \tag{4.27}$$

We have thus proven that the infimum in (4.14) is achievable if and only if $d = 0$ and $G(z)$ has no poles on the unit circle. Our claims follow from (4.21), (4.19) and the definition of $\mathcal{K}_\varepsilon\{\cdot\}$.

If $d \neq 0$ or the plant has poles on the unit circle, then we have that, since for every $K(z) \in \mathcal{S}$ $||T_{drv}(z)\Omega_{dr}(z)||_2^2 < \infty$, the second term in (4.19) is a continuous function of σ_q^2 for every fixed $K(z) \in \mathcal{S}$ and, in particular, for $K(z)$ as in Part 1 of Corollary 4.1 with $M(z) \in \mathcal{RH}_\infty$ and $Q(z)$ as in (4.15) with $\epsilon \in (0, 1]$. Also, it is clear that for any such $K(z)$,

$$\lim_{\sigma_q^2 \to \infty} \frac{||T_{drv}(z)\Omega_{dr}(z)||_2^2}{\sigma_q^2} = 0. \tag{4.28}$$

Our remaining claims now follow immediately upon using Lemma A.3 to show that the second term in (4.21) can be made arbitrarily small if one chooses $Q(z)$ as in (4.15) and $\epsilon \to 0^+$. □□□

Remark 4.1 (Stable plant models) *If $G(z)$ is stable, then $K(z) = 0 \in \mathcal{S}$. Thus, $\gamma_{\inf} = 0$ and can be achieved with $K(z) = 0$ and any $\sigma_q^2 \in \mathbb{R}^+$.* □□

Theorem 4.1 establishes a closed form for the minimal stationary channel signal-to-noise ratio compatible with MSS in the feedback system of Figure 4.3. We also provided a characterization of the filter $K(z)$ and the channel noise variance σ_q^2 that allows one to achieve any signal-to-noise ratio arbitrarily close to the identified limit. An immediate consequence of Theorem 4.1 is the following:

[4]Equivalently, there exists $K(z) \in \mathcal{S}$ such that the last term in (4.21) equals zero.

Corollary 4.2 (Conditions on Γ) *Consider the feedback system in Figure 4.3, where q is the noise in an additive i.i.d. noise channel with signal-to-noise ratio constraint Γ, and suppose that Assumption 4.1 holds. If $G(z)$ is unstable, and $d \neq 0$ or $G(z)$ has poles on the unit circle, then $\Gamma > \gamma_{\inf}$ is necessary and sufficient to be able to find $K(z)$ and σ_q^2 such that MSS is guaranteed and the channel signal-to-noise ratio constraint is satisfied (i.e., to be able to find $K(z) \in \mathcal{S}$ and $\sigma_q^2 \in \mathbb{R}_0^+$ such that $\gamma \leq \Gamma$). If $G(z)$ is stable, or $d = 0$ and $G(z)$ has no poles on the unit circle, then the condition on Γ reads $\Gamma \geq \gamma_{\inf}$.*

Proof: Immediate from Theorem 4.1 and Definition 3.1. □□□

In the present case, the minimal signal-to-noise ratio compatible with MSS is a function of the unstable plant poles only (as opposed to the situation studied in Chapter 3, where the unstable poles of the in-that-situation-fixed controller played a role). Other plant features such as, e.g., NMP zero locations or relative degree,[5] do not play a role. It is thus clear that, when compared with the situation studied in Chapter 3, the general control and coding architecture proposed in this chapter has the ability to enlarge the class of plants that are stabilizable over an additive i.i.d. noise channel with a given signal-to-noise ratio constraint Γ. It is important to realize that the key constraining features in the architecture of Chapter 3 are the fact that we considered a perfect reconstruction coding scheme and a pre-designed admissible controller (that, as such, has different pole location patterns depending on whether the plant is strongly stabilizable or not). In this chapter we do not impose any of such constraints. On the contrary, we exploit all degrees-of-freedom that the considered architecture has to offer. Consistent with the results of Chapter 3, our results here show that the availability of additional degrees of freedom is fundamental to overcome limitations that architectural constraints impose (see also [26, 28]).

Remark 4.2 (Previous work) *As already mentioned in Remark 3.4, [18] also studies minimal channel signal-to-noise ratio requirements for MSS. However, the authors of [18] limit their analysis to a case where $F(z) = [0 \;\; 1]$ and $M(z) = C(z)[1 \;\; -1]$. This choice does not allow one to achieve the channel signal-to-noise ratio identified in Theorem 4.1 (except when the plant has relative degree one and is marginally-MP; see also Theorem 4.2 below).* □□

The next theorem studies control and coding architectures without feedback. To that end, we define

$$\Delta_G \triangleq \eta + \delta, \tag{4.29}$$

[5] This encompasses the issue of strong stabilizability or not.

where (see also Equation (34) in [18])

$$\eta \triangleq \sum_{i=1}^{n_c}\sum_{j=1}^{n_c} \frac{\gamma_i \overline{\gamma}_j}{(c_i \bar{c}_j - 1)}, \quad \gamma_i \triangleq (1-|c_i|^2)\left(\xi_p(c_i) - \sum_{j=0}^{m-1} \beta_j c_i^{-j}\right)\prod_{\substack{j=1\\ j\neq i}}^{n_c} \frac{1-c_i\bar{c}_j}{c_i - c_j}, \tag{4.30a}$$

$$\beta_k \triangleq \left\{\frac{1}{k!}\frac{d^k}{dz^k}\xi_p(z)^{-1}\right\}\Big|_{z=0}, \quad \delta \triangleq \begin{cases} 0, & \text{if } m=1 \\ \sum_{i=1}^{m-1}|\beta_k|^2, & \text{if } m>1 \end{cases}, \tag{4.30b}$$

m is the relative degree of $G(z)$, and $\{c_1, \cdots, c_{n_c}\}$ and $\xi_p(z)$ are defined as in Section 2.2.

Theorem 4.2 (Minimal signal-to-noise ratio for MSS (no feedback)) *Consider the feedback system in Figure 4.3, where q is the noise in an additive i.i.d. noise channel, and suppose that Assumption 4.1 holds. If, in addition, the plant is unstable and has only simple finite strictly-NMP zeros, then*

$$\gamma_{\inf,C} \triangleq \inf_{\substack{K(z)\in\mathcal{S}_C\\ \sigma_q^2\in\mathbb{R}_0^+}} \gamma = \left(\prod_{i=1}^{n_p}|p_i|^2\right) - 1 + \Delta_G, \tag{4.31}$$

where $\Delta_G = 0$ if and only if the plant is marginally-MP and has relative degree one.

Unless $d=0$ and $G(z)$ has no poles or zeros on the unit circle, $\gamma_{\inf,C}$ is not achievable, but can be approached with arbitrary precision with a sufficiently large (but finite) $\sigma_q^2 \in \mathbb{R}_0^+$ and $K(z)$ as in Part 2 of Corollary 4.1, with $F_2(z) \in \mathcal{U}_\infty$ and

$$Q(z) = \left[Q_1(z) \quad \mathcal{K}_\varepsilon\left\{\xi_c(z)\xi_p(z)N(z)D(z)\right\}^{-1}\left(\left\{[H(z)]_{\mathcal{H}_2^\perp}\right\}\Big|_{z=0} + [H(z)]_{\mathcal{H}_2}\right)\right], \tag{4.32}$$

where $Q_1(z) \in \mathcal{RH}_\infty$, $N(z)$ and $D(z)$ form a coprime factorization over $\mathcal{RH}_\infty$ of $G(z)F_2(z)$ with $G(z)F_2(z) = N(z)D(z)^{-1}$,

$$H(z) \triangleq \xi_c(z)\xi_p(z)N(z)Y_i(z)\begin{bmatrix}0\\1\end{bmatrix}, \tag{4.33}$$

and ε is a sufficiently small (but non-zero) real in $(0,1]$. Moreover, if the plant has no poles or zeros on the unit circle, then $\mathcal{K}_\varepsilon\{\cdot\}$ in (4.32) is redundant and $\gamma_{\inf,C}$ is achievable if and only if $d=0$. To do so, it suffices to choose $K(z)$ as above, with $Q_1(z)\Omega_r(z)=0$ and any $\sigma_q^2 \in \mathbb{R}^+$.

Proof: Consider $F(z) = [0 \quad F_2(z)]$, with $F_2(z) \in \mathcal{R}_p$ having a detectable and stabilizable underlying realization such that (4.10) is detectable and stabilizable. Also consider $N(z), D(z) \in \mathcal{RH}_\infty$, coprime in $\mathcal{RH}_\infty$ and with $D(z)$ biproper, such that $G(z)F_2(z) = N(z)D(z)^{-1}$. Then,

$$N_d(z) = \begin{bmatrix}0\\N(z)\end{bmatrix}, \quad D_d(z) = D(z), \quad N_i(z) = \begin{bmatrix}0\\N(z)\end{bmatrix}, \quad D_i(z) = \begin{bmatrix}1 & 0\\0 & D(z)\end{bmatrix} \tag{4.34}$$

are such that $D_d(z)$ and $D_i(z)$ are biproper, (2.42) is satisfied with $P_{22}(z)$ as in (4.12), and there exist $X_i(z)$, $Y_i(z)$, $X_d(z)$, $Y_d(z) \in \mathcal{RH}_\infty$ satisfying (2.43). As a consequence, Theorem 2.2 allows one to conclude that all transfer functions from q to v in Figure 4.3 that are achievable with $K(z) \in \mathcal{S}_C$ are given by

$$T_{qv}(z) = N(z)Y_i(z)\begin{bmatrix}0\\1\end{bmatrix} - N(z)Q(z)\begin{bmatrix}0\\D(z)\end{bmatrix}, \tag{4.35}$$

where $Q(z) \in \mathcal{RH}_\infty$ is a free parameter. Defining

$$\bar{Y}(z) \triangleq Y_i(z)\begin{bmatrix}0\\1\end{bmatrix}, \quad \bar{Q}(z) \triangleq Q(z)\begin{bmatrix}0\\1\end{bmatrix}, \tag{4.36}$$

we have

$$T_{qv}(z) = N(z)\bar{Y}(z) - N(z)\bar{Q}(z)D(z). \tag{4.37}$$

Proceeding as in the proof of Theorem 4.1 (and using the notation introduced there), it is easy to see that, if $K(z) \in \mathcal{S}_C$ and $\sigma_q^2 \in \mathbb{R}_0^+$, then σ_v^2 exists and is finite. Thus, (4.19) makes sense in the present case (recall that $T_{drv}(z)$ is the transfer function from $[d\ r]^T$ to v, which, in this case, equals $T_{drw}(z)$ in (6.25)). Therefore, we have, for any $K(z) \in \mathcal{S}_C$ and $\sigma_q^2 \in \mathbb{R}_0^+$,

$$\gamma \geq ||T_{qv}(z)||_2^2 = \left|\left|N(z)\bar{Y}(z) - N(z)\bar{Q}(z)D(z)\right|\right|_2^2. \tag{4.38}$$

We can rewrite the right hand side in (4.38) as follows (recall (4.33) and the notation introduced in Section 2.2):

$$\begin{aligned}\left|\left|N(z)\bar{Y}(z) - N(z)\bar{Q}(z)D(z)\right|\right|_2^2 &\\ &\overset{(a)}{=} \left|\left|H(z) - \xi_c(z)\xi_p(z)N(z)\bar{Q}(z)D(z)\right|\right|_2^2 \\ &\overset{(b)}{=} \left|\left|[H(z)]_{\mathcal{H}_2^\perp} - \left\{[H(z)]_{\mathcal{H}_2^\perp}\right\}\Big|_{z=0}\right|\right|_2^2 + \\ &\qquad \left|\left|\left\{[H(z)]_{\mathcal{H}_2^\perp}\right\}\Big|_{z=0} + [H(z)]_{\mathcal{H}_2} - \xi_c(z)\xi_p(z)N(z)\bar{Q}(z)D(z)\right|\right|_2^2,\end{aligned} \tag{4.39}$$

where (a) follows from the fact that both $\xi_p(z)$ and $\xi_c(z)$ are unitary, (b) follows from the fact that $\xi_c(z)\xi_p(z)N(z)D(z) \in \mathcal{RH}_\infty$, and from Parts 3 and 4 in Fact A.1 (see Appendix A). A straightforward modification of the proof of Theorem III.2 in [18] shows that the right hand side of (4.39) increases whenever $F_2(z)$ has strictly-NMP zeros, strictly unstable poles, or a relative degree greater than zero. Since we are interested in minimizing γ, we thus conclude that $F_2(z)$ must be chosen in $\overline{\mathcal{U}}_\infty$. For such a choice for $F_2(z)$, we have from (4.38), (4.39), Lemma A.3, and Theorem III.2 in [18] that

$$||T_{qv}(z)||_2^2 \geq \left(\prod_{i=1}^{n_p} |p_i|^2\right) - 1 + \Delta_G. \tag{4.40}$$

From the results in [18] we also have that $\Delta_G = 0$ if and only if $G(z)$ has no strictly-NMP zeros and has relative degree one. Thus,

$$\gamma \overset{(a)}{\geq} ||T_{qv}(z)||_2^2 \overset{(b)}{\geq} \left(\prod_{i=1}^{n_p} |p_i|^2\right) - 1 + \Delta_G = \gamma_{\inf,C}. \tag{4.41}$$

for every $K(z) \in \mathcal{S}_C$ and every $\sigma_q^2 \in \mathbb{R}_0^+$. It remains to prove that γ can be made arbitrarily close to $\gamma_{\inf,C}$ by means of an appropriate choice for $K(z) \in \mathcal{S}_C$ and $\sigma_q^2 \in \mathbb{R}_0^+$.

To complete the proof, it suffices to proceed as in the proof of Theorem 4.1 and note that, in this case, the following holds: equality in (a) in (4.41) holds for $\sigma_q^2 \in \mathbb{R}_0^+$ if and only if $r = d = 0$, or $d = 0$ and $M_1(z) = 0$. It is straightforward to see from (2.43) and (4.34) that $Y_i(z) = \begin{bmatrix} 0 & \hat{Y}(z) \end{bmatrix}$. Therefore, (4.11) implies that $M_1(z) = 0$ if and only if $Q(z) = \begin{bmatrix} 0 & \bar{Q}(z) \end{bmatrix}$. On the other hand, Lemma A.3 implies that, for $F_2(z) \in \overline{\mathcal{U}}_\infty$, there exists $\bar{Q}(z) \in \mathcal{RH}_\infty$ such that equality in (b) holds if and only if $\xi_c(z)N(z)\xi_p(z)D(z)$ has no zeros on the unit circle, i.e., if and only if the plant has no poles or zeros on the unit circle and $F_2(z) \in \mathcal{U}_\infty$. □□□

Remark 4.3 (More general plant models) *It is possible to extend Theorem 4.2 to the case of plant models that have finite strictly-NMP zeros with arbitrary multiplicity at the expense of additional algebraic effort (see also [18]).*

If $G(z)$ is stable, then $K(z) = 0 \in \mathcal{S}_C$. Thus, $\gamma_{\inf,C} = 0$ and can be achieved with $K(z) = 0$ and any $\sigma_q^2 \in \mathbb{R}^+$. □□

Theorem 4.2 is the counterpart of Theorem 4.1 for the case of control and coding architectures without feedback. Comparing both results, it follows that, unless $G(z)$ is marginally-MP and has relative degree one, the constraint on $F(z)$ limits the achievable channel signal-to-noise ratio that is compatible with MSS. For completeness, we note that the following holds:

Corollary 4.3 (Conditions on Γ (no feedback)) *Consider the feedback system in Figure 4.3, where q is the noise in an additive i.i.d. noise channel with signal-to-noise ratio constraint Γ, $F(z) = [0 \; F_2(z)]$, and suppose that Assumption 4.1 holds. If the plant is unstable, has only simple finite strictly-NMP zeros, and $d \neq 0$ or $G(z)$ has poles or zeros on the unit circle, then $\Gamma > \gamma_{\inf,C}$ is necessary and sufficient to be able to find $K(z)$ and σ_q^2 such that MSS is guaranteed* *and the channel signal-to-noise ratio constraint is satisfied (i.e., to be able to find* $K(z) \in \mathcal{S}_C$ *and* $\sigma_q^2 \in \mathbb{R}_0^+$ *such that* $\gamma \leq \Gamma$*). If* $G(z)$ *is stable, or* $d = 0$ *and* $G(z)$ *has no poles or zeros on the unit circle, then the condition on* Γ *reads* $\Gamma \geq \gamma_{\text{inf},C}$*.*

Proof: Immediate from Theorem 4.2 and Definition 3.1. □□□

Remark 4.4 (Other architectures) *It is easy to see from our previous results in this section (and their proofs) that the following holds:*

- *If* $M(z)$ *is any fixed transfer function in* $\mathcal{RH}_\infty$ *that satisfies the conditions in Part 1 of Corollary 4.1, then Theorem 4.1 and Corollary 4.2 still apply. In particular, our results still apply if* $M(z) = [0 \ \ m_2]$ *or* $M(z) = [m_1 \ \ m_2]$*, where* $m_1, m_2 \in \mathbb{R}$*,* $m_2 \neq 0$*.*
- *If* $F(z) = [0 \ \ F_2(z)]$ *and* $F_2(z)$ *is any fixed transfer function in* $\mathcal{U}_\infty$ *satisfying the conditions in Part 2 of Corollary 4.1, then Theorem 4.2 and Corollary 4.3 still apply. In particular, our results apply to the case* $F(z) = [0 \ \ f]$*, where* $f \in \mathbb{R} \setminus \{0\}$*.*
- *If* $K(z)$ *is constrained to yield a one-dof control architecture, i.e.,* $F(z) = [0 \ \ 1]$ *and* $M(z) = C(z)[1 \ \ -1]$*, then Theorem 4.2 and Corollary 4.3 still apply (see also [18]).*

□□

Remark 4.5 (Previous work) *Note that the condition on the minimal signal-to-noise ratio compatible with MSS derived in Theorem 4.2 is the same condition obtained in [18], even though our setup is more general than that of [18] (see also the third bullet in Remark 4.4). This implies that a simple one-dof architecture is sufficient to be able to achieve MSS at any signal-to-noise ratio* γ *greater than* $\gamma_{\text{inf},C}$*. As foreshadowed in Remark 3.4, this also implies that feedback around the channel is the key element in reducing the signal-to-noise ratio required for MSS.*[6] *If no feedback is used, then, no matter how many (additional) degrees of freedom one employs, the minimal signal-to-noise ratio required for MSS is as in a one-dof architecture.* □□

From Theorems 4.1 and 4.2 we see that in order to approach the minimal signal-to-noise ratio γ_{inf} (or $\gamma_{\text{inf},C}$) one needs, for most cases of interest, a channel noise variance that grows unbounded. As a consequence, the performance of the resulting networked control system will be

[6] Recall that in our framework we allow only for feedback around the channel, but we rule out other communication channels. Similar results and conclusions apply if, for example, no feedback around the channel is available, but it is possible to access the plant input at the encoder side.

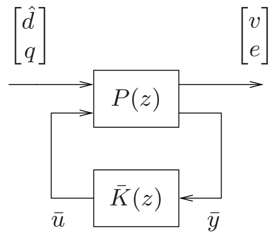

Figure 4.4: General feedback configuration for signal-to-noise ratio minimization.

severely compromised. This reveals that the study of minimal signal-to-noise ratio requirements for MSS is insufficient, and motivates the extension of the presented framework to situations where guaranteeing acceptable performance levels is necessary. This is done in the next two sections.

4.4 A General Approach to Design with Signal-to-noise Ratio Constraints

This section studies a general optimization problem that will enable us to characterize the minimal channel signal-to-noise ratio that guarantees a certain performance level, and also the best performance that is achievable with a given signal-to-noise ratio constraint.

Consider the feedback system of Figure 4.4, where $\hat{d}$ and q are external stochastic (vector) processes, and $P(z)$ and $\bar{K}(z)$ are LTI systems in $\mathcal{R}_p$. Assume that $P(z)$ is partitioned in a way such that

$$\begin{bmatrix} v \\ e \\ \hline \bar{y} \end{bmatrix} = \left[\begin{array}{c:c} P_{v11}(z) & P_{v12}(z) \\ P_{e11}(z) & P_{e12}(z) \\ \hdashline P_{21}(z) & P_{22}(z) \end{array}\right] \begin{bmatrix} \hat{d} \\ q \\ \hline \bar{u} \end{bmatrix}. \tag{4.42}$$

Throughout this section we will assume that the following holds:

Assumption 4.2

(a) $P(z)$ belongs to $\mathcal{R}_p$, is stabilizable, such that $P_{22}(z) \in \mathcal{R}_{sp}$, and such that $P_{21}(z)^T, P_{v12}(z)$ and $P_{e12}(z)$ have full column normal rank (i.e., full column rank for almost every $z \in \mathbb{C}$; see, e.g., [87]).

(b) The signals $\hat{d}$ and q are second order mutually uncorrelated wss processes that have rational PSDs. The process $\hat{d}$ admits a spectral factor in $\mathcal{U}_\infty$ (hence, $\hat{d}$ is non-zero). The process q is scalar, white, has zero mean and its variance $\sigma_q^2 \in \mathbb{R}_0^+$ is a design parameter.

(c) The initial states of both $P(z)$ and $\bar{K}(z)$ are jointly second order random variables.

□□

Assumption 4.2 has, essentially, the same spirit as Assumption 4.1. The last part of Assumption 4.2(a) is, however, somewhat arbitrary and restrictive. Nevertheless, we stress that the results in this section are just tools that will be used in Section 4.5 to deal with the real problems of interest. Assumption 4.2 will prove general enough to encompass all those cases.

If the stationary variance of v in the feedback system of Figure 4.4 exists, then we define the ratio

$$\bar{\gamma} \triangleq \frac{\sigma_v^2}{\sigma_q^2}. \tag{4.43}$$

We are interest in finding, within the class of all $\bar{K}(z)$ and σ_q^2 that render the feedback system in Figure 4.4 MSS, the controller $\bar{K}(z)$ and the noise variance σ_q^2 that satisfy one of the following requirements:

(i) Minimize $\bar{\gamma}$ while satisfying a performance constraint in the form of a bound on the stationary variance of e (i.e., on σ_e^2; see Figure 4.4). That is, we are interested in finding $\bar{K}(z)$ and σ_q^2 that achieve a signal-to-noise ratio $\bar{\gamma}$ equal (or arbitrarily close) to

$$\bar{\gamma}_D \triangleq \inf_{\substack{\sigma_q^2 \in \mathbb{R}_0^+ \\ \bar{K}(z) \in \bar{\mathcal{S}} \\ \sigma_e^2 \leq D}} \frac{\sigma_v^2}{\sigma_q^2}, \tag{4.44}$$

where $D \in [0, \infty)$ is the desired performance level and $\bar{\mathcal{S}}$ the set of all admissible controllers $\bar{K}(z)$ for $P(z)$ (recall Section 2.4).

(ii) Minimize the stationary variance of e, i.e., optimize performance, while satisfying a constraint on the signal-to-noise ratio $\bar{\gamma}$. That is, we are interested in finding $\bar{K}(z)$ and σ_q^2 that achieve a stationary variance of e equal (or arbitrarily close) to

$$\left[\sigma_e^2\right]_\Gamma \triangleq \inf_{\substack{\sigma_q^2 \in \mathbb{R}_0^+ \\ \bar{K}(z) \in \bar{\mathcal{S}} \\ \gamma \leq \Gamma}} \sigma_e^2, \tag{4.45}$$

where $\Gamma \in [0, \infty)$ is the maximum allowable signal-to-noise ratio and $\bar{\mathcal{S}}$ is as before.

It is straightforward to see from Figure 4.4 that, provided Assumption 4.2 holds and $\bar{K}(z) \in \bar{\mathcal{S}}$, both variances σ_v^2 and σ_e^2 are well defined, finite and satisfy

$$\sigma_v^2 = \left|\left|T_{\bar{d}v}(z)\Omega_{\bar{d}}(z)\right|\right|_2^2, \quad \sigma_e^2 = \left|\left|T_{\bar{d}e}(z)\Omega_{\bar{d}}(z)\right|\right|_2^2, \tag{4.46}$$

where

$$\bar{d} \triangleq \begin{bmatrix} \hat{d} & q \end{bmatrix}^T, \quad \Omega_{\bar{d}}(z) \triangleq \mathrm{diag}\left\{\Omega_{\hat{d}}(z), \sigma_q\right\}, \tag{4.47}$$

and $T_{\bar{d}v}(z), T_{\bar{d}e}(z)$ denote the transfer functions from $\bar{d}$ to v and e, respectively. Using Theorem 2.2, we can write all transfer functions $T_{\bar{d}v}(z)$ and $T_{\bar{d}e}(z)$ that are achievable with $\bar{K}(z) \in \bar{\mathcal{S}}$ as

$$T_{\bar{d}v}(z) = T^o_{\bar{d}v}(z) - P_{v12}(z)D_d(z)Q(z)D_i(z)P_{21}(z), \tag{4.48a}$$

$$T_{\bar{d}e}(z) = T^o_{\bar{d}e}(z) - P_{e12}(z)D_d(z)Q(z)D_i(z)P_{21}(z), \tag{4.48b}$$

where

$$T^o_{\bar{d}v}(z) \triangleq P_{v11}(z) + P_{v12}(z)D_d(z)Y_i(z)P_{21}(z), \tag{4.49a}$$

$$T^o_{\bar{d}e}(z) \triangleq P_{e11}(z) + P_{e12}(z)D_d(z)Y_i(z)P_{21}(z), \tag{4.49b}$$

$X_i(z)$, $Y_i(z)$, $N_i(z)$, $D_i(z)$, $N_d(z)$, $D_d(z)$ are defined as in Theorem 2.2, and $Q(z)$ is a free parameter in $\mathcal{RH}_\infty$ (since we assume that $P_{22}(z)$ is strictly proper, $\det\left(X_i(\infty) - Q(\infty)N_i(\infty)\right) \neq 0$, as required). Also, all $\bar{K}(z) \in \bar{\mathcal{S}}$ obey

$$\bar{K}(z) = (X_i(z) - Q(z)N_i(z))^{-1}(Y_i(z) - Q(z)D_i(z)), \tag{4.50}$$

with $Q(z)$ as above.

With the aid of (4.48) and (4.46) it is possible to write (4.44) and (4.45) in the equivalent form

$$\bar{\gamma}_D = \inf_{\sigma_q^2 \in \mathbb{R}_0^+} \frac{1}{\sigma_q^2} \inf_{\substack{Q(z)\in\mathcal{RH}_\infty \\ J_{\sigma_q^2}(Q(z)) \leq D}} R_{\sigma_q^2}(Q(z)), \tag{4.51a}$$

$$\left[\sigma_e^2\right]_\Gamma = \inf_{\sigma_q^2 \in \mathbb{R}_0^+} \inf_{\substack{Q(z)\in\mathcal{RH}_\infty \\ R_{\sigma_q^2}(Q(z)) \leq \Gamma\sigma_q^2}} J_{\sigma_q^2}(Q(z)), \tag{4.51b}$$

where

$$J_{\sigma_q^2}(Q(z)) \triangleq \left|\left|T_{\bar{d}e}(z)\Omega_{\bar{d}}(z)\right|\right|_2^2, \tag{4.52a}$$

$$R_{\sigma_q^2}(Q(z)) \triangleq \left|\left|T_{\bar{d}v}(z)\Omega_{\bar{d}}(z)\right|\right|_2^2. \tag{4.52b}$$

$J_{\sigma_q^2}(Q(z))$ is the stationary variance of e as a function of $Q(z)$ with σ_q^2 as a parameter; analogously, $R_{\sigma_q^2}(Q(z))$ is the corresponding stationary variance of v. We also define

$$\left(\bar{\sigma}_D^2, \bar{Q}_D(z)\right) \triangleq \arg \inf_{\substack{Q(z)\in\mathcal{RH}_\infty \\ \sigma_q^2\in\mathbb{R}_0^+ \\ J_{\sigma_q^2}(Q(z))\leq D}} \frac{R_{\sigma_q^2}(Q(z))}{\sigma_q^2}, \tag{4.53}$$

$$\left(\bar{\sigma}_\Gamma^2, \bar{Q}_\Gamma(z)\right) \triangleq \arg \inf_{\substack{\sigma_q^2\in\mathbb{R}_0^+ \\ Q(z)\in\mathcal{RH}_\infty \\ R_{\sigma_q^2}(Q(z))\leq\Gamma\sigma_q^2}} J_{\sigma_q^2}(Q(z)), \tag{4.54}$$

and assume that the following holds:

Assumption 4.3 *The performance level D (resp. the signal-to-noise ratio constraint Γ) is such that the optimization problem in* (4.51a) *(resp.* (4.51b)*) is feasible and the minimal signal-to-noise ratio $\bar{\gamma}_D$ (resp. minimal stationary variance $\left[\sigma_e^2\right]_\Gamma$) is achieved with $\sigma_q^2 \in \mathbb{R}^+$, i.e., $0 < \bar{\sigma}_D^2 < \infty$ (resp. $0 < \bar{\sigma}_\Gamma^2 < \infty$).* □□

Again, Assumption 4.3 is restrictive, but is sufficient for our purposes.

It is easy to see that the following holds:

Corollary 4.4 (Strict convexity) *Consider the feedback system in Figure 4.4 and assume that Assumption 4.2(a-b) holds. Then, $J_{\sigma_q^2}$ and $R_{\sigma_q^2}$ in* (4.52) *are strictly convex functions of $Q(z)$ for every fixed $\sigma_q^2 \in \mathbb{R}^+$.*

Proof: By construction, $D_i(z)$ and $D_d(z)$ are square and non-singular (almost everywhere). Moreover, by Assumption 4.2 $\Omega_{\bar{d}}(z) \in \mathcal{U}_\infty$ for every $\sigma_q^2 \in \mathbb{R}^+$, and $P_{v12}(z)$, $P_{e12}(z)$ and $P_{21}(z)^T$ have full column normal rank. Therefore, $P_{v12}(z)D_d(z)$, $P_{e12}(z)D_d(z)$ and $(D_i(z)P_{21}(z)\Omega_{\bar{d}}(z))^T$ have full column normal rank and the result follows from Lemma A.2 in Appendix A. □□□

The next subsection is aimed at providing preliminary results that will be used to characterize both $\bar{\gamma}_D$ and $\left[\sigma_e^2\right]_\Gamma$.

4.4.1 Performance limits

In this section we explore performance limits for the feedback system of Figure 4.4 when $\sigma_q^2 \in \mathbb{R}^+$ is *fixed*. In particular, we will focus on elucidating, for every $\sigma_q^2 \in \mathbb{R}^+$, the *optimal trade-off curves* in the $J_{\sigma_q^2}$ *versus* $R_{\sigma_q^2}$ plane (see, e.g., Section 4.7 in [16]). These insights will then be used in Section 4.5 to provide a characterization of $\bar{\gamma}_D$ and $\left[\sigma_e^2\right]_\Gamma$.

The set of all achievable pairs $(J_{\sigma_q^2}, R_{\sigma_q^2})$ is given by

$$\mathcal{F}_{\sigma_q^2} \triangleq \Big\{(\alpha_e, \alpha_v) \in \mathbb{R}^2 : \; J_{\sigma_q^2}(Q(z)) \le \alpha_e \text{ and } R_{\sigma_q^2}(Q(z)) \le \alpha_v \text{ for some } Q(z) \in \mathcal{RH}_\infty \Big\}. \quad (4.55)$$

It is clear that, in general, $J_{\sigma_q^2}$ and $R_{\sigma_q^2}$ are competing objectives, i.e., one cannot simultaneously minimize both $J_{\sigma_q^2}$ and $R_{\sigma_q^2}$. As a consequence, the set $\mathcal{F}_{\sigma_q^2}$ has no optimal point, i.e., there exist no $(\alpha_e, \alpha_v) \in \mathcal{F}_{\sigma_q^2}$ such that $\alpha_e \le J_{\sigma_q^2}(Q(z))$ and, simultaneously, $\alpha_v \le R_{\sigma_q^2}(Q(z))$ for every $Q(z) \in \mathcal{RH}_\infty$. Nevertheless, we can consider the set $\mathcal{P}_{\sigma_q^2} \subseteq \mathcal{F}_{\sigma_q^2}$ containing the points in the $(J_{\sigma_q^2}, R_{\sigma_q^2})$ plane that provide the *best* trade-off in the following sense: $(\theta_e, \theta_v) \in \mathcal{P}_{\sigma_q^2}$ if and only if, for every $(\alpha_e, \alpha_v) \in \mathcal{F}_{\sigma_q^2}$, $\alpha_e \le \theta_e$ and $\alpha_v \le \theta_v$ implies $\alpha_e = \theta_e$ and $\alpha_v = \theta_v$. Roughly speaking, $\mathcal{P}_{\sigma_q^2}$ contains the points in $\mathcal{F}_{\sigma_q^2}$ that *cannot be improved in both components simultaneously*. The set $\mathcal{P}_{\sigma_q^2}$ contains the so-called Pareto optimal points of the vector valued criterion $\begin{bmatrix} J_{\sigma_q^2} & R_{\sigma_q^2} \end{bmatrix}$ (see, e.g., [16, 92]).

Paraphrasing [16] and [92], we conclude that if one is interested in *good* solutions, i.e., solutions that, in some sense, minimize both $J_{\sigma_q^2}$ and $R_{\sigma_q^2}$, then one should focus on the points of the $(J_{\sigma_q^2}, R_{\sigma_q^2})$ plane that belong to $\mathcal{P}_{\sigma_q^2}$. Indeed, there exist Youla parameters that achieve a smaller value for $J_{\sigma_q^2}$ at the expense of a smaller value for $R_{\sigma_q^2}$ for every point in $\mathcal{F}_{\sigma_q^2}$ that does not belong to $\mathcal{P}_{\sigma_q^2}$. In our case, $\mathcal{P}_{\sigma_q^2}$ defines a curve in $\mathbb{R}^2$, the *optimal trade-off curve*. The next lemma characterizes $\mathcal{P}_{\sigma_q^2}$.

Lemma 4.1 (Characterization of $\mathcal{P}_{\sigma_q^2}$) *Consider the feedback system in Figure 4.4 and suppose that Assumption 4.2(a-b) holds. Define, for every $\sigma_q^2 \in \mathbb{R}^+$ and every $\epsilon \in [0,1]$,*

$$L_{\sigma_q^2,\epsilon}(Q(z)) \triangleq \epsilon J_{\sigma_q^2}(Q(z)) + (1-\epsilon) R_{\sigma_q^2}(Q(z)), \quad (4.56)$$

$$Q_{\sigma_q^2,\epsilon}(z) \triangleq \arg \inf_{Q(z)\in\mathcal{RH}_\infty} L_{\sigma_q^2,\epsilon}(Q(z)), \quad (4.57)$$

where $J_{\sigma_q^2}$ and $R_{\sigma_q^2}$ are as in (4.52)*. Then, $(\alpha_e, \alpha_v) \in \mathcal{P}_{\sigma_q^2}$ if and only if*

$$(\alpha_e, \alpha_v) = \Big(J_{\sigma_q^2}(Q_{\sigma_q^2,\epsilon}(z)), R_{\sigma_q^2}(Q_{\sigma_q^2,\epsilon}(z)) \Big) \quad (4.58)$$

for some $\epsilon \in [0,1]$.

Proof: Given our assumptions, Corollary 4.4 guarantees the strict convexity of $J_{\sigma_q^2}$ and $R_{\sigma_q^2}$. Therefore, $L_{\sigma_q^2,\epsilon}$ is also strictly convex (see Section 3.2.1 in [16]) and the optimization problem in (4.57) admits, at most, one solution (see [16, 102]). We thus conclude from Theorem 2.1 in [92] (see also Section 4.7 in [16]) that $\Big(J_{\sigma_q^2}(Q_p(z)), R_{\sigma_q^2}(Q_p(z))\Big) \in \mathcal{P}_{\sigma_q^2}$ if and only if there exist $\lambda_1, \lambda_2 \ge 0$, $\lambda_1 + \lambda_2 = 1$, such that

$$Q_p(z) \triangleq \arg \inf_{Q(z)\in\mathcal{RH}_\infty} \lambda_1 J_{\sigma_q^2}(Q(z)) + \lambda_2 R_{\sigma_q^2}(Q(z)). \quad (4.59)$$

The result now follows by defining $\epsilon \triangleq \lambda_1$. □□□

With the aid of Lemma 4.1 we can find, for every $\sigma_q^2 \in \mathbb{R}^+$, an optimal trade-off curve in the $J_{\sigma_q^2}$ *versus* $R_{\sigma_q^2}$ plane. To do so, it suffices to find $Q_{\sigma_q^2,\epsilon}(z)$ for $\epsilon \in [0,1]$ and use (4.48) and (4.52). The (strict) convexity of both $J_{\sigma_q^2}$ and $R_{\sigma_q^2}$ guarantees that $L_{\sigma_q^2,\epsilon}$ is also a (strictly) convex functional.[7] Therefore, finding $Q_{\sigma_q^2,\epsilon}(z)$ is an easy problem, both from a theoretical and numerical perspective (see, e.g., [16] and also [13, 15, 149] for a general treatment of convex optimization problems in control theory with emphasis on linear matrix inequalities (LMIs)). Given the simplicity of $L_{\sigma_q^2,\epsilon}$, it is possible to minimize $L_{\sigma_q^2,\epsilon}$ using direct methods (that do not rely on LMIs). Indeed, one can use standard $\mathcal{H}_2$ optimization techniques to end up with a Ricatti equation based solution (see, e.g., [194]), or use the alternative analytic $\mathcal{H}_2$ optimization machinery illustrated in, e.g., [26, 113, 163, 178] (see also Appendix A and [46, 182]). We adopt the latter approach here:

Theorem 4.3 (Closed form characterization of $Q_{\sigma_q^2,\epsilon}(z)$) *Consider the feedback system in Figure 4.4, and assume that Assumption 4.2(a-b) holds. Define*

$$A(z) \triangleq \begin{bmatrix} \sqrt{\epsilon}\, T^o_{\bar{d}e}(z) \\ \sqrt{1-\epsilon}\, T^o_{\bar{d}v}(z) \end{bmatrix} \Omega_{\bar{d}}(z), \quad E_\epsilon(z) \triangleq \begin{bmatrix} \sqrt{\epsilon}\, P_{e12}(z) \\ \sqrt{1-\epsilon}\, P_{v12}(z) \end{bmatrix}, \tag{4.60}$$

$$A_v(z) \triangleq \text{vec}\left\{A(z)\right\}, \qquad B_v(z) \triangleq \left(D_i(z)P_{21}(z)\Omega_{\bar{d}}(z)\right)^T \otimes E_\epsilon(z)D_d(z), \tag{4.61}$$

where all symbols are as defined previously.[8] *Consider an inner-outer factorization of $B_v(z)$ given by $B_v(z) = B_i(z)B_o(z)$, where $B_i(z)$ is inner and $B_o(z)$ is outer (see Appendix A). Then, for any $\sigma_q^2 \in \mathbb{R}^+$ and $\epsilon \in [0,1]$, we have*

$$Q_{\sigma_q^2,\epsilon}(z) = \text{vec}^{-1}\left\{B_o(z)^{-1}\left(\left\{[B_i(z)^\sim A_v(z)]_{\mathcal{H}_2^\perp}\right\}\Big|_{z=0} + [B_i(z)^\sim A_v(z)]_{\mathcal{H}_2}\right)\right\}. \tag{4.62}$$

Moreover, $Q_{\sigma_q^2,\epsilon}(z) \in \mathcal{RH}_\infty$ if $P_{22}(z)$ has no poles on the unit circle and both $P_{21}(z)$ and $E_\epsilon(z)$ have no zeros on the unit circle.

Proof: Using (4.56), (4.52), (4.48) and basic properties of the 2-norm, it is possible to write

$$\begin{aligned} L_{\sigma_q^2,\epsilon}(Q(z)) &= \left|\left|\begin{bmatrix} \sqrt{\epsilon}\ T^o_{\bar{d}e}(z) \\ \sqrt{1-\epsilon}\ T^o_{\bar{d}v}(z) \end{bmatrix}\Omega_{\bar{d}}(z) - \begin{bmatrix} \sqrt{\epsilon}\ P_{e12}(z) \\ \sqrt{1-\epsilon}\ P_{v12}(z) \end{bmatrix} D_d(z)Q(z)D_i(z)P_{21}(z)\Omega_{\bar{d}}(z)\right|\right|_2^2 \\ &= ||A_v(z) - B_v(z)\text{vec}\left\{Q(z)\right\}||_2^2, \end{aligned} \tag{4.63}$$

[7] This guarantees that the optimal trade-off curves are also strictly convex (see [16]).

[8] vec $\{\cdot\}$ denotes the column stacking operator and $\otimes$ is the Kronecker product (see, e.g., [11, 19]).

where the last equality follows from Lemma A.1 in Appendix A.

Since $P_{21}(z)^T, P_{e12}(z)$ and $P_{v12}(z)$ have full column normal rank, it is immediate to conclude from the proof of Corollary 4.4 and properties of the Kronecker product (see Chapter 7 in [11] and, in particular, Fact 7.4.20), that $B_v(z)$ has full column normal rank. Therefore, basic properties of inner-outer factorizations (see Theorem A.1 in Appendix A) imply that $B_o(z) \in \overline{\mathcal{U}}_\infty$ and all its marginally-MP zeros are the marginally-MP zeros of $E_\epsilon(z)D_d(z)$ and $D_i(z)P_{21}(z)\Omega_{\bar{d}}(z)$.

Consider the unitary matrix

$$\phi(z) \triangleq \begin{bmatrix} B_i(z)^\sim \\ I - B_i(z)B_i(z)^\sim \end{bmatrix} \tag{4.64}$$

and define

$$H(z) \triangleq B_i(z)^\sim A_v(z), \quad H_u(z) \triangleq [B_i(z)^\sim A_v(z)]_{\mathcal{H}_2^\perp}, \quad H_s(z) \triangleq [B_i(z)^\sim A_v(z)]_{\mathcal{H}_2}. \tag{4.65}$$

Since $\phi(z)$ is unitary and $B_i(z)^\sim B_i(z) = I$, we have from (4.63) that

$$\begin{aligned} L_{\sigma_q^2,\epsilon}(Q(z)) &= ||\phi(z)\left(A_v(z) - B_v(z)\text{vec}\left\{Q(z)\right\}\right)|| \\ &= ||\left(I - B_i(z)B_i(z)^\sim\right)A_v(z)||_2^2 + ||H(z) - B_o(z)\text{vec}\left\{Q(z)\right\}||_2^2 \\ &= ||\left(I - B_i(z)B_i(z)^\sim\right)A_v(z)||_2^2 + ||H_u(z) - H_u(0)||_2^2 + \\ &\qquad ||H_u(0) + H_s(z) - B_o(z)\text{vec}\left\{Q(z)\right\}||_2^2, \end{aligned} \tag{4.66}$$

were the last equality follows from Part 4 in Fact A.1 in Appendix A. Since $B_o(z) \in \overline{\mathcal{U}}_\infty$ and, by definition, $H_s(z) \in \mathcal{RH}_2$, the first part of the result follows from (4.66).

To prove the remaining claim, we note that if $B_o(z) \in \mathcal{U}_\infty$, then $Q_{\sigma_q^2,\epsilon}(z) \in \mathcal{RH}_\infty$. (In general, this condition is not necessary since $H_u(0) + H_s(z)$ may contain zeros on the unit circle that coincide [in location and direction] with all the marginally-MP zeros of $B_o(z)$.) Since $\Omega_{\bar{d}}(z) \in \mathcal{U}_\infty$ by assumption, we conclude that $P_{22}(z)$ with no poles on the unit circle (i.e., $D_d(z)$ and $D_i(z)$ with no marginally-MP zeros) and both $P_{21}(z)$ and $E_\epsilon(z)$ with no marginally-MP zeros imply that $B_o(z) \in \mathcal{U}_\infty$. □□□

The crucial point in calculating $Q_{\sigma_q^2,\epsilon}(z)$ is the inner-outer factorization of $B_v(z)$. This can be made with the aid of the algorithms described in, e.g., [46, 123].

As pointed out in Theorem 4.3, in general the Youla parameter $Q_{\sigma_q^2,\epsilon}(z) \in \overline{\mathcal{RH}_\infty}$. We will sometimes require an explicit form for a family of Youla parameters in $\mathcal{RH}_\infty$ that make $L_{\sigma_q^2,\epsilon}$ arbitrarily close to $L_{\sigma_q^2,\epsilon}(Q_{\sigma_q^2,\epsilon}(z))$. To characterize such a family, we note that Theorem 4.3

gives an expression for $Q_{\sigma_q^2,\epsilon}(z)$ in terms of $B_o(z)$, $B_i(z)$ and $A_v(z)$, and that $B_o(z) \in \mathcal{U}_\infty$ implies $Q_{\sigma_q^2,\epsilon}(z) \in \mathcal{RH}_\infty$. Accordingly, we will define, for any $Q_{\sigma_q^2,\epsilon}(z)$, the related transfer function

$$Q^{\varepsilon}_{\sigma_q^2,\epsilon}(z) \triangleq \text{vec}^{-1}\left\{\mathcal{K}_\varepsilon\left\{B_o(z)\right\}^{-1}\left(\left\{[B_i(z)^{\sim}A_v(z)]_{\mathcal{H}_2^\perp}\right\}\Big|_{z=0} + [B_i(z)^{\sim}A_v(z)]_{\mathcal{H}_2}\right)\right\}, \tag{4.67}$$

where $\varepsilon \in (0,1]$. By construction, $Q^{\varepsilon}_{\sigma_q^2,\epsilon}(z) \in \mathcal{RH}_\infty$ and

$$\lim_{\varepsilon\to 0^+} L_{\sigma_q^2,\epsilon}(Q^{\varepsilon}_{\sigma_q^2,\epsilon}(z)) = L_{\sigma_q^2,\epsilon}(Q_{\sigma_q^2,\epsilon}(z)), \tag{4.68}$$

as required. (If $B_o(z) \in \mathcal{U}_\infty$, then the operator $\mathcal{K}_\varepsilon\{\cdot\}$ in (4.67) is redundant and $Q^{\varepsilon}_{\sigma_q^2,\epsilon}(z) = Q_{\sigma_q^2,\epsilon}(z)$.)

Remark 4.6 ($\bar{\gamma}$ *versus* σ_e^2) *From the results presented above, one immediately obtains, for every fixed $\sigma_q^2 \in \mathbb{R}^+$, a characterization of the optimal trade-off curve in the σ_e^2 (i.e., $J_{\sigma_q^2}$)* versus *$\bar{\gamma}$ (i.e., $\sigma_q^{-2}R_{\sigma_q^2}$) plane.* □□

4.4.2 Optimal designs

In this section, we use the results of Section 4.4.1 to provide an exact characterization of the minimal signal-to-noise ratio $\bar{\gamma}_D$ and the best performance $\left[\sigma_e^2\right]_\Gamma$. We will start by establishing bounds on these quantities:

Lemma 4.2 (Bounds on $\bar{\gamma}_D$ and $\left[\sigma_e^2\right]_\Gamma$) *Consider the setup of Figure 4.4 and suppose that Assumptions 4.2 and 4.3 hold. Then:*

1. *$\bar{\gamma}_D$ is lower bounded by*

$$\bar{\gamma}_{\text{inf}} \triangleq \inf_{\substack{Q(z)\in\mathcal{RH}_\infty \\ \sigma_q^2\in\mathbb{R}_0^+}} \frac{R_{\sigma_q^2}(Q(z))}{\sigma_q^2} \leq \bar{\gamma}_D \tag{4.69}$$

and, if there exists $\sigma_q^{2*} \in \mathbb{R}_0^+$ such that $J_{\sigma_q^{2*}}(Q_{\sigma_q^{2*},0}(z)) = D$, then

$$\bar{\gamma}_D \leq \frac{R_{\sigma_q^{2*}}(Q_{\sigma_q^{2*},0}(z))}{\sigma_q^{2*}}. \tag{4.70}$$

2. *$\left[\sigma_e^2\right]_\Gamma$ is lower bounded by*

$$D_{\text{inf}} \triangleq \inf_{\substack{Q(z)\in\mathcal{RH}_\infty \\ \sigma_q^2=0}} J_{\sigma_q^2}(Q(z)) \leq \left[\sigma_e^2\right]_\Gamma \tag{4.71}$$

and, if there exists $\sigma_q^{2\#} \in \mathbb{R}_0^+$ such that $R_{\sigma_q^{2\#}}(Q_{\sigma_q^{2\#},1}(z)) = \Gamma\sigma_q^{2\#}$, then

$$\left[\sigma_e^2\right]_\Gamma \leq \frac{J_{\sigma_q^{2\#}}(Q_{\sigma_q^{2\#},1}(z))}{\sigma_q^{2\#}}. \tag{4.72}$$

Proof:

1. To prove that $\bar{\gamma}_{\inf}$ constitutes a lower bound on $\bar{\gamma}_D$, it suffices to note that $\bar{\gamma}_{\inf}$ is the minimal signal-to-noise ratio compatible with MSS (compare with the results in Section 4.3). Therefore, if one adds the constraint $J_{\sigma_q^2}(Q(z)) \leq D < \infty$, then the constrained minimal signal-to-noise ratio must necessarily increase.

   We next consider the upper bound on $\bar{\gamma}_D$. By definition, σ_q^{2*} is such that $J_{\sigma_{q^*}^2}(Q_{\sigma_{q^*}^2,0}(z)) \leq D$ and, hence, the pair $(\sigma_{q^*}^2, Q_{\sigma_{q^*}^2,0}(z))$ is feasible except for the fact that $Q_{\sigma_{q^*}^2,0}(z)$ may belong to $\overline{\mathcal{RH}}_\infty$ and not to $\mathcal{RH}_\infty$; see Theorem 4.3. Therefore, $\sigma_{q^*}^{-2} R_{\sigma_q^{2*}}(Q_{\sigma_q^{2*},0}(z))$ must necessarily be an upper bound on $\bar{\gamma}_D$.

2. Analogous to the proof of Part 1 above.

□□□

The bounds on $\bar{\gamma}_D$ and $\left[\sigma_e^2\right]_\Gamma$ established in Lemma 4.2 are loose, but can be used to obtain a first indication of the required signal-to-noise ratio that guarantees a performance level D or, alternatively, the minimal stationary variance of e that is achievable with a given signal-to-noise ratio constraint. We note that $\bar{\gamma}_{\inf}$ is the minimal signal-to-noise ratio $\bar{\gamma}$ that is achievable when no attention is paid to the variance of e, i.e., when $D \to \infty$. In other words, $\bar{\gamma}_{\inf}$ is the minimal signal-to-noise ratio $\bar{\gamma}$ that guarantees MSS. As such, $\bar{\gamma}_{\inf}$ can be evaluated using the techniques used in Section 4.3. Analogously, $D_{\inf}$ is the minimal stationary variance of e that is achievable when $q = 0$, i.e., when no constraints on $\bar{\gamma}$ are enforced. $D_{\inf}$ can be found using standard $\mathcal{H}_2$ optimization techniques (see, e.g., [113, 194]).

The next two theorems present the main results of this section:

Theorem 4.4 (Characterization of $\bar{\gamma}_D$) *Consider the setup in Figure 4.4 and suppose that Assumptions 4.2 and 4.3 hold. Define the set*

$$\Sigma_D \triangleq \left\{ \sigma_q^2 \in \mathbb{R}^+ : \inf_{Q(z) \in \mathcal{RH}_\infty} J_{\sigma_q^2}(Q(z)) \leq D \right\} \tag{4.73}$$

and the function $f_D : \Sigma_D \to [0,1]$ *defined via*

$$f_D(\sigma_q^2) = \begin{cases} 0 & \text{if } D > J_{\sigma_q^2}(Q_{\sigma_q^2,0}(z)), \\ \epsilon & \text{if } \epsilon \in [0,1] \text{ is such that } D = J_{\sigma_q^2}(Q_{\sigma_q^2,\epsilon}(z)). \end{cases} \tag{4.74}$$

Then:

1. *The optimal noise variance* $\bar{\sigma}_D^2$ *and the optimal Youla parameter* $\bar{Q}_D(z)$ *are given by*

$$\bar{\sigma}_D^2 = \arg \min_{\sigma_q^2 \in \Sigma_D} \frac{1}{\sigma_q^2} R_{\sigma_q^2}(Q_{\sigma_q^2, f_D(\sigma_q^2)}(z)), \quad \bar{Q}_D(z) = Q_{\bar{\sigma}_D^2, f_D(\bar{\sigma}_D^2)}(z). \tag{4.75}$$

Thus, the minimal signal-to-noise ratio $\bar{\gamma}_D$ *satisfies*

$$\bar{\gamma}_D = \frac{1}{\bar{\sigma}_D^2} \, R_{\bar{\sigma}_D^2}(\bar{Q}_D(z)). \tag{4.76}$$

2. *It is possible to get arbitrarily close to* $\bar{\gamma}_D$*, while violating the performance constraint as little as desired, upon choosing* $\sigma_q^2 = \bar{\sigma}_D^2$ *and* $\bar{K}(z)$ *as in* (4.50) *with*

$$Q(z) = Q^{\varepsilon}_{\bar{\sigma}_D^2, f_D(\bar{\sigma}_D^2)}(z) \tag{4.77}$$

and $\varepsilon \in (0,1]$ *sufficiently small. Moreover, if* $P_{22}(z)$ *has no poles on the unit circle and both* $P_{21}(z)$ *and* $E_{f_D(\bar{\sigma}_D^2)}(z)$ *have no zeros on the unit circle, then* $Q^{\varepsilon}_{\bar{\sigma}_D^2, f_D(\bar{\sigma}_D^2)}(z) = \bar{Q}_D(z) \in \mathcal{RH}_\infty$*. In those cases,* $\bar{\gamma}_D$ *can be achieved with performance level no worse than* D*.*

Proof:

1. Since Assumption 4.3 holds, the problem of finding $\bar{\gamma}_D$ is feasible and $\bar{\sigma}_D^2 \in \mathbb{R}^+$. Also, it is clear that the inner optimization problem in (4.51a) is feasible if and only if $\inf_{Q(z)\in\mathcal{R}_\infty} J_{\sigma_q^2}(Q(z)) \leq D$. Therefore, it follows that all σ_q^2 of interest lie in Σ_D. We will thus focus on any fixed $\sigma_q^2 \in \Sigma_D$. Define

$$Q_{\sigma_q^2}(z) \triangleq \arg \inf_{\substack{Q(z)\in\mathcal{RH}_\infty \\ J_{\sigma_q^2}(Q(z))\leq D}} R_{\sigma_q^2}(Q(z)). \tag{4.78}$$

The well known KKT optimality conditions (see, e.g., [16,102]) state that $Q_{\sigma_q^2}(z)$ must be a stationary point of

$$\hat{L}(Q(z)) \triangleq \lambda_1 J_{\sigma_q^2}(Q(z)) + \lambda_2 R_{\sigma_q^2}(Q(z)), \tag{4.79}$$

for some non-negative λ_1 and λ_2, satisfying $\lambda_1 + \lambda_2 > 0$ and $\lambda_1 \left(J(Q_{\sigma_q^2}(z)) - D \right) = 0$. Since $\lambda_1 + \lambda_2 > 0$, we can define $\epsilon \triangleq \lambda_1(\lambda_1+\lambda_2)^{-1}$ (which is clearly in $[0,1]$). We thus conclude that $Q_{\sigma_q^2}(z)$ must be a stationary point of $L_{\sigma_q^2,\epsilon}$ for some $\epsilon \in [0,1]$ (see (4.56)). Moreover, since $L_{\sigma_q^2,\epsilon}$ is strictly convex in $Q(z)$, it has only one stationary point at $Q_{\sigma_q^2,\epsilon}(z)$ (see (4.57)). As a consequence, $Q_{\sigma_q^2}(z) = Q_{\sigma_q^2,\epsilon}(z)$ for some $\epsilon \in [0,1]$ such that $\epsilon \left(J_{\sigma_q^2}(Q_{\sigma_q^2}(z)) - D \right) = 0$.

If $J_{\sigma_q^2}(Q_{\sigma_q^2,0}) \geq D$, then, by convexity, there exists $\epsilon \in [0,1]$ such that $J_{\sigma_q^2}(Q_{\sigma_q^2,\epsilon}(z)) = D$. In these cases, it is immediate to see that the inequality constraint in (4.78) is active at the optimum. As a consequence, we have that, if $J_{\sigma_q^2}(Q_{\sigma_q^2,0}) \geq D$, then (recall the definition of f_D in (4.74))

$$Q_{\sigma_q^2}(z) = Q_{\sigma_q^2, f_D(\sigma_q^2)}(z). \tag{4.80}$$

If $J_{\sigma_q^2}(Q_{\sigma_q^2,0}) < D$, then, again by convexity, $J_{\sigma_q^2}(Q_{\sigma_q^2,\epsilon}(z)) < D$ for every $\epsilon \in [0,1]$ and the constraint on $J_{\sigma_q^2}$ is redundant. In those cases, one can choose $Q_{\sigma_q^2}(z) = Q_{\sigma_q^2,0}(z) = Q_{\sigma_q^2,f_D(\sigma_q^2)}(z)$, where we used the definition of f_D. (We note that in this case there are many $Q(z)$ that satisfy (4.78).)

The previous analysis provides an expression for the parameter $Q(z)$ that solves the inner problem in (4.51a) for every $\sigma_q^2 \in \Sigma_D$. To solve (4.51a), it suffices to pick the value of σ_q^2 that minimizes $\sigma_q^{-2} R_{\sigma_q^2}(Q_{\sigma_q^2}) = \sigma_q^{-2} R_{\sigma_q^2}(Q_{\sigma_q^2,f_D(\sigma_q^2)}(z))$. This is achieved solving the optimization problem in (4.75) which, by Assumption 4.3, always admits a solution. The result is now immediate.

2. This result follows from the proof of Part 1 above, the definition of $Q^{\varepsilon}_{\sigma_q^2,\epsilon}(z)$ in (4.67) and Theorem 4.3.

□□□

Theorem 4.5 (Characterization of $\left[\sigma_e^2\right]_\Gamma$) *Consider the setup in Figure 4.4 and suppose that Assumptions 4.2 and 4.3 hold. Define the set*

$$\Sigma_\Gamma \triangleq \left\{ \sigma_q^2 \in \mathbb{R}^+ : \inf_{Q(z)\in\mathcal{RH}_\infty} R_{\sigma_q^2}(Q(z)) \leq \Gamma\sigma_q^2 \right\} \tag{4.81}$$

and the function $f_\Gamma : \Sigma_\Gamma \to [0,1]$ defined via

$$f_\Gamma(\sigma_q^2) = \begin{cases} 1 & \text{if } \Gamma\sigma_q^2 > R_{\sigma_q^2}(Q_{\sigma_q^2,1}(z)), \\ \epsilon & \text{if } \epsilon \in [0,1] \text{ is such that } \Gamma\sigma_q^2 = R_{\sigma_q^2}(Q_{\sigma_q^2,\epsilon}(z)). \end{cases} \tag{4.82}$$

Then:

1. *The optimal noise variance $\bar{\sigma}_\Gamma^2$ and the optimal Youla parameter $\bar{Q}_\Gamma(z)$ are given by*

$$\bar{\sigma}_\Gamma^2 = \arg\min_{\sigma_q^2\in\Sigma_\Gamma} J_{\sigma_q^2}(Q_{\sigma_q^2,f_\Gamma(\sigma_q^2)}(z)), \quad \bar{Q}_\Gamma(z) = Q_{\bar{\sigma}_\Gamma^2,f_\Gamma(\bar{\sigma}_\Gamma^2)}(z). \tag{4.83}$$

Thus, the minimal stationary variance $\left[\sigma_e^2\right]_\Gamma$ satisfies

$$\left[\sigma_e^2\right]_\Gamma = J_{\bar{\sigma}_\Gamma^2}(\bar{Q}_\Gamma(z)). \tag{4.84}$$

2. *It is possible to get arbitrarily close to* $\left[\sigma_e^2\right]_\Gamma$*, while violating the signal-to-noise ratio constraint as little as desired, upon choosing* $\sigma_q^2 = \bar{\sigma}_\Gamma^2$ *and* $\bar{K}(z)$ *as in* (4.50) *with*

$$Q(z) = Q^{\varepsilon}_{\bar{\sigma}_\Gamma^2, f_\Gamma(\bar{\sigma}_\Gamma^2)}(z) \tag{4.85}$$

and $\varepsilon \in (0, 1]$ *sufficiently small. Moreover, if* $P_{22}(z)$ *has no poles on the unit circle and both* $P_{21}(z)$ *and* $E_{f_\Gamma(\bar{\sigma}_\Gamma^2)}(z)$ *have no zeros on the unit circle, then* $Q^{\varepsilon}_{\bar{\sigma}_\Gamma^2, f_\Gamma(\bar{\sigma}_\Gamma^2)}(z) = \bar{Q}_\Gamma(z) \in \mathcal{RH}_\infty$*. In those cases,* $\left[\sigma_e^2\right]_\Gamma$ *can be achieved with a signal-to-noise ratio no greater than* Γ*.*

Proof: The result follows from straightforward modifications of the proof of Theorem 4.4.

□□□

Theorems 4.4 and 4.5 provide an analytic characterization of the minimal signal-to-noise ratio $\bar{\gamma}$ that allows one to achieve a given performance level D, and of the minimal variance σ_e^2 that is achievable with a given signal-to-noise ratio constraint. Our results also provide a characterization of the noise variances $\sigma_q^2 \in \mathbb{R}^+$ and the controllers $\bar{K}(z) \in \mathcal{S}$ that allow one to achieve signal-to-noise ratios as close to $\bar{\gamma}_D$ as desired, and variances of e as close to $\left[\sigma_e^2\right]_\Gamma$ as desired, while violating the corresponding constraint by an arbitrarily small amount. (We also provided a characterization of cases where the constraints can be actually satisfied.)

The results of Theorems 4.4 and 4.5, although analytical, are not explicit. Nevertheless, they can be used as the basis of simple numerical algorithms to approximate $\bar{\gamma}_D$ or $\left[\sigma_e^2\right]_\Gamma$, and the associated optimal noise variances and controllers. The following conceptual algorithm illustrates the idea:

Algorithm 4.1 (Finding $\bar{\gamma}_D$ (resp. $\left[\sigma_e^2\right]_\Gamma$))

- Step 1: *Pick a grid* $\mathcal{G} = \{\sigma_1^2, \cdots, \sigma_n^2\}$*, where* $\sigma_i^2 \in \Sigma_D$ *(resp.* $\sigma_i^2 \in \Sigma_\Gamma$*)* $\forall i \in \{1, \cdots, n\}$*.*
- Step 2: *For every* $\sigma_i^2 \in \mathcal{G}$*, use Lemma 4.1 and Theorem 4.3 to determine the associated optimal trade-off curve in the* $J_{\sigma_q^2}$ versus $R_{\sigma_q^2}$ *plane. Find* $f_D(\sigma_i^2)$ *(resp.* $f_\Gamma(\sigma_i^2)$*).*
- Step 3: *For every* $\sigma_i^2 \in \mathcal{G}$ *evaluate* $\bar{\gamma}_i \triangleq \sigma_i^{-2} R_{\sigma_i^2}(Q_{\sigma_i^2, f_D(\sigma_i^2)}(z))$ *(resp.* $\left[\sigma_e^2\right]_i \triangleq J_{\sigma_i^2}(Q_{\sigma_i^2, f_\Gamma(\sigma_i^2)}(z))$*). Define* i^* *as the index* i *associated to the smallest* $\bar{\gamma}_i$ *(resp. smallest* $\left[\sigma_e^2\right]_i$*). Approximate* $\bar{\gamma}_D$*,* $\bar{\sigma}_D^2$ *and* $Q_{\bar{\sigma}_D^2, f_D(\bar{\sigma}_D^2)}(z)$ *(resp.* $\left[\sigma_e^2\right]_\Gamma$*,* $\bar{\sigma}_\Gamma^2$ *and* $Q_{\bar{\sigma}_\Gamma^2, f_\Gamma(\bar{\sigma}_\Gamma^2)}(z)$*) by* γ_{i^*}*,* $\bar{\sigma}_{i^*}^2$ *and* $Q_{\sigma_{i^*}^2, f_D(\sigma_{i^*}^2)}(z)$ *(resp.* $Q_{\sigma_{i^*}^2, f_\Gamma(\sigma_{i^*}^2)}(z)$*), respectively.*

□□

We note that both $f_D(\sigma_i^2)$ and $f_\Gamma(\sigma_i^2)$ can be found using standard bisection or any other line search algorithm. Algorithm 4.1 can be much improved if the grid $\mathcal{G}$ is iteratively modified as in, e.g., the so-called golden section optimization procedure (see, e.g., [127]), and Steps 2 and 3 are replaced by, e.g., an LMI based optimization procedure to immediately obtain $R_{\sigma_i^2}(Q_{\sigma_i^2, f_D(\sigma_i^2)}(z))$ or $J_{\sigma_i^2}(Q_{\sigma_i^2, f_\Gamma(\sigma_i^2)}(z))$ (see, e.g., [13, 15, 149]).

We end this section by identifying cases where the constraints in (4.51a) and (4.51b) are active at the optima. To that end, we denote by $T_{qe}(z)$ (resp. $T_{\hat{d}v}(z)$) the transfer function from q to e (resp. from $\hat{d}$ to v) in Figure 4.4.

Corollary 4.5 (Active constraints) *Consider the setup in Figure 4.4 and suppose that Assumptions 4.2 and 4.3 hold. If the optimal filter $\bar{K}_D(z)$ (resp. $\bar{K}_\Gamma(z)$) is such that $T_{qe}(z) \neq 0$ and $T_{\hat{d}v}(z)\Omega_{\hat{d}}(z) \neq 0$, then the inequality constraint in* (4.51a) *(resp.* (4.51b)*) is active at the optimum.*

Proof: We will prove our claim for the case of $\bar{K}_D(z)$ only; the case of $\bar{K}_\Gamma(z)$ follows using a similar argument.

Assume that $\bar{K}_o$ and σ_o^2 are optimal and that the constraint in (4.51a) is not active, i.e., $\bar{K}_o$ and σ_o^2 are such that

$$\bar{\gamma}_D = \bar{\gamma}\big|_{\substack{\bar{K}(z)=\bar{K}_o \\ \sigma_q^2=\sigma_o^2}}, \quad \sigma_e^2\big|_{\substack{\bar{K}(z)=\bar{K}_o \\ \sigma_q^2=\sigma_o^2}} < D. \tag{4.86}$$

Fix $\bar{K}(z) = \bar{K}_o$ and consider $\sigma_1^2 > \sigma_o^2$ such that

$$\sigma_e^2\big|_{\substack{\bar{K}(z)=\bar{K}_o \\ \sigma_q^2=\sigma_1^2}} < D \tag{4.87}$$

(since both Assumption 4.3 and (4.86) hold, this is always possible). If, for $\bar{K}(z) = \bar{K}_o$ and $\sigma_q^2 = \sigma_o^2$, we have $T_{qe}(z) \neq 0$ and $T_{\hat{d}v}(z)\Omega_{\hat{d}}(z) \neq 0$, then it follows from (4.46) that

$$\bar{\gamma}_D = \bar{\gamma}\big|_{\substack{\bar{K}(z)=\bar{K}_o \\ \sigma_q^2=\sigma_o^2}} > \bar{\gamma}\big|_{\substack{\bar{K}(z)=\bar{K}_o \\ \sigma_q^2=\sigma_1^2}}, \quad \sigma_e^2\big|_{\substack{\bar{K}(z)=\bar{K}_o \\ \sigma_q^2=\sigma_o^2}} < \sigma_e^2\big|_{\substack{\bar{K}(z)=\bar{K}_o \\ \sigma_q^2=\sigma_1^2}} < D. \tag{4.88}$$

This means that the pair $\bar{K}(z) = \bar{K}_o(z)$ and $\sigma_q^2 = \sigma_1^2$ is feasible and such that the corresponding signal-to-noise ratio is smaller than the optimal one. This contradicts the fact that $\bar{K}_o$ and σ_o^2 satisfy (4.86) and we conclude that the constraint must be active at the optimum. □□□

4.5 Design for Performance

In Section 4.4 we studied, in a general prototype setting, the problem of signal-to-noise ratio minimization subject to a performance constraint, and the problem of performance optimization subject to a signal-to-noise ratio constraint. In this section we will use those results to tackle the same problems in the networked control architecture introduced in Section 4.2 (see Figure 4.3). To that end we define the tracking error via

$$e \triangleq r - y, \tag{4.89}$$

and use its stationary variance σ_e^2 as performance measure. In particular, we will study the following:

Problem 4.1 (Optimal performance and minimal signal-to-noise ratio) *Consider the networked control situation in Figure 4.1, where the control and coding architecture is as in Figure 4.2, the channel is a signal-to-noise ratio constrained additive i.i.d. noise channel, and Assumption 4.1 holds.*

1. *Find (or prove the problem to be infeasible)*

$$\gamma_D \triangleq \inf_{\substack{K(z)\in\mathcal{S} \\ \sigma_q^2\in\mathbb{R}_0^+ \\ \sigma_e^2\leq D}} \gamma, \tag{4.90}$$

   where $0 \leq D < \infty$ *is the desired performance level, and* $\mathcal{S}$ *and* γ *are defined as in* (4.5) *and* (4.4)*, respectively.*[9] *If* γ_D *exists, then find the filter* $K(z) \in \mathcal{S}$ *and the channel noise variance* $\sigma_q^2 \in \mathbb{R}_0^+$ *that achieve* γ_D *(or approximate* γ_D *with arbitrary precision).*

2. *Find (or prove the problem to be infeasible)*

$$\left[\sigma_e^2\right]_\Gamma \triangleq \inf_{\substack{K(z)\in\mathcal{S} \\ \sigma_q^2\in\mathbb{R}_0^+ \\ \gamma\leq\Gamma}} \sigma_e^2, \tag{4.91}$$

   where $0 \leq \Gamma < \infty$ *is the signal-to-noise ratio constraint in the channel. If* $\left[\sigma_e^2\right]_\Gamma$ *exists, then find the filter* $K(z) \in \mathcal{S}$ *and the channel noise variance* $\sigma_q^2 \in \mathbb{R}_0^+$ *that achieve* $\left[\sigma_e^2\right]_\Gamma$ *(or approximate it with arbitrary precision).*

□□

[9] We note that, under Assumption 4.1, the constraints on $K(z)$ and σ_q^2 are necessary and sufficient to achieve MSS.

For future reference, we also define

$$\left(\sigma_D^2, K_D(z)\right) \triangleq \arg \inf_{\substack{K(z)\in\mathcal{S}\\ \sigma_q^2\in\mathbb{R}_0^+\\ \sigma_e^2\leq D}} \frac{\sigma_v^2}{\sigma_q^2}, \qquad \left(\sigma_\Gamma^2, K_\Gamma(z)\right) \triangleq \arg \inf_{\substack{K(z)\in\mathcal{S}\\ \sigma_q^2\in\mathbb{R}_0^+\\ \gamma\leq\Gamma}} \sigma_e^2. \tag{4.92}$$

As will become evident below, there exist cases where the *infima* in Problem 4.1 are not achievable with $K(z) \in \mathcal{S}$. In such cases, we will sometimes need to deal with families of controllers in $\mathcal{S}$ that allow one to get arbitrarily close to the minimal values of the considered functionals. To that end, we define the families $\{K_D^\varepsilon(z)\}_{\varepsilon\in(0,1]}$ and $\{K_\Gamma^\varepsilon(z)\}_{\varepsilon\in(0,1]}$, which are such that $K_D^\varepsilon(z)$ and $K_\Gamma^\varepsilon(z)$ belong to $\mathcal{S}$ for every $\varepsilon \in (0,1]$, and

$$\lim_{\varepsilon\to 0^+} \sigma_e^2\big|_{\substack{K(z)=K_D^\varepsilon(z)\\ \sigma_q^2=\sigma_D^2}} \leq D, \qquad \lim_{\varepsilon\to 0^+} \gamma\big|_{\substack{K(z)=K_D^\varepsilon(z)\\ \sigma_q^2=\sigma_D^2}} = \gamma_D, \tag{4.93a}$$

$$\lim_{\varepsilon\to 0^+} \sigma_e^2\big|_{\substack{K(z)=K_\Gamma^\varepsilon(z)\\ \sigma_q^2=\sigma_\Gamma^2}} = \left[\sigma_e^2\right]_\Gamma, \qquad \lim_{\varepsilon\to 0^+} \gamma\big|_{\substack{K(z)=K_\Gamma^\varepsilon(z)\\ \sigma_q^2=\sigma_\Gamma^2}} \leq \Gamma. \tag{4.93b}$$

In the remainder of this section we will show how to use the results of Section 4.4 to give an answer to Problem 4.1. We will start by considering a one-dof architecture in Section 4.5.1. We will address the general case in Section 4.5.2.

4.5.1 One degree-of-freedom architecture

In this case, we limit our analysis to $K(z) \in \mathcal{S}_1$, where (recall the definition of $K(z)$ in (4.3))

$$\mathcal{S}_1 \triangleq \left\{K(z) \in \mathcal{S} : M(z) = C(z)\begin{bmatrix}1 & -1\end{bmatrix}, F(z) = \begin{bmatrix}0 & 1\end{bmatrix}, C(z) \in \mathcal{R}_p\right\}. \tag{4.94}$$

Provided Assumption 4.1 holds, $K(z) \in \mathcal{S}_1$ and $\sigma_q^2 \in \mathbb{R}_0^+$, both σ_v^2 and σ_e^2 are well defined, finite and satisfy

$$\sigma_e^2 = \left|\left|\begin{bmatrix}\Omega_r(z)\\ \Omega_d(z)\end{bmatrix} S(z)\right|\right|_2^2 + \sigma_q^2 \,||T(z)||_2^2, \tag{4.95a}$$

$$\sigma_v^2 = \left|\left|\begin{bmatrix}T(z)\Omega_r(z)\\ S(z)\Omega_d(z)\end{bmatrix}\right|\right|_2^2 + \sigma_q^2 \,||T(z)||_2^2, \tag{4.95b}$$

where

$$S(z) \triangleq \frac{1}{1+G(z)C(z)}, \quad T(z) \triangleq 1 - S(z). \tag{4.96}$$

Define

$$D_{\inf} \triangleq \inf_{\substack{K(z)\in\mathcal{S}_1\\ \sigma_q^2=0}} \sigma_e^2, \quad D_{\sup} \triangleq \begin{cases} ||\Omega_r(z)||_2^2 + ||\Omega_d(z)||_2^2 & \text{if } G(z) \text{ is stable,}\\ \infty & \text{if } G(z) \text{ is unstable.}\end{cases} \tag{4.97}$$

$D_{\inf}$ corresponds to the best performance that is achievable in a one-dof control loop when transparent communication channels are used. As such, $D_{\inf}$ can be evaluated using standard $\mathcal{H}_2$ optimization techniques (see, e.g, [113, 194]) and bounds from below the achievable performance in the considered networked situation. $D_{\sup}$ corresponds to the stationary tracking error variance that is achieved when the plant is left in open loop. We also recall from Theorem 4.2 and Remark 4.4 that, when $K(z)$ is constrained to belong to $\mathcal{S}_1$, the minimal stationary signal-to-noise ratio compatible with MSS, i.e., the minimal signal-to-noise ratio that is achievable when no attention is paid to the size of σ_e^2 (hence, $D \to \infty$), is given by

$$\gamma_{\inf} = \begin{cases} 0 & \text{if } G(z) \text{ is stable,} \\ \left(\prod_{i=1}^{n_p} |p_i|^2\right) - 1 + \Delta_G & \text{if } G(z) \text{ is unstable.} \end{cases} \tag{4.98}$$

With the previous definitions it is easy to see that the following holds:

Fact 4.1 (Extreme cases of Problem 4.1) *Consider the networked control situation in Figure 4.1, where the control and coding architecture is as in Figure 4.2 with $K(z) \in \mathcal{S}_1$, the channel is an additive i.i.d. noise channel with signal-to-noise ratio constraint Γ, and Assumption 4.1 holds. Then:*

1. *If the plant is unstable and $D \to D_{\sup} = \infty$ (resp. $D \geq D_{\sup}$ in the case of stable plant models), then $\gamma_D \to \gamma_{\inf}$ (resp. $\gamma_D = \gamma_{\inf} = 0$). If $D \to D_{\inf}$, then $\gamma_D \to \infty$ unless the plant is stable, $d = 0$ and the controller that achieves $D_{\inf}$ is identically zero.*[10]

2. *If $\Gamma \to \gamma_{\inf}$, then $\sigma_e^2 \to D_{\sup}$. If $\Gamma \to \infty$, then $\sigma_e^2 \to D_{\inf}$.*

Proof: Immediate from (4.95) and the definitions of $D_{\inf}$, $D_{\sup}$ and $\gamma_{\inf}$. □□□

Fact 4.1 (and the discussion leading to it) allow one to see that Theorem 4.2, Remark 4.4 and standard $\mathcal{H}_2$ optimization techniques provide characterizations of σ_D^2 and $\left[\sigma_e^2\right]_\Gamma$, when $D \in \{D_{\inf}, D_{\sup}\}$ and $G \in \{\gamma_{\inf}, \infty\}$. Accordingly, in the sequel we will focus on performance levels and signal-to-noise ratio constraints that satisfy $D_{\inf} < D < D_{\sup}$ and $\gamma_{\inf} < G < \infty$, respectively.

If neither r nor d are zero, then it is straightforward to see that, when $K(z)$ is constrained

[10] This is a pathological case.

to belong to $\mathcal{S}_1$, Problem 4.1 fits into the setting examined in Section 4.4 where

$$P(z) = \left[\begin{array}{ccc:c} 0 & 1 & 0 & G(z) \\ 1 & -1 & 0 & -G(z) \\ \hdashline 1 & -1 & -1 & -G(z) \end{array}\right], \tag{4.99}$$

$\bar{K}(z) = C(z)$, $\hat{d} = \begin{bmatrix} r & d \end{bmatrix}^T$, $\bar{u} = u$, and $\bar{y} = r - q - d - G(z)u$. If either r or d is zero,[11] then $\hat{d}$ equals the non-zero signal, and $P(z)$ and $\bar{y}$ need to be modified accordingly. In either case, it is immediate to see that if the conditions of Problem 4.1 are satisfied, then Assumption 4.2 also holds.

In principle, the characterizations of γ_D and $\left[\sigma_e^2\right]_\Gamma$ follow immediately from the results in Section 4.5. There is, however, a technical detail that we need to clarify. The results in Section 4.5 assume that D (resp. Γ) is such that the problem of finding γ_D (resp. $\left[\sigma_e^2\right]_\Gamma$) is feasible and that the optimal noise variance (optimal channel noise variance in our case), i.e., σ_D^2 (resp. σ_Γ^2), belongs to $\mathbb{R}^+$ (see Assumption 4.3). We next show that this is actually the case for all cases of interest:

Lemma 4.3 *Consider the networked control situation in Figure 4.1, where the control and coding architecture is as in Figure 4.2 with $K(z) \in \mathcal{S}_1$, the channel is an additive i.i.d. noise channel with signal-to-noise ratio constraint Γ, and Assumption 4.1 holds. Then:*

1. *If $D_{\inf} < D < D_{\sup}$, then the problem of finding γ_D is feasible, $\sigma_D^2 \in \mathbb{R}^+$ and the constraint in* (4.90) *is active at the optimum.*

2. *If $\gamma_{\inf} < \Gamma < \infty$, then the problem of finding $\left[\sigma_e^2\right]_\Gamma$ is feasible, $\sigma_\Gamma^2 \in \mathbb{R}^+$ and the constraint in* (4.91) *is active at the optimum.*

Proof:

1. It is immediate to see from the definition of $D_{\inf}$ that $D > D_{\inf}$ guarantees feasibility.

   To see that $\sigma_D^2 \in \mathbb{R}^+$, we will first discard the case $\sigma_D^2 \to \infty$. Since $D < D_{\sup}$, it follows that σ_q^2 cannot be arbitrarily large in order to satisfy the performance constraint. Thus, necessarily, $\sigma_D^2 < \infty$. On the other hand, if $\sigma_q^2 = 0$ then $\gamma \to \infty$ (unless the plant is stable and left in open loop, which, since $D < D_{\sup}$, is not feasible). Since $D > D_{\inf}$, there exists $\sigma_q^2 > 0$ such that the performance constraint is satisfied and $\gamma < \infty$. Thus, $\sigma_D^2 > 0$.

[11] Assumption 4.1 precludes both signals to be simultaneously zero.

To see that the constraint must be active at the optimum, we will use Corollary 4.5. It is easy to see from (4.95) that, in this case, $T_{qe}(z) = T(z) = 0$ and $T_{\hat{d}v}(z)\Omega_{\hat{d}}(z) = T(z)\Omega_r(z) + S(z)\Omega_d(z) = 0$ if and only if $C(z) = 0$ and $d = 0$ (recall that we assume that either r and d are not simultaneously zero), which is acceptable only if $G(z)$ is stable. Thus, the constraint might not be active only if $G(z)$ is stable. But, since $D < D_{\text{sup}}$, the controller $C(z) = 0$ is not feasible even for stable plants and the result follows.

2. Analogous to the proof of Part 1 above.

□□□

From Lemma 4.3 we see that $D_{\text{inf}} < D < D_{\text{sup}}$ and $\gamma_{\text{inf}} < \Gamma < \infty$ guarantee that Assumption 4.3 is satisfied. Hence, the results in Section 4.4 yield an immediate characterization of both γ_D and $\left[\sigma_e^2\right]_\Gamma$, the optimal quantization noise variances σ_D^2 and σ_Γ^2, and the optimal filters $K_D(z)$ and $K_\Gamma(z)$ for all non-trivial cases of interest. We note that the results in Section 4.4 characterize the transfer function $\bar{K}(z) = C(z)$ in the optimal filters $K_D(z)$ and $K_\Gamma(z)$ (see (4.94)) in terms of a Youla parameter that, in general, belongs to $\overline{\mathcal{RH}}_\infty$. Thus, $K_D(z)$ and $K_\Gamma(z)$ do not belong to $\mathcal{S}_1$ in general. To construct families of functions in $\mathcal{S}_1$ that allow one to approximate either $K_D(z)$ or $K_\Gamma(z)$ to any degree of accuracy, it suffices to use Part 2 of Theorems 4.4 and 4.5, respectively.

Remark 4.7 (Relationship to other work) *The recent work [48] also addresses the problem of optimal networked control system design in one-dof architectures, where additive i.i.d. noise channels with signal-to-noise ratio constraints are employed in the feedback path. However, the authors of [48] limit their analysis to a case where the plant is MP, has relative degree one, and* $r = 0$. *The latter assumption greatly simplifies the solution to the problem. Indeed, if* $r = 0$, *then the output of the plant is both the signal whose variance needs to be minimized and also the signal where the communication constraint manifest itself. More precisely, the setup in [48] leads to an optimization problem where the Lagrangian equals the objective function (in the jargon of Section 4.4,* $L_{\sigma_q^2,\epsilon} = \lambda J_{\sigma_q^2}$, *where* λ *is a multiplier). The results in this section are, hence, an extension of the results in [48]. It is worth mentioning that the results in [48] show that the signal-to-noise ratio constraint is active at the optimum, which is consistent with our results.*

□□

Remark 4.8 (Other architectures) *We note that our previous results apply,* mutatis mutandis*, to control and coding architectures where either* $F(z)$ *or* $M(z)$ *are fixed, and the non-fixed* *transfer function has to be designed.* □□

4.5.2 The general architecture

In this section we return to the control architecture in Figure 4.3 without constraints on $M(z)$ or $F(z)$, i.e., we allow $K(z)$ to be chosen freely in $\mathcal{S}$. In principle, one could think of proceeding as in the one-dof architecture by applying the results of Section 4.4 to this case. However, there exists a fundamental difficulty when trying to do so. Indeed, provided neither r nor d are zero,[12] the setup in this section corresponds to the setup examined in Section 4.4 with

$$\bar{K}(z) \triangleq \begin{bmatrix} 0 & M_2(z) & M_1(z) \\ F_2(z) & F_1(z) & 0 \end{bmatrix}, \quad P_{22}(z) = \begin{bmatrix} G(z) & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}. \tag{4.100}$$

We note that the condition $\bar{K}_{11}(z) = \bar{K}_{23}(z) = 0$ originates in the communication constraint: the plant input u can only depend on r and w, and the channel input v can only depend on y and w. Having either $\bar{K}_{11}(z) \neq 0$ or $\bar{K}_{23}(z) \neq 0$ would necessitate the use of additional communication channels (which has been ruled out in our framework). We thus conclude that, in the present case, the synthesis of $\bar{K}(z)$ in Figure 4.4 is equivalent to the problem of synthesizing a *structurally constrained* controller. A general solution to such problems remains open. Several alternative procedures have been explored focusing mainly on the case of diagonal controllers (see, e.g., [20, 23, 63, 64, 90, 144, 146] and the many references therein).

In our case, it is fairly simple to characterize the set of all admissible controllers $\bar{K}(z)$. Indeed, Corollary 4.1 provides a simple characterization of all admissible $M(z)$ and $F(z)$. The problem with that characterization is that it is not explicit since it relies on an initial choice for $M(z)$. This limitation was dealt with in Section 4.3 by means of obtaining explicit analytic solutions to the optimization problems of interest. Those explicit characterizations allowed us to optimally choose $M(z)$ and $F(z)$. However, in the situation of interest in this section, it seems very difficult to understand how the initial choice for $M(z)$ affects the minimal signal-to-noise ratio or the optimal performance. Therefore, it seems difficult to see how to choose the best initial $M(z)$.

The problems mentioned in last paragraph may be overcome if it were possible to obtain an explicit characterization of all admissible $\bar{K}(z)$ satisfying (4.100). This could be achieved by extending the results in Section II of [169]. Unfortunately, the resulting characterization will, in

[12] Similar conclusions apply when either r or d are zero.

general, be non-convex[13] and hence difficult to use (see, e.g., [128, 140, 169] and also [14] for related results). This leads to the long-standing problem of what is the broadest class of structural constraints that lead to a convex characterization of the set of admissible controllers that have a given structure. To the best of the author's knowledge, the broadest known class of structurally constrained control problems that are amenable to be stated as convex problems was identified in [140]. That work defines the notion of *quadratic invariance* of a constraint on the controller structure with respect to a given plant model. If quadratic invariance holds for the problem at hand, then all admissible structurally constrained controllers can be parameterized as in the standard Youla parametrization with a parameter that inherits the structure imposed on the controller (note that the condition is only sufficient; see also [161] for additional results applicable to sparsity constraints). If the set of all transfer functions that have the desired controller structure is convex, then the resulting parametrization is convex in the Youla parameter.

We will next check whether the results in [140] are useful in our case. Denote by $\mathcal{M}$ the set of all controllers $\bar{K}(z)$ satisfying (4.100). According to [140], the Youla parameter associated with any admissible $\bar{K}(z) \in \mathcal{M}$ would inherit the structure of the controller if and only if, for every $\bar{K}(z) \in \mathcal{M}$,

$$\bar{K}(z)\begin{bmatrix} G(z) & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}\bar{K}(z) = \begin{bmatrix} M_2(z)F_2(z) & M_2(z)F_1(z) & 0 \\ F_1(z)F_2(z) & F_2(z)G(z)M_2(z) + F_1(z)^2 & F_2(z)G(z)M_1(z) \end{bmatrix} \in \mathcal{M} \tag{4.101}$$

i.e., if and only if $M_2(z)F_2(z) = 0$ and $F_2(z)G(z)M_1(z) = 0$, which is equivalent to leaving the plant in open loop. This case is, unfortunately, of very little interest and is compatible with MSS if and only if $G(z)$ is stable. We thus conclude that, given the state of knowledge on structurally constrained optimal control problems, the problem of optimally designing $\bar{K}(z)$ in Figure 4.4 subject to (4.100) is not currently amenable to be stated as a convex problem. As a consequence, the study of suboptimal synthesis procedures emerges as a relevant task.

We next present a simple iterative algorithm that allows one to approximate the solution to Problem 4.1 in the present case:

Algorithm 4.2 (Finding γ_D (resp. $\left[\sigma_e^2\right]_\Gamma$)in the general architecture)

- Step 1: *Pick a tolerance value $\rho > 0$ and set $M(z) = M_o(z)$, where $M_o(z) \in \mathcal{R}_p$ has a detectable and stabilizable underlying realization such that the realization* (4.7) *is detectable*

[13] And indeed is non-convex in the present case.

and stabilizable, and also satisfies

$$\inf_{\substack{K(z)\in\mathcal{S} \\ \sigma_q^2=0 \\ M(z)=M_o(z)}} \sigma_e^2 \leq D. \tag{4.102}$$

(resp. satisfies

$$\inf_{\substack{K(z)\in\mathcal{S} \\ \sigma_q^2\in\mathbb{R}_0^+ \\ M(z)=M_o(z)}} \gamma \leq \Gamma.) \tag{4.103}$$

Set $i = 0$ and $\gamma_i = \infty$ (resp. $\sigma_i^2 = \infty$).

- Step 2: *Set $i = i + 1$ and fix $M(z)$ at its current value. Choose σ_q^2 and $F(z)$ so as to minimize γ (resp. σ_e^2) subject to MSS and the constraint $\sigma_e^2 \leq D$ (resp. $\gamma \leq \Gamma$). Denote the minimal γ by γ_i (resp. the minimal σ_e^2 by σ_i^2).*
- Step 3: *Set $i = i + 1$ and fix $F(z)$ at its current value. Choose σ_q^2 and $M(z)$ so as to minimize γ (resp. σ_e^2) subject to MSS and $\sigma_e^2 \leq D$ (resp. $\gamma \leq \Gamma$). Denote the minimal γ by γ_i (resp. the minimal σ_e^2 by σ_i^2).*
- Step 4: *If $|\gamma_i - \gamma_{i-1}|\,\gamma_i^{-1} \leq \rho$ (resp. $\left|\sigma_i^2 - \sigma_{i-1}^2\right| \sigma_i^{-2} \leq \rho$), then* Stop. *Else, return to* Step 2.

□□

It is clear that Algorithm 4.2 is such that $0 \leq \gamma_i \leq \gamma_{i-1}$ (or, alternatively, $0 \leq \sigma_i^2 \leq \sigma_{i-1}^2$) for every $i \geq 1$. Therefore, Algorithm 4.2 is guaranteed to converge to a local minimum and, hence, it can be used to find approximations (or good candidates) for γ_D and $\left[\sigma_e^2\right]_\Gamma$.

Once a suitable $M_o(z)$ has been found, the optimization problems in Steps 2 and 3 of Algorithm 4.2 are guaranteed to be feasible. These problems are special cases of the problem studied in Section 4.4. In particular, we have the following relationships:

- *Step 2 (choosing $F(z)$ and σ_q^2):* If both r and d are non zero, then the problem at this stage is equivalent to the problem in Section 4.4 with

$$P(z) = \left[\begin{array}{ccc:c} 0 & 0 & 0 & 1 \\ (1 - G(z)M_1(z)) & -1 & -G(z)M_2(z) & -G(z)M_2(z) \\ \hdashline G(z)M_1(z) & 1 & G(z)M_2(z) & G(z)M_2(z) \\ 0 & 0 & z^{-1} & z^{-1} \end{array}\right], \tag{4.104}$$

$$\bar{K}(z) = F(z)\begin{bmatrix} z & 0 \\ 0 & 1 \end{bmatrix}, \tag{4.105}$$

$\hat{d} = \begin{bmatrix} r & d \end{bmatrix}^T$, $\bar{u} = v$, and $\bar{y} = \begin{bmatrix} y & z^{-1}(q+v) \end{bmatrix}^T$.

- *Step 3 (choosing $M(z)$ and σ_q^2):* If both r and d are non zero, then the problem at this stage is equivalent to the problem in Section 4.4 with

$$P(z) = \left[\begin{array}{cccc} 0 & F_2(z)S_1(z) & F_1(z)S_1(z) & F_2(z)G(z)S_1(z) \\ 1 & -1 & 0 & -G(z) \\ \hline 0 & F_2(z)S_1(z) & S_1(z) & F_2(z)G(z)S_1(z) \\ 1 & 0 & 0 & 0 \end{array}\right], \tag{4.106}$$

$S_1(z) \triangleq (1 - F_1(z))^{-1}$, $\bar{K}(z) = M(z)$, $\hat{d} = \begin{bmatrix} r & d \end{bmatrix}^T$, $\bar{u} = u$, and $\bar{y} = \begin{bmatrix} w & r \end{bmatrix}^T$.

As was the case for the one-dof architecture studied before, one needs to modify $P(z)$, $\hat{d}$ and $\bar{y}$ if either r or d are zero in order for the results of Section 4.4 to apply. Also, it is straightforward to adapt the reasoning in Section 4.5.1 to show that, for all non-trivial cases of interest, the optimization problems in both Steps 2 and 3 satisfy Assumption 4.3. Accordingly, the results in Section 4.4 yield an immediate characterization of both γ_i and σ_i^2 at each iteration step. Thus, use of Algorithm 4.2 and the results of Section 4.4 allow one to approximate γ_D and $\left[\sigma_e^2\right]_\Gamma$, the corresponding optimal filters $K_D(z)$ and $K_\Gamma(z)$, and the optimal quantization noise variances σ_D^2 and σ_Γ^2. As was the case for the one-dof architecture considered before, the (algorithmic) characterization of $K_D(z)$ and $K_\Gamma(z)$ described above does not, in general, lead to filters in $\mathcal{S}$. However, in this case, it suffices to use Part 2 of Theorems 4.4 and 4.5 to construct families of functions in $\mathcal{S}$ satisfying (4.93).

Remark 4.9 (Other architectures) *The previous ideas apply,* mutatis mutandis*, to any control and coding architecture where structural constraints are imposed on $F(z)$ or $M(z)$. In particular, they apply to the architecture without feedback described in Section 4.3.* □□

4.6 An Example

In this section we provide a simple example to illustrate the results of this chapter. Consider the unstable plant model

$$G(z) = \frac{z - 0.8}{z(z-2)}, \tag{4.107}$$

and assume that $d = 0$ and that the reference r has a spectral factor given by

$$\Omega_r(z) = \frac{0.1z}{z - 0.9}. \tag{4.108}$$

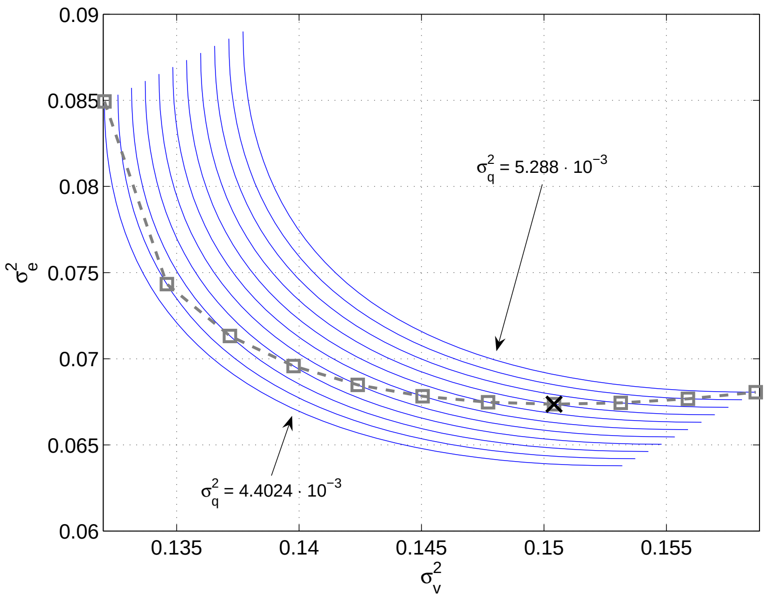

Figure 4.5: Optimal trade-off curves in the σ_e^2 versus σ_v^2 plane for $\Gamma = 30$ and $\sigma_q^2 \in [4.402 \cdot 10^{-3}, 5.288 \cdot 10^{-3}]$ (see text for details).

We will design a one-dof architecture so as to minimize the stationary tracking error variance in the face of a signal-to-noise ratio constraint. For this plant and architecture, straightforward calculations based on the results of this chapter lead to

$$\gamma_{\text{inf}} = 3, \quad D_{\text{inf}} = 0.04, \quad D_{\text{sup}} = \infty. \tag{4.109}$$

We will start by examining optimal trade-off curves in the $J = \sigma_e^2$ versus $R = \sigma_v^2$ plane for different values of σ_q^2, when the maximum allowable signal-to-noise ratio is given by $\Gamma = 10 \cdot \gamma_{\text{inf}} = 30$. For this situation,

$$\Sigma_\Gamma = \Sigma_{30} = \{\sigma_q^2 \in \mathbb{R}_0^+ : \sigma_q^2 \geq 4.402 \cdot 10^{-3}\}. \tag{4.110}$$

We also note that $R(Q_{\sigma_q^2,1}(z)) \leq \Gamma\sigma_q^2$ for all $\sigma_q^2 \leq 5.288 \cdot 10^{-3}$. Therefore, and given the fact that we know that the constraint is active at the optimum in this case, we will focus on $\sigma_q^2 \in [4.402 \cdot 10^{-3}, 5.288 \cdot 10^{-3}]$. Figure 4.5 shows optimal trade-off curves for ten values of σ_q^2 (chosen so that the values for σ_q are uniformly distributed in $[\sqrt{4.402 \cdot 10^{-3}}, \sqrt{5.288 \cdot 10^{-3}}]$). The values of σ_v^2 at the squares on each of the optimal trade-off curves, correspond to the upper limit on the admissible values for σ_v^2 for each value of σ_q^2, i.e., they are equal to $\Gamma\sigma_q^2$. On the other hand, the values for σ_e^2 at the squares correspond to the minimal values of σ_e^2 that are

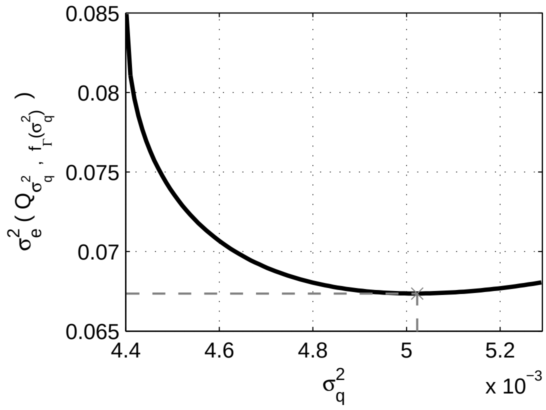

Figure 4.6: Minimal value of σ_e^2 that is achievable with $\gamma \leq \Gamma = 30$, and a fixed value for σ_q^2, as a function of σ_q^2.

consistent with the signal-to-noise ratio constraint for the corresponding value of σ_q^2, i.e., they are equal to $\sigma_e^2(Q_{\sigma_q^2, f_\Gamma(\sigma_q^2)}(z))$. That is, the values of σ_e^2 at the squares in Figure 4.5 correspond to candidates to be equal to the minimum of σ_e^2 when the channel signal-to-noise ratio is upper bounded by $\Gamma = 30$, i.e., are candidates to be equal to $\left[\sigma_e^2\right]_\Gamma = \left[\sigma_e^2\right]_{30}$. Indeed, the crossed square indicates the minimal value of σ_e^2 for the considered grid for σ_q^2.

We next consider a one hundred points grid for σ_q^2, and plot $\sigma_e^2(Q_{\sigma_q^2, f_\Gamma(\sigma_q^2)}(z))$ as a funtion of σ_q^2. The results are shown in Figure 4.6. The cross in Figure 4.6 marks the minimal value of $\sigma_e^2(Q_{\sigma_q^2, f_\Gamma(\sigma_q^2)}(z))$, i.e., the best performance $\left[\sigma_e^2\right]_\Gamma = \left[\sigma_e^2\right]_{30}$, which turns out to be given by

$$\left[\sigma_e^2\right]_{30} = 6.736 \cdot 10^{-2}. \tag{4.111}$$

The corresponding optimal channel noise variance σ_Γ^2 and optimal controller $C_\Gamma(z)$ are given by

$$\sigma_{30}^2 = 5.023 \cdot 10^{-3}, \quad C_{30}(z) = \frac{2.1871z(z-0.801)}{(z-0.9389)(z-0.8)}. \tag{4.112}$$

We also computed $[\sigma_e^2]_\Gamma$ for several values of $\Gamma \in [\gamma_{\text{inf}}, \gamma_{\text{inf}} \cdot 10^4] = [3, 30000]$. The results are shown in Figure 4.7, where we also include a plot of σ_Γ^2 as function of Γ. As expected, $[\sigma_e^2]_\Gamma$ is a decreasing function of Γ that tends to D_{inf} as $\Gamma \to \infty$, and that grows unbounded when $\Gamma \to \gamma_{\text{inf}} = 3$ (see horizontal and vertical lines in Figure 4.7 (left), respectively). We also note that for moderate values of Γ, say $\Gamma = 100 \cdot \gamma_{\text{inf}} = 300$, the best achievable performance is almost indistinguishable from the best non-networked performance D_{inf}. On the other hand (and, again, as expected), σ_Γ^2 decreases with Γ and tends to zero when $\Gamma \to \infty$, while growing unbounded when $\Gamma \to \gamma_{\text{inf}}$ (see vertical asymptote in Figure 4.7 (right)). For completeness,

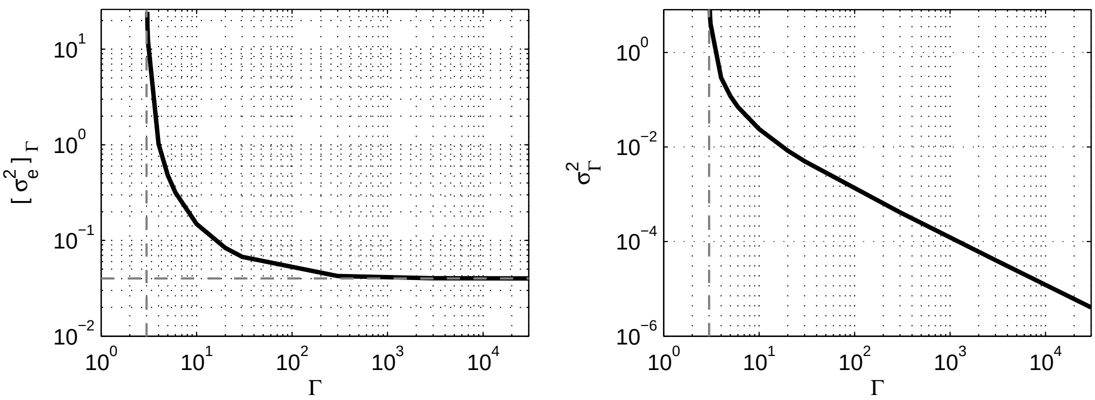

Figure 4.7: Optimal performance $\left[\sigma_e^2\right]_\Gamma$ and optimal channel noise variance σ_Γ^2 as a function of Γ (vertical lines are at $\Gamma = \gamma_{\text{inf}} = 3$; the horizontal line on the left is located at $\left[\sigma_e^2\right]_\Gamma = \left[\sigma_e^2\right]_\infty = D_{\text{inf}}$).

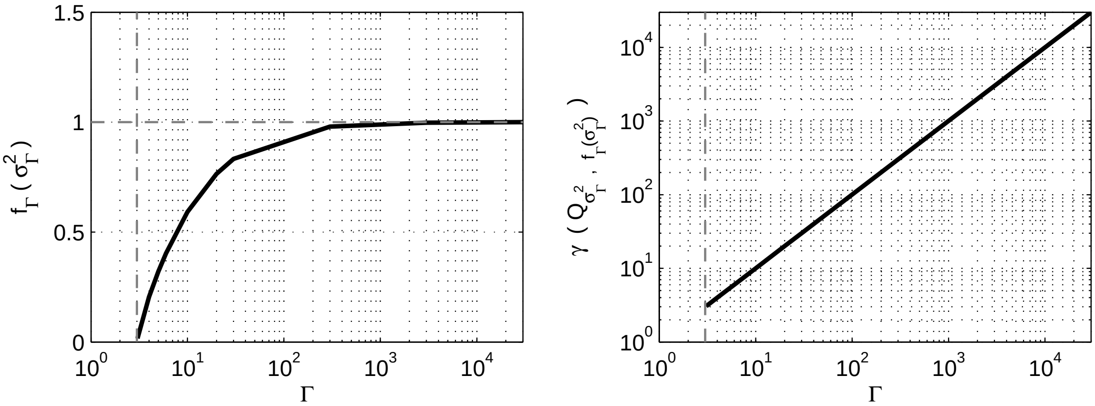

Figure 4.8: $f_\Gamma(\sigma_\Gamma^2)$ and values of γ at the optima as functions of Γ (vertical lines are at $\Gamma = \gamma_{\text{inf}} = 3$).

Figure 4.8 shows both $f_\Gamma(\sigma_\Gamma^2)$ and the values of γ at the optima as functions of Γ. It is seen that $f_\Gamma(\sigma_\Gamma^2) \to 1$ as $\Gamma \to \infty$ and that $f_\Gamma(\sigma_\Gamma^2) \to 0$ as $\Gamma \to \gamma_{\text{inf}}$. This behavior should be intuitively clear: if $\Gamma \to \infty$, then the constraint on γ is redundant and, accordingly, the optimization does not need to pay attention to the size of σ_v^2 but only to the size of σ_e^2. That is, one can choose $\epsilon = f_\Gamma(\sigma_\Gamma^2) = 1$ in the associated Lagrangian. (A similar argument justifies the behavior $f_\Gamma(\sigma_\Gamma^2) \to 0$ as $\Gamma \to \gamma_{\text{inf}}$.) Figure 4.8 (left) also shows that the constraint on γ is always active, as predicted by our results.

4.7 Summary

This chapter has presented a general framework to deal with control problems with signal-to-noise ratio constraints. In particular, we have focused on studying the interplay between closed loop performance and signal-to-noise ratio constraints. In the most general case, our results give an algorithmic characterization of the best performance that is achievable when a signal-to-noise ratio constraint is imposed, and also of the minimal signal-to-noise ratio that is required to attain a desired performance level. For simple situations, our results admit closed form expressions and properties of the corresponding solutions have been presented.

We have also studied the minimal signal-to-noise ratio needed to achieve MSS in the architecture of interest. Our results show that the use of coding architectures with feedback are essential to recover, for any plant model and any combination of external noise sources, the bound on the channel signal-to-noise ratio for MSS presented in [18] for the noiseless static state feedback case.

The results in this chapter serve as backbone for the rest of the thesis, where we will show that two important networked control design problems, namely control design with either average data-rate constraints or i.i.d. data dropouts, can be written as control problems with signal-to-noise ratio constraints.

Extensions of the results presented in this chapter lie in the consideration of MIMO plant models and architectures with multiple channels. (For example, one could consider a situation where several parallel channels are used to send plant measurements to the controller (see also [133]), or, alternatively, a situation where both measurements and control signals are transmitted over non-transparent communication channels.) Given the fact that our framework relies on convex-optimization-related ideas, these extensions seem possible at the cost of (not too much) additional effort.

Chapter 5

Average Data-Rate Limits

5.1 Introduction

This chapter is intended to be a link between information theory and control theory. Our aim is to obtain a framework that allows one to analyze and design control systems subject to average data-rate constraints.

To motivate the problem of interest, we will start by considering the basic communication scheme illustrated in Figure 5.1. In that figure, v is a sequence of continuous random variables that needs to be transmitted over a physical medium. We will refer to the physical medium as the *underlying channel*.[1] We will focus on digital communications, i.e., on a situation where the information conveyed through the underlying channel belongs to a discrete set. In general, the underlying channel will introduce errors and, hence, the need arises to carefully choose what to transmit in order to be able to reliably communicate with the receiving end (see, e.g., Chapter 7 in [31]). Moreover, since we assume v to be a continuous random variable, the need to quantize arises (see, e.g., Chapter 10 in [31]). We thus conclude that, in order to communicate v through the underlying channel, *one needs to code v* in an appropriate way. It was shown by Shannon that, if mild conditions are satisfied, then the problem of optimally coding v can be separated into two independent problems (see, e.g., [31, 51, 156]): *source coding* and *channel coding* (see the individual blocks in Figure 5.1). The problem of source coding deals with the parsimonious representation of continuous random variables, using a finite or countable set of symbols, and introducing the smallest possible error as measured by a suitable distortion measure (also called

[1] As will become clear as we proceed, this notion is not to be confused with that of *channel* (without qualifier).

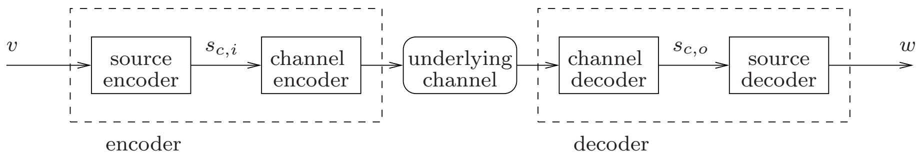

Figure 5.1: Communication system with separate source and channel coding.

fidelity criterion). In other words, the source coding problem is an *optimal quantization problem.* This problem lies at the very core of rate distortion theory (see, e.g., [10, 31, 157]). On the other hand, the problem of channel coding studies ways to reliably communicate over a given underlying channel, i.e., methodologies to make the transmission error probability as small as possible (see [31]). It should be clear that, necessarily, *source coding extracts redundancy* from the source v, whereas *channel coding adds redundancy* to what is intended to be sent over the channel in order to be able to, somehow, counteract channel errors. It is worth mentioning that enforcing a separation between channel and source coding may not be always convenient from a practical point of view. This is due to the fact that, when doing so, the resulting schemes usually become very complex and may introduce unbounded delays. This has motivated the study of situations where coding is not actually necessary [51].[2]

As foreshadowed above, rate distortion theory deals with optimal quantization problems and, consistent with Shannon's separation theorem, *does not take channel coding into account.* That is, rate distortion theory assumes that the link between $s_{c,i}$ and $s_{c,o}$ in Figure 5.1 is *transparent* and focuses on how to quantize efficiently [10, 31, 51, 59]. Source encoders (or quantizers[3]) can have a fixed number of output levels, each one with the same length, or can have variable length output symbols. The first case is, in general, difficult to analyze (see references in [59]), whereas the latter case lends itself to fairly easy treatment using the concept of entropy introduced by Shannon in [156] (see also [31]).

In this chapter we adopt a purely source coding perspective, and consider source encoders with variable length output symbols. In this context, the average rate at which the quantized symbols (i.e., $s_{c,i}$ in Figure 5.1) are generated or, in short, the *average data-rate*, is defined as the average length of the quantizer output symbols (measured in bits or nats). We note

[2] Needless to say, if our objective is to communicate analog data using digital words, then coding becomes essential.

[3] It is correct to think of a source encoder as a quantizer. However, it is usually necessary to make a distinction between the device that actually quantizes (say, a rounding-off function) and the encoder itself, which may encompass additional building aimed at reducing the transmission rate (see, e.g., [59]).

that within this framework, guaranteeing a bounded average data-rate does not, necessarily, guarantee bounded instantaneous data-rates (see, e.g., [31, 59]). This constitutes a drawback of variable length quantization. Thus, it can be argued that, from a practical perspective, variable length quantization schemes are idealized devices. Nonetheless, this paradigm, which stands at the foundations of information theory, yields interesting results that can be used as benchmarks in practical situations.

The remainder of this chapter is organized as follows: Section 5.2 presents the basic definitions and facts upon which the subsequent sections are based. Section 5.3 studies lower bounds on average data-rates that are valid when general source encoders and decoders are used. These bounds highlight the role of the information available at the encoder and decoder. In Section 5.4 we define the notion of i.i.d. source coding scheme and show that, for this class of source coding schemes, the lower bounds on average data-rates obtained in Section 5.3 can, in turn, be bounded by quantities that are related to the spectral characteristics of the input and output of the source coding scheme in a simple manner. These results provide a precise link between information theory and control theory. Their implications will be explored in Chapter 6. Section 5.5 reviews entropy coded dithered quantizers (see, e.g., [190]), which are i.i.d. source coding schemes. Section 5.6 draws conclusions. (We suggest that the reader review the notation given in the beginning of this thesis, and in Appendix C.)

5.2 Background

As foreshadowed above, this chapter focuses on source coding problems. For brevity, we will refer jointly to the source encoder and the source decoder as the source coding scheme. Following [59], we will consider that source coding schemes consist of three components: a lossy encoder (in short, the *encoder*), a reproduction decoder (in short, the *decoder*) and a lossless encoder-decoder pair (also called *entropy coder-decoder* pair); see Figure 5.2. We next elaborate on each of these components.

The (lossy) *encoder* $\mathcal{E}$ corresponds to the component of the source coding scheme that actually quantizes data. Since we are interested in source coding for control systems, we will focus on causal encoders. That is, the output of the encoder at any time instant k, say $s(k)$, may depend on a subset of the information pattern given by[4]

$$I_{\mathcal{E}}^{\max}(k) \triangleq \{s^{k-1}, v^k\} \tag{5.1}$$

[4] x_i^k is shorthand for $x(i), x(i+1), \cdots, x(k)$, and x^k is shorthand for $x(0), \cdots, x(k)$.

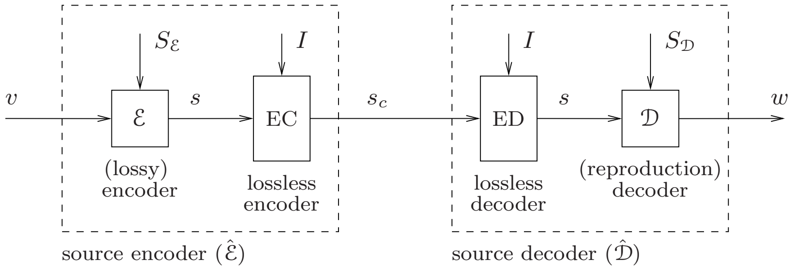

Figure 5.2: The general structure of a source coding scheme.

and on an additional information set, say $S_{\mathcal{E}}(k)$, referred to as side information.[5] Thus,

$$s(k) = \mathcal{E}_k(I_{\mathcal{E}}(k), S_{\mathcal{E}}(k)), \tag{5.2}$$

where $I_{\mathcal{E}}(k) \subseteq I_{\mathcal{E}}^{\max}(k)$, and $\mathcal{E}_k$ is a deterministic mapping that may depend explicitly upon time and that has a *countable range*, i.e., for every $k \in \mathbb{N}_0$, $s(k) \in \mathcal{A}$, with $\mathcal{A}$ a countable set.

The (reproduction) *decoder* $\mathcal{D}$ corresponds to the component of the source coding scheme that uses the symbols generated by the encoder to construct the output of the source decoder w. Again, we will focus on causal decoders and, accordingly, the decoder output at any time instant k may depend on a subset of

$$I_{\mathcal{D}}^{\max}(k) \triangleq \{s^k\} \tag{5.3}$$

and on additional side information contained in the set $S_{\mathcal{D}}(k)$. That is,

$$w(k) = \mathcal{D}_k(I_{\mathcal{D}}(k), S_{\mathcal{D}}(k)), \tag{5.4}$$

where $I_{\mathcal{D}}(k) \subseteq I_{\mathcal{D}}^{\max}(k)$, and $\mathcal{D}_k$ is a deterministic mapping that may depend explicitly upon time and that has (a subset of) the real line as range.

Remark 5.1 (Deterministic encoders and decoders) *In our definitions, the side information contained in $S_{\mathcal{E}}(k)$ and $S_{\mathcal{D}}(k)$ may be random. Hence, our definitions encompass stochastic encoders and decoders. We also note that, if both the encoder and decoder are deterministic, then $S_{\mathcal{E}}(k)$ and $S_{\mathcal{D}}(k)$ do not play any role and, accordingly, do not need to be defined. In such cases, we simplify* (5.2) *and* (5.4) *to*

$$s(k) = \mathcal{E}_k(I_{\mathcal{E}}(k)), \quad w(k) = \mathcal{D}_k(I_{\mathcal{D}}(k)), \tag{5.5}$$

respectively. □□

[5] More will be said about $S_{\mathcal{E}}(k)$ later (see Assumption 5.1).

A direct way to obtain a binary representation of the symbol $s(k)$ generated by the encoder is to define any arbitrary rule to map $s(k)$ into a binary words. The problem with this approach is that it does not take into account that, depending on the statistics of the input v, some encoder output values $s(k)$ are likely to be more frequent than others. (Exploiting this fact is crucial when trying to represent data using the least possible amount of bits.) Moreover, it is very likely that an arbitrarily chosen map between the symbols $s(k)$ and binary words is such that some binary words appear as a prefix to others. This makes the decoding of long sequences of binary words impossible, i.e., the resulting binary sequences may not be *uniquely decodable*. A systematic way to overcome the above problems is provided by loss-less data compression theory (see Chapter 5 in [31]). The basic idea is to assign short binary words to very frequent encoder output values and long binary words to infrequent ones. Moreover, these binary words should be chosen in a way such that no word is a prefix of another (*prefix-free condition*). The previous tasks are performed by so-called lossless encoders or *entropy coders* (EC). The inverse task, i.e., translating binary words back into symbols in $\mathcal{A}$, is performed by a lossless decoder or *entropy decoder* (ED); see Figure 5.2.

In standard source coding theory it is customary to consider the entropy coding of successive (very) long blocks of data (block entropy coding). This reduces the number of bits that, on average, is needed to represent the data. However, this also introduces delays that are potentially harmful in control systems. Accordingly, we will focus on ECs that work on a sample by sample basis, and that can be written as time varying deterministic mappings $\mathcal{H}_k : \mathcal{A} \times \mathcal{I}(k) \to \mathcal{A}_b$. Such ECs use the current output of the encoder $s(k)$ and some additional information $I(k) \in \mathcal{I}(k)$ to generate the binary word $s_c(k) \in \mathcal{A}_b$, where $\mathcal{A}_b$ is a collection of prefix-free binary words (i.e., binary words that satisfy the prefix-free condition mentioned above). Analogously, we will consider EDs that can be written as time varying deterministic mappings $\mathcal{H}_k^{-1} : \mathcal{A}_b \times \mathcal{I}(k) \to \mathcal{A}$. The EC and the ED act as just a bridge between the encoder output alphabet $\mathcal{A}$ and the binary words in $\mathcal{A}_b$. Accordingly, it is natural to impose the following constraints, which stem from the fact that the EC and ED operation is lossless and, therefore, invertible:

$$\mathcal{H}_k^{-1}(\mathcal{H}_k(a, I(k)), I(k)) = a, \quad \mathcal{H}_k(\mathcal{H}_k^{-1}(a_b, I(k)), I(k)) = a_b, \tag{5.6}$$

for every $a \in \mathcal{A}$, $a_b \in \mathcal{A}_b$, $I(k) \in \mathcal{I}(k)$, and $k \in \mathbb{N}_0$. Equation (5.6) also makes explicit the fact that the ECs and EDs we are considering operate without delay. It also explains why we have used the same symbol to denote both the input to $\mathcal{D}$ and the output of $\mathcal{E}$ in Figure 5.2.

At each time instant $k \in \mathbb{N}_0$, the information that, before the reception of $s_c(k)$, is available at the decoder $\mathcal{D}$ is contained in $(I_{\mathcal{D}}(k) \setminus \{s(k)\}) \cup S_{\mathcal{D}}(k)$. This corresponds to the information that is already known at the receiving end and that, hence, does not need to be coded in $s_c(k)$. It thus follows from fundamental information theoretic ideas that, since $\mathcal{A}$ is countable, the *expected length* $R(k)$ of any binary description of $s(k)$ is necessarily lower bounded by (see, e.g., [31, 156])[6]

$$R(k) \geq H(s(k)|\,(I_{\mathcal{D}}(k) \setminus \{s(k)\}) \cup S_{\mathcal{D}}(k)), \tag{5.7}$$

where $H(\cdot|\cdot)$ denotes conditional discrete entropy, and $R(k)$ is measured in bits or nats (depending on the basis used to define entropy).

The gap between the left and right hand sides in (5.7) depends on how efficient the EC is, i.e., on how good the EC is at exploiting $(I_{\mathcal{D}}(k) \setminus \{s(k)\}) \cup S_{\mathcal{D}}(k)$ to remove the redundancy in $s(k)$. It can be shown that if $I(k) = (I_{\mathcal{D}}(k) \setminus \{s(k)\}) \cup S_{\mathcal{D}}(k)$, i.e., if $\mathcal{H}_k$ is allowed to explicitly use the conditioning in (5.7), then it is possible to design $\mathcal{H}_k$ in a way such that (see Chapter 5 in [31] and also [97, 191])

$$H(s(k)|\,(I_{\mathcal{D}}(k) \setminus \{s(k)\}) \cup S_{\mathcal{D}}(k)) \leq R(k) \leq \\ H(s(k)|\,(I_{\mathcal{D}}(k) \setminus \{s(k)\}) \cup S_{\mathcal{D}}(k)) + \ln 2, \tag{5.8}$$

where $R(k)$ is measured in nats. That is, the gap between the conditional entropy of the input to the EC, i.e., $s(k)$, and the expected length of the corresponding binary description $s_c(k)$ can be made so that it is no greater than $\ln 2$ nats (1 bit). We note that making $I(k) = (I_{\mathcal{D}}(k) \setminus \{s(k)\}) \cup S_{\mathcal{D}}(k)$ amounts to knowing $(I_{\mathcal{D}}(k) \setminus \{s(k)\}) \cup S_{\mathcal{D}}(k)$ at the encoder side. Since the encoder generates s, it is clear that $I_{\mathcal{D}}(k) \setminus \{s(k)\}$ can be stored at the encoder side. In a purely theoretical setting one can assume that $S_{\mathcal{D}}(k) \subseteq S_{\mathcal{E}}(k)$ so that knowledge of $S_{\mathcal{D}}(k)$ is available at the encoder side. However, there may exist practical cases where the availability of $S_{\mathcal{D}}(k)$ at the encoder side is not a sensible assumption (see also Remark 5.8 in Section 5.5).

We can now introduce the following definition:

Definition 5.1 (Average data-rate) *Consider the source coding scheme in Figure 5.2, where $\mathcal{E}$, $\mathcal{D}$, EC and ED are as defined above. If $R(k)$ denotes the expected length of the binary word $s_c(k)$, then the* average data-rate *of the source coding scheme is given by*

$$\bar{R} \triangleq \lim_{k\to\infty} \frac{1}{k} \sum_{i=0}^{k-1} R(i) \tag{5.9}$$

[6] We note that our notation in (5.7) is a little abusive: the conditioning must be understood as the random variables contained in the set $(I_{\mathcal{D}}(k) \setminus \{s(k)\}) \cup S_{\mathcal{D}}(k)$. The same logic will be used later where the meaning is clear from the context. This logic has been adopted to avoid the definition of a notation for the elements of a given set.

(provided the limit exists). □□

The quantity $\bar{R}$ is the average rate at which the source encoder generates binary words and, as such, is a measure of the *information flow* at the physical level. We will provide lower bounds on $\bar{R}$ in the next section.

We end this section by clarifying the relationship between the general communication scheme of Figure 5.1, and the source coding scheme described above and shown in Figure 5.2. As mentioned in Section 5.1, the problem of source coding disregards channel coding issues, i.e., assumes that $s_{c,i}$ equals $s_{c,o}$ in Figure 5.1. It thus immediately follows by simple comparison that the signal s_c in Figure 5.2 satisfies

$$s_c(k) = s_{c,i}(k) = s_{c,o}(k), \quad \forall k \in \mathbb{N}_0. \tag{5.10}$$

That is, the ideal link between the output of the EC and the input to the ED in Figure 5.2 is intended to model an idealized channel coding scheme and the corresponding underlying channel. Accordingly, we define the following notion:

Definition 5.2 (Ideal digital channel) *For any source coding scheme as in Figure 5.2, where $\mathcal{E}$, $\mathcal{D}$, EC and ED are as defined above, we define an* ideal digital channel *as a device with input $s_{c,i}$ and output $s_{c,o}$ such that*

$$s_{c,o}(k) = s_{c,i}(k), \quad \forall k \in \mathbb{N}_0, \quad \forall s_{c,i}(k) \in \mathcal{A}_b, \tag{5.11}$$

where $\mathcal{A}_b$ is the collection of all prefix-free binary words that the EC can output. The ideal digital channel average data-rate is defined as the average data-rate of the corresponding source coding scheme (see Definition 5.1). □□

An ideal digital channel is an abstraction and, as such, does not correspond to any physical underlying channel that one could envisage. It corresponds to the idealized link that rate distortion theory assumes to exist between the output of the lossless encoder (EC) and the input to the lossless decoder (ED) (recall our discussion on Shannon's separation theorem in Section 5.1). In the sequel, we will make reference to ideal digital channels. In such cases, it will be important to keep in mind that, by construction, such a channel is defined in terms of a source coding scheme and that the average rate at which data is sent through the (ideal digital) channel is defined as the average data-rate of the corresponding source coding scheme.

5.3 Lower Bounds on Average Data-rates

A key question that arises when considering any source coding scheme is how to characterize lower bounds on the achievable average data-rates. In this section, we answer this question when the source coding scheme belongs to a broad class. (Upper bounds are explored in Section 5.5 for a specific source coding scheme.)

We start by introducing the following definition:

Definition 5.3 (Memoryless and full-memory source coding schemes) *Consider the source coding scheme in Figure 5.2, where $\mathcal{E}$, $\mathcal{D}$, EC and ED are as defined previously.*[7]

1. *The source coding scheme is said to be* memoryless *if and only if* $I_{\mathcal{D}}(k) = \{s(k)\}$ *and* $I_{\mathcal{E}}(k) = \{v(k)\}$.
2. *The source coding scheme is said to be* full-memory *if and only if* $I_{\mathcal{D}}(k) = I_{\mathcal{D}}^{\max}(k)$ *and* $I_{\mathcal{E}}(k) = I_{\mathcal{E}}^{\max}(k)$.

□□

We note that (5.7) reduces to

$$R(k) \geq H(s(k)|S_{\mathcal{D}}(k)) \tag{5.12}$$

in the case of source coding schemes where $I_{\mathcal{D}}(k) = \{s(k)\}$ and to

$$R(k) \geq H(s(k)|s^{k-1}, S_{\mathcal{D}}(k)) \tag{5.13}$$

when the source coding scheme is such that $I_{\mathcal{D}}(k) = \{s^k\}$.

In this chapter we will consider source coding schemes that satisfy the following assumptions:

Assumption 5.1 (Source coding scheme) *The encoder $\mathcal{E}$ and decoder $\mathcal{D}$ in the source coding scheme of Figure 5.2 are such that:*

(a) $v(k) \in I_{\mathcal{E}}(k)$, $s(k) \in I_{\mathcal{D}}(k)$.

(b) If $I_{\mathcal{D}}(k) = \{s(k)\}$, *then* $v(k)$ *and* $S_{\mathcal{D}}(k)$ *are independent for every* $k \in \mathbb{N}_0$.

(c) If $I_{\mathcal{D}}(k) = I_{\mathcal{D}}^{\max}(k)$, *then:*

  (c.i) The sequence of mappings $\{\mathcal{D}_i\}_{i\in\{0,\cdots,k\}}$ *in (5.4) is such that,* $\forall k \in \mathbb{N}_0$, *one can recover* s^k *from* w^k *and, upon knowledge of* $S_{\mathcal{D}}(k)$, *one can also derive* w^k *from* s^k.

[7] In these definitions, the side information in $S_{\mathcal{E}}(k)$ and $S_{\mathcal{D}}(k)$ does not play any role.

(c.ii) $\{v_\ell^k\} \setminus S_{\mathcal{D}}(k) = \{v_\ell^k\}$, $\forall k, \ell \in \mathbb{N}_0$, $\ell \leq k$.

(c.iii) *Define* $\ell \triangleq \infty$, *if* $v(i) \in I_{\mathcal{E}}(k)$ *for all* $i \leq k$ *and every* $k \in \mathbb{N}_0$. *Otherwise, define* ℓ *as follows:*

$$\ell \triangleq \max\{i \in \mathbb{N}_0 : v(j) \in I_{\mathcal{E}}(k),\ \forall j \in \{k-i, k-i+1, \cdots, k\} \text{ and } \forall k \geq i\ \}.$$

Then $S_{\mathcal{D}}(k)$ *and* $v^k_{\max\{0,k-\ell\}}$ *are conditionally independent given* w^{k-1} *for every* $k \in \mathbb{N}_0$, *i.e.,*[8]

$$v^k_{\max\{0,k-\ell\}} \leftrightarrow w^{k-1} \leftrightarrow S_{\mathcal{D}}(k). \tag{5.14}$$

□□

Assumption 5.1(a) is by no means restrictive. Indeed, if $\mathcal{E}$ does not have access to $v(k)$, then one can always redefine v as the delayed version of the signal to be encoded. On the other hand, since the link between the EC output and the ED input is assumed to be transparent, it is sensible to assume that $s(k)$ is available instantly at the decoder. If Assumption 5.1(b) is not satisfied, then there must be some "leakage" of information from the encoder side to decoder side. Clearly, this would not make sense in our setting where communication between encoder and decoder is made only through s_c. If the first part of Assumption 5.1($c.i$) is not satisfied, then $w(k)$ is not uniquely determined by the current and past received symbols (and may not be able to be determined at all). If this were the case, then it would be possible to reduce the rate at which symbols are transmitted by exploiting this redundancy. Since we are interested in source coding schemes that achieve minimal data-rates, it follows that it is sensible to restrict our attention to the cases where w^k is uniquely determined by s^k. The second part of Assumption 5.1($c.i$) states that the side information at any time instant k is enough to be able recover w^k from s^k. Since nothing impedes the definition of $S_{\mathcal{D}}(k)$ in a way such that $S_{\mathcal{D}}(k) \subseteq S_{\mathcal{D}}(k+1)$, this assumption does not limit the class of considered source coding schemes. On the other hand, Assumption 5.1($c.ii$) is reasonable since knowledge of v_ℓ^k at the decoding end would imply that quantization is actually not necessary, and the source coding problem would admit a trivial solution. Finally, Assumption 5.1($c.iii$) is a rather technical requirement that is satisfied by the source coding schemes considered in this thesis. This assumption says that, once the past values of the output of the decoder (i.e., w^{k-1}) are known, knowledge of the past and present

[8] $x \leftrightarrow y \leftrightarrow z$ means that x, y and z form a Markov chain (in that order); see Appendix C. We note that the notation in (5.14) makes use of the convention mentioned in Footnote 5: the elements of $S_{\mathcal{D}}(k)$ (and not the set itself) are independent of $v^k_{\max\{0,k-\ell\}}$ when conditioned upon w^{k-1}.

values of the input to the encoder (i.e., $v^k_{\max\{0,k-\ell\}}$) does not help in order to describe the side information at the decoder (i.e., $S_{\mathcal{D}}(k)$). Heuristically speaking, this is reasonable: w is the signal that contains the immediate effects of $S_{\mathcal{D}}$ and, even if $S_{\mathcal{D}}$ affects v via strictly causal feedback (as it would in a networked control system), then $v^k_{\max\{0,k-\ell\}}$ is unlikely to contain better information on $S_{\mathcal{D}}(k)$ than w^{k-1}.

Remark 5.2 *If both the encoder $\mathcal{E}$ and the decoder $\mathcal{D}$ are deterministic (recall Remark 5.1), then Assumption 5.1 reduces to Parts* (a) *and* (c.i) *(with an obvious modification: in the present case, s^k should be recoverable from w^k and* vice versa*).* □□

The following definition will prove useful:

Definition 5.4 (Average mutual information) *Consider two random processes v and w. We define (if the defining limits exist) the* average scalar mutual information *between v and w as*[9]

$$I_\infty(v;w) \triangleq \lim_{k\to\infty} \frac{1}{k} \sum_{i=0}^{k-1} I(v(i);w(i)), \tag{5.15}$$

and the average directed mutual information between v and w *as*[10] *(see also [40, 105, 108])*

$$I_\infty(v \to w) \triangleq \lim_{k\to\infty} \frac{1}{k} \sum_{i=0}^{k-1} I(w(i);v^i|w^{i-1}). \tag{5.16}$$

We also define the auxiliary quantity

$$I_\infty(v \xrightarrow{\ell} w) \triangleq \lim_{k\to\infty} \frac{1}{k} \sum_{i=0}^{k-1} I(w(i);v^i_{\max\{0,i-\ell\}}|w^{i-1}), \tag{5.17}$$

where $\ell \in \mathbb{N}_0 \cup \{\infty\}$ is a fixed number. (For $\ell = \infty$, this definition reduces to (5.16)*)* □□

Using Definition 5.4 it is possible to state the following result that bounds average data-rates from below:

Theorem 5.1 *Consider the source coding scheme in Figure 5.2, where $\mathcal{E}$, $\mathcal{D}$, EC and ED are as defined above, and suppose that Assumption 5.1 holds. Then:*

1. *If $I_{\mathcal{D}}(k) = \{s(k)\}$, then $\bar{R} \geq I_\infty(v;w)$.*
2. *If $I_{\mathcal{D}}(k) = I^{\max}_{\mathcal{D}}(k)$, then $\bar{R} \geq I_\infty(v \xrightarrow{\ell} w)$, where ℓ is as in Assumption 5.1. In particular, if the source coding scheme is a full-memory one, then $\bar{R} \geq I_\infty(v \to w)$.*

---

[9] $I(\cdot;\cdot)$ refers to mutual information and $I(\cdot;\cdot|\cdot)$ to conditional mutual information (see Appendix C).

[10] This quantity is sometimes also called directed mutual information rate.

Proof: The proof uses the same ideas used in the proof of the so-called *converse* to the rate distortion theorem (see Section 10.4 in [31]). We note that we make use of the conventions in Footnote 5 extensively.

1. From (5.12) we have that (recall the definitions in Section 5.2):

$$\begin{aligned} R(i) &\geq H(s(i)|S_{\mathcal{D}}(i)) \\ &\overset{(b)}{=} H(s(i)|S_{\mathcal{D}}(i)) - H(s(i)|I_{\mathcal{E}}(i), S_{\mathcal{E}}(i), S_{\mathcal{D}}(i)) \\ &\overset{(c)}{=} I(s(i); \{I_{\mathcal{E}}(i), S_{\mathcal{E}}(i)\} \setminus S_{\mathcal{D}}(i)|S_{\mathcal{D}}(i)) \end{aligned} \tag{5.18}$$

where (b) follows from the fact that $s(i)$ is deterministic function of $I_{\mathcal{E}}(i)$ and $S_{\mathcal{E}}(i)$ (hence the second term is zero), and (c) follows from Remark C.2 in Appendix C. Since $w(i)$ is a deterministic function of $s(i)$ and $S_{\mathcal{D}}(i)$, we obtain the following Markov chain:

$$\{I_{\mathcal{E}}(i), S_{\mathcal{E}}(i)\} \setminus S_{\mathcal{D}}(i)\big|_{S_{\mathcal{D}}(i)} \leftrightarrow s(i)\big|_{S_{\mathcal{D}}(i)} \leftrightarrow w(i)\big|_{S_{\mathcal{D}}(i)}. \tag{5.19}$$

Thus, the data processing inequality (see Theorem C.1 and Remark C.3 in Appendix C) and (5.18) yield

$$R(i) \geq I(w(i); \{I_{\mathcal{E}}(i), S_{\mathcal{E}}(i)\} \setminus S_{\mathcal{D}}(i)|S_{\mathcal{D}}(i)) \tag{5.20}$$

with equality if and only if (5.19) still holds when the last two terms are interchanged. On the other hand, it follows from Assumption 5.1(*a*-*b*) and the definition of $I_{\mathcal{E}}(k)$ that $v(i) \in \{I_{\mathcal{E}}(i), S_{\mathcal{E}}(i)\} \setminus S_{\mathcal{D}}(i)$. Hence,

$$w(i)\big|_{S_{\mathcal{D}}(i)} \leftrightarrow \{I_{\mathcal{E}}(i), S_{\mathcal{E}}(i)\} \setminus S_{\mathcal{D}}(i)\big|_{S_{\mathcal{D}}(i)} \leftrightarrow v(i)\big|_{S_{\mathcal{D}}(i)} \tag{5.21}$$

and (5.20) yields

$$\begin{aligned} R(i) &\overset{(a)}{\geq} I(w(i); v(i)|S_{\mathcal{D}}(i)) \\ &\overset{(b)}{=} I(S_{\mathcal{D}}(i), w(i); v(i)) - I(v(i); S_{\mathcal{D}}(i)) \\ &\overset{(c)}{=} I(S_{\mathcal{D}}(i), w(i); v(i)) \end{aligned} \tag{5.22}$$

where (b) and (c) follow from Facts C.2 and C.3, and the fact that, by Assumption 5.1(b), $v(i)$ and $S_{\mathcal{D}}(i)$ are independent. We note that equality in (a) holds if and only if the Markov chain (5.21) still holds when the last two terms are interchanged. Since, clearly,

$$v(i) \leftrightarrow S_{\mathcal{D}}(i), w(i) \leftrightarrow w(i), \tag{5.23}$$

we conclude from (5.22) that

$$R(i) \geq I(v(i); w(i)) \tag{5.24}$$

with equality if and only if the Markov chain (5.23) still holds when the last two terms are interchanged. The result follows using (5.24) in (5.9) and the definition of $I_\infty(v; w)$.

2. Proceeding exactly as in the previous part, we have from (5.13) that

$$\begin{aligned} R(i) &\geq H(s(i)|s^{i-1}, S_{\mathcal{D}}(i)) \\ &= H(s(i)|s^{i-1}, S_{\mathcal{D}}(i)) - H(s(i)|I_{\mathcal{E}}(i), S_{\mathcal{E}}(i), s^{i-1}, S_{\mathcal{D}}(i)) \\ &= I\left(s(i); \{I_{\mathcal{E}}(i), S_{\mathcal{E}}(i)\} \setminus \left\{s^{i-1}, S_{\mathcal{D}}(i)\right\} |s^{i-1}, S_{\mathcal{D}}(i)\right). \end{aligned} \tag{5.25}$$

As before, $w(i)$ is a deterministic function of s^i and $S_{\mathcal{D}}(i)$. Thus,

$$\{I_{\mathcal{E}}(i), S_{\mathcal{E}}(i)\} \setminus \left\{s^{i-1}, S_{\mathcal{D}}(i)\right\} \big|_{s^{i-1}, S_{\mathcal{D}}(i)} \leftrightarrow s(i)\big|_{s^{i-1}, S_{\mathcal{D}}(i)} \leftrightarrow w(i)\big|_{s^{i-1}, S_{\mathcal{D}}(i)} \tag{5.26}$$

and (5.25) yields

$$R(i) \geq I\left(w(i); \{I_{\mathcal{E}}(i), S_{\mathcal{E}}(i)\} \setminus \left\{s^{i-1}, S_{\mathcal{D}}(i)\right\} |s^{i-1}, S_{\mathcal{D}}(i)\right) \tag{5.27}$$

with equality if and only if the Markov chain (5.26) still holds when the last two terms are interchanged. Assumption 5.1(*c.ii*) and the definition of $I_{\mathcal{E}}(i)$ imply that $\{v^i_{\max\{0,i-\ell\}}\} \subseteq \{I_{\mathcal{E}}(i), S_{\mathcal{E}}(i)\} \setminus \left\{s^{i-1}, S_{\mathcal{D}}(i)\right\}$. Therefore,

$$\begin{aligned} &w(i)\big|_{s^{i-1}, S_{\mathcal{D}}(i)} \leftrightarrow \\ &\qquad \{I_{\mathcal{E}}(i), S_{\mathcal{E}}(i)\} \setminus \left\{s^{i-1}, S_{\mathcal{D}}(i)\right\} \big|_{s^{i-1}, S_{\mathcal{D}}(i)} \leftrightarrow v^i_{\max\{0,i-\ell\}}\big|_{s^{i-1}, S_{\mathcal{D}}(i)} \end{aligned} \tag{5.28}$$

and (5.27) yields

$$R(i) \geq I\left(w(i); v^i_{\max\{0,i-\ell\}}|s^{i-1}, S_{\mathcal{D}}(i)\right) \tag{5.29}$$

with equality if and only if the Markov chain (5.28) still holds when the last two terms are interchanged. Since Assumption 5.1(*c.i*) holds, knowledge of $S_{\mathcal{D}}(i)$ allows one to deterministically obtain s^{i-1} from w^{i-1} and *vice-versa*. Thus, (5.29) yields

$$\begin{aligned} R(i) &\geq I\left(w(i); v^i_{\max\{0,i-\ell\}}|w^{i-1}, S_{\mathcal{D}}(i)\right) \\ &= I(w^i, S_{\mathcal{D}}(i); v^i_{\max\{0,i-\ell\}}) - I(w^{i-1}, S_{\mathcal{D}}(i); v^i_{\max\{0,i-\ell\}}), \end{aligned} \tag{5.30}$$

where the last equality follows from Fact C.2. Since, clearly,

$$v^i_{\max\{0,i-\ell\}} \leftrightarrow w^i, S_{\mathcal{D}}(i) \leftrightarrow w^i, \tag{5.31}$$

we have from (5.30) that

$$\begin{aligned} R(i) &\overset{(a)}{\geq} I(w^i; v^i_{\max\{0,i-\ell\}}) - I(w^{i-1}, S_{\mathcal{D}}(i); v^i_{\max\{0,i-\ell\}}) \\ &\overset{(b)}{=} I(w^{i-1}; v^i_{\max\{0,i-\ell\}}) + I(w(i); v^i_{\max\{0,i-\ell\}}|w^{i-1}) - \\ &\qquad\qquad I(w^{i-1}; v^i_{\max\{0,i-\ell\}}) - I(S_{\mathcal{D}}(i); v^i_{\max\{0,i-\ell\}}|w^{i-1}) \\ &\overset{(c)}{=} I(w(i); v^i_{\max\{0,i-\ell\}}|w^{i-1}), \end{aligned} \tag{5.32}$$

where (b) follows from Fact C.2, and (c) follows from Assumption 5.1($c.iii$) and Fact C.3. Equality in (a) holds if and only if the Markov chain (5.31) still holds when the last two terms are interchanged. Our claim follows using (5.32) in (5.9) and the definition of $I_\infty(v \to w)$.

□□□

The relevance of Theorem 5.1 lies in the fact that it relates a physical quantity, namely average data-rate, to information theoretic quantities such as average directed mutual information and average scalar mutual information. It is important to note that *different* bounds on the average data-rate arise depending on the *information available* at the encoder and decoder. Another interesting consequence of Theorem 5.1 is that, no matter what the information available at the encoder is, the information available at the decoder is central in determining bounds on the average data-rate. In particular, if the decoder is such that $I_{\mathcal{D}}(k) = \{s(k)\}$, then the bound on average data-rate is the average scalar mutual information for any information available at the encoder $I_{\mathcal{E}}(k)$. This implies that, if $I_{\mathcal{D}}(k) = \{s(k)\}$, then one can assume without loss of generality that $I_{\mathcal{E}}(k) = \{v(k)\}$.[11] On the other hand, if $I_{\mathcal{D}}(k) = \{s^k\}$, then different average-directed-mutual-information-like bounds on the average data-rate arise depending on the information available at the decoder. The most favorable case is, of course, the case where $I_{\mathcal{E}}(k) = I_{\mathcal{E}}^{\max}(k)$.

Remark 5.3 (Partial memory) *In Theorem 5.1 we have only focused on the case where the decoder has no memory or has complete knowledge of all past received symbols. From a practical point of view, it would be interesting to examine the effect of partial knowledge of the past*

[11] Since Assumption 5.1 holds, this is always possible.

symbols. For example, one could analyze cases where $I_{\mathfrak{D}}(k) = \{s^k_{\max\{0,k-i\}}\}$ for some $i \in \mathbb{N}_0$. In this thesis we will not consider such cases. Instead, we will only focus on the two extreme cases considered in Definition 5.3. This is motivated by the fact that the full-memory case will yield the best possible results (within the class of source coding schemes under consideration) and, on the other hand, the memoryless case will allow one to study worst cases. Interestingly, we will show that, in some cases of interest, no loss occurs when considering memoryless source coding schemes. □□

Remark 5.4 (Relationship to previous work) *The concept of directed mutual information was first defined in [108]. Reference [108] proved that directed mutual information is the appropriate lower bound on average data-rate when causal processing is sought (as is the case of feedback systems) and no additional constraints are imposed on the coding schemes. Therefore, the results of Theorem 5.1 that are related to full-memory source coding schemes are not surprising. However, our results also show that different bounds may arise if constraints are imposed on the source coding scheme memory. (See also Remark 5.7 below.)* □□

5.4 A Class of Coding Schemes

In this thesis we will limit our attention to a specific class of source coding schemes. The main reason for this is simplicity. We will not address arbitrarily complex source coding schemes, but, instead, focus on source coding schemes that are simple enough to allow one to design control systems subject to average data-rate limits using standard tools. We will adopt the following definition:

Definition 5.5 (I.i.d. source coding scheme) *We will say that the source coding scheme in Figure 5.2 is i.i.d. if and only if Assumption 5.1 holds, and the quantization noise (or coding noise) q, defined via*

$$q \triangleq w - v, \tag{5.33}$$

is an independent second order and zero-mean i.i.d. sequence. □□

Remark 5.5 *We note that the "i.i.d. -ness" of a source coding scheme is related to the statistical properties of the coding noise it introduces. It is not related, in any way, to the availability or otherwise of memory at the encoder or decoder. An i.i.d. source coding scheme can be memoryless, have full-memory, or have partial memory.* □□

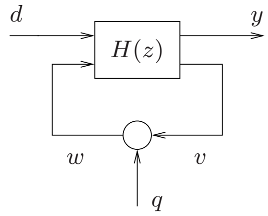

Figure 5.3: General feedback system closed over an i.i.d. source coding scheme.

The reader may wonder whether Definition 5.5 encompasses any practical, or at least conceptually plausible, source coding scheme. The answer is affirmative and, indeed, a specific source coding scheme will be studied in Section 5.5.

Given the fact that our interest in this thesis lies in control systems, we will study i.i.d. source coding schemes embedded in feedback loops, as depicted in Figure 5.3. In that figure, d is a (possibly vector) random sequence, $H(z)$ is the transfer function of an LTI system, and q is the quantization noise in an i.i.d. source coding scheme. We will assume that the following holds:

Assumption 5.2 (Exogenous signal, initial state and $H(z)$)

(a) The transfer function $H(z)$ belongs to $\mathcal{R}_p$, has an underlying stabilizable and detectable realization, is such that the transfer function from w to v belongs to $\mathcal{R}_{sp}^{1\times 1}$, and is such that the closed loop in Figure 5.3 is internally stable and well-posed.

(b) The signal d is a second order wss process.

(c) The initial state of $H(z)$, say $x_H(0)$, is a second order random variable.

□□

We start by noting that there exist cases where the memory in an i.i.d. source coding scheme plays no role in the achievable average data-rate:

Lemma 5.1 (Sometimes, memory does not help) *Consider the setup of Figure 5.3, where q is the quantization noise in an i.i.d. source coding scheme, and suppose that Assumption 5.2 holds. Then, $I_\infty(v \to w) \leq I_\infty(v, w)$ with equality if and only if w is i.i.d..*

Proof: Note that, by definition, q is an i.i.d. sequence which is independent of d and of the initial state $x_H(0)$. Since $H(z)$ has at least one step delay from w to v, we also have that

$q(k)$ is independent of v^k and w^{k-1}. Therefore,

$$\begin{aligned} I(w(k); v^k | w^{k-1}) &\overset{(a)}{=} h(w(k)|w^{k-1}) - h(q(k) + v(k)|v^k, w^{k-1}) \\ &\overset{(b)}{=} h(w(k)|w^{k-1}) - h(q(k) + v(k)|v(k)) \\ &\overset{(c)}{\leq} h(w(k)) - h(w(k)|v(k)) \\ &\overset{(d)}{=} I(w(k); v(k)), \end{aligned} \tag{5.34}$$

where (a) follows from Fact C.2 and the definition of q, (b) follows from the fact that $q(k)$ is independent of v^k and w^{k-1} (indeed, $h(q(k)+v(k)|v(k)) = h(q(k)))$, (c) follows from Fact C.1 and the definition of q, and (d) follows from Fact C.2. From Fact C.1 we also know that equality in (c) is achieved if and only if w is i.i.d.. The result follows upon using (5.34) and Definition 5.4. □□□

Lemma 5.1 suggests that, as expected, the use of full-memory i.i.d. source coding schemes allows one to achieve average data-rates that are smaller than (or at least equal to) the average data-rates achievable with a memoryless i.i.d. source coding scheme. However, the implementation of source coding schemes with memory is, in general, very complicated (see Remark 5.10 in Section 5.5). In this respect, Lemma 5.1 is encouraging since it states that, if w is (somehow) made i.i.d., then the use of memoryless i.i.d. source coding schemes will, in principle,[12] not limit, the achievable average data-rates. In other words, if w is i.i.d., then using an i.i.d. source coding scheme with memory does not provide any advantage over a (much simpler) memoryless one.

A key feature of i.i.d. source coding schemes is that the average (scalar or directed) mutual information across them can be bounded from above by the average mutual information that would arise if all random variables were replaced by their *Gaussian counterparts*. To be precise, we introduce the following definition:

Definition 5.6 (The Gaussian case) *Consider the feedback system of Figure 5.3, where q is the quantization noise in an i.i.d. source coding scheme, and suppose that Assumption 5.2 holds. We will use the wording* Gaussian case *to refer to the case where d, q and $x_H(0)$ are replaced by jointly Gaussian random variables (or sequences of random variables) having the same first and second order moments (and cross-moments) and which maintain the same statistical dependence relationships as in the original situation. We will add a "G" subscript to the symbols that refer to signals and variables in the Gaussian case. (For example, v_G and w_G denote the processes that would arise as input and output of the source coding scheme in the Gaussian case; v_G and w_G are the* Gaussian counterparts *of v and w.)* □□

[12] Recall that $I_\infty(v \to w)$ and $I_\infty(v, w)$ are bounds on $\bar{R}$, and are thus not necessarily equal to $\bar{R}$.

Lemma 5.2 (Using the Gaussian case to bound mutual information) *Consider the setup of Figure 5.3, where q is the quantization noise in an i.i.d. source coding scheme, and suppose that Assumption 5.2 holds. Then (we use the notation introduced in Definition 5.6):*

1. $I_\infty(v;w) \leq I_\infty(v_G;w_G) + D(q(k)||q_G(k))$, *where* $D(q(k)||q_G(k))$ *is the divergence of the distribution of* $q(k)$ *from its Gaussian counterpart (see Appendix C).*

2. $I_\infty(v \rightarrow w) \leq I_\infty(v_G \rightarrow w_G) + D(q(k)||q_G(k))$.

3. *If, in addition,* d *and* $x_H(0)$ *are jointly Gaussian, then*

   (a) $I_\infty(v_G;w_G) \leq I_\infty(v;w)$.

   (b) $I_\infty(v_G \rightarrow w_G) \leq I_\infty(v \rightarrow w)$.

Proof: We start by recalling that $q(k)$ is independent of v^k and w^{k-1} (see proof of Lemma 5.1). Analogously, we have, by definition of Gaussian counterparts given above, that $q_G(k)$ is independent of v_G^k and w_G^{k-1}.

1. This part follows from Part 1 of Lemma C.1 (see Appendix C), (5.15) and the fact that q and q_G are i.i.d.. (Actually, the proof follows immediately from (5.37) [see below] and Fact C.4.)

2. We note that the following holds:

$$\begin{aligned} I(w(k);v(k)|w^{k-1}) &\stackrel{(a)}{=} h(w(k)|w^{k-1}) - h(q(k)+v(k)|w^{k-1},v(k)) \\ &\stackrel{(b)}{=} h(w(k)|w^{k-1}) - h(q(k)+v(k)|w^{k-1},v^k) \\ &\stackrel{(c)}{=} I(w(k);v^k|w^{k-1}), \end{aligned} \tag{5.35}$$

where (a) and (c) follow from Fact C.2 and the definition of $q(k)$, and (b) follows from Fact C.1 and the independence of v^k and $q(k)$. The same applies if one replaces all variables by their Gaussian counterparts. Therefore,

$$
\begin{aligned}
&I(w(k); v^k|w^{k-1}) - I(w_G(k); v_G^k|w_G^{k-1}) \\
&\quad = I(w(k); v(k)|w^{k-1}) - I(w_G(k); v_G(k)|w_G^{k-1}) \\
&\quad \overset{(a)}{=} h(v(k)+q(k)|w^{k-1}) - h(v(k)+q(k)|v(k), w^{k-1}) - \\
&\qquad\qquad h(v_G(k)+q_G(k)|w_G^{k-1}) + h(v_G(k)+q_G(k)|v_G(k), w_G^{k-1}) \\
&\quad \overset{(b)}{=} h(q_G(k)) - h(q(k)) + h(v(k)+q(k)|w^{k-1}) - h(v_G(k)+q_G(k)|w_G^{k-1}) \\
&\quad \overset{(c)}{=} D(q(k)||q_G(k)) - D(w(k)|w^{k-1}||w_G(k)|w_G^{k-1}) \\
&\quad \overset{(d)}{\leq} D(q(k)||q_G(k)),
\end{aligned}
\tag{5.36}
$$

where (a) follows from Fact C.2 and the definition of $q(k)$, (b) follows from Fact C.1 and the independence of $q(k)$ (resp. $q_G(k)$) from v^k and w^{k-1} (resp. v_G^k and w_G^{k-1}), (c) follows from Fact C.4 and Remark C.4, and (d) follows from Remark C.4. The result now follows from (5.36), (5.16) and the fact that q and q_G are i.i.d..

3. We proceed by parts.

   (a) We first note that

$$
\begin{aligned}
I(v(k); w(k)) - I(v_G(k); w_G(k)) &\overset{(a)}{=} h(v(k)+q(k)) - h(v(k)+q(k)|v(k)) - \\
&\qquad h(v_G(k)+q_G(k)) + h(v_G(k)+q_G(k)|v_G(k)) \\
&\overset{(b)}{=} h(q_G(k)) - h(q(k)) - \\
&\qquad h(v_G(k)+q_G(k)) + h(v(k)+q(k)) \\
&\overset{(c)}{=} D(q(k)||q_G(k)) - D(v(k)+q(k)||v_G(k)+q_G(k)) \\
&\triangleq M
\end{aligned}
\tag{5.37}
$$

   where (a) follows from Fact C.2 and the definition of $w(k)$, (b) follows from the independence of $v(k)$ (resp. $v_G(k)$) from $q(k)$ (resp. $q_G(k)$), and (c) follows from Fact C.4 and the definition of v_G and q_G. If $D(v(k)||v_G(k)) \leq D(q(k)||q_G(k))$, then Part 3 of Lemma C.1 would imply $M \geq 0$ and the result would follow from (5.15) and (5.37). We next show that $D(v(k)||v_G(k)) \leq D(q(k)||q_G(k))$ actually holds.

   From Figure 5.3 it is immediate to see that, if $(A, \begin{bmatrix} B_d & B_w \end{bmatrix}, C_v, \begin{bmatrix} D_d & 0 \end{bmatrix})$ is a realization of the open loop transfer function from $\begin{bmatrix} d & w \end{bmatrix}^T$ to v (recall that the transfer function from w to v is strictly proper), then

$$
v(k) = v_{0,d}(k) + v_q(k),
\tag{5.38}
$$

where

$$v_{0,d}(k) \triangleq C_v(A + B_w C_v)^k x_H(0) + D_d d(k) + \sum_{i=1}^{k-1} C_v(A + B_w C_v)^{k-1-i} B_d d(i), \quad (5.39)$$

$$v_q(k) \triangleq \sum_{i=0}^{k-1} M_i^k q(i), \quad M_i^k \triangleq C_v(A + B_w C_v)^{k-1-i} B_w. \quad (5.40)$$

We note that, since the closed loop is assumed to be stable, then M_i^k is guaranteed to be finite for every $i, k \in \mathbb{N}_0$. Since both q and q_G are i.i.d., Fact C.4 implies

$$D(M_0^k q(0)||M_0^k q_G(0)) = D(q(0)||q_G(0)) = D(q(1)||q_G(1)) = D(M_1^k q(1)||M_1^k q_G(1)). \quad (5.41)$$

Therefore, Part 3 of Lemma C.1 and the fact that q is i.i.d. imply

$$D(M_0^k q(0) + M_1^k q(1)||M_0^k q_G(0) + M_1^k q_G(1)) \leq D(q(1)||q_G(1)) = D(q(2)||q_G(2)) = D(M_2^k q(2)||M_2^k q_G(2)). \quad (5.42)$$

Using the previous argument in a recursive fashion, it follows that

$$D(v_q(k)||v_{q_G}(k)) \leq D(q(k)||q_G(k)), \quad (5.43)$$

where

$$v_{q_G}(k) \triangleq \sum_{i=0}^{k-1} M_i^k q_G(i). \quad (5.44)$$

Since $x_H(0)$ and d are jointly Gaussian, so is $v_{0,d}$. Also, since q is independent of d and $x_H(0)$, then $v_q(k)$ is independent of $v_{0,d}(k)$. Therefore, Part 3 of Lemma C.1 and (5.43) imply

$$D(v(k)||v_G(k)) \leq D(v_q(k)||v_{q_G}(k)) \leq D(q(k)||q_G(k)), \quad (5.45)$$

as claimed.

(b) In this case, we take an indirect approach. It is straightforward to rewrite Figure 5.3 as shown in Figure 5.4, where $T_{dw}(z), T_{qw}(z) \in \mathcal{RH}_\infty$ denote the transfer functions from d to w and from q to w, respectively. Since $H(z)$ has at least one step delay from w to v, it follows that $1 - T_{qw}(\infty) = 0$, i.e., $1 - T_{qw}(z) \in \mathcal{RH}_2$. Accordingly, Lemma C.2 in Appendix C guarantees that the following two relations hold:

$$\bar{I}_\infty(\tilde{d}; w) = I_\infty(v \to w) + K,$$
$$\bar{I}_\infty(\tilde{d}_G; w_G) = I_\infty(v_G \to w_G) + K, \quad (5.46)$$

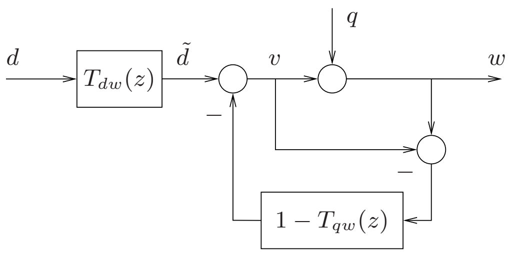

Figure 5.4: Equivalent rewriting of Figure 5.3.

where $\bar{I}_\infty(\cdot;\cdot)$ is defined as in Definition C.6, and K is a constant (identified in Lemma C.2). On the other hand, since d and $x_H(0)$ are Gaussian, so is $\tilde{d}$ and, hence, $\tilde{d}_G = \tilde{d}$. Thus, Part 2 of Lemma C.1 implies that $\bar{I}(\tilde{d}_G; w_G) = \bar{I}(\tilde{d}; w_G) \leq \bar{I}(\tilde{d}; w)$ and, as a consequence, (5.46) implies

$$I_\infty(v_G \to w_G) \leq I_\infty(v \to w). \tag{5.47}$$

□□□

As anticipated, Lemma 5.2 states that the Gaussian case provides upper bounds for the average (scalar or directed) mutual information for general i.i.d. source coding schemes.[13] The importance of this result lies in the fact that the characterization of average mutual information in the Gaussian case is straightforward and, hence, it is possible to easily obtain explicit bounds on average mutual information for general i.i.d. source coding schemes. This is done in Theorem 5.2 below. Furthermore, Lemma 5.2 also states that if the external signal d and the initial state of the linear system $H(z)$ are jointly Gaussian, then the best i.i.d. source coding scheme (in the sense of minimizing average mutual information) is such that the quantization noise is Gaussian. This observation, together with Theorem 5.1, yields an absolute lower bound on the average data rate for i.i.d. source coding schemes embedded in an otherwise LTI feedback system, when the external signals and the initial states are jointly Gaussian.

Theorem 5.2 (Bounds for mutual information in i.i.d. source coding schemes) *Consider the setup of Figure 5.3, where q is the quantization noise in an i.i.d. source coding scheme, and suppose that Assumption 5.2 holds. Then:*

[13] It is interesting to note that there exists a plethora of situations where Gaussianity assumptions lead to upper bounds on quantities that involve random variables or processes. (As another example, see Remark C.5 in Appendix C.)

1. *The average scalar mutual information between v and w is bounded from above by*

$$I_\infty(v;w) \leq \frac{1}{2}\ln(1+\gamma) + D(q(k)||q_G(k)), \quad \gamma \triangleq \frac{\sigma_v^2}{\sigma_q^2}, \tag{5.48}$$

where σ_v^2 is the stationary variance of v and, accordingly, γ is the stationary signal-to-noise ratio *of the source coding scheme.*

2. *The average directed mutual information between v and w is bounded from above by*

$$I_\infty(v \to w) \leq \frac{1}{4\pi}\int_{-\pi}^{\pi} \ln \frac{S_w(e^{j\omega})}{\sigma_q^2} d\omega + D(q(k)||q_G(k)) \overset{(a)}{\leq} \frac{1}{2}\ln(1+\gamma) + D(q(k)||q_G(k)), \tag{5.49}$$

where $S_w(e^{j\omega})$ is the stationary PSD of w, and equality in (a) holds if and only if $S_w(e^{j\omega})/\sigma_q^2$ is constant.

Proof:

1. Recall Definition 5.6. Given Lemma 5.2 we just need to prove that $I_\infty(v_G;w_G) = \frac{1}{2}\ln(1+\gamma)$. We start by noting that the following holds for every $i \in \mathbb{N}_0$:

$$\begin{aligned} I(v_G(i);w_G(i)) &\overset{(a)}{=} h(w_G(i)) - h(q_G(i)+v_G(i)|v_G(i)) \\ &\overset{(b)}{=} h(w_G(i)) - h(q_G(i)|v_G(i)) \\ &\overset{(c)}{=} h(w_G(i)) - h(q_G(i)), \end{aligned} \tag{5.50}$$

where (a) and (b) follow from Facts C.2 and C.1 and the definition of $q_G(k)$, and (c) follows from the fact that $q_G(i)$ is independent of v_G^i (and in particular independent of $v_G(i)$; see proof of Lemma 5.1). Since d_G, $x_{H,G}(0)$ and q_G are jointly Gaussian, so is $w_G(i)$ for every i. Therefore, we conclude from (5.50) and Example C.1, that

$$I(v_G(i);w_G(i)) = \frac{1}{2}\ln\frac{\sigma_{w_G}^2(i)}{\sigma_{q_G}^2} = \frac{1}{2}\ln\left(1+\frac{\sigma_{v_G}^2(i)}{\sigma_{q_G}^2}\right), \tag{5.51}$$

where the last equality follows form the independence of $q_G(i)$ and $v_G(i)$. Since the loop is stable and d_G is stationary,[14] then v_G is a second order asymptotically stationary Gaussian process. It follows from the Césaro mean convergence theorem (see, e.g., [31]), (5.51) and (5.15) that

$$I_\infty(v_G;w_G) = \lim_{k\to\infty}\frac{1}{k}\sum_{i=0}^{k-1}\frac{1}{2}\ln\left(1+\frac{\sigma_{v_G}^2(i)}{\sigma_{q_G}^2}\right) = \frac{1}{2}\ln\left(1+\frac{\sigma_{v_G}^2}{\sigma_{q_G}^2}\right), \tag{5.52}$$

[14]By Assumption 5.2, d is a second order wss process. Hence, Definition 5.6 implies that d_G is a wss second order Gaussian process, i.e., d_G is a second order stationary Gaussian process.

where $\sigma_{v_G}^2$ denotes the steady state variance of v_G (which, given the fact that v_G is a second order asymptotically stationary process, exists and is finite). By Definition 5.6, $\sigma_q^2 = \sigma_{q_G}^2$ and, given the fact that the stationary variance of any signal in an internally stable linear system depends only on the first and second order moments of the external signals involved, we also have that $\sigma_v^2 = \sigma_{v_G}^2$. The result thus follows from (5.52).

2. In this case we need to prove that $I_\infty(v_G \to w_G) = \frac{1}{4\pi}\int_{-\pi}^{\pi} \ln \frac{S_w(e^{j\omega})}{\sigma_q^2} d\omega$. Proceeding as above and exploiting the fact that $q_G(i)$ is independent of $v_G(i)$ and w_G^{i-1} (see proof of Lemma 5.1), we conclude that

$$I(w_G(i); v_G^i | w_G^{i-1}) = h(w_G(i)|w_G^{i-1}) - h(q_G(i)). \tag{5.53}$$

Thus,

$$I_\infty(v_G \to w_G) \stackrel{(a)}{=} \lim_{k\to\infty} \frac{1}{k} h(w_G^{k-1}) - \frac{1}{2}\ln 2\pi e \sigma_{q_G}^2, \tag{5.54}$$

where (a) follows from (5.16) and Fact C.1. Since d_G, $x_{H,G}(0)$ and q_G are jointly Gaussian, d_G is a second order and stationary process, and the loop is stable, then w_G is a second order asymptotically stationary Gaussian processes that admit a stationary PSD. Therefore, the first part of the result follows using Definition C.5, Theorem C.2 and Remark C.5 in (5.54), and the fact that $S_{w_G}(e^{j\omega}) = S_w(e^{j\omega})$ and that $\sigma_{q_G}^2 = \sigma_q^2$. Inequality (a) in (5.49) follows from Jensen's inequality (see, e.g., [31]), the independence of $q(i)$ and $v(i)$, and the concavity of the ln function.

□□□

Remark 5.6 *Note that inequality* (a) *in* (5.49) *is consistent with Lemma 5.1.* □□

Theorem 5.2 provides explicit upper bounds for the average (scalar or directed) mutual information across any i.i.d. source coding scheme embedded in a stable feedback loop. These bounds are stated in terms of expressions that involve only spectral characteristics of the input and output of the source coding scheme and the corresponding quantization noise. As such, they can be used as the basis of synthesis and design procedures (see Chapter 6).

In particular, Part 1 of Theorem 5.2 states that there exists a one-to-one correspondence between the signal-to-noise ratio of an i.i.d. source coding scheme and a bound on the corresponding average scalar mutual information. This is a key result and, by virtue of Theorem 5.1, suggests[15] that studying minimal signal-to-noise ratio requirements to guarantee a certain closed loop property (e.g., stability or a certain level of performance) could provide an upper bound on the average data-rate needed to guarantee the same property, when a memoryless i.i.d. source coding scheme is employed.

From Part 2 of Theorem 5.2 we see that the signal-to-noise ratio of an i.i.d. source coding scheme not only provides an upper bound on the associated average scalar mutual information, but also on the corresponding average directed mutual information. Again, this suggests that studying minimal requirements on the source coding scheme signal-to-noise ratio that guarantee a certain closed loop property could help in establishing upper bounds on the average data-rate needed to guarantee the same property, when full-memory i.i.d. source coding schemes are employed. Furthermore, we see that, if it were possible to assume, without loss of generality, that $S_w(e^{j\omega})/\sigma_q^2$ is constant, then signal-to-noise ratio considerations would provide our best upper bound on average directed mutual information.

From the previous discussion we conclude that, for the setting studied in Theorem 5.2, there exists a *strong link* between the stationary signal-to-noise ratio of a source coding scheme and the associated average data-rate. We think that the insights provided by these results are of fundamental importance and, to the best of our knowledge, new.

We end this section noting that the following holds:

Corollary 5.1 (Bounds on average mutual information (Gaussian case)) *Consider the setup of Figure 5.3, where q is the quantization noise in an i.i.d. source coding scheme, and suppose that Assumption 5.2 holds. If, in addition, d and the initial state of $H(z)$ are jointly Gaussian, then:*

1. *The average scalar mutual information between v and w is bounded by*

$$\frac{1}{2}\ln(1+\gamma) \leq I_\infty(v;w) \leq \frac{1}{2}\ln(1+\gamma) + D(q(k)||q_G(k)). \tag{5.55}$$

2. *The average directed mutual information between v and w is bounded by*

$$\frac{1}{4\pi}\int_{-\pi}^{\pi}\ln\frac{S_w(e^{j\omega})}{\sigma_q^2}d\omega \leq I_\infty(v\to w) \leq \frac{1}{4\pi}\int_{-\pi}^{\pi}\ln\frac{S_w(e^{j\omega})}{\sigma_q^2}d\omega + D(q(k)||q_G(k)). \tag{5.56}$$

Proof: Immediate from Theorem 5.2, its proof, and Lemma 5.2. □□□

[15] We note that, for the moment, we cannot be more precise because Theorems 5.2 and 5.1 involve inequalities in opposite directions.

We thus conclude that, if the conditions of Corollary 5.1 are met, then all our previous conclusions can be made stronger in the sense that we can use spectral properties of v and w to provide not only upper bounds, but also lower bounds on the average mutual information across i.i.d. source coding schemes. (It also follows from Corollary 5.1 that if, in addition to the conditions given there, q is Gaussian, then average [scalar and directed] mutual information can be evaluated by using only spectral properties of v and w.)

Remark 5.7 (Previous work) *In [40] the author establishes an equivalence between the stabilization of LTI systems over additive white Gaussian noise channels, and a communication scheme (with feedback) that reliably communicates the initial state of the plant over the same additive white Gaussian noise channel. The average data-rate of the communication scheme is shown to be equal to the Bode integral of the sensitivity of the associated LTI feedback loop (hence, to the sum of the logarithm of the unstable plant poles), and also to the average directed mutual information across the channel. We note that the analysis in [40] implicitly assumes strongly stabilizable plant models, and focuses on a situation where the only noise source is the Gaussian channel noise. Our results can be regarded as an extension of the results in [40] to more general cases (and collapse to those in [40] if q is assumed Gaussian and d is assumed zero). However, a fundamental difference arises when one realizes that [40] deals with analog communications, whereas our results were constructed assuming a digital communications setting. The fact that our results are consistent with those in [40] serves to emphasize the fundamental relevance of directed mutual information as a bound on the average data-rate at which communication systems with feedback convey information.*

In related work, [105] studies fundamental information flow inequalities that make explicit the interplay between average directed mutual information, disturbance attenuation (as measured by Bode-integral-like quantities) and channel capacity. The assumptions considered in [105] are more general than ours, but the setting and aims are different: [105] does not focus on digital communications, and places emphasis on fundamental limitations rather than design-facilitating results. In our case, and given the fact that we are focusing on a restricted class of source coding schemes, we are able to obtain simple characterizations of average data-rates in terms of well known quantities, for which results like those in [105] played an essential role (see also Remark C.5 in Appendix C). Related work that focuses on situations where the channel is not part of a feedback loop can be found in [107]. □□

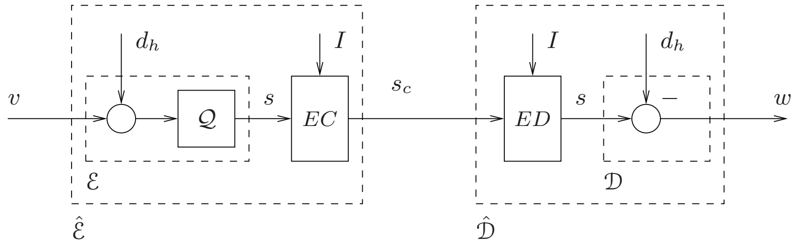

Figure 5.5: Entropy coded dithered quantizer.

5.5 Entropy Coded Dithered Quantizers

So far we have considered generic i.i.d. source coding schemes. This section focuses on a specific class of i.i.d. source coding schemes: entropy coded dithered quantizers (ECDQs; see, e.g., [190, 191]). ECDQs constitute *conceptually simple* source coding schemes, that have convenient properties that make them suitable for use as building blocks when dealing with average data-rate limits. We also note, in passing, that it has been shown that, in some situations, ECDQs are (almost) optimal from a rate-distortion point of view (see, e.g., [190–192]).

The general structure of an ECDQ is shown in Figure 5.5. In that figure, the output of the encoder $\mathcal{E}$ satisfies

$$s(k) = \mathcal{Q}(v(k) + d_h(k)), \tag{5.57}$$

where d_h is a signal that is assumed known at both the encoder $\mathcal{E}$ and decoder $\mathcal{D}$ (the *dither signal*), and $\mathcal{Q}$ corresponds to a (lattice) uniform quantizer, i.e.,

$$\mathcal{Q}(x) \triangleq i\Delta, \quad \text{for } \left(i - \frac{1}{2}\right)\Delta \leq x < \left(i + \frac{1}{2}\right)\Delta, \quad i \in \mathbb{Z}, \tag{5.58}$$

where $\Delta > 0$ is the quantization step. The output of the encoder is then losslessly coded by the EC to generate the binary words s_c. On the receiving end, the ED recovers $s(k)$ and the decoder $\mathcal{D}$ constructs its output w via

$$w(k) = s(k) - d_h(k). \tag{5.59}$$

A standard result states that, if the dither signal d_h is properly chosen, then the quantization noise q (defined as in (5.33)) can be made independent of v and $q(k) \sim -d_h(k)$ (see, e.g., [60, 152, 190, 191]). The many proofs of that result assume (implicitly) that there exists no feedback from w to v. When there exists feedback around the dithered quantizer, it is not immediately obvious whether a similar result holds or not. This issue is explored in the next theorem:

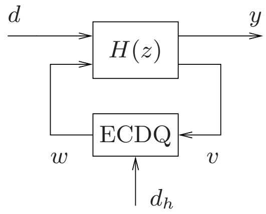

Figure 5.6: Entropy coded dithered quantizer (ECDQ) inside a feedback loop.

Theorem 5.3 (Noise in dithered uniform quantizers with feedback) *Consider the setup in Figure 5.6, where the ECDQ is as described above and has a finite quantization step* Δ*, and suppose that Assumption 5.2 holds. If* d_h *is an independent i.i.d. sequence such that*[16] $d_h(k) \sim U\left(-\frac{\Delta}{2}, \frac{\Delta}{2}\right)$*, then the quantization noise* $q \triangleq w - v$ *is an independent i.i.d. sequence and* $q(k) \sim U\left(-\frac{\Delta}{2}, \frac{\Delta}{2}\right)$*.*

Proof: (The proof proceeds along the same lines as the proof of Theorem 1 in [191] with some changes due to the fact that, in the present case, there exists feedback around the quantizer.)

Define $M \triangleq (q^{i-1}, x_H(0), d)$. It suffices to show that $\mathcal{P}\{q(i) \leq \alpha | M\}$ has the same value regardless of M, and that $q(i)$ is uniformly distributed in $\left(-\frac{\Delta}{2}, \frac{\Delta}{2}\right)$.

Define the (indicator) function $I_\alpha(\cdot)$ via

$$I_\alpha(x) = \begin{cases} 1 & \text{if } \mathcal{Q}(x) - x \leq \alpha, \\ 0 & \text{if } \mathcal{Q}(x) - x > \alpha. \end{cases} \tag{5.60}$$

Thus,

$$\begin{aligned} \mathcal{P}\{q(i) \leq \alpha | M\} &= \mathcal{P}\{\mathcal{Q}(v(i) + d_h(i)) - (v(i) + d_h(i)) \leq \alpha | M\} \\ &\stackrel{(a)}{=} \mathcal{E}\{I_\alpha(v(i) + d_h(i)) | M\} \\ &\stackrel{(b)}{=} \mathcal{E}\{\mathcal{E}\{I_\alpha(v(i) + d_h(i)) | M, v(i)\} | M\}, \end{aligned} \tag{5.61}$$

where (a) follows from the definition of $I_\alpha(\cdot)$, and (b) from the properties of iterated expectations (see, e.g., [39, 69]). To proceed further we note that:

- The i^{th} sample of d_h, i.e., $d_h(i)$, is independent of $(v(i), M)$ (hence, independent of $v(i)$ and independent of M). Indeed, $(v(i), M)$ is a deterministic function of $x_H(0), d, d_h^{i-1}$ (recall

[16] $U\left(-\frac{\Delta}{2}, \frac{\Delta}{2}\right)$ denotes a uniform distribution with support $\left(-\frac{\Delta}{2}, \frac{\Delta}{2}\right)$.

that there is not feed-through from w to v). Hence, since d_h is independent and i.i.d., our claim follows. (We note that the same argument shows that $d_h(i)$ is independent of v^i.)

- The quantizer has a lattice structure, namely, each quantization cell (or interval) P_i is generated via $P_i = \ell_i + P_0$, where $P_0 \triangleq \left(-\frac{\Delta}{2}, \frac{\Delta}{2}\right)$ and $\ell_i \triangleq i\Delta$, $i \in \mathbb{Z}$. The set $\{\ell_i\}_{i\in\mathbb{Z}}$ corresponds to the set of all reconstruction points. Accordingly,

$$Q(x) - x = Q(x + \ell_i) - (x + \ell_i), \quad \forall x \in \mathbb{R}, \quad \forall i \in \mathbb{Z}, \tag{5.62}$$

  i.e., the function $Q(\cdot) - (\cdot)$ has *lattice periodicity*. The same applies to $I_\alpha(\cdot)$.

If $v(i) \in P_j$, then $v_0 \triangleq v(i) - \ell_j \in P_0$. Assume, without loss of generality that $v_0 \geq 0$. Let $f(u)$ (resp. $f(u|M, v(i))$) denote the distribution of $d_h(i)$ (resp. conditional distribution of $d_h(i)$ given $(M, v(i))$). Then,

$$\begin{aligned}
\mathcal{E}\left\{I_\alpha\left(v(i) + d_h(i)\right)|M, v(i)\right\} &\overset{(a)}{=} \int_{\mathbb{R}} I_\alpha\left(v(i) + u\right) f(u|M, v(i)) du \\
&\overset{(b)}{=} \int_{\mathbb{R}} I_\alpha\left(v(i) + u\right) f(u) du \\
&\overset{(c)}{=} \frac{1}{\Delta} \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} I_\alpha\left(v(i) + u\right) du \\
&\overset{(d)}{=} \frac{1}{\Delta} \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} I_\alpha\left(v_0 + u\right) du \\
&\overset{(e)}{=} \frac{1}{\Delta} \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2} - v_0} I_\alpha\left(v_0 + u\right) du + \frac{1}{\Delta} \int_{\frac{\Delta}{2} - v_0}^{\frac{\Delta}{2}} I_\alpha\left(v_0 - \Delta + u\right) du \\
&\overset{(f)}{=} \frac{1}{\Delta} \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} I_\alpha(u) du \triangleq \kappa.
\end{aligned} \tag{5.63}$$

where (a) follows from the definition of conditional expectation, (b) follows from the fact that $d_h(i)$ is independent of $(M, v(i))$, (c) follows from the fact that $d_h(i)$ is uniformly distributed on $\left(-\frac{\Delta}{2}, \frac{\Delta}{2}\right)$, (d) and (e) follow from the periodicity if $I_\alpha(\cdot)$ and the definition of v_0, and (f) follows using a simple change of variables. Since, by definition, κ is a constant that depends only on the (fixed) characteristics of the quantizer, it is immediate to see from (5.63) and (5.61) that $\mathcal{P}\left\{q(i) \leq \alpha|M\right\}$ has the same value regardless of M, as required.

To complete the proof we need to show that $q(i)$ is uniformly distributed in $\left(-\frac{\Delta}{2}, \frac{\Delta}{2}\right)$. We have:

$$\begin{aligned}
\mathcal{P}\left\{q(i) \leq \alpha\right\} &\overset{(a)}{=} \mathcal{P}\left\{\mathcal{Q}(v(i) + d_h(i)) - (v(i) + d_h(i)) \leq \alpha\right\} \\
&\overset{(b)}{=} \mathcal{P}\left\{\mathcal{Q}(d_h(i)) - d_h(i) \leq \alpha\right\} \\
&\overset{(c)}{=} \mathcal{P}\left\{-d_h(i) \leq \alpha\right\},
\end{aligned} \tag{5.64}$$

where (a) follows from the definition of q, and (c) follows upon noting that $\mathcal{Q}(d_h(i)) = 0$. To show that (b) holds we recall that $q(i)$ is independent of M, and note that $v(i)$ is a deterministic function of M. Hence, $q(i)$ is independent of $v(i)$. Thus, we can pick, without loss of generality, $v(i) = 0$ and (b) follows. We just showed that $q(i) \sim U\left(-\frac{\Delta}{2}, \frac{\Delta}{2}\right)$ and the proof is thus complete.

□□□

From Theorem 5.3 we see that the input-output relationship in an appropriately dithered uniform quantizer is (essentially) as in the classical additive white noise model for uniform quantization (see Appendix B and also [83, 151]). The interesting point here is that this simple model becomes *exact* in the case of dithered uniform quantizers.

In the sequel, we will say that an ECDQ *is appropriately dithered* if and only if the dither signal d_h satisfies the conditions in Theorem 5.3.

Remark 5.8 (Availability of the dither) *In practice, implementation of dithered quantizers is not trivial since it requires the availability of the dither signal d_h at both the encoder and decoder [59] (which usually are physically separated). A simple solution to this would be to generate the dither signal using pseudo-random number generators that are initialized with the same seeds. This procedure introduces, however, approximations that may render the theoretical properties of dithered quantizers invalid.*

In the literature, so called non-substractive dithered quantizers have also been proposed. In such a quantizer, dither is added to the signal being quantized, but it is not substracted at the decoding side (see, e.g., [60, 185]). Even though such a quantization scheme is more practically appealing, the quantization noise cannot be rendered independent, and only moment-independence *can be achieved [185].* □□

Remark 5.9 (The dither PDF) *We note that the dither PDF does not need to be uniform to guarantee the nice properties of the quantization noise in a dithered quantizer. Indeed, [152] (see also [60]) used a reasoning based on characteristic functions[17] to characterize the class of all dither PDFs that allow one to get independence between the quantization noise and the quantizer input (in open loop). Uniform turns out to be, not only the simplest PDF that belongs to the class, but also the one that generates quantization noise of the smallest variance.* □□

We now turn to the analysis of average data-rates in ECDQs. To that end, and consistent with Remark 5.3, we will consider just two cases:

[17] Fourier transforms of PDF functions.

Definition 5.7 (Memoryless ECDQs and ECDQs with memory)

1. *An ECDQ is said to be memoryless if and only if*

$$I_{\mathcal{E}}(k) = \{v(k)\}, \quad I_{\mathcal{D}}(k) = \{s(k)\}, \quad S_{\mathcal{E}}(k) = S_{\mathcal{D}}(k) = \{d_h(k)\}. \tag{5.65}$$

2. *An ECDQ is said to have memory if and only if*

$$I_{\mathcal{E}}(k) = \{v(k), s^{k-1}\}, \quad I_{\mathcal{D}}(k) = \{s^k\}, \quad S_{\mathcal{E}}(k) = S_{\mathcal{D}}(k) = \{d_h^k\}. \tag{5.66}$$

□□

We note that, given the fact that d_h is assumed to be known at both the encoder and decoder side, $(I_{\mathcal{D}}(k) - \{s(k)\}) \cup S_{\mathcal{D}}(k)$ is actually known at the encoder side for both memoryless ECDQs and ECDQs with memory (provided storage is available in the case of ECDQs with memory).

In order to arrive at our main results, we next verify that both memoryless ECDQs and ECDQs with memory are i.i.d. source coding schemes:

Corollary 5.2 (ECDQs are i.i.d. source coding schemes) *If an ECDQ (with or without memory) is appropriately dithered and the associated quantization step is finite, then it is an i.i.d. source coding scheme (in open loop and also in internally stable and well posed feedback loops with strictly causal feedback).*

Proof: By definition of d_h and $\mathcal{Q}$ we have that in an ECDQ

$$s(k) = \mathcal{Q}(w(k)), \quad d_h(k) = \mathcal{Q}(w(k)) - w(k), \quad w(k) = s(k) - d_h(k), \tag{5.67}$$

i.e., knowledge of $w(k)$ (resp. w^k) is equivalent to knowledge of $d_h(k)$ and $s(k)$ (resp. d_h^k and s^k). Moreover, we know from the proof of Theorem 5.3 that $d_h(k)$ is independent of $v(k)$ for every k. It is thus straightforward to verify that both memoryless ECDQs and ECDQs with memory satisfy Assumption 5.1. The proof can be completed using Theorem 5.3. □□□

We will also need the following additional result:

Corollary 5.3 (Entropy of $s(k)$) *Consider the setup in Figure 5.6, where the ECDQ is as described above and has a finite quantization step Δ, and suppose that Assumption 5.2 holds. Assume that the information available to the EC and the ED is given by $I(k) = (I_{\mathcal{D}}(k) - \{s(k)\}) \cup S_{\mathcal{D}}(k)$, and that the ECDQ is appropriately dithered. Then:*

1. *For memoryless ECDQs,* $H(s(k)|I(k)) = I(v(k); w(k))$.

2. *For ECDQs with memory,* $H(s(k)|I(k)) = I(w(k); v(k)|w^{k-1}) = I(w(k); v^k|w^{k-1})$.

Proof: From Corollary 5.2, its proof, and Definition 5.5, we see that ECDQs satisfy the conditions of Theorem 5.1 and that (5.67) holds. Also, by virtue of the proof of Theorem 5.3, v^k (and, in particular, $v(k)$) is independent of $d_h(k)$.

1. In this case $I(k) = \{d_h(k)\}$ (see Definition 5.7) and the results follow using the same argument as in, e.g., the proof of Theorem 2 in [191]. Alternatively, one can also use the proof of Theorem 5.1 to prove the result. Indeed, it is immediate to see from that proof that $H(s(k)|I(k)) = I(v(k); w(k))$ if and only if the following Markov chains hold:

$$\begin{aligned} \{I_{\mathcal{E}}(i), S_{\mathcal{E}}(i)\} \setminus S_{\mathcal{D}}(i)\big|_{S_{\mathcal{D}}(i)} &\leftrightarrow w(i)\big|_{S_{\mathcal{D}}(i)} \leftrightarrow s(i)\big|_{S_{\mathcal{D}}(i)}, \\ w(i)\big|_{S_{\mathcal{D}}(i)} &\leftrightarrow v(i)\big|_{S_{\mathcal{D}}(i)} \leftrightarrow \{I_{\mathcal{E}}(i), S_{\mathcal{E}}(i)\} \setminus S_{\mathcal{D}}(i)\big|_{S_{\mathcal{D}}(i)}, \\ v(i) &\leftrightarrow w(i) \leftrightarrow S_{\mathcal{D}}(i), w(i). \end{aligned} \tag{5.68}$$

The definition of a memoryless ECDQ and (5.67) allows one to easily see that all these Markov chains hold and hence the result follows.

2. In this case $I(k) = \{d_h^k, s^{k-1}\}$ (see Definition 5.7). From the proof of Theorem 5.1 we conclude that $H(s(k)|d_h^k, s^{k-1}) = I(w(k); v(k)|w^{k-1})$ if and only if

$$\begin{aligned} \{I_{\mathcal{E}}(i), S_{\mathcal{E}}(i)\} \setminus \left\{s^{i-1}, S_{\mathcal{D}}(i)\right\}\big|_{s^{i-1}, S_{\mathcal{D}}(i)} &\leftrightarrow w(i)\big|_{s^{i-1}, S_{\mathcal{D}}(i)} \leftrightarrow s(i)\big|_{s^{i-1}, S_{\mathcal{D}}(i)}, \\ w(i)\big|_{s^{i-1}, S_{\mathcal{D}}(i)} \leftrightarrow v^i_{\max\{0, i-\ell\}}\big|_{s^{i-1}, S_{\mathcal{D}}(i)} &\leftrightarrow \{I_{\mathcal{E}}(i), S_{\mathcal{E}}(i)\} \setminus \left\{s^{i-1}, S_{\mathcal{D}}(i)\right\}\big|_{s^{i-1}, S_{\mathcal{D}}(i)}, \\ v^i_{\max\{0, i-\ell\}} &\leftrightarrow w^i \leftrightarrow w^i, S_{\mathcal{D}}(i). \end{aligned} \tag{5.69}$$

The definition of a full-memory ECDQ and (5.67) allows one to easily see all these Markov chains hold and, therefore, the first part of the result follows. To conclude the proof, we note that, since $q(k)$ is independent of v^k, we have $I(w(k); v(k)|w^{k-1}) = I(w(k); v^k|w^{k-1})$ (see the proof of Lemma 5.2).

□□□

From (5.7), (5.9), Definition 5.4 and Corollary 5.3, we conclude that, when an appropriately dithered ECDQ is employed, the only thing that precludes equality in the results of Theorem 5.1 is the inefficiency of the lossless coder-decoder pair (EC and ED) embedded in the ECDQ.

The next corollary provides a bound on the gap between the left and right hand sides of the inequalities of Theorem 5.1 for the case of ECDQs.

Corollary 5.4 (Average data-rates in ECDQs) *Consider the setup in Figure 5.6, where the ECDQ is as described above and has a finite quantization step Δ, and suppose that Assumption 5.2 holds. If the ECDQ is appropriately dithered, then there exists an EC and ED pair such that:*

1. *For memoryless ECDQs,* $I_\infty(v; w) \leq \bar{R} \leq I_\infty(v; w) + \ln 2$.
2. *For ECDQs with memory,* $I_\infty(v \to w) \leq \bar{R} \leq I_\infty(v \to w) + \ln 2$.

Proof: This result is an immediate consequence of (5.8), (5.9), Corollary 5.3 and Definition 5.4. □□□

Remark 5.10 (Design of EC and ED) *Corollary 5.4 is an existence-type result. It does not give any clue as how to actually design the EC or the ED. To illustrate the difficulties in the design of the EC let us consider any source coding scheme with memory (the same comments apply, of course, to ECDQs with memory). From the results in Chapter 5 in [31] it follows that for* (5.8) *to hold at every time instant k, not only knowledge of $(I_{\mathfrak{D}}(k) - \{s(k)\}) \cup S_{\mathfrak{D}}(k)$ is required, but also of the conditional distribution of $s(k)$ given $(I_{\mathfrak{D}}(k) - \{s(k)\}) \cup S_{\mathfrak{D}}(k)$. In theory, one can assume that this knowledge is available. In practice, however, the much simpler problem of just characterizing the distribution of $s(k)$ at every time instant is an awkward problem. Another practical problem arises when one notices that, in general, knowledge of $(I_{\mathfrak{D}}(k) - \{s(k)\}) \cup S_{\mathfrak{D}}(k)$ requires an amount of memory that grows unbounded with time. Much of the previous complications are alleviated if one considers a memoryless source coding scheme. For example, use of a memoryless ECDQ does not require any storage and, in addition, the design of the EC requires knowledge of only the conditional distribution of $s(k)$ given the current value of the dither signal.*

In this thesis we will content ourselves with the guarantee of the existence of appropriate ECs and EDs provided by Corollary 5.4. However, we note, for completeness, that the definition of ECs, EDs and the results of Theorem 5.3, imply that, irrespective of how the EC and ED are actually chosen, the quantization noise in an ECDQ is independent, i.i.d. and uniformly distributed. This has implementation-related implications:

(i) If the EC is designed assuming that $s(k)$ has any (finite entropy) distribution sharing the same alphabet as that of the true distribution of $s(k)$,[18] *then the relationship between v*

[18]Note that this alphabet depends only on the choice of Δ in the quantizer.

and w established in Theorem 5.3 does not change. Accordingly, any system connected to the ECDQ will not be affected by the mismatch. The only effect of considering a wrong distribution for $s(k)$ is that the average data-rate of the resulting ECDQ may be far from the one predicted by our theory.

(ii) It follows from our previous comment that the statistics of any signal in the feedback loop of Figure 5.6 are not dependent *on the choice of EC and ED. This is not a minor observation. On the contrary, this fact makes the design of ECDQs inside strictly causal feedback loops possible: the statistics of v, and hence the statistics of the encoder output s, depend only on the statistics of d and the initial state, and on the uniformity and i.i.d.-ness of the quantization noise. This allows one to characterize the distribution of $s(k)$ given the current value of the dither $d_h(k)$ and, using this knowledge, to design the EC and the ED in a memoryless ECDQ.*

(iii) The design of the EC and the ED can be carried out considering the stationary distribution of $s(k)$ (which, given our previous discussion, is independent of the actual EC and ED design). This will ensure that the expected length of the binary words generated by the EC satisfies (5.8) *in steady state and, accordingly, the average data-rate will be as predicted by our results. We note that this conclusion implicitly assumes that the transients do not incur binary words with unbounded expected lengths. The latter can be avoided if the ECDQ is embedded in an asymptotically stable and well-posed feedback loop, with second order initial states and exogenous inputs.*

We would also like to emphasize, once again, that the use of any source coding with finite average data-rate does not guarantee finite instantaneous data-rates. That is, even though the instantaneous expected length of the binary words generated by the EC is bounded[19]*, their instantaneous lengths may be unbounded. If bounded instantaneous binary word lengths need to be guaranteed, then one can saturate the quantizer input values (see also Appendix B). However, this may lead to instability if the ECDQ is placed inside a feedback loop. Needless to say, if the support of v is finite, then it immediately follows that the alphabet of s will be finite and, hence, all binary words will be of bounded length.* □□

From Corollary 5.4 we see that ECDQs provide a simple way to achieve average data-rates that are close to the absolute lower bounds established in Theorem 5.1. The gap, namely $\ln 2$

[19] This is guaranteed by the fact that the (conditional) entropy of its input is bounded (otherwise, the source coder would not have finite average data-rate).

nats (1 bit) is intrinsic to any static lossless coder and cannot, in principle, be circumvented unless one uses block entropy coding (see, e.g., [31, 156]), special time varying policies (see, e.g., [116,175]), or focuses on the high-rate regime (i.e., when $\bar{R} \to \infty$; [59]). We note, however, that this gap corresponds to a worst case rate loss. In practice the gap may be much smaller than $\ln 2$ nats (see, e.g., [59]).

Care must be taken when interpreting the conclusions of the previous paragraph: ECDQs are such that the quantization noise q is i.i.d. and uniform and, as such, corresponds to a special case of quantization noise that is not guaranteed to minimize average (scalar or directed) mutual information within the class of all noises generated by causal source coding schemes. We will see in Chapter 6 that, in some cases, it is possible to guarantee the optimality of i.i.d. quantization noise, but the fact that an ECDQ generates uniform quantization noise is the source of an additional (small) rate loss even in those cases.

We are now ready to present the final result in this chapter:

Corollary 5.5 (Explicit bounds on average data-rates in ECDQs) *Consider the setup in Figure 5.6, where the ECDQ is as described above and has a finite quantization step Δ, and suppose that Assumption 5.2 holds. If the ECDQ is appropriately dithered, then it is possible to design EC and ED in a way such that:*

1. *For a memoryless ECDQ,*

$$\bar{R} \leq \frac{1}{2} \ln (1 + \gamma) + \frac{1}{2} \ln \left(\frac{2\pi e}{12} \right) + \ln 2, \tag{5.70}$$

where, consistent with our previous definition, γ is the stationary signal-to-noise ratio of the ECDQ , i.e., $\gamma \triangleq \frac{\sigma_v^2}{\sigma_q^2}$, where σ_v^2 is the stationary variance of the input v to the ECDQ, and σ_q^2 is the variance of the corresponding quantization noise q.

2. *For an ECDQ with memory,*

$$\bar{R} \leq \frac{1}{4\pi} \int_{-\pi}^{\pi} \ln \frac{S_w(e^{j\omega})}{\sigma_q^2} d\omega + \frac{1}{2} \ln \left(\frac{2\pi e}{12} \right) + \ln 2, \tag{5.71}$$

where $S_w(e^{j\omega})$ is the stationary PSD of the output w of the ECDQ.

3. *If, in addition, d and the initial state of $H(z)$ are jointly Gaussian, then:*

   (a) *For memoryless ECDQs,*

$$\bar{R} \geq \frac{1}{2} \ln (1 + \gamma). \tag{5.72}$$

(b) For ECDQs with memory,

$$\bar{R} \geq \frac{1}{4\pi} \int_{-\pi}^{\pi} \ln \frac{S_w(e^{j\omega})}{\sigma_q^2} d\omega. \tag{5.73}$$

Proof: The result follows immediately from Theorem 5.2, Corollaries 5.1 and 5.4, and Example C.2. □□□

Corollary 5.5 provides an explicit and analytic characterization of an upper bound on the average data-rate in ECDQs (with and without memory) in terms of spectral properties of its input and output. As such, Corollary 5.5 is a fundamental result that will play an essential role in Chapter 6 when we deal with control systems subject to average data-rate limits. We note that the second term in the upper bounds provided by Corollary 5.5, namely $\frac{1}{2}\ln\left(\frac{2\pi e}{12}\right)$, corresponds to the divergence of the ECDQ quantization noise from Gaussianity and arises because ECDQs generate uniform quantization noise only (and not Gaussian noise). As mentioned earlier, the third term in the bounds provided by Corollary 5.5, namely $\ln 2$, arises due to the inefficiency of the EC embedded in the ECDQ and cannot be removed, unless one assumes $\bar{R} \to \infty$, uses block entropy coding, or other sophisticated source coding schemes.

5.6 Summary

In this chapter we have used fundamental information theoretic concepts to arrive at a framework that is suitable for analyzing and designing control systems subject to average data-rate constraints. The main results in this chapter are related to the fact that, when focusing on i.i.d. source coding schemes, average data-rates across ideal digital channels can be bounded from above by the signal-to-noise ratio of a related additive i.i.d. noise channel. This implies that the tools and insights developed in Chapter 4 are, in principle, useful to tackle control problems with average data-rate constraints. This will be explored in Chapter 6.

Future work related to the material in this chapter should focus on the case of channels with vector inputs, and also on situations where multiple channels are present. It would be also interesting to establish conditions under which one can restrict the attention to i.i.d. source coding schemes without any loss of generality. We suspect that this is rarely the case, but a characterization of such cases would be interesting (see also the recent thesis [35], where constrained source coding problems are studied).

Chapter 6

Control with Average Data-Rate Limits

6.1 Introduction

In this chapter we study networked control systems where the feedback communication link comprises an ideal digital channel (see Definition 5.2). Instead of considering coding and control architectures of arbitrary complexity (as in, e.g., [116, 118, 175]), we will focus on coding schemes that employ only LTI filters and i.i.d. source coding schemes (see Definition 5.5). Within this setup, we study the interplay between average data rate limits and MSS and performance of the resulting networked control system. When doing so, the results in Chapter 5 are shown to play an essential role: they enable us to address these questions by using the results developed in Chapter 4 for the case of signal-to-noise ratio constrained additive i.i.d. noise channels.

Using the constrained control and coding architectures described above, we show that it is possible to achieve average data-rates that guarantee MSS and are larger than the lowest achievable rate identified in [116] by the sum of two terms: a term due to the divergence of the distribution of quantization noise from Gaussianity ($\frac{1}{2}\ln\frac{2\pi e}{12}$ nats), and a term that originates from the inefficiency of the loss-less coding scheme employed to generate the channel symbols ($\ln 2$ nats). In our results, the information available to the coder and to the decoder plays an essential role (which is not surprising given the results in [175, 176]). In particular, if in the proposed scheme one constrains the encoder to use only the current plant output measurement, and the decoder to use only the last received symbol, then the average data-rate will be, in some cases, much greater than the bound in [116]. We also provide an explicit worst-case characterization of the additional rate in this case. This result constitutes an improvement over the results of [175], which guarantee bounded (but otherwise unknown) rates for stability when the encoder has only access to the current plant output measurement. An important feature of our results is that the insights provided by the proofs of necessity, are of fundamental importance when constructing a scheme that achieves average data-rates close to the necessary bounds. This is not the case for most approaches found in the literature (see, e.g., [116, 118, 175]).

This chapter also addresses performance related questions. In particular, we give a guaranteed upper bound on the minimal average data-rate that is needed to attain a desired performance level. This bound is given in terms of the minimal signal-to-noise ratio of a specific additive i.i.d. noise channel that guarantees the same performance level. As such, our results rely heavily on the results of Chapter 4. These results are believed to be important since they give, for the specific class of control and coding architectures considered here, a partial answer to the question of what is the smallest average data-rate that allows one to achieve a given performance level.

The remainder of this chapter is organized as follows: Section 6.2 presents a description of the considered architecture and the proposed control and coding scheme. Section 6.3 studies the interplay between average data-rate limits and the MSS of the proposed architecture. We put special emphasis on the role played by the degrees of freedom of our proposal and by the memory in the i.i.d. source coding scheme. Section 6.4 addresses performance related questions. Section 6.5 includes a simple simulation study and, in Section 6.6, we present concluding remarks.

6.2 Problem Definition

In this chapter we will focus on the networked control situation depicted in Figure 6.1. In that figure, d is a disturbance signal, r is a reference signal and the channel is an ideal digital channel, as described in Definition 5.2.

As was the case in previous chapters, we will work under the assumption that it is impossible to use any communication channels other than those explicitly shown in Figure 6.1. We also constrain ourselves to control and coding architectures that use only LTI filters and i.i.d. source coding schemes. With these constraints in mind, we propose the control and coding architecture depicted in Figure 6.2, where $M(z)$ and $F(z)$ are LTI systems in $\mathcal{R}_p^{1\times 2}$ that are to be designed, and $\hat{\mathcal{E}}$ and $\hat{\mathcal{D}}$ constitute an i.i.d. source coding scheme.

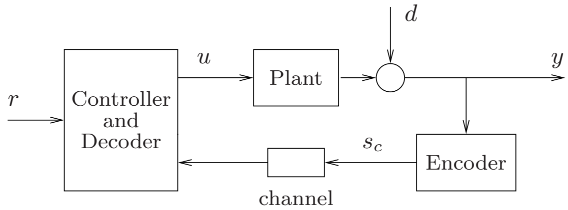

Figure 6.1: General networked control system with ideal digital channel in the feedback path.

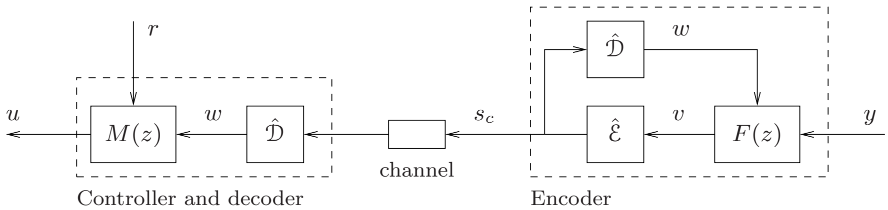

Figure 6.2: Proposed coding scheme.

The proposed control and coding architecture is the extension of our proposal in Chapter 4 to the case of average bit-rate limited channels.[1] We will thus maintain the notation introduced in Chapter 4, which we repeat here for ease of reference:

$$v = F(z) \begin{bmatrix} w \\ y \end{bmatrix} \triangleq F_1(z)w + F_2(z)y, \qquad u = M(z) \begin{bmatrix} r \\ w \end{bmatrix} \triangleq M_1(z)r + M_2(z)w, \tag{6.1}$$

and

$$K(z) \triangleq \begin{bmatrix} F(z) & M(z) \end{bmatrix}. \tag{6.2}$$

We note that, in the present case, the constraint $F_1(z) \in \mathcal{R}_{sp}$ arises naturally. Indeed, if it was not satisfied, then the feedback loop at the encoder side would be not well-posed (see also Chapter 4 in [122]). It is also possible to verify that our proposal constitutes the most general architecture that uses only LTI filters, an i.i.d. source coding scheme, and no additional communication channels that are not explicitly shown in Figure 6.1.

Given the fact that we restrict our attention to i.i.d. source coding schemes, it is straightforward to see that the networked control system that arises when the proposed coding scheme is used in the networked control system of Figure 6.1 can be *exactly modeled* by the linear feedback loop in Figure 6.3. In that figure, $G(z)$ is the plant transfer function and q models quantization

[1] We note that in the present situation we do not need to assume the existence of feedback from the output of the channel to the encoder side: since the channel is an ideal digital one, knowing its input amounts to knowing its output.

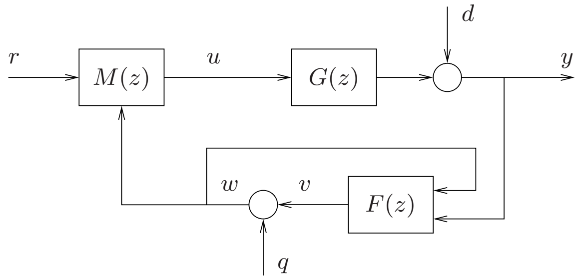

Figure 6.3: (Exact) Linear model for the networked control system that arises when employing the proposed coding scheme in the feedback loop of Figure 6.1.

noise in the i.i.d. source coding scheme. The variance of q, namely σ_q^2, depends on the way in which the source coding scheme is designed. Thus, it is natural to consider σ_q^2 as a design parameter.

The linear model of the networked control system under study is (essentially) identical to the networked control system considered in Chapter 4. It thus follows that all the results in that chapter may, in principle, be given an interpretation in terms of average data-rate limitations. This is the focus of the current chapter. We will start by studying average data-rate requirements for MSS. Later, we will address the problem of minimal average data-rates that guarantee a certain level of performance. We will be mostly concerned with the control and coding architecture in Figure 6.2, but we will also explicitly consider a case where no feedback around the encoder $\hat{\mathcal{E}}$ in Figure 6.2 is used, i.e., where $F(z) = [0 \;\; F_2(z)]$. This will highlight the importance of *degrees of freedom* in networked control systems.

For future reference, we recall the definition of the set $\mathcal{S}$ in Chapter 4, namely

$$\mathcal{S} \triangleq \{K(z) \in \mathcal{R}_{sp} \times \mathcal{R}_p \times \mathcal{R}_p \times \mathcal{R}_p : \text{the feedback loop in Figure 6.3 is internally stable and well-posed}\}, \quad (6.3)$$

and make the following assumption regarding $G(z)$ (its validity has been discussed in previous chapters):

Assumption 6.1 (Plant model) *The plant transfer function $G(z)$ belongs to $\mathcal{R}_{sp}^{1\times 1}$, is non-zero, and has a stabilizable and detectable underlying realization $(A_G, B_G, C_G, 0)$.* □□

We also remind the reader that the feedback loop in Figure 6.3 is MSS (hence, the networked control system under study is MSS) if and only if $K(z) \in \mathcal{S}$ and $\sigma_q^2 \in \mathbb{R}_0^+$ (see Section 2.5).

6.3 Mean Square Stability

In this section we characterize the average data-rates across the channel that, for the architectures of interest, guarantee MSS. We will start by considering an idealized setting, where Gaussianity conditions are imposed. Later, we will use these results to draw conclusions that are valid in more general cases.

Given the fact that the problem of stabilization is trivial for stable plant models,[2] we will focus only on unstable plant models.

6.3.1 Necessary bounds for mean square stability in the Gaussian case

We will assume throughout this sub-section that the following holds (see Figure 6.3):

Assumption 6.2 (Gaussianity)

(a) The i.i.d. source coding scheme is such that the quantization noise introduced, i.e., q, is Gaussian.

(b) Both r and d are jointly second order, stationary, Gaussian and mutually independent random processes with rational PSDs. At least one signal, r or d, is non-zero. If r or d are non-zero, then they admit spectral factors in $\mathcal{U}_\infty$.

(c) The initial states of all LTI filters in Figure 6.3 (including the plant) are jointly second order and Gaussian random variables.

□□

A. Coding schemes with feedback

We will first consider the case where $K(z)$ can be chosen freely in $\mathcal{S}$ (see (6.2) and (6.3)) and study lower bounds on the average data-rate for MSS, when either full-memory or memoryless i.i.d. source coding schemes are employed. Surprisingly, it turns out that both situations give (for most cases of interest) the same lower bound on the average data-rate.

We start by studying the full-memory case. For this case we know that the average data-rate is lower bounded by the average directed mutual information across the coding scheme (see Theorem 5.1). Thus, we first examine directed mutual information requirements for MSS:

[2] If the plant is stable and left in open loop, then one achieves MSS with zero data-rate.

Theorem 6.1 (Minimal directed mutual information for MSS) *Consider the networked control situation of Figure 6.1, where the control and coding architecture is as in Figure 6.2, the blocks $\hat{\mathcal{E}}$ and $\hat{\mathcal{D}}$ form an i.i.d. source coding scheme, and Assumptions 6.1 and 6.2 hold. If $G(z)$ is unstable, then*[3]

$$I_\infty^{\inf}(v \to w) \triangleq \inf_{\substack{K(z)\in\mathcal{S}\\ \sigma_q^2\in\mathbb{R}_0^+}} I_\infty(v \to w) = \sum_{i=1}^{n_p} \ln|p_i|. \tag{6.4}$$

Unless $d = 0$, $I_\infty^{\inf}(v \to w)$ is not achievable, but it can be approached with arbitrary precision with a sufficiently large (but finite) $\sigma_q^2 \in \mathbb{R}_0^+$ and, e.g., any $K(z)$ such that

$$M(z) = \begin{bmatrix} 0 & -1 \end{bmatrix}, \quad F(z) = K(zI - A_G + JC_G)^{-1}\begin{bmatrix} -B_G \\ J \end{bmatrix}, \tag{6.5}$$

where J and K are such that $A_G - JC_G$ and $A_G - B_GK$ are Hurwitz matrices.[4] *If $d = 0$, then $I_\infty^{\inf}(v \to w)$ can be achieved by choosing $K(z)$ as above and any $\sigma_q^2 \in \mathbb{R}^+$.*

Proof: It is straightforward to see from Figure 6.3 that

$$w = T_{qw}(z)q + T_{drw}(z)\begin{bmatrix} d \\ r \end{bmatrix}, \tag{6.6}$$

where

$$T_{qw}(z) \triangleq \frac{1}{1 - F_1(z) - F_2(z)M_2(z)G(z)}, \quad T_{drw}(z) \triangleq F_2(z)T_{qw}(z)\begin{bmatrix} 1 & G(z)M_1(z) \end{bmatrix}. \tag{6.7}$$

Since $G(z)$ and $F_1(z)$ are strictly proper, and Assumption 6.2 holds, it follows from Corollary 5.1 and Equation (6.6) that, for any $K(z) \in \mathcal{S}$ and $\sigma_q^2 \in \mathbb{R}_0^+$,

$$\begin{aligned} I_\infty(v \to w) &= \frac{1}{4\pi}\int_{-\pi}^{\pi} \ln\left(\left|T_{qw}(e^{j\omega})\right|^2 + \frac{T_{drw}(e^{j\omega})S_{dr}(e^{j\omega})T_{drw}(e^{j\omega})^H}{\sigma_q^2}\right)d\omega \\ &\overset{(a)}{\geq} \frac{1}{2\pi}\int_{-\pi}^{\pi} \ln\left|T_{qw}(e^{j\omega})\right|d\omega \\ &\overset{(b)}{\geq} \sum_{i=1}^{n_p} \ln|p_i|, \end{aligned} \tag{6.8}$$

where we have used the independence of q from r and d, $S_{dr}(e^{j\omega})$ denotes the PSD of $[d\ r]^T$, and where (a) follows form the positive definiteness of spectral densities, and (b) follows from the Bode Integral Theorem (see, e.g., Theorem 3.4.4 in [155]), the fact that $G(z), F_1(z) \in \mathcal{R}_{sp}$

[3] Recall the notation in Section 2.2.

[4] I.e., matrices with all their eigenvalues strictly inside the unit circle.

imply $T_{qw}(\infty) = 1$, and that a necessary condition for stability is that $T_{qw}(z)$ contains as NMP zeros (at least) the unstable plant poles (see, e.g., [53,155]). It remains to prove that $I_\infty(v \to w)$ can be made arbitrarily close to $I_\infty^{\inf}(v \to w)$ by means of an appropriate choice for $K(z) \in \mathcal{S}$ and $\sigma_q^2 \in \mathbb{R}_0^+$.

Equality in (a) and (b) in (6.8) holds if and only if

$$\frac{T_{drw}(e^{j\omega})S_{dr}(e^{j\omega})T_{drw}(e^{j\omega})^H}{\sigma_q^2} = 0 \tag{6.9}$$

and $T_{qw}(z)$ has as strictly-NMP zeros the unstable plant poles only. Since $G(z)$ is unstable and r and d are uncorrelated, $K(z) \in \mathcal{S}$ and $\sigma_q^2 \in \mathbb{R}_0^+$ imply that (6.9) holds if and only if $r = d = 0$, or $d = 0$ and $M_1(z) = 0$. On the other hand, $T_{qw}(z)$ has as strictly-NMP zeros only the unstable plant poles if and only if $F(z)$ and $M_2(z)$ are (marginally) stable. Given the choice for J, it is immediate to see that $F(z)$ in (6.5) is stable. Also, $F(z)$ in (6.5) is simply the series connection of an observer and a static gain which, by construction, are such that the feedback loop in Figure 6.3 is internally stable and well-posed when $M(z) = \begin{bmatrix} 0 & -1 \end{bmatrix}$. As a consequence, our claims regarding the achievability of $I_\infty^{\inf}(v \to w)$ when $d = 0$ follow.

If $d \neq 0$, then the previous discussion allows one to conclude that inequality (a) in (6.8) is strict for every $K(z) \in \mathcal{S}$ and every $\sigma_q^2 \in \mathbb{R}_0^+$. Thus, the infimum in (6.4) is not achievable. Nevertheless, since, for every $K(z) \in \mathcal{S}$, $||T_{drw}(z)\Omega_{dr}(z)||_2^2 < \infty$, it follows that, for every fixed $K(z) \in \mathcal{S}$, the right hand side of the equality in (6.8) is a continuous function of σ_q^2 such that

$$\lim_{\sigma_q^2 \to \infty} \frac{1}{4\pi} \int_{-\pi}^{\pi} \ln \left(\left| T_{qw}(e^{j\omega}) \right|^2 + \frac{T_{drw}(e^{j\omega})S_{dr}(e^{j\omega})T_{drw}(e^{j\omega})^H}{\sigma_q^2} \right) d\omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln \left| T_{qw}(e^{j\omega}) \right| d\omega. \tag{6.10}$$

In addition, the suggested $K(z)$ is stable and, therefore, equality holds in (b) in (6.8). We thus conclude that for such a $K(z)$ and a sufficiently large (but finite) σ_q^2, it is possible to make $I_\infty(v \to w)$ arbitrarily close to $I_\infty^{\inf}(v \to w)$. This completes the proof. □□□

Theorem 6.1 establishes a bound on the average directed mutual information across the coding scheme whose satisfaction is necessary and sufficient to be able to achieve MSS, when a coding scheme with feedback is employed. It also provides a characterization of a stabilizing filter $K(z)$ and the noise variance σ_q^2 that allow one to achieve an average directed mutual information arbitrarily close to the bound established. These results, when combined with Theorem 5.1, allow one to immediately conclude that the following holds:

Corollary 6.1 (Necessary condition for MSS (I)) *Consider the networked control situation of Figure 6.1, where the control and coding architecture is as in Figure 6.2, the blocks $\hat{\mathcal{E}}$ and $\hat{\mathcal{D}}$ form an i.i.d. source coding scheme, and Assumptions 6.1 and 6.2 hold. If $G(z)$ is unstable, $d \neq 0$ and the i.i.d. source coding scheme is a full-memory one, then*

$$\bar{R} > \bar{R}_{\text{inf}} \triangleq \sum_{i=1}^{n_p} \ln |p_i| \tag{6.11}$$

is a necessary condition for MSS.

Proof: Immediate from Theorem 6.1 and Parts 2 of Theorem 5.1 and of Corollary 5.1. □□□

Remark 6.1 (Zero disturbances) *If $d = 0$, then inequality* (6.11) *becomes non-strict. We feel, however, that this case is not interesting since, in practice, there will always be some source of external noise. Indeed, in the sequel, we will be mostly interested in the case where d is not zero.* □□

Corollary 6.1 establishes an inequality that average data-rates in MSS feedback loops necessarily satisfy, when the proposed control and coding architecture is employed. Since we have recovered the absolute lower bound on average data-rates for MSS derived in [116], we conclude that the assumptions in Theorem 6.1 and Corollary 6.1 do not limit, in principle, the minimal data-rates for MSS. We note, however, that we have not, as yet, shown that the bound in (6.11) is achievable within the class of considered coding schemes. This issue will be explored in Section 6.3.2.

We next turn to the consideration of the memoryless case. Here, the average data-rate is lower bounded by the scalar mutual information across the coding scheme (see Theorem 5.1). We have the following result:

Corollary 6.2 (Minimal scalar mutual information for MSS) *Consider the networked control situation of Figure 6.1, where the control and coding architecture is as in Figure 6.2, the blocks $\hat{\mathcal{E}}$ and $\hat{\mathcal{D}}$ form an i.i.d. source coding scheme, and Assumptions 6.1 and 6.2 hold. If $G(z)$ is unstable, then*

$$I_{\infty}^{\text{inf}}(v; w) \triangleq \inf_{\substack{K(z)\in\mathcal{S} \\ \sigma_q^2 \in \mathbb{R}_0^+}} I_{\infty}(v; w) = \sum_{i=1}^{n_p} \ln |p_i|. \tag{6.12}$$

Unless $d = 0$ *and* $G(z)$ *has no poles on the unit circle,* $I_\infty^{\inf}(v; w)$ *is not achievable, but can be approached with arbitrary precision with a sufficiently large (but finite)* $\sigma_q^2 \in \mathbb{R}_0^+$ *and* $K(z)$ *as in Part 1 of Corollary 4.1 with* $M(z) \in \mathcal{RH}_\infty$ *and*[5]

$$Q(z) = -z\mathcal{K}_\varepsilon \left\{\xi_p(z)D(z)\right\}^{-1} \left(\xi_p(\infty) - X_i(z)\xi_p(z)D(z)\right) \begin{bmatrix} 1 & 0 \end{bmatrix}, \tag{6.13}$$

where $D(z) \in \mathcal{RH}_\infty$ *is any denominator of* $G(z)M_2(z)$ *in a coprime factorization of* $G(z)M_2(z)$ *over* $\mathcal{RH}_\infty$*, and* ε *is a sufficiently small (but non-zero) real in* $(0, 1]$*. If the plant has no poles on the unit circle, then* $\mathcal{K}_\varepsilon \{\cdot\}$ *in* (6.13) *is redundant and* $I_\infty^{\inf}(v; w)$ *is achievable if and only if* $d = 0$*. To do so, it suffices to choose any* $\sigma_q^2 \in \mathbb{R}^+$ *and* $K(z)$ *as above, with* $M_1(z)\Omega_r(z) = 0$*.*

Proof: Immediate from Corollary 5.1 and Theorem 4.1. □□□

Corollary 6.2 establishes a closed form expression for the minimal average scalar mutual information across the coding scheme that is compatible with MSS in the feedback system of Figure 6.3. It also provides a characterization of filters $K(z)$ and the quantization noise variance σ_q^2 that allow one to achieve an average scalar mutual information arbitrarily close to the identified limit. It is remarkable that we obtained the same minimum as in Theorem 6.1, where we explored minimal average directed mutual information for MSS. To examine the implications of this, we start by noting that the following holds:

Corollary 6.3 (Necessary condition for MSS (II)) *Consider the networked control situation of Figure 6.1, where the control and coding architecture is as in Figure 6.2, the blocks* $\hat{\mathcal{E}}$ *and* $\hat{\mathcal{D}}$ *form an i.i.d. source coding scheme, and Assumptions 6.1 and 6.2 hold. If* $G(z)$ *is unstable,* $d \neq 0$ *and the coding scheme is an i.i.d. memoryless one, then*

$$\bar{R} > \bar{R}_{\inf}, \tag{6.14}$$

where $\bar{R}_{\inf}$ *is defined in* (6.11)*, is a necessary condition for MSS.*

Proof: Immediate from Corollary 6.2 and Parts 1 of Theorem 5.1 and Corollary 5.1. □□□

As was the case in Corollary 6.1, we recover the absolute lower bound on average data-rates for MSS derived in [116]. Thus, we see from Corollaries 6.1 and 6.3 that, not only it is true that the assumptions in Theorem 6.1 and Corollary 6.2 do not limit the minimal average data-rate for MSS, but also that the use of the proposed control and coding architecture with a

[5] Recall the definition of the operator $\mathcal{K}_\varepsilon \{\cdot\}$ in Section 2.2.

memoryless i.i.d. source coding scheme does not constrain the achievable average data-rates for MSS. In other words, the *memory in the LTI filter* $K(z)$ *is sufficient to handle all the memory that, in principle, the EC within* $\hat{\mathcal{E}}$ *should have in order to minimize the average data-rate* (see also the fourth remark after Theorem 2 in [192]). Given the fact that the implementation of full-memory (or partial-memory) coding schemes is not trivial (recall Remark 5.10), we feel that this result is important. We stress, once more, that we have not shown that the bound in (6.14) is achievable within the class of considered control and coding architectures. For the moment, we can conjecture that control and coding architectures that achieve MSS at average data-rates *close* to $\bar{R}_{\text{inf}}$ do not, necessarily, need to rely on coding schemes with memory (see also Section 6.3.2).

Remark 6.2 *The results of Corollaries 6.1 and 6.3 verify the necessity of the bounds on average data-rates for MSS established in [116]. As such, our results are not surprising and, at first sight, may be considered redundant. The key point of Corollaries 6.1 and 6.3 is that, together with Theorem 6.1, Corollary 6.2, and the results in Section 5.5, allow one to easily envisage source coding schemes that achieve average data-rates close to* $\bar{R}_{\text{inf}}$*. This issue is explored in Section 6.3.2 below.* □□

We finish this section by showing that the facts discussed in the previous paragraph are manifestations of a more general result:

Theorem 6.2 (Relationship between different problems (with feedback)) *Consider the networked control situation of Figure 6.1, where the control and coding architecture is as in Figure 6.2, the blocks* $\hat{\mathcal{E}}$ *and* $\hat{\mathcal{D}}$ *form an i.i.d. source coding scheme, and Assumptions 6.1 and 6.2 hold. If* $G(z)$ *is unstable, then the problem of finding the minimal stationary coding scheme signal-to-noise ratio that guarantees MSS is* equivalent[6] *to the problem of finding the average scalar mutual information across the coding scheme that guarantees MSS, and* also provides a solution *to the problem of finding the minimal average directed mutual information across the coding scheme that guarantees MSS.*

Proof: The equivalence between minimizing the stationary source coding scheme signal-to-noise ratio and minimizing the corresponding average scalar mutual information follows immediately from Part 1 in Corollary 5.1.

To prove our remaining claim, recall the notation used in the proof of Theorems 6.1 and Corollary 6.2 (actually, Theorem 4.1) and note that the transfer function from q to w satisfies

[6] In the sense that a solution of either problem follows readily from the solution of the other (see, e.g, [16]).

$T_{qw}(z) = T_{qv}(z) + 1$. Use of (6.13) in (4.17) and the definitions of $\bar{\xi}_p^\varepsilon(z)$, $\bar{\xi}_p(z)$ and $\xi_p(z)$ allow one to conclude that: (i) if $G(z)$ has no poles on the unit circle, then $T_{qw}(z) = \xi_p(z)^{-1}\xi_p(\infty)$; (ii) if $G(z)$ has poles on the unit circle, then

$$T_{qw}(z) = X_i(z)D(z)\left(1 - \bar{\xi}_p^\varepsilon(z)\bar{\xi}_p(z)^{-1}\right) + \bar{\xi}_p^\varepsilon(z)\left(\bar{\xi}_p(z)\xi_p(z)\right)^{-1}\xi_p(\infty) \tag{6.15}$$

and $\lim_{\varepsilon\to 0^+} T_{qw}(z) = \xi_p(z)^{-1}\xi_p(\infty)$. As a consequence, we conclude that $T_{qw}(z)$ becomes an all-pass filter when $Q(z)$ is as in (6.13) and the parameter ε becomes arbitrary small.

From the above we conclude that all $K(z)$ as in Corollary 6.2 (which, of course, belong to $\mathcal{S}$) are such that

$$\lim_{\substack{\varepsilon\to 0^+\\ \sigma_q^2\to\infty}} \frac{S_w(e^{j\omega})}{\sigma_q^2} = \lim_{\substack{\varepsilon\to 0^+\\ \sigma_q^2\to\infty}} \left(\left|T_{qw}(e^{j\omega})\right|^2 + \frac{T_{drw}(e^{j\omega})S_{dr}(e^{j\omega})T_{drw}(e^{j\omega})^H}{\sigma_q^2}\right) = \left|\xi_p(z)^{-1}\xi_p(\infty)\right|^2 = \left|\xi_p(\infty)\right|^2, \tag{6.16}$$

i.e., are such that $S_w(e^{j\omega})/\sigma_q^2$ becomes asymptotically constant as $\varepsilon \to 0^+$ and $\sigma_q^2 \to \infty$ (see also Part 2 in Theorem 5.2). Thus, we immediately conclude from (6.16) that

$$\lim_{\substack{\varepsilon\to 0^+\\ \sigma_q^2\to\infty}} \frac{1}{4\pi}\int_{-\pi}^{\pi} \ln\frac{S_w(e^{j\omega})}{\sigma_q^2}d\omega = \sum_{i=1}^{n_p}\ln|p_i|. \tag{6.17}$$

Therefore, any $K(z) \in \mathcal{S}$ that achieves $I_\infty(v;w)$ arbitrarily close to $I_\infty^{\inf}(v;w)$ is also such that it makes $I_\infty(v \to w)$ arbitrarily close to $I_\infty^{\inf}(v \to w)$. This proves our claim. □□□

Theorem 6.2 is one of the main results in this thesis. It essentially states that, for the considered architecture and under our working assumptions, if one focuses on MSS, then stationary coding scheme signal-to-noise ratio and average scalar mutual information are two equivalent ways of looking at the same communication constraint. Moreover, if one is concerned about the minimal average directed mutual information across the coding scheme that guarantees MSS, then it suffices to study the minimal stationary coding scheme signal-to-noise ratio for MSS. By virtue of Theorem 5.1, it thus follows that determining the minimal stationary coding scheme signal-to-noise ratio for MSS provides a solution to the problem of determining a lower bound on the average data-rates that allow one to achieve MSS. The key enabling feature that allowed us to establish this connection is the fact that we consider a control and coding architecture that has *sufficient degrees of freedom.* This allows one to make $S_w(e^{j\omega})\sigma_q^{-2}$ constant in the limit as $\sigma_q^2 \to \infty$ and $\varepsilon \to 0^+$ (see (6.16)), without affecting the minimum achievable average (scalar or directed) mutual information across the coding scheme, nor the corresponding minimal achievable signal-to-noise ratio. (By virtue of our assumptions and Theorem 5.2 and Corollary 5.1, this implies that $I_\infty(v;w)$ approaches $I_\infty(v \to w)$ in the limit as $\sigma_q^2 \to \infty$ and $\varepsilon \to 0^+$.) The importance of having enough degrees of freedom will be made explicit in the following section that deals with a constrained version of the proposed control and coding scheme.

Remark 6.3 (Other architectures) *It is easy to see from our previous results that, if $M(z)$ is fixed at any point in $\mathcal{RH}_\infty$ that satisfies the conditions in Part 1 of Corollary 4.1, then all the previous results still apply (with minor and obvious changes). In particular, the results still hold with the simple choices $M(z) = [0 \;\; m_2]$ or $M(z) = [m_1 \;\; m_2]$, where $m_1, m_2 \in \mathbb{R}$, $m_2 \neq 0$.* □□

B. Coding schemes without feedback

In this section we focus our attention to a particular instance of the scheme in Figure 6.3, namely where $F(z) = \begin{bmatrix} 0 & F_2(z) \end{bmatrix}$. We will refer to this control and coding architecture as being *without feedback*. We will see that, due to this constraint, many of the nice results in Section 6.3.1A do not apply unless the plant model satisfies some additional conditions. For future reference we define (recall (4.3) and compare to (4.6))

$$\mathcal{S}_C \triangleq \left\{ K(z) \in \mathcal{S} : F(z) = \begin{bmatrix} 0 & F_2(z) \end{bmatrix} \right\}. \tag{6.18}$$

We will mirror the approach taken in Section 6.3.1. Accordingly, we start by studying the full-memory case and, in particular, minimal directed mutual information for MSS.

Theorem 6.3 (Minimal directed mutual information for MSS (no feedback)) *Consider the networked control situation of Figure 6.1, where the control and coding architecture is as in Figure 6.2, the blocks $\hat{\mathcal{E}}$ and $\hat{\mathcal{D}}$ form an i.i.d. source coding scheme, and Assumptions 6.1 and 6.2 hold. If $G(z)$ is unstable, then*[7]

$$I_\infty^{\inf,C}(v \to w) \triangleq \inf_{\substack{K(z)\in\mathcal{S}_C \\ \sigma_q^2 \in \mathbb{R}_0^+}} I_\infty(v \to w) = \sum_{i=1}^{n_p} \ln|p_i| + \sum_{\substack{i=1 \\ c_i \in \mathcal{I}}}^{n_c} \ln|c_i|. \tag{6.19}$$

Unless $d = 0$ and $G(z)$ is strongly stabilizable, $I_\infty^{\inf,C}(v \to w)$ is not achievable, but can be approached with arbitrary precision with a sufficiently large (but finite) $\sigma_q^2 \in \mathbb{R}_0^+$ and $K(z)$ such that

$$M(z) = \begin{bmatrix} 0 & M_2(z) \end{bmatrix}, \quad F(z) = \begin{bmatrix} 0 & 1 \end{bmatrix}, \tag{6.20}$$

[7] We remind the reader of the notation introduced in Section 2.2.

where

$$M_2(z) = \overline{M}_2(z) \prod_{\substack{i=1 \\ c_i \in \mathcal{I}}}^{n_c} \frac{z}{z - c_i(1+\varepsilon)}, \tag{6.21}$$

$\overline{M}_2(z) \in \mathcal{RH}_\infty$ *is any one-dof admissible controller for*

$$G(z) \prod_{\substack{i=1 \\ c_i \in \mathcal{I}}}^{n_c} \frac{z}{z - c_i(1+\varepsilon)}, \tag{6.22}$$

and $\varepsilon > 0$ *is a sufficiently small (but non-zero) real. If* $G(z)$ *is strongly stabilizable, then*

$$I_\infty^{\inf,C}(v \to w) = \sum_{i=1}^{n_p} \ln |p_i| \tag{6.23}$$

is achievable if and only if $d = 0$*. To do so, it suffices to pick* $K(z)$ *as above*[8] *and any* $\sigma_q^2 \in \mathbb{R}^+$*.*

Proof: From Figure 6.3 we see that for $F_1(z) = 0$,

$$w = T_{qw}(z)q + T_{drw}(z) \begin{bmatrix} d \\ r \end{bmatrix}, \tag{6.24}$$

where

$$T_{qw}(z) \triangleq \frac{1}{1 - F_2(z)M_2(z)G(z)}, \quad T_{drw}(z) \triangleq F_2(z)T_{qw}(z) \begin{bmatrix} 1 & G(z)M_1(z) \end{bmatrix}. \tag{6.25}$$

Since $G(z) \in \mathcal{R}_{sp}$ and Assumption 6.2 holds, it follows from Corollary 5.1 and (6.24) that, for any $K(z) \in \mathcal{S}_C$ and any $\sigma_q^2 \in \mathbb{R}_0^+$, we have

$$I_\infty(v \to w) \geq \sum_{i=1}^{n_p} \ln |p_i| + \sum_{i=1}^{n_2} \ln \left| p_i^{F_2 M_2} \right|, \tag{6.26}$$

where we have followed the same argument as in the proof of Theorem 6.1, and $\{p_1^{F_2M_2}, \cdots p_{n_2}^{F_2M_2}\}_m$ is the multi-set of unstable poles of $F_2(z)M_2(z)$. Standard results regarding strongly and not strongly stabilizable plants (see, e.g., [38,182]) imply that, if $K(z) \in \mathcal{S}_C$, then $F_2(z)M_2(z)$ must necessarily have at least one unstable pole in every interval $\mathcal{I}_i^+$ and every interval $\mathcal{I}_i^-$ for which n_i^+ or n_i^- is odd. As a consequence, we conclude that, necessarily,

$$I_\infty(v \to w) \overset{(a)}{\geq} \sum_{i=1}^{n_p} \ln |p_i| + \sum_{i=1}^{n_2} \ln \left| p_i^{F_2 M_2} \right| \overset{(b)}{\geq} \sum_{i=1}^{n_p} \ln |p_i| + \sum_{\substack{i=1 \\ c_i \in \mathcal{I}}}^{n_c} \ln |c_i| = I_\infty^{\inf,C}(v \to w) \tag{6.27}$$

[8] Note that, in this case, $\mathcal{I} = \phi$ and $M_2(z) = \overline{M}_2(z)$.

holds for any $\sigma_q^2 \in \mathbb{R}_0^+$ and any $K(z) \in \mathcal{S}_C$. It remains to prove that $I_\infty(v \to w)$ can be made arbitrarily close to $I_\infty^{\inf,C}(v \to w)$ by means of an appropriate choice for $K(z) \in \mathcal{S}_C$ and $\sigma_q^2 \in \mathbb{R}_0^+$.

Proceeding as in the proof of Theorem 6.1, we conclude that equality in (a) of (6.27) holds for $\sigma_q^2 \in \mathbb{R}_0^+$ if and only if $r = d = 0$, or $d = 0$ and $M_1(z) = 0$. On the other hand, since $K(z) \in \mathcal{S}_C$ requires $F_2(z)M_2(z)$ to be a one-dof admissible controller for $G(z)$, we conclude that there exists $K(z) \in \mathcal{S}_C$ such that equality in (b) holds if and only if $G(z)$ is strongly stabilizable. This proves our claim regarding strongly stabilizable plants.

If $d \neq 0$ or if the plant is not strongly stabilizable, then it follows from the above discussion that the infimum in (6.19) is not achievable. Proceeding as in the proof of Theorem 6.1, we conclude that one can always pick a sufficiently large σ_q^2 so as to make both sides of inequality (a) in (6.27) arbitrarily close (provided $K(z) \in \mathcal{S}_C$). In addition, $K(z)$, as given by (6.20)-(6.21), belongs, by construction, to $\mathcal{S}_C$ for every sufficiently small $\varepsilon > 0$. Also, the set of unstable poles of this controller tend to the subset of NMP plant zeros contained in $\mathcal{I}$ as $\varepsilon \to 0^+$. This concludes the proof. □□□

Theorem 6.3 establishes a bound on the average directed mutual information across the coding scheme whose satisfaction is necessary and sufficient to achieve MSS, when the proposed control and coding architecture is used without feedback. We also provide a characterization of the filters $K(z)$ and the noise variances σ_q^2 that allow one to achieve any average directed mutual information arbitrarily close to the bound established. It is apparent that, unless the plant is strongly stabilizable, the requirements on the average directed mutual information in this case are more stringent than in the case of using a coding architecture with feedback (see Theorem 6.1).

We can use Theorem 6.3 to conclude that the following holds:

Corollary 6.4 (Necessary condition for MSS (no feedback I)) *Consider the networked control situation of Figure 6.1, where the control and coding architecture is as in Figure 6.2, the blocks $\hat{\mathcal{E}}$ and $\hat{\mathcal{D}}$ form an i.i.d. source coding scheme, and Assumptions 6.1 and 6.2 hold. If $G(z)$ is unstable, $d \neq 0$, $F(z) = [0 \;\; F_2(z)]$ and the i.i.d. source coding scheme has full-memory, then*

$$\bar{R} > \bar{R}_{\inf,C}^m \triangleq \sum_{i=1}^{n_p} \ln|p_i| + \sum_{\substack{i=1 \\ c_i \in \mathcal{I}}}^{n_c} \ln|c_i| \tag{6.28}$$

is a necessary condition for MSS.[9] *Moreover, $\bar{R}_{\inf,C}^m \geq \bar{R}_{\inf}$ with equality if and only if $G(z)$ is*

[9] In $\bar{R}_{\inf,C}^m$, the subscript C stands for constrained (i.e., without feedback) and the superscript m for full-

strongly stabilizable.

Proof: Immediate from Theorem 6.3, Parts 2 in Theorem 5.1 and Corollary 5.1, and the fact that $\mathcal{I}$ is empty if and only if $G(z)$ is strongly stabilizable. □□□

Corollary 6.4 establishes an inequality that average data-rates in MSS feedback loops necessarily satisfy, when the proposed control and coding architecture, with a full-memory i.i.d. source coding scheme and without feedback, is employed. As opposed to Corollary 6.1, we do not recover the absolute lower bound on average data-rates for MSS derived in [116], except in cases where the plant is strongly stabilizable. This shows that the constraint imposed on $F(z)$ limits the achievable data-rates in stable loops (at least for the class of control and coding architectures considered here).

To further examine the effects of the constraint on $F(z)$, we next turn our attention to the study of architectures without feedback and without memory in the coding scheme. In this case, the average data-rate is lower bounded by the scalar mutual information (see Theorem 5.1). We have the following result:

Corollary 6.5 (Minimal scalar mutual information for MSS (no feedback)) *Consider the networked control situation of Figure 6.1, where the control and coding architecture is as in Figure 6.2, the blocks $\hat{\mathcal{E}}$ and $\hat{\mathcal{D}}$ form an i.i.d. source coding scheme, and Assumptions 6.1 and 6.2 hold. If $G(z)$ is unstable and has only finite strictly-NMP zeros of multiplicity one, then*

$$I_\infty^{\inf,C}(v;w) \triangleq \inf_{\substack{K(z)\in\mathcal{S}_C \\ \sigma_q^2\in\mathbb{R}_0^+}} I_\infty(v;w) = \sum_{i=1}^{n_p} \ln|p_i| + \frac{1}{2}\ln\left(1+\frac{\Delta_G}{\prod_{i=1}^{n_p}|p_i|^2}\right), \tag{6.29}$$

where Δ_G is as in Theorem 4.2.

Unless $d=0$ and $G(z)$ has no poles or zeros on the unit circle, $I_\infty^{\inf,C}(v;w)$ is not achievable, but can be approached with arbitrary precision with a sufficiently large (but finite) $\sigma_q^2\in\mathbb{R}_0^+$ and $K(z)$ as in Part 2 of Corollary 4.1, with $F_2(z)\in\mathcal{U}_\infty$ and

$$Q(z) = \begin{bmatrix} Q_1(z) & \mathcal{K}_\varepsilon\left\{\xi_c(z)\xi_p(z)N(z)D(z)\right\}^{-1}\left(\left\{[H(z)]_{\mathcal{H}_2^\perp}\right\}\Big|_{z=0} + [H(z)]_{\mathcal{H}_2}\right)\end{bmatrix}, \tag{6.30}$$

where $Q_1(z)\in\mathcal{RH}_\infty$, $N(z)$ and $D(z)$ form a coprime factorization over $\mathcal{RH}_\infty$ of $G(z)F_2(z)$ with $G(z)F_2(z)=N(z)D(z)^{-1}$,

$$H(z) \triangleq \xi_c(z)\xi_p(z)N(z)Y_i(z)\begin{bmatrix}0\\1\end{bmatrix}, \tag{6.31}$$

---

memory.

and ε is a sufficiently small (but non-zero) real in $(0,1]$. Moreover, if the plant has no poles or zeros on the unit circle, then $\mathcal{K}_\varepsilon\{\cdot\}$ in (6.30) is redundant and $I_\infty^{\inf,C}(v;w)$ is achievable if and only if $d=0$. To do so, it suffices to to choose $K(z)$ as indicated above, with $Q_1(z)\Omega_r(z)=0$ and any $\sigma_q^2\in\mathbb{R}^+$.

Proof: Immediate from Corollary 5.1 and Theorem 4.2. □□□

Corollary 6.5 establishes a closed form for the minimal average scalar mutual information which is necessary and sufficient to achieve MSS, when the proposed control and coding architecture is used without feedback. It also provides a characterization of the filters $K(z)$ and the noise variances σ_q^2 that allows one to achieve an average scalar mutual information arbitrarily close to the identified limit. Comparing Corollaries 6.5 and 6.2 it follows that the constraint on $F(z)$ limits the achievable coding scheme average scalar mutual information, except when $G(z)$ is marginally-MP and has relative degree one.

We next present a lemma that will allows us to examine consequences of Corollary 6.5:

Lemma 6.1 *Consider the networked control situation of Figure 6.1, where the control and coding architecture is as in Figure 6.2, the blocks $\hat{\mathcal{E}}$ and $\hat{\mathcal{D}}$ form an i.i.d. source coding scheme, and Assumptions 6.1 and 6.2 hold. Assume that $G(z)$ is unstable and that $\varepsilon\in(0,1]$ is a variable that (continuously) parameterizes $K_\varepsilon(z)$, where $K_\varepsilon(z)\in\mathcal{S}_C$ is such that*

$$\lim_{\substack{\varepsilon\to 0^+\\ \sigma_q^2\to\infty}} I_\infty(v;w)\big|_{K(z)=K_\varepsilon(z)} = I_\infty^{\inf,C}(v;w). \tag{6.32}$$

Then, there exists a constant $\kappa\in\mathbb{R}_0^+$ such that

$$\lim_{\substack{\varepsilon\to 0^+\\ \sigma_q^2\to\infty}} \left.\frac{S_w(e^{j\omega})}{\sigma_q^2}\right|_{K(z)=K_\varepsilon(z)} = \kappa, \quad \forall\omega\in[-\pi,\pi], \tag{6.33}$$

if and only if $G(z)$ is marginally-MP and has relative degree one. If that is the case, then

$$\kappa = |\xi_p(\infty)|^2. \tag{6.34}$$

Proof: We adopt the notation used in Theorem 4.2 and its proof. By Corollary 5.1, $K_\varepsilon(z)$ satisfies (6.32) if and only if also satisfies

$$\lim_{\substack{\varepsilon\to 0^+\\ \sigma_q^2\to\infty}} \gamma\big|_{K(z)=K_\varepsilon(z)} = \gamma_{\inf,C}, \tag{6.35}$$

where γ and $\gamma_{\inf,C}$ are defined as in Theorem 4.2.

- ($\Leftarrow$) Consider $G(z)$ with no strictly-NMP zeros and having relative degree one. Pick $K_\varepsilon(z)$ as $K(z)$ in Corollary 6.5 with $F_2(z) = 1$ and $Q_1(z) = 0$ (clearly such a $K_\varepsilon(z)$ satisfies (6.35)). Then, we have from the proof of Theorem 4.2 that

$$T_{qw}(z) = T_{qv}(z) + 1 = \\ N(z)\bar{Y}(z) + 1 - N(z)D(z)\left(\bar{\xi}_c^\varepsilon(z)\bar{\xi}_p^\varepsilon(z)\left(\bar{\xi}_c(z)\bar{\xi}_p(z)\right)^{-1} Q_{\gamma_{\inf,C}}(z)\right), \quad (6.36)$$

where

$$Q_{\gamma_{\inf,C}}(z) \triangleq (\xi_c(z)\xi_p(z)N(z)D(z))^{-1}\left(\left\{[H(z)]_{\mathcal{H}_2^\perp}\right\}\Big|_{z=0} + [H(z)]_{\mathcal{H}_2}\right). \quad (6.37)$$

Given the fact that the plant is strictly proper and has no strictly-NMP zeros, we have $\xi_c(z) = z$. Therefore, (2.43) implies that $H(z)$ in (6.31) satisfies

$$H(z) = z\xi_p(z)N(z)\bar{Y}(z) = z(\xi_p(\infty) - \xi_p(z)) + z(\xi_p(z)X_i(z)D(z) - \xi_p(\infty)). \quad (6.38)$$

Since $G(z)$ is strictly proper we also know from (2.43) that

$$X_i(\infty)D(\infty) - 1 = \bar{Y}(\infty)N(\infty) = 0. \quad (6.39)$$

Thus, a straightforward manipulation shows that (see (6.37))

$$Q_{\gamma_{\inf},C}(z) = (\xi_p(z)N(z)D(z))^{-1}(\xi_p(z)X_i(z)D(z) - \xi_p(\infty)) \quad (6.40)$$

and, accordingly, (6.36) can be written as

$$T_{qw}(z) = X_i(z)D(z)\left(1 - \bar{\xi}_c^\varepsilon(z)\bar{\xi}_p^\varepsilon(z)\left(\bar{\xi}_c(z)\bar{\xi}_p(z)\right)^{-1}\right) + \\ \bar{\xi}_c^\varepsilon(z)\bar{\xi}_p^\varepsilon(z)\left(\bar{\xi}_c(z)\bar{\xi}_p(z)\right)^{-1}\xi_p(z)^{-1}\xi_p(\infty). \quad (6.41)$$

Clearly, $\lim_{\varepsilon\to 0^+}|T_{qw}(z)|^2 = |\xi_p(\infty)|^2$. Therefore, (6.33) and (6.34) follow.

- ($\Rightarrow$) Define

$$T_{qv}^\varepsilon(z) \triangleq T_{qv}(z)|_{K(z)=K_\varepsilon(z)}, \quad T_{qw}^\varepsilon(z) \triangleq T_{qw}(z)|_{K(z)=K_\varepsilon(z)}. \quad (6.42)$$

Since, by assumption, $K_\varepsilon(z) \in \mathcal{S}_C$ for every $\varepsilon \in (0, 1]$, it follows that (6.35) and (6.33) are equivalent to

$$\lim_{\varepsilon\to 0^+}\left|\left|T_{qv}^\varepsilon(z)\right|\right|_2^2 = \gamma_{\inf,C} \quad (6.43)$$

and

$$\lim_{\varepsilon\to 0^+} \left|T^{\varepsilon}_{qw}(e^{j\omega})\right|^2 = \kappa, \quad \forall\omega \in (0,1], \tag{6.44}$$

respectively. Since the plant is strictly proper, it follows from (6.44) that, necessarily,[10]

$$T^{\varepsilon}_{qw}(z) = \left(\prod_{i=1}^{n_p+n_2+\bar{n}} -p_i\right)\left(\prod_{i=1}^{n_p+n_2} \frac{z-p_i}{1-zp_i}\right)\left(\prod_{i=n_p+n_2+1}^{n_p+n_2+\bar{n}} \frac{z-p_i}{(1-\varepsilon)-zp_i}\right), \tag{6.45}$$

where $\{p_1,\cdots,p_{n_p}\}_m$ is the multi-set of strictly unstable poles of the plant, $\{p_{n_p+1},\cdots,p_{n_p+n_2}\}_m$ is the multi-set of strictly unstable poles of $F_2(z)M_2(z)$ and $\{p_{n_p+n_2+1},\cdots,p_{n_p+n_2+\bar{n}}\}_m$ is the multi-set of marginally stable poles of $G(z)$ and $F_2(z)M_2(z)$. Use of the Residue Theorem in (6.45) yields (see, e.g., Appendix A in [155] and also the proof of Theorem 4.1)

$$\lim_{\varepsilon\to 0^+} \left|\left|T^{\varepsilon}_{qv}(z)\right|\right|_2^2 = \left(\prod_{i=1}^{n_p+n_2} |p_i|^2\right) - 1. \tag{6.46}$$

Since $K_\varepsilon(z) \in \mathcal{S}$ is such that (6.43) holds, we see from (6.46) that $K_\varepsilon(z)$ is such that, in the limit as $\varepsilon \to 0^+$, the associated transfer functions $F_2(z)M_2(z)$ must have as few unstable poles, and as close to the unit circle, as possible.

Suppose that $G(z)$ is not strongly stabilizable and has (at least) one strictly-NMP zero that belongs to $\mathcal{I}$. The previous reasoning implies that $K_\varepsilon(z)$ is such that, in the limit, the associated $F_2(z)M_2(z)$ cancels (at least) one strictly-NMP plant zero. But this contradicts the fact that, as the proof of Theorem 4.1 reveals, the optimizing $F_2(z)M_2(z)$'s are such that, in the limit as $\varepsilon \to 0^+$, at most marginally-MP zeros of the plant model are cancelled. We thus conclude that for not strongly stabilizable plants that have (at least) one strictly-NMP zero that belongs to $\mathcal{I}$, it is impossible to satisfy both (6.35) and (6.33).

We now consider plant models that are strongly stabilizable, or that are such that all their zeros in $\mathcal{I}$ lie on the unit circle. In these cases, $F_2(z)M_2(z)$ can be chosen stable except for some unstable poles that, as $\varepsilon \to 0^+$, tend to cancel the zeros of the plant model that are marginally-MP and belong to $\mathcal{I}$. Since $K_\varepsilon(z)$ satisfies (6.43), we thus conclude that (6.46) reduces to

$$\lim_{\varepsilon\to 0^+} \left|\left|T^{\varepsilon}_{qv}(z)\right|\right|_2^2 = \left(\prod_{i=1}^{n_p} |p_i|^2\right) - 1 = \gamma_{\inf,C}. \tag{6.47}$$

[10] Actually, the last factor in (6.45) may be chosen differently. However, its asymptotic behavior as $\varepsilon \to 0^+$ (which is what we will exploit) must be as in the proposed factor.

Therefore, $\Delta_G = 0$ and Theorem 4.2 implies that $G(z)$ has no strictly-NMP zeros and relative degree one. This completes the proof.

□□□

We are now ready to establish the following corollary:

Corollary 6.6 (Necessary condition for MSS (no feedback II)) *Consider the networked control situation of Figure 6.1, where the control and coding architecture is as in Figure 6.2, the blocks $\hat{\mathcal{E}}$ and $\hat{\mathcal{D}}$ form an i.i.d. source coding scheme, and Assumptions 6.1 and 6.2 hold. If $G(z)$ is unstable, $d \neq 0$, $F(z) = [0 \;\; F_2(z)]$, and the i.i.d. source coding scheme is memoryless, then*[11]

$$\bar{R} > \bar{R}^{\not m}_{\text{inf},C} \triangleq \sum_{i=1}^{n_p} \ln |p_i| + \frac{1}{2} \ln \left(1 + \frac{\Delta_G}{\prod_{i=1}^{n_p} |p_i|^2} \right) \tag{6.48}$$

is a necessary condition for MSS. Moreover, $R^{\not m}_{\text{inf},C} \geq R^{m}_{\text{inf},C}$ with equality if and only if $G(z)$ has no strictly-NMP zeros and relative degree one.

Proof: Immediate from Corollary 6.5, Parts 1 in Theorem 5.1 and Corollary 5.1, Part 2 of Theorem 5.2, and Lemma 6.1. □□□

Corollary 6.6 gives a bound that average data-rates in MSS feedback loops necessarily satisfy, when the proposed control and coding architecture, without feedback and with a memoryless i.i.d. source coding scheme, is employed. Comparing this result with Corollary 6.4, we see that, as opposed to the case where architectures with feedback are employed (see Section 6.3.1A), using a memoryless coding scheme when $F(z)$ is constrained, in general, imposes additional constraints on the achievable average data-rate. It thus follows that the complexity of using full-memory coding schemes pays off, when no feedback is used. It is worth noting that, if $G(z)$ is marginally-MP and has relative degree one, then we have the following chain of equalities (see Corollaries 6.4 and 6.6):

$$\bar{R}^{\not m}_{\text{inf},C} = \bar{R}^{m}_{\text{inf},C} = \bar{R}_{\text{inf}}. \tag{6.49}$$

That is, for such plant models neither the constraint $F(z) = \begin{bmatrix} 0 & F_2(z) \end{bmatrix}$, nor the constraint of employing a memoryless i.i.d. source coding scheme, limits the minimal achievable data-rate for

[11] In $R^{\not m}_{\text{inf},C}$, $\not m$ indicates that the source coding scheme has no memory; C is for constrained (i.e., without feedback).

MSS. (Given our previous discussion, this is also the only case in which *the two* equalities in (6.49) hold.)

We end this section with a counterpart to Theorem 6.2, which is valid when using coding schemes without feedback:

Theorem 6.4 (Relationship between different problems (no feedback)) *Consider the networked control situation of Figure 6.1, where the control and coding architecture is as in Figure 6.2, the blocks $\hat{\mathcal{E}}$ and $\hat{\mathcal{D}}$ form an i.i.d. source coding scheme, and Assumptions 6.1 and 6.2 hold. If $G(z)$ is unstable and $F(z) = \begin{bmatrix} 0 & F_2(z) \end{bmatrix}$, then the problem of finding the minimal average scalar mutual information across the coding scheme that guarantees MSS is* equivalent *to the problem of finding the minimal stationary coding scheme signal-to-noise ratio that guarantees MSS. Also, a solution to either of the above problems provides a solution to the problem of finding the minimal average directed mutual information across the coding scheme if and only if $G(z)$ is marginally-MP and has relative degree one.*

Proof: Immediate from Part 1 in Corollary 5.1, Lemma 6.1 and Corollary 6.6. □□□

Theorem 6.4 states that, as opposed to the situation considered in Section 6.3.1, studying stationary coding scheme signal-to-noise ratio for MSS (equivalently, average scalar mutual information for MSS) *does not always* give a solution to the problem of what is the minimal average directed mutual information across the coding scheme that guarantees MSS. The key impediment for this to happen is the *lack of degrees of freedom* in coding schemes without feedback (i.e., the fact that in such schemes we impose $F(z) = \begin{bmatrix} 0 & F_2(z) \end{bmatrix}$).

Remark 6.4 (Other architectures) *It is easy to see from our previous results that the following holds:*

- *If $F(z) = [0 \;\; F_2(z)]$ and $F_2(z)$ is fixed at any point in $\mathcal{U}_\infty$ satisfying the conditions in Part 2 of Corollary 4.1, then Theorems 6.3 and 6.4, and Corollaries 6.4, 6.5 and 6.6 still apply. In particular, they apply to the case $F(z) = [0 \; f]$, where $f \in \mathbb{R} \setminus \{0\}$.*
- *If $K(z)$ is constrained to yield a one-dof control architecture, i.e., $F(z) = [0 \; 1]$ and $M(z) = C(z)[1 \;\; -1]$, then Theorems 6.3 and 6.4, and Corollaries 6.4, 6.5 and 6.6 still apply.*

□□

Remark 6.5 (Channel capacity) *The results in Corollaries 6.1, 6.3, 6.4 and 6.6 give necessary conditions on the average data-rate needed to stabilize a given SISO LTI plant model, when* *specific control and coding architectures are employed. By definition of channel capacity (see, e.g., [31]), each one of these conditions can be trivially translated into a necessary condition on the capacity of the* underlying channel.[12] *If one does so, then one should keep in mind that our results consider a purely source coding perspective, where the interconnection of the channel coding scheme and the underlying channel has been replaced by an ideal digital channel. This implies that, even though the lower bounds on channel capacity that one can obtain from the above corollaries are correct, they are too optimistic in practice: achieving anything close to ideal digital communications on a given underlying channel implies the transmission at rates well below capacity. In other words, if our results require a certain average data-rate to stabilize a given system, then the underlying channel must have a capacity much higher than that rate in order to be able to safely assume that the channel coding scheme achieves near ideal transmission. We note that these comments also apply to other results in the literature that consider ideal digital channels (see, e.g., [116, 118, 175]).*

An interesting consequence of the above caveats is that, in some cases, it is preferably to use analog communications instead of digital communications. Indeed, one might assume that the underlying channel in a certain networked control situation is an additive white Gaussian noise channel with an input power constraint (see [31]; these channels have been considered in, e.g., [18, 48, 174]) and that pre- and post scaling of the channel input and output is available. The capacity of such a channel, say C_{awgn}, is given by (see, e.g., [31])

$$C_{\text{awgn}} = \frac{1}{2} \ln (1 + \gamma_{\text{awgn}}), \quad \gamma_{\text{awgn}} \triangleq \frac{\hat{\sigma}_x^2}{\sigma_n^2}, \tag{6.50}$$

where $\hat{\sigma}_x^2$ corresponds to the greatest admissible channel input power, and σ_n^2 corresponds to the channel noise variance. It is easy to see that γ_{awgn} is also equal to the signal to noise ratio of the signal-to-noise ratio constrained additive white noise channel that arises due to the scaling of the input and output of the Gaussian channel (see Section 1.1). In this situation, it is immediate to see that the results in Section 4.3 (see also [18]) allow one to identify necessary and sufficient conditions *on C_{awgn} for the existence of stabilizing control and coding schemes that make use of the Gaussian channel in an* analog *fashion (i.e., without using quantization). These conditions amount to those in Corollaries 6.3 and 6.6, where $\bar{R}$ is replaced by C_{awgn}.*[13] *On the other hand, if one uses the Gaussian channel in a digital fashion, i.e., one decides to use a source coding scheme and a channel coding scheme designed using Shannon's separation theorem, then, and as mentioned above, the conditions in the corollaries translate into only necessary conditions on the*

[12] Recall Section 5.1.

[13] Which result actually applies depends on the considered control and coding architecture.

underlying channel capacities that guarantee stability. Viewed from this perspective, it follows that, in these cases, it is preferable not to code (see also [51]). □□

6.3.2 Guaranteed upper bounds on average data-rates for mean square stability

In Section 6.3.1 we established necessary lower bounds on average data-rates that allow one to guarantee MSS, provided Assumption 6.2 holds. However, we have not provided any proof of the achievability of these bounds, nor extensions to the non-Gaussian case. In this section we present results that address these issues.

It is well known that there exist coding schemes (time varying and with memory) that allow one to achieve the bound identified for the case of control and coding architectures that use feedback (see, e.g., [116,175]). Nevertheless, it is not clear whether the coding schemes proposed in [116, 175] can be interpreted as a specific instance of the coding scheme proposed in this chapter (recall Section 6.2). This justifies the study of guaranteed upper bounds on the average data-rate which is compatible with MSS in the proposed control and coding architecture.

Throughout this section we will relax Assumption 6.2 and return to assumptions that are aligned with those in Chapters 3 and 4, namely:

Assumption 6.3 (Signals and initial states)

(a) Both r and d are second order mutually uncorrelated wss scalar processes with rational PSDs. At least one signal, r or d, is non-zero. If r or d are non-zero, then they admit spectral factors in $\mathcal{U}_\infty$.

(b) The initial states of all LTI filters in Figure 6.3 (including plant) are jointly second order random variables.

□□

We start by considering the case of coding schemes with feedback:

Theorem 6.5 *Consider the networked control situation of Figure 6.1, where the control and coding architecture is as in Figure 6.2 and the blocks $\hat{\mathcal{E}}$ and $\hat{\mathcal{D}}$ form an ECDQ with quantization step $\Delta < \infty$ (see Section 5.5). Suppose that Assumptions 6.1 and 6.3 hold, and that the dither signal d_h is an independent sequence of i.i.d. random variables, uniformly distributed over $\left(-\frac{\Delta}{2}, \frac{\Delta}{2}\right)$. If the coding scheme uses feedback, $d \neq 0$ and $K(z)$ is chosen as in Corollary 6.2,* *then, irrespective of whether the ECDQ has memory or not, it is possible to choose the EC and the ED in the ECDQ in a way such that MSS is achieved and*

$$\bar{R} \leq \sum_{i=1}^{n_p} \ln|p_i| + \frac{1}{2}\ln\left(\frac{2\pi e}{12}\right) + \ln 2 + F_u(\Delta, \varepsilon), \tag{6.51}$$

where $F_u : \mathbb{R}^+ \times (0,1] \to \mathbb{R}^+$ *is a continuous function such that*

$$\lim_{\substack{\Delta\to\infty \\ \varepsilon\to 0^+}} F_u(\Delta, \varepsilon) = 0. \tag{6.52}$$

Proof: We will prove the result for the case of coding schemes with a memoryless ECDQ. The result for the case that uses an ECDQ with memory follows immediately from the argument below and the proof of Theorem 6.2.

The choice of d_h is such that the ECDQ is appropriately dithered. Therefore, the corresponding quantization noise is independent and i.i.d. (see Theorem 5.3) and, accordingly, as long as $\Delta < \infty$ and $K(z) \in \mathcal{S}$ (which happens whenever $\varepsilon \in (0,1]$), MSS is guaranteed.

Assumptions 6.1 and 6.3, and the fact that $K(z) \in \mathcal{S}$ imply that Assumption 5.2 holds in our case. Since, in addition, the dither has been properly chosen, we can use Corollary 5.5 to conclude that

$$\bar{R} \leq \frac{1}{2}\ln(1+\gamma) + \frac{1}{2}\ln\left(\frac{2\pi e}{12}\right) + \ln 2. \tag{6.53}$$

On the other hand, the choice for $K(z)$ in Corollary 6.2 is such that $\gamma \to \gamma_{\text{inf}}$ if both $\varepsilon \to 0^+$ and $\sigma_q^2 \to \infty \Leftrightarrow \Delta \to \infty$ (recall that in an appropriately dithered ECDQ, we have $\sigma_q^2 = \frac{\Delta^2}{12}$; see Theorem 5.3). As a consequence, (6.53) can be written as in (6.51) and the result follows. □□□

Remark 6.6 *If the* $\mathcal{K}_\varepsilon\{\cdot\}$ *operator in the characterization of* $K(z)$ *in Corollary 6.2 is redundant, then* $F_u(\cdot,\cdot)$ *becomes a function of* Δ *only.* □□

Theorem 6.5 is one of the main results in this thesis. It establishes guaranteed upper bounds on the average data-rates that guarantee MSS and that are achievable with the proposed control and coding architecture, when no constraints on $F(z)$ are imposed. In particular we see that, for the proposed choice of control and coding scheme parameters, and if the quantization step Δ is chosen sufficiently large (but finite), then one will be able to achieve MSS at an average data-rate that will be arbitrarily close to the absolute bound identified in [116, 174] plus $\frac{1}{2}\ln\left(\frac{2\pi e}{12}\right) + \ln 2$ nats (i.e., 1.255 bits). We feel that this extra rate is a fair price to be paid if one constrains oneself to the conceptually simple coding schemes considered in this thesis. Another interesting point is that, in the situation under study, ECDQs with memory do not provide any advantage over memoryless ones (see also the comments made after Corollary 6.3).

We now turn to the study of control and coding schemes without feedback:

Theorem 6.6 *Consider the networked control situation of Figure 6.1, where the encoder, controller and decoder are as in Figure 6.2 and the blocks $\hat{\mathcal{E}}$ and $\hat{\mathcal{D}}$ form an ECDQ with quantization step $\Delta < \infty$. Suppose that Assumptions 6.1 and 6.3 hold, and that the dither signal d_h is an independent sequence of i.i.d. random variables, uniformly distributed over $\left(-\frac{\Delta}{2}, \frac{\Delta}{2}\right)$. If the coding scheme uses no feedback and $d \neq 0$, then:*

1. *If $K(z)$ is chosen as in Theorem 6.3 and the ECDQ has memory, then it is possible to choose the EC and the ED in the ECDQ in a way such that MSS is achieved and*

$$\bar{R} \leq \sum_{i=1}^{n_p} \ln |p_i| + \sum_{\substack{i=1 \\ c_i \in \mathcal{I}}}^{n_c} \ln |c_i| + \frac{1}{2} \ln \left(\frac{2\pi e}{12} \right) + \ln 2 + F_c^m(\Delta, \varepsilon), \tag{6.54}$$

where $F_c^m : \mathbb{R}^+ \times (0, 1] \to \mathbb{R}^+$ is a continuous function such that

$$\lim_{\substack{\Delta \to \infty \\ \varepsilon \to 0^+}} F_c^m(\Delta, \varepsilon) = 0. \tag{6.55}$$

2. *If $K(z)$ is chosen as in Corollary 6.5 and the ECDQ has no memory, then it is possible to choose the EC and the ED in the ECDQ in a way such that MSS is achieved and*

$$\bar{R} \leq \sum_{i=1}^{n_p} \ln |p_i| + \frac{1}{2} \ln \left(1 + \frac{\Delta_G}{\prod_{i=1}^{n_p} |p_i|^2} \right) + \frac{1}{2} \ln \left(\frac{2\pi e}{12} \right) + \ln 2 + F_c^{\not m}(\Delta, \varepsilon), \tag{6.56}$$

where $F_c^{\not m} : \mathbb{R}^+ \times (0, 1] \to \mathbb{R}^+$ is a continuous function such that

$$\lim_{\substack{\Delta \to \infty \\ \varepsilon \to 0^+}} F_c^{\not m}(\Delta, \varepsilon) = 0. \tag{6.57}$$

Proof: The proof is analogous to the proof of Theorem 6.5. □□□

Remark 6.7 *We note that, if $G(z)$ is strongly stabilizable, then $F_c^m(\cdot, \cdot)$ becomes a function of Δ only. Also, if the $\mathcal{K}_\varepsilon \{\cdot\}$ operator in the characterization of $K(z)$ in Corollary 6.5 is redundant, then $F_c^{\not m}(\cdot, \cdot)$ becomes a function of Δ only (see, e.g., [59]).* □□

Remark 6.8 (Bounds are conservative) *As already mentioned in the paragraph after Remark 5.10, the upper bounds provided by Theorems 6.5 and 6.6 are conservative. Indeed, the* *actual average data-rates that are achievable when the proposed coding schemes are designed as suggested, may be close to the expressions in Theorems 6.5 and 6.6 with the term* $\ln 2$ *omitted.* □□

Theorem 6.6 is the counterpart of Theorem 6.5 when the coding scheme uses no feedback. As was the case in Theorem 6.5, we see that the necessary conditions on average data-rates for MSS developed for the Gaussian case in Section 6.3.1, provide a fairly accurate indication of the actual data-rates that can be achieved using a properly designed control and coding architecture. In particular, we have shown that use of properly designed ECDQs can make the gap as close to $\frac{1}{2}\ln\left(\frac{2\pi e}{12}\right) + \ln 2$ nats (i.e., 1.255 bits) as desired, provided the quantization step is sufficiently large.

Remark 6.9 (Previous work (I)) *To the best of the author's knowledge, the earlier results that are most closely related to the results in this chapter are those reported in [180]. In that work the authors give a sufficient condition on the average data-rate that guarantees the existence of a BIBO stabilizing LTI controller for an LTI SISO plant model. The authors of [180] focus on a one-dof architecture and make use of two assumptions that are not needed in our results: firstly, they assume that the quantization step is small (and it is not clear whether their results can be extended to the general quantization step case [180]). Secondly, they assume the existence of block entropy coders that allow one to attain average data-rates arbitrarily close to the entropy of the output of the quantizer. This assumption is, in principle, sensible, but disregards the fact that block entropy coders may incur arbitrarily long delays (see, e.g., [31]). The results in [180] are for arbitrarily complex coding schemes and no attention is given to simpler architectures. Hence, it is not surprising that the sufficient condition on the average data-rate that guarantees the existence of a stabilizing controller given in [180] turns out to be essentially equal to the condition given by Part 1 of Theorem 6.6. A detailed comparison between our results and those in [180] is left for future work.* □□

Remark 6.10 (Previous work (II)) *The results in [175] guarantee that, when using memoryless encoders (referred to as Type-II encoders in [175]), asymptotic stability can be achieved with a bounded (but otherwise unknown) average data-rate. Even though our results focus on a different stability notion (namely MSS), they give a precise upper bound on the average data-rates that guarantee MSS and that can be achieved using simple building blocks without memory. (We also characterized the plant models for which use of memory does not provide any advantage.) Thus, our results shed light onto the open question of what is the smallest average data-rate* *that guarantees stability (more precisely, MSS) and which is achievable with memoryless (but otherwise arbitrarily complex) coding schemes.* □□

6.4 Performance Issues

We now turn to the question of performance. As foreshadowed at the end of Section 6.2, we will go beyond stability issues and examine average data-rate requirements to achieve a certain level of performance. To that end we define the tracking error e via

$$e \triangleq r - y, \tag{6.58}$$

and use its stationary variance σ_e^2 as performance measure (see Figure 6.1).

We will constrain ourselves to the case of memoryless i.i.d. source coding schemes. For this class of coding schemes, the average data-rate $\bar{R}$ is lower bounded by the average scalar mutual information across the coding scheme $I_\infty(v; w)$ (see Theorem 5.1). Therefore, it makes sense to start our study with an investigation of the minimal average scalar mutual information across the coding scheme that guarantees not only MSS, but also a desired performance level. Formally, we can state the problem of interest as follows:

Problem 6.1 (Rate distortion-like problem) *Consider the networked control situation in Figure 6.1, where the coding scheme is as in Figure 6.2, the blocks $\hat{\mathcal{E}}$ and $\hat{\mathcal{D}}$ form a memoryless i.i.d. source coding scheme, and Assumptions 6.1 and 6.3 hold. Find (or prove that the problem is infeasible)*

$$I_\infty^D(v; w) \triangleq \inf_{\substack{K(z)\in\mathcal{S} \\ \sigma_q^2 \in \mathbb{R}_0^+ \\ \sigma_e^2 \leq D}} I_\infty(v; w), \tag{6.59}$$

where $D < \infty$ is the desired performance level and $\mathcal{S}$ is defined as before (see (6.3)*).* □□

Problem 6.1 resembles the standard rate distortion problem in information theory (see, e.g., Chapter 10 in [31]). There are, however, two important differences. Firstly, we restrict our attention to architectures that use only LTI filters and memoryless i.i.d. source coding schemes, whilst information theoretic results do not impose constraints on the coding scheme. Therefore, our results will only provide an upper bound on the absolute minimal average data-rate that allows one to attain the desired performance level. Secondly, provided one can find a causal memoryless i.i.d. source coding scheme that achieves (or gets close to) the bound on average data-rate that a solution to Problem 6.1 provides,[14] the well-posedness (and, in particular, the

[14] Candidates for this are ECDQs (see Section 5.5).

causality) of the resulting feedback system is guaranteed by the constraints on $K(z)$. This stands in stark contrast to the *standard* results in information theory that do not take causality into account (and yield causal solutions only if restrictive conditions are satisfied; see also [98, 119]).

In Chapter 5 (see Theorem 5.2) we showed that there exists a one-to-one correspondence between an upper bound on average scalar mutual information across an i.i.d. source coding scheme and the corresponding stationary signal-to-noise ratio. Thus, it is immediate to see that the results related to performance in Chapter 4 can be used to obtain upper bounds on $I_\infty^D(v;w)$. We recall (and slightly modify) the following definitions from Chapter 4:

$$\gamma_D \triangleq \inf_{\substack{K(z)\in\mathcal{S}\cap\mathcal{O}\\ \sigma_q^2\in\mathbb{R}_0^+\\ \sigma_e^2\leq D}} \frac{\sigma_v^2}{\sigma_q^2}, \tag{6.60}$$

and

$$\left(\sigma_D^2, K_D(z)\right) \triangleq \arg \inf_{\substack{K(z)\in\mathcal{S}\cap\mathcal{O}\\ \sigma_q^2\in\mathbb{R}_0^+\\ \sigma_e^2\leq D}} \frac{\sigma_v^2}{\sigma_q^2}, \tag{6.61}$$

where the set $\mathcal{O}$ contains transfer functions $K(z)$ in $\mathcal{S}$ with a specific structure. For example, $\mathcal{O} = \mathcal{S}_1$ if the considered control and coding architecture is a one-dof one (see Section 4.5.1), $\mathcal{O} = \mathcal{S}_C$ if one focuses on a control and coding architecture without feedback (see Section 4.3), $\mathcal{O} = \mathcal{S}$ if one focuses on a general architecture (i.e., control and coding architectures with feedback), etc..

Irrespective of the constraint set $\mathcal{O}$, the results of Chapter 4 allow one to obtain an (algorithmic) characterization of the minimal stationary coding scheme signal-to-noise ratio γ_D, the corresponding optimal quantization noise variance σ_D^2 and the optimal filter $K_D(z)$. As discussed in Chapter 4, $K_D(z)$ does not, in general, belong to $\mathcal{S}$. Nevertheless, one can always define (with the help of Part 2 of Theorem 4.4) a family of functions $\{K_D^\varepsilon(z)\}_{\varepsilon\in(0,1]}$, where $K_D^\varepsilon(z) \in \mathcal{S}$ for every $\varepsilon \in (0,1]$, such that γ_D is approximated to any desired degree of accuracy when $\varepsilon \to 0^+$, while violating the constraint on σ_e^2 by an arbitrarily small amount. We also recall from Chapter 4 that (again, the definition is slightly modified to encompass different structural constraints on $K(z)$)

$$D_{\inf} \triangleq \inf_{\substack{K(z)\in\mathcal{S}\cap\mathcal{O}\\ \sigma_q^2=0}} \sigma_e^2. \tag{6.62}$$

We start with the following corollary of previous results:

Corollary 6.7 (Bound on $I^D_\infty(v;w)$) *Consider the networked control situation in Figure 6.1, where the coding scheme is as in Figure 6.2, the blocks $\hat{\mathfrak{E}}$ and $\hat{\mathfrak{D}}$ form a memoryless i.i.d. source coding scheme, and Assumptions 6.1 and 6.3 hold. Then:*

1. *If $D < D_{\text{inf}}$, then Problem 6.1 is infeasible.*

2. *For every performance level D satisfying $D_{\text{inf}} \leq D < \infty$,*

$$I^D_\infty(v;w) \leq \frac{1}{2}\ln(1+\gamma_D) + D(q(k)||q_G(k)), \tag{6.63}$$

   where $D(q(k)||q_G(k))$ is the divergence of the coding scheme quantization noise from its Gaussian counterpart (see Definition 5.6).

3. *If, in addition, Assumption 6.3 is replaced by Parts (b) and (c) in Assumption 6.2,*[15] *then*

$$I^D_\infty(v;w) \geq \frac{1}{2}\ln(1+\gamma_D). \tag{6.64}$$

Proof:

1. Immediate from the definition of D_{inf}.

2. Immediate from Theorem 5.2 and the definitions of both $I^D_\infty(v;w)$ and γ_D.

3. Immediate from Corollary 5.1 and the definitions of $I^D_\infty(v;w)$ and γ_D.

□□□

Corollary 6.7 provides bounds on the minimal average scalar mutual information across the coding scheme that guarantees a given performance level D. This characterization is given in terms of the minimal stationary coding scheme signal-to-noise ratio required to attain the desired performance level, i.e., it is given in terms of γ_D. From a numerical perspective, it suffices to use the ideas in Chapter 4 to obtain an approximation to γ_D and, as a consequence, obtain an estimate of the bounds on $I^D_\infty(v;w)$ that Corollary 6.7 provides.

We are now in a position to state the main result of this section:

Theorem 6.7 (Upper bound on $\bar{R}$ for a given performance level D) *Consider the networked control situation in Figure 6.1, where the control and coding architecture is as in Figure 6.2, the blocks $\hat{\mathfrak{E}}$ and $\hat{\mathfrak{D}}$ form a memoryless ECDQ with quantization step $\Delta < \infty$ and dither signal d_h which is an independent sequence of i.i.d. random variables, uniformly distributed over*

[15] Of course, the results in Part 2 still hold in this case.

$\left(-\frac{\Delta}{2}, \frac{\Delta}{2}\right)$. *Also suppose that Assumptions 6.1 and 6.3 hold and that the desired performance level* D *satisfies* $D_{\inf} \leq D < \infty$. *If* $K(z) = K_D^{\varepsilon}(z)$, $\varepsilon \in (0, 1]$ *(see paragraph before (6.62)), and the quantization step in the ECDQ is chosen as* $\Delta = \sqrt{12\sigma_D^2}$, *then it is possible to choose the EC and the ED in the ECDQ in a way such that MSS is achieved, and*

$$\sigma_e^2 = D + F_e(\varepsilon), \tag{6.65}$$

$$\bar{R} \leq \frac{1}{2}\ln(1 + \gamma_D) + \frac{1}{2}\ln\left(\frac{2\pi e}{12}\right) + \ln 2 + F_{\bar{R}}(\varepsilon), \tag{6.66}$$

where $F_e : (0, 1] \to \mathbb{R}^+$ *and* $F_{\bar{R}} : (0, 1] \to \mathbb{R}^+$ *are continuous function such that*

$$\lim_{\varepsilon \to 0^+} F_{\bar{R}}(\varepsilon) = \lim_{\varepsilon \to 0^+} F_e(\varepsilon) = 0. \tag{6.67}$$

Moreover, if $K_D^{\varepsilon}(z) = K_D(z)$, *then* $F_{\bar{R}}(\varepsilon) = F_e(\varepsilon) = 0$ *for every* ε.

Proof: The choice of d_h is such that the ECDQ is appropriately dithered. Therefore, the corresponding quantization noise is independent and i.i.d. (see Theorem 5.3) and, accordingly, as long as $\Delta < \infty$ and $K(z) \in \mathcal{S}$, MSS is guaranteed. By construction, the choice $K(z) = K_D^{\varepsilon}(z)$ belongs to $\mathcal{S}$ for any $\varepsilon \in (0, 1]$.

Assumptions 6.1 and 6.3, and the fact that $K(z) \in \mathcal{S}$ imply that Assumption 5.2 holds in our case. Since, in addition, the dither has been properly chosen, we can use Corollary 5.5 to conclude that

$$\bar{R} \leq \frac{1}{2}\ln(1 + \gamma) + \frac{1}{2}\ln\left(\frac{2\pi e}{12}\right) + \ln 2. \tag{6.68}$$

On the other hand, in an appropriately dithered ECDQ, we have $\sigma_q^2 = \frac{\Delta^2}{12}$ (see Theorem 5.3) and, by Lemma 4.3 (or its trivial extension to other cases), we know that $\sigma_D^2 < \infty$ for every finite D. By construction, the suggested choice for $K(z)$ is such that $\gamma \to \gamma_D$ if both $\varepsilon \to 0^+$ and $\Delta = \sqrt{12\sigma_D^2}$. As a consequence, (6.68) can be written as in (6.66) and the result follows.
□□□

Theorem 6.7 is one of the main results of this thesis. It provides an analytic (and easily computable) expression for an upper bound on the average channel data-rate needed to attain a specified performance level, when the proposed control and coding architecture is used, and the underlying coding scheme is memoryless. The bound is given in terms of the minimal coding scheme signal-to-noise ratio needed to achieve the desired performance level, plus a second term that arises due to the fact that the quantization noise in an ECDQ is uniform and not Gaussian, a third term that originates in the inefficiency of practical ECs, and a fourth term that, if non-zero, can be always made arbitrarily small by means of choosing a sufficiently small (but non-zero) value for the parameter ε. We note that, by construction, the bound in Theorem 6.7 is achievable. Also, the bound is pessimistic in the same sense as the bounds in Section 6.3.2 (see Remark 6.8).

To compute the bound in Theorem 6.7 we have focused on control architectures that are constrained to be formed by LTI filters and memoryless i.i.d. source coding scheme only. Of course, our bound also applies when more general architectures are employed. In those cases, however, the bound provided by Theorem 6.7 may be very conservative.

Remark 6.11 (Best performance for a given average data-rate) *It is straightforward to use the results of Chapter 4 to extend the results in this section and establish upper bounds on the best performance that is achievable with a given average data-rate constraint.* □□

6.5 An Example

To illustrate the results of this chapter, we will consider a one-dof architecture where the feedback path comprises an ideal digital channel. Assuming that the plant, and reference and disturbance signals, are as in the example given in Section 4.6, i.e., assuming that $d = 0$,

$$G(z) = \frac{z - 0.8}{z(z-2)}, \qquad \Omega_r(z) = \frac{0.1z}{z - 0.9}, \tag{6.69}$$

we will study average data-rate requirements for MSS and a specified level of performance.

The results in this chapter imply that, under Gaussianity assumptions and irrespective of whether or not the source coding scheme has memory, the minimal average data-rate $\bar{R}$ that is compatible with MSS is lower bounded in this case by

$$\bar{R} > \bar{R}_{\text{inf}} = \ln 2 = 0.693 \ \text{[nat]} \quad \text{(i.e., 1 [bit])}. \tag{6.70}$$

On the other hand, if one uses the controller

$$C(z) = \frac{1.5z}{(z - 0.8)} \tag{6.71}$$

and an appropriately designed ECDQ as source encoder, our results guarantee that one can design the EC and the ED inside the ECDQ so as to achieve MSS at an average data-rate that is upper bounded by

$$\begin{aligned}\bar{R} &\leq \ln 2 + \frac{1}{2}\ln\left(\frac{2\pi e}{12}\right) + \ln 2 + F(\Delta) \\ &= 1.563 + F(\Delta) \ \text{[nat]} \quad \text{i.e., } (2.255 + \frac{1}{\ln 2}F(\Delta) \ \text{[bit]}),\end{aligned} \tag{6.72}$$

where Δ is the ECDQ quantization step and $F(\Delta) \to 0$ when $\Delta \to \infty$. (We note that this results holds irrespective of the distribution of the initial states and reference, and irrespective of whether or not the ECDQ has memory.)

Figure 6.4 shows simulated results that confirm the validity of our theoretical analysis. The upper plot in Figure 6.4 shows the theoretical and measured[16] ECDQ signal-to-noise ratio γ as a function of Δ. As expected from the results of Chapter 4, $\gamma \to \gamma_{\text{inf}} = 3$ when $\Delta \to \infty$. It is worth noting that γ is within 1.75% of γ_{inf} for all $\Delta \geq 10$. The lower plot in Figure 6.4 shows both the measured average length of the words generated by the EC[17] and the upper bound in (6.72) as functions of Δ. The results show that our upper bound is, in this case, somewhat conservative. It is interesting to note that, for $\Delta \geq 10$, the measured average data-rate is within 0.35% of the limiting value achieved when $\Delta \to \infty$. That limiting value equals $\bar{R} = 1.623$ [bit], whilst the theoretical upper bound tends to $\bar{R} = 2.255$ [bit] when $\Delta \to \infty$.

[16] All simulation results reported in this section are averages over 100 simulations, each one 10^4 samples long, and considering different realizations for the initial state and reference signal.

[17] We used a memoryless EC with Huffman coding; see, e.g., [31].

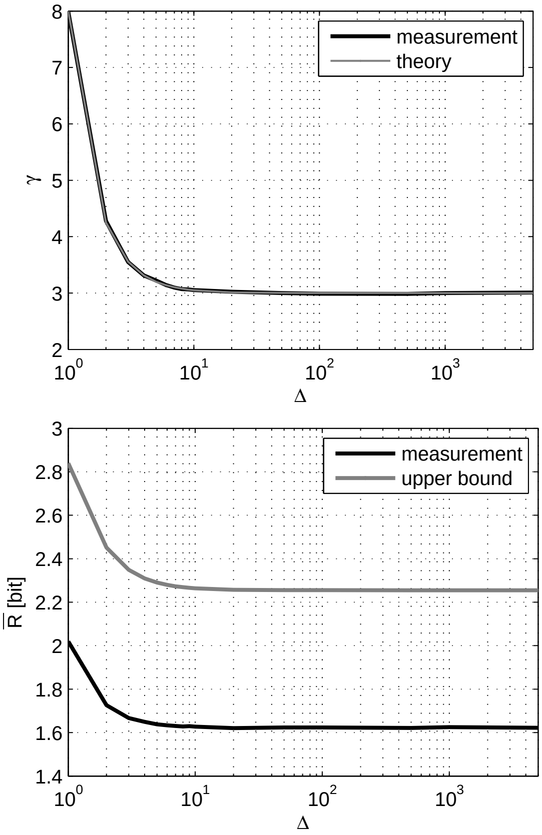

Figure 6.4: Top: Measured and theoretical ECDQ signal-to-noise ratio as a function the quantization step Δ. Bottom: Measured average length of the binary words generated by the EC and theoretical upper bound on the average data-rate as functions of Δ.

We now explore average data-rate requirements for a specified level of performance. Recall from Section 4.6 that, for the present plant and control architecture, the best non-networked performance D_{inf} satisfies

$$D_{\text{inf}} = 0.04. \tag{6.73}$$

Using the results in Chapter 4, we explored the minimal ECDQ signal-to-noise ratios that guarantee performance levels $D \in [1.01 \cdot D_{\text{inf}}, 10^3 \cdot D_{\text{inf}}] = [4.040 \cdot 10^{-2}, 40]$. Then, we used the results in the present chapter to characterize an upper bound on the minimal average data-rate $\bar{R}$ needed to achieve the desired performance levels. Both analytic and simulation results are shown in Figure 6.5. As expected, $\gamma \to \infty$ when $D \to D_{\text{inf}}$, whilst $\gamma \to \gamma_{\text{inf}} = 3$ when $D \to \infty$ (see vertical and horizontal lines in the upper plot in Figure 6.5, respectively). On the other

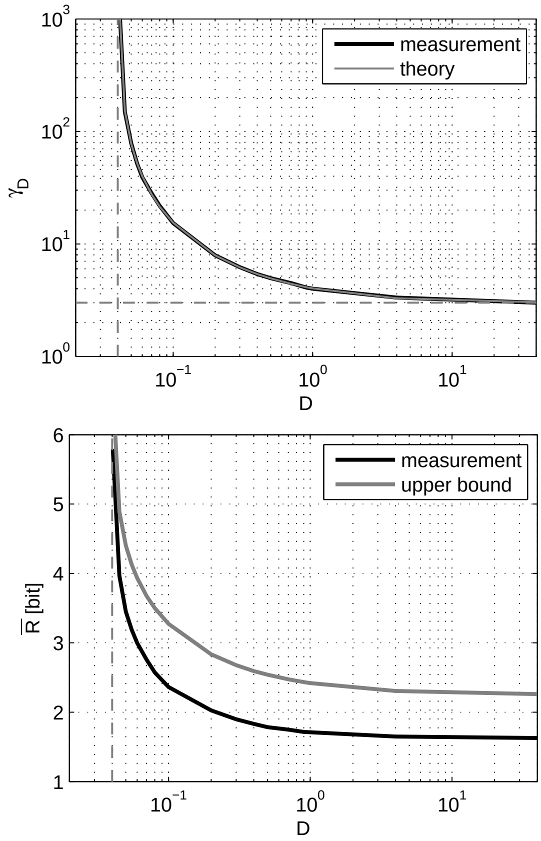

Figure 6.5: Top: Measured ECDQ signal-to-noise ratio and corresponding optimal signal-to-noise ratio as a function the desired performance level D. Bottom: Measured average length of the binary words generated by the EC, and theoretical upper bound on the optimal average data-rate for performance level D, as functions of D.

hand, we see in the lower plot in Figure 6.5 that, as was the case of the situation studied in Figure 6.4, our theoretical upper bound is pessimistic. The results show that $\bar{R} \to \infty$ when $D \to D_{\text{inf}}$ and that the measured average data-rate is such that $\bar{R} \to 1.627$ [bit] $\geq \bar{R}_{\text{inf}} = 1$ [bit] when $D \to \infty$. All this is consistent with our results and also with intuition.

6.6 Summary

This chapter has studied control systems subject to average data-rate limits. We have constrained ourselves to coding schemes that employ LTI filters and i.i.d. source coding schemes. With these building blocks, we have shown that it is possible to achieve average data-rates that guarantee MSS and that are *close* to the absolute lower bound on the data-rate for MSS identified in [116]. We have also given a precise characterization of the additional data-rate that our proposal incurs. That gap is given by the sum of two terms: one due to the divergence of the distribution of quantization noise from Gaussianity, and the other due to the inefficiency of the loss-less coding scheme employed to generate the channel symbols. We also extended our framework to the case of performance related questions. In that case, our results are not in a closed form, but lead to a simple characterization of the smallest average data-rate that allows one to attain a specific level of performance.

We would also like to highlight the fact that, in our results, the insights gained when establishing necessary conditions on the data-rates which, for example, guarantee MSS, are crucial when constructing actual coding schemes that achieve data-rates close to those necessary bounds. This stands in stark contrast to the approach in, e.g., [116].

Future work could focus on the extension of the results in this chapter to the multivariate case, and to cases with multiple channels. It is also worth noting that our results related to performance were derived for memoryless i.i.d. source coders only. The consideration of full-memory source coders is also open for future research. Another interesting open question is that of exploring whether the bounds given by our results are representative of the limits that arise when arbitrarily complex control and coding architectures are considered. We have shown that, if one uses a simple architecture with enough degrees of freedom, then the guaranteed bounds provided by our results are reasonably close to the absolute lower bounds identified in [116]. The analogous statement for cases where the information available to the encoders and decoders is limited, or where the question is related to performance, requires additional effort.

Chapter 7

Control with Data-Dropouts

7.1 Introduction

This chapter studies networked control systems closed over analog i.i.d. erasure channels. In the spirit of the average data-rate limited case considered in Chapter 6, we focus on establishing links between networked control systems that use erasure channels and networked situations that use signal-to-noise ratio constrained additive i.i.d. noise channels. One of the main results of this chapter is to show that there exists a *second-order moments equivalence* between both cases. This result is proved for a general networked control situation and, as such, extends the results in [100], where a special architecture was studied and no explicit mention to signal-to-noise ratio constrained channel was made.

We use the aforementioned equivalence to study conditions on the successful transmission probability that allow one to design a general LTI control and coding architecture that guarantees MSS. In our results, and as expected from the results in [78, 148], the existence or absence of acknowledgements that testify successful transmissions plays an important role. Interestingly, our results show that the bounds on dropout probabilities compatible with MSS obtained in [148], where time-varying policies were considered, are still valid in the LTI case. Our results go beyond the results in [100], where no detailed study of the interplay between MSS and dropout probability is performed. In [41], the author also studies minimal successful transmission probabilities for MSS, but focuses on the noiseless static state feedback case. Our results are valid in the dynamic output feedback case, and take arbitrary noise sources into account.

This chapter also addresses performance-related questions. The equivalence between i.i.d. erasure channels and signal-to-noise ratio constrained additive i.i.d. noise channels allows one to immediately use the results in Chapter 4 to design LTI control and coding architectures when data-dropouts are present. For brevity, we focus on the simplest situation where a one-dof controller has to be designed. However, other cases follow *mutatis mutandis*.

The remainder of this chapter is organized as follows: Section 7.2 presents background on Markov jump linear systems. These results are instrumental in developing the results of this chapter. Section 7.3 presents the equivalence between analog i.i.d. erasure channels and signal-to-noise ratio constrained additive i.i.d. noise channels. In Section 7.4 we exploit the results of Section 7.3 to study the interplay between channel dropout probability and MSS and performance of networked control systems that use analog i.i.d. erasure channels. Section 7.5 presents a simple example. We draw conclusions in Section 7.6.

7.2 Background on Markov Jump Linear Systems

In this section we will focus on Markov jump linear systems (MJLSs) of the form

$$x(k+1) = A(d_r(k))x(k) + B(d_r(k))\bar{d}(k), \quad k \in \mathbb{N}_0, \quad x(0) = x_o, \quad d_r(0) = d_o, \tag{7.1}$$

where x ($x(k) \in \mathbb{R}^{n_x}$) is the state vector, $\bar{d}$ ($\bar{d}(k) \in \mathbb{R}^{n_{\bar{d}}}$) is a random process, d_r is a scalar binary random process taking values in $\{0, 1\}$, and $A(i)$, $B(i)$, $i \in \{0, 1\}$, are constant matrices of appropriate dimensions. We will assume that the following holds:

Assumption 7.1 (Assumptions for MJLS analysis)

(a) The signal $\bar{d}$ is an independent second order wss process having rational PSD $S_{\bar{d}}(e^{j\omega})$.

(b) The initial state x_o is a second order random variable.

(c) The process d_r is an independent sequence of i.i.d. Bernoulli random variables with parameter $1-p$, $p \in [0, 1]$ (i.e., $\mathcal{P}\{d_r(k) = 1\} = 1 - p$, $\forall k \in \mathbb{N}_0$).

□□

Since the state space description of the system in (7.1) includes stochastically varying matrices, standard notions of stability for LTI systems are not applicable. Hence, as in, e.g., [29,30,84], we will adopt the notion of mean square stability (MSS) as the stability criterion. In the context of MJLSs, this notion is the natural extension of the notion of MSS for LTI systems introduced in Chapter 2.

Definition 7.1 (MSS and AWSS [29, 30])

1. *The system in* (7.1) *is said to be* mean square stable *(MSS) if and only if, for every x_o, d_o and $\bar{d}$ satisfying Assumption 7.1, there exist $\mu_x \in \mathbb{R}^{n_x}$ and $M_x \in \mathbb{R}^{n_x \times n_x}$, $M_x \geq 0$, both not depending on x_o or d_o, such that*

$$\lim_{k\to\infty} \mathcal{E}\left\{x(k)\right\} = \mu_x, \quad \lim_{k\to\infty} \mathcal{E}\left\{x(k)x(k)^T\right\} = M_x. \tag{7.2}$$

2. *The system in* (7.1) *is said to be* asymptotically wide sense stationary *(AWSS) if and only if, for every x_o, d_o and $\bar{d}$ satisfying Assumption 7.1, there exist $\mu_x \in \mathbb{R}^{n_x}$ and $\{Q_x(\tau) \in \mathbb{R}^{n_x \times n_x} : \tau \in \mathbb{N}_0\}$, both not depending on x_o or d_o, such that*

$$\lim_{k\to\infty} \mathcal{E}\left\{x(k)\right\} = \mu_x, \quad \lim_{k\to\infty} \mathcal{E}\left\{x(k+\tau)x(k)^T\right\} = Q_x(\tau). \tag{7.3}$$

□□

In other words, Definition 7.1 states that the system in (7.1) is MSS if and only if its state x has well defined and finite stationary mean and stationary second order moments matrix. On the other hand, system (7.1) is AWSS if and only if its state x is an asymptotically wss process. The next theorem, adapted from [29, 30], presents necessary and sufficient conditions for both MSS and AWSS:

Theorem 7.1 (MSS and AWSS) *Consider the system in* (7.1) *and suppose that Assumption 7.1 holds. Define two auxiliary MJLSs, S_1 and S_2, as follows:*

$$S_1: \quad x_1(k+1) = A(d_r(k))x_1(k), \quad x_1(0) = x_o, \quad d_r(0) = d_o \tag{7.4}$$

$$S_2: \quad x_2(k+1) = A(d_r(k))^H x_2(k), \quad x_2(0) = x_o, \quad d_r(0) = d_o, \tag{7.5}$$

where x_o, d_r and $A(\cdot)$ are as in (7.1)*. Then, the following statements are equivalent:*

1. *The system in* (7.1) *is MSS.*
2. *The system in* (7.1) *is AWSS.*
3. *The auxiliary system S_1 is MSS.*
4. *The auxiliary system S_2 is MSS.*
5. *There exists $P > 0$ such that*

$$P - pA(0)^H PA(0) - (1-p)A(1)^H PA(1) > 0. \tag{7.6}$$

6. *For every $M > 0$ there exists a unique $P > 0$ such that*

$$P - pA(0)^H PA(0) - (1-p)A(1)^H PA(1) = M. \tag{7.7}$$

Proof: If $p \in (0,1)$, then the Markov chain describing the Bernoulli process d_r is ergodic (see, e.g., [62]). Therefore, the results in Chapter 3 in [30] apply and the equivalence between Statements 1,2,3,5 and 6 follows (see also [29]). To see that Statement 4 is equivalent to all the others, we note that use of Statement 6 in system S_2 allows one to conclude that system S_2 is MSS if and only if, for every $M_2 > 0$, there exists an unique $P_2 > 0$ such that

$$P_2 - pA(0)P_2A(0)^H - (1-p)A(1)P_2A(1)^H = M_2. \tag{7.8}$$

Equation (7.8) is equivalent to

$$H\text{vec}\left\{P_2\right\} = \text{vec}\left\{M_2\right\}, \tag{7.9}$$

where[1]

$$H \triangleq \left(I - p\bar{A}(0) \otimes A(0) - (1-p)\bar{A}(1) \otimes A(1)\right). \tag{7.10}$$

In a similar way, the corresponding Lyapunov-like equation for System S_1 may be written as

$$H^H\text{vec}\left\{P_1\right\} = \text{vec}\left\{M_1\right\}, \tag{7.11}$$

where P_1 and M_1 play the role of P and M in (7.7), respectively. It is then immediate to see that there exists a unique solution for (7.8) (i.e., System S_2 is MSS) if and only there exists a unique solution to (7.11) (i.e., System S_1 is MSS). This proves our remaining claim for $p \in (0,1)$.

If $p \in \{0,1\}$, then the system in (7.1) reduces to the standard LTI system considered in Chapter 2. The result thus follows from Theorem 2.3, Corollary 2.1, and standard results on Lyapunov equations (see, e.g., Section 21.1 in [194]). □□□

An immediate corollary to Theorem 7.1 is stated next:

Corollary 7.1 (Stationary properties of linear combinations of the state) *Consider the system in* (7.1). *Define the sequence y_ℓ (here, ℓ is just an index) via*

$$y_\ell(k) \triangleq C_\ell(d_r(k))x(k) + D_\ell(d_r(k))\bar{d}(k), \tag{7.12}$$

[1] Recall the notation introduced at the beginning of this thesis.

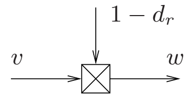

Figure 7.1: Erasure channel.

where $y_\ell(k) \in \mathbb{R}^{n_{y_\ell}}$*, and* $C_\ell(i)$ *and* $D_\ell(i)$*,* $i \in \{0,1\}$*, are constant matrices of appropriate dimensions. If system* (7.1) *is MSS and Assumption 7.1 holds, then* y_ℓ *is a second order asymptotically wss process. Moreover, any two signals* y_i *and* y_j *of the form* (7.12) *are jointly second order asymptotically wss processes.*

Proof: The result follow from a straightforward derivation that uses Definition 7.1, the assumptions on d_r and $\bar{d}$, and the results in Theorem 7.1. □□□

Theorem 7.1 and Corollary 7.1 serve as a basis for our subsequent developments.

7.3 Equivalence Between i.i.d. Data-Dropouts and Signal-to-noise Ratio Constraints

This section presents one of the main results in this chapter. We establish a connection between spectral and MSS-related properties of a specific class of networked control systems that use unreliable channels, and those of a standard LTI feedback system (subject to some constraints). We will focus on unreliable channels that can be modeled as *analog i.i.d erasure channels.* Such a channel is formally defined as follows:

Definition 7.2 (Analog i.i.d. erasure channel) *Consider the channel in Figure 7.1, where* v *is the input,* w *is the output,* d_r *is a binary process that models data dropouts (namely,* $d_r(k) \in \{0,1\}$ $\forall k \in \mathbb{N}_0$*), and*

$$w(k) \triangleq (1-d_r(k))v(k), \quad \forall k \in \mathbb{N}_0, \quad \forall v(k) \in \mathbb{R}. \tag{7.13}$$

We say that the channel is an analog i.i.d. erasure channel *if and only if* d_r *is an independent sequence of i.i.d. Bernoulli random variables with parameter* $1-p$*.* □□

Remark 7.1 *In the channel of Definition 7.2, data-dropouts at any two different time instants are assumed to be statistically independent. More general cases have been explored in, e.g., [99].* □□

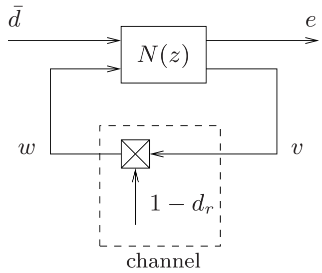

Figure 7.2: LTI system with feedback over analog i.i.d. erasure channel.

In this section we consider the general setup depicted in Figure 7.2. In that figure, $N(z) \in \mathcal{R}_p$ is a given LTI system whose transfer function has been partitioned in a way such that

$$\begin{bmatrix} e \\ v \end{bmatrix} = \begin{bmatrix} N_{11}(z) & N_{12}(z) \\ N_{21}(z) & N_{22}(z) \end{bmatrix} \begin{bmatrix} \bar{d} \\ w \end{bmatrix}, \tag{7.14}$$

and the feedback from the output v to the input w is made over an analog i.i.d erasure channel. In order to ensure (among other things) that the system in Figure 7.2 is well-posed for every $N(z)$, we will assume that the following holds:

Assumption 7.2 (Conditions on $N(z)$) *The transfer function $N(z)$ belongs to $\mathcal{R}_p$, has a stabilizable and detectable underlying realization, and $N_{22}(z) \in \mathcal{R}_{sp}^{1\times 1}$. Accordingly, we denote the realization of $N(z)$ by* $\left(A, \begin{bmatrix} B_1 & B_2 \end{bmatrix}, \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}, \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & 0 \end{bmatrix}\right)$. □□

In Assumption 7.2, we also included the condition that both v and w be scalar signals. This assumption may, in principle, be removed at the expense of additional technicalities.

Provided Assumption 7.2 holds, it is easy to see that the state of $N(z)$ in Figure 7.2, say x, evolves according to (7.1) with

$$A(d_r(k)) \triangleq A + (1 - d_r(k))\, B_2 C_2, \quad B(d_r(k)) = B_1 + (1 - d_r(k))\, B_2 D_{21}. \tag{7.15}$$

We will also consider an alternative situation where the analog i.i.d. erasure channel in Figure 7.2 is replaced by an additive noise channel plus a gain equal to the probability of successful transmission over the erasure channel, i.e., p (see Figure 7.3). In Figure 7.3, q is an exogenous noise source satisfying the following:

Assumption 7.3 (Additive noise channel) *The signal q is an independent sequence of i.i.d. random variables having zero mean and a variance σ_q^2 that satisfies*

$$\sigma_q^2 = p(1-p)\sigma_v^2, \tag{7.16}$$

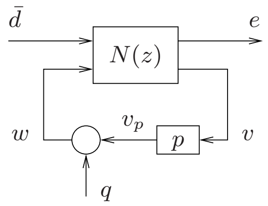

Figure 7.3: LTI system with feedback over additive i.i.d. noise channel, plus a gain.

provided the stationary variance of v exists and is finite. □□

In other words, Assumption 7.3 states that, for any given successful transmission probability p, the feedback channel in Figure 7.3 is an additive i.i.d. noise channel with an equality stationary signal-to-noise ratio constraint, plus a gain.

If Assumption 7.2 holds, then a state space description of the system of Figure 7.3 is given by (recall the notation in (7.15))

$$x(k+1) = \hat{A}x(k) + \hat{B}\bar{d}(k) + B_2 q(k), \tag{7.17}$$

where

$$\hat{A} \triangleq pA(0) + (1-p)A(1) = A + pB_2C_2, \tag{7.18}$$

$$\hat{B} \triangleq pB(0) + (1-p)B(1) = B_1 + pB_2D_{21}. \tag{7.19}$$

In the remainder of this section we will show that the LTI system of Figure 7.3 and the MJLS of Figure 7.2 are such that:

(i) The MSS of the MJLS system is equivalent to the internal stability of the LTI system and an additional inequality constraint.

(ii) If the MJLS is MSS, then the stationary PSDs of the signals in the MJLS and in the LTI system are equal.

Our first result is stated as follows:

Theorem 7.2 (Spectral Equivalence) *Consider the MJLS in Figure 7.2, the LTI system in Figure 7.3, and suppose that Assumptions 7.1, 7.2 and 7.3 hold. If the MJLS is MSS and the LTI system is internally stable, then the stationary PSDs of v, e and w are equal in both situations.*

Proof: (This proof closely follows the ideas in [100] [see Remark 7.2]) Since Assumption 7.1(*a*-*b*) and 7.2 hold, the MJLS is MSS, and the LTI system is internally stable, we can assume, without loss of generality, that the initial state of $N(z)$, i.e., x_o, is zero and that the external signal $\bar{d}$ has zero mean.

- We will start by considering the MJLS. If the MJLS is MSS, then e, v and w are all asymptotically wss and pairwise jointly asymptotically wss (see Corollary 7.1). Therefore, the stationary PSDs of these signals, and the stationary cross-PSDs among them, are well defined and the stationary means of e, v and w, i.e., and μ_e, μ_v and μ_w, equal zero (recall the comments made above). Therefore, we can proceed as follows: from (7.14) we immediately have that

$$S_{ew}(z) = N_{11}(z)S_{\bar{d}w}(z) + N_{12}(z)S_w(z), \tag{7.20}$$

$$S_{vw}(z) = N_{21}(z)S_{\bar{d}w}(z) + N_{22}(z)S_w(z), \tag{7.21}$$

$$S_e(z) = N_{11}(z)S_{\bar{d}}(z)N_{11}(z)^\sim + N_{12}(z)S_w(z)N_{12}(z)^\sim + N_{11}(z)S_{\bar{d}w}(z)N_{12}(z)^\sim + N_{12}(z)S_{w\bar{d}}(z)N_{11}(z)^\sim, \tag{7.22}$$

$$S_v(z) = N_{21}(z)S_{\bar{d}}(z)N_{21}(z)^\sim + N_{22}(z)S_w(z)N_{22}(z)^\sim + N_{21}(z)S_{\bar{d}w}(z)N_{22}(z)^\sim + N_{22}(z)S_{w\bar{d}}(z)N_{21}(z)^\sim. \tag{7.23}$$

Moreover, we have that (recall that $N_{22}(z)$ is strictly proper and that we work under the assumption that $\mu_{\bar{d}} = 0$ and that the initial state of $N(z)$ is zero)

$$\begin{aligned} R_{w\bar{d}}(\tau) &= \lim_{k\to\infty} \mathcal{E}\left\{(1-d_r(k+\tau))\left(\sum_{i=0}^{\infty} n_{21}(i)\bar{d}(k+\tau-i) + \sum_{i=1}^{\infty} n_{22}(i)w(k+\tau-i)\right)\bar{d}(k)^H\right\} \\ &= \lim_{k\to\infty} \sum_{i=0}^{\infty} n_{21}(i)\mathcal{E}\left\{(1-d_r(k+\tau))\,\bar{d}(k+\tau-i)\bar{d}(k)^H\right\} \\ &\qquad + \lim_{k\to\infty} \sum_{i=1}^{\infty} n_{22}(i)\mathcal{E}\left\{(1-d_r(k+\tau))\,w(k+\tau-i)\bar{d}(k)^H\right\} \\ &= p\sum_{i=0}^{\infty} n_{21}(i)R_{\bar{d}}(\tau-i) + p\sum_{i=1}^{\infty} n_{22}(i)R_{w\bar{d}}(\tau-i), \end{aligned} \tag{7.24}$$

where n_{ij} refers to the impulse response of $N_{ij}(z)$, and where we have used the fact that d_r is scalar, that d_r is independent of $\bar{d}$, and that, since $N(z) \in \mathcal{R}_p$, $d_r(k+\tau)$ is independent of $w(k+\tau-i)$ for $i \geq 1$. Therefore, $S_{w\bar{d}}(z) = pN_{21}(z)S_{\bar{d}}(z) + pN_{22}(z)S_{w\bar{d}}(z)$ which is equivalent to

$$S_{w\bar{d}}(z) = p\left(1 - pN_{22}(z)\right)^{-1} N_{21}(z)S_{\bar{d}}(z). \tag{7.25}$$

In addition, we also have that for $\tau > 0$

$$R_w(\tau) = \lim_{k\to\infty} \mathcal{E}\left\{(1-d_r(k+\tau))v(k+\tau)w(k)\right\} = pR_{vw}(\tau), \tag{7.26}$$

where we used the fact that, since $N_{22}(z)$ is strictly proper, $w(k)$ is independent of $d_r(k+\tau)$ (only) for $\tau > 0$. Therefore,[2]

$$S_w^+(z) = pS_{vw}^+(z) \tag{7.27}$$

and, as a consequence,

$$S_w^+(z) = p\left(S_{vw}(z) - S_{vw}^-(z) - R_{vw}(0)\right). \tag{7.28}$$

We now calculate $S_{vw}^-(z)$: proceeding as above, we see that only for $\tau < 0$

$$R_{vw}(\tau) = pR_v(\tau) \quad \Rightarrow \quad S_{vw}^-(z) = pS_v^-(z). \tag{7.29}$$

Using (7.29) in (7.28) yields

$$S_w^+(z) = p\left(S_{vw}(z) - pS_v^-(z) - R_{vw}(0)\right). \tag{7.30}$$

As a consequence,

$$\begin{aligned} S_w(z) &= S_w^+(z) + S_w^-(z) + R_w(0) \\ &= p\left(S_{vw}(z) - pS_v^-(z) - R_{vw}(0)\right) \\ &\qquad\qquad + p\left(S_{vw}(z)^\sim - pS_v^+(z) - R_{vw}(0)^H\right) + R_w(0) \\ &= p\left(S_{vw}(z) + S_{vw}(z)^\sim\right) - p^2\left(S_v^+(z) + S_v^-(z) + R_v(0)\right) \\ &\qquad\qquad + p^2R_v(0) + R_w(0) - p\left(R_{vw}(0) + R_{vw}(0)^H\right). \end{aligned} \tag{7.31}$$

Now, since $N_{22}(z)$ is strictly proper, the following results hold:

$$R_w(0) = \lim_{k\to\infty} \mathcal{E}\left\{(1-d_r(k))v(k)(1-d_r(k))v(k)^H\right\} = pR_v(0), \tag{7.32}$$

$$R_{vw}(0) = \lim_{k\to\infty} \mathcal{E}\left\{v(k)(1-d_r(k))v(k)^H\right\} = pR_v(0). \tag{7.33}$$

Since $R_v(0)$ is real and symmetric, these two expressions allow one to simplify (7.31) to

$$S_w(z) = p\left(S_{vw}(z) + S_{vw}(z)^\sim\right) - p^2S_v(z) + p(1-p)R_v(0). \tag{7.34}$$

Equations (7.20)-(7.23), (7.25) and (7.34) allow one to calculate $S_v(z), S_w(z)$ and $S_e(z)$. (Note that there are exactly 6 linear equations for the 6 unknown stationary (cross-) power spectral densities.)

[2] Recall the notation introduced at the end of Section 2.3.

- We will now consider the LTI system. To prove our claim, we will show that the equations that allow one to calculate $S_v(z), S_w(z)$ and $S_e(z)$ in this case are exactly the same as those derived above for the MJLS. If the LTI system is asymptotically stable, then v, e and w are asymptotically wss and pairwise jointly asymptotically wss, with zero stationary mean (recall that we work under the assumptions made at the beginning of this proof; see also Section 2.5). Therefore, one can proceed as follows: we first note that (7.20)-(7.23) also hold in this case and that, as a consequence, we simply need to prove that (7.25) and (7.34) hold.

  From Figure 7.3 we have that (recall that we assume $N_{22}(z) \in \mathcal{R}_{sp}$, zero initial conditions, and $\mu_{\bar{d}} = 0$)

$$\begin{aligned} R_{w\bar{d}}(\tau) &= \lim_{k\to\infty} \mathcal{E}\left\{\left(q(k+\tau) + p\left(\sum_{i=0}^{\infty} n_{21}(i)\bar{d}(k+\tau-i) + \sum_{i=1}^{\infty} n_{22}(i)w(k+\tau-i)\right)\right)\bar{d}(k)^H\right\} \\ &= \lim_{k\to\infty} \mathcal{E}\left\{q(k+\tau)\bar{d}(k)^H\right\} + \lim_{k\to\infty} p\sum_{i=0}^{\infty} n_{21}(i)\mathcal{E}\left\{\bar{d}(k+\tau-i)\bar{d}(k)^H\right\} \\ &\qquad + \lim_{k\to\infty} p\sum_{i=1}^{\infty} n_{22}(i)\mathcal{E}\left\{w(k+\tau-i)\bar{d}(k)^H\right\}. \end{aligned} \tag{7.35}$$

  Since q has zero mean and is independent of $\bar{d}$, we have that (7.35) implies (7.25).

  It also follows directly from Figure 7.3 that, for $\tau > 0$

$$R_w(\tau) = \lim_{k\to\infty} \mathcal{E}\left\{q(k+\tau)w(k)\right\} + \lim_{k\to\infty} p\mathcal{E}\left\{v(k+\tau)w(k)\right\} = pR_{vw}(\tau), \tag{7.36}$$

  where we used the fact that $N_{22}(z) \in \mathcal{R}_{sp}$ and that q being i.i.d. and independent of $\bar{d}$ implies that $w(k)$ is independent of $q(k+\tau)$ for $\tau > 0$. Therefore,

$$S_w^+(z) = pS_{vw}^+(z). \tag{7.37}$$

  Proceeding as above we also have that, for $\tau < 0$,

$$R_{vw}(\tau) = \lim_{k\to\infty} \mathcal{E}\left\{v(k+\tau)q(k)\right\} + \lim_{k\to\infty} p\mathcal{E}\left\{v(k+\tau)v(k)\right\} = pR_v(\tau). \tag{7.38}$$

  Accordingly,

$$S_{vw}^-(z) = pS_v^-(z). \tag{7.39}$$

  It also follows directly from (7.16) and Figure 7.3 that

$$R_w(0) = R_{vw}(0) = pR_v(0), \tag{7.40}$$

where the existence of $R_v(0)$ is guaranteed by the asymptotic stability of the LTI system. Proceeding as in the MJLS, one can conclude that (7.34) also holds in this case and the result follows.

□□□

Theorem 7.2 implies that in order to analyze spectral properties of the MJLS of Figure 7.2, one can equivalently utilize the LTI system of Figure 7.3 (at least from an input-output point of view). Clearly, this simplifies the analysis of the MJLS. An important consequence of Theorem 7.2 is stated next:

Corollary 7.2 (Constraint on successful transmission probability p) *Consider the MJLS in Figure 7.2, the LTI system in Figure 7.3, and suppose that Assumptions 7.1, 7.2 and 7.3 hold. If the MJLS is MSS and the LTI system is internally stable, then*

$$1 > p(1-p)\left|\left|\left(1-pN_{22}(z)\right)^{-1}N_{22}(z)\right|\right|_2^2. \tag{7.41}$$

Moreover, if $p \in (0,1)$, then (7.41) is equivalent to $\sigma_q^2 \in \mathbb{R}_0^+$ and equivalent to

$$\frac{p}{1-p} > ||T_p(z)||_2^2, \tag{7.42}$$

where $T_p(z)$ is the transfer function from q to $v_p \triangleq pv$ in Figure 7.3, namely

$$T_p(z) \triangleq \left(1-pN_{22}(z)\right)^{-1}pN_{22}(z). \tag{7.43}$$

Proof: Since both the MJLS and the LTI systems are stable, we can use Theorem 7.2 to infer spectral properties of the MJLS using linear analysis for the system of Figure 7.3. A straightforward manipulation shows that, since the LTI system is internally stable and Assumptions 7.1(*a*-*b*), 7.2 and 7.3 hold, σ_v^2 exists, is finite and obeys

$$\sigma_v^2\left(1-p(1-p)\left|\left|\left(1-pN_{22}(z)\right)^{-1}N_{22}(z)\right|\right|_2^2\right) = \left|\left|\left(1-pN_{22}(z)\right)^{-1}N_{21}(z)\Omega_{\bar{d}}(z)\right|\right|_2^2. \tag{7.44}$$

Since σ_v^2 is well defined, finite and, by definition, unique, the results follow immediately from (7.44) and (7.16). □□□

In the light of Corollary 7.2 we conclude that, if the LTI system in Figure 7.3 is internally stable, then (7.41) (or (7.42) if $p \in (0,1)$) is necessary for the MSS of the MJLS. The next theorem shows that this is a special instance of a stronger result:

Theorem 7.3 (Stability equivalence) *Consider the MJLS in Figure 7.2, the LTI system in Figure 7.3, and suppose that Assumptions 7.1, 7.2 and 7.3 hold. Then, the MJLS is MSS if and only if the LTI system is asymptotically stable and* (7.41) *holds.*

Proof: The result is immediate for $p \in \{0,1\}$ (recall Corollary 7.2). To prove the result for $p \in (0,1)$ we will use the state space descriptions of both the MJLS and the LTI system introduced before (see (7.1), (7.15) and (7.17)). With that notation, a realization of the transfer function between q and v in Figure 7.3, namely $(1 - pN_{22}(z))^{-1} N_{22}(z)$, is given by $(\hat{A}, B_2, C_2, 0)$.

- ($\Leftarrow$) (This part of the proof closely follows the ideas in [100].) If the LTI system is internally stable (hence $\hat{A}$ has all its eigenvalues strictly inside the unit circle), then for every $M > 0$ there exists a unique and positive definite $\hat{P}$ such that (see, e.g., Lemmas 21.1 and 21.2 in [194])

$$\hat{A}\hat{P}\hat{A}^H - \hat{P} + M = 0 \tag{7.45}$$

and, moreover, we can write (recall that $N_{22}(z)$ is scalar and see, e.g., Remark 21.6 in [194])

$$\left|\left|(1 - pN_{22}(z))^{-1} N_{22}(z)\right|\right|_2^2 = C_2 L C_2^H, \tag{7.46}$$

where L is the unique and positive semi-definite solution of

$$\hat{A}L\hat{A}^H - L + B_2 B_2^H = 0. \tag{7.47}$$

As a consequence, if the LTI system is stable and (7.41) holds, then (recall that we consider only $p \in (0,1)$)

$$C_2 L C_2^H < \frac{1}{p(1-p)}. \tag{7.48}$$

Therefore, if $\delta > 0$ is sufficiently small, then we also have that

$$C_2 \left(L + \delta\hat{P}\right) C_2^H < \frac{1}{p(1-p)}. \tag{7.49}$$

To complete the proof, we note that straightforward manipulations show that, since $\hat{A} = (1-p)A(1) + pA(0)$ and, for every matrix P of appropriate dimensions,

$$\left(A(1) - A(0)\right) P \left(A(1) - A(0)\right)^H = B_2 C_2 P C_2^H B_2^H, \tag{7.50}$$

the following holds:

$$(1-p)A(1)PA(1)^H + pA(0)PA(0)^H = p(1-p)B_2 C_2 P C_2^H B_2^H + \hat{A}P\hat{A}^H. \tag{7.51}$$

If one takes $P = L + \delta\hat{P}$ in (7.51), then use of (7.45), (7.49) and the fact that $M > 0$ and $\delta > 0$ imply that

$$\begin{aligned}(1-p)A(1)\left(L+\delta\hat{P}\right)A(1)^H + pA(0)\left(L+\delta\hat{P}\right)A(0)^H &< B_2B_2^H + \hat{A}L\hat{A}^H + \delta\hat{A}\hat{P}\hat{A}^H \\ &= L + \delta\hat{P} - \delta M \\ &< L + \delta\hat{P}. \end{aligned} \quad (7.52)$$

The result thus follows upon using Theorem 7.1.

- ($\Rightarrow$) The mean of the state of $N(z)$ in the MJLS of Figure 7.2 evolves according to (see (7.1) and (7.15))

$$\mathcal{E}\left\{x(k+1)\right\} = \mathcal{E}\left\{A(d_r(k))x(k)\right\} + \mathcal{E}\left\{B(d_r(k))\bar{d}(k)\right\} = \hat{A}\mathcal{E}\left\{x(k)\right\}, \quad (7.53)$$

where we have used the definition of $\hat{A}$ in (7.18) and we assumed, without loss of generality, that $\mu_{\bar{d}} = 0$. Accordingly,

$$\mathcal{E}\left\{x(k)\right\} = \hat{A}^k\mathcal{E}\left\{x_o\right\}. \quad (7.54)$$

Since the MJLS is MSS, then μ_x is well defined (and independent of x_o). Therefore, $\hat{A}$ has to have all its eigenvalues inside the unit circle, i.e., the LTI system in Figure 7.3 must be internally stable. Inequality (7.41) then follows from Corollary 7.2.

□□□

We conclude from Theorems 7.2 and 7.3 that spectral and MSS-related properties of the MJLS in Figure 7.2 can be studied by means of the LTI system in Figure 7.3, where the unreliable channel has been replaced by an additive i.i.d. noise channel, having gain p and a fixed signal-to-noise ratio. This fundamental insight can be exploited to analyze and design networked control systems that are closed over channels that drop data in an i.i.d. fashion.

Remark 7.2 (Previous work) *In the light of Proposition 3.6 in [30] the necessity part of Theorem 7.3 is no surprise: for any MJLS, the asymptotic stability of the LTI system that governs the dynamics of the mean of the state, say $\mathcal{S}_{\mathcal{E}}$, is necessary for the MSS of the MJLS. The interesting aspect of our result lies in the fact that it is not generally true that the asymptotic stability of $\mathcal{S}_{\mathcal{E}}$ is sufficient for the MSS of the MJLS (see remark 3.7 in [30]). Theorem 7.3 shows, for a class of control relevant MJLSs, that if $\mathcal{S}_{\mathcal{E}}$ is asymptotically stable and an additional*

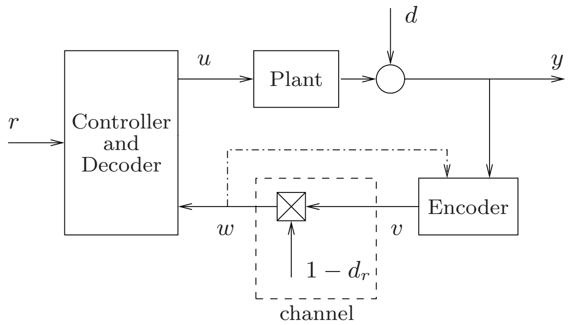

Figure 7.4: Networked control system closed over an erasure channel.

inequality constraint is satisfied, then the MSS of the MJLS is guaranteed. We note that a particular instance *of this result (and of Theorem 7.2 and Corollary 7.2) was proven in [99, 100], but neither the connection to* $\mathcal{S}_{\mathcal{E}}$ *nor the connection to signal-to-noise ratio constrained channels was made in [99, 100].* □□

7.4 Applications to Control System Design

In this section we apply the result of Section 7.3 to control system design. We will focus on the networked control situation depicted in Figure 7.4. In that figure, r is a reference signal, d is a disturbance, and the channel is an analog i.i.d erasure channel (see Definition 7.2).

As Figure 7.4 suggests, we assume that it may be possible to know the output of the channel at the encoder side. This is a reasonable assumption if the communication protocol is such that the sending end receives an acknowledgement every time a successful transmission takes place. (This is the case for, so-called, *TCP-like* protocols; see, e.g., [78,148].) If such acknowledgements are not available, then we can not assume that the output of the channel is known at the sending end. (This happens when, so-called, *UDP-like* protocols are used.) Of course, it is not sensible to assume instantaneous acknowledgements. Hence, we will assume that the output of the channel, if known at the encoder side, is known with one sample delay. This assumption is common in the literature (see, e.g., [78, 148]).

As in previous chapters, we will work under the assumption that it is impossible to use any communication channels other than those explicitly shown in Figure 7.4. We also restrict our attention to LTI control and coding architectures. With these constraints in mind, and considering that the output of the channel may (in some cases) be known at the encoder side, we propose to use the control and coding architecture illustrated in Figure 7.5. This architecture

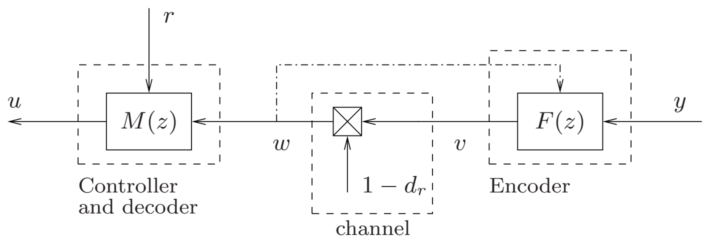

Figure 7.5: Proposed control and coding architecture.

is the straightforward extension of the architecture studied in Chapter 4 to the case of control systems that are closed over unreliable channels. We will maintain the notation introduced in Chapter 4, which we repeat here for ease of reference: the filters $F(z)$ and $M(z)$ belong to $\mathcal{R}_p^{1\times 2}$ and are such that

$$v = F(z)\begin{bmatrix} w \\ y \end{bmatrix} \triangleq F_1(z)w + F_2(z)y, \qquad u = M(z)\begin{bmatrix} r \\ w \end{bmatrix} \triangleq M_1(z)r + M_2(z)w; \tag{7.55}$$

we also define

$$K(z) \triangleq \begin{bmatrix} F(z) & M(z) \end{bmatrix}. \tag{7.56}$$

We note that, in the present case, the constraint $F_1(z) \in \mathcal{R}_{sp}$ arises naturally (if that were not the case, then instantaneous acknowledgements would be available). Our proposal constitutes the most general architecture that uses only LTI filters and no communication channels other than those explicitly shown in Figure 7.4.

By virtue of the results of Section 7.3, we can (provided mild conditions are satisfied) reduce the study of the networked control system that arises when the proposed control and coding architecture is employed in the networked control system of Figure 7.4, to the study of the linear system depicted in Figure 7.6. In that figure, $G(z)$ is the plant transfer function, p is the probability of successful transmission over the erasure channel, and q satisfies Assumption 7.3.

Throughout this section we will assume that the following holds:

Assumption 7.4 (Plant, signals and initial states)

(a) The plant transfer function $G(z)$ belongs to $\mathcal{R}_{sp}^{1\times 1}$, is non-zero, and has a stabilizable and detectable underlying realization.

(b) The signals r and d are second order jointly independent and mutually uncorrelated wss scalar processes with rational PSDs. At least one signal, r or d, is non-zero. If r or d are non-zero, then they admit spectral factors in $\mathcal{U}_\infty$.

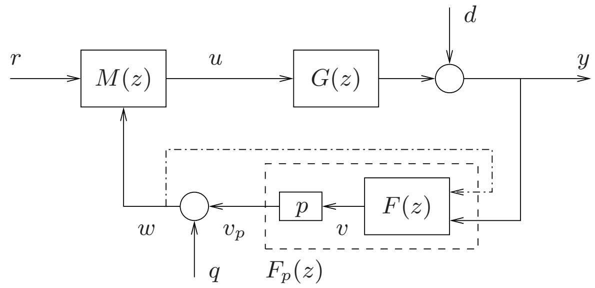

Figure 7.6: Model of the networked control situation under study.

(c) The initial states of all the filters in Figure 7.6 (including plant model) are jointly second order random variables.

□□

Assumption 7.4 is almost identical to the assumptions made in previous chapters. The only difference lies in that, here, we assume that r and d jointly independent. This ensures that Assumption 7.1(a) holds for $\bar{d} \triangleq [r\ d]^T$. We also note that, since $F_1(z) \in \mathcal{R}_{sp}$, $G(z) \in \mathcal{R}_{sp}$ guarantees that the transfer function that in Figure 7.6 plays the role of $N(z)$ in Section 7.3 satisfies Assumption 7.2.

7.4.1 Mean square stability

In this section we study conditions on the successful transmission probability p that ensure that there exist $K(z)$ such that the networked control system in Figure 7.4 is MSS, when the control and coding architecture is as in Figure 7.5. As the reader may have already noticed, our results easily follow from our analysis given earlier in Chapter 4.

Theorem 7.4 (Minimal p for MSS) *Consider the networked control system of Figure 7.4, where the control and coding architecture is as in Figure 7.5, and Assumptions 7.1(c) and 7.4 hold. If $G(z)$ is unstable, then:*

1. *For the case of TCP-like protocols, it is possible to choose $K(z)$ such that MSS holds if and only if*

$$p > p_{\inf} \triangleq 1 - \frac{1}{\prod_{i=1}^{n_p} |p_i|^2}. \tag{7.57}$$

For any $p > p_{\inf}$, MSS is guaranteed if one chooses

$$K(z) = K_p(z)\,\mathrm{diag}\left\{p^{-1}, 1\right\}, \tag{7.58}$$

where $K_p(z)$ satisfies the properties of $K(z)$ in Theorem 4.1.

2. *For the case of UDP-like protocols, it is possible to choose $K(z)$ such that MSS holds if and only if*

$$p > p_{\text{inf}}^U \triangleq 1 - \frac{1}{\left(\prod_{i=1}^{n_p} |p_i|^2\right) + \Delta_G}, \tag{7.59}$$

where $\Delta_G \geq 0$ is defined as in (4.29) (recall that $\Delta_G = 0$ if and only if the plant is marginally-MP and has relative degree one). For any $p > p_{\text{inf}}^U$, MSS is guaranteed if one chooses $K(z)$ as in (7.58) and $K_p(z)$ satisfies the properties of $K(z)$ in Theorem 4.2.

Proof: Define $\bar{d} \triangleq [r\ d]^T$. Since $G(z), F_1(z) \in \mathcal{R}_{sp}$ (note that in the UDP-like case, $F_1(z) = 0 \in \mathcal{R}_{sp}$), and Assumptions 7.1($c$) and 7.4 hold, it follows that, irrespective of the protocol employed, the results in Section 7.3 apply to the case under study (i.e., Assumptions 7.1 and 7.2 hold). Hence, we will focus on the linear feedback system of Figure 7.6, where q satisfies Assumption 7.3. Without loss of generality, we will focus on $0 < p < 1$.[3] From Theorem 7.3 we know that the networked control system of interest is MSS if and only if $K(z)$ is such that the linear system in Figure 7.6 is well-posed and internally stable, and the transfer function from q to v_p in Figure 7.6, say $T_p(z)$, satisfies (7.42). (Note that the right hand side in (7.42) is an increasing function of p.)

Since $p > 0$, we can define $F_p(z) \triangleq pF(z)$, $K_p(z) \triangleq [F_p(z)\ M(z)]$ and, for any given $F_p(z)$, recover $F(z)$ and *vice versa.*

1. In this case, $K(z)$ is such that the linear system in Figure 7.6 is well-posed and internally stable if and only if $K_p(z) \in \mathcal{S}$, where $\mathcal{S}$ is defined as in (4.5) with the obvious changes (the definition should refer to Figure 7.6). It is immediate to see from the proof of Theorem 4.1 that

$$\inf_{K_p(z) \in \mathcal{S}} ||T_p(z)||_2^2 = \left(\prod_{i=1}^{n_p} |p_i|^2\right) - 1. \tag{7.60}$$

The result follows immediately from (7.42), (7.60) and Theorem 4.1.

2. The proof of this part is completely analogous to the proof of Part 1 above. (In this case, one needs to use Theorem 4.2 instead of Theorem 4.1.)

[3] Since $G(z)$ is assumed unstable, if $p = 0$, then the resulting networked control system is not MSS for any $K(z)$. If $p = 1$, we have a standard control problem and it immediately follows from, e.g., Corollary 4.1 that there exists $K(z)$ such that MSS is achieved.

□□□

Remark 7.3 (Minimal p for stable plants) *If $G(z)$ is stable, then it is clear that, for any $p \in [0, 1]$, there exists $K(z)$ such that the resulting loop is MSS (just choose $K(z) = 0$). Thus, $p_{\text{inf}} = p_{\text{inf}}^U = 0$ in that case.* □□

Remark 7.4 (Other architectures) *We would like to stress that, as the reader may readily verify, the following holds:*

- *If one considers TCP-like protocols and forces $M(z) = [1 \ \ -1]$, then (7.57) is also necessary and sufficient for one to be able to find $F(z)$ that achieves MSS.*
- *If one chooses $F(z) = [0 \ \ 1]$ and $M(z) = C(z)[1 \ \ -1]$, then (7.59) is also necessary and sufficient for one to be able to find $C(z)$ that achieves MSS (regardless of the protocol employed). This case corresponds to a standard one-dof control architecture where the feedback signal is sent (without any pre- or post-processing) over an analog erasure channel.*

□□

Theorem 7.4 gives explicit characterizations of the minimal successful transmission probability that guarantees the existence of $K(z)$ such that the networked control system of interest is MSS. Our results state that, as intuition suggest, plants that have large magnitude unstable poles require more reliable channels than those plants having moderate unstable poles. When TCP-like protocols are used, unstable poles are the only source of limitations on the admissible successful transmission probability. However, when using UDP-like protocols the requirements on the admissible successful transmission probability are, in general, more stringent than in situations that employ TCP-like protocols. In the UDP-like case, the plant strictly-NMP zeros and relative degree also play a significative role. Interestingly, the requirements in both situations are the same if and only if the plant is marginally-MP and has relative degree one.

Remark 7.5 (Relationship to previous work) *In the TCP-like protocol case, our results are consistent with the results in [78, 148] (see, in particular, Theorems 5.4 and 5.6 in [148]). The interesting aspect of our results is that they hold for a constrained class of control and coding architectures that uses LTI filters only. Stated otherwise, for any given i.i.d. analog erasure channel, the class of SISO plant models for which the time varying schemes studied in [148] guarantee MSS is the same class for which our proposal guarantees MSS. This suggests* *that, when focusing on MSS, the consideration of time varying control and coding architectures is actually not necessary.*

In the UDP-like protocol case, the results of Theorem 6.1 in [148][4] *give explicit necessary conditions on the dropout probabilities that guarantee MSS for the case of plant models that have a square and invertible state-to-output matrix ("C" matrix). That condition turns out to be also sufficient in the case of systems that, in addition, have a square and invertible input-to-state matrix ("B" matrix; see also [78]). In the SISO case, having both a square input-to-state and state-to-output matrix is tantamount to having a scalar plant model. Our results are consistent with those reported in [148], but they provide stronger necessary and sufficient conditions that are valid for any SISO plant model when the control and coding scheme is constrained to be LTI.*

Our results also extend the results in Section 8 in [41]. In that work, the author gives conditions on the successful transmission probability for one-degree of freedom control architectures (with no pre- or post-processing) when the plant state is noiselessly available for measurement. It turns out that, in that restricted case, (7.57) *is necessary and sufficient for one to be able to find a feedback law that guarantees MSS. In the light of the results of Section 7.3 and the results in [18] that deal with state feedback, the results in [41] are not surprising. Our results extend the framework in Section 8 of [41] to the output feedback case and show that, in that case, not only the unstable plant poles limit the channel reliability needed to achieve MSS, but also the plant NMP zeros and relative degree.* □□

We end this section noting that the following holds:

Corollary 7.3 (Getting close to p_{inf}) *Consider the networked control system of Figure 7.4, where the control and coding architecture is as in Figure 7.5 and Assumptions 7.1*(c) *and 7.4 hold. Assume that a TCP-like protocol is employed and that $G(z)$ is unstable. If the successful transmission probability $p \to p_{\text{inf}}$, then $\sigma_v^2 \to \infty$ for any choice of $K(z)$ (unless $d = 0$ and $M_1(z) = 0$).*

Proof: Immediate from the proof of Corollary 7.2 and Figure 7.6. □□□

From Corollary 7.3 we see that if p approaches its minimal admissible value p_{inf}, then the performance of the networked control system under study will be, in general, severely compromised. In other words, in order to attain acceptable performance levels, one needs to consider success

[4]Note that equation (40) in Theorem 6.1 in [148] contains a typo that is corrected in the proof of that theorem.

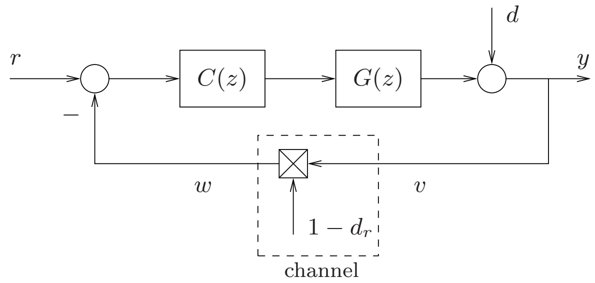

Figure 7.7: One degree of freedom control loop with feedback over unreliable channel.

probabilities that are (considerably) higher than p_{inf}. This motivates the study of performance related questions pursued in the next section.

Remark 7.6 *If one focuses on UDP-like protocols, then a result similar to Corollary 7.3 holds.*

□□

7.4.2 Design for performance

In this section we show how to use the results of Section 7.3 to solve performance related control problems. We will focus on a simple case, where a one-dof controller has to be designed to optimize performance in a networked situation employing an analog erasure channel in its feedback path. That is, we will focus on the networked control system that arises when the control and coding architecture of Figure 7.5 is used in the scheme depicted in Figure 7.4, and

$$F(z) = \begin{bmatrix} 0 & 1 \end{bmatrix}, \quad M(z) = C(z) \begin{bmatrix} 1 & -1 \end{bmatrix}. \tag{7.61}$$

As the reader can easily verify, the ideas in Chapter 4 allow one to generalize the methodology presented below to more general cases. (Hence, we decided not to bother her/him with excessive repetitions.) For ease of reference, we present a diagram of the situation under study in Figure 7.7.

We can state the problem of interest as follows:

Problem 7.1 (Optimal performance for given p) *Consider the networked control situation depicted in Figure 7.7, where d_r satisfies Assumption 7.1(c), and Assumption 7.4 holds. For any given dropout probability $p \in (0, 1)$, find (or prove the problem infeasible)*

$$\left[\sigma_e^2\right]_p \triangleq \inf_{C(z) \in \mathcal{S}_p} \sigma_e^2, \tag{7.62}$$

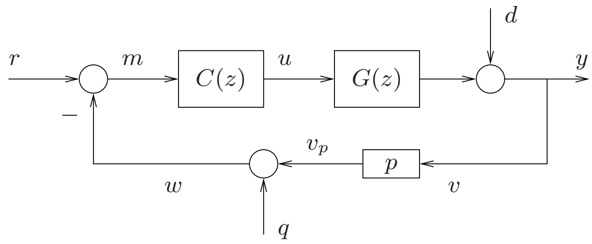

Figure 7.8: Model of the networked situation in Figure 7.7.

where

$$\mathcal{S}_p \triangleq \{C(z) \in \mathcal{R}_p : \textit{the feedback loop in Figure 7.7 is MSS}\}. \tag{7.63}$$

Also find the controller $C(z)$ *that allows one to achieve* $[\sigma_e^2]_p$ *(or get arbitrarily close to* $[\sigma_e^2]_p$*).* □□

In Problem 7.1, we left the case $p \in \{0, 1\}$ aside. These cases can be studied using standard control theoretic ideas and, as such, are not of interest here. For future reference we define

$$C_p(z) \triangleq \arg \inf_{C(z) \in \mathcal{S}_p} \sigma_e^2. \tag{7.64}$$

The rest of this section is devoted to showing that Problem 7.1 can be recast as a special case of the general problem studied in Section 4.4. To that end, we use the results in Section 7.3 and focus on the LTI system of Figure 7.8, where q satisfies Assumption 7.3. Provided the conditions of Problem 7.1 are satisfied, Theorem 7.3 and Corollary 7.2 allow one to conclude that the networked situation in Figure 7.7 is MSS if and only if $C(z) \in \mathcal{S}$, where

$$\mathcal{S} \triangleq \{C(z) \in \mathcal{R}_p : \text{the loop in Figure 7.8 is internally stable and well-posed}\}, \tag{7.65}$$

and $\sigma_q^2 \in \mathbb{R}_0^+$. For these controllers and noise variances, Theorem 7.2 ensures that the PSDs of both e and v in the feedback systems of Figures 7.7 and 7.8 are the same. Therefore, $[\sigma_e^2]_p$ can be written in the equivalent form

$$[\sigma_e^2]_p = \inf_{\substack{C(z)\in\mathcal{S} \\ \sigma_q^2 \in \mathbb{R}_0^+ \\ \gamma = \frac{p}{1-p}}} \sigma_e^2, \quad \gamma \triangleq \frac{\sigma_{v_p}^2}{\sigma_q^2}, \tag{7.66}$$

where v_p is the output of the static gain block in Figure 7.8.

The first two constraints in the optimization problem in (7.66) originate from MSS requirements, as discussed earlier. The third constraint, namely an equality signal-to-noise ratio constraint, is a consequence of the fact that the LTI system of Figure 7.8 models the networked situation of Figure 7.7 if and only if q satisfies Assumption 7.3. If this constraint were an inequality one, then $\left[\sigma_e^2\right]_p$ could be found using the techniques in Section 4.4. Thus, one can think of relaxing the equality constraint and replacing it by an inequality one, i.e., replacing the third constraint in (7.66) by

$$\gamma \leq \frac{p}{1-p}. \tag{7.67}$$

However, since p is assumed fixed, replacing the equality constraint in (7.66) by (7.67) will yield a solution to the original problem if and only if the constraint is active at the optimum (if this is not the case, then the additive noise model has, simply, no relationship with the erasure channel we intend to utilize).

If neither r nor d are zero, then it is straightforward to see that the problem of finding $\left[\sigma_e^2\right]_p$ in (7.66), when the third constraint is replaced by (7.67) (and no attention is paid to the fact that this constraint *must* be active at the optimum), fits into the setting examined in Section 4.4 with

$$P(z) = \left[\begin{array}{ccc:c} 0 & p & 0 & pG(z) \\ 1 & -1 & 0 & -G(z) \\ \hdashline 1 & -p & -1 & -pG(z) \end{array}\right], \tag{7.68}$$

$\bar{K}(z) = C(z)$, $\hat{d} = [r\ d]^T$, $\bar{u} = u$, $\bar{y} = m$, and $v = v_p$ (see Figure 7.8). If either r or d is zero,[5] then $\hat{d}$ equals the non-zero signal, and $P(z)$ and $\bar{y}$ need to be modified accordingly. It is immediate to see that, in either case, if $p \in (0,1)$ and Assumption 7.4 holds, then Assumption 4.2 also holds.

In order to apply the results of Section 4.4, and to ensure that they provide a meaningful solution in the present case, we note that the following holds:

Lemma 7.1 *Consider the feedback system in Figure 7.8 and suppose that Assumptions 7.3 and 7.4 hold. If* $p \in (p_{\text{inf}}^U, 1)$*, then the optimization problem in* (7.66)*, when the third constraint is replaced by* (7.67)*, is feasible, the optimal variance of* q *belongs to* $\mathbb{R}^+$*, and the constraint* (7.67) *is active at the optimum.*

Proof: The feasibility of the problem follows from the definition of p_{inf}^U. The fact that the optimal variance of q belongs to $\mathbb{R}^+$, and that the constraint is active at the optimum, follow from a straightforward modification of the proof of Lemma 4.3. (Note that, since $p > p_{\text{inf}}^U \geq 0$ (see

[5] Assumption 7.4 precludes both signals to be simultaneously zero.

Theorem 7.4) and $p < 1$, the upper bound on γ in (7.67), say Γ, satisfies $\frac{p^U_{\text{inf}}}{1-p^U_{\text{inf}}} < \Gamma < \infty$.) □□□

From Lemma 7.1 we see that, for every $p \in (p^U_{\text{inf}}, 1)$, i.e., for every $p \in (0, 1)$ such that the existence of $C(z)$ that guarantees the MSS of the networked control system of Figure 7.7 is guaranteed (see Theorem 7.4, Remark (7.3) and the second bullet in Remark 7.4(b)), the results of Section 4.4 provide a straightforward characterization of the solution of Problem 7.1. That characterization, although not explicit, allows one to use simple numerical methods to find the best performance and the corresponding optimal controller $C_p(z)$.

Remark 7.7 (Previous work) *It should be clear that the performance achieved by the controller that our methodology suggests, is only optimal within the class of control and coding architectures that are LTI and satisfy* (7.61)*. Obviously, better performance may be attained if one considers time varying control and coding architectures as proposed in, e.g., [130, 132, 148, 154]. Indeed, it has been shown in [148] that optimal control policies are, in general, time varying and, in the case of UDP-like protocols, non-linear. Our approach, although suboptimal, may allow one to attain acceptable levels of performance with simple LTI architectures. In the case considered, our characterization of the optimal LTI controller is simpler than the characterization provided in [166], and relies on standard optimization ideas. We note, however, that the work in [166] focuses on architectures with two unreliable channels (one for the controller-to-actuator link, and one for the sensor-to-controller link) without any coding. It would be interesting to extend the ideas in this chapter to the situation studied in [166]. We leave this for future research.* □□

Remark 7.8 (Switching strategies) *Our approach can also be adapted to deal with a class of switching control and coding architectures. To illustrate how this can be achieved, we consider the architecture in Figure 7.9, and assume that the receiving end can distinguish valid data from corrupted data (i.e., the receiving end knows, at time k, $d_r(k)$). In Figure 7.9, the block K_s is a causal switching system characterized via*

$$u(k) = \sum_{i=0}^{\infty} k_{11}(i) r(k-i) + \sum_{i=0}^{\infty} k_{12}(i) \hat{v}(k-i), \tag{7.69a}$$

$$\hat{v}(k) = (1 - d_r(k)) w(k) + d_r(k) \left(\sum_{i=0}^{\infty} k_{21}(i) r(k-i) + \sum_{i=1}^{\infty} k_{22}(i) \hat{v}(k-i) \right), \tag{7.69b}$$

where k_{ij} are the impulse responses of causal LTI filters (note that the impulse response k_{22} is strictly causal). In other words, $u(k)$ depends on the whole history of reference signal values and the whole history of an auxiliary signal $\hat{v}$. This signal is such that it equals the signal

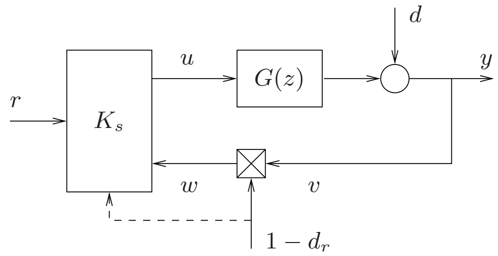

Figure 7.9: Feedback control architecture over erasure channel with switching controller.

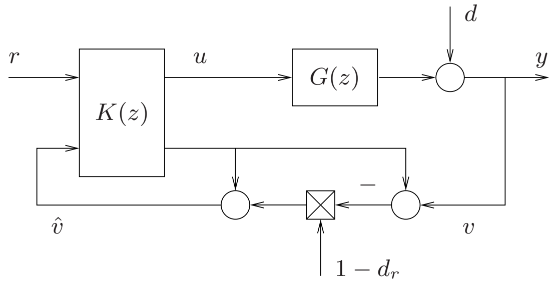

Figure 7.10: Equivalent redrawing of Figure 7.9 when K_s obeys (7.69).

sent over the channel at every successful transmission instant and, when data is lost, equals an estimate *of the value sent over the channel. This estimate is constructed using the history of the reference signal, the past estimates, and all past data that has been successfully received. The interconnection of $G(z)$ and K_s can be represented as shown in Figure 7.10, where*

$$K(z) \triangleq \begin{bmatrix} K_{11}(z) & K_{21}(z) \\ K_{21}(z) & K_{22}(z) \end{bmatrix} \in \mathcal{R}_p, \quad K_{22}(z) \in \mathcal{R}_{sp}, \tag{7.70}$$

and $K_{ij}(z)$ denotes the Z-transform of k_{ij}. From that figure it is immediate to see that, since $K_{22}(z) \in \mathcal{R}_{sp}$, and provided $G(z) \in \mathcal{R}_{sp}$ and suitable additional assumptions hold, the results in Section 7.3 apply. Thus, we can use signal-to-noise ratio related results to analyze and design the proposed switching control architecture. We note that this architecture is a generalization of the one presented in [100], where the authors constrained themselves to the case $K_{11}(z) = -K_{12}(z) = C(z)$, $K_{21}(z) = 0$, and $K_{22}(z) \in \mathcal{RH}_2$, where $C(z)$ is fixed and given, and $K_{22}(z)$ is the decision variable. □□

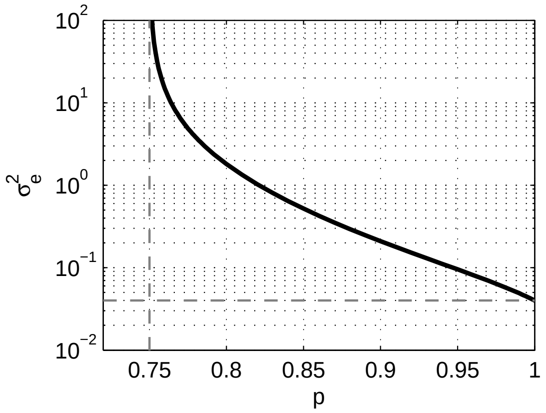

Figure 7.11: Best achievable performance as a function of the successful transmission probability p.

7.5 An Example

This section provides a simple example to illustrate the results in this chapter. As has been the case of the simulation studies in Chapters 4 and 6, we focus on a one-dof control architecture where the feedback path comprises a non transparent channel. In this case, we consider an analog i.i.d. erasure channel in the feedback path and, as before, we assume that d is zero and that

$$G(z) = \frac{z - 0.8}{z(z-2)}, \qquad \Omega_r(z) = \frac{0.1z}{z - 0.9}. \tag{7.71}$$

We will study the best achievable performance as a function of the channel successful transmission probability p.

For the considered plant and architecture, Theorem 7.4 and Remark 7.4 imply that one will be able to find a controller $C(z)$ that renders the feedback loop MSS if and only if

$$p > p_{\text{inf}}^U = 0.75. \tag{7.72}$$

We will thus focus on successful transmission probabilities $p \in (0.75, 1]$. Using the ideas in this chapter, we calculated the best achievable performance for such values of p. The results are presented in Figure 7.11. As expected, the best achievable performance is lower bounded by the best non-networked performance $D_{\text{inf}} = 0.04$ for every p. We also see that $\sigma_e^2 \to \infty$ when $p \to p_{\text{inf}}^U = 0.75$, and that $\sigma_e^2 \to D_{\text{inf}} = 0.04$ when $p \to 1$ (see vertical and horizontal lines in Figure 7.11, respectively). This behavior is consistent with our results.

7.6 Summary

This chapter has studied networked control problems where the communication constraint of importance is random data-loss. Focusing on LTI control and coding architectures and on the simplest model for data-dropouts (namely, analog i.i.d. erasure channels), we were able to show that the analysis and design of such networked control systems can be carried out using the tools which were developed in Chapter 4 for the case of signal-to-noise ratio constrained additive i.i.d. noise channels. These insights allowed us to conclude that, if the focus is on MSS only, then the restriction to LTI control and coding architectures posses no additional constraints on the successful transmission probability when compared with alternative time-varying policies previously reported in the literature. Our approach is also useful to design LTI control and coding architectures, and we have illustrated this via a simple special case.

Future work could focus on extending the ideas of this chapter to multiple channels and to the MIMO case. Another interesting extension is the consideration of unreliable digital channels. This would require a mixture between the techniques developed in this chapter and those in Chapters 5 and 6.

Chapter 8

Conclusions

8.1 Overview

This thesis has studied control problems with communication constraints. We have focused on a single channel architecture were the plant is assumed to be LTI SISO, and the communication constraints arise as either signal-to-noise ratio constraints, average data-rate constraints, or i.i.d. data-dropouts. For each of these scenarios, we have provided analysis and design methodologies. In particular, our results make the interplay between control objectives and communication constraints explicit, and also enable one to reduce the impact of communication constraints on closed loop performance by means of systematic design procedures.

In this thesis we have constrained ourselves to simple control and coding architectures that consist of LTI filters and, in the case of average data-rate limits, an i.i.d. source coding scheme. Within this setup, we have shown that both the analysis and design of networked control systems subject to the communication constraints detailed above can be carried out in a *unified manner* that exploits insights and results from linear systems theory. This nice result stems from the fact that we were able to reduce networked control problems subject to either average data-rate constraints or data-dropouts, to control problems with signal-to-noise ratio constraints. In the former case, we showed that one can use signal-to-noise ratio constraints to bound, from above, average data-rate requirements, whereas in the latter case, we were able to establish a precise second order moments equivalence between a class of unreliable channels and additive signal-to-noise ratio constrained channels. We believe these results to be novel and important. They allow one to easily gain insight into networked control problems, the associated fundamental limitations and, in passing, they shed light into many fundamental results previously presented in the literature. On the other hand, they have the potential to enable the analysis and design of more complex networked situations with several channels, MIMO plant models, and multiple communication constraints.

8.2 Summary of Contributions

We believe this thesis has made several contributions. We detail them below:

- **Chapter 3:** This chapter studied a situation where a non-networked controller design has to be implemented over a signal-to-noise ratio constrained additive noise channel. To that end, we proposed a feedback-based coding architecture and studied signal-to-noise requirements for MSS and optimal synthesis procedures. One of the main conclusions of this chapter is that the use of feedback-based coding schemes (in other words, architectures with additional degrees of freedom) can drastically reduce the signal-to-noise ratio requirements for MSS. We also showed that the proposed architecture can provide performance gains when compared with situations where no coding is in place. We also extended our results to a two-by-two MIMO case where additional non-transparent channels are used to enrich an existing decentralized control architecture. In this context, we showed that if the channels are of poor quality, then making use of them may be harmful from the loop performance point of view.

- **Chapter 4:** In this chapter we studied a general control and coding architecture aimed at controlling SISO plants over signal-to-noise ratio constrained additive noise channels. We first studied signal-to-noise ratio requirements for MSS and, as expected form our results in Chapter 3, we showed that the use of feedback-based coding schemes is the key to reduce signal-to-noise ratio requirements for MSS. Our results provide explicit characterizations of these requirements for a general architecture and many variants thereof. In a second stage, we showed that it is possible to analyze the interplay between closed loop performance and signal-to-noise ratio constraints using well-known convex optimization tools. This allowed us to give analytic characterizations of the best performance that is achievable with a given signal-to-noise ratio constraint, as well as of the minimal signal-to-noise ratio required to guarantee a given performance level.

- **Chapter 5:** This chapter established a bridge between information theory and control theory. Focusing on a specific class of source coding schemes, we were able to show that average data-rate limits can be enforced by imposing a signal-to-noise ratio constraint in a related additive noise communication channel. This is a key fact that opens the door to using the result of Chapter 4 to deal with control systems subject to average data-rate limits.

- **Chapter 6:** This chapter dealt with networked control systems subject to average data-rate limits. Using the insights of Chapter 5, we were able to bound, from above, the average data-rate needed to guarantee MSS, when a general control and coding architecture based on LTI filters and i.i.d. source coding schemes is employed. Our results guarantee MSS at average data-rates that are reasonably close to the absolute lower bounds previously identified in the literature. In contrast to existing results in the literature, our results rely on standard control and information theory. They also make explicit the interplay between minimal average data-rates for MSS, the degrees of freedom that the LTI filters in our proposal provide, and the memory of the source coding scheme. We also extended our results to performance questions. In particular, we provided an analytic characterization of the minimal average data-rate needed to achieve a given performance level when the proposed control and coding architecture is employed.

- **Chapter 7:** In this chapter we studied control systems subject to random data-loss in the feedback path. We showed, for a general case, that there exists a second order moments equivalence between situations that comprise LTI systems connected over an analog i.i.d. erasure channel, and the same situation where the unreliable channel has been replaced by a signal-to-noise ratio constrained additive i.i.d. noise channel, plus a static gain. This key result allowed us to use the results of Chapter 4 to study the interplay between dropout probabilities and MSS and performance of networked control systems closed over unreliable channels. In particular, we characterized the minimal admissible dropout probabilities that guarantee MSS for a general LTI control and coding architecture. Our results show, as previously reported in the literature, that the use of feedback based coding schemes (in other words, the presence of packet acknowledgements) allows one to reduce the channel reliability needed to achieve MSS. A key feature of our results is that they are based on LTI architectures and, nevertheless, recover results found in the literature which employ more general and complex control and coding architectures. We also provided a methodology to optimally design LTI controllers in networked situations that use unreliable channels.

8.3 Future Work

The results of this thesis have the potential to be extended in several directions:

- **Analytical performance bounds:** Most analytical results in this thesis give answer to stability related questions. On the other hand, performance related results are algorithmic and do not allow one to gain insight into the fundamental performance limits that, for example, signal-to-noise ratio constraints impose. It would be interesting to establish, at least for certain special cases, an analytical characterization of the cost of employing non-transparent communication channels.

- **Multiple channel architectures:** Our results are valid for networked control systems with one non-transparent channel. This assumption is acceptable in a theoretical setting where the focus is on fundamental questions and not necessarily practicality. In a realistic setting, one should consider at least two channels: one for the plant to controller link, and one for the controller to plant link. We also note that any of these links may encompass several channels. This opens the door to study the possible benefits of multiple parallel-channel communication architectures, where channels with different characteristics are used to transmit different parts of the signals of interest depending on, for example, the frequency content of these signals, its relevance, etc. (see, e.g., [82, 86, 142]). Preliminary work on this area has been carried out by the author and co-workers in [131, 133].

- **MIMO plant models:** In this thesis we have focused (mainly) on SISO plant models. An immediate question is whether or not our results can be extended to the MIMO case. Given the fact that our techniques and proofs seldom exploit the dimensions of the signals and systems involved, we believe that such an extension is indeed easily accomplishable. However, we feel that just a MIMO restatement of our results is not of fundamental importance. Much more interesting questions go along the lines of our results in Chapter 3, i.e., focus on how to use networked control architectures to overcome the limitations that decentralized control architectures impose (see also [80]).

- **Multiple communication constraints:** We have focused on channels where either signal-to-noise ratio constraints, or average data-rate constraints, or random data-dropouts are present. It would be interesting to explore situation were some of these constraints occur simultaneously. In order to do so, one would need to appropriately combine the techniques and results presented in this thesis. We are currently working on such extensions (see related work in [132]).

- **General control and coding architectures:** Our results are based on simple control and coding architectures. We believe that this is an interesting framework and may be sufficient from a practical point of view. However, it is worth exploring whether the limitations identified in the LTI case are representative of the limitations that govern the general case. This is specially relevant in the case of channels subject to average data-rate limits, where the consideration of i.i.d. source coding schemes is clearly restrictive.

- **Robustness:** A question that we have left aside form this thesis is that of robustness. In a networked control framework, robustness issues may arise due to the fact both the plant and channel models are inaccurate. Given the nature of our approach (which reduces communication constraints to additive noise plus suitable constraints), it follows that plant model inaccuracies can, in principle, be taken into account using standard techniques from the robust control field (see, e.g., [194]). The issue of channel inaccuracies may require additional effort that is left for future study.

Appendix A

Tools for Analytic Optimization in $\mathcal{H}_2$

This appendix presents definitions and results that are repeatedly used throughout the thesis when analytic expressions for the solution of $\mathcal{H}_2$ optimization problems are sought. We use the notation introduced in Section 2.2.

A.1 The spaces $(\mathcal{R})\mathcal{L}_2$, $(\mathcal{R})\mathcal{H}_2$ and $(\mathcal{R})\mathcal{H}_2^\perp$

This section follows the presentation in Section 6.2 of [182] with a slight change of notation (see also Chapter 2 in [46], Chapter 17 in [141] and [113]).

The set of all functions F of the complex variable z, defined for $|z| < 1$, analytic there and such that

$$||F||_\perp^2 \triangleq \sup_{r\in(0,1)} \text{trace}\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} F(re^{j\omega})F(re^{j\omega})^H d\omega\right\} < \infty, \tag{A.1}$$

is called $\hat{\mathcal{H}}_2^\perp$.

The set of all functions F of the complex variable z, defined on $|z| = 1$, and such that

$$||F||_2^2 \triangleq \text{trace}\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} F(e^{j\omega})F(e^{j\omega})^H d\omega\right\} < \infty, \tag{A.2}$$

is called $\mathcal{L}_2$. If one equips $\mathcal{L}_2$ with the inner product $\langle\cdot,\cdot\rangle$ defined via

$$\langle F, G\rangle \triangleq \text{trace}\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} F(e^{j\omega})G(e^{j\omega})^H d\omega\right\}, \tag{A.3}$$

then $\mathcal{L}_2$ becomes a Hilbert space (see, e.g, [102, 141]). The set $\{e^{j\omega n}\}_{n=-\infty}^{\infty}$ is an orthonormal basis in $\mathcal{L}_2$.

If $\hat{F}(z) \in \hat{\mathcal{H}}_2^{\perp}$, then the function F defined via $F(e^{j\omega}) \triangleq \lim_{z\to e^{j\omega}} \hat{F}(z)$,[1] belongs to $\mathcal{L}_2$ and is such that

$$||\hat{F}||_{\perp}^2 = ||F||_2^2 . \tag{A.4}$$

In other words, for any $\hat{F} \in \hat{\mathcal{H}}_2$ one can always define a function $F \in \mathcal{L}_2$ that is the natural extension of $\hat{F}$ to the unit circle. F is the *boundary function* (or value) of $\hat{F}$. (For most practical purposes, $\hat{F}$ and F are indistinguishable.)

Denote by $\mathcal{H}_2^{\perp} \subset \mathcal{L}_2$ the set of all boundary functions of functions in $\hat{\mathcal{H}}_2^{\perp}$. $\mathcal{H}_2^{\perp}$ is a closed subspace of $\mathcal{L}_2$ (hence, $\mathcal{H}_2^{\perp}$ is a Hilbert space). The orthogonal complement of $\mathcal{H}_2^{\perp}$ in $\mathcal{L}_2$ is called $\mathcal{H}_2$. Thus, for every $F \in \mathcal{L}_2$, there exist $[F]_{\mathcal{H}_2} \in \mathcal{H}_2$ and $[F]_{\mathcal{H}_2^{\perp}} \in \mathcal{H}_2^{\perp}$ such that

$$F = [F]_{\mathcal{H}_2} + [F]_{\mathcal{H}_2^{\perp}} . \tag{A.5}$$

It is fairly easy to find $[F]_{\mathcal{H}_2}$ and $[F]_{\mathcal{H}_2^{\perp}}$. Indeed, one can start by using $\{e^{j\omega i}\}_{i=-\infty}^{\infty}$ to write

$$F(e^{j\omega}) = \sum_{i=-\infty}^{\infty} \alpha_i e^{jwi} = \sum_{i=-\infty}^{-1} \alpha_i e^{jwi} + \sum_{i=0}^{\infty} \alpha_i e^{jwi}. \tag{A.6}$$

The first term in the right hand side of (A.6) corresponds to $[F]_{\mathcal{H}_2}$, and the second one corresponds to $[F]_{\mathcal{H}_2^{\perp}}$. It is clear that, if $F \in \mathcal{H}_2$ (resp. $F \in \mathcal{H}_2^{\perp}$), then $[F]_{\mathcal{H}_2^{\perp}} = 0$ (resp. $[F]_{\mathcal{H}_2} = 0$).

In this thesis we constrain ourselves to real rational transfer functions defined for $z \in \mathbb{C}$, and we use the notation $F(z)$ (rather than [just] F) to refer to them. If we say that $F(z) \in \mathcal{L}_2$, then we mean that if we restrict $F(z)$ to $z = e^{j\omega}$, then the resulting function belongs to $\mathcal{L}_2$. The same convention applies to $\mathcal{H}_2$ and $\mathcal{H}_2^{\perp}$. In a similar spirit, we use the notation $||F(z)||_2$ (rather than $||F||_2$) to refer to the $\mathcal{L}_2$-norm of $F(z) \in \mathcal{L}_2$. To make explicit the fact that we are working with real rational functions, we usually add an $\mathcal{R}$ prefix to the symbols $\mathcal{L}_2$, $\mathcal{H}_2$ and $\mathcal{H}_2^{\perp}$. From the discussion at the end of the previous paragraph, we conclude that, in the real rational case, the characterization of functions in any of the considered spaces is very simple: $F(z) \in \mathcal{RL}_2$ if and only if $F(z)$ has no poles on the unit circle (i.e., no poles in $\{z \in \mathbb{C} : |z| = 1\}$); $F(z) \in \mathcal{RH}_2$ if and only if $F(z)$ is a stable and strictly proper transfer function (i.e., all its poles are in $\{z \in \mathbb{C} : |z| < 1\}$ and $\lim_{z\to\infty} F(z) = 0$). $F(z) \in \mathcal{RH}_2^{\perp}$ if and only if $F(z)$ has no stable poles (i.e, no poles in $\{z \in \mathbb{C} : |z| < 1\}$). Also, for any $F(z) \in \mathcal{RL}_2$, $[F(z)]_{\mathcal{H}_2^{\perp}}$ and $[F(z)]_{\mathcal{H}_2}$

[1] Here, the limit is a non-tangential one, and it can be shown that $F(e^{j\omega})$ is well defined almost everywhere in $[-\pi, \pi]$ (see [141]).

are easily identified upon using a partial fraction expansion of $F(z)$: $[F(z)]_{\mathcal{H}_2}$ corresponds to all strictly proper and stable terms in the expansion, and $[F(z)]_{\mathcal{H}_2^\perp}$ corresponds to the remaining terms (including constants and improper terms).

A.2 Inner-Outer Factorizations

Definition A.1 (Unitary transfer function) *A transfer function $A(z) \in \mathcal{R}$ is said to be* unitary *if and only if $A(z)^\sim A(z) = I$.* □□

Definition A.2 (Inner and outer functions (see, e.g., [46, 123, 194]))

1. *$A(z) \in \mathcal{R}$ is said to be* inner *if and only if $A(z) \in \mathcal{RH}_\infty$ and $A(z)$ is unitary.*

2. *$A(z) \in \mathcal{R}$ is said to be* outer *if and only if $A(z) \in \mathcal{RH}_\infty$ and has full row rank for every $|z| > 1$.*

□□

Note that all matrices in $\overline{\mathcal{U}}_\infty$ are outer functions. Also, all square outer functions are in $\overline{\mathcal{U}}_\infty$, and all square outer functions that have full rank for every $|z| \geq 1$ are in $\mathcal{U}_\infty$.

Theorem A.1 (See [123] and Section 7.4 in [46]) *If $P(z) \in \mathcal{RH}_\infty$, then*

1. *There exists an inner function $P_i(z)$ and an outer function $P_o(z)$ such that $P(z) = P_i(z)P_o(z)$ (such factorization of $P(z)$ is called* inner-outer factorization*).*

2. *If $P(z)$ has full column normal rank, then $P_o(z) \in \overline{\mathcal{U}}_\infty$ and has, as marginally-MP zeros, the marginally-MP zeros of $P(z)$ only.*

□□

A.3 Properties of the 2-norm

This section summarizes key properties of the 2-norm.

Fact A.1 (Properties of $||\cdot||_2$) *For every $X(z)$ and $Y(z)$ in $\mathcal{RL}_2$ of appropriate dimensions, the following holds:*

1. *$||X(z)||_2^2 = ||X(z)^\sim||_2^2$. If, in addition, $X(z) \in \mathcal{RL}_2^{n\times 1}$, then $||X(z)||_2^2 = \left|\left|\sqrt{X(z)^\sim X(z)}\right|\right|_2^2$.*

2. $||[X(z) \;\; Y(z)]||_2^2 = ||X(z)||_2^2 + ||Y(z)||_2^2 = \left|\left|\begin{bmatrix} X(z) \\ Y(z) \end{bmatrix}\right|\right|_2^2$.

3. If $X(z) \in \mathcal{RH}_2^\perp$ *and* $Y(z) \in \mathcal{RH}_2$*, then*

$$||X(z) + Y(z)||_2^2 = ||X(z)||_2^2 + ||Y(z)||_2^2 . \tag{A.7}$$

4. If $X(z) \in \mathcal{RH}_2^\perp$ *and* $Y(z) \in \mathcal{RH}_\infty$*, then*

$$||X(z) - Y(z)||_2^2 = ||X(z) - X(0)||_2^2 + ||X(0) - Y(z)||_2^2 . \tag{A.8}$$

Proof: Parts 1-3 follow immediately from the definitions of $||\cdot||_2$, $\mathcal{H}_2^\perp$ and $\mathcal{H}_2$. To prove Part 4 we start by writing

$$\frac{X(z)}{z} = \frac{X(0)}{z} + E(z). \tag{A.9}$$

Since $X(z) \in \mathcal{RH}_2^\perp$, then $E(z) \in \mathcal{RH}_2^\perp$ and $\frac{X(0)}{z} \in \mathcal{RH}_2$. On the other hand, $\frac{Y(z)}{z} \in \mathcal{RH}_2$ which implies that

$$||X(z) - Y(z)||_2^2 \overset{(a)}{=} \left|\left|\frac{X(z) - Y(z)}{z}\right|\right|_2^2 = \left|\left|\frac{X(0)}{z} + E(z) - \frac{Y(z)}{z}\right|\right|_2^2 \overset{(b)}{=} ||E(z)||_2^2 + \left|\left|\frac{X(0) - Y(z)}{z}\right|\right|_2^2 , \tag{A.10}$$

where (a) follows form the fact that z is unitary and (b) from Part 3. The result now follows using (A.9) in (A.10). □□□

Lemma A.1 (See proof of Theorem 29 in [140]) *If* $A(z), B(z), C(z)$ *and* $X(z)$ *are matrices of appropriate dimensions,* $\otimes$ *denotes the Kronecker product and* $\text{vec}\{\cdot\}$ *is the column-stacking operator (see, e.g., [11, 19]), then*

$$||A(z) - B(z)X(z)C(z)||_2^2 = \left|\left|\text{vec}\{A(z)\} - \left(C(z)^T \otimes B(z)\right)\text{vec}\{X(z)\}\right|\right|_2^2 . \tag{A.11}$$

□□

A.4 Optimization in $\mathcal{H}_2$

Lemma A.2 *Define, for every* $X(z) \in \mathcal{RH}_\infty$*,*

$$J(X(z)) \triangleq ||A(z) - B(z)X(z)C(z)||_2^2 , \tag{A.12}$$

where $A(z), B(z)$ *and* $C(z)$ *are given transfer matrices in* $\mathcal{RH}_\infty$*. Then:*

1. *J is convex in $X(z)$.*

2. *If $B(z)$ has full column normal rank*[2] *and $C(z)$ has full row normal rank, then J is strictly convex in $X(z)$.*

Proof:

- This is a consequence of the triangle inequality and standard results regarding composition of functions and convexity (see, e.g., Section 3.2.4 in [16]).

- To show strict convexity we need to show that, if $\alpha \in [0,1]$, $X_1(z), X_2(z) \in \mathcal{RH}_\infty$, and $B(z)$ and $C(z)$ are as in the statement of this lemma, then

$$J(\alpha X_1(z) + (1-\alpha)X_2(z)) = \alpha J(X_1(z)) + (1-\alpha)J(X_2(z)) \tag{A.13}$$

if and only if $X_1(z) = X_2(z)$ or $\alpha \in \{0,1\}$.

By definition of the 2-norm,

$$\begin{aligned} &J(\alpha X_1(z) + (1-\alpha)X_2(z)) = \\ &\quad ||\alpha(A(z) - B(z)X_1(z)C(z)) + (1-\alpha)(A(z) - B(z)X_2(z)C(z))||_2^2 = \\ &\quad \alpha^2 J(X_1(z)) + (1-\alpha)^2 J(X_2(z)) - \\ &\quad 2\alpha(\alpha-1)\,\mathrm{Re}\left\{\mathrm{trace}\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} M(e^{j\omega})d\omega\right\}\right\}, \end{aligned} \tag{A.14}$$

where

$$M(e^{j\omega}) \triangleq (A(e^{j\omega}) - B(e^{j\omega})X_1(e^{j\omega})C(e^{j\omega}))^H(A(e^{j\omega}) - B(e^{j\omega})X_2(e^{j\omega})C(e^{j\omega})). \tag{A.15}$$

Therefore, (A.13) holds if and only if

$$\begin{aligned} &\alpha(\alpha-1)\left(J(X_1(z)) + J(X_2(z)) - 2\,\mathrm{Re}\left\{\mathrm{trace}\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} M(e^{j\omega})d\omega\right\}\right\}\right) = \\ &\quad \alpha(\alpha-1)\,||B(z)(X_2(z) - X_1(z))C(z)||_2^2 = 0, \end{aligned} \tag{A.16}$$

i.e., if and only if $\alpha \in \{0,1\}$ or $B(z)(X_2(z) - X_1(z))C(z) = 0$. Since $B(z)$ has full normal column rank and $C(z)$ has full normal row rank, it follows that $B(z)(X_2(z)-X_1(z))C(z) = 0 \Leftrightarrow X_1(z) = X_2(z)$. This completes the proof.

[2] i.e., has full column rank for almost every $z \in \mathbb{C}$ (see, e.g., [87]).

□□□

Consider $D(z) \in \mathcal{RH}_\infty^{1\times 1}$, with multi-set of NMP zeros given by $\{c_1^D, \cdots, c_{n_c^D+\bar{n}_c^D}^D\}_m$, where $\left|c_i^D\right| > 1$ for every $i \leq n_c$ and $\left|c_i^D\right| = 1$ for $i > n_c$. Consider the products[3]

$$\xi_c^D(z) \triangleq \prod_{i=1}^{n_c^D} \frac{1 - zc_i^D}{z - c_i^D}, \quad \bar{\xi}_c^D(z) \triangleq \prod_{i=n_c^D+1}^{n_c^D+\bar{n}_c^D} \frac{z}{z - c_i^D}, \quad \bar{\xi}_c^{\varepsilon,D}(z) \triangleq \prod_{i=n_c^D+1}^{n_c^D+\bar{n}_c^D} \frac{z}{z - c_i^D(1-\varepsilon)}, \tag{A.17}$$

where $\varepsilon \in (0, 1]$, and recall from Section 2.2 that $\xi_c^D(z)$ and $\bar{\xi}_c^D(z)$ are biproper and unstable with stable inverses, $\bar{\xi}_c^{\varepsilon,D}(z) \in \mathcal{U}_\infty$, and

$$\lim_{\varepsilon \to 0^+} \bar{\xi}_c^D(z)^{-1} \bar{\xi}_c^{\varepsilon,D}(z) = 1. \tag{A.18}$$

Also recall from Section 2.2 the definition of the operator $\mathcal{K}_\varepsilon \{\cdot\}$. The following lemma is an adaptation of Lemma 10, p. 171, in [182]:

Lemma A.3 *Define*

$$J(X(z)) \triangleq \left|\left|A(z) - \xi_c^D(z)D(z)X(z)\right|\right|_2^2, \tag{A.19}$$

where $A(z) \in \mathcal{RH}_\infty^{1\times n}$ *and* $D(z) \in \mathcal{RH}_\infty^{1\times 1}$ *are given transfer matrices. Then,*

$$J_o \triangleq \inf_{X(z)\in\mathcal{RH}_\infty} J(X(z)) = 0 \tag{A.20}$$

and

$$X_o(z) \triangleq \arg \inf_{X(z)\in\mathcal{RH}_\infty} J(X(z)) = (\xi_D(z)D(z))^{-1} A(z). \tag{A.21}$$

Moreover, $X_o(z) \in \mathcal{RH}_\infty$ *(i.e., the infimum in* (A.20) *is achieved) if and only if* $\bar{n}_c^D = 0$. *If* $\bar{n}_c^D > 0$, *then*

$$X_\varepsilon(z) = \mathcal{K}_\varepsilon \{\xi_D(z)D(z)\}^{-1} A(z) \tag{A.22}$$

belongs to $\mathcal{RH}_\infty$ *for every* $\varepsilon \in (0, 1]$ *and*

$$\lim_{\varepsilon \to 0^+} J(X_\varepsilon(z)) = 0. \tag{A.23}$$

Proof: By definition of norm, $J(X(z)) \geq 0$ for every $X(z) \in \mathcal{RH}_\infty$. Therefore, it suffices to prove that $J(X(z))$ can be made arbitrarily small by an appropriate choice for $X(z)$. If $\bar{n}_c^D = 0$, then the definition of $\xi_c^D(z)$ implies that $\xi_c^D(z)D(z) \in \mathcal{U}_\infty$ and the result follows. If

[3] We adopt the convention that $\prod_{i=n_1}^{n_2} x_i = 1$ whenever $n_1 > n_2$.

$\bar{n}_c^D > 0$, then $\xi_c^D(z)D(z)$ is not invertible in $\mathcal{RH}_\infty$, which proves that the infimum in (A.20) is not achievable in this case. By definition of $\mathcal{K}_\varepsilon\{\cdot\}$ (see Section 2.2), we have that $X_\varepsilon(z) \in \mathcal{RH}_\infty$ for every $\varepsilon \in (0,1]$ and that

$$J(X_\varepsilon(z)) = \left|\left|A(z) - \bar{\xi}_c^D(z)^{-1}\bar{\xi}_c^{\varepsilon,D}(z)A(z)\right|\right|_2^2. \tag{A.24}$$

Thus, (A.23) follows using (A.18). □□□

Remark A.1 *We note that the results in Lemma A.3 still apply,* mutatis mutandis, *if $A(z) \in \mathcal{RH}_\infty^{n\times m}$ and $D(z) \in \mathcal{RH}_\infty^{n\times n}$. In that case, however, one needs the multivariable analogues of $\xi_c^D(z)$, $\bar{\xi}_c^D(z)$ and $\bar{\xi}_D^{\varepsilon,D}(z)$. These functions can be constructed using the ideas in, e.g., [163,178]. (Note that the definition of $\mathcal{K}_\varepsilon\{\cdot\}$ in Section 2.2 encompasses $n \times n$ transfer functions.)* □□

Appendix B

An Alternative Heuristic View of Quantization

This appendix summarizes a useful (but heuristic) model for uniform quantization often employed in the literature. We do not employ this model in this thesis (except in the simulation studies of Chapter 3), but it is presented here since it provides simple guidelines that are sometimes useful in practice.

Consider the scheme of Figure B.1, where v is a zero mean random process and $\mathcal{Q}$ denotes a *finite* uniform quantizer, i.e.,

$$\mathcal{Q}(x) \triangleq i^* \Delta, \tag{B.1}$$

where

$$i^* \triangleq \arg \min_{i \in \{-\frac{(L-1)}{2}, \cdots, -1, 0, 1, \cdots, \frac{L-1}{2}\}} |x - i\Delta| \,, \tag{B.2}$$

$\Delta > 0$ is the quantization step and $L \triangleq 2n+1$, $n \in \mathbb{N}_0$, is the number of quantizer levels. Define the quantizer *dynamic range* V via $V \triangleq \frac{L\Delta}{2}$. The quantizer $\mathcal{Q}$ is said to be *overloaded* if and only if $|v(k)| > V$ for some $k \in \mathbb{N}_0$. If the quantizer does not overload, then the quantization noise, defined via

$$q \triangleq w - v, \tag{B.3}$$

is such that $|q(k)| \leq \frac{\Delta}{2}$ for every k.

Quantization is a deterministic non-linear operation and, hence, the exact analysis of quantized systems is difficult (see, e.g., [34, 65, 116, 129]). As a consequence, it has become standard,

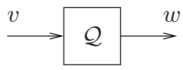

Figure B.1: Uniform quantizer.

particularly in the signal processing literature (see, e.g., [12, 60, 83, 104, 151, 181]), to model quantization noise as follows:

Assumption B.1 (Simplified quantization noise model) *The quantization noise signal q, defined in* (B.3)*, is a white noise sequence, uniformly distributed in $\left(-\frac{\Delta}{2}, \frac{\Delta}{2}\right)$, and uncorrelated with v.* □□

Remark B.1 (Variants of the model) *Sometimes, q is assumed to be an i.i.d. sequence that, in addition, is independent of the input to the quantizer v. In the case of quantizers embedded in feedback loops, it is sometimes more convenient to assume that q is uncorrelated with (or independent of) the exogenous signals in the loop.* □□

Assumption B.1 is valid only in the high resolution regime (i.e., when Δ is small compared with the magnitude of the signal being quantized), and if the quantizer does not overload and the input v has a smooth probability density (see, e.g., [9] and also [58, 59, 170]). Nevertheless, one can make use of dithered quantizers to render the model in Assumption B.1 exact, provided no overload occurs (see, e.g., [59, 60, 185, 191]). Despite the above, the predictions made using the model in Assumption B.1 are sometimes surprisingly accurate (see, e.g., simulation studies in [37,55,160]). Indeed, Assumption B.1 has had a huge impact on the way commercial (feedback based) quantization schemes are designed (see, e.g., [122, 151]).

In order to guarantee that the quantizer does not overload, in principle one needs to consider infinitely many quantization levels (or assume that the quantizer input is deterministically bounded, which may not be the case when exogenous signals are stochastic). In practice, it is standard to choose a dynamic range V such that the probability of overload is negligible (see, e.g, [83]). Indeed, if v is wss and β is any positive real, then one can always find a finite α such that choosing $V = \alpha\sigma_v$ guarantees that the probability of overload is less than β; α is called the quantizer *loading factor*.[1] With such a choice for the loading factor, it is immediate to see that

$$\frac{\sigma_v^2}{\sigma_q^2} = \frac{12 \cdot \sigma_v^2}{\Delta^2} = \frac{3}{\alpha^2} L^2, \tag{B.4}$$

[1]For example, if v is Gaussian and wss, then $\alpha = 4$ guarantees an overload probability of $6.33 \cdot 10^{-5}$.

where we have used the fact that, according to Assumption B.1, $\sigma_q^2 = \frac{\Delta^2}{12}$. This justifies the following additional assumption (see also [83, 181]):

Assumption B.2 (Fixed signal-to-noise ratio) *For a fixed number of quantization levels, the variance of the quantization noise is proportional to the variance of the signal being quantized, i.e., the quantizer has a fixed signal-to-noise ratio γ defined via*

$$\gamma \triangleq \frac{\sigma_v^2}{\sigma_q^2}. \tag{B.5}$$

□□

Assumption B.2 is a key constraint that originates from the desire to avoid quantizer overload. In other words, enforcing Assumption B.2 helps guarantee that the quantizer is always properly scaled (at least in steady state). This assumptions has also a *regularizing* effect on the optimization based design of feedback quantizers as discussed in [37, 55].

Using (B.4) one can paraphrase Assumption B.2 as follows: if the quantizer can have *at most L levels*, then

$$\gamma = \frac{\sigma_v^2}{\sigma_q^2} \leq \frac{3}{\alpha^2} L^2. \tag{B.6}$$

That is, within the simplified model for quantization introduced above, a limit on the number of quantization levels translates into an upper bound on the quantizer signal-to-noise ratio γ. This suggests that a constraint on the number of bits per sample that the quantizer uses to represent its output values is, roughly speaking, equivalent to a constraint on the corresponding quantizer signal-to-noise ratio. In particular, if one assumes that the quantizer defined in (B.1) has b bits, then we have that

$$\gamma = \frac{3}{\alpha^2} \left(2^b - 1\right)^2, \tag{B.7}$$

where, given the fact that the quantizer in (B.1) has an odd number of levels, we have considered that one of the (hypothetical) 2^b levels that can be achieved with b bits is not used.

Remark B.2 *If the quantizer has infinitely many levels, then the quantization noise is always bounded for any input v. In this case, however,* (B.4) *does not make any sense since L is infinite.*

□□

Remark B.3 *In some situations, quantizer overload may become the dominant quantization effect. This is specially true when the quantizer is embedded in a feedback system. Indeed,* *quantizer overload may trigger limit cycle oscillations that, of course, are not predicted by the linear model for quantization described above (see, e.g., [45, 122, 129]). Careful design of the quantizer loading factor may act as a safeguard against quantizer overload, but does not provide guaranteed overload-free operation if the input to the quantizer has infinite support; see also Section III.B in [118]. (Of course, if* v *is deterministically bounded, then it is always possible to pick an appropriate quantizer loading factor that guarantees zero overload probability).* □□

Appendix C

Background on Information Theory

This appendix presents a brief summary of definitions and basic results in Information Theory. Unless explicitly stated otherwise, all proofs and detailed discussions can be found in [31].

C.1 Basics

Throughout this section, unless otherwise stated, x, x_i, $i \in \mathbb{N}_0$, y, z and n are continuous random variables taking values in appropriate subsets of $\mathbb{R}^n$. We assume that they all have well defined probability density functions (PDFs), which we denote by f_x, f_{x_i}, f_y, f_z and f_n, respectively, and well defined joint PDFs denoted by f_{xy}, f_{xz}, etc.[1] We also use the notation $f_{x|y}$ to refer to the conditional PDF of x, given y.

Definition C.1 (Differential entropy) *The differential entropy of x is defined via*[2]

$$h(x) \triangleq -\int f_x(u) \ln f_x(u) du. \tag{C.1}$$

The conditional differential entropy of x, given y, is defined via

$$h(x|y) \triangleq -\int f_{xy}(u,v) \ln f_{x|y}(u,v) du\, dv. \tag{C.2}$$

□□

[1] We will seldom need to make a distinction between a random variable and its realization values. Thus, we introduce at this moment no additional notation for the *values* of x, y, z or n.

[2] It is understood that the integrals are defined over the support of the functions involved.

Example C.1 *If x is a second order scalar Gaussian random variable with variance $\sigma_x^2 > 0$, then $h(x) = \frac{1}{2}\ln 2\pi e\sigma_x^2$.* □□

The differential entropy has the following properties:

Fact C.1 (Properties of h)

- $h(x|y) \leq h(x)$ *with equality if and only if x and y are independent.*
- $h(x+y|y) = h(x|y)$.
- *If $a \in \mathbb{R} \setminus \{0\}$, then $h(ax) = h(x) + \ln|a|$.*
- $h(x_0, \cdots, x_{n-1}) = \sum_{i=0}^{n-1} h(x_i|x_0, \cdots, x_{i-1})$ *(this property is called* chain rule for differential entropy*).*

□□□

Remark C.1 (Discrete entropy) *Sometimes, we will require a version of entropy that applies to discrete valued random variables. If x is a discrete random variable having probability mass function p_x, then we define the discrete entropy of x via*

$$H(x) \triangleq -\sum p_x(u)\ln p_x(u). \tag{C.3}$$

The conditional entropy of x, given another discrete random variable y, is defined via,

$$H(x|y) \triangleq -\sum p_{xy}(u,v)\ln p_{x|y}(u,v) \tag{C.4}$$

where $p_{x|y}$ denotes the conditional probability mass function of x, given y. It is immediate to see that $H(x) \geq 0$ with equality if and only if x is deterministic, and that $H(x|y) \geq 0$ with equality if and only if x is a deterministic function of y. □□

Definition C.2 (Mutual information) *The mutual information between x and y is defined via*

$$I(x;y) \triangleq \int f_{xy}(u,v)\ln\frac{f_{xy}(u,v)}{f_x(u)f_y(v)}du\,dv \tag{C.5}$$

The conditional mutual information between x and y, given z, is defined via

$$I(x;y|z) \triangleq \int f_{xyz}(u,v,w)\ln\frac{f_{xyz}(u,v,w)f_z(w)}{f_{xz}(u,w)f_{yz}(v,w)}du\,dv\,dw. \tag{C.6}$$

□□

Mutual information has the following properties:

Fact C.2 (Properties of I)

1. $I(x;y) = h(x) - h(x|y) = h(y) - h(y|x) = I(y;x)$.

2. $I(x;y|z) = h(x|z) - h(x|y,z) = h(y|z) - h(y|x,z) = I(y;x|z)$.

3. $I(x;y) \geq 0$ *with equality if and only if* x *and* y *are independent.*

4. $I(x,y;z) = I(x;z) + I(y;z|x)$ *(*chain rule of mutual information*).*

□□□

Remark C.2 *It is also possible to define the notion of (conditional) mutual information between discrete random variables. Both the continuous random variable and the discrete random variable notions are compatible (see [31]) and, hence, no additional symbol needs to be defined. We just note that, if* x, y *and* z *are discrete random variables, then* $I(x;y|z) = H(x|z) - H(x|y,z)$. □□

Definition C.3 (Markov chain) *The random variables* x, y *and* z *are said to form a Markov chain (in that order) if and only if* $f(x,z|y) = f(x|y)f(z|y)$, *i.e., if and only if* x *and* z *are conditionally independent given* y. *If that is the case, we write*

$$x \leftrightarrow y \leftrightarrow z. \tag{C.7}$$

□□

Theorem C.1 (Data processing inequality) *If* $x \leftrightarrow y \leftrightarrow z$, *then* $I(x;y) \geq I(x;z)$. *Equality holds if and only if, in addition,* $x \leftrightarrow z \leftrightarrow y$. □□□

Some useful facts involving the data processing inequality and Markov chains are the following:

Fact C.3 (Markov chains)

1. $x \leftrightarrow y \leftrightarrow z$ *if and only if* $z \leftrightarrow y \leftrightarrow x$ *(this justifies the double arrows in our notation).*

2. $x \leftrightarrow y \leftrightarrow z$ *is equivalent to* $f(x,y,z) = f(x)f(y|x)f(z|y)$ *and also to* $f(z|x,y) = f(z|y)$.

3. $x \leftrightarrow y \leftrightarrow z$ *if and only if* $I(x;z|y) = 0$.

4. *If* z *is a deterministic function of* y, *then* $x \leftrightarrow y \leftrightarrow z$.

□□□

Remark C.3 (Conditional Markov chains) *It is possible to define the notion of conditional Markov chain and to state the corresponding data processing inequality (see, e.g., [173]). In such cases, we write*

$$x|_w \leftrightarrow y|_w \leftrightarrow z|_w \tag{C.8}$$

to mean that, conditioned upon w, the variables x, y and z form a Markov chain. If that is the case, then $I(x;y|w) \geq I(x;z|w)$ with equality if and only if $x|_w \leftrightarrow z|_w \leftrightarrow y|_w$. □□

Definition C.4 (Divergence between PDFs) *The divergence of the distribution of x with respect to the distribution of y (in short, the divergence between x and y) is defined by*[3]

$$D(x||y) \triangleq \int f_x(u) \ln \frac{f_x(u)}{f_y(u)} du. \tag{C.9}$$

□□

Example C.2 *If x and y are scalar random variables, y is zero-mean Gaussian with variance σ^2, x is uniform with support $\left(-\frac{\Delta}{2}, \frac{\Delta}{2}\right)$, and $\frac{\Delta^2}{12} = \sigma^2$, then*

$$D(x||y) = \frac{1}{2} \ln \left(\frac{2\pi e}{12} \right) \approx 0.255. \tag{C.10}$$

□□

Relevant properties of $D(\cdot||\cdot)$ are summarized below:

Fact C.4 (Properties of D)

- *$D(x||y) \geq 0$ with equality if and only if $f_x = f_y$ almost everywhere*[4] *(a.e.).*
- *If x_G is a second order Gaussian random variable and x is any other random variable with the same mean and covariance matrix, then*

$$D(x||x_G) = h(x_G) - h(x) = D(ax||ax_G), \tag{C.11}$$

where $a \in \mathbb{R} \setminus \{0\}$ is any real number.

□□□

[3] Also called the Kullback-Leibler "distance" between the distribution of x and the distribution of y.

[4] i.e., $f_x(u) = f_y(u)$ except (perhaps) on a countable set of reals.

Remark C.4 (Conditional divergence) *It will prove useful to consider an extension of the definition of divergence. Given two joint distributions f_{xy} and f_{wz}, we define the conditional divergence between them*[5] *via*

$$D(x|y||w|z) \triangleq \int f_{xy}(u,v) \ln \frac{f_{x|y}(u,v)}{f_{w|z}(u,v)} dudv. \tag{C.12}$$

It is possible to show that the following holds:

- $D(x|y||w|z) \geq 0$.
- *If x_G and y_G are jointly Gaussian random variables having joint PDF $f_{x_G y_G}$, and x and y are arbitrary random variables having a joint PDF f_{xy} with the same first and second order moments as $f_{x_G y_G}$, then*

$$D(x|y||x_G|y_G) = h(x_G|y_G) - h(x|y). \tag{C.13}$$

□□

We end this section with an extension of the notion of differential entropy to random processes.

Definition C.5 *(Differential entropy rate) Consider an asymptotically stationary process x. The differential entropy rate of x is defined by*[6]

$$\bar{h}(x) \triangleq \lim_{k\to\infty} \frac{h(x^{k-1})}{k}. \tag{C.14}$$

□□

If x is stationary, then it is clear that $\bar{h}(x) \leq h(x(k))$, with equality if and only if x is a sequence of independent random variables (recall Fact C.1).

Theorem C.2 (Differential entropy rate (see, e.g., [125, 156])) *If a stationary process $\hat{x}$ is filtered by a stable filter having frequency response $H(e^{j\omega})$, then the filter output x has an entropy rate given by*

$$\bar{h}(x) = \bar{h}(\hat{x}) + \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln \left| H(e^{j\omega}) \right| d\omega. \tag{C.15}$$

□□□

[5] Also known as conditional relative entropy (see [31]).

[6] x^k is shorthand for $x(0), x(1), \cdots, x(k)$.

As an immediate consequence of Theorem C.2 we have that, if x is Gaussian, stationary and has PSD $S_x(e^{j\omega})$, then

$$\bar{h}(x) = \frac{1}{2}\ln(2\pi e) + \frac{1}{4\pi}\int_{-\pi}^{\pi} \ln S_x(e^{j\omega})d\omega. \tag{C.16}$$

If x is, in addition, white, then we have that (C.16) yields

$$\bar{h}(x) = h(x(k)) = \frac{1}{2}\ln 2\pi e\sigma_x^2, \tag{C.17}$$

i.e., the entropy rate of a stationary Gaussian white noise process equals the differential entropy of any of the Gaussian random variables in the sequence.

Remark C.5 (Gaussian process maximizes entropy rate) *We note that* (C.16) *still holds if x an asymptotically stationary Gaussian process having stationary PSD $S_x(e^{j\omega})$. On the other hand, if x is any asymptotically wss process, then it can be shown that (see [107])*

$$\bar{h}(x) \leq \frac{1}{2}\ln(2\pi e) + \frac{1}{4\pi}\int_{-\pi}^{\pi} \ln S_x(e^{j\omega})d\omega, \tag{C.18}$$

where $S_x(e^{j\omega})$ denotes the stationary PSD of x. □□

C.2 Two Technical Lemmas

The following two Lemmas were originally proved in [35, 36] (in particular, see Lemma 4.10 and Theorem 4.12 in [35]), or follow from straightforward modifications of the results there.

Lemma C.1 (Inequalities involving mutual information) *Consider the situation depicted in Figure C.1, where x and n are n-dimensional random variables with arbitrary distributions.*

1. *If x and n are independent, and x_G and n_G denote independent n-dimensional Gaussian random variables having the same means and covariances as x and n, respectively, then*

$$I(x; y) \leq I(x_G; y_G) + D(n||n_G), \tag{C.19}$$

with equality if and only if x and n are jointly Gaussian.

2. *If x is Gaussian and n_G denotes an n-dimensional Gaussian random variable, jointly Gaussian with x, having the same mean and covariance as n, and such that the cross-covariance between n and x equals the cross-covariance between n_G and x, then*

$$I(x; x + n_G) \leq I(x; x + n), \tag{C.20}$$

with equality if the covariance matrix of $x + n$ is non-singular, and n is Gaussian and jointly Gaussian with x.

3. *If x and n are independent and scalar, x_G and n_G denote independent scalar Gaussian random variables having the same means and covariances as x and n, and $D(x||x_G) \leq D(n||n_G)$, then*

$$D(x+n||x_G+n_G) \leq D(n||n_G) \quad \textit{and} \quad I(x_G; x_G+n_G) \leq I(x; x+n), \tag{C.21}$$

with equality if and only if x and n are jointly Gaussian.

□□□

Proof:

1. Using Facts C.2 and C.1, the independence of x, n and x_G, n_G, and the definition of $D(\cdot||\cdot)$, it is easy to see that

$$\begin{aligned}
I(x;y) - I(x_G;y_G) &= h(y) - h(y|x) - h(y_G) + h(y_G|x_G) \\
&= h(x+n) - h(x+n|x) - h(x_G+n_G) + h(x_G+n_G|x_G) \\
&= h(n_G) - h(n) - h(x_G+n_G) + h(x+n) \\
&\overset{(a)}{=} D(n||n_G) - D(x+n||x_G+n_G) \\
&\leq D(n||n_G),
\end{aligned} \tag{C.22}$$

where the last inequality follows from Fact C.4. The result now follows from (C.22) and Fact C.4.

2. Using Fact C.2 it is possible to write

$$I(x; x+n) - I(x; x+n_G) = h(x|x+n_G) - h(x|x+n). \tag{C.23}$$

Use of the facts in Remark C.4 the first part of the result follows. It should be clear that, if n is Gaussian, then equality holds in (C.20). The proof of the converse can be found in [35].

3. If the right hand side in equality (a) in (C.22) were positive, then the result would be true. Thus, we will start examining the difference $D(n||n_G) - D(x+n||x_G+n_G)$:

$$\begin{aligned}
D(n||n_G) - D(x+n||x_G+n_G) &= h(n_G) - h(n) - h(x_G+n_G) + h(x+n) \\
&= h(x+n) - h(n) + \frac{1}{2}\ln\frac{2\pi e\sigma_{n_G}^2}{2\pi e\left(\sigma_{x_G}^2 + \sigma_{n_G}^2\right)} \\
&= h(x+n) - h(n) - \frac{1}{2}\ln\left(1 + \frac{\sigma_{x_G}^2}{\sigma_{n_G}^2}\right),
\end{aligned} \tag{C.24}$$

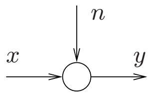

Figure C.1: Additive channel.

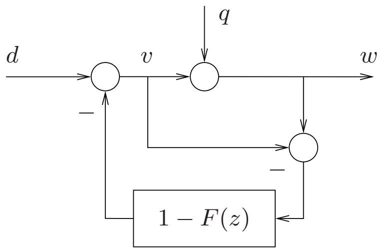

Figure C.2: Feedback system in Lemma C.2.

where we have used Fact C.4, the independence of x_G, n_G and Gaussianity. On the other hand, the entropy power inequality allows one to conclude that, since x, n are independent,

$$h(x+n) - h(n) \geq \frac{1}{2}\ln\left(e^{2h(x)} + e^{2h(y)}\right) - h(n) = \frac{1}{2}\ln\left(1 + \frac{e^{2h(x)}}{e^{2h(n)}}\right) \tag{C.25}$$

Use of (C.25) in (C.24) yields

$$D(n||n_G) - D(x+n||x_G+n_G) \geq M \triangleq \frac{1}{2}\ln\left(1 + \frac{e^{2h(x)}}{e^{2h(n)}}\right) - \frac{1}{2}\ln\left(1 + \frac{\sigma_{x_G}^2}{\sigma_{n_G}^2}\right) \tag{C.26}$$

and, since the variance of the Gaussian and non-Gaussian random variables is the same, we have from (C.26) that

$$M \geq 0 \Leftrightarrow \frac{e^{2h(x)}}{\sigma_x^2} \geq \frac{e^{2h(n)}}{\sigma_n^2} \overset{(a)}{\Leftrightarrow} h\left(\frac{x}{\sigma_x}\right) \geq h\left(\frac{n}{\sigma_n}\right) \overset{(b)}{\Leftrightarrow} \frac{1}{2}\ln 2\pi e - D(x||x_G) \geq \frac{1}{2}\ln 2\pi e - D(n||n_G) \Leftrightarrow D(x||x_G) \leq D(n||n_G), \tag{C.27}$$

where (a) follows from Fact C.1 and (b) from Fact C.4 and C.1, and the fact that the variance of the Gaussian and non-Gaussian random variables is the same. The result follows using (C.27) and (C.26) in equality (a) in (C.22).

□□□

Definition C.6 (Mutual information rate) *Consider two random processes $\tilde{d}$ and w. We define (if the defining limit exists) the* mutual information rate *between $\tilde{d}$ and w as*

$$\bar{I}_\infty(\tilde{d}; w) \triangleq \lim_{k\to\infty} \frac{1}{k} I(\tilde{d}^{k-1}; w^{k-1}). \tag{C.28}$$

□□

Lemma C.2 (Directed information in terms of mutual information rate) *Consider the feedback system in Figure C.2, where $1 - F(z) \in \mathcal{RH}_2$, $\tilde{d}$ is a random process, and q is an i.i.d. sequence that is independent of $\tilde{d}$ and of the initial state of $F(z)$. Then,*

$$\bar{I}_\infty(\tilde{d}; w) = I_\infty(v \to w) - \sum_{i=1}^{n_F} \log \left| p_i^F \right|, \tag{C.29}$$

where $\{p_1^F, \cdots, p_{n_F}^F\}$ denotes the set of NMP zeros of $F(z)$, and $I_\infty(v \to w)$ is the average directed mutual information between v and w (see Definition 5.4). □□□

Proof: By definition and the chain rule of mutual information (see, e.g., [31]),

$$\bar{I}_\infty(\tilde{d}; w) = \lim_{k\to\infty} \frac{1}{k} I(\tilde{d}^{k-1}; w^{k-1}) = \lim_{k\to\infty} \frac{1}{k} \sum_{i=0}^{k-1} I(w(i); \tilde{d}^{k-1} | w^{i-1}). \tag{C.30}$$

Since w depends causally on $\tilde{d}$, it follows that $I(w(i); \tilde{d}^{k-1}|w^{i-1}) = I(w(i); \tilde{d}^i|w^{i-1})$. Thus,

$$\bar{I}_\infty(\tilde{d}; w) = I_\infty(\tilde{d} \to w). \tag{C.31}$$

Also, define $n \triangleq w - \tilde{d}$ and note that

$$n = F(z)q. \tag{C.32}$$

We first note that

$$\begin{aligned} I(w(i); \tilde{d}^i|w^{i-1}) - I(w(i); v^i|w^{i-1}) &\overset{(a)}{=} h(w(i)|w^{i-1}; v^i) - h(w(i)|w^{i-1}, \tilde{d}^i) \\ &\overset{(b)}{=} h(w(i)|w^{i-1}; v^i) - h(n(i)|w^{i-1}, \tilde{d}^i) \\ &\overset{(c)}{=} h(w(i)|w^{i-1}; v^i) - h(n(i)|n^{i-1}, \tilde{d}^i) \\ &\overset{(c)}{=} h(w(i)|w^{i-1}; v^i) - h(n(i)|n^{i-1}), \end{aligned} \tag{C.33}$$

where (a) follows from Fact C.2, (b) follows from the definition of n and Fact C.1, (c) follows from the fact that, by definition of n, $M \leftrightarrow (w^{i-1}, \tilde{d}^i) \leftrightarrow (n^{i-1}, \tilde{d}^i)$ for every M, (d) follows from the fact that, since $\tilde{d}$ is independent of n and of the initial state of $F(z)$, $\tilde{d}$ is independent of n and, thus, $n(i) \leftrightarrow n^{i-1} \leftrightarrow \tilde{d}^i$ holds.

We also have that

$$\begin{aligned} h(w(i)|w^{i-1}, v^i) &\overset{(a)}{=} h(v(i) + q(i)|w^{i-1}, v^i) \\ &\overset{(b)}{=} h(q(i)|q^{i-1}, v^i) \\ &\overset{(c)}{=} h(q(i)|q^{i-1}), \end{aligned} \tag{C.34}$$

where (a) follows from the definition of variables in Figure C.2, (b) follows from Fact C.1, and (c) follows from the fact that both the initial state of $F(z)$ and $\tilde{d}$ being independent of q, and q being i.i.d., guarantees that $q(i) \leftrightarrow q^{i-1} \leftrightarrow v^i$.

From (C.31), (C.33), (C.34) and Fact C.1 it follows that

$$\bar{I}_\infty(\tilde{d}; w) - I_\infty(v \to w) = \bar{h}(q) - \bar{h}(n). \tag{C.35}$$

Use of Theorem C.2, (C.32) and the Bode integral theorem (see, e.g., [155]) yields the result. □□□

Remark C.6 *We note that Lemma C.2 is consistent with Lemmas 3.2 and 4.1 in [105] (see also related results in [107]), but establishes an equality relationship between both sides of* (C.29) *and not an inequality one. This is due to the fact that we focus on noise sequences q that are independent and i.i.d., whereas [105] focuses on a general architecture (see also additional results in Chapter 4 in [35]).* □□

Bibliography

[1] B.D.O. Anderson and M. Deistler. Identifiability in dynamic errors-in-variables models. *Journal of Time Series Analysis*, 5(1):1–13, 1984.

[2] B.D.O. Anderson and J.B. Moore. *Optimal filtering.* Prentice Hall, Englewood Cliffs, N.J., 1979.

[3] P. Antsaklis and J. Baillieul. Special issue on networked control systems. *IEEE Transactions on Automatic Control*, 49(9):1421–1423, September 2004.

[4] P. Antsaklis and J. Baillieul. Special issue on technology of networked control systems. *Proceedings of the IEEE*, 95(1):5–8, January 2007.

[5] K. J. Åström and B. Wittenmark. *Computer controlled systems. Theory and design.* Prentice Hall, Englewood Cliffs, N.J., third edition, 1997.

[6] K.J. Åström. *Introduction to Stochastic Control Theory.* Academic Press, New York, 1970.

[7] J. Baillieul. Feedback Designs in Information Based Control. *Stochastic Theory and Control: Proceedings of a Workshop Held in Lawrence, Kansas*, 2002.

[8] Y. Bar-Shalom and E. Tse. Dual effect, certainty equivalence, and separation in stochastic control. *IEEE Transactions on Automatic Control*, 19(5):494– 500, October 1974.

[9] W.R. Bennet. Spectra of quantized signals. *Bell Syst. Tech. J.*, 27(4):446–472, 1948.

[10] T. Berger. *Rate distortion theory: a mathematical basis for data compression.* Englewood Cliffs, NJ: Prentice-Hall, 1971.

[11] D.S. Bernstein. *Matrix Mathematics.* Princeton University Press, 2005.

[12] H. Bölcskei and F. Hlawatsch. Noise reduction in oversampled filter banks using predictive quantization. *IEEE Transactions on Information Theory*, 47(1):155–172, January 2001.

[13] S. Boyd and C. Barratt. *Linear Controller Design: Limits of Performance*. Prentice Hall, New Jersey, 1991.

[14] S. Boyd, C. Barratt, and S. Norman. Linear controller design: Limits of performance via convex optimization. *Proceedings of the IEEE*, 78(3):529–573, March 1990.

[15] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*. SIAM, 1994.

[16] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[17] J.H. Braslavsky, R.H. Middleton, and J.S. Freudenberg. Feedback stabilisation over signal-to-noise ratio constrained channels. In *Proceedings of the American Control Conference*, Boston, USA, 2004.

[18] J.H. Braslavsky, R.H. Middleton, and J.S. Freudenberg. Feedback stabilization over signal-to-noise ratio constrained channels. *IEEE Transactions on Automatic Control*, 52(8):1391–1403, 2007.

[19] J.W. Brewer. Kronecker products and matrix calculus in system theory. *IEEE Transactions on Circuits and Systems*, 25(9):772–781, September 1978.

[20] E. H. Bristol. On a new measure of interaction for multivariable process control. *IEEE Transactions on Automatic Control*, 11(1):133–134, 1966.

[21] R. Brockett and D. Liberzon. Quantized feedback stabilization of linear systems. *IEEE Transactions on Automatic Control*, 45(7):1279–1289, June 2000.

[22] L. Bushnell (Guest Editor). Special section on networks and control. *IEEE Control Systems Magazine*, 21:57–99, February 2001.

[23] P. Campo and M. Morari. Achievable closed-loop properties of systems under decentralized control: Conditions involving the steady-state gain. *IEEE Transactions on Automatic Control*, 39(5):932–943, May 1994.

[24] C. Canudas de Wit, J. Jaglin, and K.C. Vega. Entropy coding for networked controlled systems. In *Proceedings of the European Control Conference*, Kos, Greece, 2007.

[25] C. Canudas de Wit, F. Rodríguez, J. Fornés, and F. Gómez-Estern. Differential coding in networked controlled linear systems. In *Proceedings of the American Control Conference*, Minneapolis, USA, 2006.

[26] J. Chen, S. Hara, and G. Chen. Best tracking and regulation performance under control effort constraint. *IEEE Transactions on Automatic Control*, 48(8):1320–1380, August 2003.

[27] J. Chen, L. Qiu, and O. Toker. Limitations on maximal tracking accuracy. *IEEE Transactions on Automatic Control*, 45(2):326–331, 2000.

[28] J. Chen, O. Toker, and L. Qiu. Limitations on maximal tracking accuracy. *IEEE Transactions on Automatic Control*, 45(2):326–331, February 2000.

[29] O.L.V. Costa and M.D. Fragoso. Stability results for discrete-time linear systems with Markovian jumping parameters. *Journal of Mathematical Analysis and Applications*, 179:154–178, 1993.

[30] O.L.V. Costa, M.D. Fragoso, and R.P. Marques. *Discrete Time Markov Jump Linear Systems*. Springer, 2005.

[31] T.M. Cover and J.A. Thomas. *Elements of Information Theory*. John Wiley and Sons, Inc., 2nd edition, 2006.

[32] R.E. Curry. *Estimation and Control with Quantized Measurements*. M.I.T. Press, Cambridge, MA, 1970.

[33] C.E. de Souza and M.D. Fragoso. $\mathcal{H}_\infty$ filtering for discrete-time linear systems with markovian jumping parameters. In *Proceedings of the 36th IEEE Conference on Decision and Control*, San Diego, USA, 1997.

[34] D.F. Delchamps. Stabilizing a linear system with quantized state feedback. *IEEE Transactions on Automatic Control*, 35(8):916–924, August 1990.

[35] M.S. Derpich. *Optimal Source Coding with Signal Transfer Function Constraints*. PhD thesis, School of Electrical Engineering and Computer Science, The University of Newcastle, Australia, 2008.

[36] M.S. Derpich. Some mutual information inequalities. Technical report, School of Electrical Engineering and Computer Science, The University of Newcastle, Australia, 2008 (available at `http://altair.elo.utfsm.cl/public_elo/uploads/pdf/Derpich_INEQ_TR.pdf`).

[37] M.S. Derpich, E.I. Silva, D.E. Quevedo, and G.C. Goodwin. On optimal perfect reconstruction feedback quantizers. *IEEE Transactions on Signal Processing*, 56(8):3871–3890, August 2008.

[38] J.C. Doyle, B.A. Francis, and A. Tannenbaum. *Feedback Control Theory.* Macmillan Publishing Company, New York, 1992.

[39] R. Durrett. *Probability: Theory and Examples.* Wadsworth & Brooks/Cole, 1991.

[40] N. Elia. When Bode meets Shannon: Control oriented feedback communication schemes. *IEEE Transactions on Automatic Control*, 49(9):1477–1488, 2004.

[41] N. Elia. Remote stabilization over fading channels. *Systems & Control Letters*, 54(3):237–249, 2005.

[42] N. Elia and S. Mitter. Stabilization of linear systems with limited information. *IEEE Transactions on Automatic Control*, 46(9):1384–1400, 2001.

[43] F. Fagnani and S. Zampieri. Stability analysis and synthesis for scalar linear systems with a quantized feedback. *IEEE Transactions on Automatic Control*, 48(9):1569–1584, September 2003.

[44] F. Fagnani and S. Zampieri. Quantized stabilization of linear systems: complexity versus performance. *IEEE Transactions on Automatic Control*, 49(9):1534–1548, 2004.

[45] O. Feely. Nonlinear dynamics of discrete-time circuits: A survey. *Int. J. Circuit Theory Appl.*, 35(5-6):515–531, Sept.-Dec. 2007.

[46] B. Francis. *A Course on H_∞ Control Theory.* Springer, New York, 1987.

[47] J.S. Freudenberg, J.H. Braslavsky, and R.H. Middleton. Control over signal-to-noise ratio constrained channels: Stabilization and performance. In *Proceedings of the 44th IEEE Conference on Decission and Control and 2005 European Control Conference*, Sevilla, Spain, 2005.

[48] 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*, New Orleans, USA, 2007.

[49] J.S. Freudenberg, R.H. Middleton, and V. Solo. The minimal signal-to-noise ratio requirements to stabilize over a noisy channel. In *Proceedings of the American Control Conference*, Minneapolis, USA, 2006.

[50] M. Fu and L. Xie. The sector bound approach to quantized feedback control. *IEEE Transactions on Automatic Control*, 50(11):1698–1711, November 2005.

[51] M. Gastpar, B. Rimoldi, and M. Vetterli. To code, or not to code: lossy source-channel communication revisited. *IEEE Transactions on Information Theory*, 49(5):1147–1158, 2003.

[52] A. Goldsmith. *Wireless Communications*. Cambridge University Press, 2005.

[53] G.C. Goodwin, S.F. Graebe, and M.E. Salgado. *Control System Design*. Prentice Hall, New Jersey, 2001.

[54] G.C. Goodwin, H. Haimovich, D.E. Quevedo, and J. Welsh. A moving horizon approach to networked control system design. *IEEE Transactions on Automatic Control*, 49(9):1427–1445, September 2004.

[55] G.C. Goodwin, D.E. Quevedo, and E.I. Silva. Architectures and coder design for networked control systems. *Automatica*, 44(1):248–257, 2008.

[56] G.C. Goodwin, M.E. Salgado, and E.I. Silva. Time-domain performance limitations arising from decentralized architectures and their relationship to the RGA. *International Journal of Control*, 78(13):1045–1062, September 2005.

[57] M. Grant and S. Boyd. *CVX: Matlab software for disciplined convex programming (web page and software). http://stanford.edu/ boyd/cvx*. June 2008.

[58] R.M. Gray. Quantization noise spectra. *IEEE Transactions on Information Theory*, 36(6):1220–1244, November 1990.

[59] R.M. Gray and D.L. Neuhoff. Quantization. *IEEE Transactions on Information Theory*, 44(6):2325–2383, October 1998.

[60] R.M. Gray and T.G. Stockham (Jr.). Dithered quantizers. *IEEE Transactions on Information Theory*, 39(3):805–812, May 1993.

[61] M. Grimble. *Robust industrial control: Optimal design approach for polynomial systems*. Prentice Hall, 1994.

[62] C.M Grinstead and J.L. Snell. *Introduction to Probability*. AMS books, 1997.

[63] A.N. Guclu and A.B. Ozguler. Diagonal stabilization of linear multivariable systems. *International Journal of Control*, 43(3):965–980, 1986.

[64] A.N. Gündeş and M.G. Kabuli. Reliable decentralized integral action controller design. *IEEE Transactions on Automatic Control*, 46(2):296 – 301, 2001.

[65] C.S. Güntürk. One-bit sigma-delta quantization with exponential accuracy. *Commun. Pure Appl. Math.*, 56(11):1608–1630, 2003.

[66] Y. Guo, D.J. Hill, and Y. Wang. Nonlinear decentralized control of large-scale power systems. *Automatica*, 36(9):1275–1289, 2000.

[67] V. Gupta, B. Hassibi, and R.M. Murray. Optimal LQG control across packet-dropping links. *Systems & Control Letters*, 56(6):439–446, 2007.

[68] V. Gupta, D. Spanos, B. Hassibi, and R.M. Murray. On LQG control across a stochastic packet-dropping link. In *Proceedings of the American Control Conference*, Portland, USA, 2005.

[69] A. Gut. *Probability: A Graduate Course*. Springer, 2005.

[70] C.N. Hadjicostis and R. Touri. Feedback control utilizing packet dropping network links. In *Proceedings of the 41st IEEE Conference on Decision and Control*, Las Vegas, USA, 2002.

[71] H. Haimovich. *Quantisation Issues in Feedback Control*. PhD thesis, Department of Electrical Engineering, The University of Newcastle, 2006.

[72] E.J. Hannan. Rational transfer function approximation. *Statistical Science*, 2(2):135–151, 1987.

[73] A. Hassibi, S. Boyd, and J.P. How. Control of asynchronous dynamical systems with rate constraints on events. In *Proceedings of the 38th IEEE Conference on Decision and Control*, Phoenix, USA, 1999.

[74] K. Havre. *Studies on Controllability Analysis and Control Structure Design*. PhD thesis, Department of Chemical Engineering, Norwegian University of Science and Technology, 1998.

[75] J. Hespanha, A. Ortega, and L. Vasudevan. Towards the control of a linear system with minimum bit-rate. In *Proceeding of the 15th International Symposium on Mathematical Theory of Networks and Systems (MTNS)*, Universty of Notre Dame, 2002.

[76] J.P. Hespanha, P. Naghshtabrizi, and Y. Xu. A survey of recent results in networked control systems. *Proceedings of the IEEE*, 95(1):138–162, January 2007.

[77] D. Hristu-Varsakelis and W. Levine (Eds.). *Handbook of Networked and Embedded Systems.* Birkhäuser, 2005.

[78] O.C. Imer, S. Yüksel, and T. Başar. Optimal control of LTI systems over unreliable communication links. *Automatica*, 42:1429 – 1439, 2006.

[79] H. Ishii and B. A. Francis. *Limited Data Rate in Control Systems with Networks.* Springer, 2002.

[80] H. Ishii and B.A. Francis. Stabilization with control networks. *Automatica*, 38:1745 – 1751, 2002.

[81] H. Ishii and B.A. Francis. Stabilizing a linear system by switching control with dwell time. *IEEE Transactions on Automatic Control*, 47(12):1962–1973, December 2002.

[82] H. Ishii and S. Hara. A subband coding approach to networked control. In *Proceedings of the 17th International Symposium on Mathematical Theory of Networks and Systems (MTNS)*, Kyoto, Japan, 2006.

[83] N. Jayant and P. Noll. *Digital Coding of Waveforms. Principles and Approaches to Speech and Video.* Prentice Hall, 1984.

[84] Y. Ji, H.J. Chizeck, X. Feng, and K.A. Loparo. Stability and control of discrete-time jump linear systems. *Control Theory Adv. Technol.*, 7(2):247–270, 1991.

[85] S. Jiang and P.G. Voulgaris. Cooperative control over link limited and packet dropping networks. In *Proceedings of the European Control Conference*, Kos, Greece, 2007.

[86] Z. Jin, V. Gupta, and R.M. Murray. State estimation over packet dropping networks using multiple description coding. *Automatica*, 42(9):1441–1452, September 2006.

[87] Thomas Kailath. *Linear Systems.* Prentice-Hall, Englewood Cliffs, N.J., 1980.

[88] T. Kameneva and D. Nesic. Further results on robustness of linear control systems with quantized feedback. In *Proceeding of the American Control Conference*, New York, USA, 2007.

[89] G. Kaplan. Ethernet's winning ways. *IEEE Spectrum*, 38(1):113–115, 2001.

[90] V. Kariwala. *Multi-loop controller design and performance analysis.* PhD thesis, Chemical and Materials Engineering, University of Alberta, Canada, 2004.

[91] V. Kariwala. Fundamental limitation on achievable decentralized performance. *Automatica*, 43(10):1849–1854, 2007.

[92] P.P. Khargonekar and M.A. Rotea. Multiple objective optimal control of linear systems: the quadratic norm case. *IEEE Transactions on Automatic Control*, 36(1):14–24, 1991.

[93] F. Kozin. A survey of stability of stochastic systems. *Automatica*, 5:95–112, 1969.

[94] H. Kwakernaak and R. Sivan. *Linear Optimal Control Systems.* Wiley–Interscience, New York, 1972.

[95] F. Lian, J. Moyne, and D. Tilbury. Modelling and optimal controller design of networked control systems with mutliple delays. *International Journal of Control*, 76(6):591–606, 2003.

[96] F. Lian, J.R. Moyne, and D.M. Tilbury. Performance evaluation of control networks: Ethernet, ControlNet, and DeviceNet. *IEEE Control Systems Magazine*, 21(1):66–83, February 2001.

[97] T. Linder, V. Tarokh, and K. Zeger. Existence of optimal prefix codes for infinite source alphabets. *IEEE Transactions on Information Theory*, 43(6):2026–2028, 1997.

[98] T. Linder and R. Zamir. Causal coding of stationary sources and individual sequences with high resolution. *IEEE Transactions on Information Theory*, 52(2):662–680, 2006.

[99] Q. Ling. *Stability and performance of control systems with limited feedback information.* PhD thesis, University of Notre Dame, Notre Dame, USA, 2005.

[100] Q. Ling and M.D. Lemmon. Power spectral analysis of networked control systems with data dropouts. *IEEE Transactions on Automatic Control*, 49(6):955–960, June 2004.

[101] I. López, C.T. Abdallah, and C. Canudas de Wit. Gain-scheduling multi-bit delta-modulator for networked controlled system. In *Proceedings of the European Control Conference*, Kos, Greece, 2007.

[102] D. Luenberger. *Optimization by Vector Space Methods.* John Wiley and Sons, Inc., London, 1969.

[103] P.M. Makila, T. Westerlund, and H.T. Toivonen. Constrained linear quadratic Gaussian control with process applications. *Automatica*, 20(1):15–29, 1984.

[104] D. Marco and D.L. Neuhoff. The validity of the additive noise model for uniform scalar quantizers. *IEEE Transactions on Information Theory*, 51(5):1739–1755, May 2005.

[105] N.C. Martins and M.A. Dahleh. Fundamental limitations of performance in the presence of finite capacity feedback. In *Proceedings of the American Control Conference*, Portland, USA, 2005.

[106] N.C. Martins and M.A. Dahleh. Feedback control in the presence of noisy channels: Bode-like fundamental limitations of performance. *IEEE Transactions on Automatic Control*, 53(7):1604–1615, August 2008.

[107] N.C. Martins, M.A. Dahleh, and J.C. Doyle. Fundamental limitations of disturbance attenuation in the presence of side information. *IEEE Transactions on Automatic Control*, 52(1):56–66, 2007.

[108] J.L. Massey. Causality, feedback and directed information. In *Proceedings of the International Symposium on Information Theory and its Applications.*, Hawaii, USA, 1990.

[109] A.S. Matveev and A.V. Savkin. Decentralized stabilization of linear systems via limited capacity communication networks. In *Proceedings of the 44th IEEE Conference on Decission and 2005 Control and European Control Conference*, Sevilla, Spain, 2005.

[110] R.K. Miller, A.N. Michel, and J.A. Farrell. Quantizer effects on steady-state error specifications of digital feedback control systems. *IEEE Transactions on Automatic Control*, 34(6):651–654, 1989.

[111] R.K. Miller, M.S. Mousa, and A.N. Michel. Quantization and overflow effects in digital implementations of linear dynamic controllers. *IEEE Transactions on Automatic Control*, 33(7):698–704, 1988.

[112] Y. Minami, S. Azuma, and T. Sugie. An optimal dynamic quantizer for feedback control with discrete-valued signal constraints. In *Proceedings of the 46th IEEE Conference on Decision and Control*, New Orleans, USA, 2007.

[113] M. Morari and E. Zafiriou. *Robust process control.* Prentice Hall, 1989.

[114] J. R. Moyne and D. M. Tilbury. The emergence of industrial control networks for manufacturing control, diagnostics, and safety data. *Proceedings of the IEEE*, 95(1):29–47, January 2007.

[115] G. Nair and R. Evans. Exponential stabilisability of finite-dimensional linear systems with limited data rates. *Automatica*, 39(4):585–593, April 2003.

[116] G. Nair and R. Evans. Stabilizability of stochastic linear systems with finite feedback data rates. *SIAM Journal on Control and Optimization*, 43(2):413–436, July 2004.

[117] G. Nair, R. Evans, I. Mareels, and W. Moran. Topological feedback entropy and nonlinear stabilization. *IEEE Transactions on Automatic Control*, 49(9):1585–1597, September 2004.

[118] G. Nair, F. Fagnani, S. Zampieri, and R. Evans. Feedback control under data rate constraints: An overview. *Proceedings of the IEEE*, 95(1):108–137, January 2007.

[119] D. Neuhoff and R. Gilbert. Causal source codes. *IEEE Transactions on Information Theory*, 28(5):701–713, 1982.

[120] J. Nilsson. *Real-time Control Systems with Delays.* PhD thesis, Lund Institute of Technology, Sweden, 1998.

[121] J. Nilsson, B. Bernhardsson, and B. Wittenmark. Stochastic analysis and control of real-time systems with random time delays. *Automatica*, 34(1):57–64, 1998.

[122] S. Norsworthy, R. Schreier, and G. Temes (Eds.). *Delta–Sigma Data Converters: Theory, Design and Simulation.* IEEE Press, Piscataway, NJ, 1997.

[123] C. Oară. Constructive solutions to spectral and inner-outer factorizations with respect to the disk. *Automatica*, 41:1855–1866, 2005.

[124] N. Ormond. Welcome to the wireless world. *IEE Computing and Control Engineering*, pages 28–31, February/March 2006.

[125] Athanasios Papoulis. *Probability, random variables and stochastic process.* McGraw Hill Book Company, New York, 3rd edition, 1991.

[126] N.J. Ploplys, P.A. Kawka, and A.G. Alleyne. Closed-loop control over wireless networks. *IEEE Control Systems Magazine*, 24(3):58–71, 2004.

[127] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing.* Cambridge University Press, 1988.

[128] X. Qi, M.V. Salapaka, P.G. Voulgaris, and M. Khammash. Structured optimal and robust control with multiple criteria: A convex solution. *IEEE Transactions on Automatic Control*, 49(10):1623 – 1640, October 2004.

[129] D.E. Quevedo, G.C. Goodwin, and J.A. De Doná. Finite constraint set receding horizon control. *International Journal of Robust and Nonlinear Control*, 14:355–377, March 2004.

[130] D.E. Quevedo, E.I. Silva, and G.C. Goodwin. Packetized predictive control over erasure channels. In *26th American Control Conference*, New York, USA, 2007.

[131] D.E. Quevedo, E.I. Silva, and G.C. Goodwin. Performance limits in multi-channel networked control system architectures. In *Proceedings of the 3rd IFAC SSSC*, 2007.

[132] D.E. Quevedo, E.I. Silva, and G.C. Goodwin. Control over unreliable networks affected by packet erasures and variable transmission delays. *IEEE Journal on Selected Areas In Communications*, 26(4):672–685, May 2008.

[133] D.E. Quevedo, E.I. Silva, and G.C. Goodwin. Subband coding for networked control systems. *To appear in International Journal of Robust and Nonlinear Control*, 2008.

[134] J.B. Rawlings and B.T. Stewart. Coordinating multiple optimization based controllers: New opportunities and challenges. In *Procedings of the 8th International IFAC DYCOPS*, Cancún, México, 2007.

[135] A.J. Rojas. *Feedback Control over Signal to Noise Ratio Constrained Communication Channels.* PhD thesis, School of Electrical Engineering and Computer Science, The University of Newcastle, Australia, 2006.

[136] A.J. Rojas, J.H. Braslavsky, and R.H. Middleton. Control over a bandwidth limited signal-to-noise ratio constrained communication channel. In *Proceedings of the 44th IEEE Conference on Decission and 2005 Control and European Control Conference*, Sevilla, Spain, 2005.

[137] A.J. Rojas, J.H. Braslavsky, and R.H. Middleton. Output feedback stabilisation over bandwidth limited, signal to noise ratio constrained communication channels. In *Proceedings of the American Control Conference*, Minneapolis, USA, 2006.

[138] A.J. Rojas, J.H. Braslavsky, and R.H. Middleton. Output feedback sensitivity functions under signal to noise ratio constraint. In *Proceedings of the American Control Conference*, New York, USA, 2007.

[139] A.J. Rojas, J.S. Freudenberg, J.H. Braslavsky, and R.H. Middleton. Optimal signal to noise ratio in feedback over communication channels with memory. In *Proceedings of the 45th IEEE Conference on Decision and Control*, San Diego, USA, 2006.

[140] M. Rotkowitz and S. Lall. A characterization of convex problems in decentralized control. *IEEE Transactions on Automatic Control*, 51(2):274–286, 2006.

[141] Walter Rudin. *Real and Complex Analysis.* McGraw Hill Book Company, New York, third edition, 1987.

[142] A Sahai. Stabilization using both noisy and noiseless feedback. In *Proceedings of the 17th International Symposium on Mathematical Theory of Networks and Systems (MTNS)*, Kyoto, Japan, 2006.

[143] A. Sahai and S. Mitter. The necessity and sufficiency of anytime capacity for control over a noisy communication link – Part I: Scalar systems. *IEEE Transactions on Information Theory*, 52(8):3369–3395, August 2006.

[144] M.E. Salgado and A. Conley. MIMO interaction measure and controller structure selection. *International Journal of Control*, 77(4):367–383, March 2004.

[145] M.E. Salgado and E.I. Silva. Robustness issues in $\mathcal{H}_2$ optimal control of unstable plants. *Systems & Control Letters*, 55(2):124 – 131, February 2006.

[146] N. Sandell, P. Varaiya, M. Athans, and M. Safonov. Survey of decentralized control methods for large scale systems. *IEEE Transactions on Automatic Control*, 23(2):108–128, 1978.

[147] A.V. Savkin. Analysis and synthesis of networked control systems: Topological entropy, observability, robustness and optimal control. *Automatica*, 42:51–62, 2006.

[148] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S.S. Sastry. Foundations of control and estimation over lossy networks. *Proceedings of the IEEE*, 95(1):163 – 187, January 2007.

[149] C. Scherer, P. Gahinet, and M. Chilali. Multiobjective output-feedback control via LMI optimization. *IEEE Transactions on Automatic Control*, 42(7):896 – 911, July 1997.

[150] G. Schickhuber and O. McCarthy. Distributed fieldbus and control network systems. *Computing & Control Engineering Journal*, 8(1):21–32, 1997.

[151] R. Schreier and G.C. Temes. *Understanding Delta Sigma Data Converters*. Wiley-IEEE Press, 2004.

[152] L. Schuchman. Dither Signals and Their Effect on Quantization Noise. *IEEE Transactions on Communications*, 12(4):162–165, 1964.

[153] P. Seiler and R. Sengupta. A bounded real lemma for jump systems. *IEEE Transactions on Automatic Control*, 48(9):1651–1654, 2003.

[154] P. Seiler and R. Sengupta. An $\mathcal{H}_\infty$ approach to networked control. *IEEE Transactions on Automatic Control*, 50(3), 2005.

[155] M.M. Seron, J.H. Braslavsky, and G.C. Goodwin. *Fundamental Limitations in Filtering and Control*. Springer, London, 1997.

[156] C.E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423, 623–656, July, October 1948.

[157] C.E. Shannon. Coding theorems for a discrete source with a fidelity criterion. *Inst. Radio Eng. Int. Conv. Rec.*, 4:12–163, 1959.

[158] P. Shi, E.K. Boukas, and R.K. Agarwal. Control of markovian jump discrete-time systems with norm bounded uncertainty and unknown delay. *IEEE Transactions on Automatic Control*, 44(11):2139–2144, Nov 1999.

[159] E.I. Silva, G.C. Goodwin, and D.E. Quevedo. On networked control architectures for MIMO plants. In *Proceedings of the 17th IFAC World Congress*, Seoul, Korea, 2008.

[160] E.I. Silva, G.C. Goodwin, D.E. Quevedo, and M.S. Derpich. Optimal noise shaping for networked control systems. In *Proceedings of the European Control Conference*, Kos, Greece, 2007.

[161] E.I. Silva, D.A. Oyarzun, and M.E. Salgado. On structurally constrained $\mathcal{H}_2$ performance bounds for stable MIMO plant models. *IET Control Theory & Applications*, 1(4):1033–1045, 2007.

[162] 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*, Seoul, Korea, 2008.

[163] E.I. Silva and M.E. Salgado. Performance bounds for feedback control of nonminimum-phase MIMO systems with arbitrary delay structure. *IEE Proceedings - Control Theory and Applications*, 152(2):211–219, March 2005.

[164] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M.I. Jordan, and S.S. Sastry. Kalman filtering with intermittent observations. *IEEE Transactions on Automatic Control*, 49(9):1453–1464, September 2004.

[165] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, and S.S. Sastry. Optimal control with unreliable communication: the TCP case. In *Proceedings of the American Control Conference*, Portland, USA, 2005.

[166] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, and S.S. Sastry. Optimal linear LQG control over lossy networks without packet acknowlegements. *Asian Journal of Control*, 10(1), January 2008.

[167] S. Skogestad and I. Postlethwaite. *Multivariable Feedback Control: Analysis and Design.* Wiley, New York, 1996.

[168] T. Söderström. *Discrete-time stochastic systems.* Prentice Hall, 1994.

[169] D. Sourlas and V. Manousiousthakis. Best achievable decentralized performance. *IEEE Transactions on Automatic Control*, 40(11):1858–1871, November 1995.

[170] A.B. Sripad and D.L. Snyder. A necessary and sufficient condition for quantization errors to be uniform and white. *IEEE Trans. Acoust., Speech, Signal Processing*, 25:442–448, October 1977.

[171] G. Stein and M. Athans. The LQG/LTR procedure for multivariable feedback control design. *IEEE Transactions on Automatic Control*, AC-31(3):105–114, 1987.

[172] P.L. Tang and C.W. de Silva. Compensation for transmission delays in an ethernet-based control network using variable-horizon predictive control. *IEEE Transactions on Control and Systems Technology*, 14(4):707–718, July 2006.

[173] S. Tatikonda. *Control under Communication Constraints.* PhD thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, USA, 2000.

[174] S. Tatikonda and S. Mitter. Control over noisy channels. *IEEE Transactions on Automatic Control*, 49(7):1196–1201, July 2004.

[175] S. Tatikonda and S. Mitter. Control under communication constraints. *IEEE Transactions on Automatic Control*, 49(7):1056–1068, July 2004.

[176] S. Tatikonda, A. Sahai, and S. Mitter. Stochastic linear control over a communication channel. *IEEE Transactions on Automatic Control*, 49(9):1549–1561, September 2004.

[177] Y. Tipsuwan and M.Y. Chow. Control methodologies in networked control. *Control Engineering Practice*, 11:1099–1111, 2003.

[178] O. Toker, J. Chen, and L. Qiu. Tracking performance limitations in LTI multivariable discrete-time systems. *IEEE Transactions on Circuits and Systems I*, 49(5):657–670, May 2002.

[179] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication*. Cambridge University Press, UK, 2005.

[180] K. Tsumura and J. Maciejowski. Stabilizability of SISO control systems under constraints of channel capacities. In *Proceedings of the 42nd IEEE Conference on Decision and Control*, Hawaii, USA, 2003.

[181] J. Tuqan and P.P. Vaidyanathan. Statistically optimum pre- and postfiltering in quantization. *IEEE Transactions on Circuits and Systems II*, 44(1):1015–1031, 1997.

[182] M. Vidyasagar. *Control systems synthesis: A Factorization Approach*. MIT Press, Cambridge, USA, 1985.

[183] G. Walsh and H. Ye. Scheduling of networked control systems. *IEEE Control Systems Magazine*, pages 57–65, February 2001.

[184] S.H. Wang and E.J. Davidson. On the stabilization of decentralized control systems. *IEEE Transactions on Automatic Control*, 18(5):473–478, October 1973.

[185] R.A. Wannamaker, S.P. Lipshitz, J. Vanderkooy, and J. N. Wright. A theory of non-subtractive dither. *IEEE Transactions on Signal Processing*, 48(2):499–516, February 2000.

[186] W.S. Wong and R.W. Brockett. Systems with finite communication bandwidth constraints - II: Stabilization with limited information feedback. *IEEE Transactions on Automatic Control*, 44(5):1049–1053, May 1999.

[187] F. Xia, X. Dai, Z. Wang, and Y. Sun. Feedback based network scheduling of networked control systems. In *International Conference on Control and Automation (ICCA)*, Budapest, Hungary, 2005.

[188] L. Xiao, M. Johansson, H. Hindi, S. Boyd, and A. Goldsmith. Joint optimization of communication rates and linear systems. *IEEE Transactions on Automatic Control*, 48(1):148–153, January 2003.

[189] S. Yüksel and T. Başar. Communication constraints for decentralized stabilizability with time-invariant policies. *IEEE Transactions on Automatic Control*, 52(6):1060–1066, 2007.

[190] R. Zamir and M. Feder. On universal quantization by randomized uniform/lattice quantizers. *IEEE Transactions on Information Theory*, 38(2):428–436, March 1992.

[191] R. Zamir and M. Feder. Rate-distortion performance in coding bandlimited sources by sampling and dithered quantization. *IEEE Transactions on Information Theory*, 41(1):141–154, 1995.

[192] R. Zamir, Y. Kochman, and M. Feder. Achieving the Gaussian rate-distortion function by prediction. *IEEE Transactions on Information Theory*, 54(7):3354 – 3364, July 2008.

[193] W. Zhang, M. Branicky, and Stephens Phillips. Stability of networked control systems. *IEEE Control Systems Magazine*, 21:84–99, February 2001.

[194] K. Zhou, J.C. Doyle, and K. Glover. *Robust and optimal control*. Prentice Hall, Englewood Cliffs, N.J., 1996.